NOTICE:
The copyright law of the United States (Title 17, United States Code) governs the making of reproductions of copyrighted material. One specified condition is that the reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses a reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

RESTRICTIONS:
This student work may be read, quoted from, cited, and reproduced for purposes of research. It may not be published in full except by permission of the author.

Jonathan M. Cook

Honors Project Thesis, February 13, 1997:

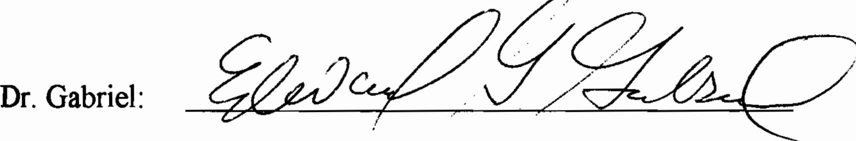
Title: Analysis of FGAR amidotransferase in *Rhizobium etli*

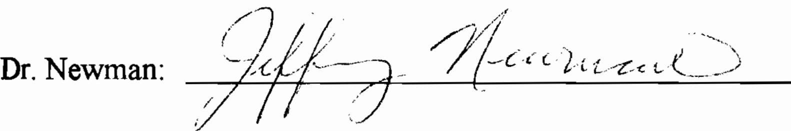
Honors Committee Members:

By placing our signatures on this cover page, we hereby deem the Honors Project of Jonathan M. Cook worthy of Biology Departmental Honors.

Dr. Bendorf:

Dr. Diehl:

Dr. Gabriel: 

Dr. Newman: 

Characterization of FGAR amidotransferase in *Rhizobium etli*

by

Jonathan M. Cook

An Honors Thesis submitted to Lycoming College in Partial Fulfillment of the Requirements for Departmental Honors in Biology

February, 1997

List of Abbreviations

AIR	5-aminoimidazole ribonucleotide
AICAR	5-aminoimidazole-4-carboxamide ribonucleotide
Amp	ampicillin
ATP	adenosine triphosphate
bp	base pairs
cAMP	cyclic adenine monophosphate
Cb	carbenicillin
CTAB	cetyltrimethylammonium bromide
FAD	flavin adenine dinucleotide
FGAM	N-formylglycinamide ribonucleotide
FGAR	N-formylglycinamide ribonucleotide
GTP	guanosine triphosphate
Kb	kilobases
Km	kanamycin
Nal	nalidixic Acid
NAD	nicotinamide adenine dinucleotide
ORF	open reading frame
ORI	origin of replication
PCR	Polymerase Chain Reaction
PRPP	5-phosphoribosyl-1-pyrophosate
STET	sucrose,TritonX-100, EDTA, Tris-HCl
Str	streptomycin
Tc	tetracycline
TE	Tris-HCl, EDTA
TPP	thiamine pyrophosphate

Characterization of FGAR amidotransferase in Rhizobium etli

Abstract

During mutation studies of several *Rhizobium etli* purine auxotrophs an odd phenotype was discovered. Sequencing of DNA flanking Tn5 insertions in *R. etli* strain CE110 revealed a mutation in *purY*, a gene of unknown function. This mutation caused a phenotype similar to mutations in another gene *purQ*, which encodes one subunit of a multi-subunit FGAR amidotransferase. Since *purY* is located in the same region as *purQ* and *purL* and shows an odd phenotype similar to that of *purQ*, it was hypothesized that *purY* encodes a third subunit of the multi-subunit FGAR amidotransferase. To test this hypothesis, the single subunit FGAR amidotransferase encoded by *purL* in *Escherichia coli* was cloned downstream of the *R. etli purL* promoter region in the broad-host-range vector pRK415. Complementation analysis via conjugation of this construct into several purine auxotrophs was completed to confirm or refute the hypothesis. The complementation study confirmed the hypothesis with all purine auxotrophs except mutations in *purL*. It is possible that another gene(s) lie downstream of *purL* to cause this mixed result. In addition, sequence analysis of three pCOS110 fragments was performed to obtain sequence data for *purY* and *purL*, and to determine whether any previously unrecognized gene(s) lie downstream of *purL*. This sequence data made it possible to identify the location of *purC*, immediately upstream of *purY*, and a gene of unknown function, *yjii*, immediately downstream of *purY*. Currently, sequence analysis of *purL* and flanking regions is being performed to obtain sequence data and to identify possible open reading frames downstream of *purL*.

Characterization of FGAR amidotransferase in Rhizobium etli

Introduction

Earlier studies by Noel *et al.* (1988) suggested that undiminished metabolic flow through de novo purine biosynthesis, or a particular intermediate in the pathway, is essential in early symbiotic interactions between *Rhizobium* bacteria and legumes. In attempts to test this, Newman *et al* (1994) studied the symbiosis of *Rhizobium* bacteria with beans, soybeans, and peas. It was found that purine auxotrophs of various *Rhizobium* species are symbiotically defective. This made them unable to initiate or complete the infection process. Before this, Newman *et al.* (1992) demonstrated that, in the *R. etli*-bean symbiosis, infection by purine auxotrophs is partially restored by supplementation of a plant medium with 5-amino-imidazole-4-carboxamide (AICA) riboside, the unphosphorylated form of the purine biosynthetic intermediate AICAR. Newman *et al.* (1995) confirmed the early studies by Noel *et al.* (1988) by testing phenotypes of various *Rhizobium* auxotrophs. Clues to the physiology of nodule development were found. In some biosynthetic pathways, intermediates, rather than end products, may be required for the bacteria to elicit nodule development or infection. It was hypothesized that purine auxotrophs fail to elicit infection threads due to a deficiency in the production of the purine precursor (AICAR).

It is important to study the purine biosynthetic pathway (Figure-1) since it is responsible for the production of essential purine nucleotides. These are used in almost every activity in the cell. Purine nucleotides serve as monomers for macromolecule synthesis of RNA and DNA. They are a component of a cell's energy currency in ATP and redox potential in NAD, FAD, and NADP. Finally, they are a major component of signal transduction mechanisms involving molecules such as GTP and cAMP.

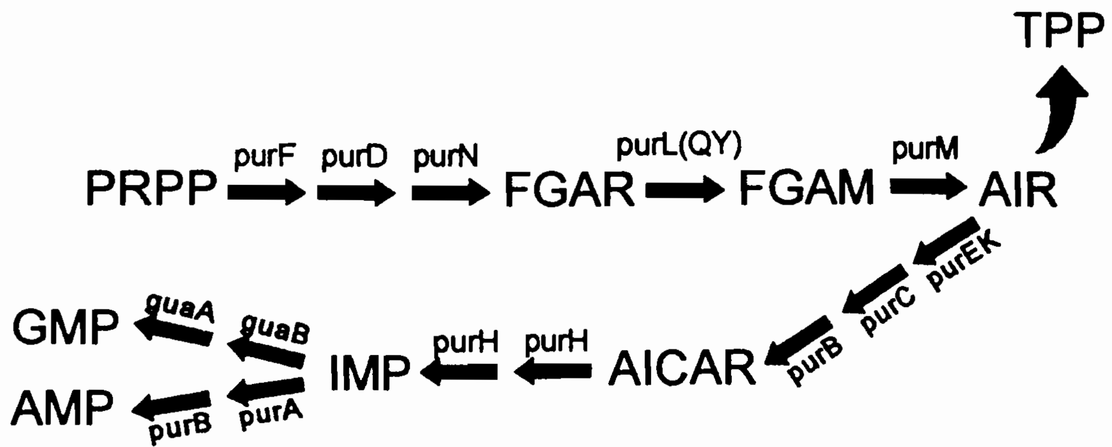


Figure-1. The Purine Biosynthetic Pathway.
Abbreviations: PRPP, 5-phosphoribosyl-1-pyrophosphate; FGAR, N-formylglycinamide ribonucleotide; FGAM, N-formylglycinamidine ribonucleotide; AIR, 5-aminoimidazole ribonucleotide; AICAR, 5-aminoimidazole-4-carboxamide ribonucleotide; TPP, thiamine pyrophosphate.

The purine biosynthetic pathway is needed by all organisms except parasitic protozoa, which must obtain purines from their environment. The series of chemical reactions represented in Figure-1 are used by all organisms capable of *de novo* purine biosynthesis. This is understandable since purines serve the same functions in all organisms (Newman, 1992).

The observations made by Newman et al (1995) in analysis of purine metabolism revealed an odd phenotype. Certain mutants used AICA riboside as a purine source which indicates the metabolic block of these mutants occurs before AICAR in the pathway. The metabolic block was further delimited to the steps before AIR because the mutants did not accumulate imidazole intermediates between AIR and AICAR. However, the mutants did not require thiamine, suggesting that the metabolic block was *after* the production of AIR. This contradiction lead us to investigate purine metabolism further.

