Genetic Modification of the *Escherichia coli* Strain DH5α to Allow the Selection of Plasmids Carrying Complementary Yeast Genes

Malachi Griffith

A thesis submitted in partial fulfillment of the requirements for,

Projects in Biology 05.4111/6

Supervised by Dr. R. D. Gietz

April 2, 2001

Abstract

The *Escherichia coli* strain DH5α is used extensively in recombinant DNA technology. This is due in part to the *endA* gene mutation, which results in the reduction of endogenous levels of nuclease activity and increases the yield and quality of purified plasmid molecules. However, DH5α lacks the *pyrF*, *leuB*, and *trpC* mutations and is therefore not convenient for the selection of plasmids carrying the yeast *LEU2*, *TRP1* and *URA3* genes that are commonly used in the two-hybrid system.

A novel method of disrupting *E. coli* genes was used to create *pyrF*, *leuB*, and *trpC* gene disruptions in the strain DH5α. The method utilizes the polymerase chain reaction (PCR) to create a gene disruption cassette. This cassette contains the kanamycin resistance gene flanked by FRT (FLP recognition target) sites, in turn flanked by DNA sequences homologous to the gene targeted for disruption. The gene disruption cassette was transformed into DH5α cells containing the helper plasmid pKD46, which has the λ Red recombinase functions for stimulation of homologous recombination at the target site. Gene disruptions were selected by kanamycin resistance and identified by screening for uracil, leucine, or tryptophan auxotrophy. The kanamycin resistance gene was then removed by induction of recombination between the flanking FRT sites through expression of the FLP recombinase from a second helper plasmid pCP20. This procedure allowed the sequential disruption of the *pyrF*, *leuB* and *trpC* genes in DH5α creating a new strain, MG107. This new strain can be used to select for plasmids containing the yeast *URA3*, *LEU2*, and *TRP1* genes and produces high quality plasmid DNA preparations.

Acknowledgements

I would first like to thank Dan Gietz for being the best supervisor a student could ask for. You consistently exceeded my expectations and I consider myself lucky to have had the opportunity of working with you. Thanks to my committee members Dr. Byard and especially Dr. Woods for referring me to Dan Gietz in the first place and providing me with advice on scientific writing. Special thanks to Dr. Moodie for his patience and consideration throughout the year. I would also like to thank Barry Wanner for providing some of the bacterial strains for this project, Shun Zhen Zhang for her technical advice and training, and Kevin Struhl for his communications. Finally I would like to thank Obi and Tabitha for their support and understanding throughout the year.

Table of Contents

Abstract ii
Acknowledgements iii
Table of Contents iv
List of Tables vi
List of Figures vi
Abbreviations vii
List of Appendices vii
Introduction 1
Escherichia coli and Recombinant DNA 1
The Yeast Two-hybrid System 1
PCR Generated Homologous Gene Disruption 6
Escherichia coli Strain DH5α 9
endA1 10
recA1 10
deoR 11
gyrA96 11
lacZΔM15 11
Objectives of the Project 11
Materials and Methods 13
Bacterial Strains 13
Plasmids 13
Selective Media 16
Minimal Media 16
Solutions 17
Plasmid DNA Preparation 18
Restriction Enzyme Digestions 18
PCR Primers 19
PCR Amplification 20
Agarose Gel Electrophoresis 20
Gel Purification and Band Elution 21
Preparation of Electrocompetent *E. coli* 21
Preparation of Electrocompetent, Recombination Proficient *E. coli* 22
Electroporation 22
Gene Disruption 22
Removal of the Kanamycin Resistance Gene 23
Phenotype Tests 23
Antibiotic Resistance 23
Auxotrophy 23
lacZ Activity 24
EndA Activity 24
Results 25
Confirmation of the Structure of the Template and Helper Plasmids 25
PCR Primer Design 25
Preparation of PCR Gene Cassettes 26

Verification of the Presence of the Helper Plasmid pKD46 27
The Results of Gene Disruption 28
Disruption of the *pyrF* gene in DH5α 29
Disruption of the *leuB* gene in MG102A (*pyrF*Δ*168*) 32
Disruption of the *trpC* gene in MG104A (*pyrF*Δ*168*, *leuB*Δ*211*) 32
Confirmation of the Genotyped MG107 33
Auxotrophic Analysis 33
Assay for Plasmid DNA Quality 34
Test for LacZ Alpha Complementation 35
Discussion 37
The Method of Gene Disruption 37
Genotyped MG107 39
Comparison of KC8, DH5α, and MG107 40
Future Directions 42
References 45

List of Tables

Table 1: Bacterial Strains……...……13

Table 2: The Sequence of Primers Used to Generate Gene Disruption Cassettes….……19

Table 3: A Summary of the Phenotypic Test Results……....41

List of Figures

Figure 1: The Yeast Two-Hybrid System……...……...….3

Figure 2: Yeast Two-Hybrid Vectors Carrying *TRP1* and *LEU2* Markers…….5

Figure 3: Gene Disruptions Occur by Homologous Recombination…….…..7

Figure 4: Features of the Template Plasmid Used to Generate the Gene Cassettes……9

Figure 5: Template and Helper Plasmids Used in the Gene Disruption System……...15

Figure 6: An Agarose Gel of the *leuB*-pKD4 PCR Product Prior to Purification…….27

Figure 7: Verification of the Presence of pKD46 in Electroporated Bacterial Cells…....28

Figure 8: An Overview of the Process of Gene Disruption……...31

Figure 9: An Assay of the Endonuclease Activity in MG107 and KC8……35

Figure 10: An Assay of LacZ Activity in Three Bacterial Strains……36

Abbreviations

amp – Ampicillin
amp^R – Ampicillin resistance
bla – Ampicillin resistance gene
CGSC – The *E. coli* Genome Stock Center
deoR – *E. coli deoR* gene
DNA – Deoxyribonucleic acid
endA – *E. coli endA* gene
$GAL4_{AD}$ – Yeast Gal4 Activating Domain
$GAL4_{BD}$ – Yeast Gal4 Binding Domain
gyrA – *E. coli gyrA* gene
HIS3 – Yeast *HIS3* gene
hisB – *E. coli hisB* gene
kan – Kanamycin or kanamycin resistance gene
kan^R – Kanamycin resistance
lacZ – *E. coli lacZ* gene
LB – Luria Bertani Medium
LEU2 – Yeast *LEU2* gene
leuB – *E. coli leuB* gene
M9 – Minimal medium containing M9 salts.
MCS – Multiple cloning site
PCR – Polymerase Chain Reaction
pyrF – *E. coli pyrF* gene
recA – *E. coli recA* gene
relA – *E. coli relA* gene
TE – Tris EDTA
TRP1 – Yeast *TRP1* gene
trpC - *E. coli trpC* gene
UAS – Upstream activating sequence
URA3 – Yeast *URA3* gene
Y2H – Yeast two-hybrid

List of Appendices

Appendix A: Summary of BLAST Results for *pyrF* and *leuB* Primers

Introduction

Escherichia coli and Recombinant DNA

E. coli has been used as the host organism for cloning DNA segments for many years. Many of the genetic markers in strains of *E. coli* are used for specific cloning strategies or DNA amplification purposes. As recombinant DNA technology evolves, novel bacterial strains are needed to serve the purposes of new strategies. Fields and Song (1989) published the first account of the yeast two-hybrid system. This system uses two different yeast-*E. coli* shuttle vectors to detect protein-protein interactions. Each of these vectors contains a different gene for selection in yeast as well as an antibiotic resistance gene for selection in *E. coli*. Recombinant plasmids are identified in yeast but transformed back into *E. coli* for further analysis.

Certain yeast genes can complement mutations in bacteria (Ratzkin and Carbon, 1977). The bacterial strain KC8 contains such mutations to take advantage of the different genetic markers found on yeast two-hybrid vectors (Struhl, 2000), however it does not produce high quality plasmid DNA preparations (Gietz, 2001). To overcome this problem a new strain was produced from DH5α, a strain used extensively for recombinant DNA technology. To generate the new strain, MG107, mutations corresponding to a number of yeast genes were created using the method developed by Datsenko and Wanner (2000).

The Yeast Two-hybrid System

The yeast two-hybrid (Y2H) system is a method for detecting protein-protein interactions using a molecular genetic methodology (Gietz *et al.*, 1997). The system has

been used to detect protein-protein interactions in insulin signaling (Hemming, *et al*, 2001), toxicology (Sumida *et al.*, 2001), cancer research (Fong *et al.*, 2000), and many other fields. The goal of this project is to produce a bacterial strain that improves the efficiency of the Y2H system.

The Y2H system makes use of the budding yeast, *Saccharomyces cerevisiae*, and involves a screening protocol based on the activities of the *GAL4* protein. Gal4p is a transcription factor containing DNA binding and a transcription activating domains. The DNA binding domain binds to the UAS element of the *GAL1* promoter and the activating domain stimulates transcription of any downstream gene. In the Y2H system the DNA binding and transcription activating domains are expressed as fusion proteins from separate plasmids (See Figure 1). Protein-protein interactions are detected when the fusion proteins bind to each other and reconstitute the transcriptional activation of Gal4p driving the expression of a reporter gene (See Figure 1). Novel protein-protein interactions can be detected by screening a GAL4 activating domain ($GAL4_{AD}$):cDNA fusion library against a GAL4 binding domain fusion of a known gene such as the Huntington disease protein (Kalchman *et al.*, 1996). The identity of the interacting protein is determined by isolating the $GAL4_{AD}$:cDNA plasmid from the Y2H positive.

