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.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CLONING, CHARACTERIZATION, EXPRESSION AND LOCALIZATION OF INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (IMPDH) cDNA/mRNA IN SOYBEAN (*GLYCINE MAX*) NODULES

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

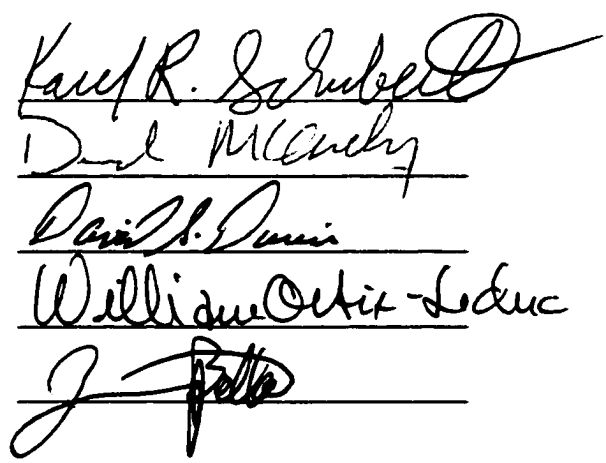
Doctor of Philosophy

By
YAJUAN CAO
Norman, Oklahoma
2000

CLONING, CHARACTERIZATION, EXPRESSION AND LOCALIZATION OF INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (IMPDH) cDNA/mRNA IN SOYBEAN (*GLYCINE MAX*) NODULES

A Dissertation APPROVED FOR THE
DEPARTMENT OF BOTANY AND MICROBIOLOGY

BY



This dissertation is dedicated to my parents, for the memory of my mother.

ACKNOWLEDGEMENTS

Special thanks are due to my major professor, Dr. Karel R. Schubert, for giving me the opportunity of working in his lab and for providing me with a research assistantship for the past three years. Thanks are due to him also for his interesting projects, insightful guidance and his smile all the time to his students.

I am grateful to my committee members, Dr. Jimmy Ballard, Dr. Alyce DeMarais, Dr. David Durica, Dr. David McCarthy and Dr. William Ortiz-Leduc for their advice and assistance. I would also like to express my appreciation to Dr. Ruzhu Chen for his technical advice and valuable discussions on this project, especially at the beginning of my work in Dr. Schubert's lab. The encouragement and support provided by Dr. Tahzeeba Hossain are also appreciated. Yinghe Li, Sandrine Ferrand, Richard Longeras, Isabelle Salles and Lily Gavilano are all acknowledged for their help.

My acknowledgements are given to the Department of Botany and Microbiology for providing me with a Teaching Assistantship from 1994 to 1997 and to Dr. Scott D. Russell for giving me the opportunity of working in his lab. Work from his laboratory resulted in two conference presentations and the publication of two research papers. The help from Greg Strout and Bill Chissoe in the Samuel Noble Electron Microscopy Laboratory, and the help from the faculty, staff and students in the Department of Botany and Microbiology are all greatly appreciated.

Forever gratitude is due to my parents, Yi-Jian Cao and Xiu-Qin Bai. As a former high school Chemistry teacher and a former elementary school teacher, my

parents were not rich, but they raised three children and sent all of them to the top universities of China. They also influenced us with their honesty, intelligence and diligence, which are priceless treasures for my whole life. Due to a heart attack in 1996, my mother was not able to see the day when I obtain my Ph.D. degree.

Thanks are due to my husband, Yunxiang Wu, for being supportive in the pursuit of my career and for giving me a relaxed and joyful family life. I also owe thanks to my two younger brothers, Yafeng and Yachao, who always made time from their busy life to take care of everything at home and let me study in the U.S. free of worry.

Other people who have helped me in my career are acknowledged here. Among them are Professor Fu-Xiong Wang, the late member of the Chinese Academy of Sciences, Professor Hui Wang and Lan-Fen Li at Beijing Normal University, and the people who helped me during my three years of teaching at Tsinghua University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A_{260}: absorbance at 260 nm
BCIP: 5-bromo-4-chloro-3-indolyl phosphate
bp: base pair
BSA: bovine serum albumin
CIAP: calf intestinal alkaline phosphatase
Ci: curie (3.7×10^{10} Bq)
cDNA: complementary DNA
CTAB: Hexadecyl Trimethylammonium Bromide
Da: Dalton
DEPC: diethylpyrocarbonate
DIG: digoxygenin
DMSO: dimethylsulfoxide
DNase: deoxyribonuclease
DTT: dithiothreitol
EDTA: ethylenediamine tetra-acetic acid
EtBr: ethidium bromide
FPLC: fast protein liquid chromatography
HPLC: high-performance liquid chromatography
IMP: inosine monophosphate
IMPDH: inosine monophosphate dehydrogenase
IPTG: isopropyl β-D-thiogalactopyranoside
kb: kilobase
Lb: leghemoglobin
LB agar / broth: Luria-Bertani agar / broth
mRNA: messenger RNA
MW: molecular weight
NAD: nicotinamide adenine dinucleotide
NBT: nitroblue tetrazolium
PAGE: polyacrylamide gel electrophoresis
PCR: polymerase chain reaction
PEG: polyethylene glycol
pI: isoelectric point
Poly(A^+): polyadenylated RNA
PRAT: PRPP amidotransferase
PRPP: phosphoribosyl pyrophosphate
PRS: PRPP synthetase
PVP: polyvinyl pyrrolidone
RNase: ribonuclease
rRNA: ribosomal RNA
SDS: sodium dodecyl sulfate
SDS-PAGE: sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SSC: saline sodium citrate
TAE: Tris-acetate-EDTA electrophoresis buffer
TBE: Tris-borate-EDTA electrophoresis buffer

TEMED: N, N, N', N'-tetramethylethylenediamine
Tris: Tris(hydroxymethyl)aminomethane
UV: ultraviolet
X-Gal: 5-bromo-4-chloro-3-indolyl-β-D-galactoside

ABSTRACT

Inosine monophosphate dehydrogenase (IMPDH, E.C. 1.1.1.205) is a key enzyme in ureide-producing nodules of tropical legumes. A soybean nodule *IMPDH* cDNA was cloned by screening a soybean nodule cDNA library (in lambda gt11) with degenerate probes. The full length of the *IMPDH* cDNA was 2016 bp with a 166 bp 5'-untranslated region and a 341 bp 3'-untranslated tail. The open reading frame encodes a polypeptide of 502 amino acids with a predicted molecular weight of 53 kDa and a pI of 5.54. The deduced amino acid sequence was 70.5% identical to that of *Arabidopsis* and contains the putative active-site region. Southern blotting of digested soybean genomic DNA suggested that there was only one copy of the *IMPDH* gene in the soybean genome. Functional complementation experiments showed that expression of soybean *IMPDH* cDNA in *E. coli* mutant strain KLC381 (*ΔguaB*) could restore IMPDH activity and allow the bacteria to grow on minimal media.

In order to obtain antibodies against soybean nodule IMPDH, soybean *IMPDH* cDNA was subcloned in a pET-32 vector and expressed in *E. coli*. Three different expression hosts, *E. coli* strain BL21(DE3)LysS, AD494(DE3) and AD494(DE3)LysS were used for expression at different temperatures. Under the optimal temperature, the recombinant IMPDH fusion protein accumulated in inclusion bodies. After inclusion body isolation, purification and solubilization, IMPDH fusion protein was purified by affinity chromatography and anion-exchange chromatography under denaturing conditions in the presence of 6 M urea. Following 12.5% SDS-PAGE and staining with 0.25 M KCl, the target

protein was sliced from the gel, lyophilized and ground into a fine powder. The powder was then used to immunize rabbits for antibody preparation. The molecular weight of the IMPDH fusion protein was ~70 kDa by SDS-PAGE.

Differential expression of the *IMPDH* gene in soybean was studied at both the transcriptional and translational levels by Northern blot analysis, Western blot analysis and IMPDH activity assays. Results showed that IMPDH is a nodulin, that is expressed specifically in nodules. IMPDH transcripts and protein were not detected in roots, epicotyls, buds, stems, leaves, flowers, pods and immature seeds. The expression of *IMPDH* in nodules was induced between 8 to 13 days after inoculation (DAI) with *Bradyrhizobium japonicium*. The expression level increased until 21 DAI and then remained fairly constant until 36 DAI. Two different immunoreactive bands were observed by Western blot analysis using anti-recombinant IMPDH antibody. These two bands were approximately 52 kDa and 48 kDa in size based on analysis by SDS-PAGE.

The technique of *in situ* hybridization was used to localize the *IMPDH* mRNA in soybean nodules. Mature nodules were fixed in 4% paraformaldehyde and 0.25% glutaraldehyde, dehydrated and embedded in Paraplast. Sections of 7-10 μm in thickness were prepared. A DIG-labeled antisense RNA probe, chemically hydrolyzed to a size of 150-200 bases, was used to hybridize to the pre-treated sections at 50-55°C. DIG-RNA:RNA hybrids were detected with AP-conjugated anti-DIG antibody. *IMPDH* gene expression was observed in soybean nodule cells infected with rhizobia, while uninfected cells and peripheral tissues were not labeled.

Chapter 1. Introduction to *De novo* Purine Biosynthesis and Ureide Biogenesis in Ureide-exporting Legume Nodules

1. NITROGEN FIXATION IN LEGUMES: BASIC INFORMATION

In the plant kingdom, a variety of pathogenic, saprophytic and symbiotic interactions occur between plants and microorganisms. Many of these interactions involve potent pathogens that are detrimental to the plant, while some microorganisms such as the diazotrophic bacteria and the mycorrhizal fungi enter into beneficial symbiotic associations with plants.

Nitrogen fixation by bacteria in symbiosis with plants is the most important contributor of fixed nitrogen to the nitrogen cycle and is of great agricultural importance by allowing certain plants to grow in nitrogen deficient soil (Verma and Delauney, 1988). Among the nitrogen-fixing symbioses, the interaction between rhizobia and leguminous plants is the most widely studied. This symbiosis results in the formation of specialized organs on the roots called nodules. In the nodules, the symbiotic bacteria are able to reduce atmospheric nitrogen to ammonia. The latter is assimilated by the plants and used as part of the nitrogen required for plant growth. In return, the host plant provides an environment and photoassimilates to support the growth and function of the nodule (Schubert, 1986).

Generally, three bacterial genera, *Rhizobium*, *Bradyrhizobium* and *Azorhizobium* are collectively referred to as rhizobia (van Rhijn and Vanderleyden, 1995). The formation of symbiosis between rhizobia and legume is a species-specific process and has been extensively reviewed (Verma and Delauney, 1988; Nap and Bisseling, 1990; Hennecke, 1990; Franssen, 1992; van

Rhijn and Vanderleyden, 1995; Van de Sande and Bisseling., 1997). In the following sections, some basic information and general concepts about the legume-rhizobia symbiosis will be discussed.

1.1. Formation of nodules:

The *Rhizobium*-legume symbiosis involves a complex set of interactions which result in a highly organized structure, the root nodule. During root nodule formation, infection and nodule organogenesis, take place concurrently. To infect the root, rhizobia induce root hair deformation and curling. Curled root hairs form so-called 'Shepherd's crooks' in which rhizobia are trapped within the curls. The entrapped rhizobia cause local hydrolysis of the root hair cell wall and invagination of the plasma membrane to enter the root (Van Spronsen *et al.*, 1994). Deposition by the plant cell of new cell wall material around this invagination leads to the formation of an *infection thread.* Concurrent with infection, root cortical cells dedifferentiate and start dividing to form the nodule primordium from which the nodule will develop. The infection thread filled with proliferating rhizobia pushes its way through the root hair and the cortical cell layers moving by tip growth in the direction of the nodule primordium. Primordial cells are infected by the rhizobia through an endocytotic process. During endocytosis, the invading bacteria are surrounded by a plant-derived membrane, the peribacteroid membrane (van de Sande and Bisseling, 1997).

After infection, the primordium develops into N_2-fixing root nodules composed of several different tissues. The central tissue consists of the *infected*

cells and interstitial *uninfected cells* dispersed between them. Infected cells contain the rhizobia that have differentiated into their symbiotic form, the so-called *bacteroids*. Within uninfected cells, peroxisomes are closely associated with the uptake of O_2 required for uricase activity. Therefore, the uninfected cells are considered to play a key role in the final steps of ureide biogenesis (Newcomb and Tandon, 1981). The peripheral tissues surround the central tissue and consist of the nodule vascular bundles within the nodule parenchyma and the nodule endodermis that separates the parenchyma from the nodule cortex. The uninfected cells and the peripheral tissues contribute to the proper physiological environment, allowing rhizobial N_2 fixation to occur (Mylona *et al.*, 1995; van de Sande and Bisseling, 1997).

1.2. Nodulins and nodulin genes:

The processes of nodule formation and N_2-fixation involve the expression of a set of organ (nodule)-specific genes in both the bacterium and the host plant. The plant proteins that are specifically induced or whose synthesis is enhanced during the formation and function of a root nodule are called *nodulins* (van Kammen, 1984). Leghemoglobins and uricase II (nodulin-35) are the most abundant and historically most extensively studied nodulins (Verma and Delauney, 1988). The nodulin genes that are markedly expressed before the onset of N_2 fixation are the *early nodulin genes*. These genes are presumably involved in root hair curling, infection and the formation of the nodule structure. ENOD2 (Franssen *et al.*, 1987; Dickstein *et al.*, 1988; van de Wiel *et al.*, 1990; Allen *et al.*, 1991; Dehio

and de Bruijn, 1992; Lauridsen *et al.*, 1993; Minami *et al.*, 1996; Wycoff *et al.*, 1998; Chen *et al.*, 1998), ENOD12 (Csanádi *et al.*, 1994; Vijn *et al.*, 1995; Christiansen *et al.*, 1996) and ENOD40 (Yang *et al.*, 1993; Crespi *et al.*, 1994; Vijn *et al.*, 1995b; van de Sande *et al.*, 1996; Creelman and Mullet, 1997; van de Sande and Bisseling, 1997; Corich *et al.*, 1998; Fang and Hirsch, 1998) are the most extensively studied early nodulins. The nodulin genes expressed shortly before or concomitantly with the onset of N_2 fixation are the *late nodulin genes* (van Kammen, 1984; Nap and Bisseling, 1989). Late nodulins include leghemoglobins, nitrogen-assimilatory enzymes (including glutamine synthetase, purine synthetic and oxidative enzymes), carbon metabolic enzymes and peribacteroid membrane nodulins (Verma and Delauney, 1988).

2. *DE NOVO* PURINE BIOSYNTHESIS AND UREIDE BIOSYNGENISIS

In soybean and many tropical legumes, fixed nitrogen is exported from the nodules primarily as the ureides, allantoin and allantoic acid. Atmospheric nitrogen is reduced to ammonium (NH_4^+) in the bacteroid and secreted into the plant cytosol. In the host plant, ammonium is assimilated and converted into ureides through a proposed pathway of *de novo* purine biosynthesis and ureide biogenesis. In the proposed pathway, the IMPDH-catalyzed oxidation of inosine monophosphate (IMP), the immediate product of *de novo* purine biosynthesis, leads to the production of ureides (Figure 1-1).

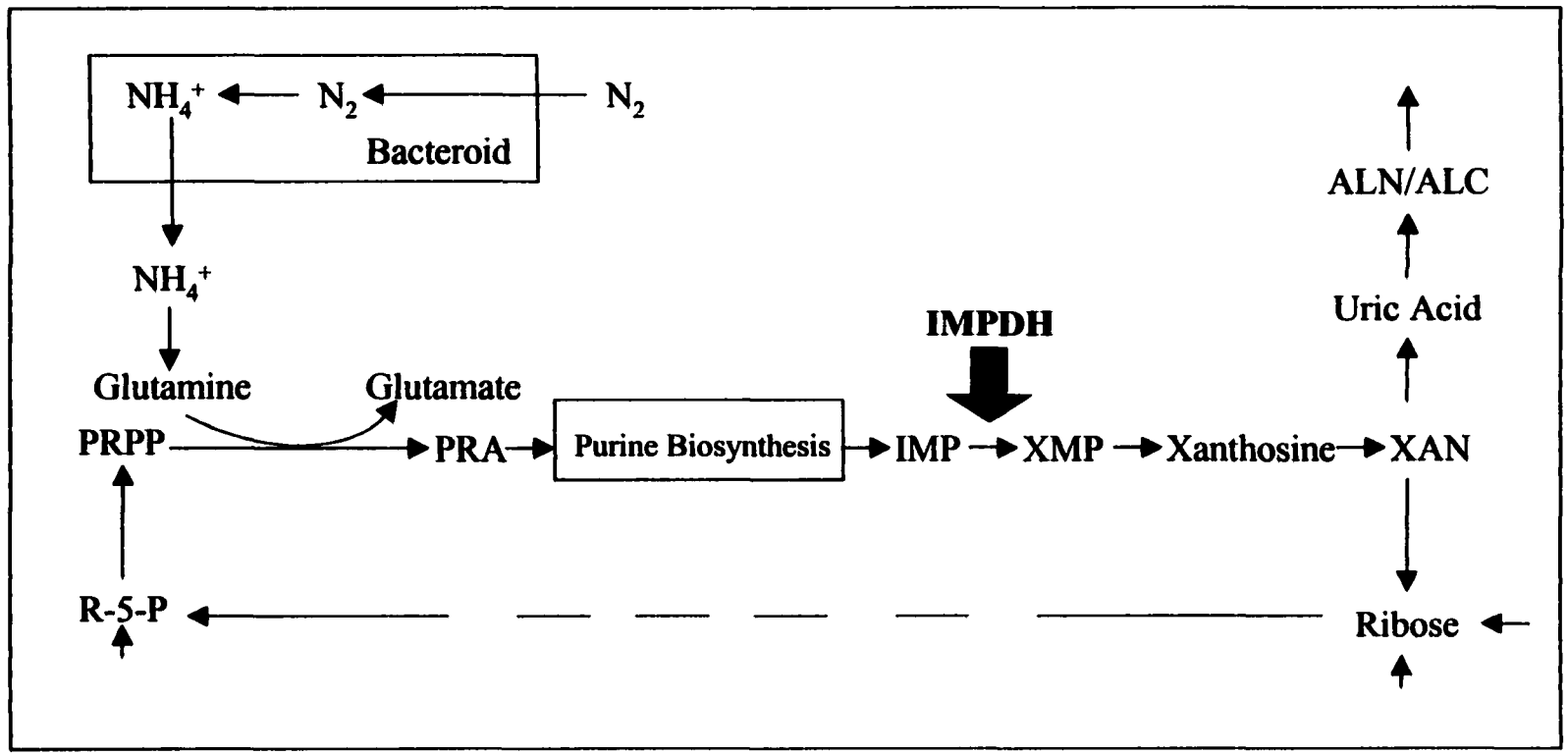


Figure 1-1. Proposed pathway for ureide biogenesis. PRPP, 5-phosphoribosyl-1-pyrophosphate; PRA, 5-phosphoribosyl-1-amine; IMP, inosine 5'-monophosphate ; XMP, xanthosine-5'-monophosphate; XAN, xanthine; ALN, allantoin; ALC, allantoic acid; R-5-P, ribose-5-phosphate. (Schubert and DeShone, 1980; Hanks *et al.*, 1981; Schubert, 1981; Schubert and Boland, 1990)

Considerable progress in understanding the process and regulation of *de novo* purine biosynthesis and ureide biogenesis in tropical legumes was made in the early 1980s and late 1990s. Studies in the 1980s principally employed biochemical techniques to investigate the enzymes involved. These techniques included tracer studies with labeled substrates ($^{15}N_2$, $^{13}N_2$, ^{14}C or ^{3}H), enzyme activity assays from different nodule fractions, inhibitor studies and enzyme purification and kinetic studies. These early studies resulted in a proposed model for purine biosynthesis and ureide biogenesis (Schubert and DeShone, 1980; Hanks *et al.*, 1981; Schubert, 1981; Schubert, 1986; Schubert and Boland, 1990). With the widespread application of the techniques of molecular biology in the 1990s, more studies on purine biosynthesis and ureide biogenesis have been carried out and several cDNA / genes for enzymes involved in these pathways have been cloned (Chapman *et al.*, 1994; Kim *et al.*, 1995; Schnorr *et al.*, 1996; Smith *et al.* 1998; Takane *et al.*, 1997a; Lee *et al.*, 1993; Kim and An, 1993; Papadopoulou *et al.*, 1995; Capote-Maínez and Sánchez, 1997). In the next sections, the model for the pathway of purine biosynthesis and ureide biogenesis will be discussed. The model includes the compartmentalization and potential sites of regulation of the pathway.

2.1. Experimental evidence for the proposed pathway of purine biosynthesis and ureide biogenesis:

Ureides (allantoin and allantoic acid), which are synthesized in the nodule, are the primary export products of recent N_2 fixation in some legumes of tropical origin.

The process of ureide formation includes assimilation of ammonia (glutamine synthetase and glutamate synthetase – dependent) followed by purine synthesis and oxidation (Schubert and Boland, 1990). Based on previous studies of the enzymes of purine biosynthesis and catabolism (Schubert and DeShone, 1980; Hanks *et al.*, 1981), a pathway for purine biosynthesis and ureide biogenesis (Figure 1-1) was proposed by Schubert (1981). Although the reaction catalyzed by PRPP synthetase (PRS) is not unique to this pathway (PRPP is also substrate in the synthesis of pyrimidine nucleotides, nicotinamide adenine dinucleotide coenzymes, histidine and tryptophan in plants), this key step along with the assimilation of ammonia and the synthesis of essential precursors for purine biosynthesis are included as part of the overall pathway (Schubert and Boland, 1990).

De novo Purine Biosynthesis:

The *de novo* purine biosynthesis pathway has been firmly established as the route for ureide biogenesis in legume nodules based on developmental studies, inhibitor studies, tracer studies and an analysis of rates of *de novo* synthesis versus turnover of purines (Schubert and Boland, 1990). Based on the pathway in animal and microbial systems, the *de novo* synthesis of the purine nucleotide IMP in ureide-exporting nodules is made up of 10 enzymatic steps (Schubert and Boland 1990; Figure 1-2). Although all of the individual steps of the pathway have not been demonstrated, evidence for the occurrence of individual steps is presented below.

Based on earlier studies, it was suggested that recently-fixed nitrogen was utilized for the *de novo* synthesis of purines that are subsequently catabolized to produce the ureides (Tajima and Yamamoto, 1975; Fujihara and Yamaguchi, 1978; Atkins *et al.*, 1980; Schubert and DeShone, 1980; Triplett *et al.*, 1980; Schubert, 1981). The activity of 5-phosphoribosyl-1-pyrophosphate (PRPP) synthetase, which catalyzes the formation of the precursor to the pathway, was examined in soybean nodule extracts (Schubert, 1981). The level of PRPP synthetase activity increased in parallel with the rates of nitrogen fixation (as estimated by C_2H_2 production) from day 18 to day 49, and then both activities declined rapidly. The close association between nitrogen fixation, PRPP synthetase activity and ureide export in soybeans supported the proposal. Phosphoribosyl-1-pyrophosphate amidotransferase (PRAT) has been detected in soybean nodule extracts and its activity increased during nodule development (Reynolds *et al.*, 1982). Evidence for the third step, catalyzed by phosphoribosyl glycinamide synthetase (GAR synthetase), relies on the incorporation of labeled glycine into purines. The fourth step, the single-carbon transfer to produce N-formyl glycinamide ribonucleotide (FGAR), and the tenth step, the single-carbon transfer to give formyl-AICAR (formyl-4-amino-5-imidazole-carboxamide ribonucleotide), have been inferred from the incorporation of label from [3-^{14}C]-serine into IMP in cell-free extracts from cowpea and soybean nodules (Atkins *et al.*, 1982; Boland and Schubert, 1983) and from the labeling of FGAR and AICAR in extracts of nodules incubated with [^{14}C] glycine. The presence of 5-aminoimidazole ribonucleotide (AIR) carboxylase, the seventh step, was

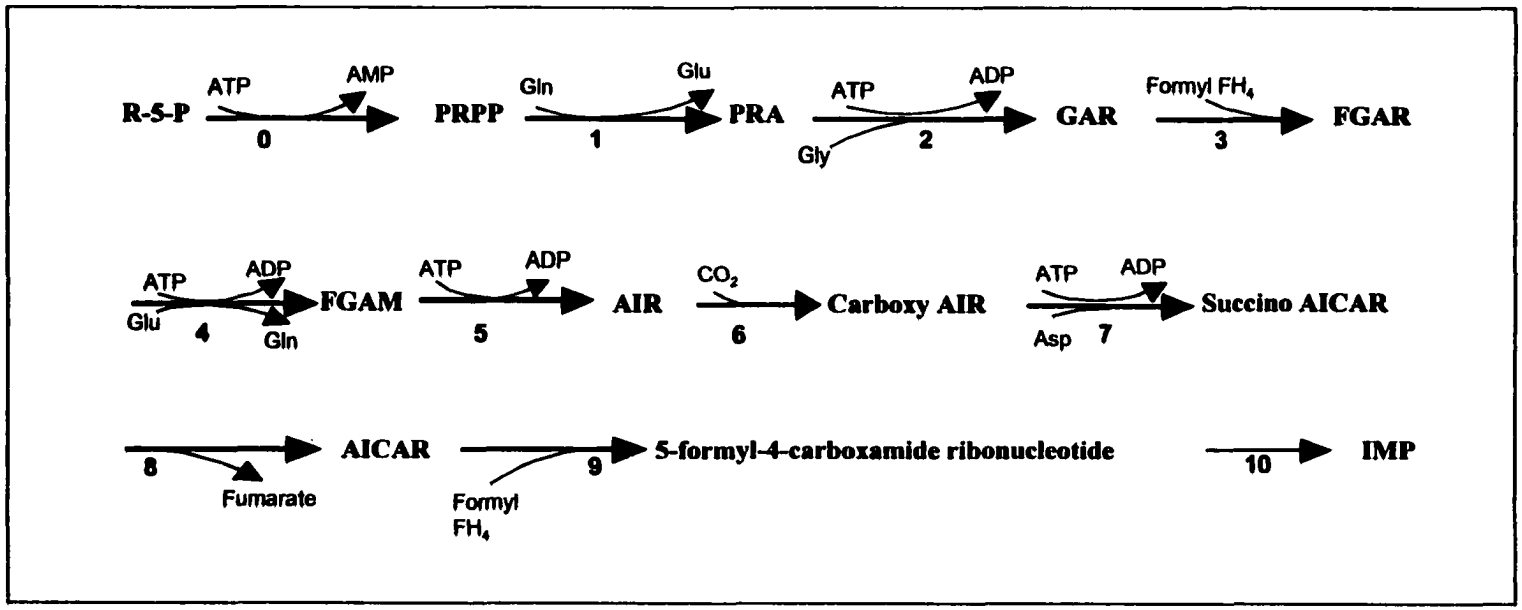


Figure 1-2. The pathway of *de novo* purine ribonucleotide biosynthesis. The enzymes catalyzing the numbered reactions are: 0) PRPP synthetase, 1) PRPP amidotransferase (PRAT), 2) GAR synthetase, 3) GAR transformylase, 4) FGAR amidotransferase, 5) AIR synthetase, 6) AIR carboxylase, 7) Succino-AICAR synthetase, 8) Adenylosuccinase, 9) AICAR transformylase, 10) IMP cyclohydrolase. *Abbrevations*: R-5-P, ribose-5-phosphate; PRPP, 5-phosphoribosyl-1-pyrophosphate; PRA, 5-phospho-ribosyl-1-amine; GAR, glycinamide ribonucleotide; FGAR, N-formylglycinamide ribonucletotide; FGAM, N-formyl-glycinamidine ribonucleotide; AIR, 5-aminoimidazole ribonucleotide; AICAR, 5-aminoimidazole-4-carboxamide ribonucleotide; IMP, inosine monophosphate (modified from Schubert and Boland, 1990)

implicated by the incorporation of label from $^{14}CO_2$ into IMP at position-6 in intact nodules (Boland and Schubert, 1982a) and into IMP/AICAR in cell-free extracts (Atkins *et al.*, 1982). Although these results confirmed that the pathway of *de novo* purine biosynthesis in nodules is similar to the pathway of purine biosynthesis in other organisms, in depth characterization of the individual enzymatic reactions in nodules and other plant tissue is missing.

Purine Oxidation and Ureide Biogenesis:

The normal pathway for purine catabolism involves removal of the phosphate and ribose moieties from IMP to produce hypoxanthine. Hypoxanthine is then successively oxidized to xanthine and uric acid by xanthine dehydrogenase (XDH). Finally, uric acid is oxidized to allantoin by uricase (Figure 1-3). In ureide-synthesizing legume nodules, an alternative pathway (Figure 1-3) has been proposed. The pathway was based on the following observations and it differs from the normal pathway in two aspects. First, the oxidation of IMP to XMP catalyzed by IMP dehydrogenase (IMPDH) was proposed as the primary route for IMP metabolism leading to ureide biogenesis in ureide-exporting legume nodules (Shelp and Atkins, 1983). The normal pathway was initially questioned because xanthine, instead of hypoxanthine, accumulated when xanthine-oxidizing activity was inhibited by allopurinol (Fujihara and Yamaguchi, 1978; Triplett *et al.*, 1980; Atkins *et al.*, 1980; Boland and Schubert, 1982a). The findings support the existence of an alternative pathway, which involves the enzyme IMPDH. IMPDH catalyzes the NAD^+-dependent conversion of IMP to XMP. In 1983, Boland and

Schubert provided the first direct evidence for the alternative pathway by demonstrating IMP dehydrogenase activity in a proplastid fraction from soybean nodules. Subsequently, Shelp and Atkins (1983) were able to detect the IMPDH activity in cowpea nodule extracts and suggested that the primary route for xanthine production involved IMPDH. Combined with results from inhibitor studies, tracer studies and IMP dehydrogenase kinetic studies, they concluded that the oxidation of IMP via IMPDH was the primary route for IMP metabolism leading to ureide biogenesis. Second, the enzyme responsible for the oxidation of xanthine was not xanthine oxidase but was a NAD^+-linked xanthine dehydrogenase (XDH) (Atkins *et al.*, 1980; Triplett *et al.*, 1980).

The oxidative enzymes involved in the conversion of purines to ureides have been well studied. These studies include the measurement of enzyme activity and determination of preliminary properties of 5'-nucleotidase (Christensen and Jochimsen, 1983) and purine nucleosidase (Christensen and Jochimsen, 1983; Atkins *et al.*, 1989). Xanthine dehydrogenase (XDH) and uricase have been purified from nodules from several species of legumes and will be discussed in later sections.

In summary, in ureide-exporting nodules, it appears that the oxidation of purines leading to the formation of the ureides occurs via this alternative pathway (Figure 1-1; 1-3). The conversion of IMP to XMP catalyzed by IMPDH is the first step of purine catabolism leading to the production of the ureides.

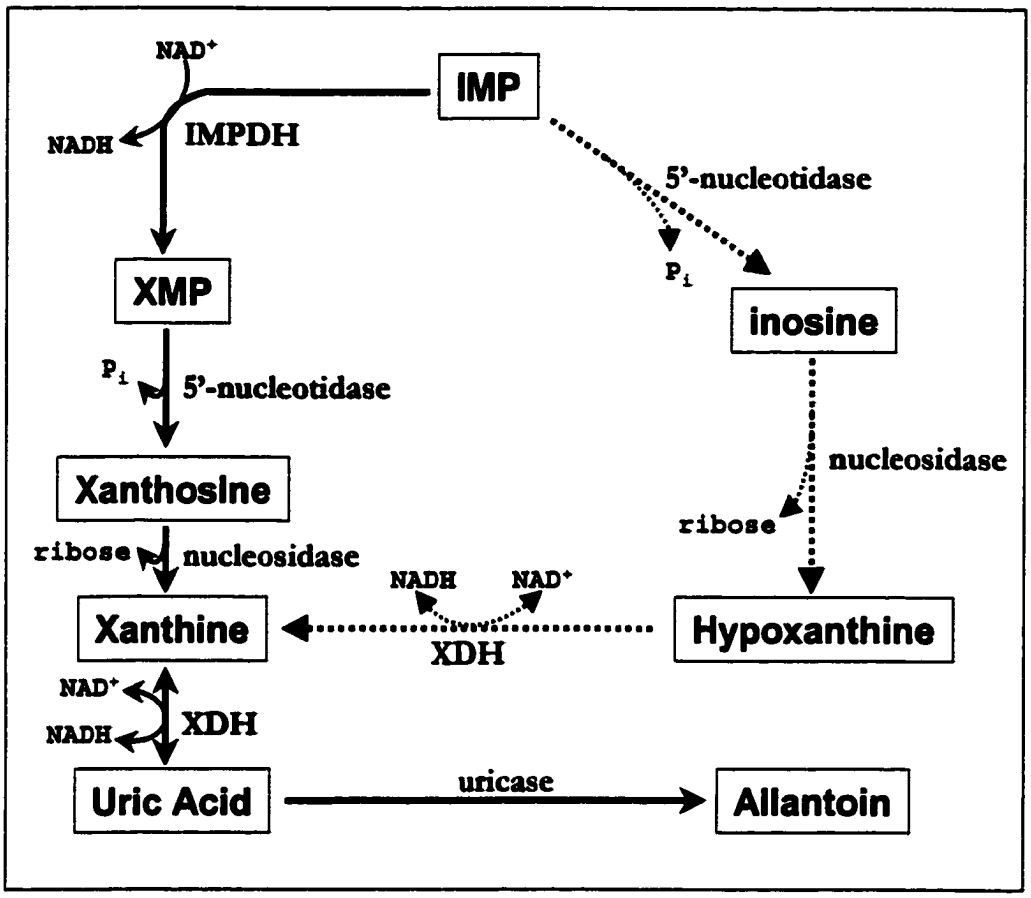


Figure 1-3. Potential pathways for the biogenesis of allantoin and allantoic acid from IMP. Experimental results suggest that an alternative pathway be involved in the oxidation of purine leading to the biogenesis of ureides. The alternative pathway goes as follows: IMP → XMP → xanthosine → xanthine → uric acid → allantoin (represented with bold lines and arrows). The normal purine oxidation pathway is represented with dashed lines. XDH, xanthine dehydrogenase.

2.2. Compartmentalization of the pathway of purine biosynthesis and ureide biogenesis:

The enzymes of ureide biogenesis and the associated reactions of ammonium assimilation and carbon metabolism have been localized to specific cells or subcellular compartments within legume nodules. Techniques for homogenizing nodule cells in an osmotically balanced medium and density gradient centrifugation to separate intact organelles were used to localize the reactions at the cellular level (Hanks *et al.*, 1981). Subsequently, Hanks *et al.* (1983) developed methods to separate different cell types. Transmission electron microscopy has been used to reveal the ultrastructure of nodules (Bergersen and Goodchild, 1973; Newcomb and Tandom, 1981; Shelp *et al.*, 1983; Selker and Newcomb, 1985; Newcomb *et al.*, 1985) and to localize the involved enzymes by immunolabeling techniques (Vandenbosh and Newcomb, 1986; Webb and Newcomb, 1987; Datta *et al.*, 1991). The cellular and subcellular compartmentalization of the ureide biogenesis pathway and the related reactions is summarized in the model presented in Figure 1-4. The localization of enzymes of *de novo* purine biosynthesis and ureide biogenesis is discussed in greater detail in the following sections.

Compartmentalization of *De novo* Purine Biosynthesis Pathway:

PRAT was found to be located in the proplastid fraction of soybean nodules (Boland *et al.*, 1982b). Based on this study, Boland *et al.* (1982b) further

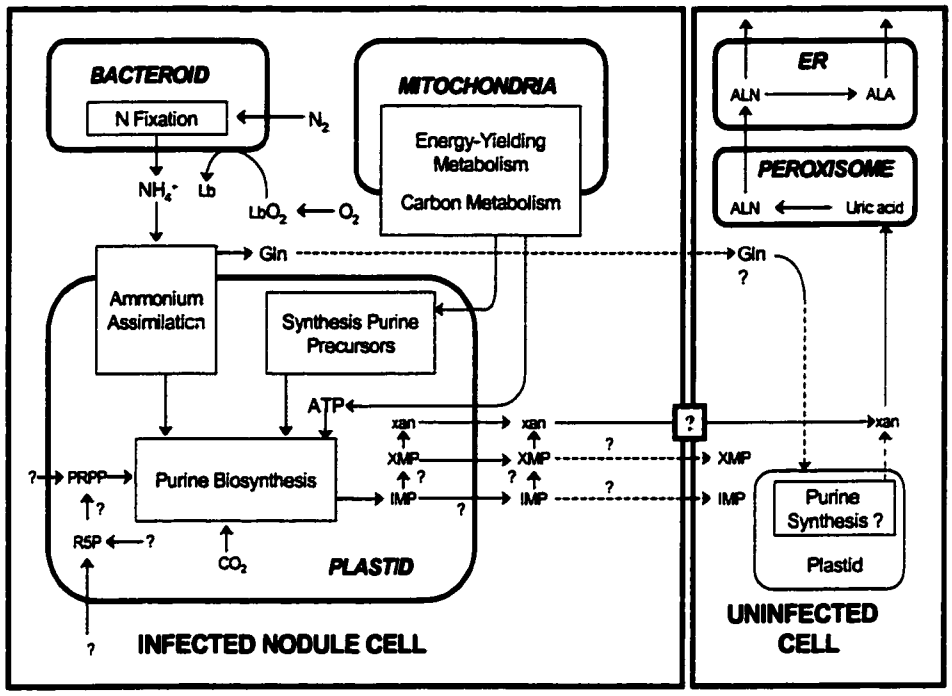


Figure 1-4. Proposed model for the cellular and subcellular compartmentalization of the reactions of nitrogen fixation, ammonium assimilation, purine synthesis and ureide biogenesis in nodules of ureide-exporting legumes. The uncertainties are indicated with question marks and/or dashed lines (modified from Schubert and Boland, 1990). Results of recent studies suggested that purine biosynthesis may take place in both plastids and mitochondria (Atkins *et al.*, 1997; Smith *et al.*, 1998). Lb, Leghemoglobin.

suggested that the purine biosynthesis pathway was localized in nodule plastids. This suggestion was subsequently supported by tracer studies using soybean (Schubert and Boland, 1982; Boland and Schubert, 1983) and cowpea nodules (Atkins and Shelp, 1983; Shelp *et al.*, 1983). PRPP synthetase, however, may exist in different subcellular compartments. According to cellular fractionation studies of Boland *et al.* (1982), PRPP synthetase was found primarily in the ground cytoplasm of soybean nodule cells. PRPP synthetase activity was not stable and very little activity remained after sucrose gradient centrifugation. In later studies, using a procedure to rapidly isolate plastids, Boland and Schubert (1983) were able to show that ribose-5-phosphate plus ATP could substitute for PRPP in plastids. In 1997, a re-examination of the intracellular localization of *de novo* purine biosynthesis in cowpea nodules (Atkins *et al.*, 1997) showed that both plastids and mitochondria have a full complement of enzymes for *de novo* purine synthesis (Atkins *et al.*, 1997). *In vitro* activities of individual component enzymes, including GAR synthetase, GAR transformylase, AIR synthetase and AICAR transformylase as well as of the whole purine pathway (from ribose-5-phosphate to IMP) were similar in the two organelles. No significant cytosolic or bacteroidal activity of any of the purine pathway enzymes was reported in this study.

In the late 1990s, molecular studies on the purine biosynthesis pathway provided more evidence for the proposed pathway and raised new questions concerning the compartmentalization of the pathway. Analysis of AIR synthetase cDNA sequence and the deduced amino acid sequence (Smith *et al.*, 1998)

showed that a signal sequence encoded by this gene had properties associated with plastid transit sequences but there was no consensus cleavage site as would be expected for a plastid targeted protein. Although the signal sequence does not have the structural features of a mitochondrial-targeted protein, it has a mitochondrial cleavage site motif close to the predicted N-terminus of the mature protein. AIR synthetase is apparently encoded by a single-copy gene, which raises the question as to how the product of this gene is targeted to the two organelles.

Compartmentalization of Enzymes of Purine Oxidation and Ureide Biosynthesis: Ultrastructural studies showed that the central region of nodules of N_2-fixing plants contains both densely cytoplasmic, infected cells and highly vacuolated, uninfected cells (Bergersen and Goodchild, 1973; Newcomb and Tandom, 1981; Shelp *et al.*, 1983; Selker and Newcomb, 1985; Newcomb *et al.*, 1985). Previous studies had suggested that infected cells were the primary site of ammonium assimilation and purine synthesis (Hanks *et al.*, 1983; Shelp *et al.*, 1983; Schubert and Boland, 1984; Schubert, 1986). The last step of ureide biogenesis, catalyzed by uricase, was found primarily in uninfected cells (Bergmann *et al.*, 1983; Hanks *et al.*, 1983; Van den Bosch and Newcomb, 1986). This difference in cellular localization between the first step and the last step of the purine biosynthesis and ureide biogenesis pathway raised a question as to where the intermediate steps take place.

IMPDH, the enzyme that catalyzes the first step of purine oxidation, was found in plastids of soybean nodules (Boland and Schubert, 1983). In contrast, Shelp and Atkins (1983) found that IMPDH was in the ground cytoplasm of cowpea nodules. Resolution of these apparent differences awaits immunocytochemical examination. At the cellular level, it is still not clear whether or not the oxidation of IMP to XMP was carried out in infected cells or in uninfected cells.

XDH was reported to be present only in soybean nodule infected cells using histochemical staining using microscopy (Triplett, 1985). Nguyen *et al.* (1986) found that XDH was localized in the uninfected cells by immunofluorescence. In 1991, immunocytochemical studies by Datta *et al.* (1991) showed that significantly more XDH labeling was observed in the uninfected cells than in infected cells. They also found that XDH was soluble and not present in any organelle or membrane. They suggested that xanthine or a precursor to xanthine (IMP, XMP or xanthosine), rather than uric acid, was the intermediate that moves from infected to uninfected cells during ureide biogenesis.

Uricase was shown to exist exclusively in the enlarged peroxisomes of uninfected nodule cells by enzyme activity studies, cell fractionation, histochemical and immunocytochemical observations (Bergmann *et al.*, 1983; Nguyen *et al.*, 1985; VandenBosch and Newcomb, 1986; Web and Newcomb, 1987). Uricase II transcripts were detected only in the uninfected cells of the central nodule tissue and mainly in the periphery of the cell by *in situ*

hybridization (Papadopoulou *et al.*, 1995). This is consistent with the previous conclusion.

3. MOLECULAR BIOLOGY AND BIOCHEMISTY OF ENZYMES INVOLVED IN *DE NOVO* PURINE BIOSYNTHESIS AND UREIDE BIOGENESIS

Studies on the pathway of *de novo* purine biosynthesis and ureide biogenesis in the 1990s have focused on cloning genes for enzymes involved in the pathways and studying their expression. Up to now, six cDNAs coding for purine biosynthetic enzymes in N_2-fixing nodules have been cloned. These enzymes include: AIR carboxylase and SAICAR synthetase in mothbean (Chapman *et al.*, 1994), PRAT in soybean and mothbean (Kim *et al.*, 1995), GAR synthetase and GAR transformylase (Schnorr *et al.*, 1996) and AIR synthetase in cowpea (Smith *et al.*, 1998). Of the purine oxidative enzymes, uricase II is the only one that has been cloned. The isolation of cDNA and / or genomic clones for uricase II from nodules of soybean (Bergman *et al.*, 1983; Nguyen *et al.*, 1985), cowpea (Lee *et al.*, 1993), jackbean (Kim and An, 1993) and bean (Sánchez *et al.*, 1987; *Papadopoulou et al.*, 1995) has been reported. No protein purification of purine biosynthetic enzymes has been reported. Proteins of three purine oxidative enzymes, IMPDH (Yang and Schubert, 1997), XDH (Boland, 1981; Triplett *et al.*, 1982; Boland *et al.*, 1983c) and allantoinase (Bell and Webb, 1995) were purified. Molecular and biochemical studies on enzymes involved in ureide biogenesis are summarized in Table 1-1. Results of these studies are presented herein.

3.1. PRAT:

PRAT catalyzes the first committed step in the *de novo* purine biosynthesis pathway. PRAT activity was first detected in developing soybean nodules ~12 days after infection reaching a peak after ~24 days. The induction of PRAT activity occurred in parallel with the development of nitrogen fixation activity (Reynolds *et al.*, 1982b). A cDNA clone for PRAT has been isolated from soybean and mothbean nodules (Kim *et al.*, 1995). In mothbean, significant expression of PRAT mRNA occurred prior to the onset of nitrogen fixation. The PRAT mRNA levels were high in fix$^-$ mutant nodules (nodules incapable of nitrogen fixation). These results suggested that the regulation of the purine synthesis pathway is independent of nitrogen fixation. The expression of the PRAT gene in mothbean was induced by treatment of nodules with 10 mM L-glutamine, suggesting that the glutamine produced during nitrogen fixation may constitute one of the signals for activating PRAT gene expression. The increased levels of PRAT mRNA in nodules suggest that one of the mechanisms of controling *de novo* purine biosynthesis in legume nodules involves transcriptional regulation.

3.2. GAR synthetase and GAR transformylase:

Glycinamide ribonucleotide (GAR) synthetase and GAR transformylase catalyze the second and third steps of the *de novo* purine biosynthetic pathway. cDNAs encoding these two enzymes (*GMpurD* and *GMpurN* respectively) in soybean have been isolated, and the expression of these genes has been studied (Schnorr *et*

al., 1996). The cDNAs were cloned by functional complementation of corresponding *E. coli purD* and *purN* mutant strains. One class of GAR synthetase and three classes of GAR transformylase cDNA clones were identified. Northern blot analysis showed that these purine biosynthetic genes were highly expressed in young (10-day-old) and mature (28-day-old) nodules and weakly expressed in roots and leaves. Levels of *GMpurD* and *GMpurN* mRNAs were not enhanced when ammonia was provided to uninfected roots.

3.3. AIR synthetase:

A cDNA clone coding for aminoimidazole (AIR) synthetase, the fifth enzyme of the *de novo* purine biosynthetic pathway, has been isolated from a cowpea nodule cDNA library by Smith *et al.* (1998). It encodes a 388 amino acid protein with a predicted molecular mass of 40.4 kDa. This gene is expressed at much higher levels in nodules than in other tissues. Because expression was detected in immature 12-day-old nodules, it is likely that the mRNA for the AIR synthetase gene (*VUpur5*) is present before nitrogenase is active. This pattern of gene expression is similar to the pattern of expression of PRAT, GAR synthetase and GAR transformylase (Chapman *et al.*, 1994; Kim *et al.*, 1995; Schnorr *et al.*, 1996) because these genes are all expressed before nitrogen fixation begins. Smith *et al.* (1998) suggested that more than one factor is involved in the enhanced

Table 1-1. Summary of molecular and biochemical studies on enzymes involved in *de novo* purine biosynthesis and ureide biogenesis in legume nodules

Enzyme Name	Plant Source	Protein Purified	cDNA / Gene Cloned	Reference
Purine Biosynthesis				
PRS	--	--	--	--
PRAT	soybean	--	cDNA	Kim *et al.*, 1995
	mothbean	--	cDNA	Kim *et al.*, 1995
GAR synthetase	soybean	--	cDNA	Schnorr *et al.*, 1996
GAR transformylase	soybean	--	cDNA	Schnorr *et al.*, 1996
FGAR amidotransferase	--	--	--	--
AIR synthetase	cowpea	--	cDNA	Smith *et al.*, 1998
AIR carboxylase	mothbean	--	cDNA	Chapman *et al.*, 1994
SAIR synthetase	mothbean	--	cDNA	Chapman *et al.*, 1994
Adenylosuccinase	--	--	--	--
AICAR transformylase	--	--	--	--
IMP cyclohydrolase	--	--	--	--
Purine Oxidation				
IMPDH	soybean	purified	--	Yang & Schubert, 1997
	cowpea	purified	--	Atkins *et al.*, 1985
XDH	bean	purified	--	Boland, 1981
	soybean	purified	--	Triplett *et al.*, 1982
Uricase II	soybean	purified	--	Legocki & Verma, 1979
		purified	--	Bergmann *et al.*, 1983
		purified	--	Lucas *et al.*, 1983
		--	cDNA /gene	Nguyen *et al.*, 1985
	cowpea	purified	--	Rainbird & Atkins, 1981
		--	cDNA	Lee *et al.*, 1993
	jackbean	--	cDNA	Kim & An, 1993
	bean	purified	cDNA	Sánchez *et al.*, 1987
			cDNA	Papadopoulou *et al.*, 1995
Allantoinase	soybean	purified		Bell and Webb, 1995

expression of these genes. Analysis of the regulation of transcription of these genes may help to identify the linkages between the processes of N_2 fixation and purine biosynthesis.

3.4. AIR carboxylase and SAICAR synthetase:

cDNAs encoding these two enzymes were cloned from *Vigna aconitifolia* by functional complementation (Chapman *et al.*, 1994). SAICAR synthetase mRNA was first detectable in 19-day-old nodules, suggesting that this gene was induced late in nodule development and after nitrogen fixation has begun. Investigations on the expression of AIR carboxylase have not been reported. Even though screening for cDNA clones by complementation of *E. coli* mutations should provide full-length clones encoding functional protein (Murray and Smith, 1996), the isolated mothbean clones for PRAT, AIR carboxylase and SAICAR synthetase lack sequences corresponding to the extreme N-terminus of the encoded proteins.

3.5. Inosine monophosphate dehydrogenase (IMPDH):

IMPDH from nodules of *Vigna unguiculata* has been purified 140-fold (Atkins *et al.*, 1985). Gel filtration chromatography indicated a molecular weight of 200 kDa and SDS-gel electrophoresis revealed a single subunit of 50 kDa. In soybean, a 289-fold purification of IMPDH from nodules has been reported (Yang, 1997). The soybean IMPDH has a molecular weight of 260 kDa and two bands of 55 and 60 kDa were revealed after SDS-PAGE of the purified enzyme.

The pI of the purified soybean nodule IMPDH was estimated to be 6.23 by IEF. Neither of the soybean enzyme nor the cowpea enzyme has been sequenced. To date, cDNA or genomic clones coding for nodule IMPDH have not been isolated.

3.6. Xanthine dehydrogenase (XDH):

XDH catalyzes the oxidation of hypoxanthine to xanthine and then to uric acid under microaerophilic conditions. The enzyme has been purified from *Phaseolus vulgaris* (Boland, 1981) and soybean (Triplett *et al.*, 1982; Boland *et al.*, 1983c). Soybean XDH showed antigenic cross-reactivity with nodule extracts from two other ureide producers, cowpea and lima bean but not from the amide producers alfalfa and lupin (Triplett, 1985). However, XDH may not be totally nodule-specific since low levels of the enzyme were immunologically detectable in soybean leaves, roots, stems and pods (Triplett, 1985). The isolation of a legume XDH cDNA or genomic clone has not been reported.

3.7. Uricase (urate oxidase):

Uricase catalyzes the oxidative cleavage of uric acid to ALN and CO_2, liberating H_2O_2. The nodule-specific uricase, uricase II, is the second most abundant nodulin in soybean nodules and has been the subject of a number of reviews (Verma *et al.*, 1986; Verma and Delaunery, 1988; Nap and Bisseling, 1989; Sánchez *et al.*, 1991). Uricase II has a native molecular weight of approximately 125 kDa and is composed of four 35-kDa subunits. Based on the subunit size, uricase II is referred to as nodulin-35. Nodulin-35 has been shown to be nodule-

specific and is believed to be responsible for the large increase in uricase activity which occurs in nodules following the onset of nitrogenase activity (Bergmann *et al.*, 1983).

Another uricase enzyme, designated as uricase I, catabolizes uric acid as part of the purine degradation pathway. Uricase I has been detected in uninfected roots and leaves and, therefore, has been considered to be ubiquitous in living organisms (Tajima and Yomomato, 1975; Bergmann *et al.*, 1983). The antibodies against uricase II showed no immunological cross-reactivity with neither uricase I nor with any other protein in uninfected roots and leaves. Confirming the nodule-specificity of uricase II, a nodulin-35 cDNA clone hybridized only to nodule RNA but not to RNA from roots or leaves (Nguyen *et al.*, 1985). Uricase II is preferentially synthesized on free polysomes during nodule development and is localized in the peroxisomes of uninfected cells of the nodules (Hanks *et al.*, 1981; Nguyen *et al.*, 1985; Van den Bosch and Newcomb, 1986; Kouchi *et al.*, 1989). Events during nodule development (e.g. nodule induction, bacteroid release, and N_2 fixation) are considered to be important for the regulation of uricase II gene expression (Padilla *et al.*, 1991).

Updated evidence from molecular biology studies demonstrated that uricase II, formerly considered to be exclusively present in nodules, was also expressed in non-symbiotic organs (Capote-Maínez and Sánchez, 1997; Takane *et al.*, 1997a, 1997b). In soybean, uricase II cDNA isolated from a cotyledon library was demonstrated to be identical to uricase II cDNA from nodules. Transcripts of uricase II were present in developing cotyledons and uricase II accumulated

during the pod-filling stage (Takane *et al.*, 1997a). In bean, uricase II was also expressed in other organs, such as roots, stems and leaves (Capote-Maínez and Sánchez, 1997). Uricase II expression is probably modulated when the plant needs to mobilize reduced N_2 (in the form of ureides), i.e. during seedling establishment (Capote-Maínez and Sánchez, 1997).

Isolation of two distinct uricase II (nodulin-35) clones, *UR2* and *UR9* from a soybean genomic library and a nodule cDNA library was reported (Takane *et al.*, 1997b). *UR9* behaves as a nodulin gene, whereas *UR2* appears to be a nonsymbiotic uricase II gene. *UR2* contains a postulated non-symbiotic motif GAAGAG in its promoter. Sequence comparisons of the promoter regions of *UR9* and *UR2* suggested that symbiotic and non-symbiotic uricase II diverge by gene duplication and that relatively small alterations in the promoter sequence enable the nodule-specific expression (Takane *et al.*, 1997b).

Up to now, uricase II has been purified from nodules of cowpea (Rainbird and Atkins, 1981), soybean (Legocki and Verma, 1979; Bergmann *et al.*, 1983; Lucas *et al.*, 1983) and bean (Sánchez *et al.*, 1987). Uricase II cDNA clones have been isolated from soybean nodule libraries (Bergman *et al.*, 1983; Nguyen *et al.*, 1985), soybean cotyledon libraries (Takane *et al.*, 1997a), cowpea (Lee *et al.*, 1993), jackbean (*Canavalia liniata*) (Kim and An, 1993) and bean (Sánchez *et al.*, 1987; Papadopoulou *et al.*, 1995) nodule libraries. Southern analysis suggested that the genes for uricase in soybean (*Glycine max*) and *Canavalia liniata* might be present as a small gene family (Nguyen et al., 1985; Kim and An, 1993).