In the purine biosynthetic pathway the *pur*Q gene encodes the glutamine amide transfer subunit of formylglycinamide ribonucleotide (FGAR) amidotransferase, an enzyme that catalyzes a reaction before the production of AIR (Ebbole and Zalkin, 1987). This enzyme provides precursors for thiamine biosynthesis as well as purine biosynthesis (Figure-1). The *purQ* mutants displayed thiamine prototrophy suggesting that the catalytic portion of the FGAR amidotransferase, encoded by *purL*, has the ability to use ammonia as a substrate to provide precursors for thiamine biosynthesis which is required at a low level (Newman, 1992). However, purine biosynthesis is eliminated due to the heavier demand for precursors which would be provided by a normal FGAR amidotransferase enzyme. Upstream of *purQ* lies a gene of unknown function, *purY* (Figure-2). Mutations in *purY* have shown the same odd phenotype as mutations in *purQ* (Newman, 1992).

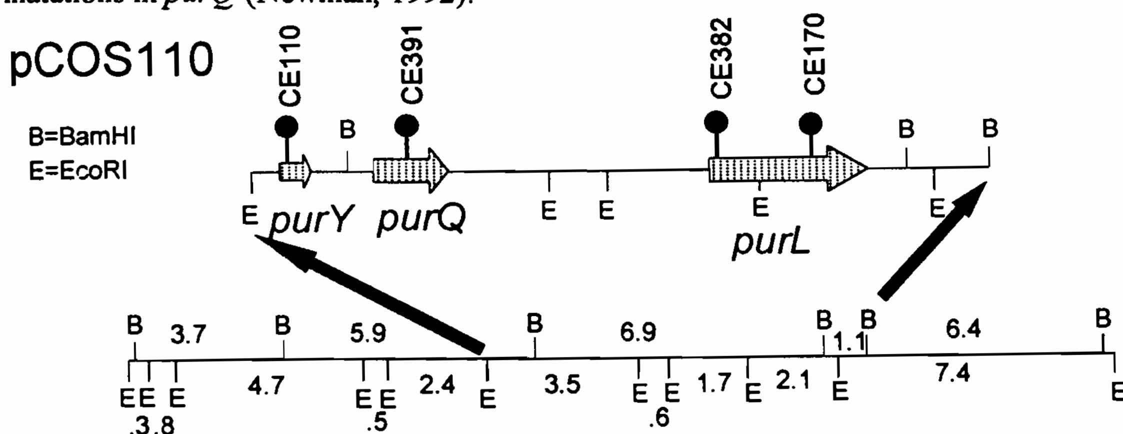


Figure-2. Map of pCOS110.
Numbers indicate sizes of restriction fragments in Kb. Ball and stick objects represent Tn5 insertions from the purine auxotrophs indicated.

It was hypothesized that since mutations of *purY* cause a similar phenotype to mutations in *purQ* and since *purY* is located in the same region as *purQ* and *purL*, that *purY* encodes for a third subunit of FGAR amidotransferase. One of the major goals of the present project was to

determine the function of the *purY* gene. The second goal was to characterize FGAR amidotransferase through DNA sequence of the genes that encode it. DNA encoding *purY* and flanking regions were to be sequenced. In addition, attempts to sequence the *purL* gene, which encodes one of the known subunits of the multi-subunit FGAR amidotransferase were made. Regions immediately upstream and downstream of *purY* and *purL* were sequenced to determine whether there are any other open reading frames within these regions.

To have an understanding of how the project goals are to be completed it is necessary to give a brief explanation of the chemistry involved in FGAR amidotransferase catalyzed reactions. Equally important is background on the evolutionary development and organization of the enzyme.

The overall chemistry for purine biosynthesis has been conserved among all organisms that possess this ability. However, the enzymes that catalyze the chemical processes show a pattern of evolutionary change from one organism to the next. Zalkin and Dixon (1992) showed an evolutionary trend in the enzymes in *de novo* purine biosynthesis suggesting that multi-functional enzymes are a more modern development.

Important to this project is the fact that some organisms have a FGAR amidotransferase enzyme consisting of a single polypeptide, which catalyzes step 4 in the pathway (Figure-1), N-formylglycinamide ribonucleotide (FGAR) to N-formylglycinamidine ribonucleotide (FGAM). In eukaryotes and the γ-group of proteobacteria which includes *Salmonella typhimurium*, *Haemophilus influenzae*, and *Esherichia coli*, FGAR amidotransferase is encoded as a single polypeptide, designated by our lab as type I. The *E. coli* FGAR amidotransferase has three domains, an ATP-binding and cleavage domain on the N-terminus, an FGAR enolization domain

in the central portion, and a glutamine amide transfer domain at the C-terminus (Figure-3).

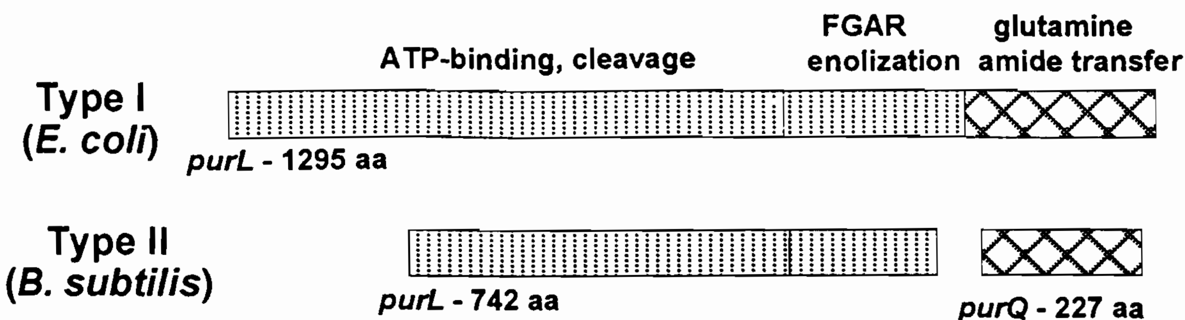


Figure-3. Organization of FGAR amidotransferase.

Ebbole and Zalkin (1987) sequenced the DNA of purine biosynthetic genes from *B. subtilis.* Comparison of this sequence to *E. coli pur* gene sequence showed 40-70% amino acid sequence identity for most of the genes. However, the amino acid identity ranged between 25-30% for the *purL* gene, demonstrating a larger difference. These findings showed that the FGAR amidotransferase of *B. subtilis* is composed of at least two subunits, designated by our lab as type II. The *purL* gene of *B. subtilis* encodes polypeptides that are very similar to the ATP- binding, FGAR-enolization, and glutamine amide transfer domains of the *E. coli* FGAR amidotransferase (Figure-3). The *B. subtilis purQ* gene overlaps the 5' end of *purL* and shows amino acid sequence homology to the glutamine amide transfer domain of *E. coli* FGAR amidotransferase. In addition, all prokaryotes other than γ-group proteobacteria, show a greater amount of similarity to *B. subtilis purL* than to the *E. coli purL* (Newman, 1992). Thus, it appears that among prokaryotes, the multi-subunit FGAR amidotransferase is more common than the single subunit form in *E. coli* and other γ-group proteobacteria bacteria.

An additional difference between *B. subtilis* and *E. coli* is the presence of *purY* in *B. subtilis* but not *E. coli*. The only information on these genes, until recently, was sequence data. Two subunits were thought to make up FGAR amidotransferase, but since *purY* is located in the same region as *purQ* and *purL*, and *Rhizobial purY* mutants show the same phenotype as *purQ* mutants, it is thought that *purY* may encode part of the type II FGAR amidotransferase in prokaryotes such as *R. etli*.

To test this hypothesis, the *R. etli purL* promoter and *E. coli purL* gene encoding type I FGAR amidotransferase were cloned into a broad-host-range vector and conjugated into several purine auxotrophs. If the construct corrects the metabolic defect in the *R. etli purY* mutant it will confirm that the *purY* mutant is defective in FGAR amidotransferase, suggesting that it is indeed a third subunit of FGAR amidotransferase. Initial results suggest that *purY* is indeed a third subunit of FGAR amidotransferase, however, purine auxotrophs mutated in *purL* were not complemented by the construct leaving question of a possible gene or genes downstream of *purL*.