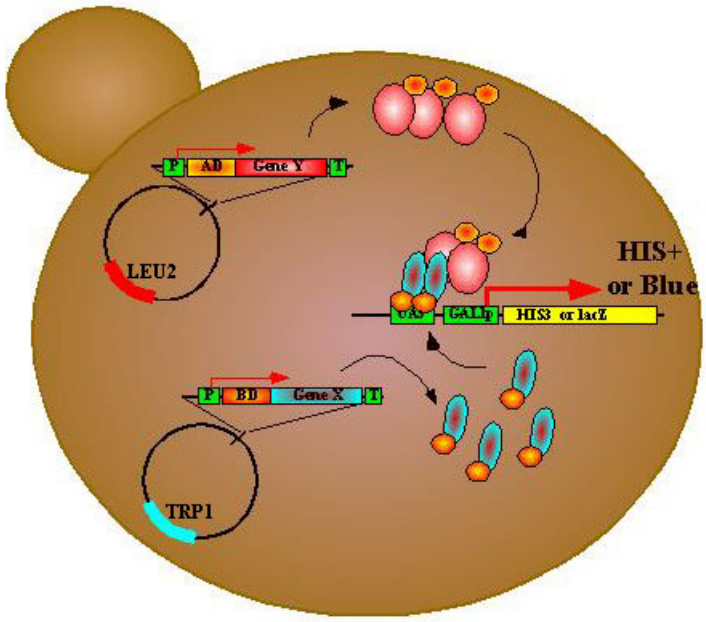


Figure 1: The Yeast Two-Hybrid System (Gietz *et al.*, 1997).
The LEU2 plasmid produces a fusion protein containing the *GAL4* activating domain and the TRP1 plasmid produces a fusion protein with the *GAL4* binding domain. The binding domain fusion protein binds to the UAS to act as the bait. When an activating domain fusion protein interacts with the bait the *GAL1* promoter is turned on and the reporter gene is expressed. The *HIS3* or *lacZ* reporters allow identification of positive interactions because when activated, the colony is conferred with the ability to grow without histidine or turn X-gal blue.

The plasmid vectors used for the Y2H system each contain a different gene that allows for selection in yeast. The yeast genes commonly used are *LEU2*, and *TRP1*. Each participates in a biosynthetic pathway to produce uracil, leucine, and tryptophan, respectively. Each of these yeast genes also has an equivalent gene in bacteria that is responsible for the same biosynthetic function. A yeast strain can only be used for plasmid transformation and maintenance when it contains mutations in these genes, which allow selection for the plasmid. The yeast strain can thus grow on synthetic

complete medium lacking uracil, leucine, or tryptophan, as long as it carries a plasmid with the complementing gene. The $GAL4_{BD}$ plasmid carries the *TRP1* gene and the $GAL4_{AD}$ plasmid usually carries the *LEU2* gene (See Figure 2).

To recover the plasmid with the $GAL4_{AD}$:cDNA fusion gene encoding the interacting protein, a yeast DNA extract, containing both plasmids is transformed by electroporation into bacteria. Ampicillin resistant colonies can be identified, however, the differentiation between the *TRP1* or *LEU2* plasmids requires specific mutations in the *E. coli* host strain. The strain KC8 can be used for identifying yeast *URA3*, *LEU2*, and *TRP1* plasmids due to the presence of the *pyrF*, *leuB* and *trpC* mutations. However, KC8 has a functional *endA* gene. The endonuclease product of this gene results in poor plasmid DNA preparations, making it necessary to transform plasmids purified from KC8 into another strain prior to sequencing or cloning.

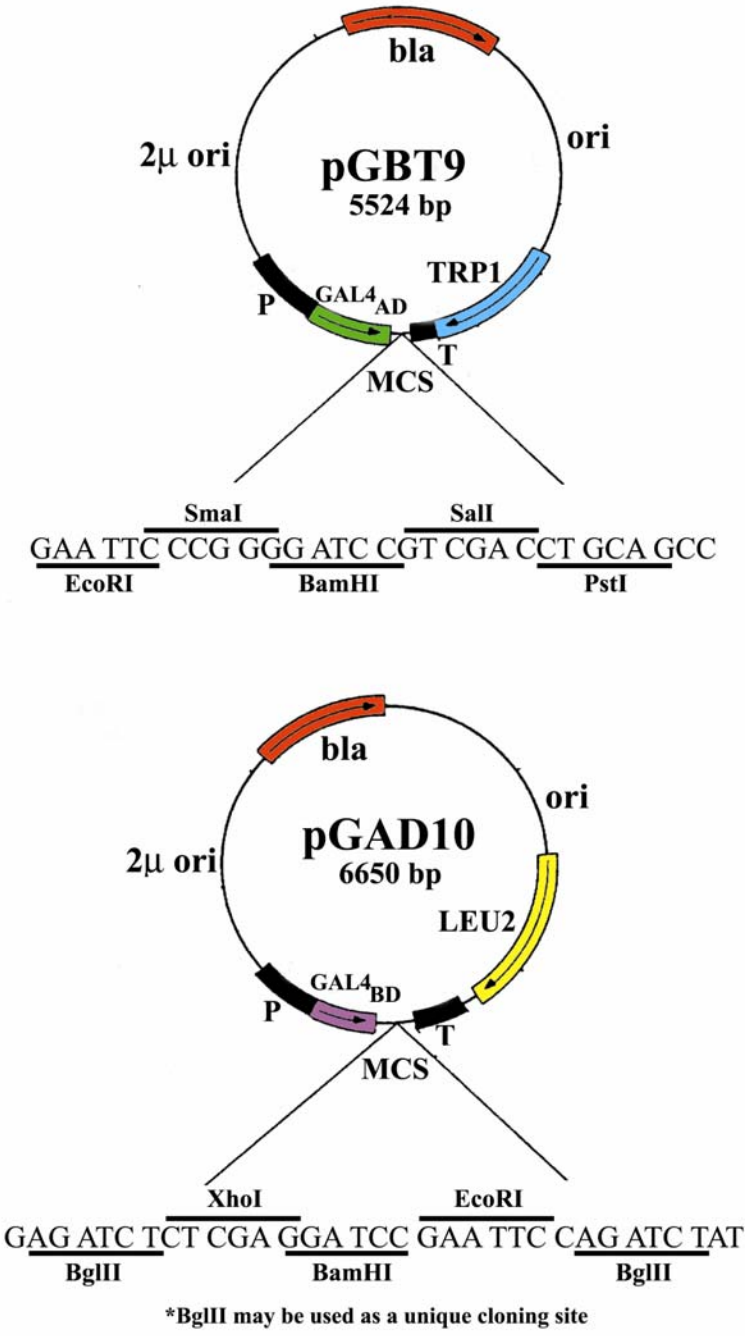


Figure 2: Yeast Two-Hybrid Vectors Carrying *TRP1* and *LEU2* Markers.
Each vector contains a multiple cloning site (MCS) with common restriction enzyme sites for insertion of DNA fragments. The vectors also contain an ampicillin resistance gene (*bla*), the bacterial origin of replication (*ori*), the 2μ yeast origin of replication (*2μ ori*), and the sequence coding for the GAL4 binding or activating domain. The GAL4 binding and activating domains are followed by the MCS and expression of this region is placed under the control of the yeast alcohol dehydrogenase 1 promoter (P) and terminator (T). (Gietz *et al.*, 1997)

PCR Generated Homologous Gene Disruption

Datsenko and Wanner (2000) recently described a method (using PCR products) for one-step inactivation of chromosomal genes in *E. coli*. The method is a refinement of previous recombination-based methods of gene disruption (Kato *et al.*, 1998; Hamilton *et al.*, 1989; Cherepanov and Wackernagel, 1995; reviewed in Zhang *et al.*, 1998). PCR is used to generate a gene disruption cassette which directs the kanamycin resistance gene, flanked by FRT (FLP recognition target) sites, to a specific gene in the *E. coli* chromosome by homologous recombination (See Figure 3). A homologous recombination event within the *E. coli* chromosome replacing the resident gene with the kanamycin resistance gene is stimulated by the λ Red recombinase, expressed from the helper plasmid pKD46. The kanamycin resistance gene is evicted from the chromosome by recombination between its flanking FRT sites (See Figure 3), stimulated by expression of the FLP recombinase from the helper plasmid, pCP20. This method can be used in sequential fashion without resulting in the accumulation of antibiotic resistances.