However, in common bean (*Phaseolus vulgairs*), uricase II has been shown to be encoded by a single copy gene (Capote-Maínez and Sánchez, 1997).

3.8. Allantoinase (allantoin amidohydrolase):

This enzyme catalyzes the conversion of allantoin to allantoic acid in the final step of ureide biogenesis. Allantoinase has been purified from soybean seeds (Webb and Lindell, 1993), cotyledons and root nodules (Bell and Webb, 1995). Comparison of the allantoinase amino acid sequences from cotyledons and nodules showed some differences, suggesting that a nodule-specific form is present in soybean (Bell and Webb, 1995). A cDNA or the gene coding for this enzyme has not been cloned.

4. PROPOSED STUDIES ON IMPDH

IMP is the direct product of *de novo* purine biosynthesis. IMPDH catalyzes the oxidation of IMP into XMP with the conversion of NAD^+ to NADH. This reaction is the rate-limiting step in the *de novo* biosynthesis of guanine monophosphate (GMP). GMP can be converted into GTP and plays a central role in cell growth and differentiation processes, such as DNA and RNA biosynthesis, signal transduction, and tubulin assembly. Thus, IMPDH activity regulates cellular guanine levels and consequently affects cell proliferation (Weber, 1983; Kerr and Hedstrom, 1997).

In soybean and many other ureide-exporting plants, nitrogen is transported into other parts of the plant in the form of the ureides. Fixed nitrogen is a major

factor limiting plant growth. The IMPDH-catalyzed oxidation of IMP into XMP was postulated to be the first step of purine metabolism in soybean nodules leading to ureide production (Schubert, 1981, 1986; Schubert and Boland, 1990), and as such, IMPDH is an important player in the regulation of the ureide biogenesis pathway.

Unfortunately, studies on plant nodule IMPDH have been very limited. It is still not clear whether IMPDH is localized in nodule infected cells or in uninfected cells, in plastids or in the cytoplasm (Figure 1-4). Although IMPDH has been isolated from the nodules of *Vigna unguiculata* (Atkins *et al.*, 1985) and *Glycine max* (Yang, 1997), the protein has not been sequenced. A cDNA or a gene coding for IMPDH has not been cloned from any plant nodules. The regulation of IMPDH gene expression as related to purine biosynthesis and ureide biogenesis has not been addressed. Tackling these issues could help clarify the compartmentalization and regulation of the ureide biogenesis pathway and to better understand the processes of nitrogen fixation, ammonium assimilation, purine biosynthesis and ureide biogenesis.

The objectives of this research were: 1). To clone the *IMPDH* cDNA from soybean nodules; 2). To characterize the cloned soybean nodule *IMPDH* cDNA; 3). To confirm the identity of IMPDH by expression of the putative *IMPDH* cDNA in the *E. coli ΔguaB* mutant (functional complementation); 4). To purify the soybean IMPDH expressed in *E. coli*, and to raise antibodies against the expressed protein; 5). To study the spatial and temporal patterns of expression of IMPDH in soybean nodules; 6). To localize the *IMPDH* mRNA and protein in

soybean nodules. The proposed studies on soybean nodule IMPDH is outlined in Figure 1-5.

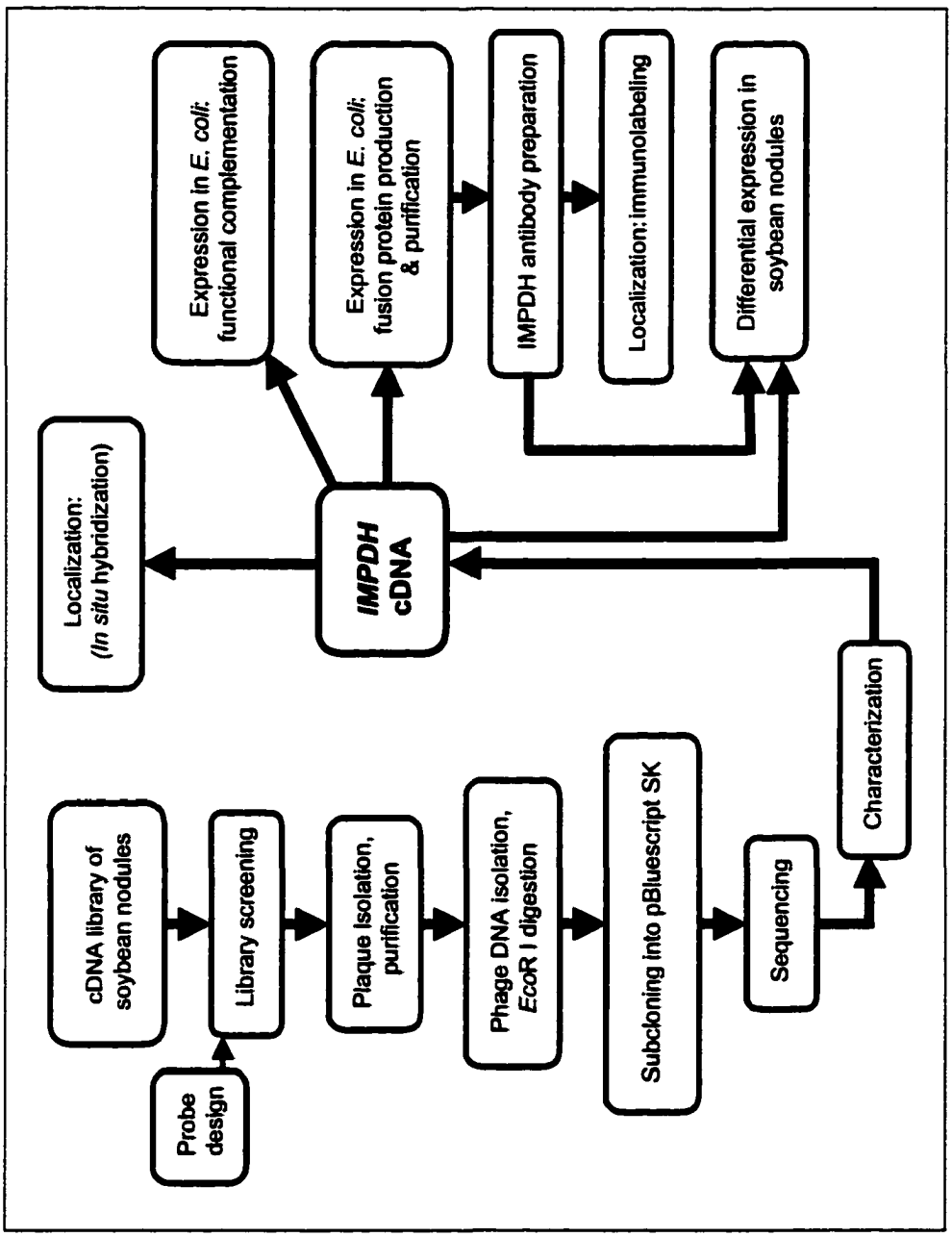


Figure 1-5. Outline of the proposed studies on soybean nodule IMPDH.

REFERENCES

Atkins, C.A., Rainbird, R. and Pate, J.S. (1980). Evidence for a purine pathway of ureide synthesis in N_2-fixing nodules of cowpea (*Vigna unguiculata* L. Walp. J.). Z. Pflanzenphysiol 97: 249-260.

Atkins, C.A., Ritchie, A., Rowe, P.B., McCairns, E. and Sauer, D. (1982). *De novo* purine synthesis in nitrogen-fixing nodules of cowpea cowpea (*Vigna unguiculata* L. Walp) and soybean (*Glycine max* Merr.). Plant Physiol 70: 55-60.

Atkins, C.A., Shelp, B.J. and Storer, P.J. (1985). Purification and properties of inosine monophosphate oxidoreductase from nitrogen-fixing nodules of cowpea (*Vigna unguiculata* L. Walp). Arch Biochem Biophys 236: 807-814.

Atkins, C.A., Storer, P.J. and Shelp, B.J. (1989). Purification and properties of purine nucleotidase from N_2-fixing nodules of cowpea (*Vigna unguiculata*). J Plant Physiol 134:447-452.

Atkins, C.A., Smith, P.M.C. and Storer, P.J. (1997). Re-examination of the intracellular localization of *de novo* purine biosynthesis in cowpea nodules. Plant Physiol 113: 127-135.

Allen, T., Raja, S. and Dunn, K. (1991). Cells expressing ENOD2 show differential spatial organization during the development of alfalfa root nodules. Mol Plant Microb Interact 4: 139-146.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struh, K. (1991). Current protocols in molecular biology. V1, 2, John Wiley and Sons, Inc., New York.

Bell, J.A. and Webb, M.A. (1995). Immunoaffinity purification and comparison of allantoinases from soybean root nodules and cotyledons. Plant Physiol 107: 435-441.

Bergmann, H., Preddie, E. and Verma, D.P.S. (1983). Nodulin-35: a subunit of specific uricase (uricase II) induced and localized in the uninfected cells of soybean nodules. EMBO J 2: 2333-2339.

Bergersen, F.J. and Goodchild, D.J. (1973a). Aeration pathways in soybean root nodules. Aust J Biol Sci 26:729-740.

Bergersen, F.J. and Goodchild, D.J. (1973b). Cellular location and concentration of leghaemoglobin in soybean root nodules. Aust J Biol Sci 26:741-756.

Boland, M.J. (1981). NAD^+: Xanthine dehydrogenase from nodules of navy beans: partial purification and properties. Biochem Int 2:567-574.

Boland, M.J. and Schubert, K.R. (1982a). Purine biosynthesis and catabolism in soybean root nodules: incorporation of $[^{14}C]CO_2$ into xanthine. Arch Biochem Biophys 213: 486-491.

Boland, M.J. and Schubert, K.R. (1982b). Purine biosynthesis by proplastids in soybean nodules. Plant Physiol 69S: 112.

Boland, M.J. and Schubert, K.R. (1983). Biosynthesis of purines by proplastid fraction from soybean nodules. Arch Biochem Biophys 220: 179-187.

Boland, M.J. and Schubert, K.R. (1983b). Phosphoglycerate dehydrogenase from soybean nodules: partial purification and some kinetic properties. Plant Physiol 71: 658-661.

Boland, M.J., Blevins, D.G. and Randall, D.D. (1983c). Soybean nodule xanthine dehydrogenase: a kinetic study. Arch Biochem Biophys 222(2): 435-441.

Chen, R., Silver, D.L. and de Bruijn, F.J. (1998). Nodule parenchyma-specific expression of the *Sesbania rostrata* early nodulin gene SrEnod2 is mediated by its 3' untranslated region. Plant Cell 10: 1585-1602.

Christensen, T.M.I.E. and Jochimsen, B.U. (1983). Enzymes of ureide synthesis in pea and soybean. Plant Physiol 72: 56-59.

Capote-Maínez, N. and Sánchez, F. (1997). Characterization of the common bean uricase II and its expression in organs other than nodules. Plant Physiol 115: 1307-1317.

Chapman, K.A., Delauney, A.J., Kim, J.H. and Verma, D.P.S. (1994). Structural organization of *de novo* purine biosynthesis enzymes in plants: 5-aminoimidazole ribonucleotide carboxylase and 5-aminoimidazole-4-*N*-succinocarboxamide ribonucleotide synthetase cDNAs from *Vigna aconitifolia*. Plant Mol Biol 24: 389-395.

Christiansen, H., Hansen, A.C., Vijn, I., Pallisgaard, N., Larsen, K., Yang, W.-C., Bisseling, T., Marcker, K.A. and Jensen, E. (1996). A novel type of DNA-binding protein interacts with a conserved sequence in an early nodulin ENOD12 promoter. Plant Mol Biol 32: 809-821.

Collart, F.R., Osipiuk, J., Trent, J., Olsen,G and Huberman, E. (1996). Cloning and characterization of the gene encoding IMP dehydrogenase from *Arabidopsis thaliana*. Gene 174: 217-220.

Corich, V., Goormachtig, S., Lievens, S., van Montagu, M. and Holsters, M. (1998). Patterns of ENOD40 gene expression in stem-borne nodules of *Sesbania rostrata*. Plant Mol Biol 37: 67-76.

Creelman, R.A. and Mullet, J.E. (1997). Oligosaccharins, brassinolides, and jasmonates: nontraditional regulators of plant growth, development, and gene expression. Plant Cell 9: 1211-1223.

Crespi, M.D., Jurkevitch, E., Poiret, M., d'Aubenton-Carafa, Y., Petrovics, G., Kondorosi, E. and Kondorosi, A. (1994). Enod 40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. EMBO J 13: 5099-1112.

Csanádi, G., Szécsi, J., Kaló, P., Kiss, P., Endre, G., Kondorosi, A., Kondorosi, E. and Kiss, G. (1994). ENOD12, an early nodulin gene, is not required for nodule formation and efficient nitrogen fixation in alfalfa. Plant Cell 6: 201-213.

Datta, D.B., Triplett, E.W. and Newcomb, E.H. (1991). Localization of xanthine dehydrogenase in cowpea root nodules: implication for the interaction between cellular compartments during ureide biogenesis. Proc Natl Acad Sci USA 88: 4700-4702.

Dehio, C. and Bruijn, F.J. (1992). The early nodulin gene SrEnod2 from *Sesbania rostrata* is inducible by cytokinin. Plant J 2: 117-128.

Dickstein, R., Bisseling, T., Reinhold, V.N. and Ausubel, F.M. (1988). Expression of nodule-specific genes in alfalfa root nodules blocked at an early stage of development. Genes Dev 2: 677-687.

Fang, Y. and Hirsch, A.M. (1998). Studying early nodulin gene ENOD40 expression and induction by nodulation factor and cytokinin in transgenic alfalfa. Plant Physiol 116: 53-68.

Franssen, H.J., Nap, J.-P., Gloudemans, T., Stiekema, W., van Dam, H., Govers, F., Louwerse, J., van Kammen, A. and Bisseling, T. (1987). Characterization of cDNA for nodulin-75 of soybean: a gene product involved in early stages of root nodule development. Proc Natl Acad Sci USA 84: 4495-4499.

Franssen, H.J., Thompson, D.V., Idler, K., Kormelink, R., van Kammen, A. and Bisseling, T. (1989). Nucleotide sequence of two soybean ENOD2 early nodulin genes encoding Ngm-75. Plant Mol Biol 14: 103-106.

Franssen, H.J., Vijn, I., Yang, W.C. and Bisseling, T. (1992). Developmental aspects of the *rhizobium*-legume symbiosis. Plant Mol Biol 19: 89-107.

Fujihara, S., Yamaguchi, M. (1978). Effects of allopurinol [4-hydroxypyrazolo (3,4-d) pyrimidine] on the metabolism of allantoin in soybean plants. Plant Physiol 62: 134-138.

Fuller, F., Kunstener, P.W., Nguyen, T. and Verma, D.P.S. (1983). Soybean nodulin genes: analysis of cDNA clones reveals several major tissue-specific sequences in nitrogen-fixing root nodules. Proc Natl Acad Sci USA 80: 2594-2598.

Hanks, J.F., Tolbert, N.E. and Schubert, K.R. (1981). Localization of ureide biosynthesis in peroxidase and microsomes of nodules. Plant Physiol 68: 65-69.

Hanks , J.F., Schubert, K.R. and Tolbert, N.E. (1983). Isolation and characterization of infected and uninfected cells from soybean nodules. Role of uninfected cells in ureide synthesis. Plant Physiol 71: 869-873.

Hennecke, H. (1990). Nitrogen fixation genes involved in the *Bradyrhizobium japoicaum*-soybean symbiosis. FEBS Lett 268: 422-426.

Kerr, K.M. and Hedstrom, L. (1997). The roles of conserved carboxylate residues in IMP dehydrogenase and identification of a transition state analog. Biochemistry 36: 13365-13373.

Kim, H.J., Delaunery, A.J. and Verma, D.P.S. (1995). Control of *de novo* purine biosynthesis genes in ureide-producing legumes: induction of glutamine phosphoribosylphosphate amidotransferase gene and characterization of its cDNA from soybean and *Vigna*. The Plant J 7(1): 77-86.

Kim, H.B. and An, C.S. (1993). Nucleotide sequence of expression of cDNA clones encoding uricase II in *Canavalia lineata*. Korean J Biol 36: 415-423.

Kouchi, H., Tsukamoto, M. and Tajima, S. (1993). Differential expression of nodule-specific (nodulin) genes in the infected, uninfected and cortical cells of soybean root nodules. J Plant Physiol 135: 608-617.

Lauridsen, P., Franssen, H., Stougaard, J., Bisseling, T. and Marcker, K.A. (1993). Conserved regulation of the soybean early nodulin ENOD2 gene promoter in determinate and indeterminate transgenic root nodules. Plant J 3: 483-492.

Lee, N., Stein, B., Suzuki, H. and Verma, D.P.S. (1993). Expression of antisense noduliln-35 RNA in *Vigna aconitifolia* transgenic root nodules retards peroxisome development and affects nitrogen availability to the plants. Plant J 3: 599-606.

Legocki, R.P. and Verma, D.P.S. (1979). A nodule specific plant protein (Nodulin-35) from soybean. Science 205: 190-192.

Lucas, K., Boland, M.J. and Schubert, K.R. (1983). Uricase from soybean root nodules: purification, properties and comparison with the enzyme from cowpea. Arch Biochem Biophys 226: 190-197.

Minami, E., Kouchi, H., Carlson, R.W., Cohn, J.R., Kolli, V.K., Day, R.B., Ogawa, T. and Stacey, G. (1996). Cooperative action of lipo-chitin nodulation signals on the induction of the early nodulin, ENOD2, in soybean roots. Mol Plant Microbe Interact 9: 574-583.

Murray, J.A.H. and Smith, A.G. (1996). Functional complementation in yeast and *E. coli*. In: Plant gene isolation (G.D. Foster and D. Twell Ed.). Jonh Wiley & Sons Ltd., Chichester, UK.

Mylona, P., Pawlowski, K. and Bisseling, T. (1995). Symbiotic nitrogen fixation. Plant Cell 7: 869-885.

Nap, J.P. and Bisseling, T. (1989). Nodulin function and nodulin gene regulation in root nodule development. In: The molecular biology of symbiotic nitrogen fixation (P.M. Gresshoff Ed.), pp181-229. CRC Press, Boca Raton, FL.

Nap, J.P. and Bisseling, T. (1990). Developmental biology of a plant-prokaryote symbiosis: the legume root nodule. Science 250 (16): 948-954.

Newcomb, E.H. and Tandon, S.R. (1981). Uninfected cells of soybean root nodules: ultrastructure suggests key roles in ureide production. Science 212 (19): 1394-1396.

Newcomb, E.H., Tandon, S.R. and Kowal, R.R. (1985). Ultrastructural specialization for ureide production in uninfected cells of soybean root nodules. Protoplasma 125: 1-12.

Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H. and Verma, D.P.S. (1985). Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. Proc Natl Acad Sci USA 82: 5040-5044.

Nguyen, J., Machal, L., Vidal, J., Perrot-Rechenmann, C. and Gadal, P. (1986). Immunochemical studies on xanthine dehydrogenase of soybean root nodules. Planta 167: 190-195.

Padilla, J.E., Miranda, J. and Sánchez, F. (1991). Nodulin regulation in common bean nodules induced by bacterial mutants. Mol Plant-Microbe Interact 4: 433-439.

Papadomoulou, K., Roussis, A., Kuin, H. and Katinakis, P. (1995). Expression pattern of uricase II gene during root nodule development in *Phaseolus vulgaris*. Experientia 51: 90-94.

Rainbird, R.M. and Atkins, C.A. (1981). Purification and some properties of urate oxidase from nitrogen-fixing nodules of cowpea. Biochim Biophys Acta 659: 132-140.

Reynolds, P.H.S., Boland, M.J., Blevins, D.G., Schubert, K.R. and Randall, D.D. (1982a). Enzymes of amide and ureide biogenesis in developing soybean nodules. Plant Physiol 69: 1334-1338.

Reynolds, P.H.S., Blevins, D.G. and Randall, D.D. (1982b). 5-Phosphoribosylpyrophosphate amidotransferase from soybean root nodules: kinetic and regulatory properties. Arch Biochem Biophys 229: 623-631.

Sánchez, F., Padilla, J.E., Pérez, H. and Lara, M. (1991). Control of nodulin genes in root-nodule development and metabolism. Ann Rev Plant Physiol Plant Mol Biol 42: 507-528.

Sánchez, F., Campos, F., Padilla, J., Bonneville, J.M., Enriquez, C. and Caput, D. (1987). Purification, cDNA cloning and developmental expression of the nodule-specific uricase from *Phaseolus vulgaris*. Plant Physiol 84:1143-1147.

Schnorr, K.M., Laloue, M. and Hirel, B. (1996). Isolation of cDNAs encoding two purine biosynthetic enzymes of soybean and expression of the corresponding transcripts in roots and root nodules. Plant Mol Biol 32: 751-757.

Schubert, K.R. and DeShone, G.M. (1980). Enzymes of purine biosynthesis and catabolism in soybean root nodules: role in ureide biosynthesis. Plant Physiol 65: S-614.

Schubert, K.R. (1981). Enzymes of purine biosynthesis and catabolism in *Glycine max*. Plant Physiol 68: 1115-1122.

Schubert, K.R. and Boland, M.J. (1990). Ureides. In: The Biochemistry of Plants (P.K. Stumpf and E.E. Conn Ed.) 16: 197-282. Academic Press, New York.

Selker, J.M.L. and Newcomb, E.H. (1985). Spatial relationship between uninfected and infected cells in root nodules of soybean. Planta 165: 446-454.

Shelp, B.J. and Atkins, C.A. (1983). Role of inosine monophosphate oxidoreductase in the formation of ureide in nitrogen-fixing nodules of cowpea (*Vigna unguiculata* L. Walp). Plant Physiol 72: 1029-1034.

Shelp, B.J., Atkins, C.A., Storer, P.J. and Canvin, D.T. (1983). Cellular and subcellular organization of pathway of ammonia assimilation and ureide synthesis in nodules of cowpea (*Vigna unguiculata* L. Walp). Arch Biochem Biophys 224: 429-441.

Smith, P.M.C., Mann, A.J., Goggin, D.E. and Atkins, C.A. (1998). AIR synthetase in cowpea nodules: a single gene product targeted to two organelles?. Plant Mol Biol 36: 811-820.

Tajima, S. and Yamamoto, Y. (1975). Enzymes of purine catabolism in soybean plants. Plant Cell Physiol 16: 271-282.

Takane, K., Tanaka, K., Tajima, S., Okazaki, K. and Kouchi, H. (1997a). Expression of a gene from uricase II (nodulin-35) in cotyledons of soybean plants. Plant Cell Physiol 38:149-154.

Takane, K, Tajima, S. and Kouchi, H. (1997b). Two distinct uricase II (nodulin 35) genes are differentially expressed in soybean plants. Mol Plant-Microbe Interact 10: 735-741.

Tayjima, S. and Yomomato, Y. (1975). Enzymes of purine catabolism in soybean plants. Plant and Cell Physiol 16: 271-282.

Triplett, E.W., Blevins, D.G. and Randall, D.D. (1980). Allantoic acid synthesis in soybean root nodule cytosol via xanthine dehydrogenase. Plant Physiol 65: 1203-1206.

Triplett, E.W., Blevins, D.G. and Randall, D.D. (1982). Purification and properties of soybean nodule xanthine dehydrogenase. Arch Biochem Biophys 219:39-46.

Triplette, E.W. (1985). Intracellular nodule localization and nodule specificity of xanthine dehydrogenase in soybean. Plant Physiol 77: 1004-1009.

Van den Bosch, K.A. and Newcomb, E.H. (1986). Immunogold localization of nodule-specific uricase in developing soybean root nodules. Planta 167: 425-436.

Van de Sande, K., Pawlowski, K., Czaja, I., Wieneke, U., Schell, J., Schmidt, J., Walden, R., Matvienko, M., Wellink, J., van Kammen, A., Franssen, H. and Bisseling, T. (1996). Modification of phytohormone response by a peptide encoded by ENOD40 of legumes and a nonlegume. Science 273: 370-373.

Van de Sande, K. and Bisseling, T. (1997). Signaling in symbiotic root nodule formation. Essays Biochem 32: 127-142.

Van de Wiel, C., Scheres, B., Franssen, H., van Lierop, M.J., van Lammeren, A., van Kammen, A. and Bisseling, T. (1990). The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. EMBO J 9: 1-7.

Van Kammen, A. (1984). Suggested nomenclature for plant genes involved in nodulation and symbiosis. Plant Mol Biol Rep 2: 43-45.

Van Rhijn, P. and Vanderleyden, J. (1995). The *rhizobium*-plant symbiosis. Microbiol Rev 59: 124-142.

Van Spronsen, P.C., Bakhuizen, R., van Brussel, A.A.N. and Kijne, J.W. (1994). Cell wall degradation during infection thread formation by the nodule bacterium *Rhizobium leguminosarum* is a two-step process. Eur J Cell Biol 64: 88-94.

Verma, D.P.S., Fortin, M.G., Stanley, J., Mauro, V.P., Purohit, S. and Morrison, N. (1986). Nodulins and nodulin genes of *Glyicine max*. Plant Mol Biol 7: 51-61.

Verma, D.P.S. and Delauney, A.J. (1988). Root nodule symbiosis: nodulins and nodulin genes. In: Plant gene research: temporal and spatial regulation of plant genes (D.P.S. Verma and R.B. Goldberg Ed.). pp169-199. Springer-Verlag, New York.

Vijn, I., Christiansen, H., Lauridsen, P., Kardailsky, I., Quandt, H.-J., Broer, I., Drenth, J., Jensen, E., van Kammen, A. and Bisseling, T. (1995a). A 200 bp region of the pea ENOD12 promoter is sufficient for nodule-specific and Nod factor induced expression. Plant Mol Biol 28: 1103-1110.

Vijn, I., Yang, W.-C., Pallisgard, N., Jensen, E., van Kammen, A. and Bisseling, T. (1995b). VsENOD5, VsENOD12 and VsENOD40 expression during *Rhizobium*-induced nodule formation on *Vicia sativa* roots. Plant Mol Biol 28: 1111-1119.

Webb, M.A. and Newcomb, E.H. (1987). Cellular compartmentation of ureide biogenesis in root nodules of cowpea (*Vigna unguiculata* (L.) Walp.). Planta 172: 162-175.

Webb, M.A. and Lindell, J.S. (1993). Purification of allantoinase from soybean seeds and characterization of anti-allantoinase antibodies. Plant Physiol 103: 1235-1241.

Weber, G. (1983). Biochemical strategy of cancer cells and design of chemotherapy. Cancer Res 43: 3466-3492.

Wycoff, K.L., Hunt, S., Gonzales, M.B., VandenBosch, K.A., Layzell, D.B. and Hirsch, A.M. (1998). Effects of oxygen on nodule physiology and expression of nodulins in alfalfa. Plant Physiol 117: 385-395.

Yang, D. (1997). Purification and characterization of inosine-5'-monophosphate dehydrogenase (IMPDH) from nitrogen-fixing nodules of soybean (*Glycine max*). Master's thesis, University of Oklahoma, Norman, OK, USA.

Yang, W.C., Katinakis, P., Hendriks, P., Smolders, A., de Vries, F., Spee, J., van Kammen, A., Bisseling, T. and Franssen, H. (1993). Characterization of GmENOD40, a gene showing novel patterns of cell-specific expression during soybean nodule development. Plant J 3: 573-585.

Chapter 2. Isolation of a cDNA Clone for Soybean Nodule IMPDH

INTRODUCTION

Inosine monophosphate dehydrogenase (IMPDH; EC 1.1.1.205) catalyzes the rate-limiting step in the *de novo* biosynthesis of guanine nucleotides (Weber, 1983) and has an essential role in providing the necessary precursors for DNA and RNA biosynthesis and in signal transduction pathways that mediate cellular differentiation (Huberman *et al.*, 1995) and transformation (Jenkins *et al.*, 1993). Because inhibition of IMPDH activity results in the cessation of DNA synthesis, IMPDH inhibitors have been used in diverse therapeutic applications including immunosuppression (Dayton *et al.*, 1994), arthritis (Godblum, 1993) and the treatment of parasitic disease (Hupe *et al.*, 1986). Because increases in IMPDH activity are associated with cell proliferation, the enzyme is also a possible target for cancer chemotherapy (Tricot *et al.*, 1990).

In soybean and many tropical legumes, fixed nitrogen is exported from the root nodules primarily as the ureides --- allantoin and allantoic acid. Recently fixed nitrogen in the form of ammonium (NH_4^+) is secreted from the bacteroid into the plant cytosol. Within the plant cytosol, ammonium is assimilated and the products of ammonium assimilation, glutamine and glutamate, are used for ureide production. Ureides are formed via purine biosynthesis and purine oxidation. The IMPDH-catalyzed oxidation of inosine monophosphate (IMP), the immediate product of *de novo* purine biosynthesis, leads to the production of ureides (review: Schubert and Boland, 1990).

The first direct evidence for the possible role of IMPDH in ureide biogenesis was the successful demonstration of IMPDH activity in a proplastid

fraction from soybean nodules by Boland and Schubert (1983). Subsequently, Shelp and Atkins (1983) found IMPDH activity in extracts from cowpea nodules and suggested that the primary route for xanthine production involved this enzyme. In their review of ureide biogenesis, Schubert and Boland (1990) proposed that the oxidation of IMP via IMPDH was the primary pathway for IMP metabolism leading to ureide biogenesis. Although IMPDH from nodules of *Vigna unguiculata* has been purified 140-fold (Atkins *et al.*, 1985), the protein has not been sequenced. The gene encoding IMPDH has not been cloned from any nitrogen-fixing plant.

Up to now, cDNAs for IMPDH have been cloned and the sequences have been published from *Arabidopsis*, human, mouse, *Drosophila*, yeast, fungus and bacteria (refer to Genebank database and Table 3-1). Among these species, *Arabidopsis thaliana* (Collart *et al.*, 1996a) is the only plant. Cloning and characterization of the *IMPDH* cDNA in soybean nodules will make the studies of IMPDH gene expression, regulation and cellular localization possible, which will lead to the clarification of the function of IMPDH in the pathway of ureide biogenesis in nitrogen-fixing plants.

MATERIALS AND METHODS

Library titering

A λgt11 cDNA library, constructed from mRNA isolated from 30 to 35-day-old soybean nodules, was provided by Dr. Robert V. Klucas from the University of Nebraska (Ji *et al.*, 1994). cDNA inserts were ligated to the vector λgt11 with *EcoR* I sites (Figure 2-1). The original cDNA library consisted of ~$3.5x10^{11}$

bacteriophage particles with 90% recombinants (Ji *et al.*, 1994). For titering, *E. coli* strain of Y1090 (r-) was cultured overnight at 37°C in LB broth with 10 mM $MgSO_4$ and 0.2% maltose. The culture was centrifuged at 2,500 x g for 10 min, and the cell pellet was resuspended and diluted with 10 mM $MgSO_4$ to an OD_{600} of 0.5. The bacterial resuspension (200 µl) was mixed with 100 µl phage at dilutions of 10^{-2}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, 10^{-8} and 10^{-9} in SM buffer (0.1 M NaCl, 8 mM $MgSO_4.7H_2O$, 50 mM Tris-HCl, pH 7.5, 0.1% gelatin). After incubation for 15 min at 37°C, 3 ml of 0.7% top agar was added to the bacteria-phage suspension and plated on LB medium. After incubation at 37°C for 7-9 h, plaques were counted to determine the titer of the library.

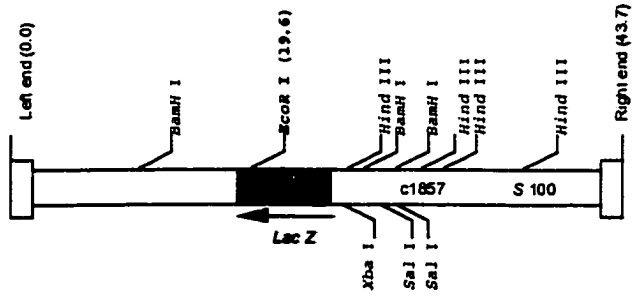


Figure 2-1. Structural map of uncut λgt11 vector (From Promega). During library construction, cDNA inserts were ligated to the vector with *EcoR* I sites.

cDNA library screening

Two degenerate oligonucleotide probes were designed by comparing the known IMPDH sequences. The sequence of the first probe, 5'-(A/T)(C/G)I(A/T)(C/G) ICC(A/T)ATGGA(C/T)ACIGT(G/T)(A/T)(C/G)IGA-3' was based on the IMPDH conserved amino acid sequence region of SSPMDTV(T/S)E. The second probe,

5'-(G/A)TG(C/A)AC(T/A)CC(T/A)CC(T/C)TC(C/A)AC(T/C)TGIGC-3', was based on the amino acid sequence AQVEGGVH. The codon preferences in soybean plants were used in the design of these degenerate probes. Oligonucleotide probes were synthesized at the Molecular Biology Resource Facility of the University of Oklahoma Health Sciences Center. Probes were radioactively labeled with γ-^{32}P-ATP using the polynucleotide kinase end-labeling method (Ausubel *et al.*, 1991).

To start the primary library screening, the host bacteria *E. coli* Y1090 (r-) was grown overnight at 37°C in a 50-ml LB broth culture containing 10 mM $MgSO_4$ and 0.2% maltose. Bacterial cells were pelleted by centrifugation at 1000 x g for 10 min, resuspended with 10 mM $MgSO_4$ and adjusted to an OD_{600} of 0.5 with 10 mM $MgSO_4$. A 600-µl volume of cell suspension was mixed with 300 µl phage dilution ($2.1x10^5$ pfu/ml) and incubated at 37°C for 15 min. After incubation, 7 ml of 0.7% top agar kept at 45°C was added to the bacterial phage mixture, vortexed for 1 min and poured onto a 132-mm LB plate. The plate was incubated at 37°C for 9-12 hours until well-formed plaques were readily visible.

Appropriately marked nitrocellulose membranes (MSI, Micron Separation Inc., Westborough, WA. Cat No. WP4HY08550) were used for transferring plaques and making lifts. When making lifts, a membrane was gently placed on the agar surface and air bubbles between the membrane and the top agar were expelled. After 3 min incubation on the top agar, the membrane bearing plaques was removed and treated successively with the denaturation solution (0.5 M NaOH, 1.5 M NaCl) and the neutralization solution (1.0 M Tris-Cl, 1.5 M NaCl,

pH 7.5) for 4 min each. Afterwards, the membranes were washed twice in 2x SSC solution (0.3 M NaCl, 0.3 M Na-citrate, pH 7.0) for 4 min per wash. After washing, lifts were air-dried and baked at 80°C for 2 hours under vacuum.

Both pre-hybridization and hybridization were carried out overnight at 42°C. The pre-hybridization solution contained 0.02 M NaH_2PO_4, 0.4% SDS, 5X Denhardt solution, 6x SSC and 0.5 mg/ml herring sperm DNA. The hybridization solution was the same as the pre-hybridization solution except that the Denhardt solution was omitted and the radiolabeled oligonucleotide probes were added. The hybridization was carried out overnight. Membranes were washed with 6x SSC + 0.1% SDS at 45°C several times until the background radioactivity was low.

Labeling was detected by autoradiography. Exposure was carried out at -80°C for 48-72 h. Signals on the X-ray films were matched to the corresponding plates. Positive plaques were picked up, put in 1 ml SM buffer containing 20 µl chloroform, vortexed briefly and stored at least overnight at 4°C before beginning the secondary screening.

The procedures used in the secondary screening were the same as those used in the primary screening except: 1) for infection, 200 µl bacterial suspension (OD_{600}=0.5) was mixed with 100 µl diluted phage from the primary screening; 2) 85-mm petri dishes were used for plaque culture; 3) duplicate lifts were made from each plate; 4) higher stringency conditions (4x SSC or 2x SSC instead of 6x SSC) were used to wash the membranes after the hybridization. After the secondary screening, positive clones that had signals in both of the duplicate lifts

were picked up. Each plaque was put in 500 µl SM buffer and 10 µl chloroform, vortexed briefly and stored at 4°C.

Phage isolation and purification

Phage isolation and purification was necessary because of their diffusion in LB plates. Positive plaques from the secondary screening, stored in SM buffer containing chloroform, were diluted 10^{-4} to 10^{-6} fold. 100 µl of the plaque dilution was incubated with 200 µl bacteria (OD_{600}=0.5) at 37°C for 20 min. After incubation, 3 ml of 0.7% top agar containing 250 µg/ml X-gal and 1 mM IPTG was added to the bacterial phage mixture and the mixture was vortexed briefly and poured onto LB plates. The plates were cultured at 37°C for 9~12 h until the plaques were visible. Single, white plaques were picked from the plates, put in 100 µl SM buffer and 1.5 µl chloroform and stored at 4°C overnight.

DNA isolation and sequencing of phage clones

Phage DNAs from purified plaques were isolated using the Promega Wizard® phage DNA isolation kit with the following modifications in lysate preparation. Fifteen µl phage stored in SM buffer containing chloroform was mixed with 1250 µl of an overnight culture of *E. coli* stain Y1090(-). After incubation for 30 min at 37°C, the mixture was added to a pre-warmed flask containing 100 ml LB broth and 1.0 ml 1 M $MgSO_4$. Cell lysis occurred after 5 hours of culture at 37°C. At that point, 500 µl of chloroform was added to the culture. After another 15 min of

growth, the phage culture was centrifuged for 10 min at 10,000 x g. The supernatant was collected as the lysate and stored at 4°C.

Ten ml of lysate was used for each phage DNA isolation. The detailed protocol for phage DNA isolation from Promega is presented in "Appendices: protocols and solutions".

Isolated phage DNAs were digested with *Eco*R I. The inserts were analyzed by electrophoresis on a 1.2% agarose gel. DNA samples yielding inserts with the desired size range were further confirmed by Southern blot hybridization using the same radiolabeled probes as those used in the library screening. Phage DNA isolated from the target clones was sequenced, initially with λgt11 sequencing primers (Promega, Cat No. Q5851, Q5861) and then by primer-walking. DNA sequencing was performed at the Recombinant DNA / Protein Resource Facility of Oklahoma State University.

Subcloning into phagemid pBluescript SK(-)

In order to facilitate DNA sequencing, inserts from the target phage clones were subcloned into phagemid pBluescript SK(-) (Stratagene, La Jolla, CA). Phage DNA was isolated and digested with *EcoR* I. As described in Appendix, inserts of ~2.0 kb were cut from the gel and cleaned with QIAEX II kit (Qiagen). The vector pBluescript SK(-) was first digested with *EcoR* I, then dephosphorylated with calf intestinal alkaline phosphatase (CIAP). Ligation was carried out at 15°C overnight. Competent cells of *E. coli* strain XL1 Blue mrf' (Stratagene) were prepared and transformed by the one-step method of Chung *et al.* (1989). The

protocol for the preparation of competent cells is presented in the Appendix. Recombinants were selected on LB-ampicillin plates and screened based on their blue / white color. Plasmid DNA from white colonies was isolated and the identity of the target cDNA was confirmed by *Eco*R I digestion and Southern blot hybridization before being used for sequencing.

Identification of IMPDH *cDNA by sequence analysis*

ExPASy (Expert Protein Analysis System), a molecular biology resource from Swiss Institute of Bioinformatics and available over the Internet (http://www.expasy.ch) was employed to analyze the sequence of the cloned cDNA. The program "Translate tool" was used to translate a nucleotide sequence into a protein sequence. The homology of the deduced amino acid sequence to other IMPDHs was identified using the program Basic Local Alignment Search (BLAST) (Altschul *et al.*, 1997). The proposed IMPDH signature sequence was identified using the PROSITE (database of protein families and domains) program (Bairoch, 1995; Hofmann *et al.*, 1999). Internal restriction sites in the cloned cDNA was searched using "Webcutter 2.0" (copyright 1997, Max Heiman) available from the Baylor College of Medical Search Launcher website (http://dot.imgen.bcm.tmc.edu:9331).

RESULTS

Library titering

Titering results were shown below in Table 2-1. The library titer was calculated from the results of the phage dilution to be $2.1X10^{10}$ pfu/ml.

Table 2-1. Soybean nodule cDNA library titering results

phage concentration	plaques
10^{-4} X	*
10^{-5} X	*
10^{-6} X	*
10^{-6} X	*
10^{-7} X	152
10^{-7} X	199
10^{-8} X	22
10^{-8} X	19
10^{-9} X	1
10^{-9} X	5
10^{-10}X	1

* Too many plaques to count.

cDNA library screening and soybean nodule IMPDH *cDNA sequencing*

The process and results of screening the soybean cDNA library for IMPDH are summarized in Figure 2-2. The process is divided into two rounds, 'Screening I' and 'Screening II'.

'Screening I' was carried out using degenerate oligonucleotide probes. Twenty-two positive signals were detected after the primary screening and 24 were detected after the secondary screening (Figure 2-2, 2-3). For the 24 clones,

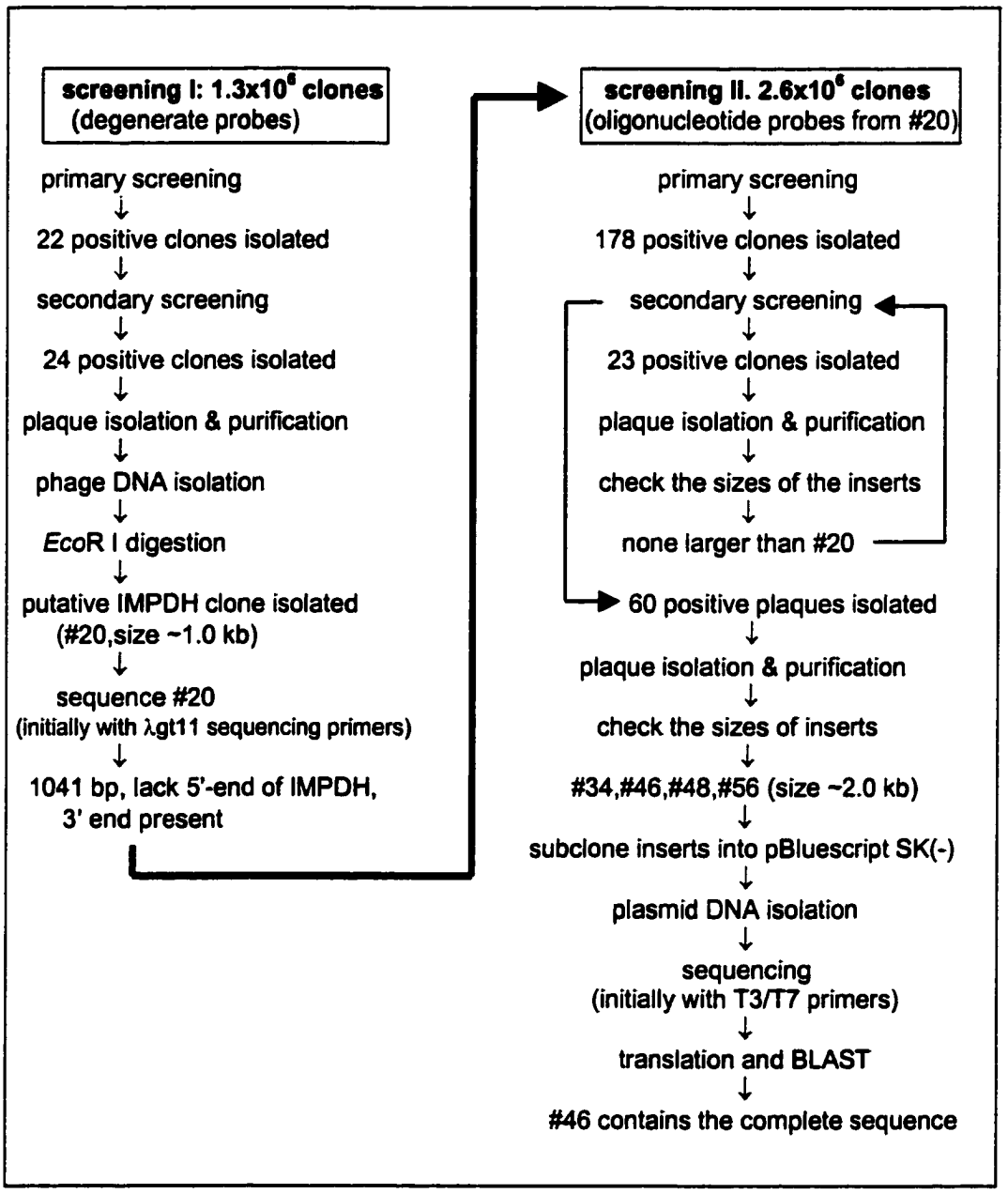


Figure 2-2. Outline of procedures and results of soybean nodule cDNA library screening for IMPDH clones.

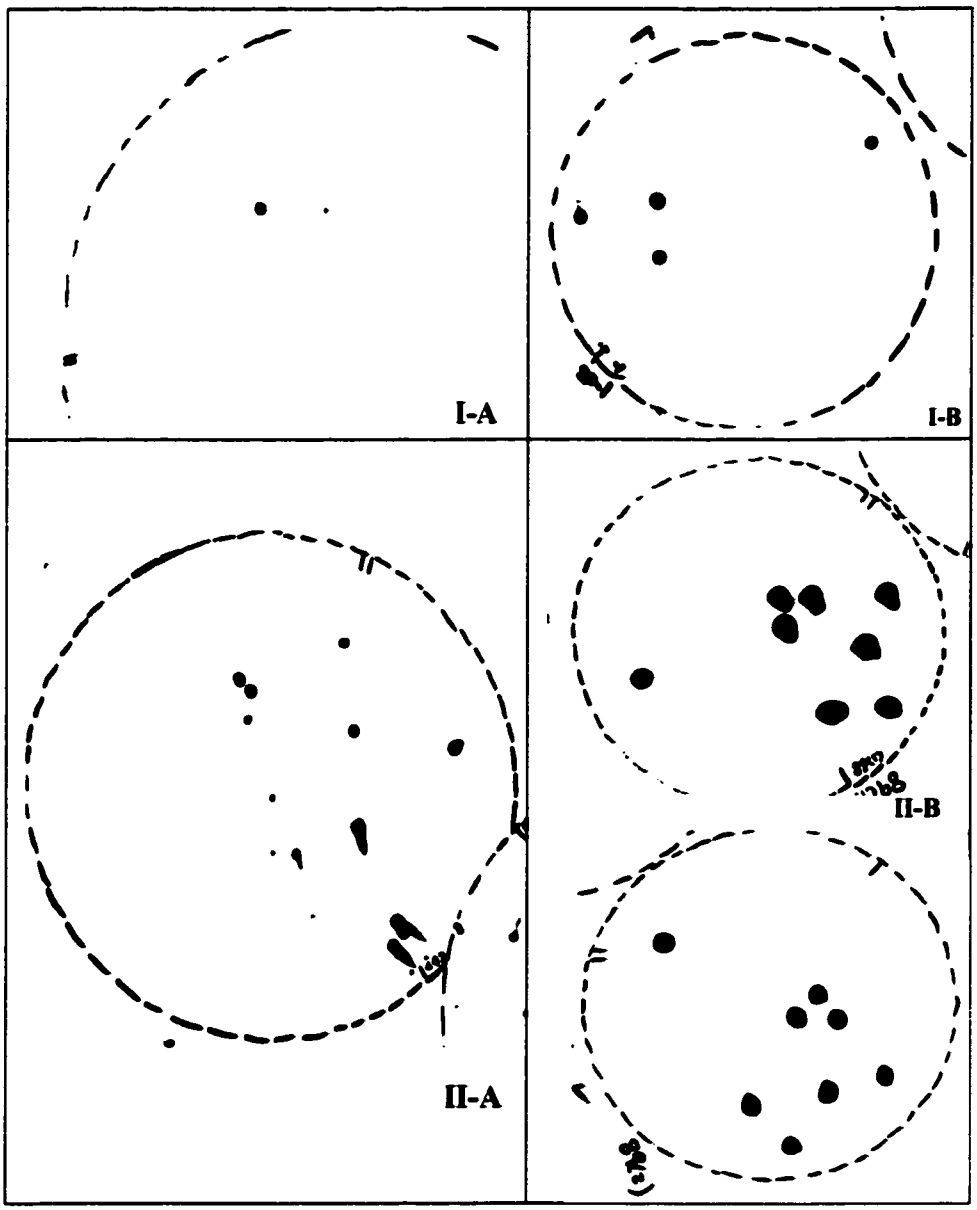


Figure 2-3. Positive signals on X-ray film from cDNA screening for IMPDH. I-A, primary screening in screening I; I-B, secondary screening in screening I; II-A, primary screening in screening II; II-B, a pair of duplicate plates from the secondary screening in screening II. Degenerate oligonucleotides were used for screening I, and oligonucleotides based on the DNA sequence of clone #20 were used for screening II.

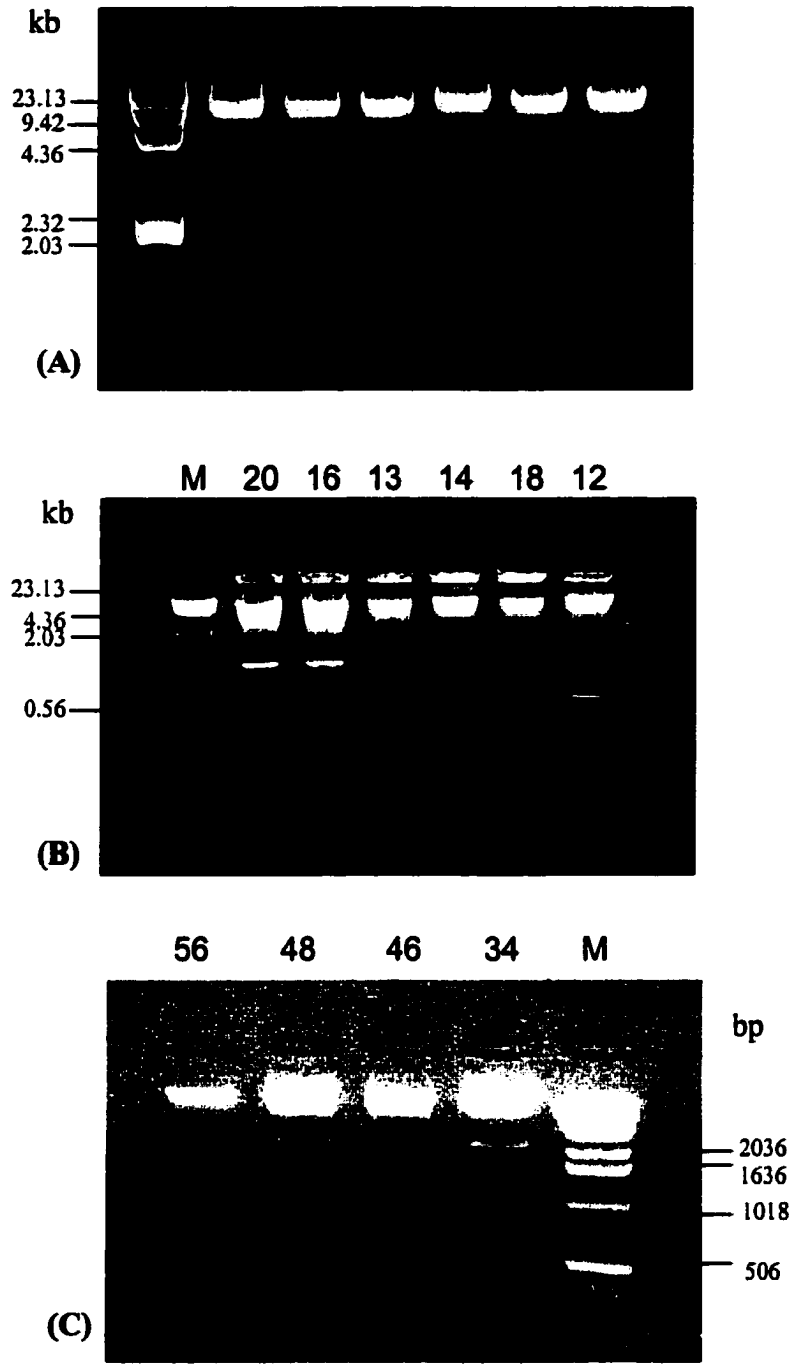


Figure 2-4. *Eco*R I digestion of phage DNA from positive clones. DNAs were run on 1.2% agarose gel. A). Phage DNA isolated from different clones, before digestion; B). *Eco*R I digestion of phage DNA from the screening I. Inserts from clone # 16 and # 20 had the largest sizes, ~1.0 kb; C). *Eco*R I digestion of phage DNA from the screening II. Clone # 56, # 48, # 46 and # 34 showed inserts of ~2.0 kb. Molecular size markers in A) and B) are lambda DNA digested with *Hind* III; marker in C) is 1 kb DNA ladder (Gibco BRL, Cat. No. 15615-016).

the phage DNA was isolated and digested with *EcoR* I. Clone #16 and #20 contained the largest inserts approximately 1.0 kb in size (Figure 2-4-B). The inserts were sequenced in both directions using the T3 and T7 primers. Based on the results of sequencing, the insert from clone #20 was 1041 bp in length, contained the stop codon and poly (A) tail at the 3' end, and contained an ORF that encoded a polypeptide of 232 amino acid residues (Figure 2-5). The DNA sequence was translated and examined for ORF's. The deduced amino acid sequence matched *A. thaliana* IMPDH starting from amino acid position 217 with 85% of the residues being identical (Figure 2-5). Thus, clone #20 apparently codes for the C-terminal fragment of soybean nodule IMPDH.

A second screening (Screening II) was used to obtain a full-length IMPDH cDNA clone. For 'screening II', two oligonucleotide probes, 5'-CGGCTTATG ATTTTGGATGGTTTT-3' and 5'-AGCTTCTTGGCTGGTATCCTTGAG-3' were synthesized based on the DNA sequence of clone #20. These two probes were 100% identical to the target clone. After the primary screening, 178 positive signals were detected (Figure 2-2, 2-3; Table 2-2). The secondary screening was carried out twice. In the first of these, 23 clones were identified, but none of these clones contained inserts larger than that of clone #20 from 'screening I'. The second time, 60 clones were identified and 19 of them were analyzed (Table 2-3). Four clones, designated as #34, #46, #48 and #56, contained inserts of about 2 kb (Figure 2-4-C). Inserts from these clones were subcloned into plasmid pBluescript SK(-) and sequenced. Sequencing results showed that clone #46 was 2016 bp in length, and its 3'-end was identical to the 3' end of the clone #20.