The second part of this project involved sequencing of DNA containing several purine biosynthetic genes. DNA sequencing of the *purY* and *purL* genes was performed through use of the Tn1000 system (Strathmann *et al.* 1991) distributed by Gold Biotechnology. The *purY* gene is located in a 1.0Kb fragment while the approximately 2.0Kb *purL* gene spans two fragments of DNA on the pCOS110 clone. As seen in Figure-2, the 5' end of *purL* starts on the latter half of a 1.7Kb DNA EcoRI fragment. The 3' end of *purL* extends onto the adjoining 2.1Kb EcoRI-BamHI fragment. The DNA fragments to be sequenced were cloned into a vector containing only an origin of replication, an ampicillin resistance gene and a multi-cloning region. Once cloned into such a vector, the plasmid was transformed into a host strain containing an episomal Tn1000.

There, the recombinant plasmid forms a mobilizable cointegrate with the episome. Transfer via conjugation into a kanamycin-resistant recipient strain and resolution on the cointegrate yielded a plasmid containing Tn1000 insertions. If the transposon inserts itself in the origin of replication or the ampicillin resistance gene the clone will not survive. Restriction analysis was done to map the positions of Tn1000 insertions, and a nested set was developed in order to cover the entire region of the gene. Primers designed to recognize DNA immediately adjacent on either end of the Tn1000 in the clones of the nested set were used to quickly sequence DNA. This system is very efficient since new primers do not have to be designed and ordered after each set of new sequence is analyzed. Gaps that remain in the nested set after primary sequencing is done can be sequenced using specifically designed primers.

In addition to gaining sequence data for the known genes, the sequence information will be able to tell whether any other open reading frames are located within the same regions or immediately upstream or downstream of the *purY and purL* genes.

DNA sequence was obtained for the 1.0Kb fragment containing *purY*. It was found that the C-terminal end of *purC* is just upstream of *purY* while the N-terminal end of *yjii* lies immediately downstream of *purY*. Little sequence was produced for the 1.7Kb and 2.1Kb fragments containing *purL*.

Methods

Bacterial Strains, Plasmids, and Antibiotics Used:

A list of bacterial strains, plasmids, antibiotics, and some information about them can be found in Table-I. *R.etli* strains were grown on TY (rich) media at 30°C for 2 to 3 days. *E. coli* strains were grown at 37°C for approximately 24 hours on LB media.

Table I. **Bacterial Strains, Plasmids, and Antibiotics**

Strain/Plasmid	Relative Characteristics	Reference
Bacterial Strains		
R.etli		
CE110	*purY* mutant	Noel *et al.* 1984
CE170	*purL* mutant	Noel *et al.* 1984
CE382	*purL* mutant	Newman, unpublished
CE390	*purQ* mutant	Newman, unpublished
CE391	*purQ* mutant	Newman, unpublished
CE292	*purD* mutant	Shultz, unpublished
CE385	*purE,K* mutant	Newman, unpublished
E. coli		
TB1	susceptible to aquiring foreign DNA	New England Biolabs
DPWC	contains episomal TN1000 transposon	Strathmann *et al.* 1991
BW26	Kanamycin resistant	Strathmann *et al.* 1991
HB101	carried pRK2013 helper plasmid	Sambrook *et al.* 1989
Plasmids		
pRK2013	helper plasmid (tra+,oriT) Km^r	Figurski and Helinski, 1979
pRK415	broad-host-range vector (tra-,oriT) derived from pRK290, Tc^r	Keen *et al.* 1988
pCOS110	pLAFR1 derivivative containing the *R. etli purY, purQ, purL* region	Noel *et al.* 1988
pMOB	Amp^r,	Strathmann *et al.* 1991
pLC1	pRK415 containing *R.etli purL* promoter and *E.coli purL* gene	Present study
pBS	pBlueScript SK+, Amp^r cloning vector Blue-white screening	Stratagen

Antibiotics
Amp=Ampicillin
Cb=Carbenicillin
Km=Kanamycin
Nal=Nalidixic Acid
Str=Streptomycin
Tc=Tetracycline

Transformation

For transformations, competent cells were prepared by inoculating 7.5mL LB liquid with *E. coli* TB1 cells. After the cells had grown into log phase (2-3 hours) they were centrifuged, and the supernatant was discarded. The cells were then resuspended in 5-7 mL of 50mM $CaCl_2$ and placed on ice for 30 minutes. The cells were centrifuged again and resuspended in 0.5ml of 50mM $CaCl_2$. DNA was added to a sterile microfuge tube and placed on ice to keep cold until competent cells were added. Competent cells were added to the DNA and incubated on ice for 30 minutes. A two minute 42°C heat shock preceded the addition of 1mL LB liquid and a mix by inversion. The tube was laid on its side in a 37°C incubator for 1-2 hours without shaking. After incubation, the cells were centrifuged for 15 seconds at 14K rpm, the supernatant decanted, and cells were resuspended in residual supernatant. The entire mixture was spread onto an LB AMP plate containing X-GAL (5-bromo-4-chloro-3-indoyl-*B*-D-galactoside) (Sambrook *et al.* 1989). The X-GAL containing medium screens for the presence of an insert via α-complementation and shows colonies with inserts as white and those lacking inserts as blue. White colonies were patched on LBamp plates for future plasmid minipreps.

Plasmid minipreps - CTAB method

Plasmid DNA was extracted by the CTAB method of Del Sal *et al* (1989). Briefly, Cetyltrimethylammonium bromide (CTAB) is a cationic detergent that forms an insoluble complex with DNA at low concentrations of NaCl, while proteins and polysaccharides remain in solution. In this procedure 1.5 mL liquid culture was centrifuged and the pellet was resuspended in STET buffer(0.1% TritonX-100, 50mM EDTA, 20mM Tris-HCL, 0.23M sucrose). This buffer serves to disrupt the cell membrane and inhibits nuclease. Lysozyme was added, and the mixture was

boiled to lyse the cells. An immediate spin allows the lysed cell parts to be separated from the desired DNA. RNaseA was added to remove RNA followed by CTAB solution to separate DNA from unwanted proteins and polysaccharides. The DNA was resuspended in 1.2 M NaCl precipitated with 95% EtOH, and washed with 70% EtOH. The washed DNA was dried and resuspended in 50ul of water for later use (Del Sal *et al.*, 1989).

Complementation study

Polymerase Chain Reaction

PCR was set up for a 50uL reaction as follows: A master mix containing 5ul10x PCR buffer, 5uL 2.5 mM dNTPs, 22.3uL water, and 0.2ul Ex Taq Polymerase was prepared. Added to the master mix was 7.5uL of each primer (Table-2) and 2.5uL of template DNA. For the *R. etli* promoter a pBS clone containing the 1.7Kb fragment from pCOS110 (Figure-2) was used as a template. The template for *E. coli purL* was pJS80 (Schendel *et al*, 1989). Finally, 50uL of mineral oil was added to act as a seal to keep the contents of the PCR from evaporating. The program used for PCR was as follows:

	PCR50
Phase 1 - 1 cycle	
Initial denaturation	5 min. @ 94°C
Primer annealing	2 min. @ 50°C
Primer extension	4 min. @ 72°C
Phase 2 - 35 cycles	
Standard denaturation	1min. @ 94°C
Primer annealing	1min. @ 50°C
Primer extension	4 min. @ 72°C
Phase 3 - 1 cycle	
Standard denaturation	1 min. @ 94°C
Primer annealing	1 min. @ 50°C
Primer extension	10 min. @ 72°C

A 4 minute extension time was used to allow extra time for taq polymerase to copy the relatively large 3.9 Kb *E. coli purL* fragment. After PCR, phenol/chlorform extraction and ethanol precipitation was performed to rid the DNA of a enzymes and other unwanted components (Sambrook *et al.* 1989).