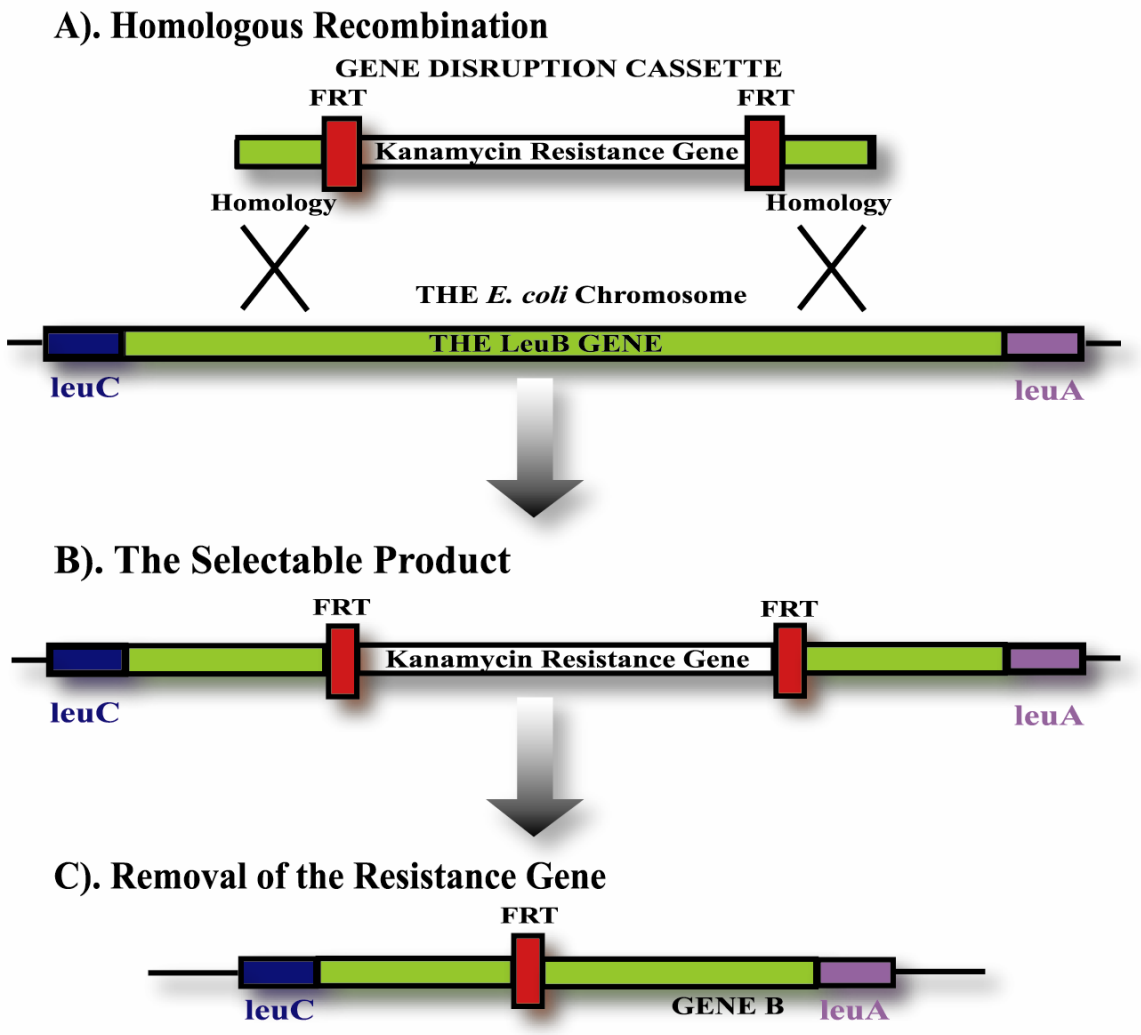


Figure 3: Gene Disruptions Occur by Homologous Recombination.
The *leuB* gene disruption is used as an example. A gene disruption cassette terminated by 39 bp regions of homology to the *leuB* gene is introduced into a bacterium. The cassette lines up with the *E. coli* chromosome and in the presence of Red recombinase a recombination event occurs. The result is a selectable colony with a disrupted *leuB* gene. Finally the kanamycin resistance is evicted by a recombination event stimulated by FLP recombinase.

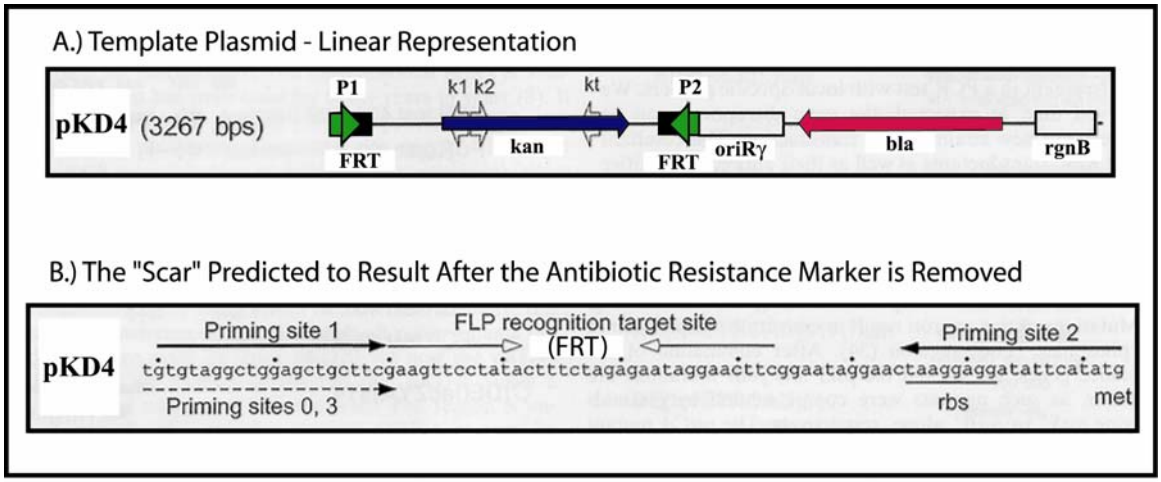
Each gene disruption cassette differs by the flanking regions homologous to the gene that is targeted for disruption. The cassettes were generated from the template plasmid pKD4 (See Figure 4 for a linear representation of this plasmid), which contains the sequence for the kanamycin resistance gene (flanked by FRT sites). Each PCR

product was treated with the *Dpn*I restriction enzyme to degrade any template plasmid present after gel purification. This treatment eliminates the template plasmid in the PCR sample because *Dpn*I only digests methylated DNA. Plasmid DNA isolated from bacteria is normally methylated; however, PCR generated DNA fragments are not methylated. Without this treatment intact template plasmids could produce kanamycin resistant colonies that represent false positives. The cassette contains the kanamycin resistance gene flanked by FRT sites that are in turn flanked by homologous regions of the gene targeted for disruption (See Figures 3 and 4). When the cassette is introduced into the *E. coli* cells a double crossover event, stimulated by the Red recombinase, replaces the target sequence with the cassette and disrupts the gene.

The λ Red recombinase, which stimulates homologous recombination, is expressed from the helper plasmid pKD46. The λ Red recombinase is composed of three functional subunits (γ, β, and *exo*) that work in concert to promote the homologous recombination. The expression of these λ factors is tightly controlled by the arabinose induced promoter, P_{BAD}. The protein product of *araC* turns on the promoter when arabinose is present and limits its activity in the absence of arabinose (Guzman *et al.*, 1995). This stimulation is necessary in strains with *recA1* mutations (Berger and Cohen, 1989; Murphy, 1998).

Removal of the kanamycin resistance gene is accomplished by site-specific recombination between the flanking FRT sites in the gene disruption cassette, stimulated by expression of the FLP recombinase. The result of this removal is an FRT scar of approximately 85 bp (See Figure 4). FLP recombinase is an enzyme from the yeast 2μm plasmid that acts as a site-specific recombinase (Chen *et al.*, 2000). Its natural function is

to assist in the replication of the yeast 2μm plasmid. The enzyme acts by a complex series of DNA cleavage and rejoining steps which may involve the formation of multiple intermediate Holliday structures depending on the type of recombination occurring (Waite and Cox, 1995).



Note: The green arrows (labeled P1 and P2) represent the forward and reverse priming sequences respectively.

Figure 4: Features of the Template Plasmid Used to Generate the Gene Cassettes
A.) Structure of template plasmid used to generate a gene cassette. B.) The scar left by removal of the kanamycin resistance gene (Datsenko and Wanner, 2000).

The approach described above was used to sequentially delete the *pyrF*, *leuB*, and *trpC* genes from DH5α to produce a new strain, MG107 (see Figure 5 for an overview of the procedure). The new strain produces good quality plasmid DNA preparations and can be used to select yeast two-hybrid vectors carrying *URA3*, *LEU2*, and *TRP1* genes. The bacterial strain DH5α was used as the starting material for the creation of MG107.

Escherichia coli Strain DH5α

The strain DH5α is one of the most commonly used strains for recombinant DNA technology. Listed below are the mutations of DH5α that the strain useful for recombinant DNA techniques:

endA1

The product of the *endA* gene is endonuclease I, which is magnesium dependent and acts as a nickase in the presence of RNA or as a double strand nuclease in the absence of RNA (Schoenfeld *et al.*, 1995). The enzyme cuts pure plasmid DNA into seven base oligonucleotides and in the presence of RNA will nick plasmids an average of once per plasmid (Schoenfeld *et al.*, 1995). Endonuclease I expression can be controlled by growing on low glucose nutrient medium, and phenol:chloroform extraction. However DNA degradation is still problematic in strains with *endA* activity. Consequently, any strain carrying the *endA1* mutation (a point mutation) produces higher quality plasmid DNA preparations.

recA1

The recA protein takes part in DNA repair, division of chromosomes, and general homologous recombination (Ullsperger and Cox, 1995). *RecA* along with a *recBCD* enzyme constitute the RecBCD pathway of homologous recombination that occurs preferentially at short DNA sequences called Chi sites (Dabert and Smith, 1997). *RecA* is also a required component in the recE and recF pathways of homologous recombination and acts as an overall regulator of all three pathways. The *recA1* mutation is desirable because any vector carrying a cloned insert may be destabilized by general recombination in the bacterial host if the recA protein is present (Hanahan *et al.*, 1991).

deoR

A *deoR* mutation allows the bacteria to grow on minimal media that contains only one source of carbon (inosine) as well as allowing efficient uptake of large pieces of DNA, useful in constructing gene libraries of very large inserts (Hanahan, 1991).

gyrA96

GyrA mutations are necessary for strains used in cloning because the product of *gyrA*, a DNA gyrase may cause deletions between direct repeats that may well occur in a cloned insert on a plasmid (Hannahan *et al.*, 1991).

lacZ*Δ*M15

The *lacZ*Δ*M15* mutation deactivates *lacZ* activity in the bacteria producing an inactive form of β-galactosidase (Yannish-Perron *et al.*, 1985). Strains with this mutation can not cleave X-gal and remain colourless on X-gal plates (X-gal turns blue when cleaved). However, if a plasmid carrying a *lacZ* alpha subunit (such as a pUC cloning plasmids) is introduced into the strain, it complements the truncated *lacZ* gene and produces β-galactosidase activity. Since the alpha subunit has been placed in frame with the multiple cloning site (MCS) of the pUC vectors, an insert at the MCS will disrupt the alpha subunit and colonies containing an insert will appear as white colonies rather than blue ones .