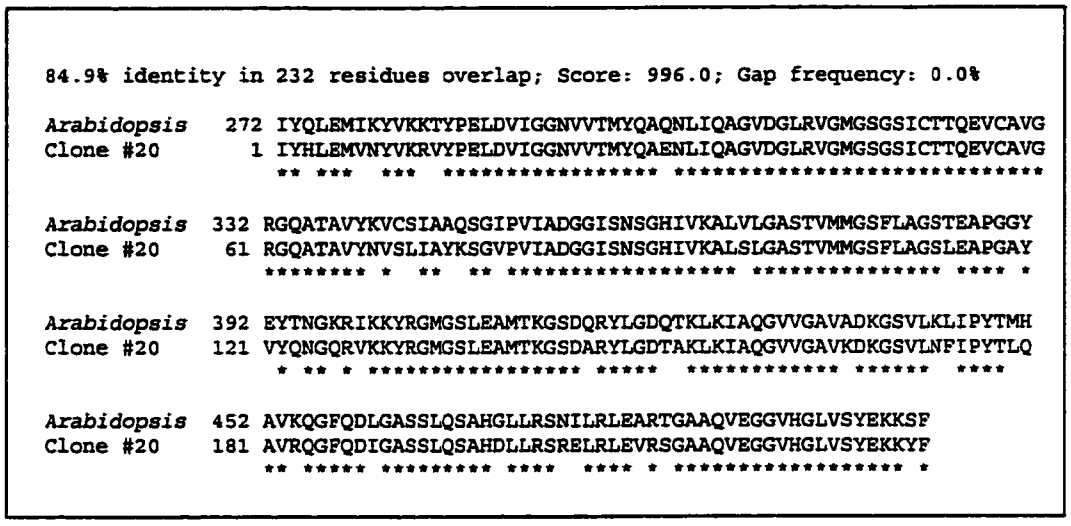


Figure 2-5. Comparison of the putative polypeptide encoded by clone #20 inserts with *Arabidopsis thaliana* IMPDH. Identical residues are marked with asterisks (*). Sequences were aligned by using the SIM program (Huang and Miller, 1991).

Identification of IMPDH cDNA by sequence analysis of clone #46

Sequence analysis of clone #46 showed that this 2016 bp-long cDNA contained a 166-bp 5'-untranslated sequence and a 341-bp 3'-untranslated sequence (Figure 2-6). The open reading frame coded for 502 amino acid residues. The typical putative polyadenylated signal in animal cDNAs, AATAAA, was not found at the 3' end of clone #46, but a match of four of the six bases appears several times. BLAST results showed that the deduced amino acid sequence of clone #46 shares 70.5% identity with IMPDH from *A. thaliana* (Figure 2-7). Internal restriction sites were searched with the program "Webcutter 2.0". Results are shown in Figure 2-8 and Table 2-4.

Table 2-2. Results of cDNA library screening II: primary screening

Plate No.	Isolated Positive Plaques (As Labeled)	Total # of Positive Signals
1	1,**2**,3,4,5	5
2	**6**,7,**8**	3
3	**9**,**10**,11,12,13,**14**,15,16,17	9
4	**18**,**19**,**20**,**21**,**22**,**23**,**24**,**25**,26,27	10
5	28,29,30,31,32	5
6	33,34,35,**36**,**37**,**38**,**39**,**40**	8
7	41,42,43,44,45,46,47,48,49,50	10
8	51,**52**,**53**,**54**,55	5
9	56,57,58,**59**,60,**61**,**62**,63	8
10	64,65,66,**67**,68,69,**70**,71	8
11	72,73,74,75,76,**77**,78,79,80,81,82,83,84,85,86,87,88	17
12	**89**,**90**,91,**92**,93,94,95,96,97,98	10
13	**99**,**100**,101,102,103,104,105,106,107,**108**,109,110, 111	13
14	112,113,114,115,116,117,118,**119**,120,121,122	11
15	**123**,**124**,125,126,127,128,129,130,131,132,**133**,134	12
16	135,136,137,**138**,139,140,141,142	8
17	143,144,145,146,147,148,149,150,151,152	10
18	**153**,**154**	2
19	155,**156**,157,158,159,160	6
20	161,162,163,164,165,166,167,168,**169**,170	10
21	**171**,172,173,174,175	5
22	**176**,177,178	3
Total		178

* Positive plaques in bold numbers in Table 2-2 represent the plaques that were picked up for the secondary screening.

** Positive plaques in bold numbers in Table 2-3 (next page) represent the plaques that were picked up for further isolation, purification, phage DNA isolation and *Eco*R I digestion.

Table 2-3. Results of cDNA library screening II: secondary screening

1st:

Plate No.	Positive Plaques (As Labeled)	Total # of Positive Signals
2 (from #1)	**2-1**, 2-2	2
6 (from #2)		0
8 (from #2)		0
10 (from #3)	**10-1**, 10-2, **10-3**, **10-4**	4
14 (from #3)	14-1	1
18 (from #4)	**18-1**, 18-2, 18-3, **18-4**, 18-5	5
19 (from #4)	**19-1**	1
20 (from #4)	**20-1**	1
21 (from #4)		0
22 (from #4)	**22-1**, 22-2, 22-3, **22-4**	4
23 (from #4)		0
24 (from #4)	24-1, **24-2**	2
25 (from #4)		0
36 (from #6)		0
37 (from #6)		0
38 (from #6)	**38-1**, **38-2**	2
39 (from #6)		0
40 (from #6)		0
53 (from #8)	**53-1**	1
Total		23

2nd:

85mm Plate No.	Positive Plaque No. (As Labeled)	Total # of Positive Signals
52 (from #8)	**#16**, #17, #18, #19(2), #20(2)	5
54 (from #8)	**#21**	1
59 (from #9)	**#22**, #23	2
61 (from #9)	**#25**, #25(2), #26(tiny)	3
62 (from #9)	**#27**, #28, #29	3
67 (from #10)	**#30**, #31(2), #32(2), #33(2)	4
70 (from #10)	**#34**, #35, #36, #37(2), #38(2)	5
77 (from #11)		0
92 (from #12)	**#39**,#40,#41,#42, #43, #44(2), #45(2)	7
90 (from #12)	**#46**	1
89 (from #12)	**#47**(?)	1
99 (from #13)		0
100(from#13)	**#48**	1
108(from#13)	**#49**, #50, #51, #52, #53(2)	5
119(from#14)	**#54**(2), #55(tiny)	2
123(from#15)		0
124(from#15)	**#56**(2)	1
133(from#15)	#57(2), **#58**	2
138(from#16)		0
153(from#18)	**#59**, #60	2
154(from#18)	**#61**, #62, #63, #64, #65, #66	6
156(from#19)		0
169(from#20)		0
171(from#21)	**#67**, #68, #69, #70, #71, #72	6
176(from#22)	**#73**, #74, #75(2)	3
Total		60

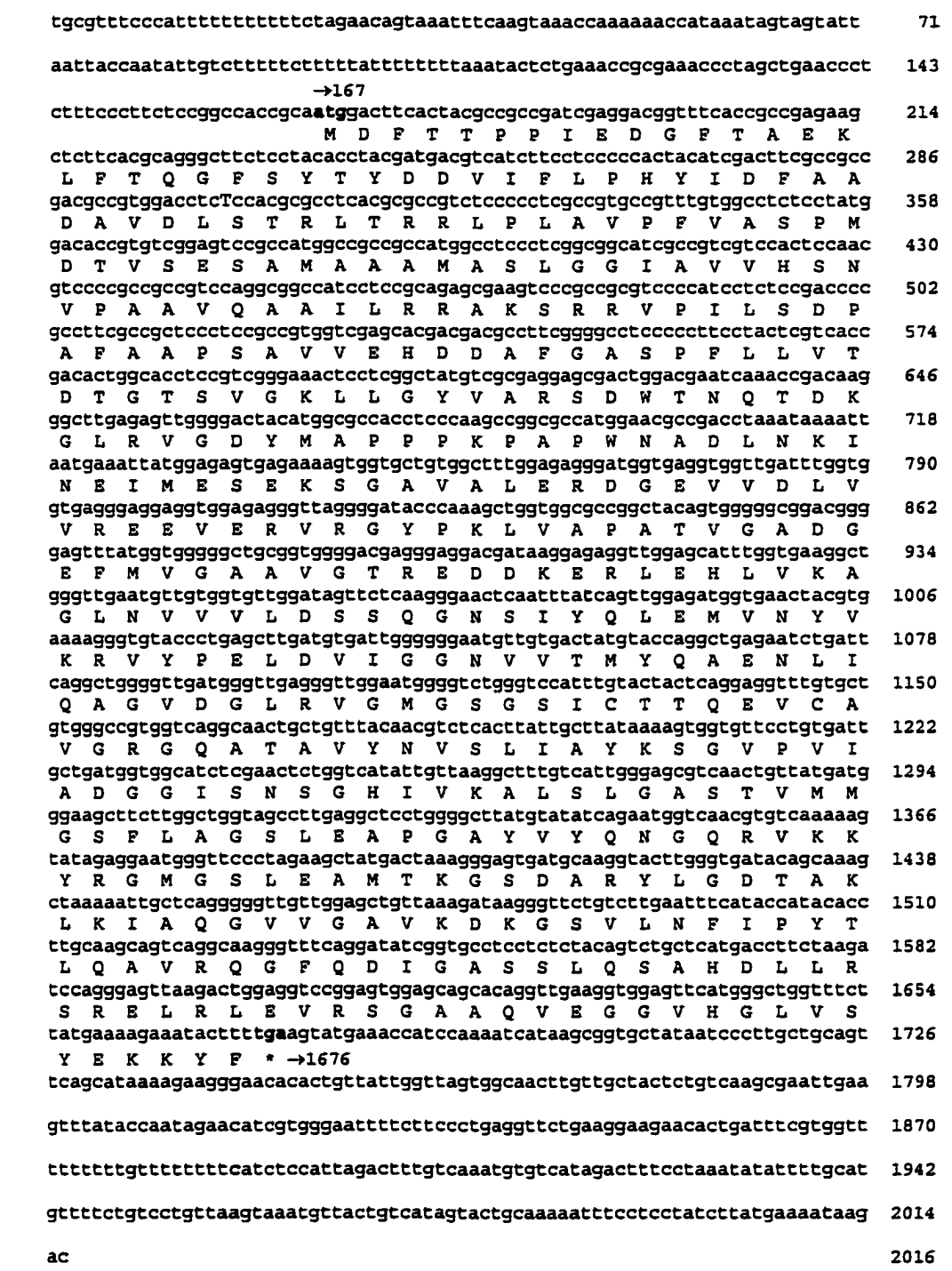


Figure 2-6. Complete sequence of soybean nodule *IMPDH* cDNA clone #46. Deduced amino acid sequence was obtained using the Internet molecular biology tools of 'Translate tool' in ExPASY program from the Swiss Institute of Bioinformatics.

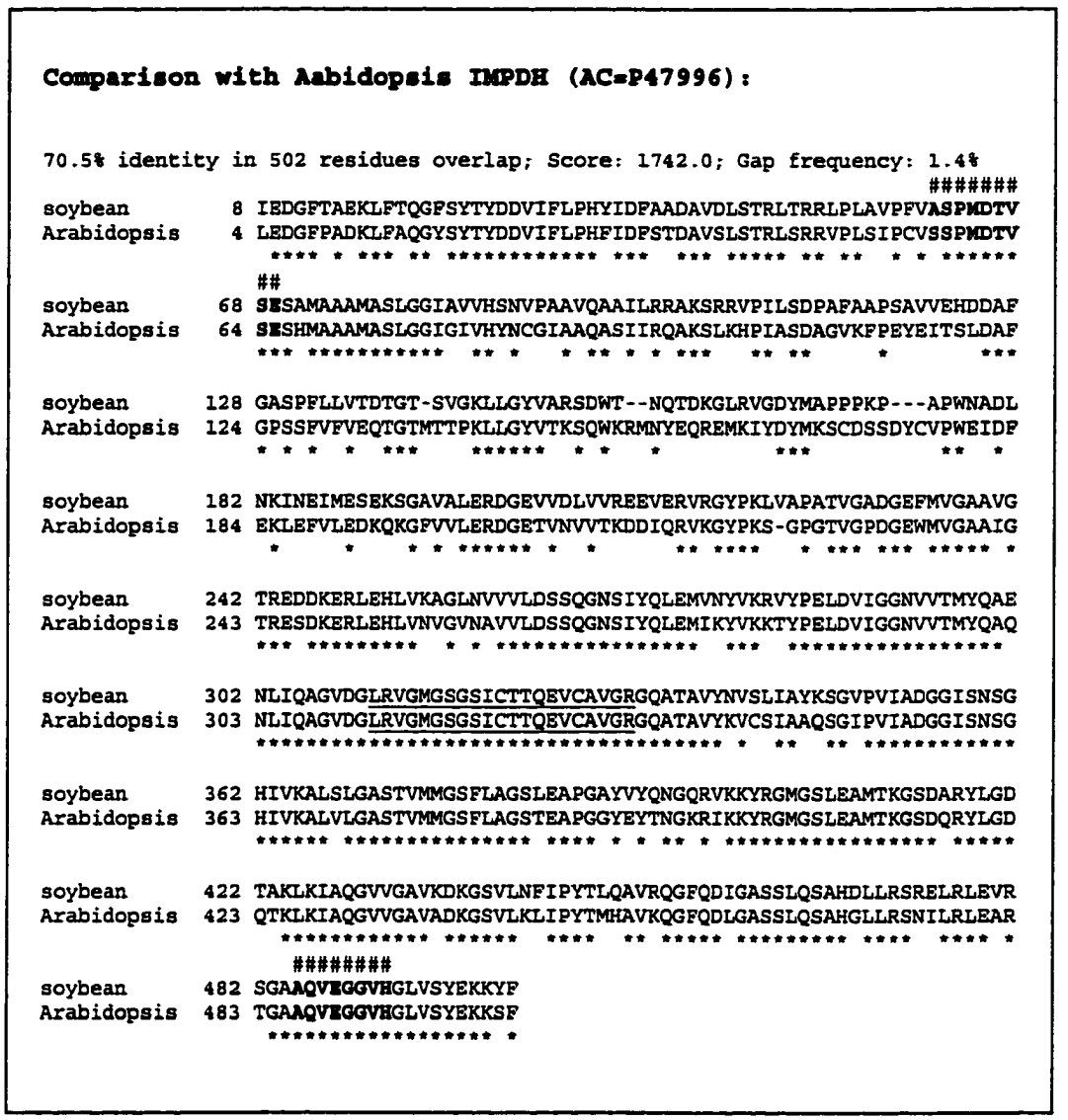


Figure 2-7. Alignment of predicted amino acid sequence of IMPDH with *Arabidopsis thaliana* IMPDH (SWISS-PROT No. P47996). Identical residues are marked with asterisks (*). The putative active-site region is underlined. The amino acid sequences for degenerate probe design and synthesis are represented in bold letters with #. Sequences were aligned using the SIM program (Huang and Miller, 1991).

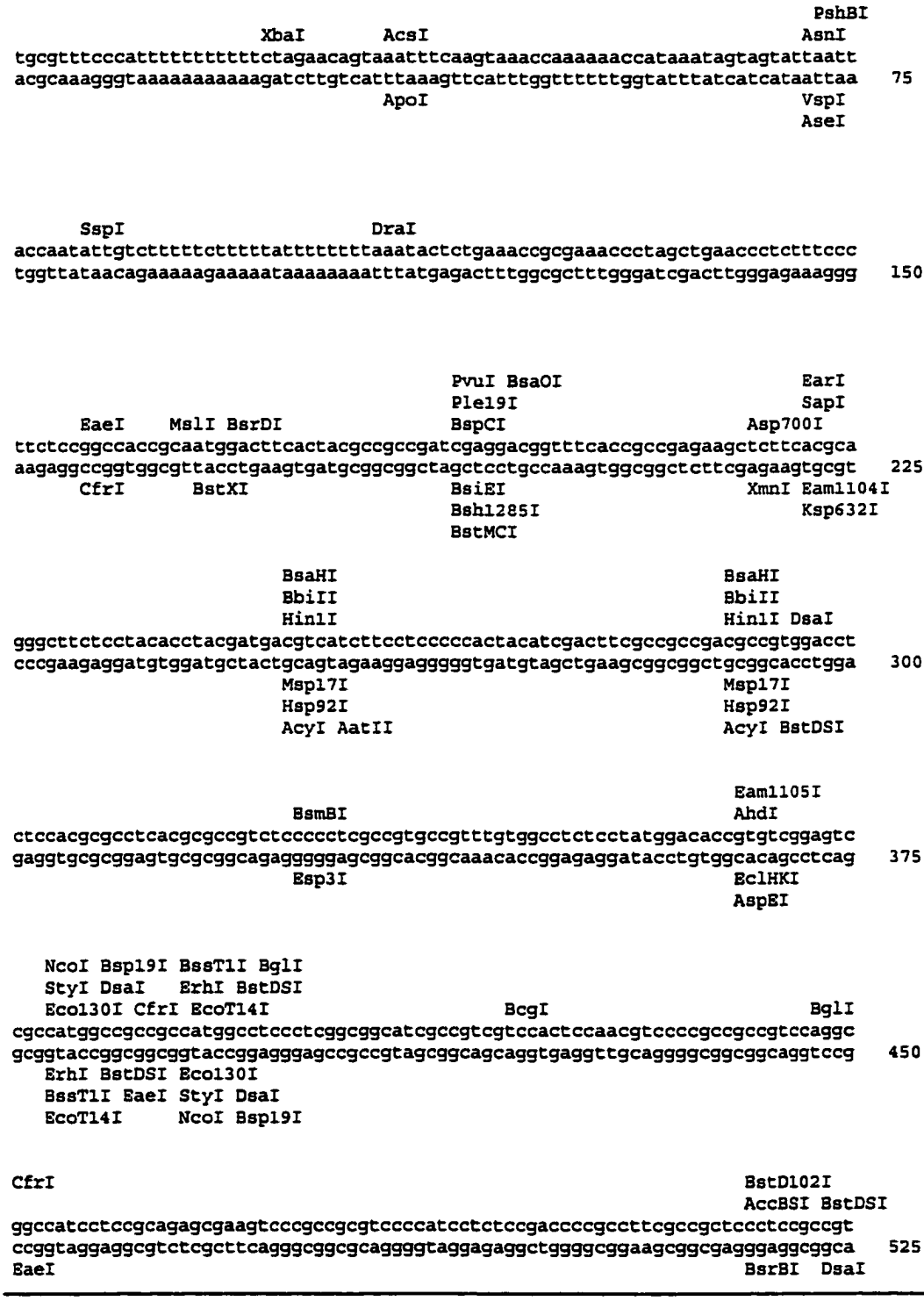


Figure 2-8. Restriction map of soybean nodule *IMPDH* cDNA. The map was constructed with the Internet molecular biology tool "Webcutter 2.0" (copyright 1997, Max Heiman).

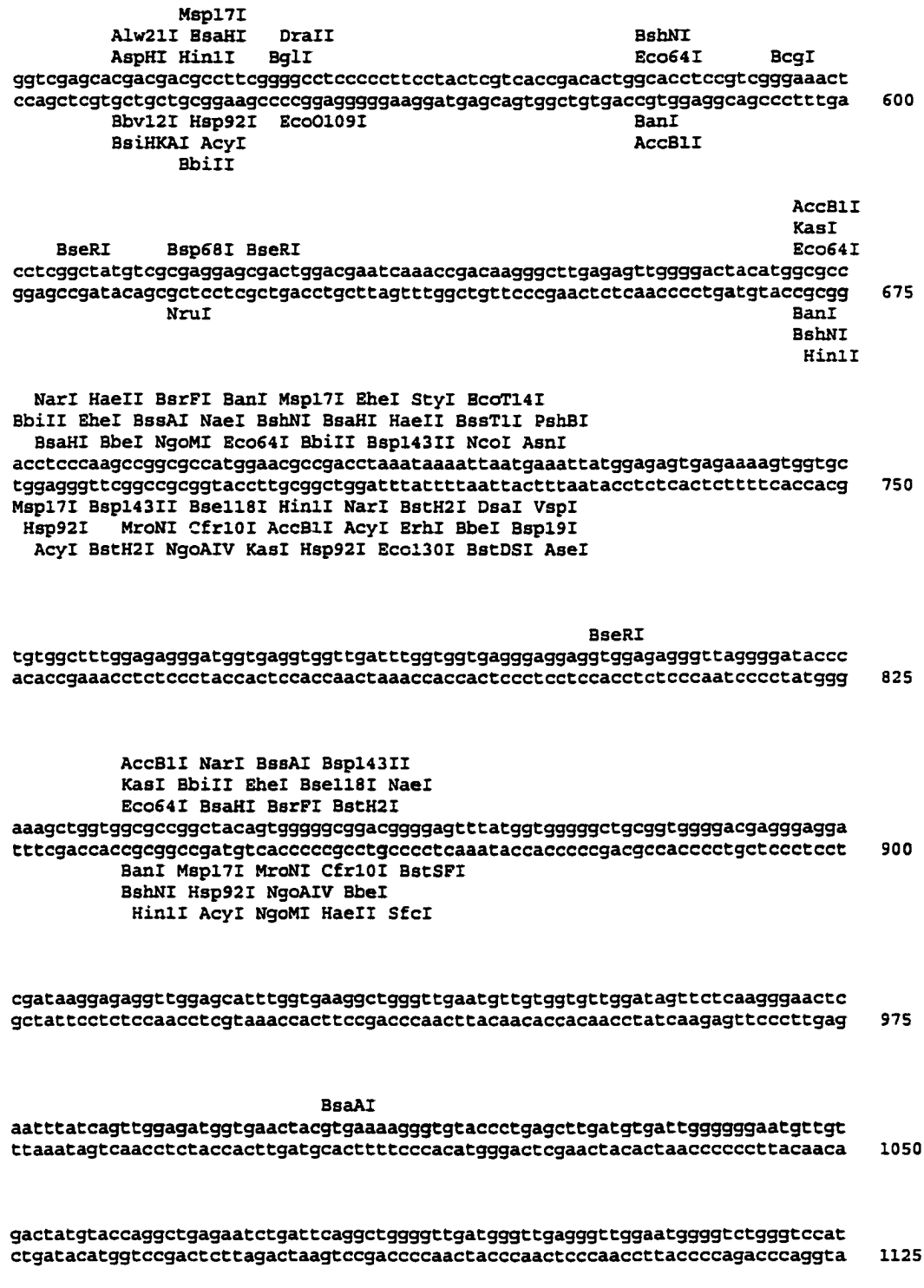


Figure 2-8 (continued). Restriction map of soybean nodule *IMPDH* cDNA. The map was constructed with the Internet molecular biology tool "Webcutter 2.0" (copyright 1997, Max Heiman).

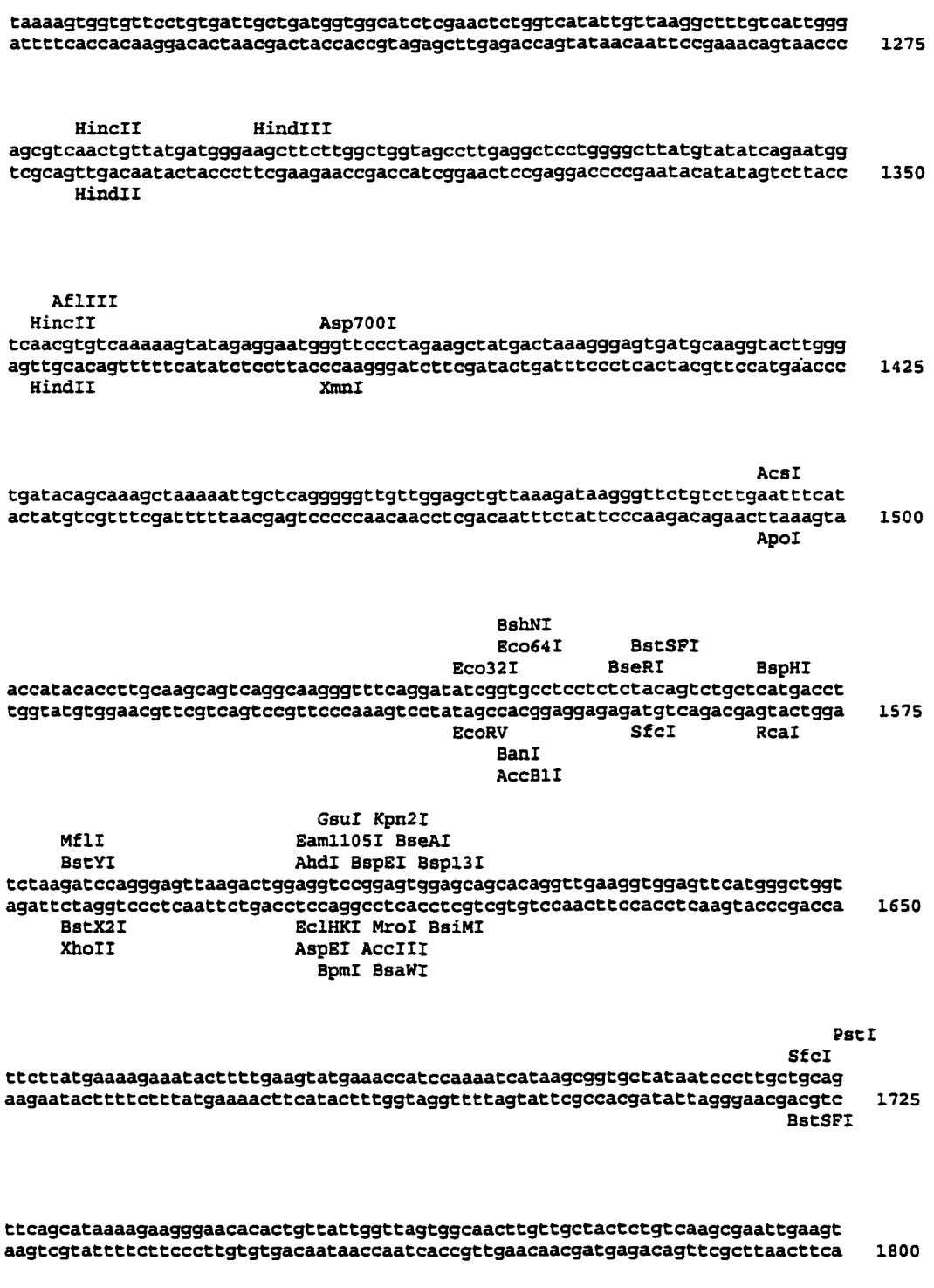


Figure 2-8 (continued). Restriction map of soybean nodule *IMPDH* cDNA. The map was constructed with the Internet molecular biology tool "Webcutter 2.0" (copyright 1997, Max Heiman).

Figure 2-8 (continued). Restriction map of soybean nodule *IMPDH* cDNA. The map was constructed with the Internet molecular biology tool "Webcutter 2.0" (copyright 1997, Max Heiman).

Table 2-4. Internal restriction sites of soybean nodule *IMPDH* cDNA

Enzyme Name	No. Cuts	Positions of Sites	Enzyme Recognition Sequence
*Aat*II	1	252	gacgt/c
*Acc*113I	1	1979	agt/act
*Acc*B1I	5	581 670 688 835 1544	g/gyrcc
*Acc*BSI	1	515	gagcgg
*Acc*III	1	1605	t/ccgga
*Acs*I	4	33 1492 1823 1987	r/aatty
*Acy*I	6	249 288 540 671 689 836	gr/cgyc
*Afl*III	1	1354	a/crygt
*Ahd*I	2	364 1601	gacnnn/nngtc
*Alw*21I	1	534	gwgcw/c
*Aoc*I	1	1835	cc/tnagg
*Apo*I	4	33 1492 1823 1987	r/aatty
*Ase*I	2	70 717	at/taat
*Asn*I	2	70 717	at/taat
*Asp*700I	2	215 1378	gaann/nnttc
*Asp*EI	2	364 1601	gacnnn/nngtc
*Asp*HI	1	534	gwgcw/c
*Asp*I	1	1901	gacn/nngtc
*Ats*I	1	1901	gacn/nngtc
*Ban*I	5	581 670 688 835 1544	g/gyrcc
*Bbe*I	3	674 692 839	ggcgc/c
*Bbi*II	6	249 288 540 671 689 836	gr/cgyc
*Bbv*12I	1	534	gwgcw/c
*Bcg*I	2	419 593	cgannnnnntgc
*Bgl*I	3	392 446 548	gccnnnn/nggc
*Bpm*I	1	1603	ctggag
*Bsa*AI	1	1003	yac/gtr
*Bsa*HI	6	249 288 540 671 689 836	gr/cgyc
*Bsa*OI	1	189	cgry/cg
*Bsa*WI	1	1605	w/ccggw
*Bse*118I	2	685 838	r/ccggy
*Bse*21I	1	1835	cc/tnagg
*Bse*AI	1	1605	t/ccgga
*Bse*RI	4	604 621 802 1554	gaggag
*Bsh*1285I	1	189	cgry/cg
*Bsh*NI	5	581 670 688 835 1544	g/gyrcc
*Bsi*EI	1	189	cgry/cg
*Bsi*HKAI	1	534	gwgcw/c
*Bsi*MI	1	1605	t/ccgga
*Bsm*BI	2	325 1188	cgtctc
*Bsp*13I	1	1605	t/ccgga
*Bsp*143II	3	674 692 839	rgcgc/y
*Bsp*19I	3	378 390 692	c/catgg
*Bsp*68I	1	614	tcg/cga
*Bsp*CI	1	189	cgat/cg
*Bsp*EI	1	1605	t/ccgga
*Bsp*HI	1	1567	t/catga
*Bsr*BI	1	515	gagcgg
*Bsr*DI	1	169	gcaatg
*Bsr*FI	2	685 838	r/ccggy

Table 2-4 (continued). Internal restriction sites of soybean nodule *IMPDH* cDNA

*Bss*AI	2	685 838	r/ccggy
*Bss*T1I	3	378 390 692	c/cwwgg
*Bst*D102I	1	515	gagcgg
*Bst*DSI	6	291 378 390 522 692 1156	c/crygg
*Bst*H2I	3	674 692 839	rgcgc/y
*Bst*MCI	1	189	cgry/cg
*Bst*SFI	3	843 1556 1720	c/tryag
*Bst*X2I	1	1580	r/gatcy
*Bst*XI	1	166	ccannnnn/ntgg
*Bst*YI	1	1580	r/gatcy
*Bsu*36I	1	1835	cc/tnagg
*Cfr*10I	2	685 838	r/ccggy
*Cfr*I	3	156 381 450	y/ggccr
*Cvn*I	1	1835	cc/tnagg
*Dra*I	1	107	ttt/aaa
*Dra*II	1	549	rg/gnccy
*Dsa*I	6	291 378 390 522 692 1156	c/crygg
*Eae*I	3	156 381 450	y/ggccr
*Eam*1104I	1	220	ctcttc
*Eam*1105I	2	364 1601	gacnnn/nngtc
*Ear*I	1	220	ctcttc
*Ecl*HKI	2	364 1601	gacnnn/nngtc
*Eco*130I	3	378 390 692	c/cwwgg
*Eco*255I	1	1979	agt/act
*Eco*32I	1	1540	gat/atc
*Eco*57I	1	1848	ctgaag
*Eco*64I	5	581 670 688 835 1544	g/gyrcc
*Eco*81I	1	1835	cc/tnagg
*Eco*O109I	1	549	rg/gnccy
*Eco*RV	1	1540	gat/atc
*Eco*T14I	3	378 390 692	c/cwwgg
*Ehe*I	3	672 690 837	ggc/gcc
*Erh*I	3	378 390 692	c/cwwgg
*Esp*3I	2	325 1188	cgtctc
*Gsu*I	1	1603	ctggag
*Hae*II	3	674 692 839	rgcgc/y
*Hin*1I	6	249 288 540 671 689 836	gr/cgyc
*Hinc*II	2	1281 1352	gty/rac
*Hind*II	2	1281 1352	gty/rac
*Hind*III	1	1297	a/agctt
*Hsp*92I	6	249 288 540 671 689 836	gr/cgyc
*Kas*I	3	670 688 835	g/gcgcc
*Kpn*2I	1	1605	t/ccgga
*Ksp*632I	1	220	ctcttc
*Mfl*I	1	1580	r/gatcy
*Mro*I	1	1605	t/ccgga
*Mro*NI	2	685 838	g/ccggc
*Msl*I	1	164	caynn/nnrtg
*Msp*17I	6	249 288 540 671 689 836	gr/cgyc
*Nae*I	2	687 840	gcc/ggc
*Nar*I	3	671 689 836	gg/cgcc
*Nco*I	3	378 390 692	c/catgg
*Ngo*AIV	2	685 838	g/ccggc

Table 2-4 (continued). Internal restriction sites of soybean nodule *IMPDH* cDNA

*Ngo*MI	2	685 838	g/ccggc
*Nru*I	1	614	tcg/cga
*Nsp*I	1	1943	rcatg/y
*Ple*19I	1	189	cgat/cg
*Psh*BI	2	70 717	at/taat
*Pst*I	1	1724	ctgca/g
*Pvu*I	1	189	cgat/cg
*Rca*I	1	1567	t/catga
*Sap*I	1	220	gctcttc
*Sca*I	1	1979	agt/act
*Sfc*I	3	843 1556 1720	c/tryag
*Ssp*I	1	81	aat/att
*Sty*I	3	378 390 692	c/cwwgg
*Tth*111I	1	1901	gacn/nngtc
*Vsp*I	2	70 717	at/taat
*Xba*I	1	22	t/ctaga
*Xho*II	1	1580	r/gatcy
*Xmn*I	2	215 1378	gaann/nnttc

The following endonucleases were selected but don't cut this sequence:

*Aat*I, *Acc*16I, *Acc*65I, *Acc*B7I, *Acc*I, *Acl*NI, *Afe*I, *Afl*II, *Age*I, *Alw*44I, *Alw*NI, *Ama*87I, *Aor*51HI, *Apa*I, *Apa*LI, *Asc*I, *Asp*718I, *Ava*I, *Avi*II, *Avr*II, *Bal*I, **BamHI**, *Ban*II, *Ban*III, *Bbr*PI, *Bbs*I, *Bbu*I, *Bbv*16II, *Bcl*I, *Bco*I, *Bfr*I, *Bgl*II, *Bln*I, *Blp*I, *Bpi*I, *Bpu*1102I, *Bpu*14I, *Bpu*AI, *Bsa*29I, *Bsa*BI, *Bsa*I, *Bsa*MI, *Bsc*I, *Bse*8I, *Bse*CI, *Bse*PI, *Bsg*I, *Bsh*1365I, *Bsi*I, *Bsi*WI, *Bsm*I, *Bso*BI, *Bsp*106I, *Bsp*119I, *Bsp*120I, *Bsp*1407I, *Bsp*1720I, *Bsp*DI, *Bsp*LU11I, *Bsp*MI, *Bsp*TI, *Bsp*XI, *Bsr*BRI, *Bsr*GI, *Bss*HII, *Bss*SI, *Bst*1107I, *Bst*98I, *Bst*BI, *Bst*EII, *Bst*I, *Bst*PI, *Bst*SNI, *Bst*ZI, *Bsu*15I, *Cci*NI, *Cel*II, *Cfr*42I, *Cfr*9I, *Cla*I, *Cpo*I, *Csp*45I, *Csp*I, *Dra*III, *Drd*I, *Eag*I, *Ecl*136II, *Ecl*XI, *Eco*105I, *Eco*147I, *Eco*24I, *Eco*31I, *Eco*47III, *Eco*52I, *Eco*72I, *Eco*88I, *Eco*91I, *Eco*ICRI, *Eco*NI, *Eco*O65I, **EcoRI**, *Eco*T22I, *Esp*1396I, *Fau*NDI, *Fba*I, *Fri*OI, *Fse*I, *Fsp*I, *Hpa*I, **KpnI**, *Ksp*22I, *Ksp*I, *Lsp*I, *Mam*I, *Mfe*I, *Mlu*I, *Mlu*NI, *Mph*1103I, *Msc*I, *Msp*A1I, *Msp*CI, *Mun*I, *Mva*1269I, *Nde*I, *Nhe*I, *Not*I, *Nsi*I, *Nsp*BII, *Nsp*V, *Pac*I, *Pae*I, *Pae*R7I, *Pfl*23II, *Pfl*MI, *Pin*AI, *Pma*CI, *Pme*55I, *Pme*I, *Pml*I, *Ppu*10I, *Ppu*MI, *Psh*AI, *Psp*124BI, *Psp*5II, *Psp*AI, *Psp*ALI, *Psp*EI, *Psp*LI, *Psp*OMI, *Pst*NHI, **PvuII**, *Rsr*II, **SacI**, *Sac*II, *Sal*I, *Sbf*I, *Sex*AI, *Sfi*I, *Sfr*274I, *Sfr*303I, *Sfu*I, *Sgf*I, *Sgr*AI, *Sma*I, *Smi*I, *Sna*BI, *Spe*I, *Sph*I, *Spl*I, *Srf*I, *Sse*8387I, *Sse*BI, *Ssp*BI, *Sst*I, *Sst*II, *Stu*I, *Sun*I, *Swa*I, *Van*91I, *Vha*464I, *Vne*I, *Xcm*I, **XhoI**, *Xma*I, *Xma*III, *Zsp*2I

* Enzymes in bold letters were used in Chapter 3 for soybean genomic DNA digestion and *IMPDH* subcloning.

DISCUSSION

cDNA library screening

Bacteriophage lambda vectors λgt10 and λgt11 were developed in the early 1980s by Davies and colleagues (Young and Davies, 1983a,b; Huynh *et al.*, 1985). λgt11 is a cloning and expression vector, with a single *Eco*R I site located within the *lac*Z gene and can accept foreign DNA of up to 7.2 kb in length (Figure 2-1). λgt11 produces blue plaques on *lac*$^-$ hosts in the presence of IPTG (isopropyl β-D-thiogalactopyranoside) and X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), whereas recombinant phage produce colorless plaques because of the inactivation of β-galactosidase.

The cDNA library used for screening was constructed in λgt11 from 30 to 35-day-old soybean nodules (Ji *et al.*, 1994). In the process of soybean nodule cDNA library screening for IMPDH, a large number of positive clones were obtained after every primary screening, and a large proportion of them were further confirmed as positive signals in the secondary screening. This was especially obvious in library screening II (Table 2-2, 2-3) in which oligonucleotide probes were synthesized based on the partial soybean nodule *IMPDH* cDNA sequence. The number of positive signals could be an indication that the *IMPDH* cDNA was relatively abundant in the library even though many of them contained only incomplete *IMPDH* cDNA fragments. The abundance of the cDNA might further indicate that the *IMPDH* gene is highly expressed in 30 to 35-day-old soybean nodules.

Because degenerate probes were used in the first screening and oligonucleotides from *IMPDH* cDNA fragments were used as probes in second screening, the hybridization in screening II was more specific, resulting in more and stronger signals (Figure 2-3).

Sequence analysis of clone #46

Due to the laborious isolation procedure and low yield of phage DNA, the target clones of λgt11 plaques were subcloned into plasmid (phagemid) pBluescript SK (-). Plasmid DNA inserts were then isolated and used for sequencing. Sequencing results showed that clone #46 contained a 2016 base pair insert. After translation and BLAST analysis, clone #46 yielded very high similarity scores with IMPDHs from *Arabidopsis*, humans, mice, fungi and bacteria. The signature of IMPDH, the putative active-site region for IMP binding, was 100% identical between clone #46 and *Arabidopsis* IMPDH. On this basis, we proposed that this cDNA coded for soybean nodule IMPDH. The plasmid pBluescript SK (-) containing soybean nodule *IMPDH* was designated as p46-10.

This 2016 bp *IMPDH* has a 166 bp 5'-untranslated region and a 341 bp 3'-untranslated sequence. The typical consensus poly (A) signal in animal cDNA, AAUAAA, was not found in the 3' end of soybean nodule *IMPDH*, but a match of four of the six bases appears several times (Figure 2-6). This is consistent with the sequence for the poly(A) signal for most plant mRNAs (Hughes, 1996). Rothie *et al.* (1994) also reported that the AAUAAA motif was not used as a plant polyadenylation signal.

There are several reasons why we consider that our assignment of the ATG-initiation codon at nucleotide position 167 is correct: 1) BLAST analysis showed that the methionine (Met) at position 167 is the closest one to the start codons of IMPDHs in other organisms like *Arabidopsis*, human, mouse and yeast; 2) It is suggested that the 3'- and 5'- flanking sequences of the ATG triplets could be crucial in determining the translation initiation site (Kozak, 1983, 1984; Joshi, 1987). A purine "A" or "G" frequently occurs three residues before the ATG codon, and a purine "G" often follows the ATG, forming a typical eukaryotic translation initiation consensus sequence of "CC(A/G)CCATGG" (Kozak, 1984). In soybean *IMPDH*, the DNA sequence coding for Met at nucleotide position 167 forms a "CCGCAATGG" with its flanking sequences. This is fairly consistent with the plant consensus initiation sequence, thus making it a reasonable initiation site for translation; 3) There were no other Met codons upstream from position 167, while the next Met codon downstream from position 167 was 63 amino acids away at position 356. The Met codon in position 356 could generate a protein product with a molecular mass around 46 kDa if this Met functions as the translational initiation site. This molecular weight is less than the reported molecular weights for IMPDH (55 and 60 kDa) from soybean nodules (Yang, 1997). Also, based on the analysis of 211 messenger RNAs from higher eukaryotic cells, Kozak (1984) concluded that "An AUG triplet that deviates from the consensus in the crucial –3 position *can nevertheless serve as the initiator codon*". The Met at position 356 has a pyrimidine (C) 3 nucleotides upstream making this site unlikely as the start codon; 4) Characterization of the *IMPDH*

cDNA in Chapter 3 will provide further support for the conclusion that the postulated initiation site at nucleotide 167 is correct.

Up to now, *A. thaliana* was the only plant from which an *IMPDH* cDNA had been cloned and sequenced (Collart *et al.*, 1996a). No other cDNA sequence or amino acid sequence for IMPDH from plants has been reported. Compared to the deduced amino acid sequence of *A. thaliana* IMPDH, soybean nodule IMPDH is one residue shorter (Figure 2-7). There was a high percentage of relatedness (70.5% identity) between soybean nodule IMPDH and *A. thaliana* IMPDH. A putative active-site region (Andrews and Guest, 1988; Antonino *et al.*, 1994) is present in soybean nodule IMPDH (position 311-331). This site displays a remarkable degree of amino acid conservation, i.e. the sequence was 100% identical between IMPDHs from soybean nodules and *A. thaliana*.

The results of an analysis for internal restriction sites are displayed in Figure 2-8 and Table 2-4. These results provide important reference for further characterization and subcloning of *IMPDH* cDNA in Chapter 3 and Chapter 4 of this dissertation. The sequence of the cloned *IMPDH* cDNA (clone #46) was registered in EMBL, Genebank and DDBJ Nucleotide Sequence Database with an accession number of AJ010201 in August, 1998. There was no evidence for heterogeneity within clones.

REFERENCES

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

Andrews, S.C. and Guest, J.R. (1988). Nucleotide sequence of the gene encoding the GMP reductase of *Escherichia coli* K12. Biochem J 255: 35-43.

Antonino, L.C., Straub, K. and Wu, J.C. (1994). Probing the active site of human IMP dehydrogenase using halogenated purine riboside 5'-monophosphates and covalent modificatin reagents. Biochemistry 33: 1760-1765.

Atkins, C.A., Shelp, B.J. and Storer, P.J. (1985). Purification and properties of inosine monophosphate oxidoreductase from nitogen-fixing nodules of cowpea (*Vigna unguiculata* L Walp). Arch Biochem Biophys 236: 807-814.

Ausubel, F.M., Brent, R., Kingston, R.E., Mooew, D.D., Seidman, J.G. and Smith, J.S. (1991). Current Protocols in Molecular Biology. Wiley, New York.

Bairoch, A. (1995). Prosite Dictionary: Release 12.2. University of Geneva, Geneva, Switzerland.

Boland, M.J. and Schubert, K.R. (1983). Biosynthesis of purines by a proplastid fraction from soybean nodules. Arch Biochem Biophys 220: 179-187.

Chung, C.T., Niemela, S.L. and Miller, R H. (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci USA 86: 2171-2175.

Collart, F.R., Osipiuk J., Trent, J., Olsen, G. and Huberman, E. (1996a). Cloning and characterization of the gene encoding IMP dehydrogenase from *Arabidopsis thaliana*. Gene 174: 217-220.

Collart, F.R., Osipiuk, J., Trent, J., Olsen, G. and Huberman, E. (1996b). Cloning, characterization and sequence comparison of the gene encoding IMP dehydrogenase from *Pyrococcus furiosus*. Gene 174: 209-216.

Dayton, J.S., Lindsten, T., Thompson, C.B. and Mitchell, B.S. (1994). Effects of human T lymphocyte activation on inosine monophosphate dehydrogenase expression. J Immunol 152: 984-991.

Goldblum, R. (1993). Therapy of rheumatoid arthritis with mycophenolate mofetil. Clin Exp Rheumatol 11: S117-S119.

Haymes, K.M. (1996). Mini-prep method suitable for a plant breeding program. Plant Molecular Biology Report 14(3): 280-284.

Heiman, M. (1997). http//wwwfirstmarketcom/cutter/cut2html. Webcutter 20, copyright 1997.

Higgins, D., Thompson, J., Gibson, T., Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W. improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

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

Huang, X. and Miller, W. A. (1991). Time-efficient, linear-space local similarity algorithm. Adv App Math 12: 337-357.

Huberman, E., Glesne, D. and Collart, F.R. (1995). Regulation and role of inosine-5'-monophosphate dehydrogenase in cell replication, malignant rransformation, and differentiation. In: Purine and pyrimidine metabolism in man, VIII (Sahota, A. and Taylor, M. Ed.). Plenum Press, New York, pp741-746.

Hughes, M.A. (1996). Plant molecular genetics. Addison Wesley Longman Limited, New York, pp15.

Hupe, D.J., Azolina, B.A. and Behrens, N.D. (1986). IMP dehydrogenase from the intracellular parasitic protozoan *Eimeria tenella* and its inhibition by mycophenolic acid. J Biol Chem 261:8363-8369.

Huynh, D.J. and Davies, R.W. (1985). DNA cloning techniques: a practical approach. pp49-78.

Jenkins, D.C., Stables, J.N., Wilkinson, J., Topley, P., Holmes, L.S., Linstead, D.J. and Rapson, E.B. (1993). A novel cell-based assay for the evaluation of anti-*ras* compounds. Br J Cancer 68: 856-861.

Ji, L., Becana, M., Sarath, G. and Klucas, R.V. (1994). Cloning and sequence analysis of a cDNA encoding ferric leghemoglobin reductase from soybean nodules. Plant Physiol 104: 453-459.

Joshi, C.P. (1987). An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Res 15:6643-6653.

Kozak, M. (1983). Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. Microbiology Rev 47: 1-45.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNAs. Nucleic Acids Res 12: 857-872.

Rothnie, H.M., Reid, J. and Hohn, T. (1994). The contribution of AAUAAA and the upstream element UUUGUA to the efficiency of mRNA 3'-end formation in plants. The EMBO J 13: 2200-2210.

Schubert, K.R. (1981). Enzymes of purine biosynthesis and catabolism in *Glycine max.* Plant Physiol 68:1115-1122.

Schubert, K.R. (1986). Products of biological nitrogen fixation in higher plants: synthesis, transport, and metabolism. Ann Rev Plant Physiol 37:539-74.

Schubert, K.R. and Boland, M.J. (1990). Ureides. In: the biochemistry of plants (Stumpf, P.K. & Conn, E.E. Ed.) 16: 197-282. Academic Press, New York.

Shelp, B.J. and Atkins, C.A. (1983). Role of inosine monophosphate oxidoreductase in the formation of ureides in nitrogen-fixing nodules of cowpea (*Vigna unguiculata LWalp*). Plant Physiol 72: 1029-1034.

Tricot, G., Jayaram, H.N., Weber, G. and Hoffman, R. (1990). Tiazofurin: biological effects and clinical uses. Int J Cell Cloning 8: 161-170.

Weber, G. (1983). Biochemical strategy of cancer cells and design of chemotherapy. Cancer Res 43:3466-3492.

Yang, D. (1997). Purification and characterization of inosine-5'-monophosphate dehydrogenase (IMPDH) from nitrogen-fixing nodules of soybean (*Glycine max*). Master's thesis, University of Oklahoma, Norman, OK, USA.

Young, R.A. and Davies, R.W. (1983a). Efficient isolation of genes by using antibody probes. Proc Natl Acad Sci USA 80: 1194-1198.

Young, R.A. and Davies, R.W. (1983b). Yeast RNA polymerase genes: isolation with antibody probes. Science 222: 778-782.

Chapter 3. Characterization of Soybean Nodule *IMPDH* cDNA

INTRODUCTION

IMPDH has been studied principally in mammals, yeast, fungi and bacteria (SWISS-PROT database; Table 3-1). Data from a range of organisms revealed that IMPDH is a multimeric enzyme composed of identical subunits of approximately 50-60 kDa (Beck *et al.*, 1993; Ikegami *et al.*, 1987; Krishnaiah, 1975; Verham *et al.*, 1987; Wilson *et al.*, 1994). Table 3-1 is a summary of the subunit molecular weight (MW) of IMPDH deduced from *IMPDH* cDNA or genomic clones.

The only *IMPDH* cDNA cloned from a plant source was isolated from *Arabidopsis thaliana* (Collart *et al.*, 1996a). Although IMPDH has been purified from the nodules of cowpea (Atkins *et al.*, 1985) and soybean (Yang, 1997), neither of these purified proteins has been sequenced. The molecular weight of IMPDH subunits was reported to be 50 kDa in cowpea. Two bands of 55 and 60 kDa respectively were observed in fractions for purified soybean IMPDH by SDS-PAGE.

As the first *IMPDH* cDNA cloned from legumes and the second from a plant source, soybean nodule *IMPDH* cDNA needs to be further characterized. Sequence analysis using molecular biology tools will provide information on the properties of soybean nodule IMPDH. This information includes the degree of homology between soybean nodule IMPDH and other IMPDHs, predicted primary and secondary structure, molecular weight, isoelectic point and subcellular localization. Southern blot analysis will help to estimate the number

of *IMPDH* genes in the soybean genome and Northern blot analysis can be used to determine the tissue-specificity of *IMPDH* gene expression.

Table 3-1. Subunit molecular weight (MW) of IMPDH deduced from the open reading frame of cloned *IMPDH* cDNAs / genes

Organism	MW (Da)	References
Prokaryotes		
Acinetobacter calcoaceticus	51530	Anderegg *et al.*, 1992
Borrelia burgdorferi	43767	Margolis *et al.*, 1994; Fraser *et al.*, 1997
Chlorobium vibrioforme	56627	Peterson *et al.*, 1996
Haemophilus influenzae	51981	Fleischmann *et al.*, 1995
Helicobacter pylori	51802	Tomb *et al.*, 1997
Leishmania donovani	55551	Wilson *et al.*, 1991
Methanococcus jannaschii	53316	Bult *et al.*, 1996
Methanopyrus kandleri (ORFX)	19347	Nolling *et al.*, 1995
Mycobacterium leprae	54814	Smith & Robison, 1994
Mycobacterium tuberculosis	54867	Oliver *et al.*, 1996
Pyrococcus furiosus	52900	Collart *et al.*. 1996b
Streptococcus pyogenes	52807	Ashbaugh & Wessels, 1995
Escherichia coli	52022	Tiedeman & Smith, 1985; Thomas & Drabble, 1985; Link *et al.*, 1997; Molloy *et al.*, 1998
Bacillus subtilis	55725	Kanzaki & Miyagawa, 1990; Ogasawara *et al.*, 1994; Antelmann *et al.*, 1997

Table 3-1 (continued). Subunit molecular weight (MW) of IMPDH deduced from the open reading frame of cloned *IMPDH* cDNAs / genes

Eukaryotes		
Arabidopsis thaliana (plant)	54194	Collart *et al.*, 1996a
Homo sapiens (human) (IMPDH-I)	55449	Natsumeda *et al.*, 1990; Hager *et al.*, 1995
Homo sapiens (human)(IMPDH-II)	55805	Collart & Huberman, 1988; Natsumeda *et al.*, 1990; Glesne & Huberman, 1994; Zimmermann *et al.*, 1995; Hager *et al.*, 1995
Mus musculus (mouse)(IMPDH-1)	55294	Dayton and Mitchell, 1993
Mus musculus (mouse) (IMPDH-II)	55785	Tiedeman & Smith, 1991; Lightfoot & Snyder, 1994
Drosophila melanogaster (fruit fly)	57829	Nash & Hu, 1992; Sifri *et al.*, 1994; Slee & Bownes, 1995
Mesocricetus auratus (IMPDH-II)	55890	Collart & Huberman, 1988
Saccharomyces cerevisiae (baker's yeast)	56530	Johnston *et al.*, 1994
Saccharomyces cerevisiae (baker's yeast)	56584	Johnston *et al.*, 1995
Saccharomyces cerevisiae (baker's yeast)	56394	Devlin *et al.*, 1994
Saccharomyces cerevisiae (baker's yeast)	44386	Bussey *et al.*, 1995
Candida albicans (yeast)	56239	Kohler *et al.*, 1997
Tritrichomonas foetus	55473	Beck *et al.*, 1993; Huete-perez *et al.*, 1995; Whitby *et al.*, 1997
Pneumocystis carinii	49522	O'Gara *et al.*, 1997
Trypanosoma brucei brucei	55708	Wilson *et al.*, 1994

MATERIALS AND METHODS

Plant materials

Soybean (*Glycine max* L. Merr. cv. Essex) seeds were obtained from Pioneer Hi-Bred International, Inc. (Des Moines, Iowa). Seeds were infected with liquid culture of *Bradyrhizobium japonicum* strain USDA 110 before planting in the greenhouse. Plants were grown in pots containing medium-grade Agricultural Vermiculite (Construction Products, W.R. Grace & Conn., 62 Whittemore Ave, Cambridge, MA 02140-1692). N-free nutrient solution, as modified by Dr. Sengupta-Gopalan from New Mexico State University at Las Cruses (personal communication), was applied daily during soybean growth. Detailed protocols are presented in the Appendix.

Nodules were harvested from 36-day-old soybean plants grown in the University of Oklahoma Biocontainment Greenhouse. After nodules were detached from the plant, they were immediately frozen in liquid nitrogen and stored at -80°C. Root and leaf tissues were obtained from uninfected soybean plants grown in a growth chamber with 14 h of light per day. Peters Professional Water-Soluble Fertilizer (Pete Light Special 20-10-20) was applied to the plants not infected with *Bradyrhizobium*.