Table 2. Primers for PCR and DNA Sequncing

Primer Name	Sequence	Characteristics
EcpurL start	CCCTCTAGAATGATGGAAATTCTGGGTGG	Start end of *E. coli purL* with XbaI site Stop
EcpurL stop	CCCGGATCCTTACCCCAACTGCTTACGCG	end of *E. coli purL* with BamHI site
RepurL 3'	CCCTCTAGAGCGGAAAGTCCCTGAC	3' end of *R. etli purL* promoter with XbaI site
M13Rev	AACAGCTATGACCATG	Vector sequence flanking multi-cloning region
T7	GCGCGTAATACGACTCAC	Vector sequence flanking multi-cloning region
G186	ATATAAACAACGAATTATCTCC	TN1000 sequence included
G187	GTATTATAATCAATAAGTTATACC	TN1000 sequence included

Cloning

The cloning strategy to reach the experimental goal first involved subcloning of the *R. etli purL* promoter fragment from pCOS110 and moving it into pBS. To accomplish this, pCOS110 containing the *R. etli purL* locus was cut with EcoRI to obtain a variety of DNA fragments which were ligated into pBS cut with EcoRI. This was done by mixing 10ul of fragment DNA with 1ul pBS cut with EcoRI, 1.5ul 10x ligation buffer, 1.5ul water and 1ul T4 DNA ligase. The mixture was incubated at room temperature for 1 hour. In the map of pCOS110 (Fig-2) several BamHI and EcoRI sites can be seen. The most desirable fragment is the 1.7Kb fragment. This would yield the *R. etli purL* promoter with very little extraneous DNA. CTAB minipreps were performed, the DNA was cut with EcoRI, and the digest was gel electrophoresed to identify the clone containing the 1.7Kbfragment. This was then used as a template for PCR amplification using M13Rev and RepurL 3' primers (Table-2). The PCR products were cleaned by

phenol/chlorform extractions (Sambrook *et al.* 1989). Next, BamHI and XbaI were used to cut the PCR products, so that sticky ends would result. Using Prep-A-Gene matrix (Bio-Rad, 1995) the resulting fragments were gel purified to prepare the DNA for ligations with pBS cut with BamHI and XbaI. Ligations were set up with 10ul of insert DNA, 1.5ul of 10x ligation buffer, 1ul of pBS cut with BamHI and XbaI, 1.5ul water, and 1ul T4 DNA ligase. The mixture was incubated at room temperature for 1 hour. This was done to incorporate the *R. etli purL* promoter fragment into pBS, then the DNA was transformed into *E. coli* TB1 cells as described above. White colonies were chosen for CTAB plasmid minipreps and DNA from the minipreps was checked by BamHI and XbaI digests and gel electrophoresis to see if the 1.1Kb fragment containing the *R. etli purL* promoter was inserted. The same basic procedure was used to obtain the *E. coli purL* open reading frame (ORF).

Once both the 1.1Kb fragment containing the *R. etli purL* promoter and the 3.9Kb fragment containing the *E. coli purL* ORF were individually cloned into pBS, they were then cloned together in pBS. This was done by cutting the pBS clones BamHI and XbaI, gel purifying the fragments using the Prep-A-Gene matrix, and ligating them together with pBS cut with only BamHI. After the entire construct was in pBS, BamHI digests were performed to cut out the promoter:ORF fusion, and it was separated from the plasmid vector by gel electrophoresis and purified with Prep-A-Gene matrix. It was then ligated with pRK415 cut by BamHI. Transformation of the construct into TB1 was followed by CTAB minipreps, BamHI digests, and subsequent gel electrophoresis to determine whether cloning was successful. The construct containing the *R. etli purL* promoter and *E. coli purL* gene in pRK415 was designated as pLC1 (Figure-4).

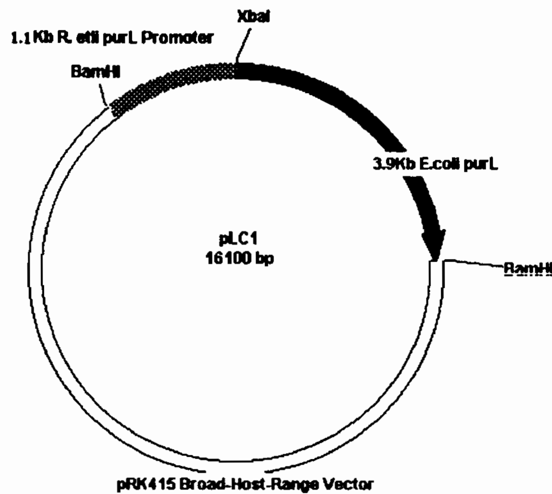


Figure 4. Map of pLC1

Conjugations

Transformation of pLC1 into TB1 was followed by a complementation experiment with several *R. etli* purine auxotrophs. One hundred microliters of TB1 culture containing pLC1 and 100ul pRK2013 containing culture were spread onto TY media. The plate was incubated at 30°C for 30 minutes to dry before spotting the plate with CE110(*purY* mutant), CE170(*purL* mutant), CE382(*purL* mutant), CE390(*purQ* mutant), CE391(*purQ* mutant), CE106(*purF* mutant), and CE385(*purEK* mutant). CE170, CE382, CE390, and CE391 were positive controls while CE106 and CE385 were negative controls. The plate was incubated at 30°C for one day. Cells were recovered from the plate, suspended in 1mL TY liquid, and transconjugants were selected by spreading 50ul of the suspension on TY plates containing Str, Km, and Nal to select the Rhizobium genetic background and Tc to select for the plasmid.

DNA Sequencing (TN1000 Method)

Subcloning of Fragments into pMOB for Sequencing:

A 1.0Kb fragment known to contain the *R. etli purY* gene and part of the *yjii* gene was previously cloned into pMOB by Kevin Ferguson. The 1.7Kb and 2.1Kb fragments containing the *R. etli purL* gene were cloned into pBS as described above. pCOS110 which is a derivative of the pLAFR1 plasmid (Friedman *et al.* 1982), contains a large portion of the operon containing *R. etli* purine genes. pBS clones containing the 1.7Kb and 2.1Kb fragments were cut with BamHI and EcoRI, or EcoRI alone, and the inserts were purified using Prep-A-Gene matrix (Bio-Rad, 1995). Two sets of enzyme digests were set up for the pMOB plasmid as follows:

10ul pMOB DNA	10ul pMOB DNA
2ul buffer	2ul buffer
1ul EcoRI	1ul EcoRI
7ul dH_20	1ul BamHI
	6ul dH_20

The 20uL digests were incubated at 37°C for 1 hour. To remove the enzymes, both digests were phenol/chloroform extracted (Sambrook et al, 1989). The pMOB plasmid DNA was finally resuspended in 10ul TE(10mM Tris-HCl pH 7.6 and 1mMEDTA pH 8.0). Once the DNA was prepared, ligations with 10ul of the purified fragments, 1ul of pMOB DNA, 1.5ul ligation buffer, 1.5ul water, and 1ul of T4 DNA ligase were set up so that the 1.7Kb and 2.1Kb fragments containing the *purL* gene were individually incorporated into pMOB. Next, transformations of pMOB into *E. coli* TB1 were performed. Recipient cells containing the recombinant pMOB plasmids were selected on LB amp plates. To confirm that pMOB contained the desired DNA fragments, plasmid minipreps were performed, and the DNA was cut with either EcoRI only or BamHI + EcoRI. Gel electrophoresis was used to detect whether the correct fragment was present.

Generation of Transposon Insertions within Cloned DNA

Mini-preps were made of the 1.0Kb, 1.7Kb, and 2.1Kb fragments in pMOB. Once prepared, the TN1000 technique of preparing DNA for sequencing was used (Strathmann et al. 1991). pMOB plasmids with the DNA fragment to be sequenced were transformed into the *E.coli* DPWC strain individually and selected for with LBamp plates. Two separate transformations were done for each of the three recombinant plasmids in case one of the chosen colonies contained a dimer plasmid. DPWC is the F+ donor host used in conjugation. Two colonies of DPWC containing the recombinant plasmid were grown in an overnight culture. Along with this, an overnight culture of *E. coli* BW26 was grown. BW26 is a kanamycin resistant F- recipient host. All three of the conjugations were set up the same way. Two mililiters of LB liquid was added to a 15mL centrifuge tube and prewarmed to 37°C. One hundred microliters of donor DPWC and 100uL of recipient BW26 was added to the LB. The tube was placed at an angle in a rotary shaker and rotated at 60 rpm at 37°C.

While conjugations were proceeding, miniprep DNA of the three recombinant plasmids was analyzed for dimers. This was done by cutting the DNA with an enzyme which cuts only once in the recombinant plasmid and running this DNA beside uncut DNA using gel electrophoresis. The linearized DNA will migrate more slowly than the monomer supercoil DNA. If the DNA is a dimer there will be two sites in the plasmid and the linearized DNA will migrate identically or even faster than the supercoil dimer. Dimers are unsuitable for conjugation experiments since transposon insertion in critical genes of the plasmid will not be lethal. This is due to the fact that they are diploid for the genes (Strathmann *et al.* 1991).

Conjugal mix was plated out at 2 hours and 4 hours after initiation of conjugation. Aliquots of 5uL and 50uL of the conjugal mix were added to either half of a kanamycin/carbenicillin plate to select for those colonies containing the recombinant plasmid with a transposon inserted in the target gene. The kanamycin selects for BW26 recipients, while carbenicillin selects for the pMOB plasmid. If the transposon inserts into either the ORI or the amp resistance gene, those cells will not survive. Also, if the cointegrate resolves itself before being transferred into BW26, it will not move the pMOB plasmid with it, and the recipients will not be carbenicillin resistant.