Objectives of the Project

The objective of this project was to create a bacterial strain that has the *pyrF*, *leuB*, and *trpC* mutations required for use in plasmid rescue for the Y2H system. The

strain must be able to produce good quality DNA preparations and allow selection of common two-hybrid vectors carrying the yeast *URA3*, *LEU2*, and *TRP1* genes

Materials and Methods

Bacterial Strains

The *Escherichia coli* strains used during this project are listed below in table 1.

Table 1: Bacterial Strains

Strain	Genotype	Plasmid	Source	Reference
DH5alpha	*F⁻, endA1, hsdR17* ($r_K^- m_K^+$), *glnV44,thi-1 deoR, gyrA96, recA1, relA1, supE44, Δ(lacZYA-argF)U169, λ-, (ϕ80dlacΔ(lacZ) M15)*	none	R. D. Gietz	(Woodcock *et al.*, 1989)
HB101	*F⁻, Δ(gpt-proA)62, leuB6, glnV44, ara-14 galK2, lacY1, Δ(mcrC-mrr), rpsL20* (Str^r) *Xyl-5, mtl-1, recA13.*	none	R. D. Gietz	(Maniatis *et al.*, 1989)
KC8	*hsdR, leuB600, trpC9830, pyr::Tn5, hisB463, lacΔX74, strA, galU, galK.* Note: Tn5 confer kan^r.	none	R. D. Gietz	(Struhl, 2000)
BW25141	*Δ(araD-araB)567, ΔlacZ4787(::rrnB-4), lacIp-4000(lacI^Q), ΔphoB-phoR)580, λ-, galU95, ΔuidA3::pir⁺, rpoS396(AM), endA9(del-ins)::FRT, rph-1, Δ(rhaD-rhaB)568, rrnB-4, hsdR514.*	pKD4	CGSC (The *coli* Genome Stock Center)	(Datsenko and Wanner, 2000)
BW25113	*Δ(araD-araB)567, ΔlacZ4787(::rrnB-4), lacIp-4000(lacI^Q), λ-, rpoS396(Am), rph-1, Δ(rhaD-rhaB)568, rrnB-4, hsdR514*	pKD46	CGSC	(Datsenko and Wanner, 2000)
BT340	Δ(*argF-lac*)*169, phi80ΔlacZΔM15, glnV44(AS), λ-, rfbD1, gyrA96(NalR) recA1, endA1, spoT1, thi-1, hsdR17.*	pCP20	CGSC	(Datsenko and Wanner, 2000)

Plasmids

Bacterial strains carrying the plasmids pKD46, pKD4, and pCP20 were obtained from the *E. coli* Genome Stock Center (CGSC) and used in the process of gene disruption (Datsenko and Wanner, 2000).

pKD4

pKD4 is the template plasmid used to generate the PCR gene cassettes. It contains the kanamycin resistance gene flanked by FRT sites. The plasmid structure can be seen in Figure 5.

pKD46

pKD46 contains the Red recombinase system. It is a low copy plasmid that contains the γ, β, and *exo* functions of the Red system, which increase the efficiency of recombination events. The plasmid also contains a temperature sensitive origin of replication that allows it to be "cured" (the plasmid is lost) from bacteria at 43°C. Bacteria containing pKD46 were grown at 30°C to maintain the plasmid. The structure is shown in Figure 5.

pCP20

pCP20 expresses the FLP recombinase which stimulates removal of the FRT flanked kanamycin resistance gene. The plasmid also contains a temperature sensitive origin of replication that allows it to be cured from bacteria at 43°C. Bacteria containing pCP20 were grown at 30°C to maintain the plasmid. The plasmid structure is shown in Figure 5.

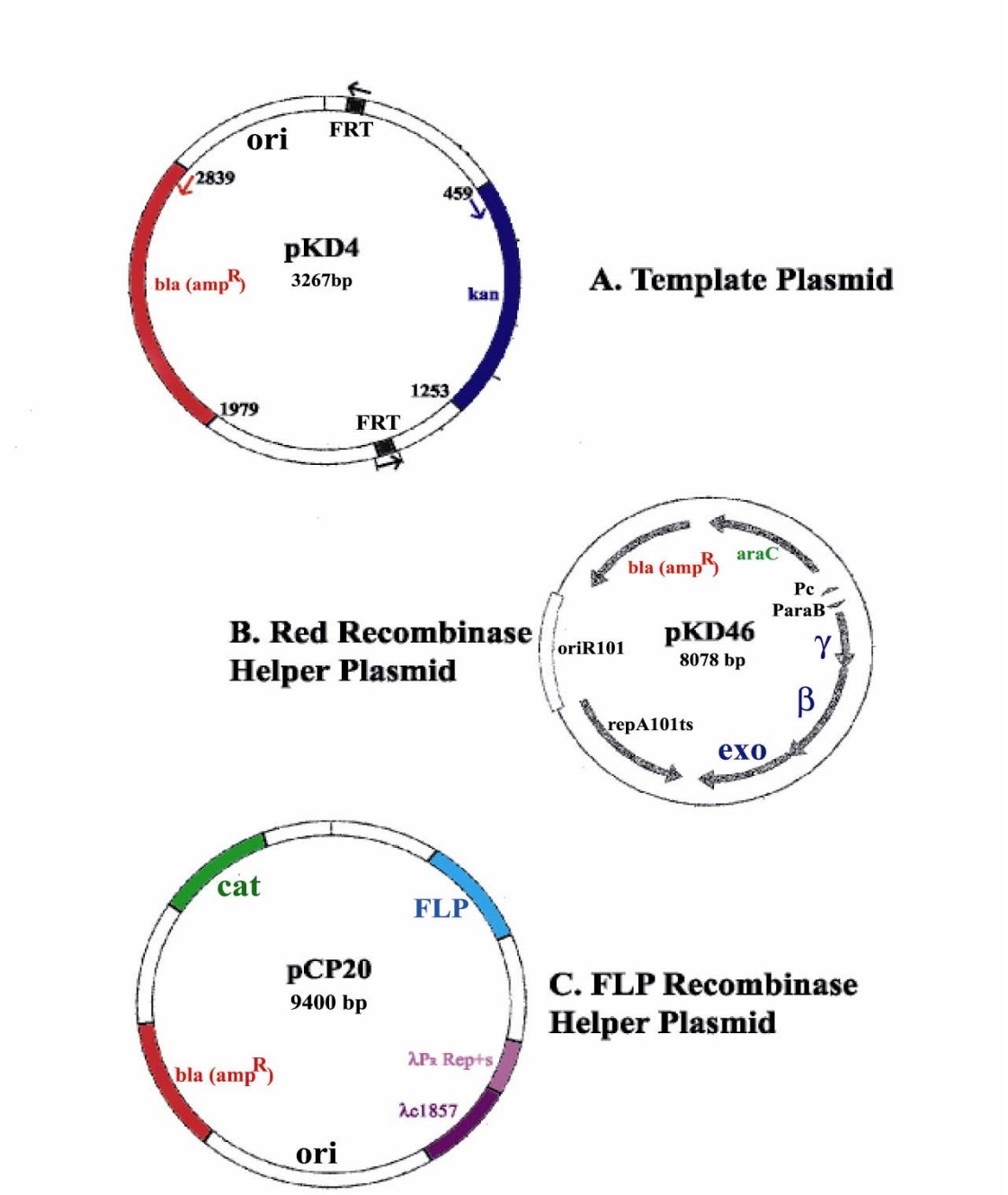


Figure 5: Template and Helper Plasmids Used in the Gene Disruption System
pKD4 contains the gene for ampicillin resistance (*bla*) and the gene for kanamycin resistance (*kan*). The kan gene is flanked by two FRT sites. The origin of replication is indicated by ori. pKD46 contains the bla gene. The Red recombinase genes are indicated as γ, β, and *exo* . The plasmid also contains the P_{BAD} promoter (indicated by *ParaB*) and the *AraC* gene. The origin of replication (oriR101) and repA101ts regions confer temperature sensitive replication. pCP20 contains the genes for ampicillin resistance (*bla*) and chloroamphenicol resistance (*cat*). It also contains the FLP recombinase gene and temperature sensitive replication origins.

YEplac112, YEplac181, and YEplac195

YEplac112, YEplac181, and YEplac195 were built from pUC19 and therefore contain the *lacZ* alpha fragment (Gietz and Sugino, 1988). Each of these plasmids contains an *E. coli* origin of replication, the *bla* gene for ampicillin resistance, and the yeast *URA3*, *LEU2*, and *TRP1* genes, respectively.

Selective Media

LB agar medium (Luria Burtani) was made by dissolving 24 g of LB mix (DIFCO) in 600 mL of nanopure water and autoclaving the solution for 25 minutes. Once the media had cooled to 60°C, kanamycin or ampicillin was added to final concentrations of 25 μg/ml and 100 μg/ml respectively. The mixture was stirred and poured into 100 x 15 mm petri plates.

Minimal Media

Minimal medium was made using the following recipe.