Comparison of deduced amino acid sequence of IMPDH *cDNAs between soybean and other organisms*

Deduced amino acid sequences of IMPDH from other organisms were obtained from the NCBI Genebank database. Comparisons between deduced soybean

nodule IMPDH and other IMPDHs were carried out and the percentage identity scores were obtained by using the SIM-local similarity program from the ExPASy molecular biology website of the Swiss Institute of Bioinformatics (SIB) (Huang and Miller, 1991). The Clustal W program (Higgins *et al.*, 1994) from the European Bioinformation Institute was used to align the predicted amino acid sequences of soybean nodule IMPDH and IMPDHs from other organisms. The PROSITE (database of protein families and domains) program (Hofmann *et al.*, 1999) was used to search for the IMPDH signature sequence from the deduced soybean nodule IMPDH amino acid sequence.

Prediction of hydrophobicity, subcelluar localization, MW and pI

The primary structure analysis program in ExPASy --- "ProScale" was used for the analysis of hydrophobicity / hydrophilicity of the deduced IMPDH sequence (Kyte and Doolittle, 1982). "PSORT" (Nakai, 1991; Nakai and Kanehisa, 1992) was used to identify potential protein sorting signals to predict subcellular localization. The program "SignalP" (Nielsen *et al.*, 1997) was used to predict potential signal peptide cleavage sites. The isoelectric point (pI) and molecular weight (MW) were predicted using the "Compute pI / MW" program (Bjellqvist *et al.*, 1993; Bjellqvist *et al.*, 1994; Wilkins *et al.*, 1998).

Southern blot analysis

Genomic DNA was isolated from 36-day-old nodules and from soybean leaves according to the method of Haymes (1996). Individual samples (10 μg) of total

DNA were digested with each of the following 7 restriction enzymes: *Bam*H I, *Eco*R I, *Kpn* I, *Pvu* II, *Sac* I, *Pst* I and *Xba* I. Digestions were carried out overnight at room temperature. DNAs were separated by electrophoresis on a 0.8% agarose gel with a running buffer of TBE (0.089 M Tris base, 0.089 M boric acid, 0.002 M EDTA). DNA in the gel was fragmented by soaking the gel in 0.25 M HCl for 30 min. The gel was rinsed with nanopure water and then the DNA was denatured by soaking the gel in 1.0 M NaCl / 0.5 M NaOH for 20 min. The denaturation step was repeated a second time. After treatment with neutralization solution (0.5 M Tris, pH 7.5; 1.5 M NaCl) two times for 20 min each, the gel was transferred to a nitrocellulose membrane using standard procedures (Ausubel *et al.*, 1991). After an overnight transfer, the membrane was washed with 10 x SSC for 15 min and baked in a vacuum oven at 80°C for 2 hours.

The *IMPDH* cDNA clone was labeled with α-^{32}P-dCTP by the random labeling method with the Ready-To-Go® labeling kit (Amersham Pharmacia Biotech, Inc., Piscataway, New Jersey). Pre-hybridization was performed overnight at 42°C in pre-hybridization solution containing 50% formamide, 5x Denhardts' solution, 5x SSC, 0.1% SDS and 100 µg/ml denatured salmon sperm DNA (Wadsworth *et al.*, 1993). Hybridization was carried out in the same solution overnight at 42°C except that the probe was added to a final concentration of 50 ng / 50 ml solution. The final stringency of the wash was 0.1 x SSC + 0.1% (w/v) SDS at 65°C. Plasmid DNA containing *IMPDH* cDNA was used as a positive control. In order to make sure that the digestion was complete, genomic DNA without restriction enzyme digestion was used as a control.

Northern blot analysis

Total RNA was isolated from 36-day-old soybean nodules, 12-day-old uninfected roots and 12-day-old leaves from uninfected plants by the guanidinium thiocyanate-phenol-chloroform extraction method (Ausubel *et al.*, 1991). Eight μg of total RNA was incubated with 4 volumes of 5x RNA loading buffer at 65°C for 3-5 min. 5x RNA loading buffer was prepared by combining 16 μl saturated bromophenol blue, 80 μl 500 mM EDTA, pH 8.0, 720 μl 37% formaldehyde, 2 ml 100% glycerol, 3084 μl formamide, 4 ml 10x FA gel buffer (pH 7.0) and RNase-free water to a final volume of 10 ml. The 10x FA gel buffer contained 200 mM MOPS (3-[N-morpholino]propane sulfonic acid), 50 mM sodium acetate and 10 mM EDTA, pH 8.0. After incubation, the denatured RNA was separated by electrophoresis on a 1.2% agarose gel containing 1x MOPS and 0.22 M formaldehyde (1.8 ml 37% formaldehyde in 100 ml solution). Electrophoresis was carried out at 80 V for 1.5-2 h.

After electrophoresis, the gel was treated with 50 mM NaOH for 15 min followed by two 20-min washes with 0.2 M NaOAc, pH 4.0. RNA was transferred to nitrocellulose membranes using the same procedure as used for the Southern blots. The Northern blot was probed with α-^{32}P-dCTP-labeled *IMPDH* cDNA. The labeling of the probe, pre-hybridization and hybridization procedures were the same as those used for Southern blot analysis. The stringency of the final wash was 0.1x SSC, 0.1% (w/v) SDS at room temperature.

RESULTS

Comparison of deduced amino acid sequence of IMPDH from soybean and other organisms

The relatedness of the deduced amino acid sequence of soybean nodule IMPDH and IMPDHs from other organisms is presented in Table 3-2. Soybean nodule IMPDH has the highest identity with *Arabidopsis thaliana* IMPDH and the lowest with IMPDH from bacteria. The soybean nodule IMPDH polypeptide is also one of the smallest of the eukaryotic IMPDH polypeptides. The alignment and comparison of the complete sequences of soybean IMPDH and IMPDHs from plants, mammals, insects, fungi and bacteria are shown in Figure 3-1.

A comparison of the consensus sequence of the proposed IMPDH catalytic domain from soybean and other organisms is shown in Figure 3-2. Twelve out of 20 residues (60%) were identical when the catalytic domain of soybean IMPDH was compared with the domain from other eukaryotes, while only three out of 20 residues (15%) were identical between soybean and all prokaryotes.

Prediction of hydrophobicity, subcellular localization, MW and pI

A hydrophobicity plot for soybean nodule IMPDH is presented in Figure 3-3. No distinctive hydrophobic or hydrophilic regions were noted. The prediction for protein sorting signals and localization site suggested that there was no N-terminal signal sequence, indicating that the protein might be localized in the cytoplasm. The predicted MW was 53 kDa and the predicted pI was 5.54.

Southern blot analysis

Results of the Southern blot analysis are shown in Figure 3-4. After hybridization with the soybean nodule *IMPDH* cDNA clone, all the soybean genomic DNA digestions gave only one band except *Kpn* I (lane 2) and *Pst* I (lane 5). Both *Kpn* I and *Pst* I gave two bands, one band was very strong while the other was very weak.

Northern blot analysis

The results of the Northern blot analysis, performed under stringent conditions, are shown in Figure 3-5. Staining with EtBr was similar in each sample suggesting that an equal amount of total RNA from nodules, roots and leaves was loaded in each lane (left side of the figure). After hybridization, a major RNA hybridizing band of approximately 2.0 kb was observed only in 36-day-old nodules, while signals in roots and leaves were too weak to be detectable (right side of the figure).

Table 3-2. Comparison of the deduced amino acid sequence of soybean nodule IMPDH with sequences of other inosine monophosphate dehydrogenases

Organism	Accession #	Length (a.a.)	Identity (%)*
Arabidopsis thaliana	P47996	503	70.5
Homo sapiens (IMPDH-I)	P20839	514	49.2
Homo sapiens (IMPDH-II)	P12268	514	49.2
Mus musculus (IMPDH-I)	P50096	514	49.1
Mus musculus (IMPDH-II)	P24547, Q61734	514	49.2
Mesocricetus auratus	P12269	514	49.0
Drosophila melanogaster	Q07152, Q26455	537	42.8
Saccharomyces cerevisiae	P38697	523	46.4
Saccharomyces cerevisiae	P50095	523	47.4
Saccharomyces cerevisiae	P50094	524	42.1
Saccharomyces cerevisiae	P39567	403	42.7
Pneumocystis carinii	Q12658	454	44.5
Candida albicans 3	O00086	521	44.1
Trypanosoma brueibrucei	P50098	512	42.1
Leishmania donovani	P21620	514	42.1
Borrelia burgdorferi	P49058	404	39.2
Helicobacter pylori	P56088	481	35.8
Actinobacter calcoaceticus	P31002	488	35.9
Methanococcus jannaschii	Q59011	496	33.2
Escherichia coli	P06981, P76574, 78202	488	34.1
Streptococcus pyogenes	P50099	493	34.4
Bacillus subtilis	P21879	513	35.0
Pyrococcus furiosus	P42851	485	32.2
Haemophilus influenzae	P44334	488	33.2
Mycobacterium tuberculosis	Q50715	529	37.8
Mycobacterium leprae	Q49725	529	36.4
Chlorobium vibrioforme	O50316	521	34.8
Tritrichomonas foetus	P50097	503	28.9

* Percentage identity was obtained by using the SIM-local similarity program from the ExPASy molecular biology website of the Swiss Institute of Bioinformatics (SIB).

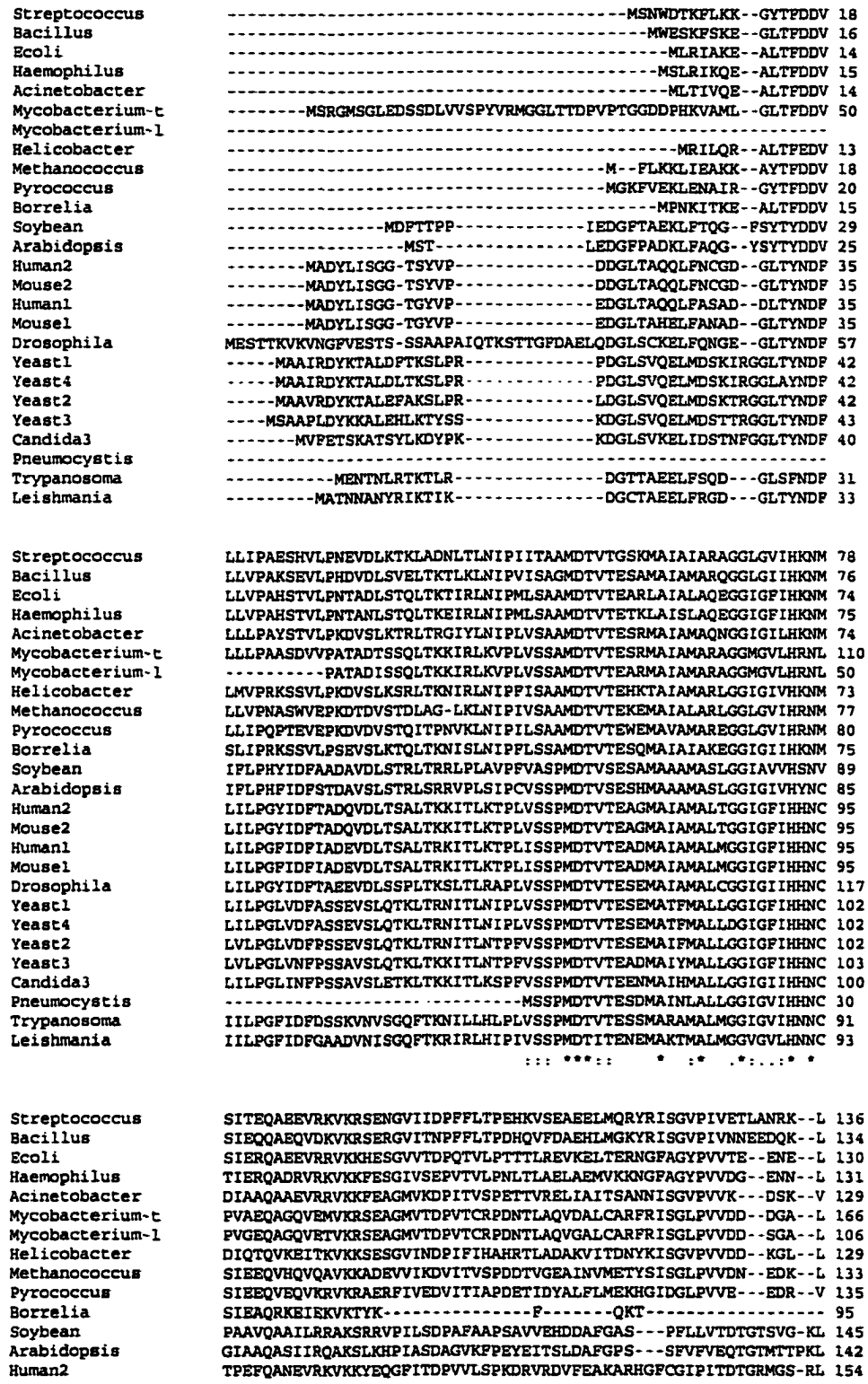


Figure 3-1. Alignment and comparison of complete sequences of IMPDHs from different organisms using the Clustal W program (Higgins *et al.*, 1994).

```
Mouse2          TPEPQANEVRKVKKYEQGPITDPVVLSPKDRVRDVFEAKARHGFCGIPITDTGRMGS-RL 154
Human1          TPEFQANEVRKVKNFEQGPITDPVVLSPSHTVGDVLEAKMRHGPSGIPITETGTMGS-KL 154
Mouse1          TPEPQANEVRKVKKPEQGPITDPVVLSPSHTVGDVLEAKIQHGPSGIPITATGTMGS-KL 154
Drosophila      TPEYQALEVHKVKKYKHGFMRDPSVMSPTNTVGDVLEARRKNGFTGYPVTENGKLGG-KL 176
Yeast1          TPEDQADMVRRVKNYENGFINNPIVISPTTTVGEAKSMKEKYGPAGFPVTTDGKRNA-KL 161
Yeast4          TPEDQADMVRRVKNYENGFINNPIVISPTTTVGEAKSMKEKYGPAGFPVTADGKRNA-KL 161
Yeast2          TPEDQADMVRRVKNYENGFINNPIVISPTTTVGEAKSMKERFGPSGFPVTEDGKRNG-KL 161
Yeast3          TPKEQASMVKKVKMFENGFINSPIVISPTTTVGEVKVMKRKPGPSGFPVTEDGKCPG-KL 162
Candida3        TSEEQAEMVRKVKKYENGFINDPVVISPEVTVGEVKKMGEVLGPTSFPVTENGKVGG-KL 159
Pneumocystis    TIEEQTEMVRKVKKFENGFITSPIVLSLNHRVRDVRRIKEELGPSGIPITDTGQLNG-KL 89
Trypanosoma     TVEQQARMVRSVKLYRNGFIMKPKSVSPDVPVSTIRNIKSEKGISGILVTEGGKYDG-KL 150
Leishmania      TVERQVEMVKSVKAYRNGFISKPKSVPPNTPISNIIRIKEEKGISGILVTENGDPHG-KL 152
                     *  :  .*                .

Streptococcus   VGIITNRDMRFI--SDYNAPISEHMT-SEHL-VTAAVGTDLETAERILHEHRIEKLPLVD 192
Bacillus        VGIITNRDLRFI--SDYSMKISDVMT-KEEL-VTASVGTTLDEAEKILQKHKIEKLPLVD 190
Ecoli           VGIITGRDVRFV--TDLNQPVSVYMTPKERL-VTVREGEAREVVLAKMHEKRVEKALVVD 187
Haemophilus     IGIITGRDTRFV--KDLSKTVSQVMTKKEDL-VTVKEGASREEILELMHQHRVEKVLVVN 188
Acinetobacter   VGIVTGRDTRFE--TNLEQPVSNIMTGQDRL-VTVREGESKENIQALLQKHRIEKVLVVG 186
Mycobacterium-t VGIITNRDMRFE--VDQSKQVAEVMT-KAPL-ITAQEGVSASAALGLLRRNKIEKLPVVD 222
Mycobacterium-l AGIITNRDMRFE--VDQSKQVAEVMT-KTPL-ITAAEGVSADAALGLLRRNKIEKLPVVD 162
Helicobacter    IGILTNRDVRFE--TDLSKKVGDVMT-KMPL-VTAHVGISLDEASDLMHKHKIEKLPIVD 185
Methanococcus   VGIITHRDVKAI--EDKTKKVKDVMT--KDV-VCAKEDVEEEEALELMYANRVERLPIVD 188
Pyrococcus      VGIITKKDIAAR--EGRT--VKELMT--REV-ITVPESVDVEEALKIMMENRIDRLPVVN 188
Borrelia        --INTNGDTNEQ--K------PEIFT--------------------AKQH-LEK----- 118
Soybean         LGYVARSDWT--NQTDKGLRVGDYMAPPPKP---APWNADLNKINEIMESEKSG-AVALE 199
Arabidopsis     LGYVTKSQWKRMNYEQREMKIYDYMKSCDSSDYCVPWEIDFEKLEFVLEDKQKG-FVVLE 201
Human2          VGIISSRDIDFLKEEEHDCFLEEIMTKREDL-VVAPAGITLKEANEILQRSKKGKLPIVN 213
Mouse2          VGIISSRDIDFLKEEEHDRFLEEIMTKREDL-VVAPAGVTLKEANEILQRSKKGKLPIVN 213
Human1          VGIVTSRDIDFLAEKDHTTLLSEVMTPRIEL-VVAPAGVTLKEANEILQRSKKGKLPIVN 213
Mouse1          VGIVTSRDIDFLAEKDHTTLLSEVMTPRVEL-VVAPAGVTLKEANEILQRSKKGKLPIVN 213
Drosophila      LGMVTSRDIDFR-ENQPEVLLADIMT--TEL-VTAPNGINLPTANAILEKSKKGKLPIVN 232
Yeast1          VGVITSRDIQFV--EDNSLLVQDVMTK--NP-VTGAQGITLSEGNEILKKIKKGRLLVVD 216
Yeast4          VGAITSRDIQFV--EDNSLLVQDVMTK--NP-VTGAQGITLSEGNEILKKIKKGRLLVVD 216
Yeast2          MGIVTSRDIQFV--EDNSLLVQDVMTK--NP-VTGAQGITLSEGNEILKKIKKGKLLIVD 216
Yeast3          VGLVTSRDIQFL--EDDSLVVSEVMTK--NP-VTGIKGITLKEGNEILKQTKKGKLLIVD 217
Candida3        VGIITSRDIQFH--EDNKSPVSEVMTK--DL-VVGKKGISLTDGNELLRSSKKGKLPIVD 214
Pneumocystis    LGIVTSRDIQFH--NNDESPLSEVITK--DL-VTGSEGIRLEEANEILRSCKKGKLPIVD 144
Trypanosoma     LGIVCTKDIDFV--KDASAPVSQYMTRRENM-TVERYPIKLEEAMDVLNRSRHGYLPVLN 207
Leishmania      LGIVCTKDIDYV--KNKDTPVSAVMTRREKM-TVERAPIQLEEAMDVLNRSRYGYLPIVN 209
                        :                :

Streptococcus   NSGRLSGLITIKDIEKVIEFPH----AAKDEF-GRLLVAAAVGVTSDTFERAEALFEAGA 247
Bacillus        DQNKLKGLITIKDIEKVIEFPN----SSKDIH-GRLIVGAAVGVTGDTMTRVKKLVEANV 245
Ecoli           DEFHLIGMITVKDFQKAERKPN----ACKDEQ-GRLRVGAAVGAGAGNEERVDALVAAGV 242
Haemophilus     DSFKLKGMITVKDFQKAEQKPN----ACKDEF-GRLRVGAAVGAGAGNEERIDALVKAGV 243
Acinetobacter   ESNELKGLITVTDFRKAESYPN----SCKDDL-GRLRVGAAVGTGADTPSRVEALVEAGV 241
Mycobacterium-t GRGRLTGLITVKDFVKTEQHPL----ATKDSD-GRLLVGAAVGVGGDAWVRAMMLVDAGV 277
Mycobacterium-l GHGRLTGLITVKDFVKTEQHPL----ATKDND-GRLLVGAAVGVGGDAWVRAMMLVDAGV 217
Helicobacter    KDNVLKGLITIKDIQKRIEYPE----ANKDDF-GRLRVGAAIGVG--QLDRAEMLVKAGV 238
Methanococcus   DENRLIGIITLRDILKRRKYPQ----AARDKK-GRLLVAAACGP--HDFERAKALIEAEV 241
Pyrococcus      EDGKLVGLITMSDLVARKKYKN----AVRNEK-GELLVAAAVSP--FDLRRAIELDRAGV 241
Borrelia        -SD------AYKNAEHKEDFPN----ACKDLN-NKLRVGAAVSIDIDTIERVEELVKAHV 166
Soybean         RDGEVVDLVVREEVERVRGYPKLVAPATVGAD-GEFMVGAAVGTREDDKERLEHLVKAGL 258
Arabidopsis     RDGETVNVVTKDDIQRVKGYPK-SGPGTVGPD-GEWMVGAAIGTRESDKERLEHLVNVGV 259
Human2          EDDELVAIIARTDLKKNRDYPL----ASKDAK-KQLLCGAAIGTHEDDKYRLDLLAQAGV 268
Mouse2          ENDELVAIIARTDLKKNRDYPL----ASKDAK-KQLLCGAAIGTHEDDKYRLDLLALAGV 268
Human1          DCDELVAIIARTDLKKNRDYPL----ASKDSQ-KQLLCGAAVGTREDDKYRLDLLTQAGV 268
Mouse1          DQDELVAIIARTDLKKNRDYPL----ASKDSH-KQLLCGAAVGTREDDKYRLDLLTQAGA 268
Drosophila      QAGELVAMIARTDLKKARSYPN----ASKDSN-KQLLVGAAIGTRSEDKARLALLVANGV 287
Yeast1          EKGNLVSMLSRTDLMKNQNYPL----ASKSANTKQLLCGASIGTMDADKERLRLLVKAGL 272
Yeast4          EKGNLVSMLSRTDLMKNQKYPL----ASKSANTKQLLWGASIGTMDADKERLRLLVKAGL 272
Yeast2          DNGNLVSMLSRTDLMKNQNYPL----ASKSATTKQLLCGAAIGTIDADKERLRLLVEAGL 272
Yeast3          DNGNLVSMLSRADLMKNQNYPL----ASKSATTKQLLCGAAIGTIEADKERLRLLVEAGL 273
Candida3        AEGNLVSLISRTDLQKNQDYPN----ASKSFHSKQLLCGAAIGTIDADRERLDKLVEAGL 270
Pneumocystis    KEGNLTALLSRSDLMKNLHFPL----ASKLPDSKQLICAQAVGTRPDDRIRLKHLVEAGL 200
Trypanosoma     DKDEVVCLCSRRDAVRARDYPN----SSLDRN-GHLLCAAATSTREADKGRVAALSEAGI 262
Leishmania      ENDEVVNLCSRRDAVRARDYPH----STLDKS-GRLICAAATSTRPEDKRRVAALADVGV 264
                              :                .     .  . : .      *  *
```

Figure 3-1 (continued). Alignment and comparison of complete sequences of IMPDHs from different organisms.

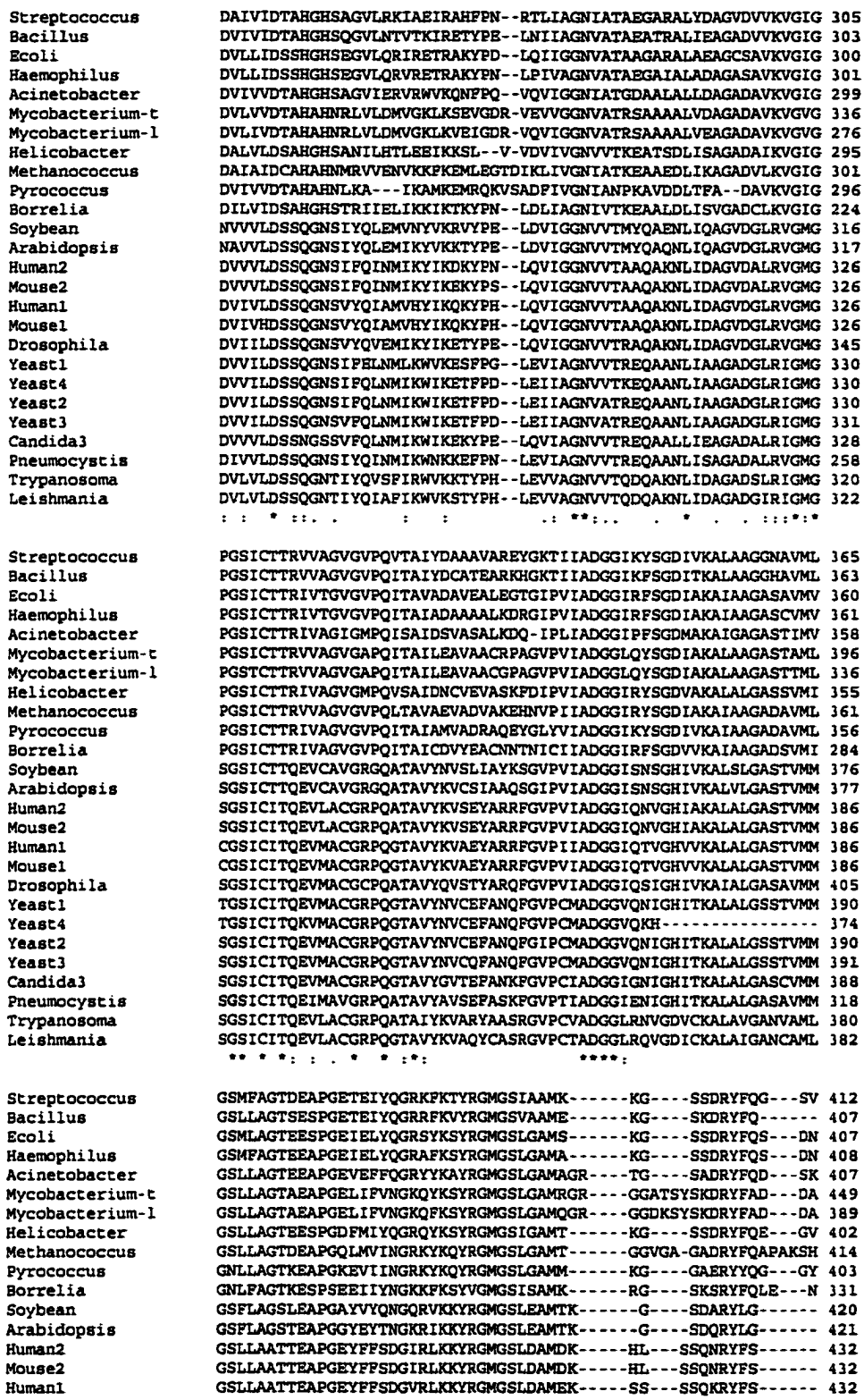


Figure 3-1 (continued). Alignment and comparison of complete sequences of IMPDHs from different organisms.

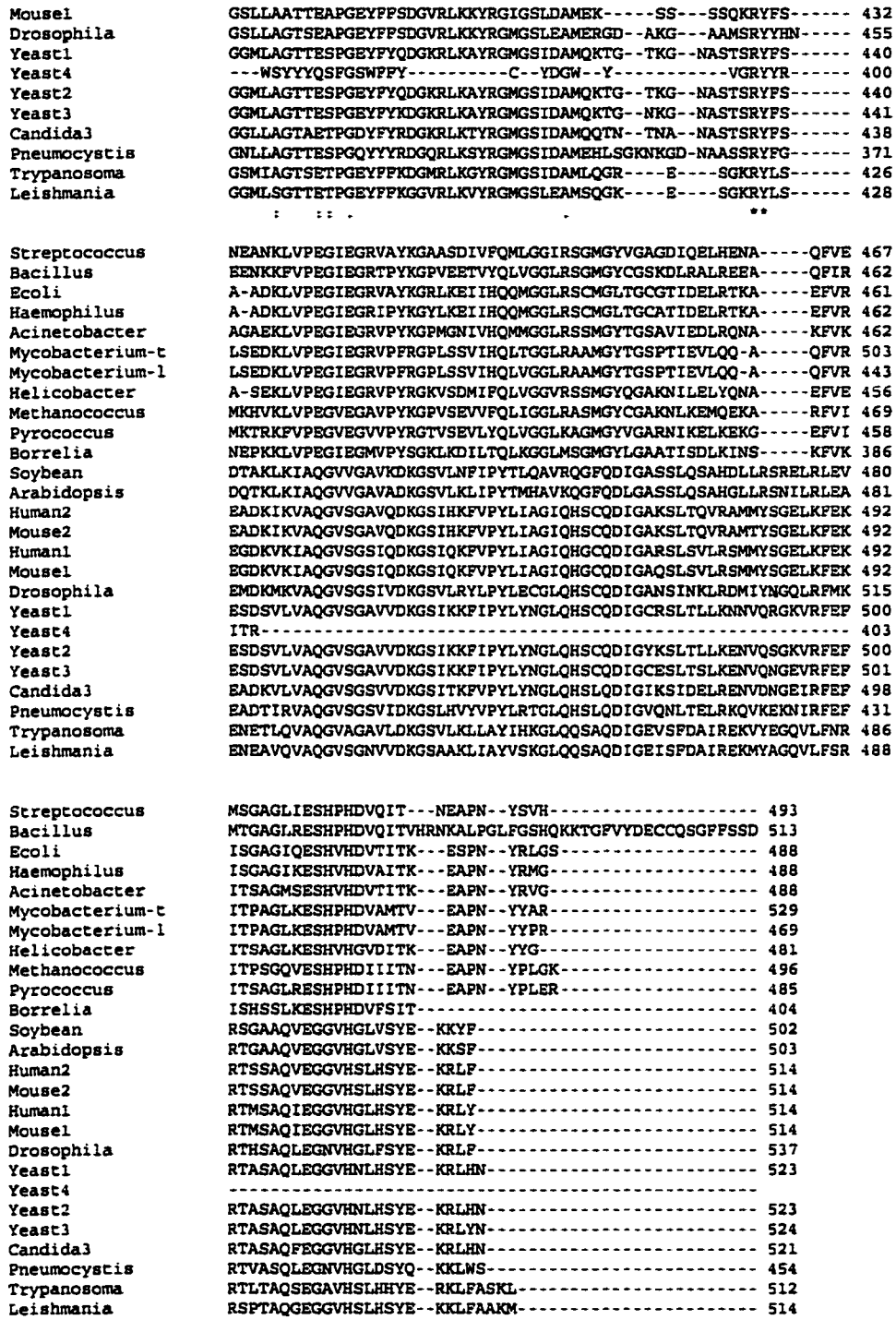


Figure 3-1 (continued). Alignment and comparison of complete sequences of IMPDHs from different organisms.

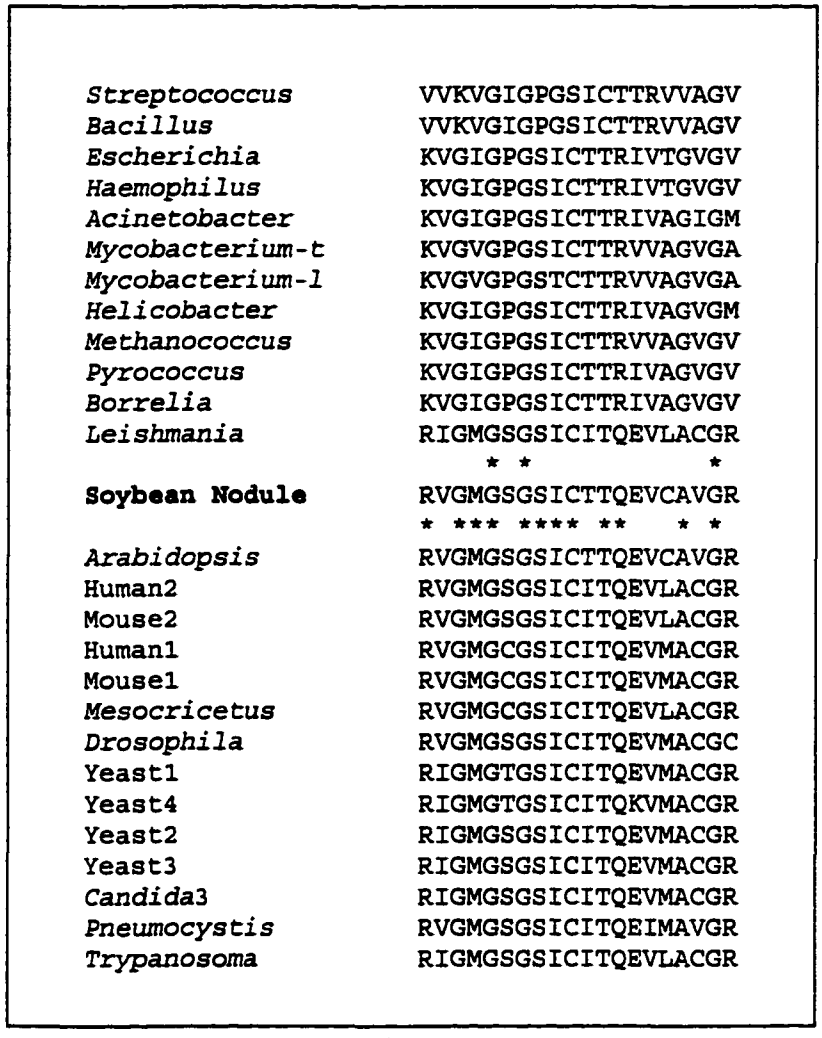


Figure 3-2. Comparison of the putative IMPDH catalytic domain from soybean nodules with that from prokaryotes (upper group) and eukaryotes (lower group). Asterisk (*) indicates identical amino acid residues. The domain is 100% identical between soybean nodule IMPDH and that of *Arabidopsis thaliana*. Sequences were aligned by using the ClustalW program (Higgins *et al.*, 1994) from the European Bioinformation Institute.

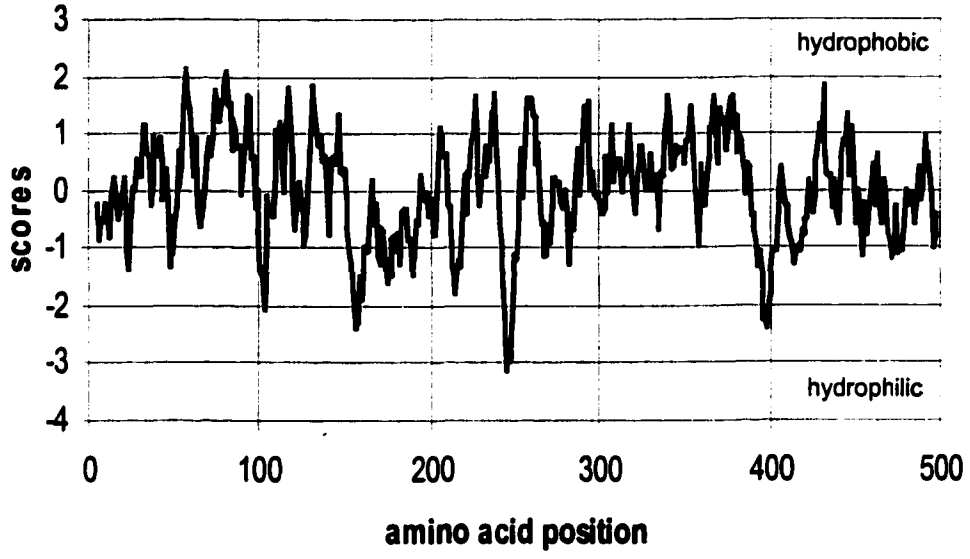


Figure 3-3. Hydropathy plot of the predicted amino acid sequence of soybean IMPDH. The plot was determined according to Kyte and Doolittle (1982) using a window size of nine amino acids. Increased hydrophobicity is indicated by positive values.

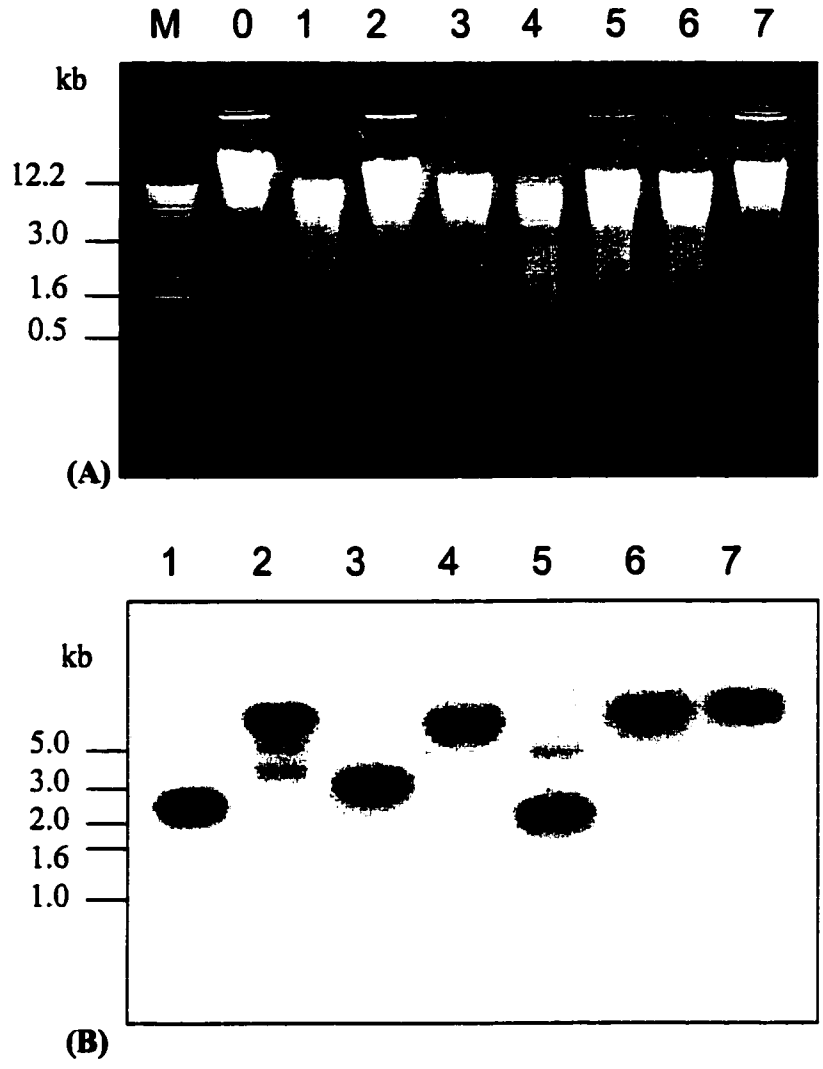


Figure 3-4. Southern blot of soybean genomic DNA. DNA was digested with *Eco*RI, *Kpn*I, *Pvu*II, *Sac*I, *Pst*I, *Bam*HI and *Xba*I (from lane 1 to 7), respectively. All these 7 enzymes do not have a restriction site in *IMPDH* cDNA except *Pst*I (at nt. 1724, lane 5) and *Xba*I (at nt. 22 from 5' end, lane 7). A full length *IMPDH* cDNA was used as a probe. Molecular size markers using the 1 kb DNA ladder (BRL) are shown on the left. Panel (A), ethidium bromide stained 0.8% agarose gel of digested and undigested soybean genomic DNA. Lane 0 represents undigested genomic DNA. Panel (B) is a scanned picture of radioactive signals detected on X-ray film after hybridization.

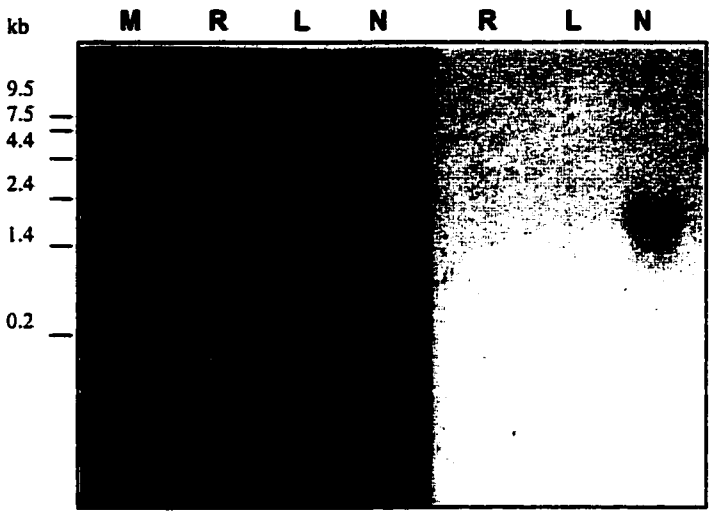


Figure 3-5. Northern blot analysis of total RNA (10 μg) from soybean roots (R), leaves (L) and nodules (N). RNA size markers are indicated on the left side of the figure. RNA was stained with ethidium bromide to show the quantity loaded in each lane (left side of the figure). Hybridization results are shown on the right side. A full length *IMPDH* cDNA was used as a probe.

DISCUSSION

Comparison of IMPDH sequences and the putative IMPDH signature motif

The length of IMPDHs from different organisms range from 403 to 537 amino acid residues. Comparison of the deduced amino acid sequences from different sources confirmed that soybean nodule IMPDH was most closely related to IMPDH from *Arabidopsis thaliana* (70.5% identity). The level of homology was the lowest for prokaryotic IMPDHs (28-42% identity) (Table 3-2).

A cysteine residue at codon 327 in human type II IMPDH (codon 328 in *E. coli*) is apparently essential for the catalytic activity of IMPDH from humans (Antonino *et al.*, 1994) and *E. coli* (Andrews and Guest, 1988). A consensus

sequence that includes this active-site cysteine has been proposed as a signature motif for IMPDH and guanosine monophosphate (GMP) reductase (Bairoch, 1995). All IMPDH active site sequences thus far characterized conform to the consensus sequence (Collart *et al.*, 1996b). The homology between the signature sequences of soybean nodule IMPDH and all other eukaryotic IMPDHs was 60%. The homology was 100% identical between soybean and *Arabidopsis* IMPDHs.

Prediction of hydrophobicity, subcellular localization, MW and pI

Conflicting results on the subcellular localization of IMPDH have been reported. In cowpea nodules, the IMPDH activity was detected in the cytoplasm (Shelp and Atkins, 1983), whereas in soybean nodules it was detected in the plastids (Schubert, 1981). IMPDH was reported to be localized in the cytosol of human cells (Holmes *et al.*, 1974). No studies on the localization of IMPDH have been reported for any organism using techniques of immunolocalization. In our current study, the prediction analysis demonstrated that the deduced amino acid sequence of IMPDH did not show any distinct hydrophobic or hydrophilic regions (Figure 3-3) and that there was no N-terminal signal sequence. These findings suggest that the protein of IMPDH is not targeted to plastids or mitochonria, but is possibly localized in the cytoplasm (Chrispeels and Tague, 1991) of soybean nodule cells.

The predicted molecular weight for soybean nodule IMPDH (53 kDa) was close to the previous estimation of 55 and 60 kDa reported by Yang (1997). It is also close to the 50-kDa subunit for IMPDH from *Vigna unguiculata* nodules

(Atkins *et al.*, 1985). The predicted pI of soybean nodule IMPDH was 5.54, while the previous estimation by IEF was 6.23 (Yang, 1997). It is still unclear how many subunits the soybean nodule IMPDH contains, or if these subunits are identical. Further biochemical studies are needed to answer these questions.

Southern blot analysis: copy number of IMPDH gene

In order to determine the copy number of *IMPDH* gene(s) in the soybean genome, the genomic DNA analysis was performed by Southern blotting. Among the seven restriction enzymes used for genomic DNA digestion, five do not have restriction sites inside the *IMPDH* cDNA. The other two, *Pst* I and *Xba* I, have an internal restriction site at nucleotide position 1724 and 22, respectively (Table 2-4). After Southern hybridization, all enzymes with no restriction sites inside *IMPDH* gave one strongly hybridizing band, suggesting that there might be only one copy of the *IMPDH* gene in the soybean genome. The results from *Pst* I and *Xba* I digestion also support the presence of a single *IMPDH* gene in the soybean genome. *Pst* I has a restriction site in nucleotide 1724, the middle of the *IMPDH* cDNA, and gave two bands. *Xba* I has a restriction site at the very beginning of the 5' untranslated region of *IMPDH* cDNA and therefore gave only a single strong hybridizing band. The presence of other weaker hybridization bands after *Kpn* I digestion suggests that either there may be introns within the soybean *IMPDH* gene which contain a *Kpn* I site or there are other unidentified *IMPDH* genes which have limited cross hybridization with the cloned *IMPDH* cDNA.

Washing with lower stringency and at lower temperature may possibly reveal more hybridization bands in the Southern blot.

Northern blot analysis

Northern blot analysis was used to examine the abundance and tissue distribution of *IMPDH* mRNA in soybean plants. A major RNA hybridizing band of approximately 2.0 kb was observed in 36-day-old nodules, while signals in roots and leaves were too weak to be detectable. These results suggested that IMPDH expression is specific to nodules and that IMPDH activity is induced or greatly enhanced by some factors related to nitrogen fixation. These findings provide further support for the proposal that the IMPDH catalyzed reaction is the principal pathway for purine oxidation and ureide biosynthesis in nodules (Schubert and Boland, 1990). Confirmation of the role of IMPDH in root nodule metabolism will require more detailed studies on the localization and regulation of expression of IMPDH in nodules.

REFERENCES

Anderegg, U., Schnuck, W.H., Asperger, O. and Kleber, H.P. (1992). IMPDH sequence of *Actinetobacter calcoaceticus*. EMBL / Genebank / DDBJ Data Banks.

Andrews, S.C. and Guest, J.R. (1988). Nucleotide sequence of the gene encoding the GMP reductase of *Escherichia coli* K12. Biochem. J. 255: 35-43.

Antelmann, H., Bernhardt, J., Schmid, R., Mach, H., Voelker, U. and Hecker, M. (1997). First steps from a two-dimensional protein index towards a response-regulation map for *Bacillus subtilis*. Electrophoresis 18: 1451-1463.

Antonino, L.C., Straub, K. and Wu, J.C. (1994). Probing the active site of human IMP dehydrogenase using halogenated purine riboside 5'-monophosphates and covalent modification reagents. Biochemistry 33: 1760-1765.

Ashbaugh, C.D. and Wessels, M.R. (1995). Cloning, sequence analysis and expression of the group A streptococcal *guaB* gene encoding inosine monophosphate dehydrogenase. Gene 165: 57-60.

Atkins, C.A., Shelp, B.J. and Storer, P.J. (1985). Purification and properties of inosine monophosphate oxidoreductase from nitogen-fixing nodules of cowpea (*Vigna unguiculata* L. Walp). Arch Biochem Biophys 236: 807-814.

Ausubel, F.M., Brent, R., Kingston, R.E., Mooew, D.D., Seidman, J.G. and Smith, J.S. (1991). Current Protocols in Molecular Biology. Wiley, New York.

Bairoch, A. (1995). Prosite Dictionary: Release 12.2. University of Geneva, Geneva, Switzerland.

Beck, J.T., Zhao, S. and Wang, C.C. (1993). Cloning, sequencing, and structural analysis of the DNA encoding inosine monophosphate dehydrogenase (E.C.1.1.1.205) from *Tritrichomonas foetus*. Exp Parasitol 78:101-112.

Bjellqvist, B., Hughes, G.J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J.C., Frutiger, S. and Hochstrasser, D.F. (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. Electrophoresis 14: 1023-1031.

Bjellqvist, B., Basse, B., Olsen, E. and Celis, J.E. (1994). Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. Electrophoresis 15: 529-539.

Boland, M.J. and Schubert, K.R. (1983). Biosynthesis of purines by a proplastid fraction from soybean nodules. Arch. Biochem. Biophys. 220: 179-187.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J.F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S.M. and Venter, J.C. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science 273: 1058-1073.

Bussey, H., Kaback, D.B., Zhong, W., Vo, D.T., Clark, M.W., Fortin, N., Hall, J., Ouellette, B.F., Keng, T., Barton, A.B. *et al.* (1995). The nucleotide sequence of chromosome I from *Saccharomydes cerevisiae*. Proc Natl Acad Sci USA 92: 3809-3813.

Chrispeels, M.J. and Tague, B.W. (1991). Protein sorting in the secretory system of plant cells. Int Rev Cytol 175: 1-45.

Collart, F.R. and Huberman, E. (1988). Cloning and sequence analysis of the human and Chinese hamster inosine-5'-monophosphate dehydrogenase cDNAs. J Biol Chem 263:15769-15772.

Collart, F.R., Osipiuk J., Trent, J., Olsen,G. and Huberman, E. (1996a). Cloning and characterization of the gene encoding IMP dehydrogenase from *Arabidopsis thaliana*. Gene 174: 217-220.

Collart, F.R., Osipiuk J., Trent, J., Olsen,G. and Huberman, E. (1996b). Cloning, characterization and sequence comparison of the gene encoding IMP dehydrogenase from *Pyrococcus furiosus*. Gene 174: 209-216.

Dayton, J.S. and Mitchell, B.S. (1993). IMPDH-I sequence. EMBL / Genebank / DDBJ Data Banks.

Devlin, K., Churcher, C., Barrell, B.G. and Rajandream, M.A. (1994). IMPDH sequence of *Saccharomyces cerevisiae* (baker's yeast). EMBL / Genebank / DDBJ Data Banks.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M. *et al.* (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496-512.

Fraser, C.M. *et al.* (1997). Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature 390: 580-586.

Glesne, D.A. and Hubrman, E. (1994). Cloning and sequence of the human type II IMP dehydrogenase gene. Biochem Biophys Res Commun 205: 537-544.

Hager, P.W., Collart, F.R., Huberman, E. and Mitchell, B.S. (1995). Recombinant human inosine monophosphate dehydrogenase type I and type II proteins. Purification and characterization fo inhibitor binding. Biochem Pharmacol 49: 1323-1329.

Haymes, K.M. (1996). Mini-prep method suitable for a plant breeding program. Plant Molecular Biology Report 14(3): 280-284.

Higgins,D., Thompson, J., Gibson, T., Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

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

Holms, E.W., Pehlke, M. and Kelley, W.N. (1974). Human IMP dehydrogenase, kinetics and regulatory properties. Biochem Biophys Acta 364: 209-217.

Huang, X. and Miller, W. (1991). A time-efficient, linear-space local similarity algorithm. Advances in Applied Mathematics 12: 337-357.

Huete-Perez, J.A., Wu, J.C., Whitby, F.G. and Wang, C.C. (1995). Identification of the IMP binding site in the IMP dehydrogenase from *Tritrichomonas foetus*. Biochemistry 34: 13889-13894.

Hupe, D.J., Azzolina, B.A. and Behrens, N.D. (1986). IMP dehydrogenase from the intracellular parasitic protozoan *Eimeria tenella* and its inhibition by mycophenolic acid. J Biol Chem 261:8363-8369.

Ikegami, I., Natsumeda, Y. and Weber, G. (1987): Purification of IMP dehydrogenase from rat hepatoma 392A. Life Sci 40: 2277-2282.

Johnston, M., Andrews, S., Brinkman, R., Cooper, J., Ding, H., Dover, J., Du, Z., Favello, A., Fulton, L., Gattung, S. *et al.* (1994). Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VIII. Science 265: 2077-2082.

Johnston, M, Andrews, S., Brinkman, R., Cooper, J., Ding, H., Dover, J., Du, Z., Favello, A., Fulton, L., Gattung, S. *et al.* (1995). IMPDH sequence of *Saccharomyces cerevisiae* (baker's yeast). EMBL/Genebank/DDBJ Data Banks.

Kanzaki, N. and Miyagawa, K.I. (1990). Nucleotide sequence of the *Bacillus subtilis* IMP dehydrogenase gene. Nucleic Acids Res 18: 6710-6710.

Kohler, G.A., White, T.C. and Agabian, N. (1997). Overexpression of a cloned IMP dehydrogenase gene of *Candida albicans* confers resistance to the specific inhibitor mycophenolic acid. J Bacteriol 179: 2331-2338.

Krishnaiah, K.V. (1975). Inosinic acid 5'-monophosphate dehydrogenase of *Escherichia coli:* purification by affinity chromatogrphy and some properties. Arch Biochem Biophys 170:567-575.

Kyte J and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol 157(1): 105-132.

Lightfoot, T. and Snyder, F.F. (1994). Gene amplification and dual point mutations of mouse IMP dehydrogenase associated with cellular resistance to mycophenolic acid. Biochim Biophys Acta 1217: 156-162.

Link, A.J., Robison, K. and Church, G.H. (1997). Comparing the predicted and observed properties of proteins encoded in the genome of *Escherichia coli* K-12. Electrophoresis 18: 1259-1313.

Margolis, N., Hogan, D., Tilly, K. and Rosa, P. (1994). Plasmid location of *Borrelia* purine biosynthesis gene homologs. J Bacteriol 176: 6427-6432.

Molloy, M.P., Herbert, B.R., Walsh, B.J. and Gooley, A.A. (1998). Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. Electrophoresis 19: 837-844.

Nakai, K. (1991). Predicting various targeting signals in amino acid sequences. Bull Inst Chem Res, Kyoto Univ (Japan) 69:269-291.

Nakai, K and Kanehisa, M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14: 897-911.

Nash, D. and Hu, S. (1992). *Drosophila* inosine monophosphate dehydrogenase is encoded at the raspberry locus. Can Fed Biol Soc 35: 72-72.

Natsumeda, Y., Ohno, S., Kawasaki, H., Konno, Y., Weber, G. and Suzuki, K. (1990). Two distinct cDNAs for human IMP dehydrogenase. J Biol Chem 265: 5292-5295.

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

Nolling, J., Pihl, T.D. and Reeve, J.N. (1995). Cloning, sequencing, and growth phase-dependent trnscription of the coenzyme F420-dependent N5, N10-methylenetetrahydromethanopterin reductase-encoding genes from *Methanobacterium thermoautotrophicum* delta H and *Methanopurus kandleri.* J Bacteriol 177: 7234-7244.

O'Gara, M.J., Lee, C.H., Wenberg, G.A., Nott, J.M. and Queener, S.F. (1997). IMP dehydrogenase from *Pneumocystis carinii* as a potential drug target. Antimicrob Agents Chemother 41: 40-48.

Ogasawara, N., Nakai, S. and yoshikawa, H. (1994). Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. DNA Res 1: 1-14.

Oliver, K., Harris, D., Barrel, B.G., Rajandream, M.A. and Walsh, S.V. (1996). IMPDH sequence of *Mycobacterium tuberculosis.* EMBL / Genebank / DDBJ Data Banks.

Peterson, B.L., Moeller, M.G., Stummann, B.M. and Henningsen, K.W. (1996). Clustering of genes with function in the biosynthesis of bacteriochlorophyll and heme in the green sulfur bacterium *Chlorobium vibrioforme.* Hereditas 125: 93-96.

Schubert, K.R. (1981). Enzymes of purine biosynthesis and catabolism in *Glycine max.* Plant Physiol. 68:1115-1122.