The next day isolated colonies were picked from each plate and patched onto another kanamycin/carbenicillin plate. To determine the site of transposon insertion, overnight cultures of the patches were grown and plasmid minipreps were performed. Miniprep DNA was cut with BamHI and gel electrophoresed. From the size of fragments produced, it could be determined approximately where the transposon was inserted.

Once the location of the transposon insertions was known, a nested set was assembled that would be used to generate overlapping sequence data. The nested set was arranged so that transposon insertions were 200-300bp apart.

Sequencing of the 1.0Kb, 1.7Kb, and 2.1Kb fragments containing purY and purL.

The DNA samples were cut with BamHI prior to setting up sequence reactions. After digestion, the DNA was then precipitated with sodium acetate and ethanol, and resuspended in water to an approximate concentration of 0.2ug/uL.

Sequence reactions were set up as follows. Template DNA was mixed with buffer, primer, Taq polymerase, and distributed among each of four tubes containing dideoxy nucleotides.

The reaction mixture was placed in a thermocycler with the following settings:

95°C for 2 minutes, then:

95°C for 30 seconds (denaturation)
42°C for 30 seconds (annealing)
70°C for 1 minute (extension)
60 cycles total for each set of reactions.

These reactions produced specific length DNA fragments by the dideoxy chain termination method. This method involves the use of 2',3'-dideoxynucleoside triphosphates (ddNTPs), which lack a 3'-hydroxyl group. The linearized DNA to be sequenced serves as a template strand. Each reaction has a low concentration of one of the four ddNTPs in addition to higher concentrations of the normal deoxynucleoside triphosphates (dNTPs). In each reaction, the ddNTP is randomly incorporated at the positions of the corresponding dNTP. This addition of ddNTP terminates polymerization because the absence of a 3' hydroxyl prevents addition of the next nucleotide. The mixture of terminated fragments from each of the four reactions can be separated by electrophoresis to read the entire DNA sequence.

A 5% polyacrylamide gel was used to run the sequencing reactions at 2000 Volts. A silver staining method by Promega (1996) was used to treat the gel. The gel was fixed in a 10% glacial-acetic acid solution. After being rinsed in purified water, the gel was stained in a silver nitrate solution. Silver ions in solution will bind to the DNA fragments in the gel. A quick rinse with purified water followed the silver staining before gels were developed using a mixture of formaldehyde, sodium carbonate, and sodium thiosulfate. These were all provided in the Promega Silver Sequencing kit. An equal volume of 10% glacial-acetic acid was added to stop development before a final rinse in purified water and subsequent drying. Dried gels were exposed to direct positive film for varying amounts of time, so that permanent records could be kept.

Sequencing gel photographs were interpreted and recorded into data files using the Lasergene Sequence analysis software package (DNAStar, 1996). Using the software, individual sequence data files were assembled into contiguous segments of DNA based on the nested set developed earlier. Gaps in sequence, and areas with only one strand sequence were filled in by additional sequence reactions and subsequent data collection. DNA sequence was translated to amino acid sequence using the software and entered into the National Center for Biotechnology Informations's Internet site (located at http://www/ncbi.nlm.nih.gov/BLAST). This site has a searching tool which takes the amino acid sequence in question and compares it to known amino acid sequences in a database. It statistically analyzes the homology of the sequence to amino acid sequences of other organisms. Open reading frames can be identified by looking at the high probability homologies.

Results

Complementation study

Before the Polymerase Chain Reactions (PCR) could be set up, primers had to be designed for the *R. etli purL* promoter and *E. coli purL* gene fragments. The primers used for the 5' end of the *R. etli purL* promoter was M13 reverse which corresponded to a sequence outside the multi-cloning region. In this way, the BamHI site of the pBS plasmid could be used for subsequent cloning. The 3' end primer (CCCTCTAGAGCGGAAAGTCCCTGAC) was designed with some sequence taken from the end of the *R. etli* promoter, an XbaI restriction site, and CCC. The cytosine rich end served to prevent the restriction enzyme from falling off the fragment. Primers for the start site end of the *E. coli purL* gene (CCCTCTAGAATGATGGAAATTCTGGGTGG) were also designed with an XbaI site 3' of a CCC and immediately 5' of the *purL* start codon. The primers

for the stop end of *purL* (CCCGGATCCTTACCCCAACTGCTTACGCG) incorporated some known sequence to allow annealing, a BamHI site, and CCC. The overall design allowed BamHI sites to be included on the extreme ends of the fragments while XbaI sites were incorporated onto the ends to be fused during ligation of the two fragments.

The 1.1Kb fragment was easily PCR amplified, but problems arose when trying to PCR amplify the 3.9Kb fragment. In Figure-5 it can be seen that the 1.1Kb fragment appears in high yields while the 3.9Kb fragment is only generated in very low yields. The plasmid, pJS80 was used to obtain the *E. coli* 3.9Kb ORF. Several PCR trials failed before a new *Taq* polymerase was used called ExTaq (TaKaRa). It is a modern DNA polymerase used to replace conventional Taq polymerase. ExTaq is supposed to produce greater processivity for PCR, have higher fidelity, and produce greater yields. Several concentrations of DNA were tested with ExTaq, bringing high yields in all cases (Figure-5).

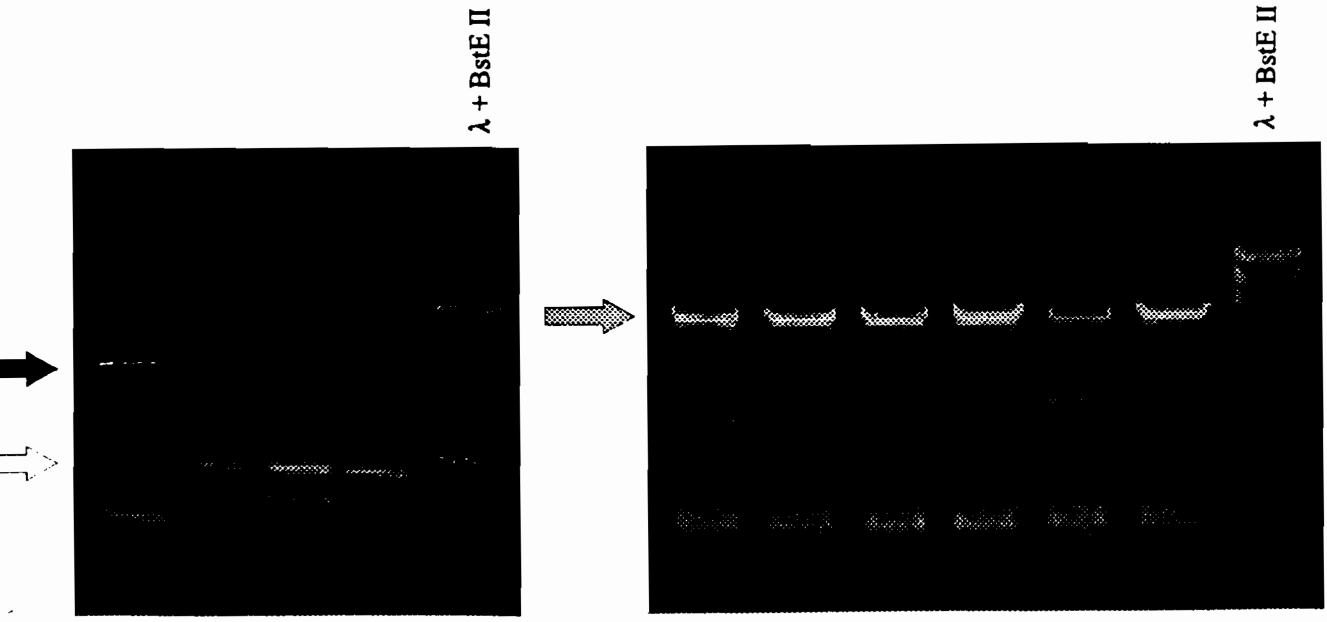


Figure -5. The white arrow indicates the 1.1Kb fragment PCR amplified, while the black arrow indicates a faint band for the 3.9Kb fragment. After using ExTaq with several concentrations of DNA, the 3.9Kb fragment was produced in high yields, as indicated by the grey arrow.

After both the 1.1Kb and 3.9Kb fragments were PCR amplified they were cut with BamHI and XbaI, gel purified, and ligated individually into pBS. The fragments were then cloned together in pBS, and finally cloned into the broad-host-range vector pRK415. This construct was designated as pLC1 and transferred into several purine auxotrophs which were subsequently tested for their ability to grow without supplementation (Table-3). The auxotrophs were mutated in one of the following genes: *purY*, *purQ*, *purL*, *purF*, and *purEK*.