M9 Minimal Medium

10x M9 salts	60 ml
Water (nanopure)	540 ml
Agar	10 g

-these components were autoclaved for 25 minutes and the following sterile components added:

$MgSO_4$ (1M)	0.6 ml
$CaCl_2$ (0.1M)	0.6 ml
Thymine (4 mg/ml)	0.5 ml
20% glucose	6.0 ml
$FeCl_3$ (0.01M)	0.15 ml
Vitamin B1 (2 mg/ml)	0.6 ml
Histidine (2 mg/ml), Tryptophan (2 mg/ml), Uracil (2 mg/ml), Leucine (2 mg/ml) → Omit One	6.0 ml each.

The mixture was stirred and poured into 100 x 15 mm petri plates.

Solutions

The solutions used in a variety of techniques are listed below.

Band Elution Buffer	
5M NaCl	400 µl
1M Tris pH 7.5	500 µl
0.5M EDTA	20 µl
20% SDS	50 µl
H_2O (nanopure)	(to 10 ml)
EndA Assay Buffer	
1M Tris-Acetate (pH 7.8)	2.5 ml
1M Potassium Acetate	10 ml
1M Magnesium Acetate	1 ml
1M DTT	100 µl
H_2O (nanopure)	86.4 ml
Loading Buffer (10X)	
Bromophenol Blue	50 mg
Xylenecyanol	50 mg
Glycerol	25 ml
H_2O (nanopure)	25 ml
Miniprep Solution I (TGE Buffer) – Final Concentrations	
50 mM Glucose	
50 mM EDTA	
25 mM Tric-HCl (pH 8.0)	
Miniprep Solution II (Lysis Buffer)	
H_2O (nanopure)	9.25 ml
20% SDS	500 µl
10N NaOH	200 µl
Miniprep Solution III	
5M Kac	60 ml
Glacial Acetic Acid	11.5 ml
H_2O (nanopure)	28.5 ml
TAE Buffer (50X)	
Tris Base	242 g
Glacial Acetic Acid	57.1 ml
0.5M EDTA (pH 8.0)	500 ml
H_2O (nanopure)	(to 1 L)

TE Buffer

0.5M Tris-HCl (pH 8.0)	10 ml
0.5M EDTA (pH 8.0)	100 ml

X-Gal (5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside) 20mg/mL

X-gal	1g
N,N-dimethylformamide	50 ml

Plasmid DNA Preparation

Plasmid DNA was extracted by a modification of the method described by Birnboim and Doly (1979). This involved preparing 2 ml overnight liquid LB bacterial cultures and centrifuging 1.5 ml of the culture at high speed (13000 rpm) to pellet the cells. The cells were resuspended in 100 μl miniprep solution I, then 200 μl solution II (to lyse the cells), and finally 150 μl solution III (to neutralize the pH and precipitate the unwanted chromosomal DNA and protein). This mixture was centrifuged at high speed for 5 minutes and the supernatant removed. The supernatant was extracted with a phenol:chloroform solution (1:1), centrifuged again, and the aqueous layer removed. Finally the aqueous layer was mixed with 50 μl sodium acetate and 2 volumes of ethanol, and the plasmid DNA was precipitated at -80°C. After at least one hour the solution was centrifuged at high speed for 15 minutes, the ethanol carefully discarded, and the DNA pellet resuspended in 50 μl TE buffer. All samples were stored at -20°C until needed.

Restriction Enzyme Digestions

Restriction enzyme digestions for analysis of plasmid DNA were performed as follows: 0.5 μl restriction enzyme (from NEB) was added to a mixture of 4μl of Plasmid DNA (~0.5 μg), 3 μl of appropriate 10x NEB buffer, and 22.5 μl nanopure H_2O. The mixture was incubated for 0.5-1 hours at 37°C.

The *Dpn*I treatments involved 50 µl of PCR product, 20 µl of NEB buffer 4, 2 µl *Dpn*I, and 128 µl of H_2O. The reaction was incubated at 37°C for 2 hours and followed by a phenol:chloroform extraction and ethanol precipitation. After resuspending the DNA in 20 µl of TE buffer, the purity was checked by running 2 µl on a gel.

PCR Primers

PCR primers were produced and desalted by GIBCO BRL at the 50 nmol scale. The primers were dissolved in TE buffer to a final concentration of 50 mM and stored at –20°C. Primer sequences are presented in Table 2.

Table 2: The Sequence of Primers Used to Generate Gene Disruption Cassettes. Each set of primers has regions homologous to an *E. coli* gene. The homologous extensions of the primers are underlined and the region of homology for the target gene is given. The primer sequence homologous to the template plasmid, pKD4 is shown (red for forward, blue for reverse). Note that for all three sets of primers, the annealing portion of the primer remained constant and corresponded to regions on the template plasmid, that flank the FRT sites.

			Position of Homology Within *E. coli* Gene	
Primer Name	**Sequence (5' to 3')**	**Tm (°C)**	**Start**	**End**
pyrF -pKD4 Forward	CGCTGTTACGAATTCTCCTGTGGTTGTTGC CCTTGA GTGTAGGCTGGAGCTGCTTC	70	28 th bp	58 th bp
pyrF -pKD4 Reverse	GCAGCAGCGACAGCGTGCGCTGCAGTGTT GGGGATA CATATGAATATCCTCCTTAG	70	227 th bp	257 th bp
leuB -pKD4 Forward	AACCGCTTTGCGATGCGCATCACCACCAGC CATTACGAT GTGTAGGCTGGAGCTGCTTC	71	93 rd bp	132 nd bp
leuB -pKD4 Reverse	GCAATGTCTGCACGCAGCGGACAGAATGC TTCCAGCCCC CATATGAATATCCTCCTTAG	70	344 th bp	383 rd bp
trpC -pKD4 Forward	TTTTAGCGAAAATCGTCGCAGACAAGGCG ATTTGGGTAG GTGTAGGCTGGAGCTGCTTC	69	13 th bp	52 nd bp
trpC -pKD4 Reverse	GGCGGCAATGCGTGCTGGATCGAAATCAT CACGGATCAC CATATGAATATCCTCCTTAG	69	190 th bp	229 th bp

PCR Amplification

A polymerase chain reaction was used to generate the PCR gene cassettes. Standard reactions were used and a custom PCR program (MG1) was entered into a PTC-100 programmable thermal controller PCR machine (MJ Research inc.) with a hot bonnet™ lid.

Reaction Mixture (52µl reactions)

37 µl	H_2O
5 µl	Taq PCR Buffer 10x
5 µl	dNTP's (2mM)
1.00 µl	Forward Primer (50 mM)
1.00 µl	Reverse Primer (50 mM)
0.75 µl	Template DNA (~10ng of pKD4)
0.25 µl	Expand High Fidelity PCR Taq Polymerase (Roche)
2.00 µl	DMSO

PCR Program MG1

Step	Time	Temperature	Function
1.	5:00 min	94.0°C	Denature DNA (Initial)
2.	30 sec	94.0°C	Denature DNA (Cycle)
3.	1:00 min	58.0°C	Primer Annealing (Cycle)
4.	1:20 min	68.0°C	Extension (Cycle) – Allows the 1600bp extension required.
5.	Repeat steps 2, 3, and 4 (30 times)		
6.	7:00 min	68.0°C	Final Extension
7.	Indefinite	4.0°C	Holds at low temp. until operator returns

Agarose Gel Electrophoresis

Gels were prepared with 0.75% w/v agarose in 1x TAE buffer (boiled and cooled). For every 100 ml of gel, 10 µl of ethidium bromide (10 mg/ml) was added while cooling. The agarose solutions were poured into taped molds and allowed to cool for at least 20 minutes. Loading buffer (10x) was added to each sample before loading. A

power source (Gelman) was used to run the samples at 150V and 50 mAmps for 30 minutes. A KB ladder (Gibco) was used to determine the approximate size of DNA molecules present.

Gel Purification and Band Elution

PCR products were electrophoresed and then gel purified using a modification of the method of Girvitz *et al.* (1980). PCR bands corresponding to the correct size of product were located using long wavelength UV light. A small piece of 3MM Whatman paper, backed by dialysis membrane (DNA dam) was inserted in front of the band into a small slit cut in the gel with a razor blade. The gel was run for an additional ten minutes, allowing the DNA dam to capture the band and thereby isolate it from DNA of different sizes. The paper was then removed and the DNA was eluted from the paper with three 100 µl aliquots of band elution buffer. The solution containing the purified DNA was then phenol:chloroform extracted, ethanol precipitated, and resuspended in 20 µl TE buffer. Finally a 2 µl sample of the solution was run on a gel to check the purity and quantity of the DNA.

Preparation of Electrocompetent *E. coli*

A single colony of bacteria was used to inoculate 2 ml of LB medium which was incubated overnight at 37°C with shaking. One ml of this culture was transferred to 100 ml of LB and grown with shaking for 2.5 hours. The media was centrifuged to pellet the cells and the cells were washed 2 times with 1 volume of ice-cold sterile water and 3 times with 1 volume of ice cold sterile 10% glycerol. The cells were then used immediately for electroporation.

Preparation of Electrocompetent, Recombination Proficient *E. coli*

When electrocompetent cells already carrying the Red helper plasmid (pKD46) were prepared, they were grown at 30°C in a 0.2% final concentration of arabinose (to activate the Red system). Ampicillin (100 μg/ml final concentration) was added to the medium to maintain the plasmid.