Schubert, K.R. and Boland, M.J. (1990). Ureides. In: the biochemistry of plants (Stumpf, P.K. & Conn, E.E. Ed.) 16: 197-282. Academic Press, New York.

Shelp, B.J. and Atkins, C.A. (1983). Role of inosine monophosphate oxidoreductase in the formation of ureides in nitrogen-fixing nodules of cowpea (*Vigna unguiculata L.Walp*). Plant Physiol 72: 1029-1034.

Sifri, C.D., Wilson, K., Smolik, S., Forte, M. and Ullman, B. (1994). Cloning and sequence analysis of a *Drosophila melanogaster* cDNA encoding IMP dehydrogenase. Biochim Biophys Acta 1217: 103-106.

Slee, R. and Bownes, M. (1995). The raspberry locus encodes *Drosophila* inosine monophosphate dehydrogenase. Mol Gen Genet 248: 755-766.

Smith, D.R. and Robson, K. (1994). IMPDH sequence of *Mycobacterium leprae.* EMBL / Genebank / DDBJ Data Banks.

Tiedeman, A.A. and Smith, J.M. (1985). Nucleotide sequence of the *guaB* locus encoding IMP dehydrogenase of *Escherichia coli* K12. Nucleic Acids Res 13: 1303-1316.

Tiedeman, A.A. and Smith, J.M. (1991). Isolation and sequence of a cDNA encoding mouse IMP dehydrogenase. Gene 97: 289-293.

Thomas, M.S. and Drabble, W.T. (1985). Nucleotide sequence and organization of the gua promoter region of *Escherichia coli*. Gene 36: 45-53.

Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzegerald, L.M., Lee, N., Adams, M.D., Venter, J.C. *et al.* (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature 388: 539-547.

Verham, R., Meek, T.D., Hedstrom, L. and Wang, C.C. (1987). Purification, characterization and kinetic analysis of inosine 5'-monophosphate dehydrogenase of *Tritrichomonas foetus*. Mol Biochem Parasitol 24: 1-12.

Wadsworth, G.J., Marmaras, S.M., Matthews, B.F. (1993). Isolation and characterization of a soybean cDNA clone encoding the plastid form of aspartate aminotransferase. Plant Mol Biol 21: 993-1009.

Whiteby, F.G., Luecke, H., Kuhn, P, Somoza, J.R., Huete-Perez, J.A., Phillips, J.D., Hill, C.P., Fletterick, R.J. and Wang, C.C. (1997). Crystal structure of *Tritrichomonas foetus* inosine-5'-monophosphate dehydrogenase and the enzyme-product complex. Biochemistry 36: 10666-10674.

Wilkins M.R., Gasteiger E., Bairoch A., Sanchez J.-C., Williams K.L., Appel R.D., Hochstrasser D.F. (1998). Protein identification and analysis tools in the ExPASy server. In: 2-D proteome analysis protocols (Link, A.J. Ed.). Humana Press, New Jersey.

Wilson, K.E., Collart, F.R., Huberman, E., Stringer, J.R. and Ullman, B. (1991). Amplification and molecular cloning of the IMP dehydrogenase gene of *Leishmania donovani*. J Biol Chem 266: 1665-1671.

Wilson, K., Berens, R.L., Sifri, C.D. and Ullman, B. (1994). Amplificaton of the inosinate dehydrogenase gene in *Trypanosoma brucei gambiense* due to an increase in chromosome copy number. J Biol Chem 269: 28979-28987.

Yang, D. (1997). Purification and characterization of inosine-5'-monophosphate dehydrogenase (IMPDH) from nitrogen-fixing nodules of soybean (*Glycine max*). Master's thesis, the University of Oklahoma, Norman, OK, USA.

Zimmermann, A.G., Spychala, J. and Mitchell, B.S. (1995). Characterization of the human inosine-5'-monophosphate dehydrogenase type II gene. J Biol Chem 270: 6808-6814.

Chapter 4. Expression of *IMPDH* cDNA in *Escherichia coli* – Functional Complementation Studies

INTRODUCTION

Functional complementation refers to the restoration of the normal phenotype of a mutant by the introduction of a wild type allele (Murray & Smith, 1996). It has been used to clone genes and to confirm the functional expression of gene clones. Up to now, three *E. coli* mutants in which the *guaB* gene coding for IMPDH has been deleted (*ΔguaB*) have been reported. These mutants include *E. coli* strain H712 (Nijkamp and Haan, 1967), strain KLC381 (Vales *et al.*, 1979) and strain SO1101 (Ashbaugh and Wessels, 1995). Among these strains, SO1101 and H712 have been used in studies on functional complementation. Transformation of competent H712 bacterial cells with *Tritrichomonas foetus IMPDH* cDNA has conferred the ability of the transformed H712 cells to grow on minimal medium without guanine (Beck *et al.*, 1994; Huete-Pérez *et al.*, 1995). Results of functional complementation studies reported by Ashbaugh and Wessels (1995) showed that the expression of an incomplete *Streptococcus* IMPDH open reading frame in *E. coli* SO1101 restored IMPDH activity.

In Chapter 2 and Chapter 3 we described the isolation of a putative IMPDH cDNA clone from a soybean nodule cDNA library. Although this putative IMPDH clone exhibits a high degree of homology with *IMPDH* genes from other sources, functional complementation studies were used to substantiate the functional identity of the putative soybean IMPDH clone. In this chapter, we will show that the expression of soybean nodule *IMPDH* cDNA in *E. coli* mutant strain KLC381 is able to restore the IMPDH activity of this mutant, conferring the ability to grow on minimal medium.

MATERIALS AND METHODS

Subcloning of p46-10 into pBluescript SK(+) with correct open reading frame

The clone p46-10, which apparently contains the complete IMPDH coding sequence, was out of frame for a correct protein translation start site driven by the T3 RNA polymerase promoter in pBluescript SK (Figure 4-1). It was necessary to modify the clone by adding an extra nucleotide before the start codon. This was accomplished using PCR (polymerase chain reaction). An oligonucleotide, 5'-CATAGAATTCCATGGACTTCACTACGCCGCCG-3', containing an extra nucleotide "C" (in bold) and a built-in *EcoR* I restriction site (underlined), was designed as the 5' primer. The 5' primer was used together with the T7 universal primer (as the 3' primer) for PCR amplification of the *IMPDH* cDNA. A 50 µl PCR reaction was set up as follows: 31.25 µl of H_2O, 5.0 µl of 10x *Pfu* buffer, 1.0 µl of dNTPs (10 mM each), 0.5 µl of DNA template (0.09 µg/µl), 1.25 µl of 5'-primer (0.1 µg/µl = 11 pmol/µl), 10 µl of T7 primer (10 ng/µl = 1.52 pmol/µl), 1.0 µl of cloned *Pfu* DNA polymerase (2.5 U/µl) (Stratagene). PCR was run for 30 cycles with 45 second denaturation at 95°C, 45 second annealing at 60°C and 2.5 min extension at 72°C.

The pBluescript SK (+) vector DNA and the product of PCR amplification (Figure 4-2) were both digested with *Eco*R I. The corresponding bands were excised from the agarose gel and cleaned using a QIAEX II kit (Qiagen). The vector and PCR amplified *IMPDH* cDNA were ligated using T4 DNA ligase. The ligation reaction was carried out at 15°C for 16-20 hours and the ligation solution

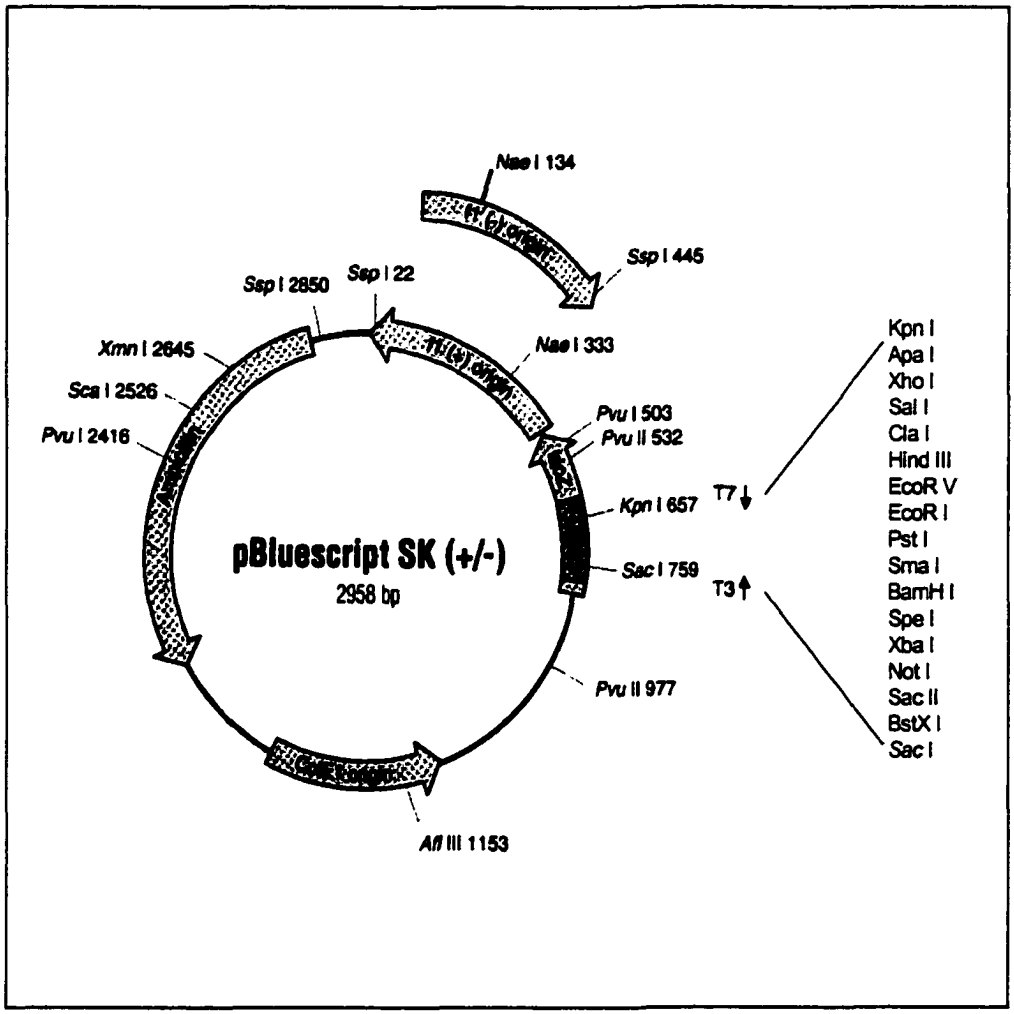


Figure 4-1. Map of vector pBluescript SK (+/-) (From Stratagene). *EcoR* I site was used for the subcloning of *IMPDH* cDNA.

was used to transform *E. coli* XL1 Blue mrf' competent cells. Recombinants were selected on LB-amp$^+$ plates with X-gal and IPTG. Single, white colonies were picked up from the plate. To confirm the presence of the correct insert, plasmid DNA was isolated, digested with *Eco*R I and separated on a 1.2% agarose gel. Plasmid DNA isolated from target clones was sequenced. Clones containing *IMPDH* cDNA with the correct open reading frame were designated as "pSBimpdh".

Functional complementation of E. coli *mutant KLC381 with the soybean IMPDH cDNA clone*

E. coli strain KLC381 (Vales *et al.*, 1979) with the *IMPDH* gene (*guaB*) deleted was obtained from the *E. coli* Genetic Stock Center at Yale University (Department of Biology, 355 OML, Yale University, P.O., New Haven, CT 06520-8104). The cells were made competent by the one-step method of Chung *et al.* (1989) and transformed with pSBimpdh plasmid DNA at 20 ng DNA / 100 µl cell culture. After incubation on ice for 45 min, the transformed cells were grown in 890 µl LB broth and 10 µl 40% glucose for 1.5 h at 37°C. At this time, the cells were pelleted by centrifugation at 13,000x g for 5 min, resuspended in 1 ml M9 minimal medium supplemented with 0.1% casamino acids (M9CA). This wash process was repeated three times to remove residual LB broth. After the final wash, the cell pellet was resuspended in 500 µl of the M9CA medium further supplemented with tryptophan (0.02 µg/ml), glucose (0.4%, w/v), thiamine (0.1

μg/ml), histidine (0.1 μg/ml), tyrosine (0.1 μg/ml), ampicillin (50 μg/ml) and IPTG (1 mM). The resuspended cells were plated on the supplemented M9CA-agar plates and incubated at 37°C for about 15 days to allow the appearance of individual cell colonies. pBluescript SK(+) without an insert was used as a negative control.

Individual colonies growing on M9CA-supplemented minimal media plates were picked and re-plated on LB-ampicillin (50 μg/ml) plates. Plasmid DNA from these colonies was isolated and digested with *Eco*R I to check for the presence and size of inserts. To confirm that the insert was the *IMPDH* cDNA clone, isolated DNA was used for a Southern blot analysis with *IMPDH* cDNA as the probe. Due to the lack of plasmid, the original KLC381 cells (untransformed) are not antibiotic resistant and do not grow on LB-amp$^+$ media.

RESULTS

Subcloning of p46-10 into pBluescript SK(+) with correct open reading frame

Plasmid DNAs from 6 recombinant colonies (from LB-amp$^+$ plates) containing a band corresponding in size to the *IMPDH* cDNA after *Eco*R I digestion were sent for sequencing. Five of these colonies contained DNA insertions in the correct direction for translation. Alignment of the 5' end of these 5 pSBimpdh clones and the original p46-10 (Figure 4-3) sequences revealed that pSBimpdh has an extra nucleotide ("C") just before the start codon ATG. The *IMPDH* cDNA was located downstream of β-galactosidase and was in the correct open reading frame

for translation. The expression of the fusion protein in these cells was under the control of the T3 promoter (Figure 4-1).

Functional complementation of E. coli *KLC381 with IMPDH clones*:

E. coli KLC381 is a mutant strain with the *IMPDH* gene deleted (*ΔguaB*). This strain can not grow on the M9 minimal medium without supplementation with purines. The *IMPDH* cDNA clones with the correct open reading frame in pBluescript SK(+) plasmids (pSBimpdh) were used to transform the competent *E. coli* strain KLC381. After plating 100 μl and 400 μl of the bacterial cells transformed with pSBimpdh, 68 individual cell colonies capable of growth on minimal agar plates without guanine supplementation were isolated. In contrast, there was no growth of cells transformed with pBluescript SK(+) without an insert under the same experimental conditions. Thus, the growth of cells transfomed with pSBimpdh on minimal media supported the conclusion that the putative *IMPDH* gene was able to complement the *gua B* deletion.

Colonies grown on minimal medium were very small, usually less than 0.5 mm in diameter. The size of the re-plated colonies growing on LB-amp$^+$ rich media was much larger (Figure 4-4). From the 68 colonies on minimal medium plates, 19 were picked and cultured on LB-ampicillin (50 μg/μl) plates for plasmid DNA isolation. After plasmid DNA isolation and *Eco*R I digestion, DNAs from all 19 isolates gave two bands. One band was ~3.0 kb, which corresponded to the size of the vector (2900 bp of pBluescript SK). The other band was ~2.0 kb, corresponding to the size of the *IMPDH* cDNA insert (Figure

4-5A). Southern hybridization with the *IMPDH* cDNA clone as a probe indicated that all 19 isolates contained the *IMPDH* cDNA insert. (Figure 4-5B).

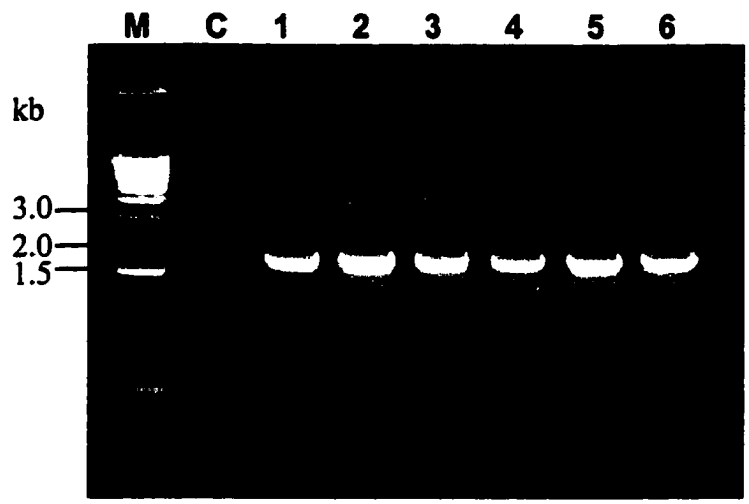


Figure 4-2. 1.2% agarose gel of PCR amplified *IMPDH* cDNA. PCR product contained a built-in *Eco*R I site and an extra nucleotide at the 5'-end. Lane C is the control reaction for PCR without the addition of template. Lane 1-6 represents 6 identical PCR reactions. The PCR product was excised from the gel, cleaned, digested with *Eco*R I and ligated with pBluescript SK (+) vector.

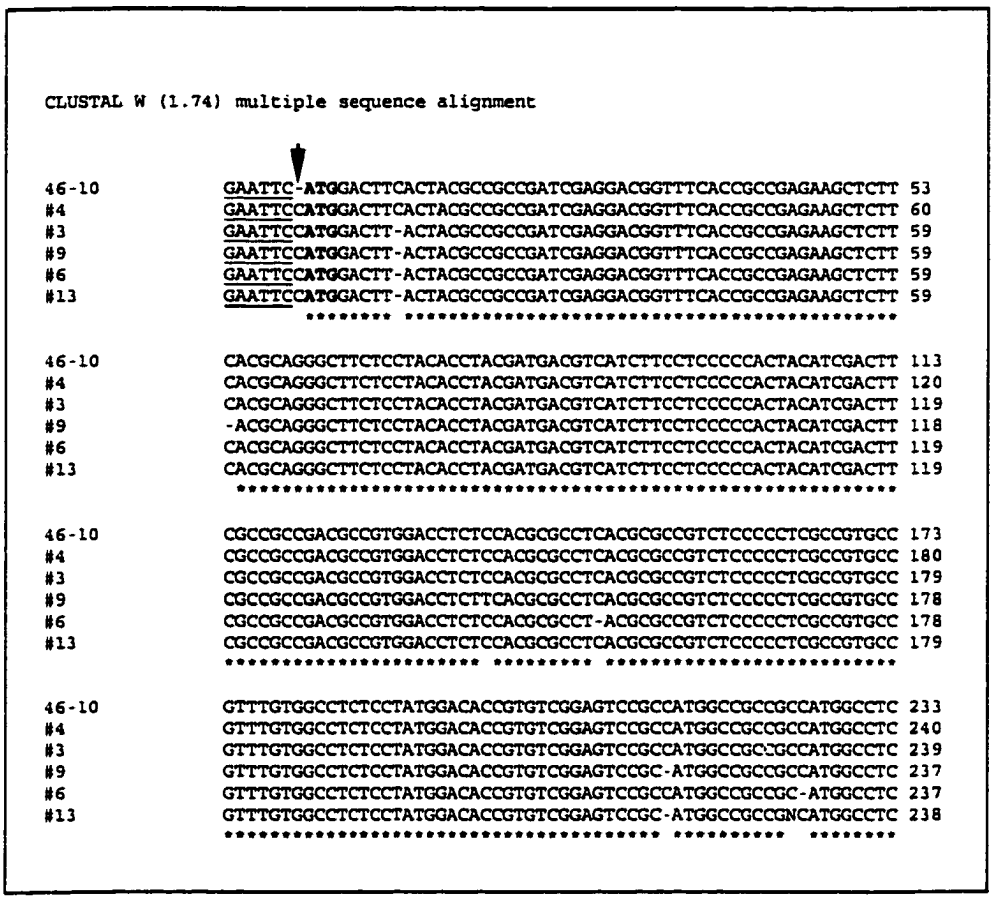


Figure 4-3. Alignment of the 5'-end (T3 end) of clone p46-10 and pSBimpdh. p46-10 was the original IMPDH cDNA clone. pSBimpdh was modified by the addition of one nucleotide ("C", arrow) just before the start codon, shifting the IMPDH cDNA into the correct open reading frame for translation driven by the T3 promoter. #3, #4, #6, #9, #13 are different clones of pSBimpdh. The clone #4 was used for functional complementation studies.

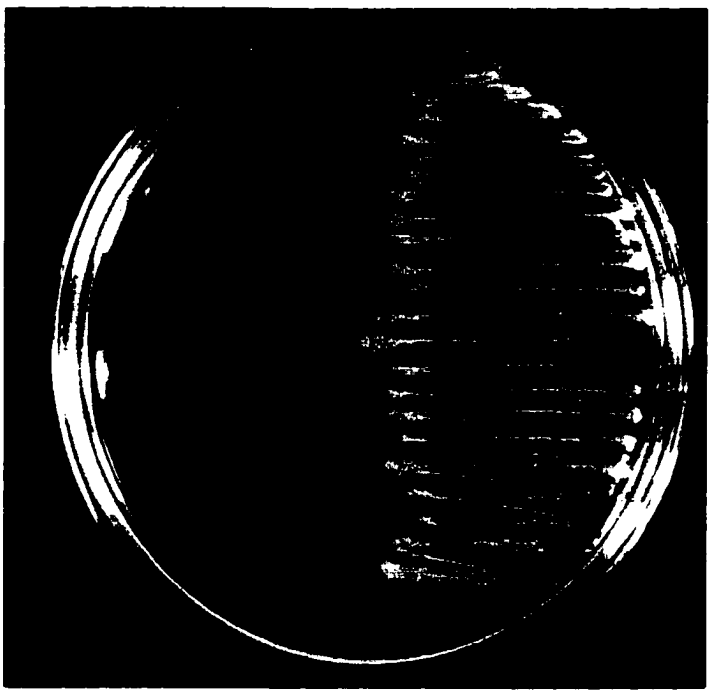


Figure 4-4. Growth of the original (left) and transformed (right) *E. coli* mutant KLC381 on LB-amp$^+$ medium. The expression of soybean nodule IMPDH cDNA in *E. coli* could restore the IMPDH activity in mutant strain KLC381, conferring it the ability of growing on minimal medium, or rich and/or minimal medium containing ampicillin. The original KLC381 strain can not grow on any of these media (left).

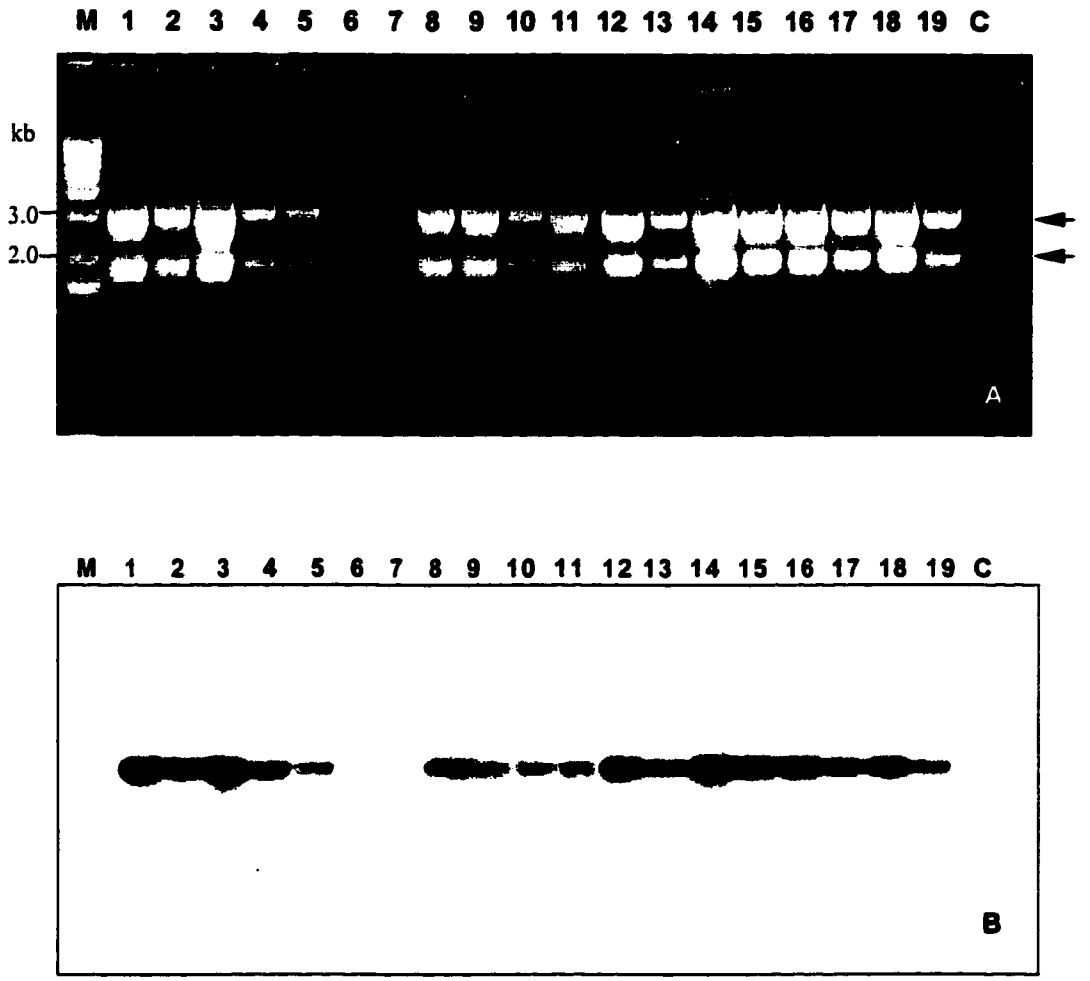


Figure 4-5. *Eco*R I digestion (Panel A) and Southern blot hybridization (Panel B) of plasmid DNA isolated from KLC381 colonies growing on minimal-amp$^+$ medium. Lanes 1 to 19 contain *Eco*R I digested DNA isolated from 19 transformant isolates from minimal medium plates. Lane "C" represents the original KLC381 strain without transformation, cultured on rich medium without antibiotics. Arrows in Panel A indicate the bands for the pBluescript SK vector (upper) and the *IMPDH* cDNA insert (lower). A much smaller amount of DNA was loaded in lane 6 and 7 as shown in panel A. This resulted in extremely weak signals in the Southern hybridization in Panel B. *IMPDH* cDNA was used as the probe for the Southern blot.

DISCUSSION

Functional complementation is a very useful technique for molecular genetics. It involves suppression of a mutant phenotype by introducing a wild-type copy of the corresponding gene (Murray & Smith, 1996). Generally this technique has two basic applications: 1). To isolate cDNAs complementing a given mutation; 2). To obtain functional information concerning cloned genes or their similarity to genes identified in other species. This is usually achieved by expressing the cloned cDNA / gene in heterologous organisms and assaying their ability to confer specific phenotypes. In functional complementation studies, one assumes that the function rather than the nucleotide sequence or epitope is conserved throughout evolution of distant organisms (Murray & Smith, 1996).

In our experiments, complementation was used to confirm the function of the putative soybean *IMPDH* cDNA clone. The soybean nodule *IMPDH* cDNA clone in the pBluescript SK (+) plasmid was used to transform the *E. coli guaB* deletion mutant KLC381. Results from the complementation experiments suggest that the protein produced by the cloned soybean *IMPDH* cDNA possesses IMPDH activity. Results from plasmid DNA isolation, restriction enzyme digestion and Southern blot hybridization showed that the transformants contained the pBluescript SK plasmid with *IMPDH* cDNA insert (Figure 4-5A,B). These results demonstrated the restoration of IMPDH activity in *E. coli* mutant KLC381 conferred by the expression of the soybean *IMPDH* cDNA in plasmids.

Soybean nodule *IMPDH* cDNA was able to complement the *E.coli* mutant KLC381 at a much lower frequency than the frequencies observed for the

transformation of the *E. coli* mutant strain H712 using IMPDH cDNA from the protozoan *Tritrichomonas foetus*. In the later case, more than 3000 colonies grew out on minimal medium. Also, the time needed for the cells to grow on minimal medium was much longer (3-7 days versus 15 days). One possible reason for the lower growth could be that the expression of soybean nodule *IMPDH* cDNA in *E. coli* was much lower. Our experiment on expression of pSBimpdh in *E. coli* strain XL1-Blue mrf' showed that the fusion protein was not detectable by SDS-PAGE electrophoresis even after 6.5 hours of IPTG induction (unpublished data). Another explanation would be in differences between the growth rates of the two *E. coli* mutants H712 and KLC318.

REFERENCES

Ashbaugh, C.D. and Wessels, M.R. (1995). Cloning, sequence analysis and expression of the group A streptococcal *guaB* gene encoding inosine monophosphate dehydrogenase. Gene 165: 57-60.

Beck, J.T., Zhao, S. and Wang, C.C. (1993). Cloning, sequencing, and structural analysis of the DNA encoding inosine monophosphate dehydrogenase (E.C.1.1.1.205) from *Tritrichomonas foetus.* Exp Parasitol 78: 101-112.

Chung, C.T., Niemela, S.L. and Miller, R. H. (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci USA 86: 2171-2175.

Hannig, G. and Makrides, S.C. (1998). Strategies for optimizing heterologous protein expression in *Escherichia coli.* Trends in Biotech 16: 54-60.

Huete-Perez, J.A., Wu, J.C., Whitby, F.G. and Wang, C.C. (1995). Identification of the IMP binding site in the IMP dehydrogenase from *Tritrichomonas foetus.* Biochemistry 34: 13889-13894.

Margolis, N., Hogan, D., Tilly, K. and Rosa, P.A. (1994). Plasmid localization of *Borrelia* purine biosynthesis gene homologs. J Bacteriology 176: 6427-6432.

Murray, J.A.H. and Smith, A.G. (1996). Functional complementation in yeast and *E. coli.* In: Plant gene isolation: principles and practice (G.D. Foster and D. Twell Ed.). pp177-211. Jonh Wiley & Sons Ltd., Chichester, UK.

Nijkamp, H.J.J. and De Haan, P.G. (1967). Genetic and biochemical studies of the guanosine 5'-monophosphate pathway in *Escherichia coli.* Biochim Biophys Acta 145: 31-40.

Vales, L.D., Chase, J.W. and Murphy, J.B. (1979). Orientation of the guanine operon of *Escherichia coli* K-12 by utilizing strains containing *guaB-xse* and *guaB-upp* deletions. J Bacteriology 139: 320- 322.

Chapter 5. Expression and Purification of IMPDH Fusion Protein from *Escherichia coli*

INTRODUCTION

Both prokaryotic and eukaryotic expression systems have been developed for protein expression. Commonly used expression systems include *E. coli*, yeast, baculovirus and mammalian cells. *E. coli* expression systems are probably the most commonly used techniques for the following reasons: 1) most investigators are familiar with techniques used to culture *E. coli*; 2) the techniques necessary to express an usable amount of protein are relatively simple; 3) the time necessary to generate an overexpressing strain is very short; 4) a familiarity with standard recombinant DNA techniques is all that is necessary to begin pilot expression experiments; 5) *E. coli* is inexpensive to culture; and 6) the vast body of knowledge about *E. coli* has made it possible to tinker intelligently with its genetics and physiology so that strains producing 30% of their total protein as the expressed gene product can often be obtained (Ausubel *et al.*, 1991). On the other hand, the *E. coli* expression system also has some disadvantages. First, eukaryotic proteins expressed in *E. coli* may be not properly folded or modified. Second, proteins expressed in large amounts in *E. coli* often precipitate as insoluble aggregates called "inclusion bodies". Proteins in inclusion bodies are normally not active. In some cases, however, active proteins can be recovered by solubilization in denaturing agents followed by careful renaturation. Protein secretion may be another way to circumvent problems associated with the formation of inclusion bodies. Although the secretion of large amounts of expressed protein from *E. coli* may be difficult, the secretion of small amounts of

protein into the periplasmic space is possible. These proteins are then recovered by osmotic shock.

Thioredoxin is a protein found in yeast, plants and mammals, as well as in bacteria. It was originally isolated from *E. coli* as a hydrogen donor for ribonuclease reductase (Holmgren, 1985). Thioredoxin is also involved in a variety of cellular functions including the reduction of protein disulfides, sulfate metabolism, as a cofactor for phage T7 DNA polymerase and in the assembly of T7 and filamentous phage (Novy *et al.*, 1995). When overexpressed in *E. coli*, thioredoxin accumulates to approximately 40% of the total cellular protein while still remaining soluble. For this reason, thioredoxin has been fused with target proteins to increase translation efficiency and solubility of eukaryotic proteins expressed in *E. coli*.

The pET-32 vector (Figure 5-1) from Novagen was designed to provide high levels of expression while retaining protein solubility. The vector codes for the 109 aa Trx·Tag™ thioredoxin upstream of the multiple cloning site. Thus, the target protein should be expressed as a (soluble) C-terminal fusion with thioredoxin. The pET-32 vector also contains sequences adjacent to the multiple cloning site that encode a number of other peptide tags including the His·Tag and S·Tag. These fusion tags facilitate detection and purification of the target protein. The thrombin and enterokinase cleavage sites make it possible to cut the fusion peptides off the purified fusion proteins. The pET-32 vector also contains translation stop codons in all three reading frames following the cloning site and tag regions as well as a downstream T7 transcription terminator.

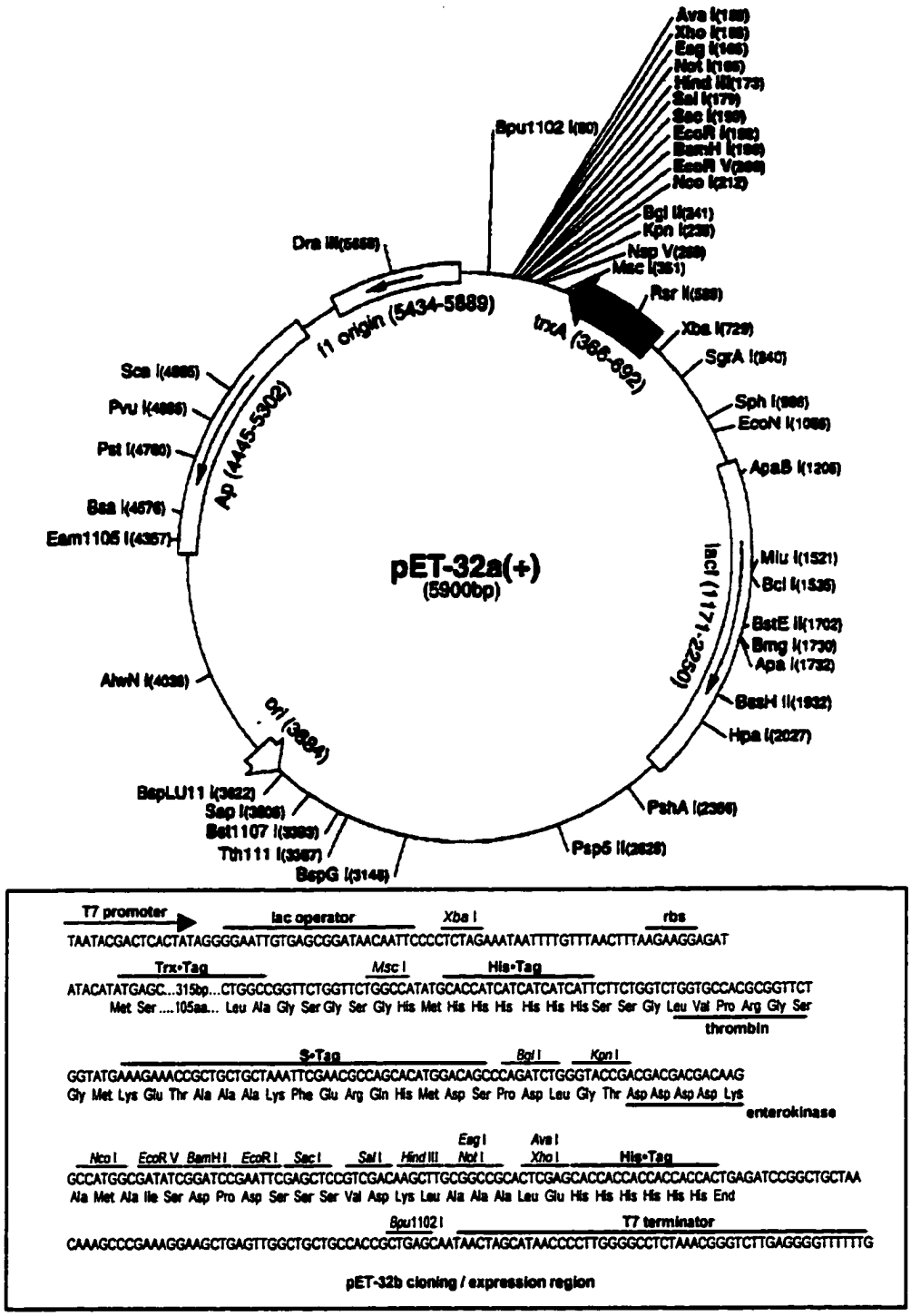


Figure 5-1. Map of the expression vector pET-32b (from Novagen).

Because of the extreme lability of plant IMPDH, the previous work on IMPDH protein purification from soybean nodules did not provide sufficient quantities of pure protein to be used for antibody preparation (Yang, 1997). The objective of the current study was to obtain sufficient quantities of IMPDH expressed in *E. coli* and to use this protein to produce polyclonal antibodies. Studies on the expression of soybean nodule IMPDH fusion protein in *E. coli* using the pET-32b vector will be described in this chapter. The fusion protein purification and antibody preparation will also be reported.

MATERIALS AND METHODS

Construction of plasmids

As indicated in Figure 5-2, the expression and purification of soybean nodule IMPDH in *E. coli* started with the plasmid construction. Soybean nodule *IMPDH* cDNA was amplified from the p46-10 plasmid DNA by PCR as described in Chapter 4. The 5' primer contained an *Eco*R I site and an extra nucleotide "C" just before the translation start codon ATG. The 3' primer was the universal T7 primer. The PCR product was digested with *Eco*R I and separated on a 1.2% agarose gel. The insert was excised from the agarose gel and cleaned using a QIAEX II kit (Qiagen).

The vector pET-32b was transformed into the *E. coli* strain NovaBlue and grown in LB medium containing 12.5 μg/ml tetracycline and 50 μg/ml ampicillin. Plasmid DNA was isolated using a Wizard Miniprep kit (Promega), digested with

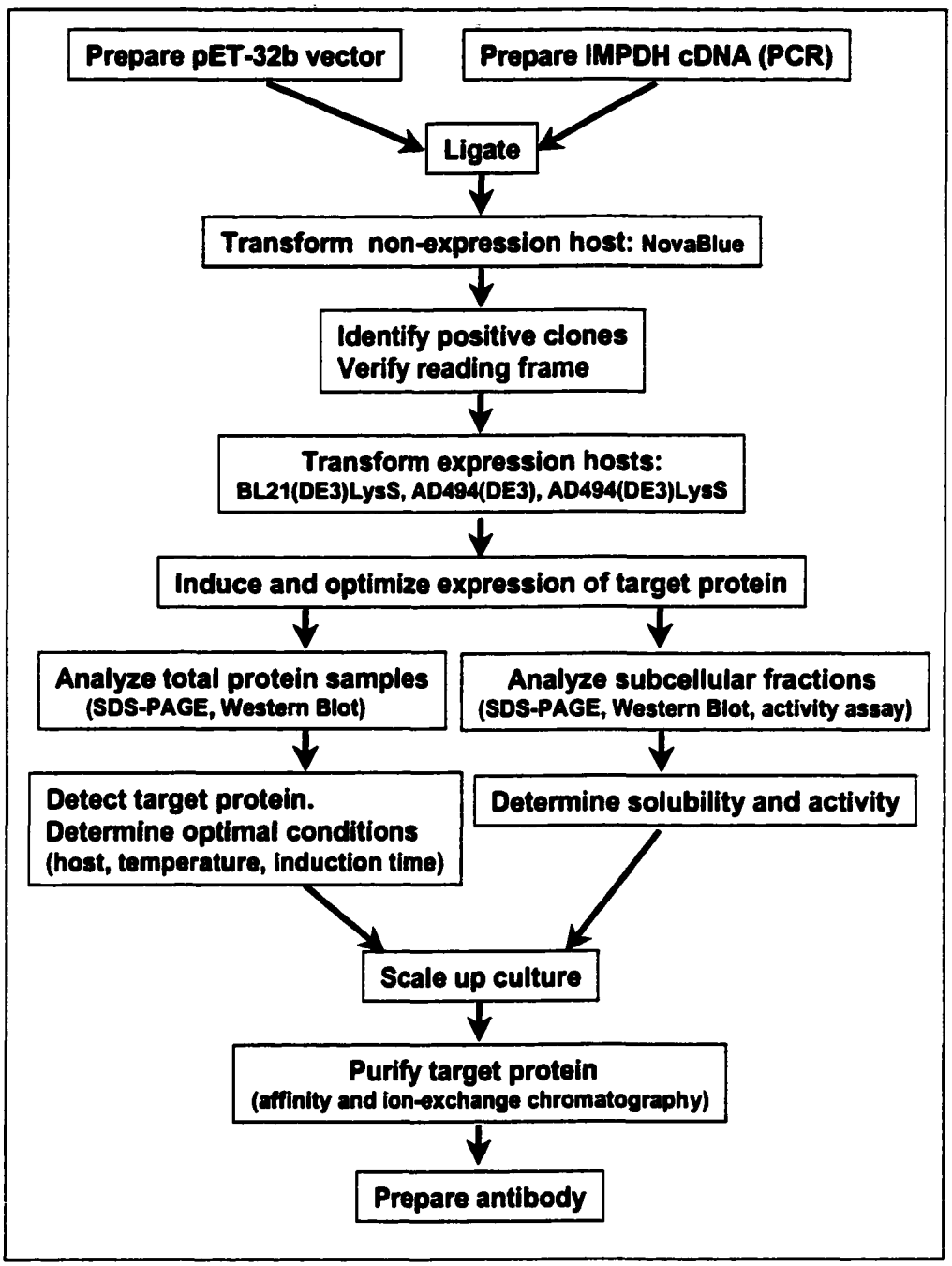


Figure 5-2. Outline of the process for the expression and purification of soybean nodule IMPDH fusion protein using the Novagen pET-32 expression system.

*Eco*R I and dephosphorylated with calf intestinal alkaline phosphatase (CIAP). The insert was gel purified as described above.

The insert and vector were ligated. The ligation reaction contained: 2 µl 10X ligase buffer (Promega), 2 µl 100 mM DTT, 1 µl 10 mM ATP, 50-100 ng prepared pET vector, 1 µl T4 DNA ligase (Promega, 0.2-0.4 U/µl), 50 ng prepared insert and nuclease-free water to bring the volume to 20 µl. The ligation solution was incubated overnight at 15°C. After ligation, the reaction mixture was used to transform the non-expression host, *E. coli recA*$^-$ strain "NovaBlue". Plasmid DNA from selected transformants was isolated, digested with *Eco*R I and applied to a 1.2% agarose gel to check the size of inserts. Plasmid DNA from two colonies, designated as "pET-IMPDH-Nova-6" and "pET-IMPDH-Nova-8", was isolated and sequenced from the 5'-end using the S·Tag primer from Novagen to verify the open reading frames.

Host strains for the expression of soybean nodule IMPDH cDNA in the pET-32b expression system

Three different strains of *E. coli* were used to express the soybean nodule IMPDH fusion protein: 1) BL21(DE3)LysS, 2) AD494(DE3), and 3) AD494(DE3)LysS. Plasmid DNA from pET-IMPDH-Nova-8 was isolated and transformed into the three different expression hosts. Transformed clones corresponding to the hosts of BL21(DE3)LysS, AD494(DE3) and AD494(DE3)LysS were designated as "pET-IMPDH-BLLys", "pET-IMPDH-AD" and "pET-IMPDH-ADLys",

respectively. The pET-32b vector DNA without an insert was also transformed into these expression hosts as a control in the expression and purification experiments.

Induction and optimization of fusion protein expression: total protein sample analysis

In order to detect the induction of expression and determine the optimal expression conditions (host strain, temperature and induction time), the total cellular proteins from the above three expression hosts were analyzed as illustrated in Figure 5-3.

Clones of pET-IMPDH-BLLys, pET-IMPDH-AD and pET-IMPDH-ADLys were cultured in LB broth containing 50 μg/ml carbenicillin at 37°C overnight. The overnight culture was diluted 100-fold with LB broth containing carbenicillin (100 μg/ml) and cultured at the desired temperatures. When the culture reached OD_{600}=0.6, IPTG was added to obtain a final concentration of 1 mM. A 1-ml sample was taken from each culture flask every 1-2 h and centrifuged at 12,000 x g for 3 min. The supernatant fluid was discarded and the pellet was resuspended in 100 μl 1X SDS gel loading buffer (50 mM Tris-Cl, pH 6.8; 100 mM DTT; 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol), and heated at 100°C for 3 min. A 20-μl volume from each sample was loaded on a 12.5% SDS-PAGE gel. After electrophoresis, the proteins were visualized with Coomassie Brilliant Blue R-250 staining.

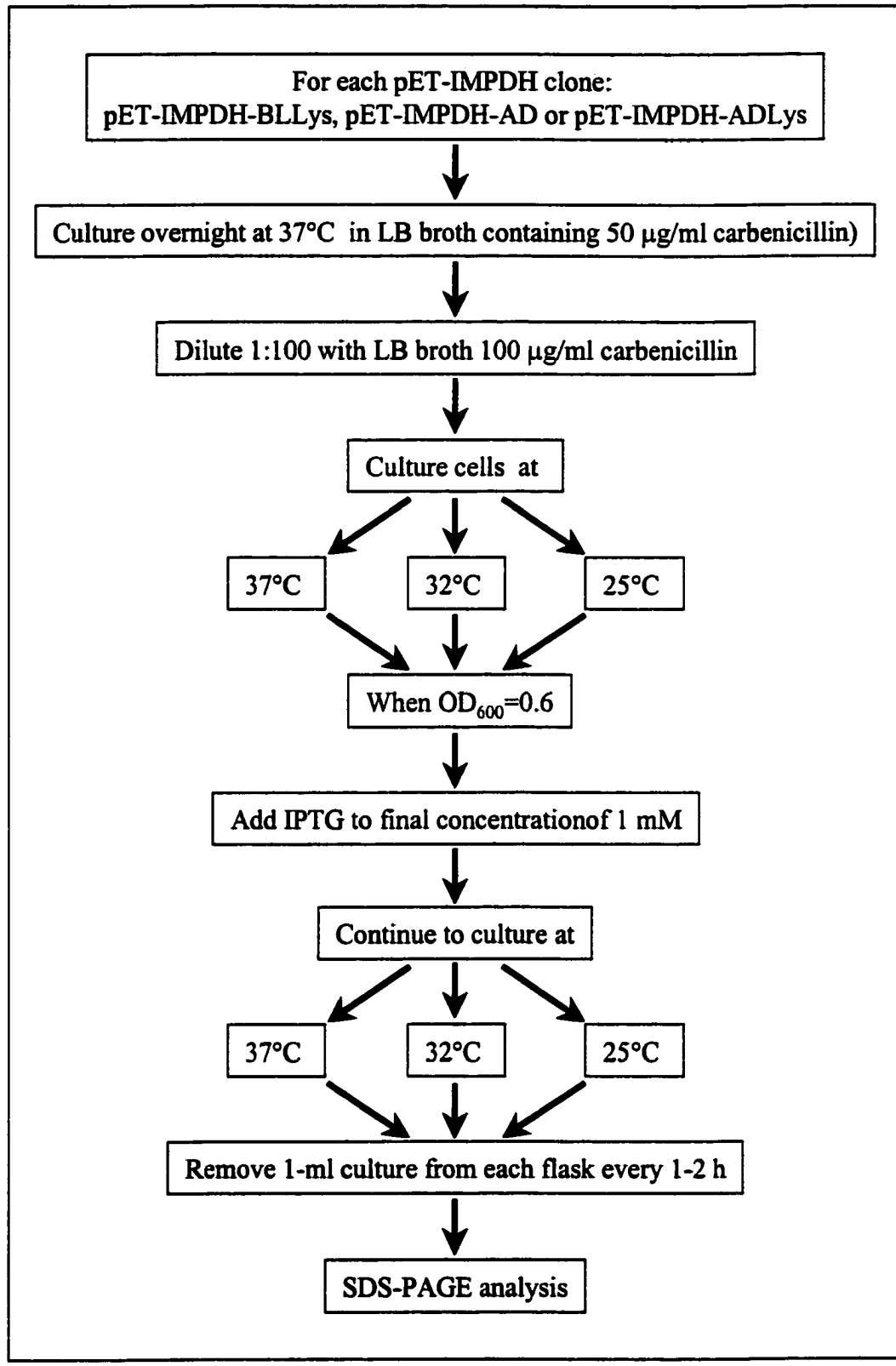


Figure 5-3. Experimental outline for the induction and optimization of fusion protein expression.

Subcellular fraction analysis

In order to determine the localization and solubility of the fusion protein, subcellular fractions from the three host strains were analyzed. After 4 hours of IPTG induction, 40-ml culture of induced cells were harvested by centrifugation at 6,500 x g for 15 min at 4°C. After centrifugation, the cell pellet was resuspended in an ice-cold solution of 20% sucrose; 2.5 mM EDTA; 20 mM Tris-HCl, pH 8.0 to a concentration of 5 OD_{550} units/ml and incubated on ice for 10 min. The osmotically shocked cells were centrifuged at 15,000 x g for 30 seconds. After centrifugation, the pellet was resuspended in the same volume (as used earlier) before of ice-cold solution containing 2.5 mM EDTA; 20 mM Tris-HCl, pH 8.0 and incubated for 10 min. The periplasmic fraction (supernatant) was separated from the other cell components by centrifugation at 15,000 x g for 10 min.

The pellet was completely resuspended in 4 ml of cold 20 mM Tris-HCl, pH 7.5, treated with 100 μg/ml lysozyme at 30°C for 10 min and sonicated. The soluble fraction (supernatant) was separated from the inclusion bodies (pellet) by centrifugation at 14,000 x g for 10 min. A 100-μl aliquot of the soluble fraction was mixed with 100 μl 2X SDS gel loading buffer and heated for 3 min at 100°C. Twenty μl were loaded onto each lane and separated by electrophoresis of a 12.5% SDS-PAGE gel. Protein bands on gels were revealed by staining with Commassie Blue.

The inclusion bodies were further purified by repeated centrifugation and washing steps as follows. The pellet was resuspended in 750 μl of 20 mM Tris-

HCl, pH 7.5, the suspension was centrifuged at 10,000 x g for 5 min. The final pellet was resuspended in 100 µl 1X SDS gel loading buffer, heated at 100°C for 3 min and a 20-µl aliquot was analyzed by SDS-PAGE. Samples used for affinity chromatography were prepared by resuspending the pellets in 5 mM imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9 and 6 M guanidine-HCl or 6 M urea.

Detection of fusion protein expression by Western blot analysis

The identity of the expression product was determined by Western blotting with S-protein AP (alkaline phosphatase) conjugate (Novagen, Cat. No.69598-3). S-protein recognizes the vector-encoded S·Tag and was used as the antibody in the Western blot to detect the S·Tag of the fusion proteins. To start the Western blot, proteins on the SDS-PAGE gel were electrophoretically transferred onto nitrocellulose membranes. Membranes bearing the proteins were then saturated with 4.5% skim milk (in 1X TBST: 20 mM Tris-HCl, pH7.5; 150 mM NaCl, 0.05% (v/v) Tween 20) for 30 min, followed by an 30-min incubation with S-protein AP conjugate (1:5,000 diluted with 1X TBST). After incubation, the blot was washed three times with 1X TBST for 10-15 min per wash and membranes were briefly rinsed twice with 1X TBS (20 mM Tris-HCl, pH7.5; 150 mM NaCl). The target protein was visualized by incubating the membrane in 20 ml alkaline phosphatase buffer (100 mM Tris-HCl, pH 9.5; 100 mM NaCl; 5 mM $MgCl_2$) containing 132 µl of 25 mg/ml NBT (nitroblue tetrazolium) and 66 µl of 50 mg/ml BCIP (5-bromo-4-chloro-3-indolyl phosphate).

IMPDH activity assay of soluble fractions

Soluble fractions were obtained according to the procedures described previously and 100 µl was used to determine IMPDH activity. For each host strain, the soluble fraction from a corresponding control (pET-32b without insert) was tested under the same conditions. IMPDH activity detected from the control represents the basal level of IMPDH activity required to maintain the essential metabolism of the cell. Therefore, the IMPDH activity of the fusion protein was obtained by subtracting basal activity.

IMPDH enzyme activity was measured by continuously monitoring the IMP-dependent reduction of NAD^+ at 340 nm (Atkins, 1985; Yang, 1997). The change in absorbance was measured at 25°C in a Beckman DU 7500 spectrophotometer. The standard reaction mixture included: 50 mM Tricine-KOH (pH 8.8); 1 mM DTT; 2 mM NAD^+; 1.35 mM IMP and 50 µl soluble fraction. The final reaction volume was 1 ml.

The reaction was initiated with the addition of soluble protein extract. The blank contained 50 mM Tricine-KOH, pH 8.8 with 1 mM DTT. H_2O was used to replace IMP in the control reaction. The change in A_{340} was monitored for 5 minutes for replicated samples. Final results for IMPDH activity were obtained from at least four replicate assays.

Purification of the IMPDH fusion protein using His·Bind chromatography under nondenaturing conditions

Approximately 5 ml of 50% His·Bind Resin (Novagen, Cat. No.69670-3) was packed into a column under gravity flow and washed with 3 volumes of sterile deionized water. The resin was charged with 5 volumes of 1X charge buffer (50 mM $NiSO_4$) and equilibrated with 3 volumes of 1X binding buffer (5 mM imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9). Extract was prepared from a 100-ml culture by solubilizing inclusion bodies in 1X binding buffer containing 6 M guanidine-HCl. The solubilized protein fraction from a 100-ml culture was loaded onto the column. The column was washed successively with 25 ml of 1X binding buffer, 15 ml of 1X wash buffer (60 mM imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9), 15 ml of 100 mM imidazole buffer (13.5 ml 1X binding buffer mixed with 1.5 ml 1X elution buffer) and 15 ml of 1X elution buffer (1 M imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9). Fractions from the wash with 0.1X and 1X elution buffer were collected.

Column fractions (500 µl) were dried in a Savant Speed-vac. The dried sample was resuspended in 100 µl 1X SDS gel loading buffer and heated at 80°C for 3 min. A 20-µl sample was loaded in each lane for SDS-PAGE analysis. After electrophoresis, gels were stained with Commassie Blue to reveal protein bands.