Table-3. Complementation.

Bacterial Strain	**Purine Auxotroph**	**Complementation**
CE110	*purY*	+
CE390, CE391	*purQ* (positive control)	+
CE170, CE382	*purL* (positive control)	-
CE106	*purF* (negative control)	-
CE385	*purE,K* (negative control)	-

Recombinant plasmid, pLC1, carrying the *R. etli* promotor and the *E. coli* *purL* gene was conjugated into several strains of *R. etli* purine auxotrophs. + or - represent either a positive or negative complementation of the metabolic defect.

As expected, both the *purY* and *purQ* mutants were complemented by pLC1 while the *purF* and *purE,K* mutants were not. However, *purL* mutants were not corrected for even though current knowledge indicates that pLC1 should correct the metabolic defect.

Sequencing of R. etli purY and purL

Before DNA was prepared for sequence reactions, the first step in the process was to map the location and orientation of the transposon within the insert of each clone. In order to do this, BamHI restriction digests of the plasmid DNA were run on an agarose gel next to a Lambda marker, and photographed. The position of BamHI sites of the vector, gene, and transposon were all known, so that the location, and in most cases, the orientation of transposon insertion could be determined. Below is a sample of some of the results for transposon insertion into the 1.0Kb Eco-Bam fragment. The same method was used for the 1.7Kb and 2.1Kb fragments.

Sample Identification	Approximate Location of transposon from BamHI site
EB1.0-7	~200bp
EB1.0-6	~300bp
EB1.0-40	~550bp
EB1.0-11	~800bp
EB1.0-29	~900bp

After the sites of transposon insertion were determined, sequence reactions were set up in a manner that would cover the entire DNA fragment to be sequenced. A map was generated to easily visualize transposon insertions. The map allowed us to determine which sample would be beneficial to sequence, which primers to use, as well as to estimate where overlapping sequence would occur. Attempts were made to sequence the 1.0Kb EcoRI-BamHI, 1.7Kb EcoRI-EcoRI, and 2.1Kb EcoRI-BamHI fragments in order to get sequence data for the *R. etli purY* and *purL* genes. Sequencing of the 1.0Kb EcoRI-BamHI fragment produced 100% single-strand coverage and 88% double-strand coverage of the entire fragment.

Next, the nucleotide sequence was translated into amino acid sequence and a BLAST search with the amino acid sequence was performed. The BLAST search showed amino acid

sequence homology of the 1.0Kb fragment with several other organisms. Statistically significant matches were found with the *purY* gene in several other organisms were revealed. In addition, it was discovered that the 3' end of *purC* is located just upstream of *purY* while the 5' end of *yjii* (Ferguson, 1996) is located just downstream of *purY*. Figure-6 below shows a map representing sequence generated by individual sequence reactions for the 1.0KB fragment. It also shows the locations of *purC*, *purY*, and *yjii* in relation to one another. Arrows indicate the 5' to 3' direction of sequence, so opposing arrows would indicated coverage on opposite DNA strands.

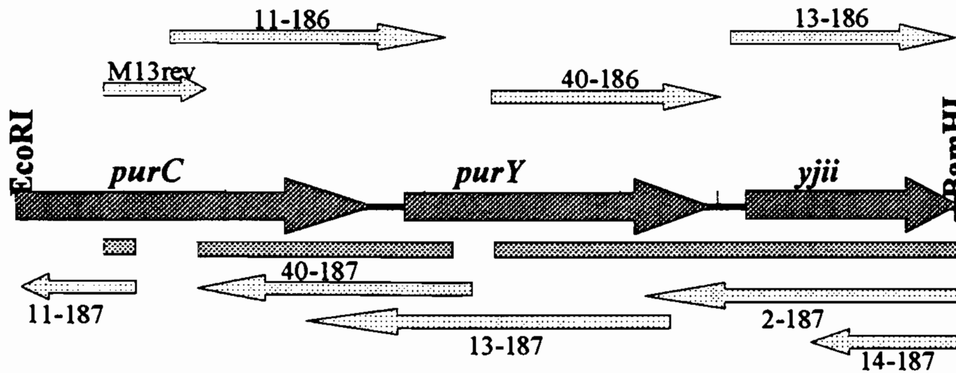


Figure-6. M13rev, -186, and -187 stand for the primer used for the sequence reaction. The arrows resting on the line indicate an open reading frame, while all other arrows indicate the direction and length of coverage by individual sequence reactions.

DNA sequence from the 1.0Kb fragment containing *purC*, *purY*, and *yjii* was entered into our sequencing software along with Kevin Ferguson's DNA sequence data from the 2.5Kb fragment containing *purQ* (Figure-2). The software joined both sets of sequence data. This sequence was then translated into amino acid sequence (Figure-7) Genes encoded by the sequence are included.

There is little sequence data for the 2.1Kb EcoRI-BamHI fragment and none generated for the 1.7Kb fragment due to several problems encountered throughout the semester. While trying to clone the 2.1Kb fragment, dimers were discovered which had to be resolved before sequencing. Transposon insertions were finally successful for both the 1.7Kb and 2.1Kb fragments, and insertions maps were made for both. In addition to cloning problems, many problems occurred with the sequence reactions. It took several sets of sequence reactions before the problem was solved. After plasmid minipreps it was found that the DNA should be thoroughly cleaned using phenol/chloroform extraction (Sambrook *et al.* 1989) The problems encountered used up a great deal of time, and as a result, only limited data was collected.

```
GAATTCTACTACAAATCCGATGCGCTCGACGATCCGATGGTCTTCGAAGAGCATATCACCGCTTTCGGCTGGGCCAATCC  80
 E  F  Y  Y  K  S  D  A  L  D  D  P  M  V  F  E  E  H  I  T  A  F  G  W  A  N  P
|------------------------------------ purC ------------------------------------
GGCCGAGCTCGACGACATCAGTGGCCCTTCGACCATCCTGAGTCACGGCTTTCAATCACTCCATGGCCTCTTCCTGGGTG  160
  A  E  L  D  D  I  S  G  P  S  T  I  L  S  H  G  F  Q  S  L  H  G  L  F  L  G
------------------------------------- purC ------------------------------------
TCGGCATCCAGCTCGTCGTCTTCAAGCTCGTTGTCGGCCGGCTCTTCGAAGGCGACATGATGCGGATCATCCTCGCGGAC  240
V  G  I  Q  L  V  V  F  K  L  V  V  G  R  L  F  E  G  D  M  M  R  I  I  L  A  D
------------------------------------- purC ------------------------------------
GAAATCTCGCCCGACAGCTGCCGGCTCTGGGATATCGAAACCCGCGAGAAGATGGACAAGGACCGCTTCCGCCGCGATCT  320
 E  I  S  P  D  S  C  R  L  W  D  I  E  T  R  E  K  M  D  K  D  R  F  R  R  D  L
------------------------------------- purC ------------------------------------
CGGCGGCTTGCTCGAAGCCTATTCCGAAGTCGCACGTCGTCTCGGCATCATCAATTAAAACGAGCCTTGTGCGCGGCACC  400
  G  G  L  L  E  A  Y  S  E  V  A  R  R  L  G  I  I  N  .
------------------------- purC -------------------------|
GGCCCGGTTCTCGTCAAGTAAGGCAGGAAAGACAAAGTGATCAAGGCTCGTGTCACCGTCACGCTGAAAAACGGCGTTCT  480
                                    V  I  K  A  R  V  T  V  T  L  K  N  G  V  L
                                   |------------------ purY -------------------
CGATCCGCAGGGCAAGGCCATTGAAGGCGCGCTCGGCGCCCTCGGCTTTTCGGGCGTCGGCCATGTAAGGCAAGGCAAGG  560
  D  P  Q  G  K  A  I  E  G  A  L  G  A  L  G  F  S  G  V  G  H  V  R  Q  G  K
------------------------------------- purY ------------------------------------
TCTTCGACCTGGAGCTCGAAGGCGCCGACAAGGCCATGGCCGAGGCAGACCTCAAGGCGATGTGCGAAAAACTACTCGCC  640
V  F  D  L  E  L  E  G  A  D  K  A  M  A  E  A  D  L  K  A  M  C  E  K  L  L  A
------------------------------------- purY ------------------------------------
AACACGGTTATCGACGTATTCCGATCGCAGATCGACTGACCATTGCAGACCTGCCCGTCATCCGGCGGGTCTTATGGATT  720
 N  T  V  I  D  V  F  R  S  Q  I  D  .
----------------- purY ----------------|
GAGCGTCATGCTGAAGGCAAAGCTGGAGACGGAAGTCTCGAAATTCATGAGCTATGTTTTGCGCCATGCGCCAGATGCCG  800
       M  L  K  A  K  L  E  T  E  V  S  K  F  M  S  Y  V  L  R  H  A  P  D  A
      |------------------------------- YJII -------------------------------
CAGGCTTGACGCTTGATGCCGAGGGTTGGGTATCGTTCGATGATCTCGAAAAGGTGCTGACGTCGAAATACGGCGTCTCC  880
A  G  L  T  L  D  A  E  G  W  V  S  F  D  D  L  E  K  V  L  T  S  K  Y  G  V  S
------------------------------------- YJII ------------------------------------
CGCGCCGATATCGTCGAGATCATCGACAACAGTCCGAAGAAACGCTTCACGCTTGTCGATAACAGGATCCGCGCCAACCA  960
 R  A  D  I  V  E  I  I  D  N  S  P  K  K  R  F  T  L  V  D  N  R  I  R  A  N  Q
------------------------------------- YJII ------------------------------------
GGGCCACAGTGTCGAGGTTGATCTGGCATTGACCCCGGTGGAACCGCCGGCAGTTCTCTTTCACGGCACGTCACTCACGA  1040
  G  H  S  V  E  V  D  L  A  L  T  P  V  E  P  P  A  V  L  F  H  G  T  S  L  T
------------------------------------- YJII ------------------------------------
ACTGGCCGTCGATCGCGCGCGAAGGCCTGAAGAAGATGGAGCGGCACCACGTCCATCTGTCGGCCGACATCGATACGGCG  1120
N  W  P  S  I  A  R  E  G  L  K  K  M  E  R  H  H  V  H  L  S  A  D  I  D  T  A
------------------------------------- YJII ------------------------------------
AAAATTGTCGCCGTGCGCCGCAAGGGCGACTATATCATTCTGCGTGTCGACGCAGCCGCCATGTTTTCAGAAGGCCATTC  1200
 K  I  V  A  V  R  R  K  G  D  Y  I  I  L  R  V  D  A  A  A  M  F  S  E  G  H  S
------------------------------------- YJII ------------------------------------
TTTCTTTGTCGCGGACAATGGAGTATGGCTCGCCGAGAGCGTTCCAGTCCAATATCTTTCGCGAAATGCGGGGACCCTAT  1280
  F  F  V  A  D  N  G  V  W  L  A  E  S  V  P  V  Q  Y  L  S  R  N  A  G  T  L
------------------------------------- YJII ------------------------------------
```