Electroporation

DNA molecules were transformed into bacteria using electroporation (Dower *et al.*, 1988). The electrocompetent bacterial cells (25 μl) were mixed with 2 μl of DNA and placed in an ice cold electroporation chamber. The cells were electroporated using a BIORAD Gene Pulser™ at 1.25 kV, 25 μF with a 400 ohm resistor in parallel with the sample. The electroporated cells were immediately resuspended in 1ml of LB broth and incubated for 0.5 to 1 hour before plating onto selective medium. When electroporation involved the introduction of a temperature sensitive plasmid the plates were incubated at 30°C, otherwise they were incubated at 37°C.

Gene Disruption

Cells expressing the Red recombinase were electroporated with 2μl of PCR gene disruption cassette DNA. The cells were incubated for 30 minutes at 37°C and half of them were spread onto kanamycin (25 μg/ml) plates to select for kanamycin resistant colonies. If none appeared after 16 hours, the remaining 500 μl was spread after standing at 25°C overnight. All kanamycin resistant colonies that appeared were transferred to fresh kanamycin plates to ensure that they were not feeder colonies. The colonies that

continued to grow well on kanamycin plates (putative positives) were tested to determine whether the desired gene disruption had occurred. This was accomplished by replica plating kanamycin resistant colonies onto M9-(uracil, leucine, or tryptophan) plates. Colonies that could not grow were recovered from the master plate, suspended in 20% glycerol and stored at -80°C. Two colonies were selected and cured to remove the helper plasmid by growing at 43°C.

Removal of the Kanamycin Resistance Gene

Mutants were electroporated with pCP20 and transformants were selected on ampicillin plates. Ten of the ampicillin resistant mutants, expressing the FLP recombinase from the helper plasmid, were grown at 43°C.

Phenotype Tests

Antibiotic Resistance

Antibiotic Resistance was tested on LB plates containing ampicillin (100 μg/ml), kanamycin (25 μg/ml), tetracycline (15 μg/ml) and carbenicillin (20 μg/ml).

Auxotrophy

To confirm the phenotype resulting from each gene disruption the strain was tested on M9 media (lacking uracil, leucine, or tryptophan). If no growth appeared after 48 hours the strain was considered a mutant for the gene in question.

lacZ Activity

A 70 µl aliquot of X-gal solution was spread onto an LB plate and allowed to dry. Bacteria were then streaked onto the plate and incubated overnight at 37°C.

EndA Activity

A modification of an assay developed by Promega was used (Schoenfeld *et al.*, 1995). Extracts were prepared from bacteria by the miniprep procedure but the phenol:chloroform extraction was omitted. Each extract was resuspended in 30 µl of *endA* assay buffer, to which 10 µL of plasmid DNA (pBR322 or pUC9) and 0.5 µl of RNase was added. The mixtures were then incubated at 37°C for varying amounts of time. The samples were analyzed for DNA degradation by gel electrophoresis.

Results

Confirmation of the Structure of the Template and Helper Plasmids

Upon receiving the bacterial strains containing the plasmids, pKD46, pKD4 and pCP20 they were streaked for single colonies on ampicillin plates and cultured for plasmid extraction. Plasmid DNA was extracted using the miniprep procedure outlined in the Materials and Methods. The presence and identity of the plasmids was determined by *EcoR*I restriction enzyme digestion and agarose gel electrophoresis. In each case the plasmid DNA gave a single band corresponding to the expected size of each plasmid.

PCR Primer Design

For each of the genes, *pyrF*, *leuB*, and *trpC* the location of the gene in the *E. coli* chromosome was determined by a search of NCBI's Entrez database of the K12 *E. coli* genome (http://www.ncbi.nlm.nih.gov:80/cgi-bin/Entrez/framik?db=Genome&gi=115). This data was contributed by the University of Wisconsin and has been indexed and cross-referenced thoroughly (Blattner *et al.*, 1997). Since the *E. coli* genome is essentially complete it is a simple matter to find a target gene in the 4,639,221 bp sequence. By searching the entrez database, the exact location and orientation (+ or -) of each target gene was determined. To expedite the process and reduce my dependence on access to the internet, I downloaded the entire genome by FTP as a single 10.7 Mb text file which could then be searched by sequence number. When the gene was located, the sequence was copied from this file and saved in the DNA analysis program DNA Strider 1.2. If the gene was in the (–) orientation the program was used to generate the complementary sequence, and then reverse the order of base pairs giving the sequence in

the forward orientation. To verify the DNA sequence I converted it to a protein sequence and compared it to the sequence of the target protein listed in the database. A printout of the gene sequence for each target gene was used to design each set of primers.

For each target gene, the homologous extensions used were designed so that the double crossover event would occur inside the target gene and replace part of its sequence with the kanamycin resistance gene disruption cassette. For each gene, 39 bp of homology were used and added onto the 20 bp of priming sequence for the template pKD4 (See Figure 4B). The forward primers had 39 bp of sequence taken from the top strand of the DNA sequence, read from left to right and starting from about 10-150 bp from the beginning of the ORF. The reverse primers used a sequence from about 200-350 bp farther into the ORF and would be read right to left from the bottom strand of the DNA sequence. The homologous extensions were 39 base pairs in length, well above the reported minimum of approximately 20 base pairs of identical sequence required for efficient recombination (Watt *et al.*, 1985). The sequence location for each primer was shifted until the GC content of both the forward and reverse primers was approximately equal, giving them similar melting temperature (Tm) values for the denaturation step of the PCR reaction (See Table 2). Each primer consisted of 59 bp, 39 of which were homologous to the target gene and 20 of which were priming sequence for the plasmid pKD4. In each case the primers designed produced the expected PCR fragment.

Preparation of PCR Gene Cassettes

The three gene disruption cassettes were generated by PCR as described in the Materials and Methods. Each pair of PCR primers (See Table 2) were mixed with approximately 50 ng of the template plasmid, pKD4, and PCR cocktail components and

then subjected to the thermo cycling program, MG1, in the MJ Research PTC-100 PCR machine. After completion of the program several reaction tubes containing crude PCR product were mixed together and ethanol precipitated to concentrate the product. The gene disruption cassettes were gel purified, *Dpn*I treated, and electroporated into DH5α or its derivatives by electroporation. Figure 6 shows an agarose gel of the *leuB*-pKD4 PCR product prior to purification. The product size appears to be approximately 1600 base pairs as expected. Similar results were obtained for the *pyrF*-pKD4 and *trpC*-pKD4 PCR products.

Figure 6: An Agarose Gel of the *leuB*-pKD4 PCR Product Prior to Purification. First Well (on right side) Contains KB Ladder. Wells 1-10 contain 4 μl of PCR reaction mixture (the reaction did not work for sample #8).

Verification of the Presence of the Helper Plasmid pKD46

The helper plasmid pKD46 was electroporated into DH5α and its derivatives to stimulate the induction of homologous recombination of the PCR generated gene disruption cassettes. Transformed cells were identified by selection on ampicillin plates and used to prepare plasmid DNA to verify the presence of pKD46. The plasmid DNA

extracted from a number of ampicillin resistant colonies was analyzed by agarose gel electrophoresis. Figure 7 shows the presence of pKD46 in each ampicillin resistant transformant. Similar results were found for the *pyrF* and *leuB* gene disruptions however this verification step was omitted for the *trpC* gene disruption.

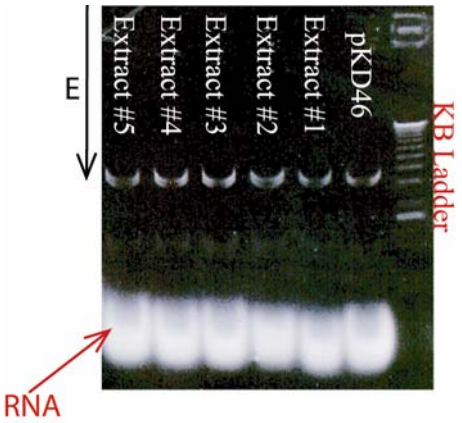


Figure 7: Verification of the Presence of pKD46 in Electroporated Bacterial Cells. Five colonies were selected on ampicillin plates and used to make plasmid DNA preparations. Each plasmid DNA sample was *EcoR*I digested and analyzed by agarose gel electrophoresis. Each lane contains 4 µl of plasmid DNA dissolved in 25 µl TE buffer and loaded with 3 µl loading buffer. The lane closest to the ladder contains the plasmid DNA originally used to electroporate the bacteria. The photograph verifies the presence of the correct plasmid.

The Results of Gene Disruption

DH5α or its subsequent derivatives containing pKD46 were electroporated with the purified PCR generated gene disruption cassettes. Kanamycin resistant bacterial colonies were selected and tested for the disruption phenotype by streaking onto M9 minimal media lacking uracil, leucine or tryptophan.

Disruption of the *pyrF* gene in DH5α

The first gene disruption to be attempted was that of the *pyrF* gene. DH5α cells containing pKD46 were used to produce electrocompetent, recombination proficient bacterial cells as described in the Materials and Methods. These cells were electroporated with the purified PCR product. The 1 ml electroporation mixture was incubated at 37°C for 1 hour. One 500 µl aliquot of the cells was spread onto a single kanamycin plate. This plate was incubated overnight at 37°C but produced no kanamycin resistant colonies. The second 500 µl aliquot was incubated overnight at room temperature before being spread onto a kanamycin plate. This was done to allow the recombination event more time to occur. This plate produced 22 large colonies after 16 hours of growth at 37°C. The kanamycin resistant phenotype of all 22 colonies was confirmed by restreaking onto a fresh kanamycin plate. All colonies were streaked onto M9-uracil plates and 12 colonies did not produce any growth. All 22 colonies were used to produce a bacterial stock however only one (MG101A) was chosen for the subsequent disruption of the *leuB* gene (of MG101A to MG101U).