Purification of the IMPDH fusion protein using His·Bind chromatography under denaturing conditions

Due to the insolubility of the target protein, purification under denaturing conditions became necessary. The purification of the target protein using His·Bind chromatography under denaturing conditions was similar to the purification under native conditions described above except that a lower imidazole concentration was used in the wash buffer. With the His·Bind resin, target proteins tend to elute at lower imidazole concentrations in the presence of 6 M urea.

After the His·Bind resin was packed into a column and charged with 50 mM $NiSO_4$, the column was equilibrated with 1X binding buffer plus 6 M urea. Extract was prepared from a 100-ml culture by solubilizing inclusion bodies in 1X binding buffer containing 6 M urea and the solubilized protein extract was loaded onto the column. The column was washed with 25 ml of 1X binding buffer plus 6 M urea, pH 7.9 and 15 ml of 20 mM imidazole buffer plus urea, pH 7.9. The protein was eluted with 400 mM imidazole buffer plus urea, pH 7.9.

Samples of the eluted fractions from the His·Bind column were concentrated with a Centricon-30 concentrator by centrifugation at 4,000 rpm and washed several times with 0.05 M Tris-HCl (containing 6 M urea, pH 9.0). Protein samples in 0.05 M Tris-HCl containing 6 M urea, pH 9.0 were ready for the anion exchange chromatography under denaturing conditions.

For SDS-PAGE analysis, eluates from the His·Bind columns were dialyzed successively against 4 M, 2 M, 1 M and 0 M urea in 20 mM Tris-HCl, pH 8.0.

Solutions were changed at 4h intervals. Dialysis was carried out at 0-4°C using SpectroPor dialysis membrane with a molecular weight cutoff (MWCO) of 3,500 Da (Spectrum, Laguna Hills, CA 92653). Fifty microliters of each fraction was dried, resuspended in 10 µl of 1X SDS gel loading buffer, heated in boiling water for 3 min and loaded onto a 12.5% SDS-PAGE gel. Gels were stained as described above.

Anion-exchange chromatography

Eluates obtained by His·Bind chromatography were further purified by anion-exchange chromatography using a Pharmacia Resource Q column and a Pharmacia FPLC system. The column was equilibrated with 0.05 M Tris-HCl, pH 9.0 containing 6 M urea (buffer A). Samples were prepared in buffer A as described above and loaded onto the column. Proteins were eluted from the column with a linear NaCl gradient by adjusting the proportion of buffer B (1 M NaCl; 0.05 M Tris-HCl, pH 9.0; 6 M urea) from 0-100%. The flow rate for the elution was 1 ml/min. The elution of the protein was monitored by recording the absorbance at 280 nm, and 1-ml fractions were collected.

Fractions were pooled and dialyzed against 4 M, 2 M, 1 M and 0 M urea in 20 mM Tris-HCl, pH 8.0 successively using a 3,500 MWCO membrane. After dialysis, 50 µl of each fraction was dried, resuspended in 10 µl 1X SDS gel loading buffer, heated for 3 min at 80°C and loaded onto a SDS-PAGE gel. The IMPDH fusion protein was identified by Western blot analysis.

Antibody preparation

Dialyzed fractions from the anion-exchange column containing the recombinant soybean nodule IMPDH protein were applied to a 12.5% SDS-PAGE gel. The gel was stained with 0.25 M cold KCl for 10 minutes, and the band containing the target protein was excised from the gel, lyophilized until dry, and ground into a fine powder. The protein-containing powder was used for injection into a rabbit for antibody preparation at the Bethyl Laboratories, Inc. (Montgomery, TX 77356). The rabbit was injected five times at 14-day intervals. A sample of preimmune serum was drawn from the animal before the immunization was begun. The total amount of IMPDH fusion proteins used for injection was 800 μg.

RESULTS

Construction of plasmids

Results of agarose gel electrophoresis showed that the constructed plasmid contained the correct insert (Figure 5-4). Plasmid DNA of pET-32b containing soybean nodule IMPDH cDNA (pET-IMPDH-Nova) was isolated and sequenced from the 5'-end using the S·Tag primer. Sequencing results confirmed that the *IMPDH* cDNA was in the correct open reading frame for translation driven by the T7 promoter.

Induction and optimization of the fusion protein expression

Different expression hosts (BL21(DE3)LysS, AD494(DE3) and AD494(DE3)LysS), different temperatures (37°C, 32°C and 25°C) and different induction times were used to determine the optimum conditions for fusion protein expression.

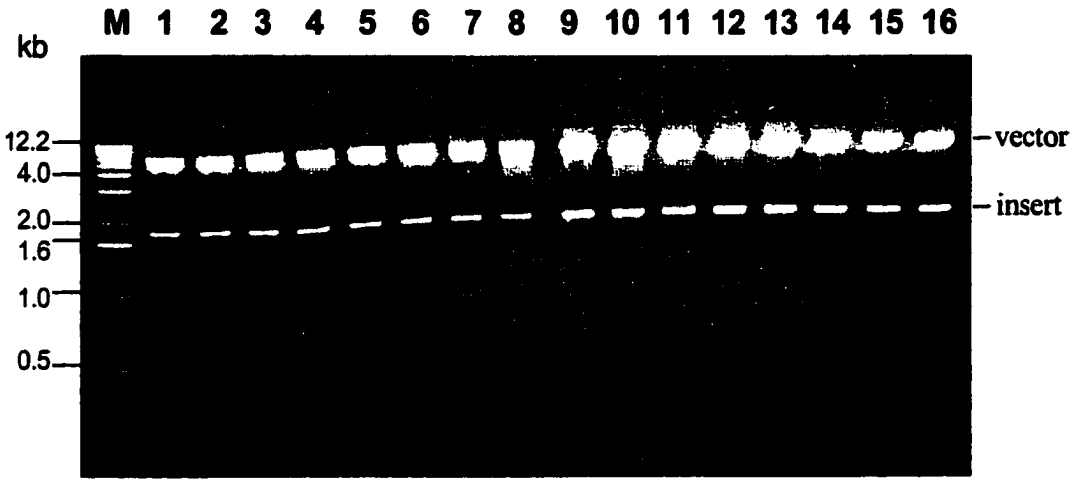


Figure 5-4. *Eco*R I digestion of selected pET-IMPDH clones. The ~1.8 kb band corresponded to the *IMPDH* cDNA insert, while the ~6.0 kb band corresponded to the vector fragment of pET-32b.

The pattern of protein expression at 37°C is presented in Figure 5-5. The expression of thioredoxin (~25 kDa, marked with arrowheads) was much more distinctive in the BL21(DE3)LysS host (A) than that in the AD494(DE3) (B) or AD494(DE3)LysS (C) host strains. Thioredoxin production increased during the first 6 hours after IPTG induction. The amount of thioredoxin that accumulated in the cells, however, decreased by the 43 h time point. The expression of the

putative IMPDH fusion protein (arrows) with a predicted size of ~70 kDa was detectable but not very distinctive in all three hosts.

The optimal temperature for expression of both the thioredoxin control and the IMPDH fusion protein in all three hosts appeared to be 32°C (Figure 5-6). Thioredoxin expression was strong and the putative IMPDH fusion protein band was very distinctive. The expressed fusion protein continued to accumulate for at least 16 hours after IPTG induction.

Expression at 25°C was not as strong as that at 32°C (Figure 5-7). The bands for thioredoxin and IMPDH fusion protein were only visible in the host strain AD494(DE3) (Figure 5-7B). In general, the IMPDH fusion protein was optimally expressed in all the three hosts at 32°C. There was no obvious difference in the expression level among the three hosts at 37°C or at 32°C. At 25°C, the expression level in AD494(DE3) appeared to be higher than in BL21(DE3)LysS and AD494(DE3)LysS.

Subcellular fraction analysis

The expression of IMPDH fusion protein in soluble fractions and inclusion bodies in the three different hosts is shown in Figure 5-8A. Although the IMPDH fusion protein was overexpressed, it existed almost exclusively in the form of inclusion bodies. In the soluble fraction, the IMPDH fusion protein was almost undetectable in the SDS-PAGE gel stained with standard Commassie Brilliant Blue R-250. In contrast, thioredoxin was detected only in the soluble fraction.

Western blot analysis was used to detect the S-Tag and thereby confirmed the identity of the overexpressed fusion proteins (Figure 5-8B). The intensely stained ~25 kDa band in the soluble fraction and the heavily stained ~70 kDa band in the inclusion body fraction were identified as S-tagged fusion proteins. Under the control of T7 RNA polymerase, these fusion proteins were highly expressed in all three host strains.

IMPDH enzyme activity from each of the soluble fractions was measured (Table 5-1). For each host, the soluble fraction from both the pET-32b (control) and the pET-IMPDH were assayed. As shown in Table 5-1, the IMPDH activity was always higher in the control. These results clearly demonstrated that the IMPDH fusion protein present in the soluble fraction was not active. One explanation for the negative net IMPDH activity (ΔOD_{340}/min) might be that the bacteria grow better in controls under the same conditions.

Table 5-1. IMPDH activities in soluble fractions of pET-32b expression host strains

ΔOD_{340}/min (Average)	BL21(DE3)LysS	AD494(DE3)	AD494(DE3)LysS
pET-32b (Control)	0.01540	0.00478	0.00212
pET-IMPDH	0.00969	0.00257	0.00184
ΔOD_{340}/min	-0.00571	-0.00221	-0.00028

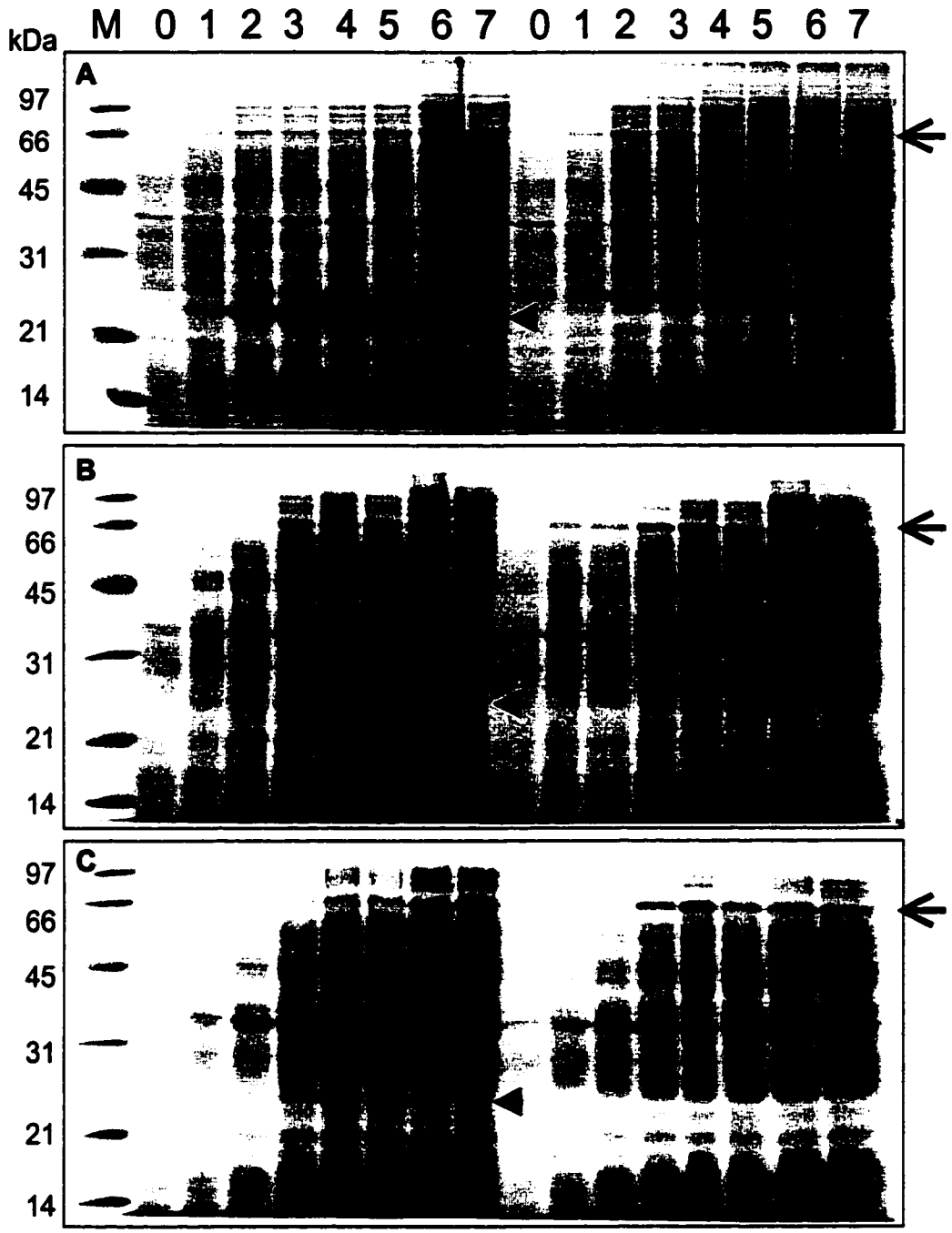


Figure 5-5. Analysis of total cellular protein by SDS-PAGE (37°C). Host cells BL21(DE3)LysS **(A)**, AD494(DE3) **(B)** and AD494(DE3)LysS **(C)** were cultured at 37°C and harvested 0, 1, 2, 3, 4, 5, 6, 43 hours (lane 0-7) after IPTG induction. In each gel, the left seven lanes reflect expressions of thioredoxin in pET-32b vector without insert. Arrowheads indicate the expressed thioredoxin (~25 kDa), while the arrows indicate the expected size of IMPDH fusion protein (~70 kDa).

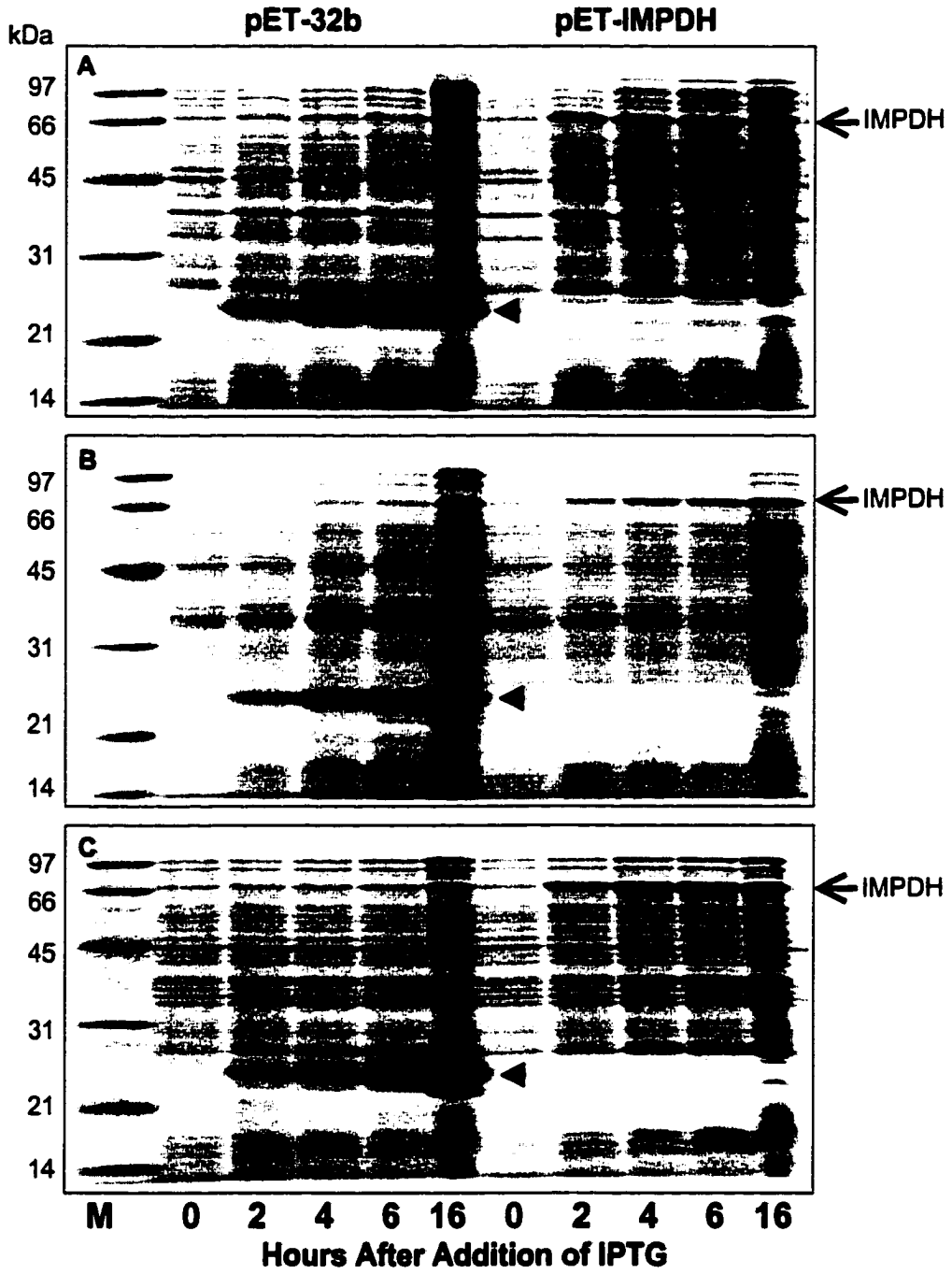


Figure 5-6. Analysis of total cellular protein by SDS-PAGE (32°C). Host cells BL21(DE3)LysS **(A)**, AD494(DE3) **(B)** and AD494(DE3)LysS **(C)** were cultured at 32°C and harvested 0, 2, 4, 6, 16 hours after IPTG induction. In each gel, the left five lanes reflect expression of thioredoxin in the pET-32b vector without insert. Arrowheads indicate the expressed thioredoxin (~25 kDa), while the arrows indicate the expected size of IMPDH fusion protein (~70 kDa).

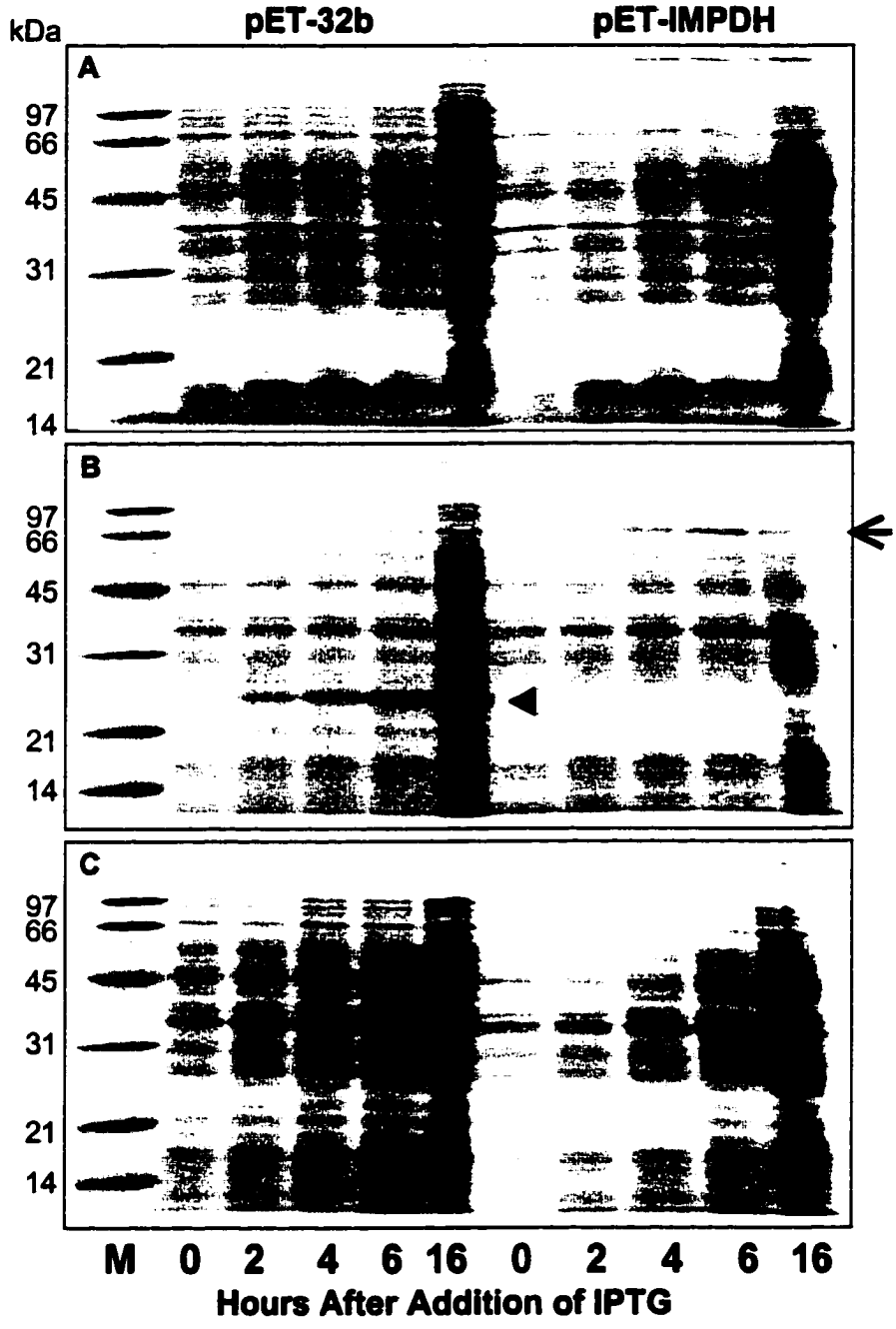


Figure 5-7. Analysis of total cellular protein by SDS-PAGE (25°C). Host cells BL21(DE3)LysS **(A)**, AD494(DE3) **(B)** and AD494(DE3)LysS **(C)** were cultured at 25°C and harvested 0, 2, 4, 6, 16 hours after IPTG induction. In each gel, the left five lanes reflect expression of thioredoxin in pET-32b vector without insert. Arrowheads indicate the expressed thioredoxin (~25 kDa), while the arrows indicate the expected size of IMPDH fusion protein (~70 kDa).

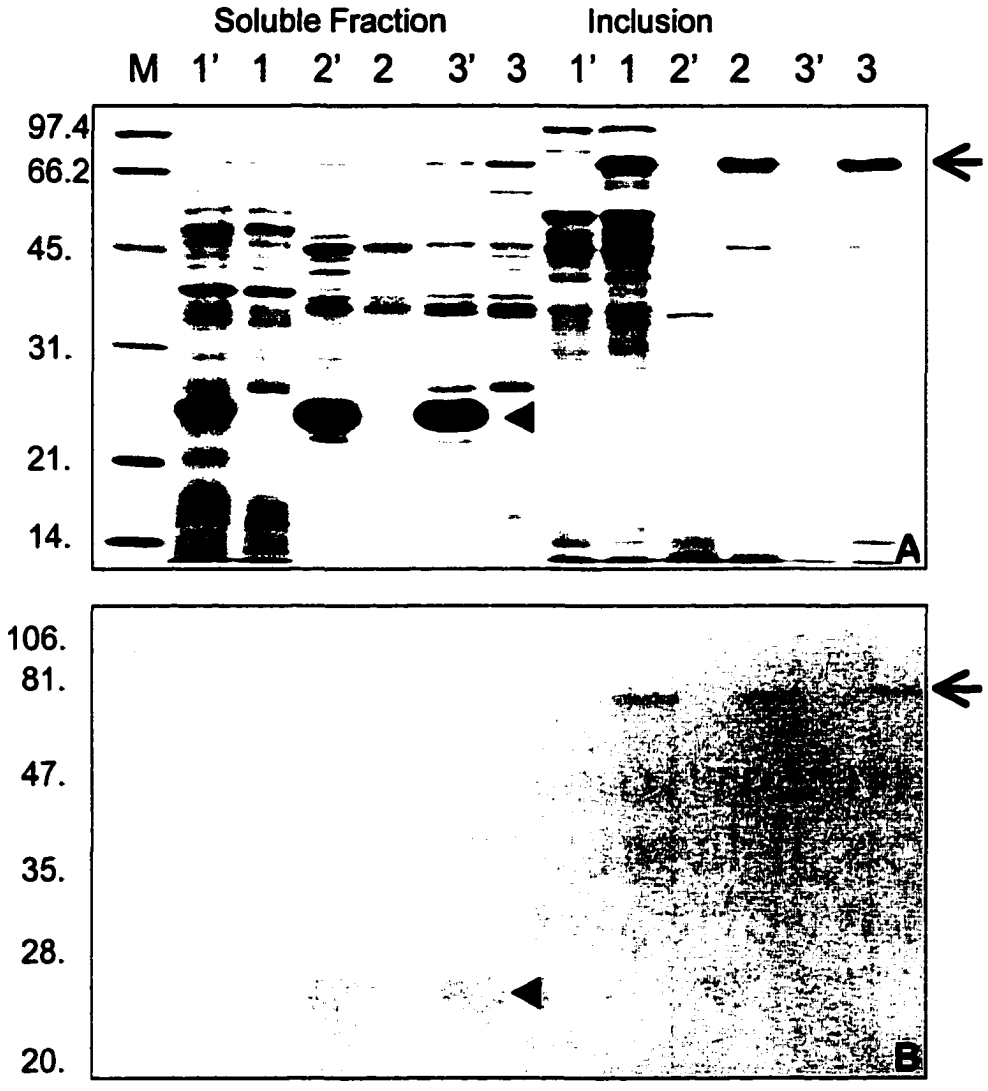


Figure 5-8. Analysis of subcellular fractions by SDS-PAGE (Panel A) and Western blotting (Panel B). Host cells were cultured at 32°C and harvested 4 hours after IPTG induction. Soluble fractions and inclusion bodies were isolated and applied to a 12.5% SDS-PAGE gel and stained with Coomassie Brilliant Blue R-250 (A). Proteins from a duplicate gel were transferred to a nitrocellulose membrane and detected with S·protein AP-conjugate in a Western blot (B). The expression of pET-IMPDH is represented by the number of 1, 2 and 3, corresponding to the host strains of BL21(DE3)LysS, AD494(DE3) and AD494(DE3)LysS, respectively. The control expression of pET-32b is indicated by the same numbers but with primes (1', 2' and 3'). The arrow indicates the putative IMPDH fusion protein, and the arrowhead indicates the thioredoxin fusion protein.

Purification of the IMPDH fusion protein by His·Bind chromatography

In order to obtain purified IMPDH fusion protein, inclusion bodies from a 100-ml culture of each of the three host strains were isolated, solubilized and loaded onto a His·Bind column. The fusion protein was eluted with a 0.1X elution buffer containing 100 mM imidazole followed by a 1X elution buffer (containing 1 M imidazole). When the chromatography was performed under nondenaturing conditions (Figure 5-9), the fusion protein was eluted in both the 100 mM imidazole and 1 M imidazole fractions. The purified fusion protein was ~70 kDa in size as determined by SDS-PAGE. For host strains BL21(DE3)LysS and AD494(DE3), the amount of fusion protein seemed to be the same in 100 mM and 1 M imidazole fractions. In AD494(DE3)LysS, almost all of the fusion protein was eluted with 100 mM imidazole. The band in the 1 M imidazole eluate was very weak.

To improve the purification, His·Bind chromatography was performed under denaturing conditions by adding 6 M urea to the binding, washing and elution buffers (Figure 5-10). The concentration of imidazole necessary to elute the target protein was lower when compared to the concentration necessary to elute the fusion protein under native conditions. Almost all of the fusion protein was eluted in the first and second fractions (corresponding to lane 6 and 7 in Figure 5-10). These two fractions were pooled and designated as the His·Bind eluates.

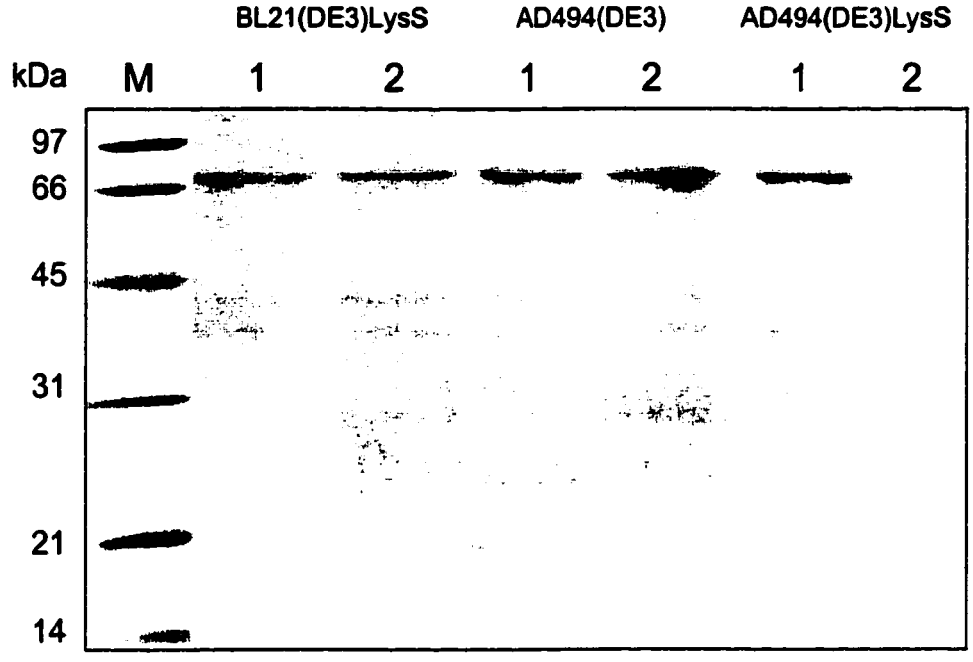


Figure 5-9. SDS-PAGE gel of fractions after His·Bind chromatography of solubilized inclusion bodies. Inclusion bodies were isolated from 100 ml culture of BL21(DE3)LysS, AD494(DE3) or AD494(DE3)LysS grown at 32°C. Isolated inclusion bodies were purified, resuspended in 5 mM imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9 and 6 M guanidine-HCl, and loaded onto the His·Bind column. **1.** Proteins eluated from 100 mM imidazole; **2.** Proteins eluated with 1 M imidazole. 100 µl from each fraction was dried, resuspended in 20 µl 1x SDS gel loading buffer, boiled and loaded in each lane. The gel was stained with Commassie Blue to reveal the protein bands.

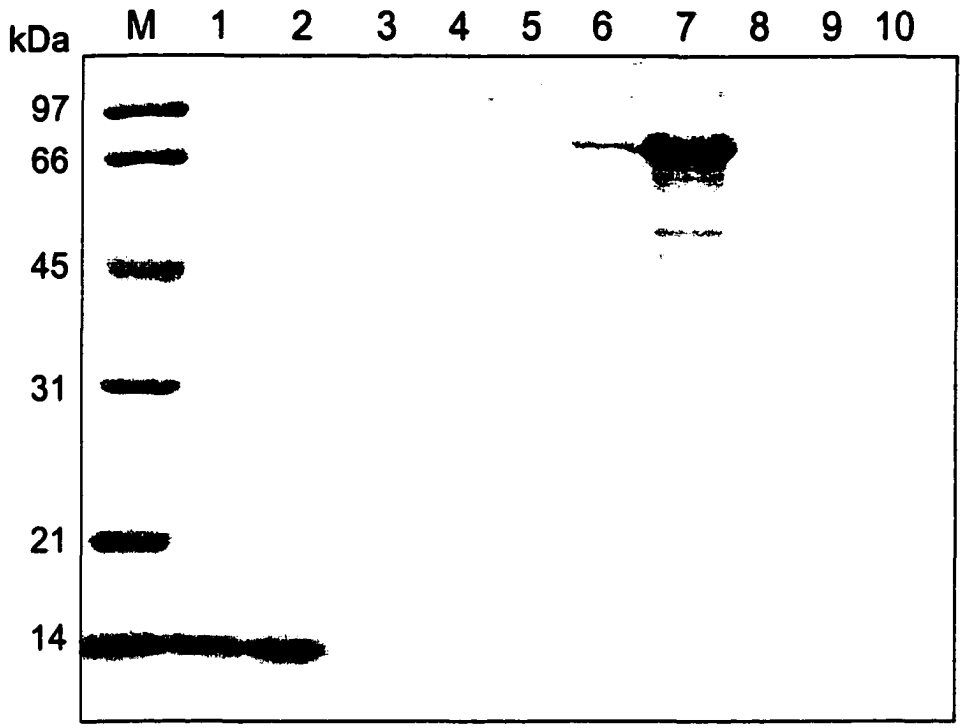


Figure 5-10. SDS-PAGE gel of fractions after His·Bind chromatography of solubilized inclusion bodies. Chromatography was performed under denaturing conditions with the inclusion of 6 M urea in the binding buffer, washing buffer and elution buffer. Inclusion bodies were isolated from 100 ml of the BL21(DE3)LysS culture grown at 32°C, purified, resuspended in 5 mM imidazole buffer (5 mM imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9) containing 6 M urea, and loaded onto the His·Bind column.

Lanes: **1.** Flowthrough; **2, 3.** The first and the last fraction from the wash with 5 mM imidazole buffer containing 6 M urea; **4, 5.** The first and the last fraction from the wash with 20 mM imidazole buffer containing 6 M urea; **6-9.** The first four fractions eluted with 400 mM imidazole buffer containing 6 M urea; **10.** First fraction from the wash with 1X strip buffer (100 mM EDTA; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9).

Fractions from the column were dialyzed successively against 4 M, 2 M, 1 M and 0 M urea in 20 mM Tris-HCl, pH 8.0 to remove imidazole, NaCl and urea. A 50-µl volume of the desalted fractions was dried, resuspended with 10 µl 1X SDS gel loading buffer, boiled and then loaded in each lane of a 12.5% SDS-PAGE gel. Gel was stained with Commassie Blue to reveal the protein bands.

Purification of IMPDH fusion protein by anion-exchange chromatography

Desalted and concentrated His·Bind eluates in buffer A (0.05 M Tris-HCl, pH 9.0 containing 6 M urea) were purified by anion-exchange chromatography on a Resource-Q column. The bound protein was eluted from the column with a linear NaCl gradient. A large A_{280}-absorbing peak (labeled as "1" in Figure 5-11A) was eluted with at 0.1 M NaCl (10% buffer B). Other smaller peaks were also observed as the NaCl concentration increased. SDS-PAGE analysis (Figure 5-11B) showed that peak 1 represented fractions of expected IMPDH fusion protein (~70 kDa). Proteins in peak 2, 3, 4, 5 also showed a principal band with an estimated molecular weight of ~70 kDa, and may still be the fusion protein.

Antigen preparation

To further isolate the IMPDH fusion protein, peak 1 samples from the anion-exchange chromatography were dialyzed to remove salt and urea, and applied to SDS-PAGE gel. The protein was visualized by soaking the gel in 0.25 M KCl (Figure 5-12). Acrylamide gel containing the fusion protein was excised from the gel, lyophilized, ground into a fine powder and used to immunize a rabbit.

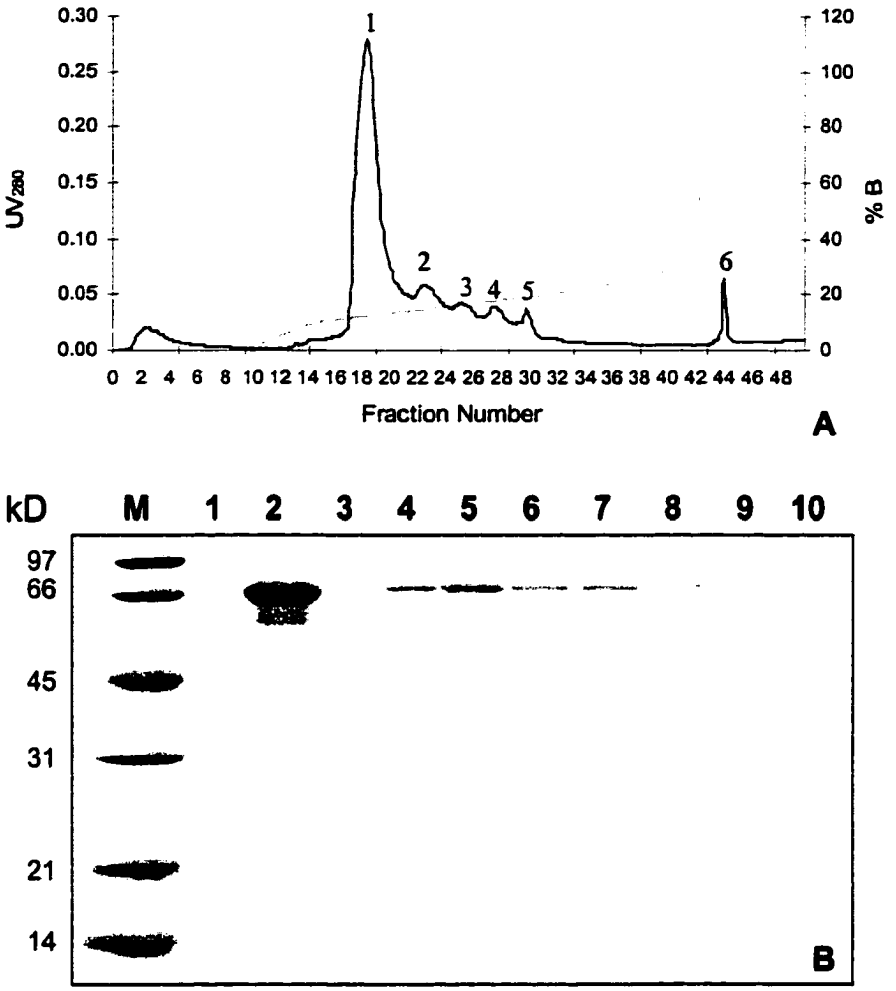


Figure 5-11. Purification of IMPDH fusion protein by anion-exchange chromatography under denaturing conditions. Eluates from the His·Bind chromatography were pooled, desalted and concentrated with Buffer A, and applied to a Resource Q column equilibrated with buffer A. Proteins were eluted with a programmed linear salt gradient. Buffer A was 0.05 M Tris-HCl, pH 9.0 containing 6 M urea was used as buffer A. Buffer B was 1 M NaCl in buffer A. **A).** Absorbance (280 nm) profile. The X-axis is the fraction number. The Y-axes are the absorbance at 280 nm (UV_{280}) (left) and the concentration of buffer B (%B) (right). **B).** 12.5% SDS-PAGE gel of selected fractions, stained with Commassie Brilliant Blue R-250. Fractions from the column were dialyzed successively against 4 M, 2M, 1M and 0 M urea in 20 mM Tris-HCl, pH 8.0 to remove imidazole, NaCl and urea. Fifty µl of the desalted fractions were dried, resuspended in 10 µl 1x SDS gel loading buffer, boiled and loaded in each lane. Lanes: **M.** molecular weight marker; **1.** Flowthrough, fractions 2,3,4,5; **2.** Peak 1, fractions 18,19,20,21; **3.** Fraction 17; **4.** Fraction 22; **5.** Fractions 23,24; **6.** Fraction 25; **7.** Fractions 26,27; **8.** Fractions 28,29; **9.** Fractions 30,31; **10.** Fractions 44, 45.

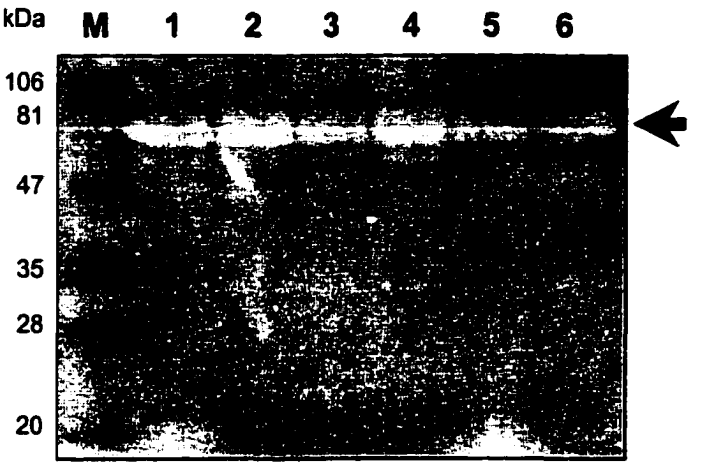


Figure 5-12. SDS-PAGE gel of purified IMPDH fusion protein after anion-exchange chromatography. Protein was detected by soaking the gel in 0.25 M KCl. A gel fragment containing the protein was excised from the gel, lypholized and the dry gel piece was ground into a fine powder. The powder was used to inject a rabbit for antibody preparation. Each lane was loaded with 129 μg of peak 1 protein after anion-exchange chromatography.

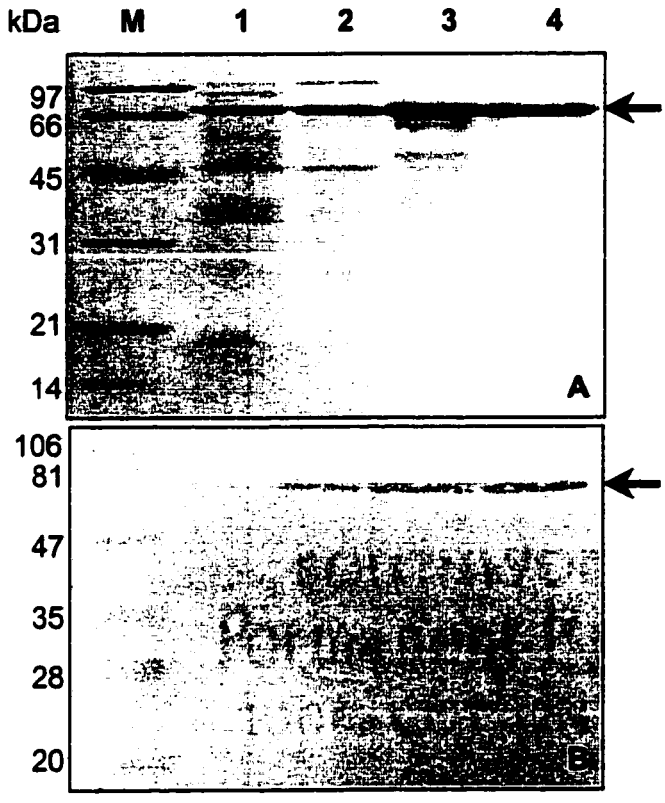


Figure 5-13. SDS-PAGE gel (A) and Western Blot (B) of IMPDH fusion protein at different purification steps. **1.** Total protein sample from the 32°C bacterial culture after IPTG induction; **2.** Purified inclusion bodies; **3.** His·Bind eluates (the first two fraction eluted with 400 mM imidazole buffer) from the Ni-column; **4.** Peak 1 fractions after anion-exchange chromatography. SDS-PAGE gel was stained with Commassie Brilliant Blue R-250 (A). For the Western Blot (B), S-protein alkaline phosphatase conjugate was used to detect the S·Tag in the fusion proteins.

DISCUSSION

Detection of IMPDH fusion protein expression

Escherichia coli is the most frequently used prokaryotic expression system for high-level production of heterologous proteins (Hannig and Makrides, 1998). A successful expression of an eukaryotic protein in *E. coli* depends on many different factors. These include but are not limited to: 1) The activity and stability of the target protein. It has been reported that the amino acid immediately following the N-terminal methionine (penultimate amino acid) influences the stability of the fusion protein (Tobias *et al.*, 1991) and the amino acid at this position affects the removal of the N-terminal fMet. Studies have suggested that leucine would be a poor choice for the penultimate position because the penultimate leucine is likely to be exposed by fMet processing and then be targeted for rapid degradation (Tobias *et al.*, 1991). 2) The secondary structure of the mRNA. The secondary structure of the transcript can interfere with the AUG translation initiation codon and the ribosome binding site (Tessier *et al.*, 1984; Looman *et al.*, 1986; Lee *et al.*, 1987). 3) The usage of rare codons. Excessive rare codon usage in the target gene has also been implicated as a cause for low level expression (Zhang *et al.*, 1991; Sorensen *et al.*, 1989), especially when multiple rare codons occur near the amino terminus (Chen and Inouye, 1990). A number of studies have also indicated that high usage of the arginine codons AGA and AGG can have severe effects on protein yield. The impact appears to be highest when these codons appear consecutively (Brinkmann *et al.*, 1989; Hua *et al.*, 1994; Schenk *et al.*, 1995; Zahn, 1996; Calderone *et al.*, 1996). However,

only a subset of codons characterized as rare in highly expressed *E. coli* genes have levels of cognate charged tRNAs that are low enough to present potential problems for translation elongation (Ikemura, 1985). 4) The occurrence of stop codons. Unexpected stop codons can be generated by mutation, especially when cloning PCR products. These stop codons cause premature termination of translation.

In the current work, the soybean nodule *IMPDH* cDNA was cloned in the Novagen pET-32 vector and expressed in *E. coli*. The successful expression of the *IMPDH* cDNA suggests that the factors discussed earlier did not appear to negatively affect expression. In the soybean nodule *IMPDH* cDNA, there are 22 arginines (R), 14 of which are coded by rare codons in *E. coli* (AGG / AGA). During the subcloning of the *IMPDH* cDNA into the pET-32 expression vector, *Pfu* DNA polymerase was used instead of *Taq* DNA polymerase for PCR in order to minimize mutations or mismatchs.

Up to now, cloned *IMPDH* cDNAs / genes from humans (Konno *et al.*, 1991; Carr *et al.*, 1993; Hager *et al.*, 1995), *Tritrichomonas foetus* (Beck *et al.*, 1994; Huete-Pérez *et al.*, 1995) and *Borrelia* (Margolis *et al.*, 1994; Zhou *et al.*, 1997) have been successfully expressed in *E. coli*. Among these, the human protein is the only IMPDH of eukaryotic origin to be successfully expressed in *E. coli*.

Solubility of the fusion protein: effects of host strain and growth temperature

Results from SDS-PAGE gels and Western blot analysis showed that the overexpressed IMPDH fusion protein was primarily found in insoluble inclusion bodies (Figure 5-8). IMPDH activity was not detected in the soluble fractions in any of the hosts (Table 5-1), indicating that there was either no active IMPDH in the soluble fractions or the amount of the enzyme was too low to be detected.

According to the published studies on heterologous gene expression, several approaches can be used to obtain greater solubility and activity: 1) Use of different expression hosts. For some unknown reason, some host strains produce soluble products while others produce insoluble protein. 2) Use of different growth temperatures. Growth at 37°C causes some proteins to accumulate as inclusion bodies, whereas incubation at 30°C leads to soluble, active protein (Schein & Noteborn, 1989; Burton *et al.*, 1991). 3) Reduce the rate of expression by using a lower concentrations of IPTG (0.01-0.1 mM) over a longer induction period. 4) Use of a specific culture medium (Blackwell and Horgan, 1991). 5) Fuse the target protein to thioredoxin. 6) Direct the heterologous proteins to the periplasm especially when attempting to isolate active, properly folded proteins containing multiple disulfide bonds. In contrast to the cytoplasm, the periplasm of *E. coli* is an oxidizing environment that contains enzymes that catalyze the formation and isomerization of disulfide bonds (Rietsch *et al.*, 1996; Raina and Missiakis, 1997; Sone *et al.*, 1997).

In the current work, the heterologous expression of soybean IMPDH in *E. coli* was carried out under different conditions. These included the use of

different host strains grown at different temperatures. None of the three hosts used expressed soluble fusion proteins (Figure 5-8, Table 5-1). Among the three temperatures used, 32°C seemed to be optimal for expression but resulted in the production of insoluble proteins (Figure 5-6, 5-8). Although expression at 25°C might lead to the production of more soluble fusion protein, but the expression level was very low even after 16 hours of induction (Figure 5-7). According to Novagen, the use of a pET-32 vector to produce thioredoxin fusion proteins may increase the yield of soluble product in the cytoplasm (Novagen pET system manual, 1999; LaVallie *et al.*, 1993; Novy *et al.*, 1995) because thioredoxin allows the formation of disulfide bonds in the *E. coli* cytoplasm, thereby affecting the solubility and/or activity of a given target protein. Stewart *et al.*(1998) have shown that disulfide bond formation in cytoplasm appears to be dependent on the presence of thioredoxins. However, this did not seem to be the case for soybean IMPDH since the expression of IMPDH fused to thioredoxin did not produce soluble recombinant IMPDH. Possible explanations for the lack of solubility and expression of active fusion protein include improper folding and disulfide bond formation because of the molecular size of IMPDH (53 kDa). Expression of small proteins (<30 kDa) in a soluble form is usually not a problem. The estimated molecular weight of the fusion protein would be greater than 70 kDa after its fusion with thioredoxin and the N-terminal tags. In order to get expression of soluble IMPDH in *E. coli*, different conditions or other expression vectors need to be tried. For example, a vector containing a signal sequence,

which enables the fusion protein to be transported to the periplasm or secreted should be considered.

In summary, the fusion protein of soybean IMPDH was successfully over-expressed in the form of insoluble inclusion bodies in *E. coli*. The recombinant protein, in the form of inclusion bodies, after solubilization and purification was suitable for use for antibody preparation.

Purification of the IMPDH fusion protein

Inclusion bodies have the following important advantages for purification: 1) they can represent the highest yielding fraction of target protein; 2) they are easy to isolate as an efficient first step in a purification scheme. Nuclease-treated, washed inclusion bodies are usually 75-95% pure target protein; 3) inclusion body formation protects the protein from proteolytic breakdown; and 4) toxic proteins may not inhibit cell growth when present in an inactive form such as inclusion bodies (Rudolph and Lilie, 1996; Mukhopadhyay, 1997).

Soybean nodule IMPDH fusion protein was expressed in inclusion bodies in our current work. Results from SDS-PAGE gel analysis showed that after several washes, the inclusion bodies contain almost pure IMPDH. This is especially true when the pET-32-IMPDH was expressed in the host strains AD494(DE3) and AD494(DE3)LysS (Figure 5-8A; Figure 5-13, lane 2). These inclusion bodies can be used directly for antibody preparation after suspension in PBS and emulsification with a suitable adjuvant (Marston, 1986; Fischer *et al.*, 1992). However, in order to avoid any non-specific reactions in antibody production, a

purer immunogen is desirable. Therefore, in our experiments, nickel column chromatography and anion-exchange chromatography under denaturing conditions were employed to further purify the IMPDH fusion protein (Figure 5-9; 5-10; 5-11; 5-13). After the final step of purification by anion-exchange chromatography, the target protein was virtually "pure" based on SDS-PAGE gel analysis. Only a few very weak bands were visible when large amounts of the fusion protein were loaded on the SDS-gel (Figure 5-11; 5-13A). In order to remove these minor contaminants and further purify the immunogen, purified fusion protein was separated by preparative SDS-PAGE (Figure 5-12) and IMPDH fusion protein was isolated by cutting the segment of the acrylamide gel containing the fusion protein from the gel. The gel pieces were dried and ground into a powder and used for antibody preparation.

Appendix: Characteristics of host strains for pET-32 system

Strain	Genotype	Description /Application	Antibiotic Resistance
NovaBlue	*endA1 hsdR17($r_{k12}^{-}m_{k12}^{+}$) supE44 thi-1 recA1 gyrA96 relA1 lac[F' $proA^{+}B^{+}$ $lacI^{q}ZΔM15$::Tn10]*	non-expression host, general purpose cloning, plasmid preps	Tetracycline (12.5 μg/ml)
BL21	F⁻*ompT* $hsdS_B(r_B^{-}m_B^{-})$ *gal dcm*	control non-expression host	none
BL21(DE3)	F⁻*ompT* $hsdS_B(r_B^{-}m_B^{-})$ *gal dcm* (DE3)	general purpose expression host	none
BL21(DE3)LysS	F⁻*ompT* $hsdS_B(r_B^{-}m_B^{-})$ *gal dcm* (DE3)LysS	high-stringency expression host	Chloramphenicol (34 μg/ml)
AD494	*Δara⁻ leu7697 ΔlacX74 ΔphoAPvuII phoR ΔmalF3* F'*[lac⁺($lacI^{q}$)pro]trxB*::kan	*trxB⁻* non-expression host; allows disulfide bond formation in *E. coli* cytoplasm	Kanamycin (15 μg/ml)
AD494(DE3)	*Δara⁻ leu7697 ΔlacX74 ΔphoAPvuII phoR ΔmalF3* F'*[lac⁺($lacI^{q}$)pro]trxB*::kan (DE3)	*trxB⁻* expression host; allows disulfide bond formation in *E. coli* cytoplasm	Kanamycin (15 μg/ml)
AD494(DE3)LysS	*Δara⁻ leu7697 ΔlacX74 ΔphoAPvuII phoR phoR ΔmalF3* F'*[lac⁺($lacI^{q}$)pro]trxB*:: ::kan(DE3)LysS	*trxB⁻* high-stringency expression host; allows disulfide bond formation in *E. coli* cytoplasm	Kanamycin (15 μg/ml) Chloramphenicol (34 μg/ml)

DE3: hosts are lysogen of bacteriophage DE3. Once a DE3 lysogen is formed, the only promoter known to direct transcription of the T7 RNA polymerase gene is the *lacUV5* promoter, which is inducible by IPTG. ***LysS***: host strains contain a compatible plasmid that provides a small amount of T7 lysozyme. T7 lysozyme cuts a specific bond in the eptidoglycan layer of the *E. coli* cell wall and it binds to T7 RNA polymerase. T7 lysozyme helps in rapid cell lysis.

REFERENCES

Atkins, C.A., Shelp, B.J. and Storer, P.J. (1985). Purification and properties of inosine monophosphate oxidoreductase from nitrogen-fixing nodules of cowpea (*Vigna unguiculata* L. Walp). Arch Biochem Biophys 236: 807-814.

Ausubel, F.M., Brent, R., Kingston, R.E., Mooew, D.D., Seidman, J.G. and Smith, J.S. (1991). Current Protocols in Molecular Biology. Wiley, New York.

Beck, J.T., Zhao, S. and Wang, W.W. (1994): Cloning, sequencing, and structural analysis of the DNA encoding inosine monophosphate dehydrogenase (EC 1.1.1.205) from *Tritrichomonas foetus*. Exp Parasitol 78: 101-112.

Blackwell, J.R. and Horgan, R. (1991). A novel strategy for production of a highly expressed recombinant protein in an active form. FEBS Lett 295: 10-12.

Brinkmann, U., Matles, R.E. and Buekel, P. (1989). High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. Gene 85: 109-114.

Burton, N., Cavallini, B., Kanno, M., Moncollin, V. and Egly, J.M. (1991). Expression in *Escherichia coli*: purification and properties of the yeast general transcription factor TFIID. Prot Exp Purif 2: 432-441.

Calderone, T.L., Stevens, R.D. and Oas, T.G. (1996). High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. J Mol Biol 262: 407-412.

Carr, S.F., Papp, E., Wu, J.C. and Natsumeda, Y. (1993). Characterization of human type I and II IMP dehydrogenase. J Biol Chem 268: 27286-27290.