Figure-7. Amino Acid Sequence of *purC*, *purY* and *yjii*

Discussion

Complementation Study

According to the hypothesis that *purY* encodes a third subunit of FGAR amidotransferase, conjugations of pLC1 into different purine auxotrophs should have corrected for mutations in *purY*, *purQ*, and *purL*. Purine auxotrophs with mutations in *purY* and *purQ* were complemented by pLC1, but *purL* mutants were not. Since *purY* was complemented, the original hypothesis was accepted. However, further studies will have to be conducted in order to find a possible explanation for the negative result for *purL* complementation. One explanation for this is the possibility of another purine biosynthetic gene located immediately downstream of *purL*. If there is another ORF, it would cause a negative result since the mutation in *purL* would disrupt any downstream genes. A gene may even overlap part of the *purL* open reading frame. To find out if there is another open reading frame, more DNA sequence of *R. etli purL* and regions immediately downstream will have to be obtained. Open reading frames can be identified through the use of sequencing software and BLAST searches on the Internet.

The distribution of bacteria containing the single subunit (type I) form of FGAR amidotransferase and those bacteria containing the multi-subunit (typeII) form of FGAR amidotransferase supports the present hypothesis. Complete genomes of several organisms have revealed that when an organism contains the type II FGAR amidotransferase, it also contains *purY*. Organisms that contain the type I enzyme, however, do not contain *purY* within their genome. For example, *Synechocystis* (cyanobacteria), *Methanococcus janaschii* (Archaeon), *Bacillis subtilis*, and *R. etli* all have a type II enzyme and contain *purY* within their genomes. *Haemophilus influenzae*, like *E. coli*, is a member of the γ-group of proteobacteria and does not

contain *purY*. It is unlikely that *E. coli* contains *purY*, but the entire genome has not been sequenced yet. Yeast also do not have *purY* within their genome. In addition, *Mycoplasma gentalium* has no purine biosynthetic genes and has no *purY*. Collectively, this supports the hypothesis that *purY* plays a role in the type II multi-subunit form of FGAR amidotransferase.

<u>*Sequencing of R. etli purY and purL*</u>

The amino acid sequences for *purY*, *yjii*, and *purC* were entered into software to compare their sequences with all amino acid sequences in the Genbank databases (Figure-8, 9, 10). Several organisms showed high levels of amino acid similarity to these genes. For *purY*, amino acid sequence of *R. etli* showed 41% and 44% identity to *B. subtilis* and *Synechocystis* respectively (Figure-8). *E. coli* and *Saccharomyces cerevisiae* respectively showed 33% and 22% amino acid sequence identity to the *R. etli yjii* gene. *R. etli purC* showed 44%, 33%, and 40% amino acid sequence identity to *B. subtilis*, *Streptococcus pneumoniae,* and *E. coli* respectively. These results suggest that the amino acid sequence for these purine biosynthetic genes have remained highly conserved throughout evolution.

It is unfortunate that more data was not obtained for the *purL* gene. Further sequencing is planned for this semester to generate more *purL* DNA sequence. When this is accomplished the sequence can be analyzed to answer some of the pending questions about the possibility of another gene or genes immediately downstream of *purL*.

```
1   M I K A - - - - - - - - - - - - R V T V T L K N G V L D P Q G K A I E G A L G A  RETLPURY.PRO
1   M - - - - - - - - - - - - Y K V K V Y V S L K E S V L D P Q G S A V Q H A L H S  BSUBPURY.PRO
1   M A D S P V R P S M S H S Y H C R I Y V T L R P S V L D P A G T A V Q S G L Q Q  SYNEPURY.PRO
1   M - - - - - - - - - - - - Y K A T V I I K L K K G V L N P E G R T I Q R A L N F  MJANPURY.PRO

29  L G F S G V G H V R Q G K V F D L E L E G A D K A M A E A D L K A M C E K L L A  RETLPURY.PRO
29  M T Y N E V Q D V R I G K Y M E L T I E K S D R D L - D V L V K R M C E K L L A  BSUBPURY.PRO
41  L G Y D G V S Q V R I G K Y I E L T L E A P D E A T A S Q Q L D T M C D Q L L A  SYNEPURY.PRO
29  L G F N N V K E V Q T Y K M I D I I M E G E N E E K V K E E V E E M C K K L L A  MJANPURY.PRO

69  N T V I D V F - - - - - - - R S Q I D .                                          RETLPURY.PRO
68  N T V I E D Y R Y E V E E V - - - V A Q                                          BSUBPURY.PRO
81  N T V I E N Y C F E I T A L E G A V T P                                          SYNEPURY.PRO
69  N P V I H D Y E I K V E K I E                                                    MJANPURY.PRO
```

Figure-8. Alignment of *purY* amino acid sequences.
Shaded residues indicate amino acids. RETLPURY = *R. etli*; BSUBPURY = *B. subtilis*; SYNEPURY = *Synechocystis sp.*; MJANPURY = *M. janaschii*.

```
1   M L K - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - A  RETLYJII.PRO
1   M W K R L C A G L S C R K F I N Q V T M T S V N K N R K S R D E N E M A K Y N E  ECOYJII.PRO
1   M R Q V L - - - - - - - - - - - - - - - - - - Q K D K R - - - - - - - - - - - -  SCERYJII.PRO

5   K L E T E V S K F M S Y V L R H A P D A A G L T L D A E G W V S F D D L E K V L  RETLYJII.PRO
41  K E L A D I S K F L S F V L R H K P E A I G I V L D R E G W A D I D K L - - I L  ECOYJII.PRO
11  - - D V Q L S K A L S Y L L R H T A V K E K L T I D S N G Y T P L K E L - - L S  SCERYJII.PRO

45  T S K Y G - - - V S R A D I V E I I D N S P K K R F - - - T L V - D N R - I R A  RETLYJII.PRO
79  C A Q K A G K R L T R A L L D T V V A T S D K K R F - - - S Y S S D G R C I R A  ECOYJII.PRO
47  H N R L K T H K C T V D D I H R I V K E N D K Q R F H I K T L G A D E E W I C A  SCERYJII.PRO

77  N Q G H S V - E V D - - - - - - - L A L T P V E P - P A V L F H G T S L T N W P S  RETLYJII.PRO
116 V Q G H S T S Q V A - - - - - - - I S F A E K T P - P Q F L Y H G T A S R F L D E  ECOYJII.PRO
87  T Q G H S I K S I Q P S D E V L V P I T E A S Q L P Q E L I H G T N L Q S V I K  SCERYJII.PRO

109 I A R E G - L K K M E R H H V H L S A D I D T A K I V - - A V R R K G D Y I I L  RETLYJII.PRO
149 I K K Q G - L I A G E R H Y V H L S A D E A T A R K V - - G A R H - G S P V I L  ECOYJII.PRO
127 I I E S G A I S P M S R N H V H L S P G M L H A K G V I S G M R S S S N V Y I F  SCERYJII.PRO

146 R V D A A A M F S E G H S F F V A D N G V W L A E S V F V Q Y L S R - - - - - -  RETLYJII.PRO
185 T V K A Q E M A K R G L P F W Q A E N G V W L T S I V A V E F L E - - - - - - -  ECOYJII.PRO
167 I D C H S P L F F Q T L K M F R S L N N V Y L S S S I P V E L I Q K V V V K G N  SCERYJII.PRO

180 - - - - - N A G T L .                                                            RETLYJII.PRO
218 - - - - - - - - - - W                                                            ECOYJII.PRO
207 L K D E E K L D T L R R I L H E R N I P L E K I                                  SCERYJII.PRO
```