Prior to disruption of the *leuB* gene it was necessary to cure the bacterial stock, MG101A, of the pKD46 plasmid (Red recombinase helper plasmid) as described in the Materials and Methods.

The final step was to evict the kanamycin resistance gene from the strain to enable subsequent use of this marker for further gene disruption by this method. This was done by electroporation of MG101A (amp^S) with the helper plasmid pCP20 (FLP recombinase) as described in the Materials and Methods. Ten ampicillin resistant colonies were streaked onto LB plates and incubated overnight at 43°C. Each stock was

tested for kanamycin and ampicillin sensitivity by streaking onto the appropriate media. All 10 stocks were ampicillin and kanamycin sensitive showing that each had lost the kanamycin resistance gene as well as the pCP20 helper plasmid. One of these stocks was chosen for continued work (MG102A). An overview of the entire process of gene disruption can be seen as Figure 8.

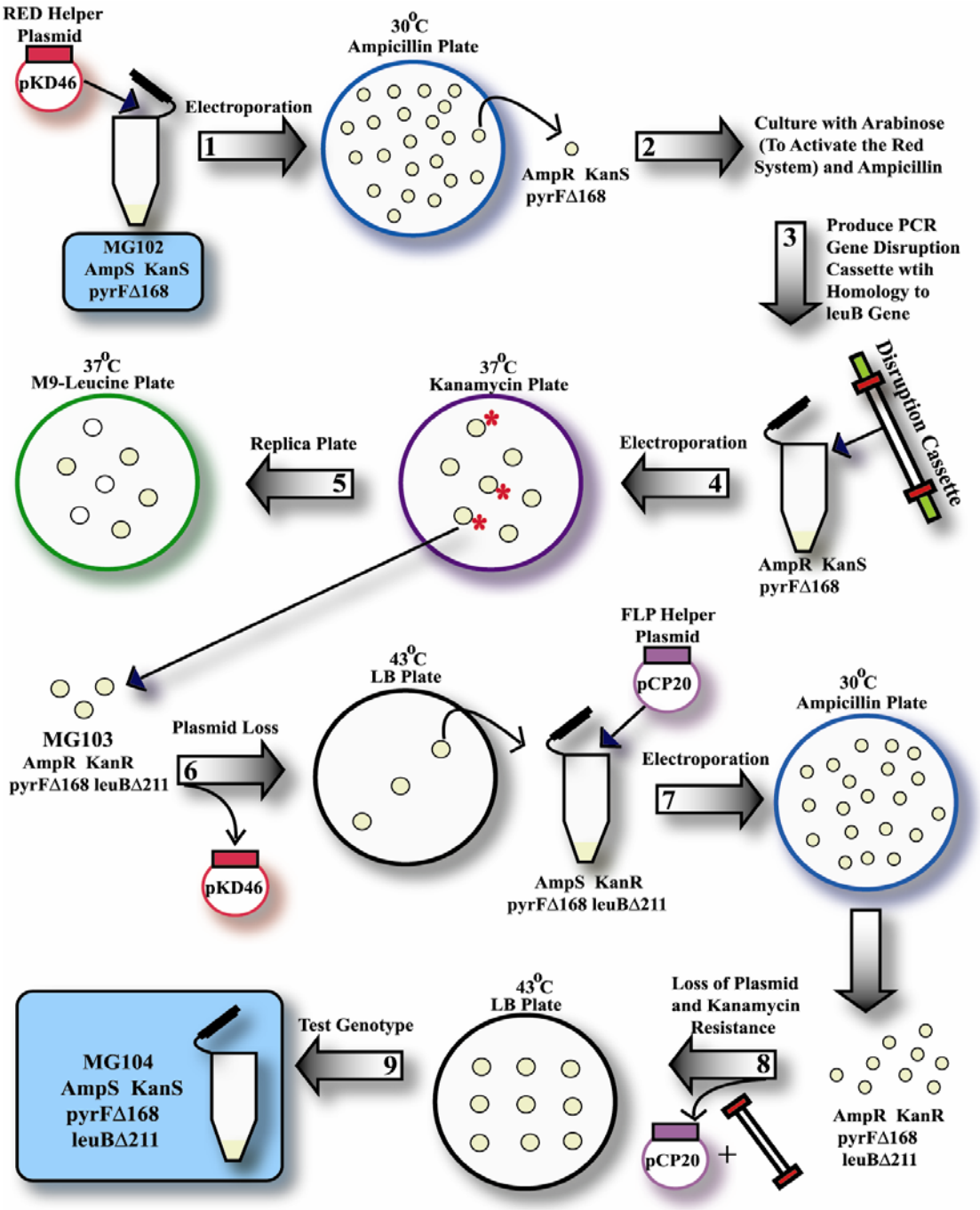


Figure 8: An Overview of the Process of Gene Disruption
The overview above illustrates the disruption of the *leuB* gene in a strain already containing the *pyrFΔ168* mutation. The process involved the electroporation of bacterial cells with the Red and FLP helper plasmids. Cells containing a helper plasmid were grown at 30°C on ampicillin plates to maintain the plasmid. Mutants containing the gene disruption cassette were selected on kanamycin plates at 37°C and the leuB genotype was tested on M9-leucine plates at 37°C. Helper plasmids were removed from the cells by growing on LB plates at 43°C.

Disruption of the *leuB* gene in MG102A (*pyrF*Δ*168*)

The DH5α derivative MG102A containing the *pyrF*Δ*168* mutation was used for deletion of the *leuB* gene. The helper plasmid pKD46 was electroporated into MG102A and the resulting stock used to prepare electrocompetent, recombination proficient bacterial cells. The *leuB*-PCR gene disruption cassette was electroporated into these cells. The electroporation mixture was treated as in the previous gene disruption. The first attempt at this disruption was unsuccessful because only small slow-growing kanamycin resistant colonies were produced. However, the second attempt of the *leuB* disruption was successful. Once again the culture incubated for only 1 hour and then grown overnight at 37°C yielded no kanamycin resistant colonies. The culture incubated for 24 hours at room temperature, then grown at 37°C yielded 44 kan^R colonies. Each was able to grow well when transferred to a new kanamycin plate. The kanamycin resistant colonies were tested for the ability to grow on M9-leucine plates and 16 did not grow. Each positive colony was used to produce a bacterial stock however only MG103A was chosen for the subsequent disruption of the *trpC* gene.

The MG103A stock was cured of the pKD46 plasmid as described in the Materials and Methods. In addition the kanamycin resistance gene used to select for the *leuB* disruption was evicted as described in the Materials and Methods. This produced the strain MG104A, which was subsequently used to disrupt the *trpC* gene.

Disruption of the *trpC* gene in MG104A (*pyrF*Δ*168*, *leuB*Δ*211*)

The DH5α derivative MG104A containing the *pyrF*Δ*168* and *leuB*Δ*211* mutations was used for deletion of the *trpC* gene. The helper plasmid pKD46 was electroporated into MG104A and the resulting stock used to prepare electrocompetent, recombination

proficient bacterial cells as described in the Materials and Methods. The *trpC*-PCR gene disruption cassette was electroporated into these cells. Cells electroporated with *trpC*-PCR product were screened for kanamycin resistance and the inability to grow on M9-tryptophan plates. The cells incubated for only 1 hour and then grown for 16 hours yielded 8 kan^R colonies. Those incubated overnight at room temperature and then grown for 16 hours yielded an additional 57 kan^R colonies. Once again, all putative positives grew well when transferred to a new kanamycin plate. All of the 65 kan^R colonies failed to grow on M9-tryptophan plates. Ten colonies were used to produce bacterial stocks however only MG105A was chosen for the subsequent manipulations.

The MG105A stock was cured of the pKD46 plasmid as described in the Materials and Methods and the kanamycin resistance gene used to select for the *trpC* disruption was evicted as described in the Materials and Methods. This produced the strain MG107, which contains the *pyrF*Δ*168*, *leuB*Δ*211*, and *trpC*Δ*137* mutations.

Confirmation of the Genotyped MG107

Auxotrophic Analysis

MG107 was tested for inability to grow on M9-uracil, M9-leucine, or M9-tryptophan media by streaking the strain onto the appropriate plates. The strain was incapable of growing on any of these media, however it grew well on LB plates.

To test the ability of yeast genes to complement the *pyrF*Δ*168*, *leuB*Δ*211*, and *trpC*Δ*137* mutations, electrocompetent MG107 cells were produced as described in Materials and Methods. The electrocompetent cells were electroporated with the yeast-*E. coli* shuttle vectors YEplac195, YEplac181, and YEplac112 (Gietz and Sugino, 1988).

Each of these plasmids contains an *E. coli* origin of replication and the *bla* gene for ampicillin resistance. YEplac 195 carries the yeast *URA3* gene, YEplac181 carries the *LEU2* gene and YEplac 112 carries the *TRP1* gene. MG107 ampicillin resistant colonies containing YEplac195 were replica plated onto M9-uracil. All colonies produced robust growth on this medium. Similar complementation was provided by YEplac181 and YEplac112 on M9-leucine and M9-tryptophan plates, respectively. This result shows that the *pyrF*Δ*168*, *leuB*Δ*211*, and *trpC*Δ*137* mutations can be complemented by the corresponding yeast genes.