Chen, G.F. and Inouye, M. (1990). Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. Nucl Acids Res 18: 1465-1473.

Fischer, L., Gerard, M., Chalut, C., Lutz, Y., Humbert, S., Kanno, M., Chambon, P. and Egly, J.M. (1992). Cloning of the 62-kilodalton component of basic transcription factor BTF2. Science 257: 1392-1395.

Hager, P.W., Collart, F.R., Huberman, E. and Mitchell, B.S. (1995). Recombinant human inosine monophosphate dehydrogenase type I and type II proteins. Purification and characterization of inhibitor binding. Biochem Pharmacol 49: 1323-1329.

Hannig, G. and Makrides, S.C. (1998). Strategies for optimizing heterologous protein expression in *Escherichia coli*. Trends in Biotech 16: 54-60.

Holmgren, A. (1985). Thioredoxin. Ann Rev Biochem 54: 237-271.

Hua, Z., Wang, H., Chen, D., Chen, Y. and Zhu, D. (1994). Enhancement of expression of human granulocyte-macrophage colony stimulating factor by argU gene product in *Escherichia coli*. Biochem Mol Biol Intl 32: 537-543.

Huete-Pérez, J.A., Wu, J.C., Whitby, F.G. and Wang, C.C. (1995). Identification of the IMP binding site in the IMP dehydrogenase from *Tritrichomonas foetus*. Biochemistry 34: 13889-13894.

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2: 13-34.

Konno, Y., Natsumeda, Y., Nagai, M., Yamaji, Y., Ohno, S., Suzuki, K. and Weber, G. (1991). Expression of human IMP dehydrogenase type I and II in *Escherichia coli* and distribution in human normal lymphocytes and leukemic cell lines. J Biol Chem 266: 506-509.

LaVallie, E.R., BiBlasio, E.A., Kovacic, S., Grant, K.L, Schendel, P.F. and McCoy, J.M. (1993). Bio/Technology 11: 187-193.

Lee, N., Zhang, S.-Q., Cozzitorto, J., Yang, J.-S. and Testa, D. (1987). Modification of mRNA secondary structure and alteration of the expression of human interferon alpha 1 in *Escherichia coli*. Gene 58: 77-86.

Looman, A.C., Bodlaender, J., De Gruyter, M., Vogelaar, A. and Van Knippenberg, P.H. (1986). Secondary structure as primary determinant of the efficiency of ribosomal binding sites in *Escherichia coli*. Nucleic Acids Res 14(13): 5481-97.

Margolis, N., Hogan, D., Tilly, K. and Rosa, P.A. (1994). Plasmid localization of *Borrelia* purine biosynthesis gene homologs. J Bacteriology 176: 6427-6432.

Marston, F.A.O. (1986). The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. Biochem J 240: 1-12.

Mukhopadhyay, A. (1997). Inclusion bodies and purification of proteins in biologically active forms. Adv Biochem Eng Biotechnol 56: 61-109.

Novy, R., Berg, J., Yaeger, K. and Mierendorf, R. (1995). pET TRX fusion system for increased solubility of protein expressed in *E. coli*. inNovations 3: 7-9.

Raina, S. and Missiakas, D. (1997). Making and breaking disulfide bonds. Ann Rev Microbiol 51:179-202.

Rietsch, A., Belin, D., Martin, N. and Beckwith, J. (1996). An *in vivo* pathway for disulfide bond isomerization in *Escherichia coli*. Proc Natl Acad Sci USA 93: 13048-13053.

Rudolph, R. and Lilie, H. (1996). *In vitro* folding of inclusion body proteins. FASEB J 10: 49-56.

Schein, C.H. and Noteborn, M.H.M. (1989). Bio/Technology 7: 1141-1148.

Schenk, P.M., Baumann, S., Matles, R. and Steinbiss, H. (1995). Improved high-level expression system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare Arg tRNAs. Biotechniques 19: 196-200.

Sone, M., Akiyama, Y. and Ito, K. (1997). Roles of disulfide bonds in bacterial alkaline phosphatase. J Biol Chem 272: 10349-10352.

Sorensen, M.A., Kurland, C.G. and Pedersen, S. (1989). Codon usage determines translation rate in *Escherichia coli*. J Mol Biol 207: 365-377.

Stewart, E.J., Aslund, F. and Beckwith, J. (1998). Disulfide bond formation in the *Escherichia coli* cytoplasm: an *in vivo* role reversal for the thioredoxins. EMBO J 17: 5543-5550.

Tessier, L.-H., Sondermeyer, P., Faure, T., Dreyer, D., Benavente, A., Villeval, D., Courtney, M. and Lecocq, J.P. (1984). The influence of mRNA primary and secondary structure on human IFN-gamma gene expression in *E. coli*. Nucleic Acids Res 12: 7663-7675.

Tobias, J.W., Shader, T.E., Rocap, G. and Varchavsky, A. (1991). The N-end rule in bacteria. Science 254: 1374-1377.

Yang, D. (1997). Purification and characterization of inosine-5'-monophosphate dehydrogenase (IMPDH) from nitrogen-fixing nodules of soybean (*Glycine max*). Master's thesis, the University of Oklahoma, Norman, OK, USA.

Zahn, K. (1996). Overexpression of a mRNA dependent on rare codons inhibits protein synthesis and cell growth. J Bacteriol 178: 2926-2933.

Zhang, S., Zubay, G. and Goldman, E. (1991). Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. Gene 105: 61-72.

Zhou, X., Cahoon, M., Rosa, P. and Hedstrom, L. (1997). Expression, purification, and characterization of inosine 5'-monophosphate dehydrogenase from *Borrelia burgdorferi*. J Biol Chem 272: 21977-21981.

Chapter 6. Differential Expression of Soybean Nodule *IMPDH* cDNA

INTRODUCTION

The formation of root nodules involves the expression of a set of nodule specific genes called nodulin genes (van Kammen, 1984). Nodulin genes that are markedly expressed before the onset of nitrogen fixation are referred to as early nodulin genes (Franssen *et al.*, 1992). About 50% of the identified early nodulin genes encode proline-rich proteins that are probably involved in plant cell wall biosynthesis (Schröder *et al.*, 1997). The nodulin genes expressed shortly before or concomitantly with the start of N_2 fixation are the late nodulin genes (Govers *et al.*, 1987; Nap & Bisseling, 1989). Late nodulins are mainly involved in the metabolic exchange between the host plant and the microsymbiont (Schröder *et al.*, 1997). Research on nodulins started with the discovery of the soybean nodulin N-35 in 1979 (Legocki and Verma, 1979). Since then, many nodulins from different legumes have been identified (Delauney and Verma, 1988; Govers and Bisseling, 1992; Verma and Miao, 1992). Leghaemoglobin and uricase II (nodulin-35) are two of the most abundant nodulins present in nodules. Therefore, these two proteins have been extensively studied.

Nodulin genes are expressed in a time and tissue-specific manner. Studies of the expression of several genes encoding enzymes involved in purine biosynthesis and ureide biogenesis have been reported. The expression of the soybean GAR synthetase and GAR transformylase genes (Schnorr *et al.*, 1996) and the cowpea AIR synthetase gene (Smith *et al.*, 1998) was nodule-specific. The PRAT gene from mothbean (Kim *et al.*, 1995), the GAR synthetase and GAR transformylase genes from soybean (Schnorr *et al.*, 1996), and the AIR synthetase

gene from cowpea (Smith *et al.*, 1998) were detected in 10 to 12-day-old nodules before nitrogenase was active. The level of SAICAR synthetase mRNA was detectable in 19-day-old nodules suggesting that this gene was induced late in nodule development following the commencement of nitrogen fixation (Chapman *et al.*, 1994).

In this chapter, results of studies on the spatial and temporal expression of the IMPDH gene are presented. These studies on the transcriptional and translational control of gene expression will address the following questions: 1) Is the expression of this gene nodule-specific, namely, is soybean nodule IMPDH a nodulin? 2) Is the expression of this gene induced during nodule development? 3) What is the pattern of *IMPDH* gene expression during nodule development?

MATERIALS AND METHODS

Plant materials

Soybean (*Glycine max* L. cv. Essex) seeds were inoculated and grown in the greenhouse as described in Chapter 2. Nodules were harvested 8, 13, 17, 21, 26, 33 and 36 days after inoculation. Nodules were frozen in liquid nitrogen and stored at –80°C. Soybean roots, epicotyls, stems, buds, leaves, flowers, young pods without seeds (2-3 weeks after flowering) and immature seeds (3 weeks after flowering) were also harvested and stored under similar conditions.

Transcriptional expression of the soybean nodule IMPDH gene

Total RNA was isolated from soybean nodules at different developmental stages and from different soybean tissues using the RNeasy RNA Isolation kit (Qiagen). A total of 0.1 g of tissue was used for each isolation. Isolated RNA was stored in diethylpyrocarbonate (DEPC)-H_2O at -80°C. The levels of *IMPDH* mRNA present in the tissue were determined using Northern blots. For each sample, 8 μg total RNA was loaded onto a 1.2 % agarose-formaldehyde gel. The concentration of RNA was determined spectrophotometrically at 260 nm. The amount of RNA loaded on each lane was determined by the RNA concentration and by the visualization with ethidium bromide. After electrophoretic separation, the RNA was transferred to a nitrocellulose membrane and dried under vacuum at 90°C for 2 hours. The Northern blot was probed at 42°C with a α–^{32}P-dCTP labeled *IMPDH* cDNA. After hybridization, membranes were washed two times with 1X SSC + 0.1% SDS at 45°C, followed by two washes with 0.1X SSC + 0.1% SDS at room temperature.

For tissue-specific expression studies, 8 μg of total RNA from different soybean tissues were loaded and run on a 1.2% agarose-formaldehyde gel. The tissues included roots, epicotyls, buds, stems, leaves, 30-day-old nodules, flowers, pods and seeds. The procedure for the Northern blot and other procedures were carried out in the same way as described for the developmental expression studies.

A soybean nodule glutamine synthetase (GS) cDNA (pGSGmD) was used as a positive control. The pGSGmD was kindly provided by Dr. Champa

Sengupta-Gopalan (Plant Genetic Engineering Laboratory, New Mexico State University, Las Cruces, NM). Membranes hybridized with the IMPDH clone were boiled in 0.1X SSC + 0.1% SDS for more than 1 h until no radioactivity was detected. Stripped membranes were probed with ^{32}P-labled pGSGmD. According to Roche *et al.* (1993), this GS clone was constitutively expressed in all soybean tissues.

Translational expression of IMPDH mRNA

Crude extracts of tissues were prepared according to the following procedures: A total of 1.0 g of frozen plant tissue was ground in an ice-cold mortar with liquid nitrogen and 0.1 g insoluble PVP (polyvinyl polypyrrolidone, Sigma P6755). Four ml of buffer (40 mM Tris-HCl, pH 7.2; 1 mM DTT; 10% (v/v) glycerol) was added to the tissue and the mixture was ground again. After grinding, the extract was transferred to a centrifuge tube and centrifuged at 30,000 x g for 30 min at 4°C. The supernatant fluid from the each crude protein extract was retained and kept at 4°C. A sample from each extract was used for SDS-PAGE analysis and for the measurement of IMPDH activity.

The total volume of each crude extract was measured with a serological pipet. For each sample, a volume (approximately 20 µl of crude extract) corresponding to the same mass (5.8 mg) was taken from the crude extract, dried, resupended with an equal volume of 1X SDS-PAGE gel loading buffer, heated in boiling water bath for 3 min and loaded onto a 12.5% SDS-PAGE gel. A duplicate gel was transferred electrophoretically onto a nitrocellulose membrane.

IMPDH was detected on Western blots using the following procedure (Promega protocol): The nitrocellulose membrane was saturated with 5% skim milk (in 1X TBST) for 30 min, and then incubated with the primary antibody (anti-IMPDH fusion protein, diluted 1:10,000 (v/v) with 1X TBST) for 60 min. The unbound antibody was removed by washing the membrane with 1X TBST three times for 10-15 min each time. The secondary antibody (anti-rabbit IgG alkaline phosphate (AP) conjugate, Sigma, Cat. No. A-3937) was diluted 1:10,000 with 1X TBST and then incubated with the membrane for 30 min. After three washes with 1X TBST, the membrane was briefly rinsed with 1X TBS twice to remove Tween 20. The secondary antibody was detected by incubating the membrane in 132 μl NBT (nitroblue tetrazolium, 25 mg/ml), 66 μl BCIP (5-bromo-4-chloro-3-indolyl phosphate, 50 mg/ml), and 20 ml alkaline phosphatase buffer (100 mM Tris-Cl, pH 9.5; 100 mM NaCl; 5 mM $MgCl_2$).

IMPDH activity assay

IMPDH enzyme activity was measured by continuously monitoring the IMP-dependent reduction of NAD^+ at 340 nm, as described previously in Chapter 5. The standard reaction mixture included: 50 mM Tricine-KOH (pH 8.8), 1 mM DTT, 2 mM NAD^+, 1.35 mM IMP, 50 μl crude extract. The final reaction volume was 1 ml. Nanopure H_2O was used to replace IMP in the control reaction mixture. The net IMPDH activity was obtained by substracting the change in absorbance for the control from the change in absorbance for the sample. For

each sample, the IMPDH activity assay was repeated at least four times. The data were averaged to obtain the final activity.

RESULTS

Northern blot analysis

To check the tissue specificity of *IMPDH* cDNA, Northern blot analysis was performed on RNA from different tissues using *IMPDH* cDNA as a probe. Total RNA from uninoculated roots, epicotyls, buds, stems, leaves, nodules (30-day-old), flowers, pods without seeds (2-3 wk after flowering), and immature seeds (3 wk after flowering) were isolated and separated by agarose gel electrophoresis. The RNA was transferred onto nitrocellulose membranes and hybridized with labeled probes. As seen in Figure 6-1, *IMPDH* hybridized only to RNA from nodules and no signals were detected in RNA obtained from other tissues. The positive control (pGSGmD) hybridized to RNA from all tissues except seeds. Also, the signal using pGSGmD as a probe was much stronger in roots and nodules than in other tissues.

To understand *IMPDH* gene expression during nodule development, Northern blot analysis was performed on RNA from soybean nodules at different developmental stages. Total RNA was isolated from nodules at different stages during development and the developmental pattern of expression was analyzed on a Northern blot analysis using the *IMPDH* cDNA as a probe. The same membranes were stripped and probed with a pGSGmD cDNA. The pGSGmD

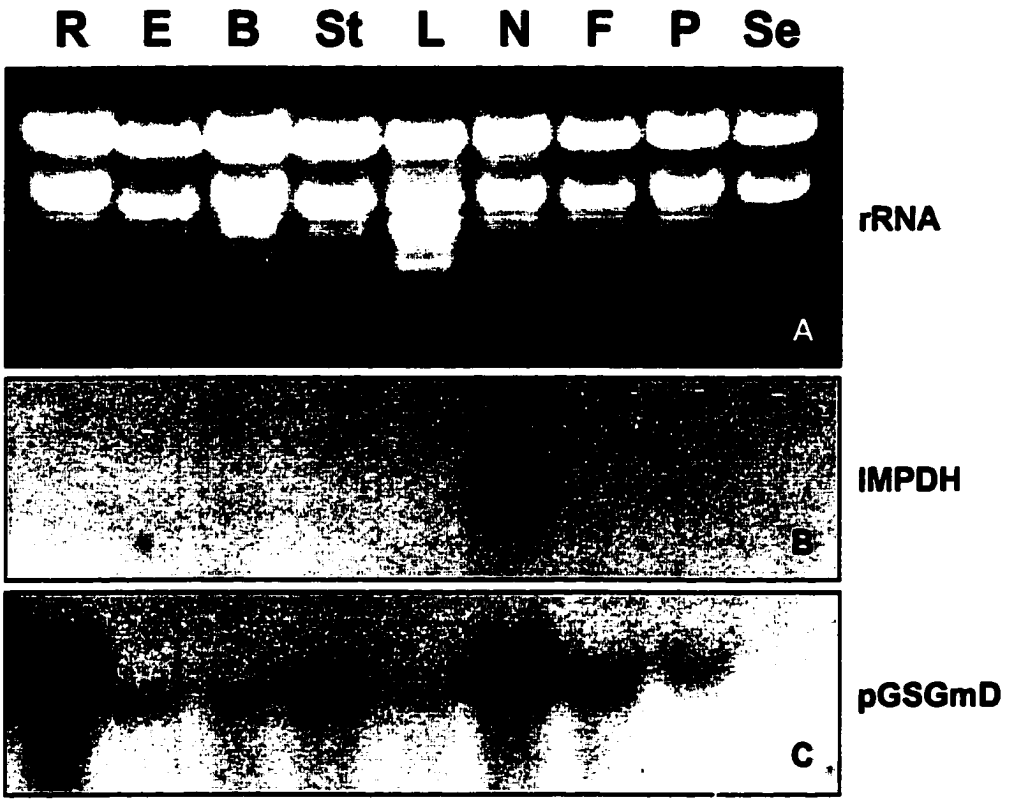


Figure 6-1. Expression of the *IMPDH* mRNA in different soybean tissues. Total RNA (8 μg) was loaded in each lane. rRNA was visualized by UV illumination as a control for the amount of RNA loaded (Panel **A**). Northern blot analysis using *IMPDH* cDNA as a probe (Panel **B**). Soybean nodule glutamine synthetase cDNA clone pGSGmD was used as a probe in the control (Panel **C**). *GSGmD* gene was reported to be constitutively expressed in soybean plants (Roche *et al.*, 1993). R-root; E-epicotyl; B-buds; St-stems; L-leaves; N-nodules; F-flowers; P-young pods without seeds; Se-immature seeds.

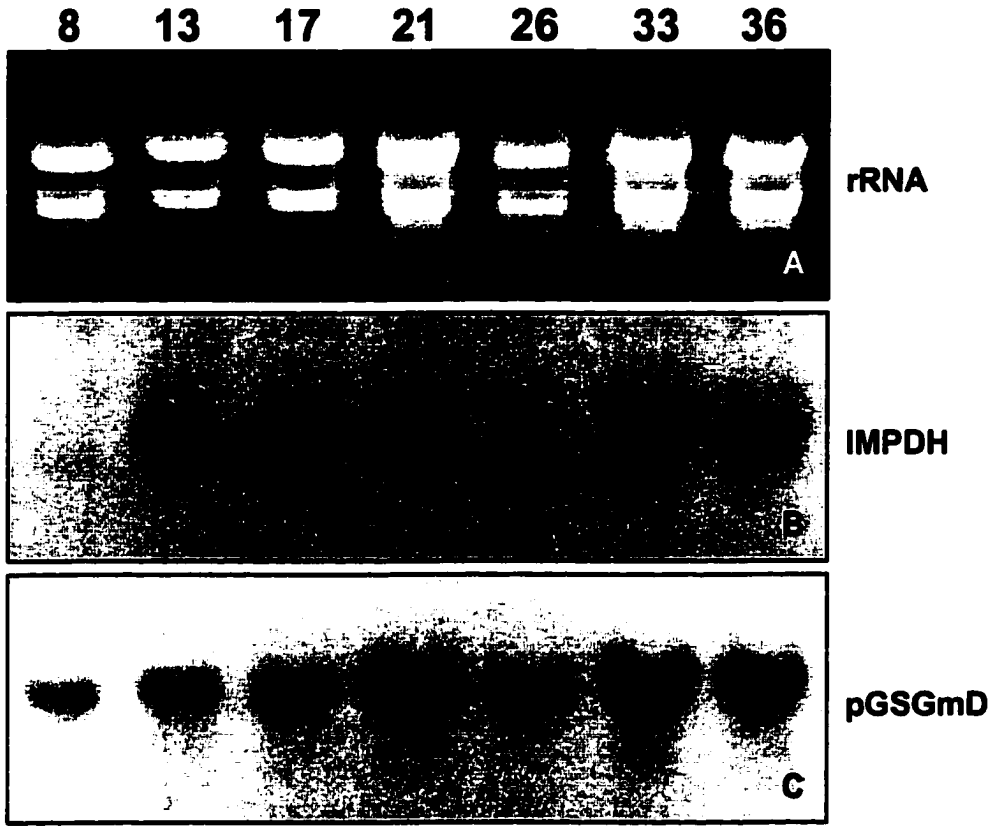


Figure 6-2. Temporal expression of *IMPDH* mRNA in developing soybean nodules. Total RNA (8 μg) was loaded in each lane. rRNA was visualized by UV illumination as a control for the amount of RNA loaded (Panel **A**). Northern blot analysis using *IMPDH* cDNA as a probe (Panel **B**). Soybean nodule glutamine synthetase cDNA clone pGSGmD was used as a probe in the control (Panel **C**). Nodules were harvested 8, 13, 17, 21, 26, 33 and 36 days after inoculation.

expression was constitutive but dramatically enhanced in nodules following the onset of N_2 fixation (Roche *et al.*, 1993). *IMPDH* mRNA was not detected in 8-day-old nodules (Figure 6-2). *IMPDH* mRNA was first detected 13 days after inoculation (DAI). The expression level appeared to increase from 13 to 21 DAI, and showed no change from 21 to 36 days except the signal in 26-day-old nodules was very low (Figure 6-2B). The hybridization signal using the pGSGmD probe was detected in nodules of all ages tested, and the signal appeared to become stronger from 8 to 21 DAI and remained constant from 21 to 36 DAI (Figure 6-2C)

IMPDH activity assay and Western blot analysis

The IMPDH activity assay and Western blot analysis were performed to examine the soybean nodule *IMPDH* gene expression at the translational level. Results showed that nodules exhibited the highest IMPDH activity (10.75 μmol/min/mg protein; 0.013 μmol/min/g nodule). IMPDH activity in stem was approximately half of that in nodules, while all other tissues had little or no activity (Figure 6-3). Total dehydrogenase activity was very high in seeds, but no IMPDH activity was detected. Samples of crude extracts were analyzed by SDS-PAGE. Gels were transferred for Western blots (Figure 6-4A). Two strong immunoreactive bands, one at ~52 kDa and one at ~48 kDa, were detected in nodules (Figure 6-4B). The 52-kDa band was also visible in immature seeds. This band was extremely weak in epitocyls, buds, stems, flowers and pods, but was not visible at all in roots and

leaves. The 48-kDa band was not detectable in soybean tissues other than nodules.

The developmental pattern of *IMPDH* gene expression was analyzed by enzyme activity assays and immunodetection on Western blots. IMPDH activity was barely detectable in 8-day-old nodules. The IMPDH activity generally increased after day 13 (Figure 6-5). In Western blot analysis, the 52-kDa band was barely visible in 8-day-old nodules and the 48-kDa band was absent. During the course of nodule development, both the 52-kDa and 48-kDa bands were detected (Figure 6-6). These bands seemed to become stronger from 13 to 21 DAI and the intensity of staining remained fairly constant thereafter.

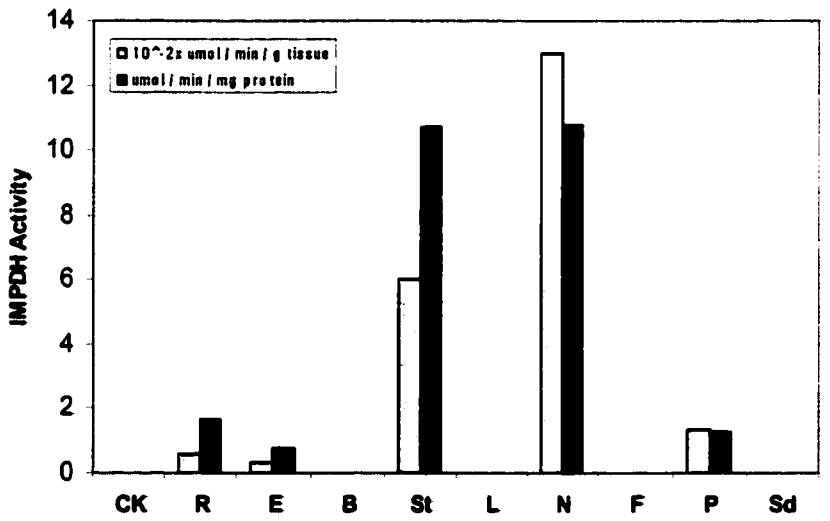


Figure 6-3. IMPDH activity in different soybean tissues. Activity was measured using the standard IMPDH assay. Crude extracts were prepared from soybean roots (R), epicotyls (E), buds (B), stems (St), leaves (L), 30-day-old nodules (N), flowers (F), pods without seeds (P) and immature seeds (Sd). Crude extract representing 29 mg of each tissue fresh weight (~100 µl) was used in each assay. Control (CK) reaction was performed without addition of protein samples.

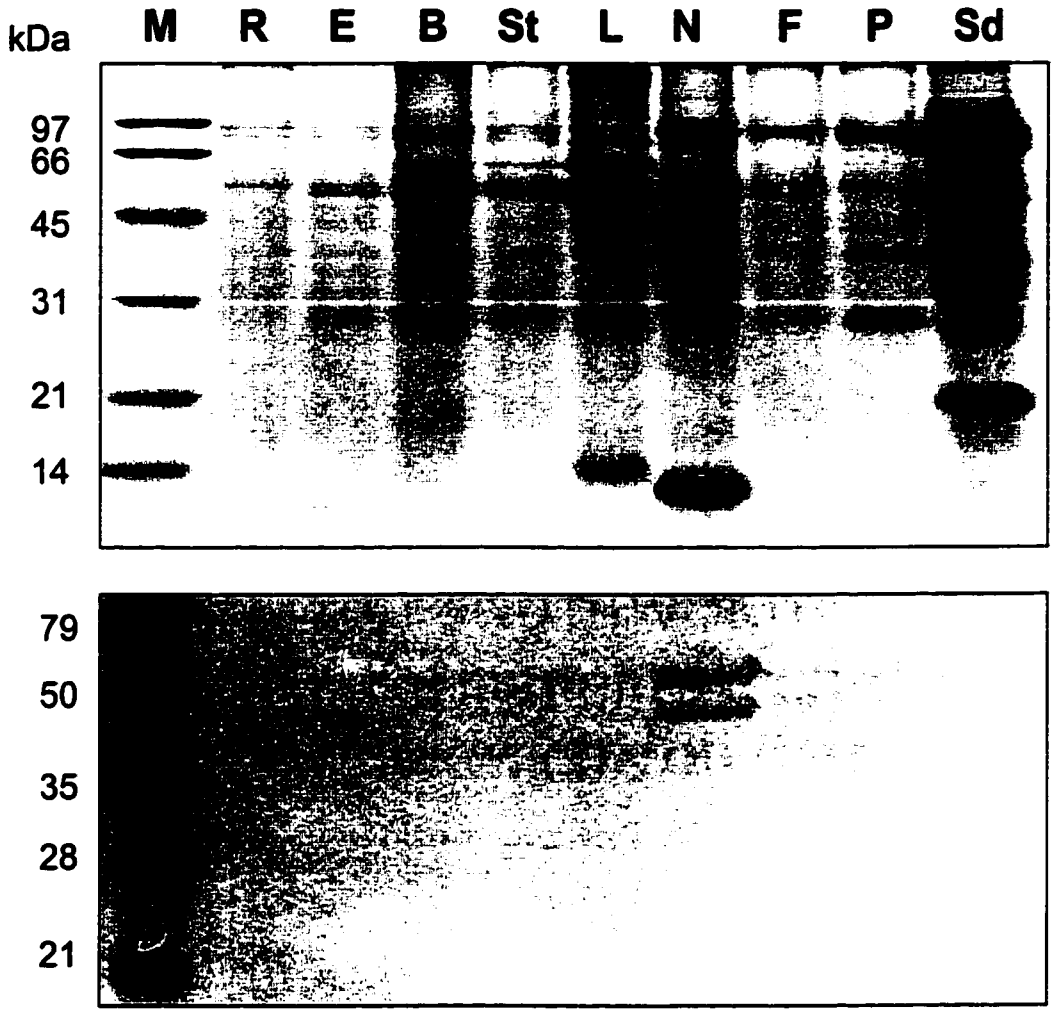


Figure 6-4. Immunodetection of IMPDH in different soybean tissues. Crude extracts were prepared from roots (R), epicotyls (E), buds (B), stems (St), leaves (L), nodules (N), flowers (F), pods without seeds (P) and immature seeds (Sd). Crude extract equivalent to 5.8 mg of frozen tissue was loaded onto a 12.5% SDS-PAGE gel. Proteins were transferred to nitrocellulose membranes and anti-IMPDH fusion protein antibody (third bleeding, 1:5,000 dilution) was used to detect IMPDH on the Western blots. Panel **(A).** SDS-PAGE gel of crude extract from different tissues. Gel was stained with Coomassie Brilliant Blue R-250. Panel **(B).** Western blot: a duplicate of the SDS-PAGE gel was transferred to a nitrocellulose membrane and IMPDH was detected with rabbit anti-IMPDH antibody. Pre-stained protein standard was used for Western blot.

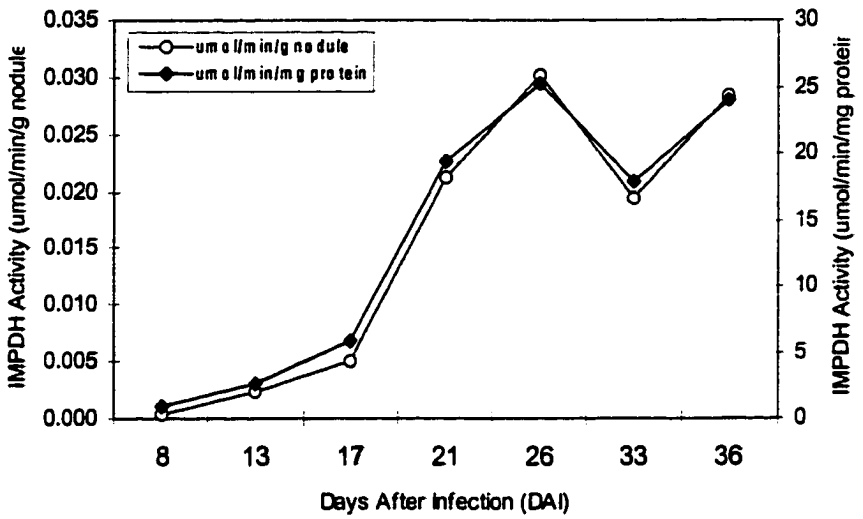


Figure 6-5. IMPDH activity in crude extracts of soybean nodules during development. Activity was measured using the standard IMPDH assay. Crude extracts were prepared from soybean nodules 8, 13, 17, 21, 26, 33 and 36 days after infection. Crude extract of equivalent to 26 mg (~100 μl) of each tissue was used in the assay.

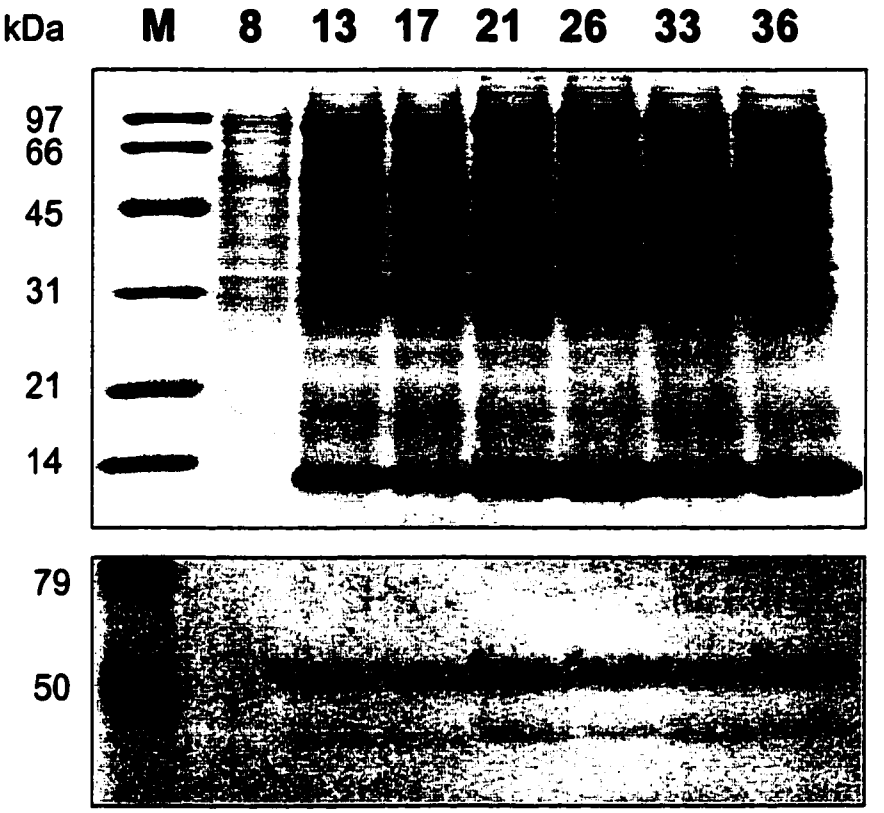


Figure 6-6. Changes in IMPDH during soybean nodule development. Crude extracts were prepared from 8, 13,17, 21, 26, 33 and 36-day-old nodules. A crude extract equivalent to 5.2 mg of nodule frozen weight was loaded in each lane of SDS-PAGE. Antibody against IMPDH fusion protein (fourth bleeding, 1:10,000 dilution) was used to detect IMPDH in Western blot analysis. **(A).** SDS-PAGE gel of crude extracts from nodules of different ages. Gel was stained with Coomassie Brilliant Blue R-250 to detect protein. **(B).** Western blot of a duplicate gel shown in panel (A) detected with IMPDH antibody. Pre-stained protein standard was used for Western blot.

DISCUSSION

Tissue specificity of soybean nodule IMPDH gene expression

Based on results of Northern blots, the *IMPDH* gene was expressed exclusively in nodules. The transcripts were not detected in any of the other tissues examined (Figure 6-1). Nodules also exhibited the highest IMPDH activity (Figure 6-3). Both immunoreactive IMPDH bands were present only in nodules (Figure 6-4). According to the nomenclature for plant genes involved in nodulation suggested by van Kammen (1984), plant genes that are exclusively or predominantly expressed during nodule formation and function are called nodulin genes, and the corresponding gene products are termed nodulins. On this basis, it appears that soybean nodule IMPDH is a nodulin.

The oxidation of IMP to XMP catalyzed by IMPDH is proposed to be the first step of purine catabolism leading to ureide biogenesis. Consistent with this function, soybean nodule IMPDH is apparently a late nodulin. Generally, late nodulin genes are expressed shortly before or concomitantly with the start of N_2 fixation and are mainly involved in the metabolic exchange between plant and microbiont (Schröder *et al.*, 1997). Late nodulins include enzymes involved in nitrogen assimilation and carbon metabolism in amide and ureide biogenesis. Leghaemoglobins and proteins present in the peribacteroid membrane also belong to this group (Sánchez *et al.*, 1991).

Induction of IMPDH *gene expression in nodules*

In soybean nodules, *IMPDH* transcripts, IMPDH activity and IMPDH protein were detected in 13-day-old nodules but not in 8-day-old nodules (Figure 6-2, 6-5, 6-6). These results suggested that there was an induction of *IMPDH* gene expression between 8 and 13 DAI. Nitrogenase activity was reported to start 11-12 days after infection (Schubert, 1981; Reynolds *et al.*, 1982; Sengupta-Gopalan and Pitas, 1986). Thus, it seems that the induction of *IMPDH* expression coincides with the onset of nitrogen fixation. Whether the induction is independent of nitrogen fixation needs to be addressed. One approach would be to examine the IMPDH in Fix^- (not capable of nitrogen fixation) nodules. The identification of other factors regulating *IMPDH* gene expression and induction needs further investigation.

Based on results from Northern blots, enzyme assays and Western blots, the expression of the *IMPDH* gene increased during the first 21 days after infection and remained generally constant during the period from 21 to 36 DAI (Figure 6-2, 6-5, 6-6). In Northern blots, the apparently lower amount of total RNA loaded on the gel may account for the lower level of *IMPDH* mRNA (and GSGmD mRNA) detected in 26-day-old nodules.

Existence of IMPDH subunits and isoforms in soybean

As discussed earlier in this chapter, soybean nodule IMPDH is a nodulin and the *IMPDH* gene is exclusively or predominantly expressed in nodules. The expression of the *IMPDH* gene starts before the onset or concomitantly with the

beginning of nitrogen fixation activity. The large amounts of nitrogen reduced in nodules need to be promptly assimilated and exported to the aboveground organs of the plant. This is accomplished through the reactions of ammonium assimilation, purine biosynthesis and ureide biogenesis. The induction of *IMPDH* gene is consistent with the increased needs for enzyme activities involved in these pathways.

Besides its importance in ureide biosynthesis, the IMPDH catalyzed oxidation of IMP to XMP is also essential for other basic physiological activities such as GMP biosynthesis for DNA and RNA biosynthesis and signal transduction (Huberman *et al.*, 1995). One way to carry out the multiple functions of an enzyme is to have isoforms for each function. It is possible that a house-keeping IMPDH isoform exists. This IMPDH is most likely expressed in all tissues of soybean plants, and functions in basic metabolic activities.

Two different immunoreactive forms of IMPDH were detected in soybean nodules (Figure 6-4, 6-6). The 48-kDa band was not detected in any of the other 8 tissues examined and appeared to be nodule specific. The 52-kDa band was very strong in nodules. This band was barely visible in eipcotyls, buds, stems, flowers, pods and immature seeds and was not detected at all in roots and leaves. A reasonable preliminary conclusion from these results is that the 48-kDa and 52-kDa bands represent two different forms of soybean nodule IMPDH.

The predicted molecular weight of soybean nodule IMPDH was 53 kDa (Chapter 3). The 52-kDa subunit detected in Western blots may represent the direct gene product of the single *IMPDH* gene. The 48-kDa subunit could result

from alternative gene splicing, alternative transcription initiation or alternative translation initiation (Small *et al.*, 1998; Smith *et al.*, 1998), since there is only one copy of the *IMPDH* gene in the soybean genome (Chapter 3). Based on the *IMPDH* cDNA sequence, the next methionine (Met) codon downstream from the proposed start codon (at position 167) is 63 amino acids away at position 356 (Figure 2-6). If this codon is used as an alternative translation start site, a protein with a molecular weight around 46 kDa will be generated. This size is almost the same as the 48-kDa band detected on Western blots. Also, the eukaryotic IMPDH genes tend to have multiple introns (Zimmermann *et al.*, 1995; Collart *et al.*, 1996). Because *Arabidopsis thaliana* IMPDH, which is 70.5% identical with the deduced amino acid sequence of soybean nodule IMPDH, has four introns and five exons in its gene (Collart *et al.*, 1996) and the human IMPDH type II gene contains 14 exons (Zimmermann *et al.*, 1995), the soybean IMPDH gene could also contain introns. Alternative transcript splicing could generate a different form of IMPDH (Gebhardt *et al.*, 1998; Small *et al.*, 1998; Smith *et al.*, 1998) from the same gene.

IMPDH gene regulation

The regulation of nodulin gene expression is a complicated but interesting subject. The induction of early nodulin genes does not require internalized bacteria or infection threads. It is stimulated by rhizobial Nod factors, the plant hormone cytokinin or synthetic lipo-chitin molecules (Mianmi *et al.*, 1996; van de Sande *et al.*, 1996, 1997; Chen *et al.*, 1998). In contrast, induction of late nodulin gene expression requires the formation of infection threads and the release of bacteria

into the infected cells (Verma and Delaunery, 1988; Verma and Miao, 1992). Some late nodulins, e.g. uricase, require nitrogen fixation for induction (Nguyen *et al.*, 1985).

Characterized as a nodulin gene, soybean nodule *IMPDH* is exclusively expressed or the expression is greatly enhanced in soybean nodules. The expression was induced between 8-13 days after infection. Because nitrogen fixation begins 11-12 DAI, it was not possible to determine if the expression of IMPDH was induced by products of nitrogen fixation. To answer this question, expression studies with soybean nodules lacking the ability to fix nitrogen (Fix^-) are needed. However, expression of most of the late nodulins is independent of nitrogen fixation (Verma and Delaunery, 1988).

The expression of nodulin genes is mainly controlled at the transcriptional level (Verma and Miao, 1992). The *cis*-regulatory sequence located in the 5' promoter region of various nodulin genes determines organ specificity and the level of expression in nodules by interacting with corresponding *trans*-acting factors. A conserved motif (the "nodulin box") was postulated to be important for nodule-specific expression of some late nodulin genes (Mauro *et al.*, 1985; Mauro and Verma, 1988; Ramlove *et al.*, 1992; Takane *et al.*, 1997). To further study the regulation of *IMPDH* gene expression will require the isolation of a genomic IMPDH clone.

REFERENCES

Andersson, C.R., Jensen, E.O., Llewellyn, D.J., Dennis, E.S and Peacock, W.J. (1996). A new hemoglobin gene from soybean: a role for hemoglobin in all plants. Proc Natl Acad Sci USA 93: 5682-5687.

Atkins, C.A., Shelp, B.J. and Storer, P.J. (1985). Purification and properties of inosine monophosphate oxidoreductase from nitrogen-fixing nodules of cowpea (*vigna unguiculasta* L. Walp). Arch Biochem Biophys 236: 807-814.

Capote-Maínez, N. and Sánchez, F. (1997). Characterization of the common bean uricase II and its expression in organs other than nodules. Plant Physiol 115: 1307-1317.

Chapman, K.A., Delauney, A.J., Kim, J.H. and Verma, D.P.S. (1994). Structural organization of *de novo* purine biosynthesis enzymes in plants: 5-aminoimidazole ribonucleotide carboxylase and 5-aminoimidazole-4-*N*-succinocarboxanmide ribonucleotide synthetase cDNAs from *Vigna aconitifolia*. Plant Mol Biol 24: 389-395.

Chen, R., Silver, D.J. and de Bruijn, F.J. (1998). Nodule parenchyma-specific expression of the *Sesbania rostrata* early nodulin gene *SrEnod2* is mediated by its 3' untranslated region. Plant Cell 10:1585-1602.

Collart, F.R., Osipiuk J., Trent, J., Olsen,G. and Huberman, E. (1996). Cloning and characterization of the gene encoding IMP dehydrogenase from *Arabidopsis thaliana*. Gene 174: 217-220.

Delaunery, A. and Verma, D.P.S. (1988). Cloned nodulin genes for symbiotic nitrogen fixation. Plant Mol Biol Rep 6: 279-285.

Franssen, A.J., Vijn, I., Yang, W.C. and Bisseling, T. (1992). Developmental aspects of the *rhizobium*-legume symbiosis. Plant Mol Biol 19: 89-107.

Gebhardt, J.S., Wadsworth, G.J. and Matthews, B.F. (1998). Characterization of a single soybean cDNA encoding cytosolic and glyoxysomal isozymes of aspartate aminotransferase. Plant Mol Biol 37: 99-108.

Govers, F. and Bisseling, T. (1992). Nodulins in root development: function and gene regulation. In: Nitrogen metabolism of plants: proceeding of the phytochemical society of Europe 33 (Menge, K. and Pilbeam, D.J. Ed.). Oxford Science Publications, New York, pp30-37.

Govers, F., Nap, J.P., van Kammen, A. and Bisseling T. (1987). Nodulins in the developing root nodules. Plant Physiol Biochem 25: 309-322.

Huberman, E., Glesne, D. and Collart, F.R. (1995). Regulation and role of inosine-5'-monophosphate dehydrogenase in cell replication, malignant rransformation, and differentiation. In: Purine and pyrimidine metabolism in man, VIII. (Sahota, A. and Taylor, M. Ed.). Plenum Press, New York, pp741-746.

Kim, H.J., Delaunery, A.J. and Verma, D.P.S. (1995). Control of *de novo* purine biosynthesis genes in ureide-producing legumes: induction of glutamine phosphoribosylphosphate amidotransferase gene and characterization of its cDNA from soybean and *Vigna*. The Plant J 7: 77-86.

Legocki, R.P. and Verma, D.P.S. (1979). A nodule specific plant protein (nodulin-35) from soybean. Science 205: 190-193.

Mauro, V.P., Nguyen, T., Katinakis, P. and Verma, D.P.S. (1985). Primary structure of the soybean nodulin-23 gene and potential regulatory elements in the 5'-flanking regions of nodulin and leghemoglobin genes. Nucleic Acids Res 13: 239-249.

Mauro, V.P. and Verma, D.P.S. (1988). Transcriptional activation in nuclei from uninfected soybean of a set of genes involved in symbiosis with *Rhizobium*. Mol Plant-Microbe Interact 1: 46-51.

Minami, E., Kouchi, H., Carlson, R.W., Cohn, J.R., Kolli, V.K., Day, R.B., Ogawa, T. and Stacey, G. (1996). Cooperative action of lipo-chitin nodulation signals on the induction of the early nodulin, ENOD2, in soybean roots. Mol Plant-Microbe Interact 7: 574-583.

Nap, J.P. and Bisseling, T. (1989). Nodulin function and nodulin gene regulation in root nodule development. In: The molecular biology of symbiotic nitrogen fixation (Gresshoff, P.M. Ed.). CRC Press, Boca Raton, FL, pp181-229.

Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H. and Verma, D.P.S. (1985). Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisome of uninfected cells of nodules. Proc Nat Acad Sci USA 82: 5040-5044.

Ramlov, K.B., Laursen, N.B., Stougaard, J. and Marcker, K.A. (1993). Site-directed mutagenesis of the organ-specific element in the soybean leghemoglobin lbc3 gene promoter. Plant J 4: 577-580.

Reynolds, P.H.S., Boland, M.J., Blevins, D.G., Schubert, K.R. and Randall, D.D. (1982). Enzymes of amide and ureide biogenesis in developing soybean nodules. Plant Physiol 69: 1334-1338.

Roche, D., Temple, S.J. and Sengupta-Gopalan, C. (1993). Two classes of differentially regulated glutamine synthetase gene are expressed in the soybean nodule: a nodule-specific class and a constitutively expressed class. Plant Mol Biol 22: 971-983.

Sánchez, F., Padilla, J.E., Pérez, H. and Lara, M. (1991). Control of nodulin genes in root-nodule development and metabolism. Ann Rev Plant Physiol Plant Mol Biol 42: 507-528.

Schnorr, K.M., Laloue, M. and Hirel, B. (1996). Isolation of cDNAs encoding two purine biosynthetic enzymes of soybean and expression of the corresponding transcripts in roots and root nodules. Plant Mol Biol 32: 751-757.

Schröder, G., Fruhling, M., Puhler, A. and Perlick, A.M. (1997). The temporal and spatial transcription pattern in root nodules of *Vicia faba* nodulin genes encoding glycine-rich proteins. Plant Mol Biol 33:113-123.

Schubert, K.R. (1981). Enzymes of purine biosynthesis and catabolism in *Glycine max*. Plant Physiol 68: 1115-1122.

Sengupta-Gopalan, C. and Pitas, J.W. (1986). Expression of nodule-specific glutamine synthetase genes during nodule development in soybeans. Plant Mol Biol 7: 189-199.

Small, I, Wintz, H., Akashi, K. and Mireau, H. (1998). Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol Biol 38: 265-277.

Smith, P.M.C., Mann, A.J., Goggin, D.E. and Atkins, C.A. (1998). AIR synthetase in cowpea nodules: a single gene product targeted to two organelles?. Plant Mol Biol 36: 811-820.

Takane, K., Tanaka, T., Tajima, S., Okazaki, K. and Kouchi, H. (1997a). Expression of a gene from uricase II (nodulin-35) in cotyledons of soybean plants. Plant Cell Physiol 38:149-154.

Takane, K, Tajima, S. and Kouchi, H. (1997b). Two distinct uricase II (nodulin 35) genes are differentially expressed in soybean plants. Mol Plant-Microbe Interact 10: 735-741.

van de Sande, K., Pawlowski, K., Czaja, I., Wieneke, U., Schell, J., Schmidt, J., Walden, R., Matvienko, M., Wellink, J., Van Kammen, A., Franssen, H. and Bisseling, T. (1996). Modification of phytohormone response by a peptide encoded by *ENOD40* of legumes and a nonlegume. Science 273:370-373.

van de Sande, K. and Bisseling, T. (1997). Signaling in symbiotic root nodule formation. Essays in Biochem 32:127-142.

van Kammen, A. (1984). Suggested nomenclature for plant genes involved in nodulation and symbiosis. Plant Mol Biol Rep 2: 43-45.

Verma, D.P.S. and Delauney, A.J. (1988). Root nodule development: nodulins and nodulin genes. In: Plant gene research – Basic knowledge and application. Temporal and spatial regulation of plant genes (Verma, D.P.S. and Goldberg, R.B. Ed.). Springer-Verlag , New York, pp170-199.

Verma, D.P.S. and Miao, G.-H. (1992). Induction of nodulin genes and root nodule symbiosis. In: Inducible plant proteins – their biochemistry and molecular biology (Wray, J.L. Ed.). Cambridge University Press, New York, pp175-204.

Yang, D. (1997). Purification and characterization of inosine-5'-monophosphate dehydrogenase (IMPDH) from nitrogen-fixing nodules of soybean (*Glycine max*). Master's thesis, the University of Oklahoma, Norman, OK, USA.

Zimmermann, A.G., Spychala, J. and Mitchell, B.S. (1995). Characterization of the human inosine-5'-monophosphate dehydrogenase type II gene. J Biol Chem 270: 6808-6814.

Chapter 7. *In situ* Localization of *IMPDH* mRNA in Soybean Root Nodules

INTRODUCTION

Root nodules formed on the roots of leguminous plants are unique organs for symbiotic nitrogen fixation by *Rhizobium*. These organized structures develop from meristems newly formed in the cortex of the root as a result of interaction with rhizobia. The mature root nodule is made up of a central tissue, containing infected and uninfected cells, surrounded by several layers of peripheral cells. The peripheral tissues include the nodule cortex, the nodule endodermis and the nodule parenchyma ("inner cortex"), which contains the vascular bundles connecting the nodule with the root stele (Newcomb, 1981; Bergersen, 1982; Franssen *et al.*, 1992; Mylona *et al.*, 1995).

Two main categories of leguminous nodules can be recognized based on their morphology: determinate and indeterminate nodules (Franssen *et al.*, 1992). Legumes such as *Pisum* (pea), *Trifolium* (clover) and *Medicago* (alfalfa) develop indeterminate nodules, whereas determinate nodules are formed on the roots of *Glycine* (soybean), *Vigna* (cowpea) and *Phaseolus* (bean). Indeterminate nodules have a persistent meristem at the apex from which cells are continuously added to the cortical and central tissues. Consequently, all tissues within these nodules increase in age from the meristem to the root attachment point. In contrast, the meristem of a determinate nodule ceases to divide two weeks after inoculation with nitrogen-fixing bacteria and it differentiates completely into different nodular tissues (Newcomb, 1981; van de Wiel *et al.*, 1990). A diagram of the structure of a determinate nodule is presented in Figure 7-1.

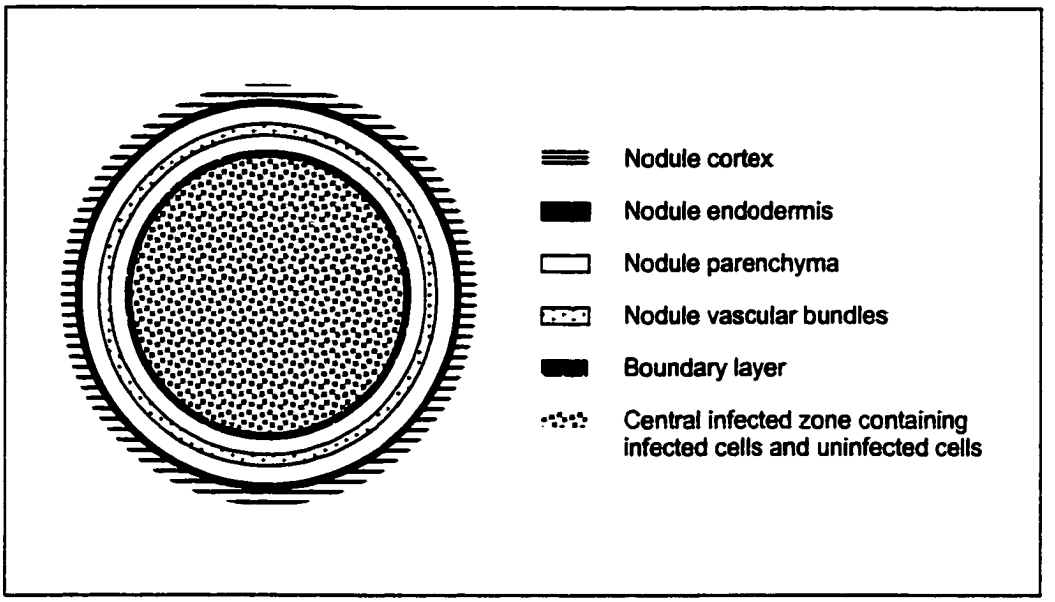


Figure 7-1. Diagrammatic representation of the structure of a determinate root nodule (transverse section) (Modified from Franssen *et al.*, 1992).

Using the technique of *in situ* hybridization, the combination of morphology and molecular biology can be employed to study gene expressions in nodules. Up to now, several nodulin transcripts have been localized in determinate or indeterminate nodules (van de Wiel *et al.*, 1990; de Billy *et al.*, 1991; Kouchi and Hata, 1993; Miao and Verma, 1993; Mativienko *et al.*, 1994; Temple *et al.*, 1995; Papadopoulou *et al.*, 1995; Shi *et al.*, 1997). For the early nodulin gene *ENOD2*, expression occurred only in the inner cortex of indeterminate pea nodules. Expression of *ENOD2* was detected in both the inner cortex and cells surrounding the vascular bundles in determinate soybean nodules (van de Wiel *et al.*, 1990). In soybean, transcripts of the early nodulin gene *GmN#36* first appeared in cells of the root pericycle near the infected cells (Kouchi and Hata, 1993). During nodule emergence, *GmN#36* transcripts were found in a few cell layers

surrounding the vascular strands connecting the nodule meristem with the root stele, and in mature nodules the transcripts were present exclusively in the pericyle cells in vascular bundles. On the other hand, *GmN#93* transcripts first appeared in the primary nodule meristem just beneath the root epidermis. In mature nodules, *GmN#93* transcripts were present only in the infected cells (Kouchi and Hata, 1993).

For late nodulin genes, alfalfa leghaemoglobin (Lb) transcripts were detected only in cells which have been invaded by the microsymbiont *Rhizobium meliloti* (de Billy *et al.*, 1991). Bean (*Phaseolus vulgaris*) uricase II (nodulin-35) transcripts were detected only in the uninfected cells of the central tissues and mainly in the periphery of the cell (Papadopoulou *et al.*, 1995). Nodulin-26, a peribacteroid membrane nodulin considered to be a channel protein, was expressed only in the infected cells (Miao and Verma, 1993).