Figure-9. Alignment of *yjii* amino acid sequences.
Shaded residues indicate identical amino acids. RETLYJII = *R. etli*; ECOLYJII = *E.coli*; SCERYJII = *S. cerevisiae*.

```
1   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   RETLPURC.PRO
1   M N I V K N E L L Y E G K A K K I Y K T D D E N T L Y V V Y K D S A T A F N G E   BSUBPURC.PRO
1   M S - - - K Q L I Y S G K A K D I Y T T E D E N L I I S T Y K D Q A T A F N G V   SPNEPURC.PRO
1   M Q - - K Q A E L Y R G K A K T V Y S T E N P D L L V L E F R N D T S A G D G A   ECOLPURC.PRO

1   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   RETLPURC.PRO
41  K K A E I S G K G R L N N E I S S L I F K H L H A K G I N N H P I E R I S E T E   BSUBPURC.PRO
38  K K E Q I A G K G V L N N Q I S S F I F E K L N A A G V A T H F V E K L S D T E   SPNEPURC.PRO
39  R I E Q F D R K G M V N N K F N Y F I M S K L A E A G I P T Q M E R L L S D T E   ECOLPURC.PRO

1   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   RETLPURC.PRO
81  Q L I K K V T I V P L E V V V R N V V A G S M S K R L G I P E G T E L E Q P I I   BSUBPURC.PRO
78  Q L N K K V K I I P L E V V L R N Y T A G S F S K R F G V D E G I A L E T P I V   SPNEPURC.PRO
79  C L V K K L D M V P V E C V V R N R A A G S L V K R L G I E E G I E L N P P L F   ECOLPURC.PRO

1   E F Y Y K S D A L D D P M V F E E H I T A F G W A N P A E L D D I S G P S T I L   RETLPURC.PRO
121 E F Y Y K D D A L G D P L I T E D H I W L L K A A T P E Q V E T I K S I T T I V   BSUBPURC.PRO
118 E F Y Y K N D D L D D P F I N D E H V K F L Q I A D D Q Q I A Y L K E E A R R I   SPNEPURC.PRO
119 D L F L K N D A M H D P M V N E S Y C E T F G W V S K E N L A R M K E L T Y K A   ECOLPURC.PRO

41  S H G F Q S L H G L P L G V G I Q L V V F K L V V G R L F E G D M M R I I L A D   RETLPURC.PRO
161 N - - - E E L Q S I F D D C H V R L I D F K L E F G L D A E G Q - - - V L L A D   BSUBPURC.PRO
158 N - - - E L L K V W F A E I G L K L I D F K L E F G F D K D G K - - - I I L A D   SPNEPURC.PRO
159 N - - - D V L K K L F D D A G L I L V D F K L E F G L Y K - G E - - - V V L G D   ECOLPURC.PRO

81  E I S P D S C R L W D I E T R E K M D K D R F R R D L G G L L E A Y S E V A R R   RETLPURC.PRO
195 E I S P D T C R L W D K E T N E K L D K D L F R R N L G S L T D A Y E E I F N R   BSUBPURC.PRO
192 E F S P D N C R L W D A D G N H - M D K D V F R R G L G E L T D V Y E I V W E K   SPNEPURC.PRO
192 E F S P D G S R L W D K E T L E K M D K D R F R Q S L G G L I E A Y E A V A R R   ECOLPURC.PRO

121 L G I I N .                                                                     RETLPURC.PRO
235 L G G I H H V                                                                   BSUBPURC.PRO
231 L Q E L K                                                                       SPNEPURC.PRO
232 L - G V Q L D                                                                   ECOLPURC.PRO
```

Figure-10. Alignment of purC amino acid sequences.
Shaded residues indicate identical amino acids. RETLPURC = *R. etli*; BSUBPURC = *B. subtilis*; SPNEPURC = *Streptococcus pneumoniae*; ECOLPURC = *E. coli*.

Works Cited

Cook, Jonathan, Kevin E. Ferguson, and Jeffrey D. Newman. Analysis of Rhizobium etli Purine Biosynthetic Gene of Unknown Function. Molecular Plant-Microbe Interactions Absract Book. Abstract H-94, 1996.

Del Sal G., Manifioletti, G., and Schneider, C. (1989) The CTAB-DNA Precipitation Method: A Common Mini-Scale Preparation of Template DNA From Phagemids, Phages, or Plasmids Suitable for Sequencing. Biotechniques. 7:514-520.

Ebbole, D.J. and Zalkin, H. (1987) Cloning and characterization of a 12-gene cluster form *Bacillius subtilis* encoding nine enymes for de novo purine synthesis. *J. Biol. Chem.* 262: 8274-8287.

Figurski, D.H., Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on plasmid function provided in trans. Proc. Natl. Acad. Sci. USA 76:1648-1652.

Newman, J.D., Diebold, R.J., Schultz, B.W. and Noel, K.D. (1994) Infection of sybean and pea nodules elicited by *Rhizobium* spp. purine auxotrophs in the presence of 5-aminoimidazole-4-carboxamide riboside. *J. Bacteriol.* 176: 3286-3294.

Newman, J.D., Rosvitz, M.J. and Noel, K.D. (1995) Requirement for rhizobial production of 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) for infection of bean. *Mol. Plant-Microbe Interact.* 8:407-414.

Newman, J.D., Schultz, B.W. and Noel, K.D. (1992) Dissection of nodule development by supplementation of *Rhizobium leguminosarum* biovar *phaseoli* purine auxotrophs with 4-aminoimidazole-5-carboxamide riboside. *Plant Physiol.* 99:401-408.

Noel, K.D., Diebold R.J., Cava, J.R. and Brink, B.A. (1988) *Rhizobial* purine and pyrimidine auxotrophs: nutrient supplementation, genetic analysis, and the symbiotic requirement for *de novo* purine biosynthesis. *Arch. Microbiol.* 149:499-506.

Quiagen Prep-A-Gene matrix Protocol. 1995.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular cloning: A laboratory manual. 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Sampei, G. and Mizobuchi, K. (1989) The organization of the *purL* gene encoding 5'-phosphoribosylformylglycinamide amidotransferase of *Escherichia coli*. *J. Biol. Chem.* 264: 21230-21238.

Schendel, F.J., Mueller, E., Stubbe, J., Shiau, A. and Smith, J.M. (1989) Formylglycinamide ribonucleotide synthetase from *Escherichia coli*: cloning, sequencing, overproduction, isolation, and characterization. *Biochemistry* 28: 2459-2471.

Strathmann, M., Hamilton, B.A., Mayeda, C.A., Simon, M.I., Meyerowitz, E.M. and Palazzolo, M.J. (1991) Transposon-facilitated DNA sequencing. *Proc. Natl Acad. Sci.* USA 88: 1247-1250.

TaKaRa Shuzo Co., LTD Ex Taq.

VandenBosch, K.A., Noel, K.D., Kaneko, Y. and Newcomb, E.H. (1985) Nodule initiation elicited by noninfective mutants of *Rhizobium phaseoli*. *J. Bacteriol.* 162: 950-959.

Zalkin, H. and Dixon, J.E. (1992) De novo purine nucleotide biosynthesis. *Prog. Nucl. Acid Res. and Mol. Biol.* 42: 259-287.