Based on results presented in Chapter 6, soybean IMPDH is a nodulin specifically expressed in nodules. In this chapter, techniques of *in situ* hybridization have been used to localize the site of *IMPDH* gene expression in mature soybean root nodules.

MATERIALS AND METHODS

Compared to radiolabeld probes (usually ^{35}S) labeled probes, DIG-labeled probes have the advantage of being safer, easier to manipulate and they can be stored for more than a year after labeling. Therefore, the non-radioactive probe was used in

the current experiments. An outline of procedures for *in situ* hybridization is presented in Figure 7-2. The detailed protocols are described as below.

Plant tissue fixation, embedding and sectioning

Tissue fixation and embedding were carried out based on the procedures of Shi *et al.* (1997). Soybean plants were grown in the greenhouse as described in Chapter 2. Nodules from 53-day-old plants were harvested and cut into 2 x 2 x 2 mm blocks. The blocks were vacuum infiltrated and fixed in 4% paraformaldehyde and 0.25% glutaraldehyde in 50 mM sodium phosphate buffer, pH 7.2, at room temperature for 15 min. Further fixation was carried out at room temperature for 2 h or at 4°C overnight. Tissues were washed two times with 50 mM sodium phosphate buffer (pH 7.2), twice with water and then dehydrated through a gradient series of ethanol concentrations (10%, 20%, 30%, 40%, 50%, 60%, 70%, 90% and 100%) for 15 min each. After a 60 min incubation in 3:1, 1:1 and 1:3 ethanol / xylene (v/v), the tissues were placed in pure xylene for 60 min. A few Paraplast chips were added to the xylene solution and tissues were infiltrated overnight at room temperature. After the addition of a few more Paraplast chips, the tissues were incubated in a 42°C waterbath for 3 h. The blocks were then transferred to embedding molds containing molten wax in a 60°C oven and incubated for 24 hours at 62°C. Paraffin blocks containing the tissue were removed from the embedding molds and stored at room temperature.

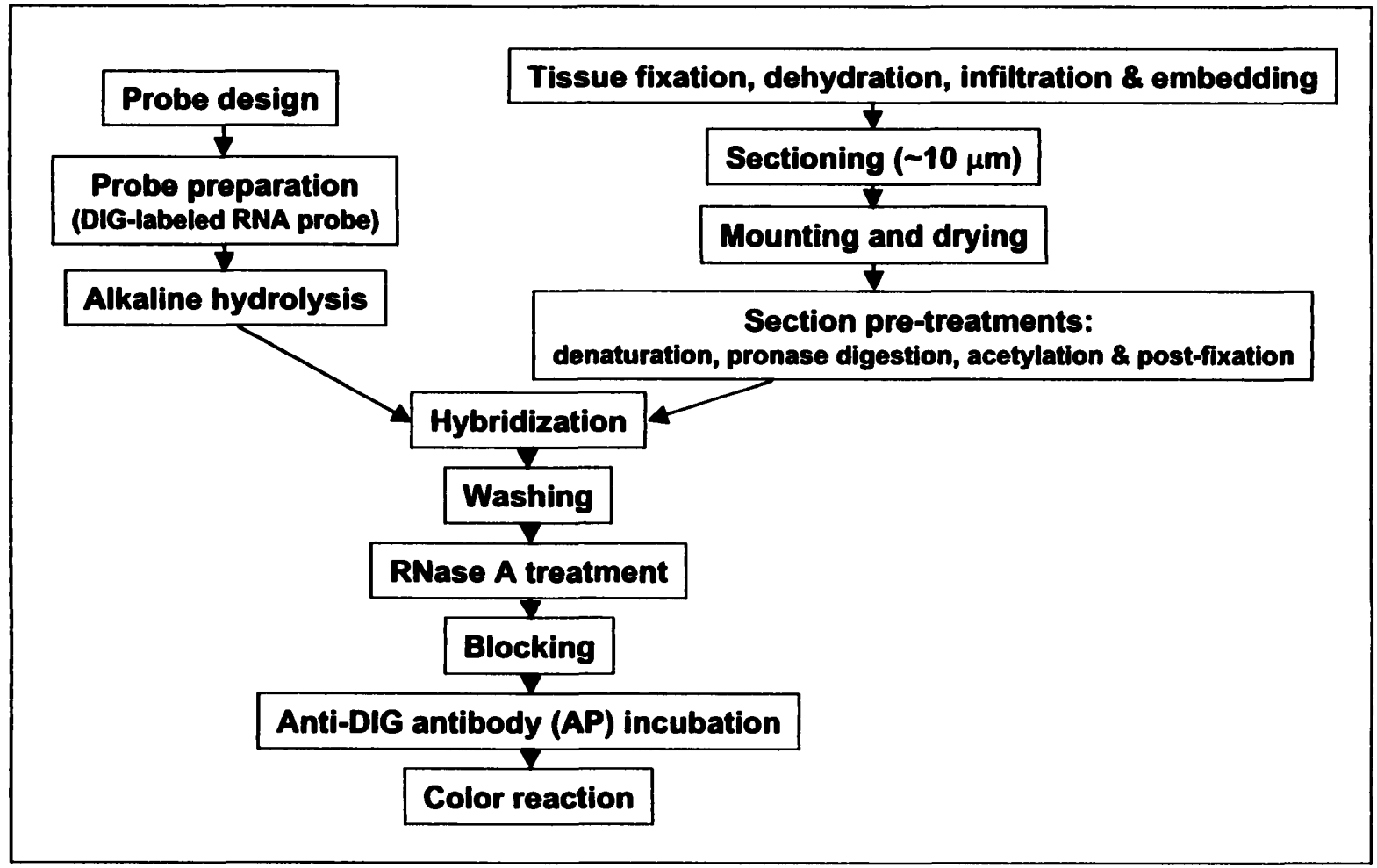


Figure 7-2. Procedure outline of *in situ* hybridization.

An AO Spencer "820" microtome equipped with a Leica 819 disposable microtome blade was used to prepare 7-10 μm thick sections. Sections were mounted on ProbeOn Plus slides (Fisher) and dried overnight in a 42°C oven.

Preparation of probes for in situ *hybridization*

DIG-labeled RNA probes were prepared by *in vitro* transcription. Plasmid DNA from pSBimpdh was linearized by restriction enzyme cleavage, cleaned and used as a DNA template for *in vitro* transcription (Figure 7-3). For the preparation of the sense probe (T3 direction), the DNA template was digested with the restriction enzymes *Pvu* II and *Xho* I, while for the preparation of antisense probe (T7 direction), *Pvu* II and *Spe* I were used. The products of *in vitro* transcription were separated on a 1.2 % agarose gel, and the bands corresponding to template DNA were excised and cleaned using a QIAEXII kit (Qiagen). DNA was resuspended in DEPC-treated water.

The probes were labeled by *in vitro* transcription using the DIG-RNA labeling kit (Boehringer Mannheim). Each reaction contained the following in order: 1 μg purified DNA template, 2 μl NTP labeling mixture (10X), 2 μl 10X transcription buffer, 1 μl RNase inhibitor, DEPC-H_2O to bring the volume to 18 μl and 2 μl RNA polymerase (T3 or T7). After incubation for 2 h at 37°C, 2 μl RNase free DNase I was added and the reaction mixture was incubated for 15 min at 37°C. The reaction was terminated with the addition of 2 μl of 0.2 M EDTA.

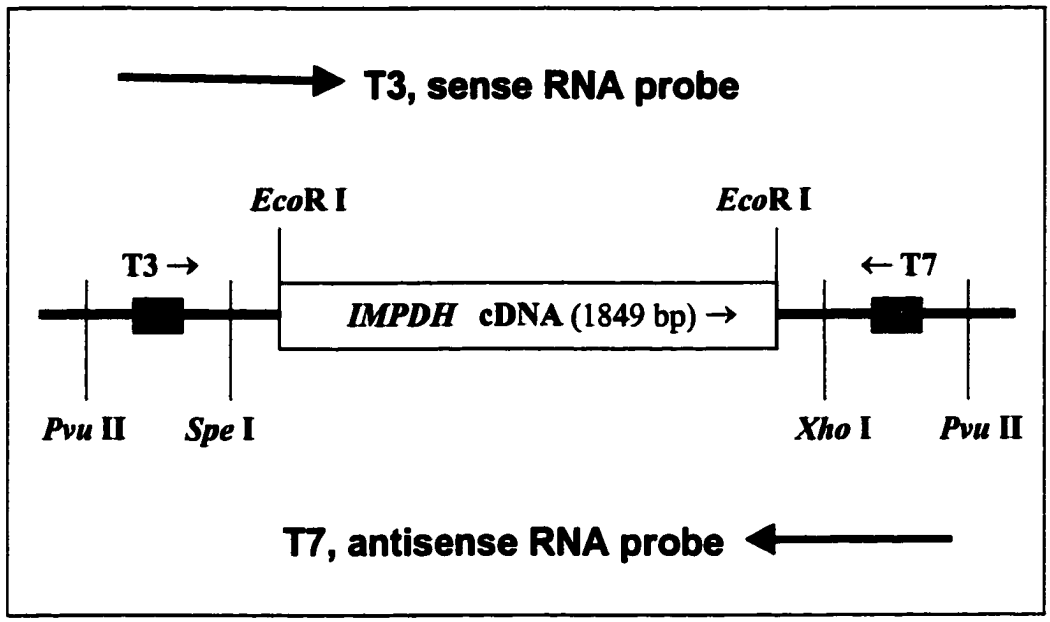


Figure 7-3. Preparation of the probe for *in situ* hybridization. The inserted 1849 bp *IMPDH* cDNA does not contain the 5'-untranslated region. *Pvu* II and *Xho* I were used to prepare the sense RNA probe; *Pvu* II and *Spe* I were used to prepare the antisense RNA probe. The 1849-bp IMPDH cDNA template was subcloned into vector pBluescript SK. Plasmid DNA was linearized with restriction enzymes. Probes were synthesized by *in vitro* transcription using the DIG-RNA labeling kit (Boehringer Mannheim).

To precipitate the labeled probes, 50 µl 3.8 M NH_4OAC and 300 µl ethanol were added and kept at -20°C overnight. After centrifugation at 14,000 rpm for 15 min at 4°C, DIG-labeled RNA pellets were collected, washed with 70% ethanol and 0.15 M NaCl and recentrifuged. Pellets were air-dried and resuspended in 20 µl DEPC-H_2O containing 10 units of RNase inhibitor (Porcine, Amersham Pharmacia Biotech., Piscataway, NJ. Cat. No. 27-0816-01). Two µl of the RNA suspension were run on a 1.2% formaldehyde-agarose gel to check the transcripts (Jones & Sutton, 1997; Dey & Harbone, 1993).

Labeled probes (1950 bases) were cut into shorter fragments (150-200 bases) by limited alkaline hydrolysis. Hydrolysis time was calculated using the following equation:

$$t = (L_0 - L_f) / (K \cdot L_0 \cdot L_f)$$

where L_0 = length in kb of the runoff transcript, L_f = desired length of hydrolysis products, K = 0.11 and t = time in min (McFadden, 1995). For each probe (20 µl), 20 µl hydrolysis buffer (80 mM Na_2CO_3; 120 mM $NaHCO_3$; 20 mM β-mercaptoethanol) was added and the mixture was incubated at 60°C for 41 min. The reaction was stopped with 40 µl of stop buffer (0.2 M NaOAc, pH 6.0; 1% glacial acetic acid; 10 mM DTT). The probe was precipitated by adding 1 µl tRNA (10 mg/ml), 10 µl 3 M NaOAc and 200 µl 100% ethanol followed by incubation at -20°C overnight. After centrifugation for 15 min at 13,000x g, the pellet was washed with 70% ethanol and 0.15 M NaCl, centrifuged again and air-dried. The pellet was resuspended in 20 µl DEPC-H_2O containing 10 units RNase inhibitor. The size of the fragment was determined by electrophoresis of a sample (3 µl) of the resuspended RNA on a 1.2% denaturing agarose gel.

Pretreatment of sections

Sections were heated in a 65°C oven for 1 hour and then placed in xylene for 1 hour to de-wax. Sections were rehydrated by passing through a series of decreasing ethanol concentration (100%, 100%, 95%, 70%, 50% and 0% ethanol (v/v)) for 2 min in each. The slides were transferred to a staining dish containing 2X SSC for 30 min at 70°C to denature the RNA. Sections were then digested

with 125 μg/ml pronase (Boehringer Mannheim, Cat. No. 165-921) for 20 min at 37°C, followed by an acetylation treatment for 10 min with freshly prepared 0.5% acetic anhydride in 20 mM triethanolamine-HCl, pH 8.0. After post-fixation with fresh 4% paraformaldehyde for 20 min, the slides were dehydrated with a gradient ethanol series and air-dried.

Hybridization

RNA probes were prepared according to the methods described above. DIG-labeled probe (2 μl) were mixed with 2 μl RNase-free H_2O and 4 μl deionized formamide, heated at 80°C for 2 min and cooled on ice. Before hybridization, 10X hybridization salts was prepared as follows: 3 M NaCl + 0.1 M Tris-HCl, pH 6.8 + 0.1 M Na_2HPO_4, pH 6.8 + 50 mM EDTA, pH 8.0. Probe was mixed with hybridization buffer (100 μl 10X hybridization salts, 400 μl formamide, 92 μl 10.9 mg/ml yeast tRNA (Sigma, Cat. No. R-8508), 10 μl 10X Denhardt's solution, and 200 μl 50% dextran sulfate). Each slide was incubated with 50 μl probe solution and covered with a sterile coverslip. The slides were placed in a hybridization container with some wash buffer (2X SSC + 50% formamide) at the bottom to keep the sections moist. Hybridization was carried out overnight at 50-55°C.

Washing and ribonuclease A treatment of sections

Slides were washed twice with wash buffer at 50°C, first for 30 min then for 90 min. Slides were then transferred to 1X NTE (NaCl-Tris-EDTA buffer: 0.5 M

NaCl; 10 mM Tris-HCl, pH 7.5; 1 mM EDTA) and incubated at 37°C twice for 5 min each time. RNase A (20 μg/ml in 1X NTE) was used to digest unbound probe by incubating the slides at 37°C for 30 min, followed by two 5 min washes in 1X NTE. Slides were then washed successively in wash buffer for 1 hour at 50°C, 1X SSC for 2 min at room temperature and PBS (phosphate-buffered saline: 0.13 M NaCl; 7 mM Na_2HPO_4; 3 mM NaH_2PO_4, pH 7.2). Slides were stored in fresh PBS overnight at 4°C.

Development and mounting of hybridized tissue for examination

Slides were placed in a shallow plastic sandwich container and incubated in 30 ml of each of the following 6 buffers at room temperature: buffer 1 (100 mM Tris-HCl, pH 7.5; 150 mM NaCl) for 15 min; buffer 2 (0.5% blocking reagent; Boehringer Mannheim, Cat. No. 1096-174) for 1 hour; buffer 3 (1% BSA; 0.3% Triton X-100®) for 1 hour; buffer 4 (1:1500 alkaline phosphatase (AP) conjugated anti-DIG antibody; Boehringer Mannheim, Cat. No. 1093-274) for 1 hour, followed by three changes with fresh buffer 4 for 20 min each; buffer 1 for 5 min; buffer 5 (100 mM Tris-HCl, pH 9.5; 100 mM NaCl; 50 mM $MgCl_2$) for 5 min; buffer 6 (15 μl 50 mg/ml BCIP (5-bromo-4-chloro-3-indolyl phosphate) in dimethylformamide; 25 μl 50 mg/ml NBT (nitroblue tetrazolium) in 70% dimethylformamide in 10 ml buffer 5) in the dark for 24 hours.

After the development, the slides were rinsed in buffer 5 and incubated successively in dH_2O, 50% ethanol, 70% ethanol, 95% ethanol, 100% ethanol, 100% ethanol and xylene for 5 min each. Slides were mounted with coverslips

and sealed with Canadian balsam (Sigma). Sections were observed using a Leitz Dialux 20 light microscope.

RESULTS

Probe preparation

According to the protocol for probe preparation presented in Figure 7-3, the length of the probe after *in vitro* transcription was about 2 kb. After limited alkaline hydrolysis, the expected size of the probe was 150-200 bases. Results from agarose electrophoresis confirmed that the RNA probe was of the expected size after *in vitro* transcription, ethanol precipitation and alkaline hydrolysis (Figure 7-4).

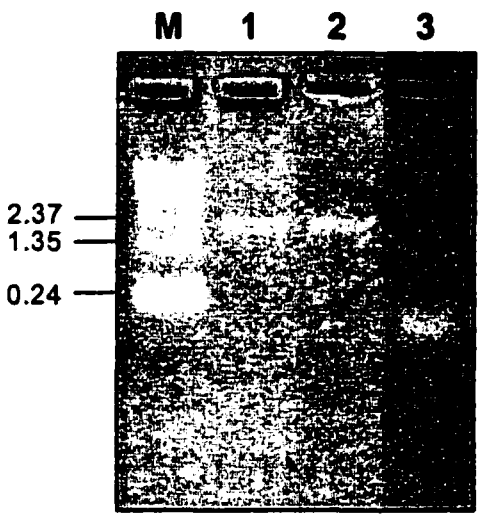


Figure 7-4. 1.2% formaldehyde-agarose gel of DIG-labeled antisense RNA probe. Probes were prepared by *in vitro* transcription using T7 RNA polymerase and precipitated. Transcription products were hydrolyzed at pH 10.2 and then precipitated. Three µl samples were taken from each of the steps mentioned above and loaded onto the gel. Lane 1. *In vitro* transcription product; Lane 2. *In vitro* transcription product after ethanol precipitation; Lane 3. RNA probe: after limited alkaline hydrolysis and ethanol precipitation. Molecular size markers of 0.24-9.5 kb RNA ladder (Gibco BRL, Cat. No. 15620-016) are shown on the left.

Localization of IMPDH *mRNA*

Three different negative controls were used in the *in situ* hybridization experiments. In control 1, incubation with the anti-DIG antibody was omitted to detect endogenous alkaline phosphatase. Control 2 was used to check possible occurrence of RNA-DNA hybrids. In this control, sections were pre-treated with 1 mg/ml RNase A at 37°C for 1 hour before hybridization. Control 3 was a probe control in which a sense RNA probe instead of an antisense RNA probe was used for hybridization.

No signals were detected in soybean nodule tissues, including the central infected region and the peripheral tissues, in controls 1 and 2. The results suggest that neither endogenous alkaline phosphatase nor DNA-RNA hybridization interfered with hybridization or detection. The endodermis was the most distinctive structure in these controls. The endodermis cells encircle the infected region. Staining of these cells was most likely due to the highly sclerified nature of the endodermis.

In control 3, using sense RNA probes (T3), no signals were detected in the infected region, including the infected cells and uninfected cells. However, the nodule vascular bundles exhibited some staining (Figure 7-6B).

When using the antisense *IMPDH* RNA probe (T7), target transcripts were detected only in infected cells in the central region of the nodule (Figure 7-5, 7-6A). In general, the infected cells appeared to be purple in color, with the greatest signal intensity in the nuclei. The uninfected cells, vascular bundles, nodule

parenchyma and nodule cortex were not stained. In mature soybean nodules, the infected cells are much bigger than the uninfected cells.

The intense labeling of the nuclei of infected cells was very interesting. At higher magnification, the dark purple staining was not evenly distributed inside the nucleus (Figure 7-6A). As an obvious comparison, even though the nuclei of infected cells were readily visible, there was no staining of the nuclei in these cells in control sections with sense probe (Figure 7-6B). Nuclei of uninfected cells and cells in peripheral tissues were not labeled with the sense or antisense probe.

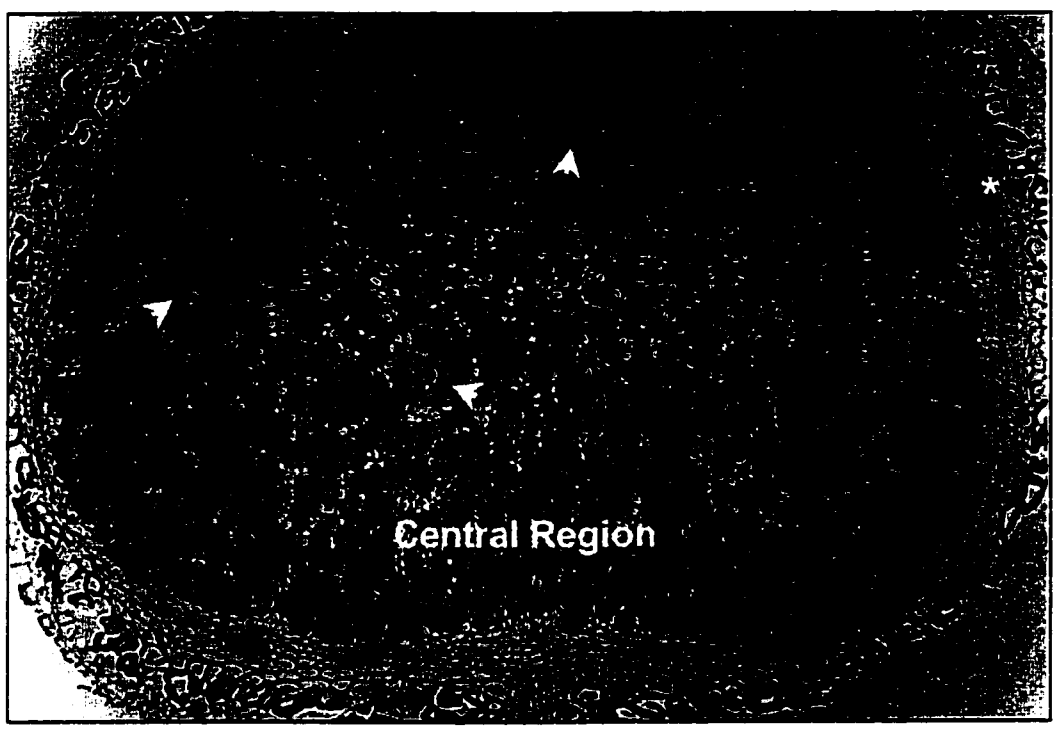


Figure 7-5. *In situ* localization of *IMPDH* transcripts in mature soybean nodules. Transverse sections. Pictures were taken with a 4x objective lense. Hybridization was carried out with DIG-labeled antisense (T7) RNA probes at 55°C. Signals were detected in infected cells within the central region of the nodule. Uninfected cells and other tissues were not stained. *, vascular bundles in the nodule parenchyma. Arrows indicate the black spots of nuclei in the infected cells.

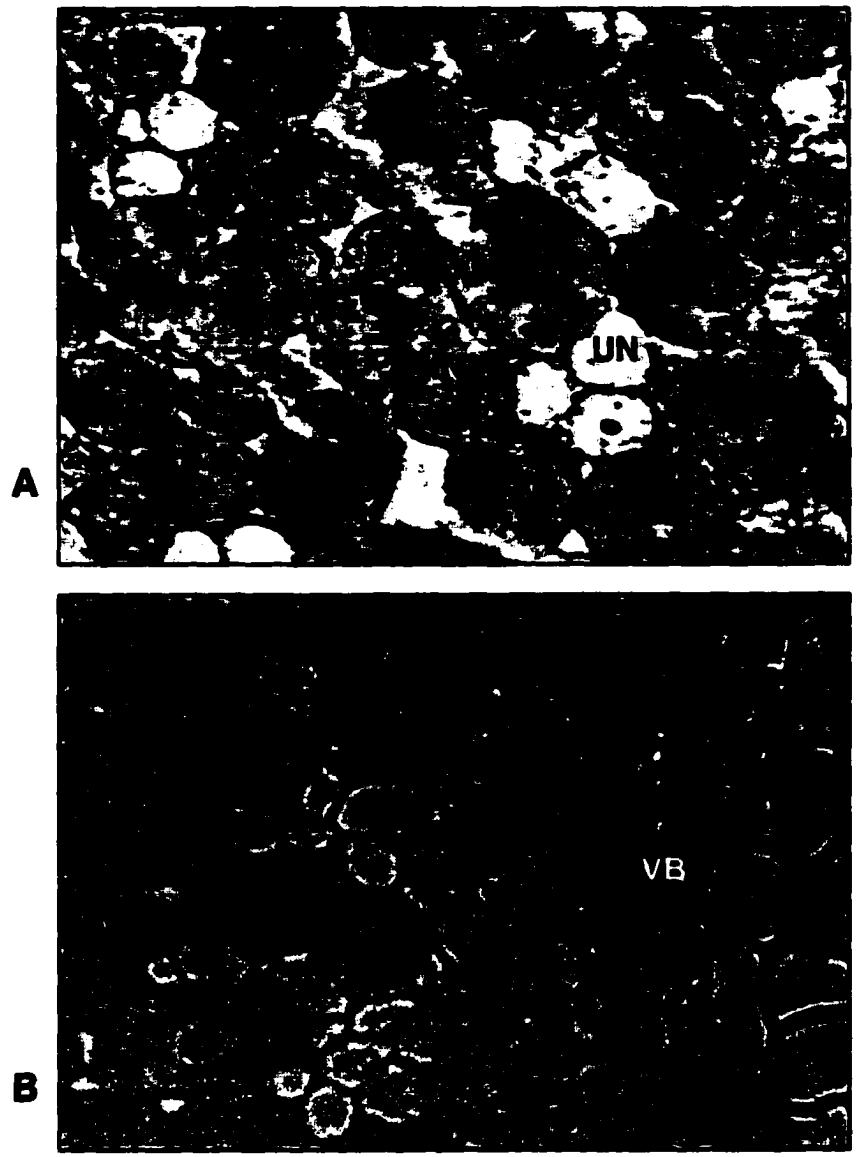


Figure 7-6. ***In situ*** **localization of** ***IMPDH*** **transcripts in mature soybean nodules with sense (T3) and antisense (T7) probes. Transverse sections.** Pictures were taken with a 10x objective lense. Panel **A.** Hybridization with antisense probe (T7). Hybridization signals were detected only in the infected cells (IN). The signal was especially intense in the nuclei. Uninfected cells (UN) were not stained. Panel **B.** Hybridization with sense probe (T3). No signals were detected in either infected cells or uninfected cells. VB, vascular bundles in the nodule parenchyma.

DISCUSSION

Localization of IMPDH transcript: A model for the compartmentalization of enzymes involved in ureide biogenesis in soybean nodules

Up to now, none of the enzymes involved in purine biosynthesis and ureide biogenesis have been localized at the transcriptional level except uricase. In the case of uricase, both mRNA and protein were detected in the uninfected cells of nodules (Bergman *et al.*, 1983; Nguyen *et al.*, 1985; VandenBosch and Newcomb, 1986; Papadopoulou *et al.*, 1995). Biochemical studies suggested that purine biosynthesis is located in the plastid of infected cells (Shelp *et al.*, 1983; Schubert and Boland, 1990). According to the proposed model (Schubert and Boland, 1990; Figure 1-4), the early steps of the purine biosynthesis and ureide biogenesis pathways occur in the infected cells and the final steps occur in the uninfected cells. It is not clear where the intermediate steps take place. Immunolocalization of xanthine dehydrogenase in cowpea showed that labeling in the uninfected cells was much greater than that in infected cells suggesting that xanthine or a precursor to xanthine, rather than uric acid, is the intermediate that moves from infected to uninfected cells during ureide biogenesis (Datta *et al.*, 1991). It is possible that IMP or XMP is the molecule transported from the infected cells to the uninfected cells. Therefore, the localization of IMPDH will help to clarify the uncertainties with respect to the compartmentalization of purine biosynthesis and ureide biogenesis.

In our current work, *in situ* hybridization was used to determine the cellular localization of *IMPDH* mRNA in mature soybean nodules. Results of these

studies demonstrated that the soybean *IMPDH* gene was expressed only in nodule cells infected by rhizobia. This result is consistent with the model for the cellular compartmentalization of purine biosynthesis and ureide biogenesis proposed by Schubert and associates (Schubert, 1986; Schubert and Boland, 1990).

Since the *IMPDH* mRNA was detected in infected cells, it is logical to imagine that the protein will be localized in the infected cells too. mRNAs are usually short-lived and not transported between cells. Confirmation of the proposed localization of IMPDH will require immunocytochemical localization of the protein by the transmission electron microscopy.

At the subcellular level, the localization of IMPDH is still an unsolved problem. IMPDH was localized in plastids of soybean nodules (Boland and Schubert, 1983). In contrast, Shelp and Atkins (1983) found that IMPDH was in the cytoplasm of cowpea nodules. Resolution of these apparent differences will require immunochemical examination combined with electron microscopy.

Labeling of the nuclei of infected cells

An interesting phenomenon observed in our studies was the intense labeling of the nuclei of infected cells. As far as we are aware, this observation has never been reported in any other *in situ* hybridization studies. In the localization of soybean nodulin *GmN#93* transcripts with DIG-labeled RNA probes, signals were detected in infected cells but the intensity of labeling in the nucleus and in the cytoplasm were almost the same (Kouchi and Hata, 1993). *In situ* localization of nodulin transcripts was performed with ^{35}S-labeled RNA probes and the nucleus was not

radiolabeled. These *in situ* hybridization studies included mRNA detection of legaehemoglobin in alfalfa (de Billy *et al.*, 1991), glutamine synthetase in alfalfa (Temple *et al.*, 1995), the early nodulin gene *ENOD 40* in pea and soybean (Matvienko *et al.*, 1994), *ENOD 2* in pea and soybean (van de Wiel *et al.*, 1990) and asparagine synthetase in alfalfa (Shi *et al.*, 1997).

For eukaryotic messenger RNA, the primary transcripts are processed in the nucleus and then transported to the cytosol as mature mRNA for translation. Since the probes should hybridize with both primary transcripts and mature mRNAs, the intense signals in the nucleus may indicate that the primary transcripts are more concentrated within the nucleus, while the mature mRNA is more dispersed in the ground cytoplasm.

As demonstrated in Chapter 6, the expression of *IMPDH* gene was induced during nodule development. *In situ* localization of *IMPDH* mRNA in developing nodules could provide information on the temporal pattern of expression in different cell types. It is possible that the intense labeling of the nuclei is a result of nodule aging. Further investigation on *in situ* localization of *IMPDH* mRNA in younger nodules is needed.

REFERENCES

Bergersen, F.J. (1982). Root nodules of legumes: Structure and function. Research Studies Press, New York.

Bergmann, H., Preddie, E. and Verma, D.P.S. (1983). Nodulin-35: a subunit of specific uricase (uricase II) induced and localized in the uninfected cells of soybean nodules. EMBO J. 2: 2333-2339.

Boland, M.J. and K.R. Schubert (1983). Biosynthesis of purines by proplastid fraction from soybean nodules. Arch Biochem Biophys 220: 179-187.

Datta, D.B., Triplett, E.W. and Newcomb, E.H. (1991). Localization of xanthine dehydrogenase in cowpea root nodules: implication for the interaction between cellular compartments during ureide biogenesis. Proc Natl Acad Sci USA 88:4700-4702.

de Billy, F., Barker, D.G., Gallusci, P. and Truchet, G. (1991). Leghaemoglobin gene transcription is triggered in a single cell layer in the indeterminate nitrogen-fixing root nodule of alfalfa. Plant J 1: 27-35.

Franssen, A.J., Vijn, I., Yang, W.C. and Bisseling, T. (1992). Developmental aspects of the rhizobium-legume symbiosis. Plant Mol Biol 19: 89-107.

Hanks, J.F., Tolbert, N.E. and Schubert, K.R. (1981). Localization of enzymes of ureide biosynthesis in peroxisomes and microsomes of nodules. Plant Physiol 68: 65-69.

Jones, P.G. and Sutton, J.M. (1997). Plant molecular biology: essential techniques. John Wiley & Sons, New York. pp111.

Kouchi, H. and Hata, S. (1993). Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. Mol Gen Genet 238: 106-119.

McFadden, G.I. (1995). *In situ* hybridization. Methods in Cell Biol 49: 165-183.

Matvienko, M., Van De Sande, K., Yang, W.-C., Van Kammen, A, Bisseling, T. and Franssen, H. (1994). Comparison of soybean and pea *ENOD40* cDNA clones representing genes expressed during both early and late stages of nodule development. Plant Mol Biol 26: 487-493.

Miao, G.-H. and Verma, D.P.S. (1993). Soybean nodulin-26 gene encoding a channel protein is expressed only in the infected cells of nodules and is regulated differently in roots of homologous and heterologous plants. Plant Cell 5:781-794.

Mylona, P., Pawlowski, K. and Bisseling, T. (1995). Symbiotic nitrogen fixation. The Plant Cell 7: 869-885.

Newcomb, W. (1981). Nodule morphogenesis and differentiation. Int Rev Cytol suppl 13: 247-298.

Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H. and Verma, D.P.S. (1985). Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. Proc Natl Acad Sci USA. 82: 5040-5044.

Papadopoulou, K., Roussis, A., Kulin, H. and Katinakis, P. (1995). Expression pattern of uricase II gene during root nodule development in *Phaseolus vulgaris*. Experientia 51: 90-94.

Shelp, B.J. and Atkins, C.A. (1983). Role of inosine monophosphate oxidoreductase in the formation of ureides in nitrogen-fixing nodules of cowpea (*Vigna unguiculata* L. Walp.). Plant Physiol 72: 1029-1034.

Shi, L, Twary, S.N., Yoshioka, H., Gregerson, R.G., Miller, S.S., Samac, D.A., Gantt, S., Unkefer, P.J. and Vance, C.P. (1997). Nitrogen assimilation in alfalfa: isolation and characterization of an asparagine synthetase gene showing enhanced expression in root nodules and dark-adapted leaves. Plant Cell 9: 1339-1356.

Tanimoto, E.Y. and Rost, T.L. (1993). Non-radioactive *in situ* RNA hybridization using digoxigenin and an application for co-localization studies with radioisotope. In: Molecular Biology (Bryant, J.A. Ed.). Academic Press, San Diego. pp141-158.

Temple, S.J., Heard, J., Ganter, G., Dunn, K. and Sengupta-Gopalan, C. (1995). Characterization of a nodule-enhanced glutamine synthetase from alfalfa: nucleotide sequence, *in situ* localization, and transcript analysis. Mol Plant Microbe Interact 2: 218-227.

VandenBosch, K.A. and Newcomb, E.H. (1988). The occurrence of leghemoglobin protein in the uninfected interstitial cells of soybean root nodules. Planta 175: 442-451.

van de Wiel, C., Scheres, B., Franssen, H., van Lierop, M.-J., van Lammeren, A., van Kammen, A. and Bisseling, T. (1990). The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. EMBO J 9: 1-7.

APPENDICES: Protocols and Solutions

GROWTH OF SOYBEAN PLANTS

The stock of *Bradyrhizobium japonicum*, strain USDA 110 stored at -80°C was recovered by streaking on MAG plates and cultured at 27°C for about 10 days. Single colonies were inoculated in MAG broth, cultured for 3 days until the OD_{600} reaches 1, and used for soaking the soybean seeds for 1~3 h. The infected soybean seeds were planted in vermiculite at a density of 6~8 seeds per pot. $MgSO_4$, $CaSO_4$ and K_2SO_4 were applied afterwards. N-free nutrient solution in an Eb and Flow tank was used for watering the plants for 20 min every other day.

'MAG' medium:

Compounds	Stock Conc.	Working Conc.*
$MgSO_4.7H_2O$	18 g / 100 ml	1 ml/L
$CaCl_2$	0.75 g / 50 ml	1 ml/L
NH_4Cl	16 g / 100ml	2 ml/L
KH_2PO_4		0.22 g/L
Na_2SO_4		0.25 g/L
$FeCl_3.6H_2O$	0.67 g / 100 ml	1 ml/L
NaMolybdate	1 g / 100 ml	1 ml/L
$NiCl_2$	1.2 g / 100 ml	0.1 ml/L
MES buffer		1.1 g/L
HEPES buffer		1.3 g/L
Arabinose		1.0 g/L
Na gluconate (gluconic acid)		1.0 g/L
Yeast extract (nutrient)		1.0 g/L

* weigh out the solids and dissolve in ~800 ml dH_2O, add the indicated volume of each stock solution, adjust the pH~6.6, bring volume to 1000 ml, add agar (15g) if required (only for plates), autoclave for 20 min.

N-free Nutrient Solution:

For a 200-liter Eb and Flow tank, add 625 ml of Macronutrient stock solution, 10 liters of $CaSO_4$ stock solution, 200 ml Micronutrient stock solution, 200 ml Sequestrene stock solution, and bring the volume to 200 liter with DI H_2O.

Macronutrient stock soluition (320X): 48 g of $MgSO_4.7H_2O$, 60 g of KCl and 56 g of KH_2PO_4 in 4 liters DI H_2O.

Calcium sulfate stock solution (20X): 26.6 g of $CaSO_4$ in 20 liters DI H_2O.

Sequestrene (Sequested Iron) stock solution (1000X): 2.9 g of Sequestrene in 4 liters DI H_2O.

Micronutrient stock solution (1000X): add 1ml of each of the following solutions and bring to a final volume of 1 liter with DI H_2O.

Compounds	Amount (g/50 ml)	MW
LiCl	0.45	42.39
$CuSO_4.5H_2O$	0.9	249.68
$ZnSO_4.7H_2O$	0.9	287.54
H_3BO_4	1.0	61.84
$Al_2(SO_4)_3.18H_2O$	0.9	665.42
KI	0.9	166.01
$MnCl_2.4H_2O$	0.9	197.92
$NiCl_2.6H_2O$	0.9	237.71
$CoCl_2.6H_2O$	3.65	99.4
KBr	0.45	119.01
$NaMoO_4.2H_2O$	1.8	241.95

λgt11 PHAGE DNA ISOLATION

The following protocol was modified from the Promega protocol for the "Wizard™ Lambda Prep DNA Purification System".

1. Lysate preparation: liquid culture method
   - Single plaques with white color were picked up from fresh plates, put in 100 µl of phage buffer and 1.0 µl chloroform and stored at 4°C overnight;
   - Start a fresh culture of *E. coli* host strain. Shake overnight at 37°C;
   - Mix 1250 µl *E. coli* culture with 15 µl phage plug elute. Incubate 30 min at 37°C;
   - Add 1 ml 1 M $MgSO_4$ in 100 ml LB broth in a 250 ml flask. Pre-warm to 37°C and add bacteria-phage mixture. Shake at 37°C for about 5 hours until cell lysis occurs. Add 500 µl chloroform, shake another 15 min;
   - Centrifuge 10 min at 10,000x g;
   - Collect the supernatant. This is the lysate for phage DNA isolation.
2. DNA isolation:
   - Transfer 10 ml lysate in a Corex centifuge tube;
   - Add 40 µl Nuclease Mixture (0.25 mg/ml RNase A + 0.25 mg/ml DNase I + 150 mM NaCl + 50% glycerol). Incubate 30 min;
   - Add 4 ml Phage Precipitant (33% PEG-8000 + 3.3 M NaCl), mix gently and put on ice for at least 30 min;
   - Centrifuge 10 min at 10,000x g;
   - Pour away the supernatant slowly. The precipitation is white in color;
   - Resuspend the precipitation with 500 µl Phage Buffer (150 mM NaCl + 40 mM Tris-HCl, pH 7.4 + 10 mM $MgSO_4$). Pipette the solution up and down several times along the wall of Corex tube until it is completely dissolved;
   - Add 50 µl 5 mg/ml proteinase K to a final concentration of 0.5 mg/ml. Incubate 15-30 min at 37°C;

- Spin 10 sec and transfer the supernatant to a new microcentrifuge tube;
- Add 1 ml thoroughly mixed Purification Resin to each tube. Mix by inverting the tube several times, then transfer into the syringe barrel which is attached to a mini-column;
- Gently press the plunger into the syringe barrel, making the resin mixture flow through the mini-column;
- Detach the plunger from the barrel, add 2 ml 80% isopropanol in the barrel and apply the plunger to the barrel again. Push in and make the solution go through the mini-column;
- Detach the syringe from the mini-column. Centrifuge the mini-column for 2 min at 10,000x g to remove the isopropanol;
- Use 100 µl pre-warmed H_2O to elute the mini-column. Spin for 20 sec;
- Phage DNA is in the eluate.

SUBCLONING OF IMPDH cDNA INTO pBluescript SK (+/-)

1. Making Competent Cells:
- Inoculate single colonies of *E. coli* strain XL1Blue mrf' into LB broth, culture at 37°C overnight;
- Dilute the overnight culture with LB broth at 1:100, shaking at 37°C until OD_{600}=0.3 (no higher);
- Centrifuge at 1000x g for 10 min at 4°C. Resuspend the pellet with transformation and storage solution (TSS) with a volume of 1/10 of the original culture. 50 ml TSS was prepared as: mix 8 ml LB broth, 1.5 g PEG 3650/4000, 0.5 ml DMSO, 0.25 ml 1 M $MgCl_2$ and 1.0 ml glycerol. Filter-sterilized;
- For storage: aliquot into 1.5 ml eppendorf tubes, and store at -80°C; For transformation: use 100 µl for each transformation tube.

2. Cloning *IMPDH* cDNA Fragments into Plasmid pBluescript SK(-): a protocol from Dr. Ruzhu Chen.
- Isolated phage/plasmid DNA was digested with *Eco*R I at 37°C for 2 h:

phage DNA (~0.1 µg/µl)	40 µl
Multicore (10X)	5 µl
*Eco*R I	4 µl
H_2O	1 µl

then concentrate the volume to ~30 µl by speed-vac;
- Linearize the plasmid with *Eco*R I:

pBluescript SK DNA (2 mg/ml)	10 µl
MultiCore (10X)	3.5 µl
EcoR I	2.5 µl
H_2O	14 µl

Incubate at 37°C for 2 h, inactivate the enzyme at 65°C for 10 min and treat with phosphatase by adding the followings:

Digested phage DNA (of the above)	30 µl
CIAP (Calf Instinal Alkaline Phosphatase)	2.5 µl
Alkline Phosphatase buffer (10X)	5.0 µl
H_2O	12.5 µl

Incubate at 37°C for 30 min.

- Run 1.2% agarose gel, cut the target bands and clean with QIAEX II kit (Qiagen). Speed-vac dry. Resuspend with H_2O in 20 µl for vector DNA and 10 µl for insert DNA.
- Ligation was set up as the following, then incubate overnight at 15°C.
  10 µl DNA fragments, 10 µl vector DNA, 3 µl 10X ligation buffer, 3 µl 1 mM ATP, 2 µl H_2O, 2 µl T4 ligase
- Mix 100 µl competent cells (thawed on ice) with 2 µl or 10 µl ligation solution. Use no DNA as negative control;
- Incubate on ice for 45 min;
- Add 890 µl LB broth + 10 µl 40% glucose for each tube, shaking (400 rpm) at 37°C for 1.5 hour;
- Plate on LB-amp$^+$ plates containing X-gal and IPTG. For each transformation, plate 10 µl, 50 µl, 100 µl, 200 µl and the remainder of the transformant culture. Grow at 37°C overnight.

  * Ampicillin was added at 50 µg/ml; 100 µl of 1:10 diluted X-gal stock (50 mg/ml in N,N' dimethylformamide) was applied to each plate and spread out; IPTG was applied in the same way as X-gal except that the concentration for the stock is 0.1 M.

PLANT GENOMIC DNA ISOLATION

- Grind plant tissue in LN_2 into powder, add 500 µl extraction buffer (100 mM Tris, pH 8.0; 1.4 M NaCl; 20 mM EDTA; 2% CTAB; 0.2% 2-mercaptoethanol), mix well, aliquot to 1.5 ml Ependorf tubes;
- Incubate at 65°C for 30 min;
- Add 100 µl chloroform / iso-amyl alcohol (24:1) into the extract and vortex briefly;
- Centrifuge 3 min, collect the supernatant in a new tube;
- Precipitate DNA with ethanol/acetate solution (96 ml 100% ethanol, 4 ml 3 M KOAc, pH 5.2), spin 3 min to collect the DNA;
- Dry the pellet and add 500 µl TE buffer to re-dissolve the pellet;
- Dissolve the DNA into 20 µl H_2O and add 1 µl RNase A (1 mg/ml). Incubate 30 min at 37°C.

TOTAL RNA ISOLATION

One-step isolation: Guanidine thiocyanate-phenol-chloroform extraction method.

- Grind 1.0 g of frozen tissue into a fine power in the presence of LN_2;
- Add 10 ml denaturing solution (working solution) and homogenize, transfer into a 50-ml polypropylene tube;

- Add 1.0 ml 2 M NaOAc, pH 4.0, mix thoroughly by inversion; 10 ml H_2O-saturated phenol, mix thoroughly; 5.0 ml chloroform/isoamyl alcohol (49:1), mix thoroughly. Incubate the suspension on ice for 15 min;
- Centrifuge at 10,000x g for 20 min. Remove the supernatant carefully and transfer to a clean tube;
- Add an equal volume of 100% isopropanol to each tube, mix and leave at -20°C for 1.5-2.0 hours to precipitate RNA;
- Centrifuge at 10,000x g for 20 min at 4°C. Discard the supernatant;
- Dissolve the RNA pellet in 2.5-3.0 volume of denaturing solution (this may require vigorous mixing and heating at 65°C), then transfer to 1.5 microcentrifuge tubes;
- Add an equal volume of 100% isoproponal to each tube, mix and precipitate RNA at -20°C for 1 hour to overnight;
- Centrifuge at 10,000x g for 20 min at 4°C, then discard the supernatant;
- Wash the pellet with 75% DEPC-treated EtOH twice: vortex and incubate 10-15 min at room temperature (to dissolve residual amounts of guanidinium contaminating the pellet), and centrifuge at 10,000x g for 20 min;
- Vacuum dry the RNA pellet for 10-15 min (never over-dry);
- Dissolve RNA in 200 µl DEPC-H_2O by heating at 65°C and vigorously vortexing;
- Measure OD_{260} and OD_{280}, calculate the RNA concentration;
- Store samples frozen at -70°C or in ethanol at -20°C;
- Quick quality assessment: run out 1-2 µg RNA on a normal non-denaturing 1% agarose gel.
-

Solutions:

1). DEPC-H_2O: 0.2% Diethyl pyrocarbonate (DEPC) in H_2O, shake vigorously and autoclave.

2). 2 M NaOAC: 16.42 g NaOAc (anhydrous), mix with 25-35 ml H_2O and 35 ml glacial acetic acid, adjust pH 4.0 with glacial acetic acid, to a final volume of 100 ml with H_2O, add 0.2 ml DEPC, shake vigorously, autoclave.

3). 0.75 M NaCitrate, pH 7.0: 22.06 g NaCitrate dissolve in H_2O, adjust pH 7.0, to a final volume of 100 ml.

4). 10% Sarkosyl: 20 g Sarkosyl solid in 200 ml H_2O.

5). Denaturing solution:
Stock: 250 g / 125 g guanidimium thiocyanate, dissolve in 293 ml / 146.5 ml, add in 17.6 ml / 8.8 ml 0.75 M NaCitrate, pH 7.0, add in 26.4 ml / 13.2 ml 10% Sarkosyl, stir at 60-65°C until thoroughly mixed (total volume is ~510 ml / 260 ml), store at RT, up to 3 months.
Working solution: 0.35 ml 2-ME in 50ml stock solution.

6). H_2O-saturated phenol:
Dissolve 100 g phenol crystals in H_2O at 60-65°C. Aspirate the upper water phase and store up to 1 month at 4°C.
Do not use buffered phenol in place of water-saturated phenol.

FORMALDEHYDE-AGAROSE (FA) GEL ELECTROPHORESIS

This protocol (from Qiagen) uses a more concentrated RNA loading buffer allowing a larger volume of RNA samples to be loaded onto the gel compared to conventional protocols.

1.2% FA Gel (100 ml):
1.2 g Agarose + 10 ml 10x FA gel buffer, make to 100 ml with RNase-free water. Microwave to melt agarose. Cool to 65°C in a waterbath. Add 1.8 ml of 37% (12.3 M) formaldehyde and 1 µl of ethidium bromide (10 mg/ml). Mix thoroughly and pour onto gel support. Prior to running the gel, equilibrate in 1x FA Gel Running Buffer for at least 30 min.

RNA Sample:
Add 1 volume of 5x RNA loading buffer per 4 volumes of RNA sample. Incubate for 3-5 min at 65°C, chill on ice, and load onto the equilibrated FA gel.

Gel Running Conditions:
Run gel at 5-7 V per cm with 1x FA gel running buffer.

Composition of FA Gel Buffers:
10x FA Gel Buffer: 200 mM MOPS (free acid), 50 mM sodium acetate, 10 mM EDTA, pH to 7.0 with NaOH.
1x FA Gel Running Buffer: 100 ml 10x FA gel buffer, 20 ml 37% formaldehyde (pH >4.0), 880 ml RNase-free water.
5x RNA Loading Buffer: 16 µl saturated bromophenol blue, 80 µl 500 mM EDTA, pH 8.0, 720 µl 37% formaldehyde, 2 ml 100% glycerol, 3084 µl formamide, 4 ml 10x FA gel buffer, add RNase-free water to 10 ml. This buffer can be kept at 4°C for ~ 3 months. When use, e.g., 40 µl RNA and 10 µl 5x RNA loading buffer, incubate for 3-5 in at 65°C, chill on ice, load onto gel.

SDS-PAGE

1) For one gel, a Separate Gel (Running Gel) was prepared as:

Components	15% gel	12.5% gel
30% acrylamide Bis (30:0.8=37.5:1)	12 ml	10 ml
3M Tris-HCl (pH8.9)	2.8 ml	2.8 ml
10% SDS	225 µl	225 µl
H_2O	9 ml	11 ml
APS (10%)	225 µl	225 µl
TEMED	7.5 µl	7.5 µl

2) Pour the gel and seal with N-butanol. Polymerize for 30 min.
3) Prepare the Stacking Gel: 3% acrylamide

Acrylamide Bis (30:1.03)	1.3 ml
0.5 M Tris-HCl (pH 6.8)	2.5 ml
H_2O	6.1 ml
10% SDS	100 µl
10% APS	100 µl
TEMED	4 µl

4) Insert the comb and pour the Stacking Gel in. Polymerize for 30 min.
5) Label the bottom of the wells, take out the comb and pour the Running Buffer. Expel air bubbles in the bottom tank.
   1000 ml Running Buffer: 28.8 g glycine, 6.0 g Tris-free base, 1.0 g SDS.
6) Mix protein sample with Loading Buffer at 1:1, boil for 3 min and load.
   Loading Buffer: 15 g sucrose, 2 g SDS and 10 mg Bromophenol Blue were dissolved in 95 ml Running buffer, store at -20°C in 950 µl aliquots. Before use, mix 950 µl aliquot with 50 µl mercaptoethanol in hood.
7) Run gel with 80-100 V for 3 h.
8) Stain the gel with Stain Solution for 2 hours to overnight.
   Stain Solution: dissolve 0.375 g Coomassie Brilliant Blue R-250 in 67 ml methanol, add 15 ml glacial acetic acid, bring the volume to 150 ml with H_2O and filter with Whatmann #1 Gravity membrane.
9) Destain the gel with Destain Solution (10% glacial acetic acid, 30% methanol).
10) If desired, Silver Staining could be carried out as the following:
    Fixation: 10% acetic acid (≥1 h, RT) → Pretreatment: 0.01% Sodium Thiosulfate (200 ml, 1 min) → Wash with dH_2O, 3 times at 5-10 min each → Staining: 0.1 g silver nitrate in 100 ml solution + 150 µl formaldehyde (200 ml, at least 20 min) → wash → development: 12% sodium carbonate + 5 ml pretreatment solution + 150 µl formaldehyde → stop the development with distilled H_2O.
11) If desired, gels could be stabilized with 3% glycerol for 30 min, then dried with gel dryer.

ELECTROPHORETIC TRANSFER OF PROTEIN SAMPLES FOR WESTERN BLOTS

1). Run SDS-PAGE, using pre-stained markers.
2). Cut nitrocellulose membrane and 3 MM filter paper to desired size. Usually the size is 9.0 cm x 14.5 cm.
3). Prepare Transfer Buffer as: 4.84 g Tris-free base, 22.52 g Glycine, 200 ml methanol. Make the volume to 2 liters with ddH_2O.
4). Set up transfer sandwich.
5). Transfer with constant current of 250 mA for 3.5-4 h.

COMMONLY USED BUFFERS AND STOCK SOLUTIONS

100x Denhardt solution: 500 ml
10 g Ficoll 400
10 g polyvinylpyrrolidone
10 g bovine serum albumin
Filter and store at -20°C in 25-ml aliquots

1 M dithiothreitol (DTT):
Dissolve 15.45 g DTT in 100 ml H_2O, aliquot and store at -20°C

1 M HCl: 1 liter
Mix in the following order: 913.8 ml H_2O, 86.2 ml concentrated HCl

10 M NaOH:
Dissolve 400 g NaOH in 450 ml H_2O, add H_2O to 1 liter

1 M Tris-Cl [tris(hydroxymethyl)aminomethane]:
Dissolve 121 g Tris base in 800 ml H_2O, adjust to desired pH with concentrated HCl, mix and add H_2O to 1 liter.

Phosphate-buffered saline (PBS):

10x stock solution, 1 liter	*Working solution, pH ~7.3*
80 g NaCl	137 mM NaCl
2 g KCl	2.7 mM KCl
11.5 g $Na_2HPO_4.7H_2O$	4.3 mM $Na_2HPO_4.7H_2O$
2 g KH_2PO_4	1.4 mM KH_2PO_4

20X SSC buffer:
3 M NaCl (175 g/liter)
0.3 M Na_3citrate.$2H_2O$ (88 g/liter)
Adjust pH to 7.0 with 1 M HCl/NaOH

TAE electrophoresis buffer:

50X stock solution, 1 liter	*Working solution*
242 g Tris base	0.04 M Tris.acetate
57.1 ml glacial acetic acid	0.002 M EDTA
37.2 g $Na_2EDTA.2H_2O$	
pH ~8.5	

TBE electrophoresis buffer

10x stock solution, 1 liter	*Working solution*
108 g Tris base	0.089 M Tris base
55 g boric acid	0.089 M boric acid
40 ml 0.5 M EDTA, pH 8.0	0.002 M EDTA

Basic 5X M9 minimal medium: 1 liter
30 g Na_2HPO_4
15 g KH_2PO_4
5 g NH_4Cl
2.5 g NaCl
15 mg $CaCl_2$ (optional)
autoclaved. Before using, concentrated media should be diluted to 1X with sterile H_2O and the following sterile solutions per liter: 1 ml 1 M $MgSO_4.7H_2O$, 10 ml 20% glucose.

LB rich medium (broth): 1 liter
5 g NaCl
10 g tryptone
5 g yeast extract
0.5 ml 2 M NaCl