Assay for Plasmid DNA Quality

To assess the quality of plasmid DNA that can be isolated from MG107 as compared to KC8 an assay of the endonuclease I activity present in cell extracts was performed. Extracts from MG107 and KC8 were prepared by a modification of the miniprep procedure as described in the Materials and Methods. The extracts contain mostly RNA and a small amount of protein. By observing the effect of each extract on added plasmid DNA over time, a qualitative comparison of the *endA* activity in KC8 and MG107 was obtained (see Figure 9). Figure 9B shows that most if not all of the super helical pUC9 DNA is converted to linear or nicked circular forms in the KC8 extract. Therefore, more nickase and non-specific degradation is apparent in the plasmid DNA added to the KC8 extract than in the plasmid DNA added to the MG107 extract.

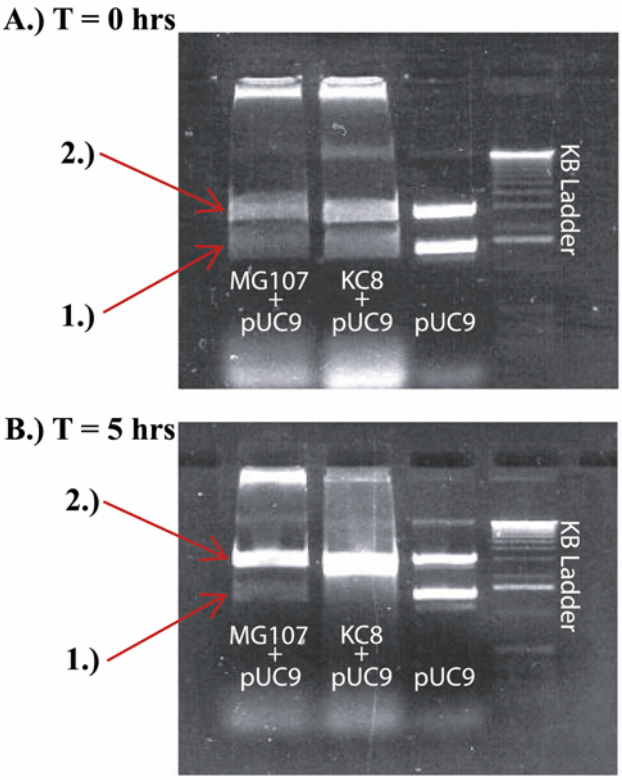


Figure 9: An Assay of the Endonuclease Activity in MG107 and KC8.
pUC9 shows two bands on the gel, a super helical circular band (1.) and a linear and nicked circular combination band (2.). A.) 0 hours incubation. MG107 and KC8 extracts were added to pUC9 and electrophoresed immediately. B.) 5 hours incubation. MG107 and KC8 extracts were added to pUC9 and incubated for 5 hours prior to electrophoresis.

Test for LacZ Alpha Complementation

One feature of DH5∝ is its inability to produce *lacZ* activity without alpha complementation. Since MG107 is a derivative of DH5∝ it was tested for *lacZ* activity in the presence and absence of alpha complementing plasmids. Figure 10 shows that DH5∝ and MG107 do not show *lacZ* activity in the absence of alpha complementing plasmids. HB101 shows high levels of *lacZ* activity without complementation. DH5∝ + pUC9 shows a reduced but evident level of *lacZ* activity due to alpha complementation by pUC9.

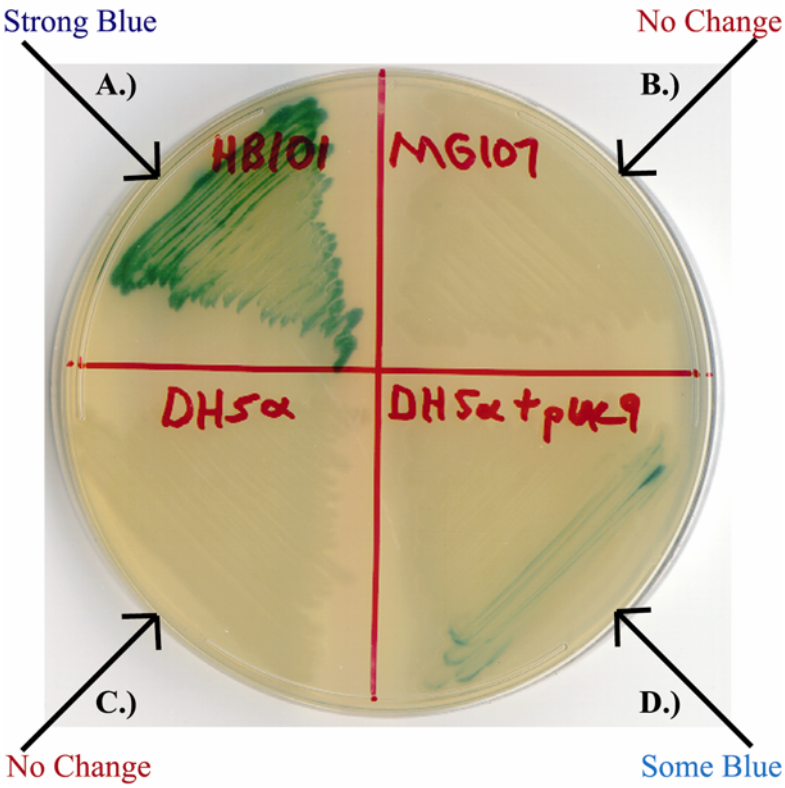


Figure 10: An Assay of LacZ Activity in Three Bacterial Strains
HB101 (A), MG107 (B), DH5α (C), and DH5α + pUC9 (D) cells were streaked onto LB+ X-gal plates, incubated overnight at 37°C and any colour change was observed.

MG107 containing either YEplac195, YEplac181, or YEplac112 showed *lacZ* activity by alpha complementation. Each of these plasmids contains the *lacZ* alpha fragment and all three allowed MG107 cells to turn x-gal blue. MG107 thus still retains the capabilities of DH5α for alpha complementation of *lacZ*.

Discussion

There are many strategies that could be employed to create the *E. coli* strain described above. Initially it seemed logical to simply produce the *endA1* mutation in KC8 by cloning it out of DH5α and inserting it into KC8. The pre-existing resistance to tetracycline and kanamycin in KC8 make it difficult to use these bacterial resistance genes as markers. This convinced us that it would be easier to start with DH5α and make it more like KC8 by disrupting the *pyrF*, *leuB*, and *trpC* genes. I was also interested in the method of gene disruption developed by Datsenko and Wanner (2000) as a quick and effective way to create a novel strain of bacteria from already existing stocks.

The Method of Gene Disruption

The one-step inactivation of *E. coli* DH5α genes was successful. There was a general increase in the efficiency of the transformation and each disruption yielded more selectable kan^R colonies than the previous one. The increase can probably be attributed to improved technique as I became more comfortable with the procedure. In particular, as I practiced the gel purification, I was able to obtain more concentrated PCR gene disruption cassette DNA. The higher concentration would allow more bacteria to take up the gene cassette and therefore result in more recombination events.

The number of recombination events that occurred at the target gene followed a less consistent pattern. Every gene disruption cassette was different with respect to its regions of homology to the *E. coli* chromosome; therefore, even if two homologous regions are the same size, a crossover at one gene target might occur at a higher frequency than at another. This could explain why the *trpC* gene disruption cassette gave

rise to trpC mutants in 100% of the kanamycin resistant colonies, whereas the other gene disruption cassettes were less efficient. The *kan*^{R} colonies that did not show disruption of the target gene (*leuB* and *pyrF*) may have incorporated the cassette at an incorrect location because of multiple partial homologies to the chromosome since it is possible for a small sequence of oligonucleotides to occur more than once in the *E. coli* chromosome. An advanced BLAST (Basic Local Alignment Search Tool) analysis did reveal a few short homologies (15 bp or smaller) to the target sequence of the gene disruption cassette other than at the target (See Appendix A). The primer portion of each cassette had no homologous matches in the *E. coli* genome (likely by design).

Another explanation for the *kan*^{R} colonies that did not have a disruption at the target gene is that the PCR gene cassette samples contained some intact template plasmid that had escaped *Dpn*I treatment and gel purification. Such plasmid molecules can remain intact and provide the colony with kanamycin resistance or be incorporated into the *E. coli* genome by unwanted recombination events (Datsenko and Wanner, 2000).

Finally, some small kanamycin resistant colonies appeared overnight and showed a slow growth phenotype. This colony type also appeared on the control plate. These colonies turned out to be false positives for gene disruption. The slow growth phenotype at first appeared to mimic leucine auxotrophy however if left long enough the bacteria did produce colonies on M9-leucine plates. The colonies also showed slow growth on LB plates. Similar slow growth colonies have been described by (Sasarman and Horodniceanu, 1967; and Datsenko and Wanner, 2000) and commonly arise on antibiotic medium that selects for small-colony-forming drug resistant mutants (*ncf*- mutants) that occur spontaneously at relatively high frequency.

All things considered, the method of disruption used seems a highly effective and relatively quick procedure for altering the *E. coli* genome in a target specific fashion. The method avoids cloning procedures, is inexpensive and unlike disruption by cloning or use of transposons, does not leave residual antibiotic resistance in the strain. The main limitation of the method is its absolute dependence on complete and accurate genome sequence information. This will limit the use of the system in some organisms (even other strains of bacteria) until the sequencing of these organisms is complete.

Genotyped MG107

The mutations I created are designated as the name of the gene involved followed by the number of base pairs removed by the gene disruption (eg. *pyrFΔ168* lacks 168 bp). The genotype presented below makes two assumptions: 1.) Each disruption has occurred in the predicted manner. 2.) All of the genotype of DH5α that was not targeted has remained unchanged.

The exact nature of each disruption could be confirmed by PCR amplification using primers that flank the target region. The PCR product generated would contain sequence surrounding the targeted region and if the disruption occurred as expected it should contain an FRT scar (See Figure 4B). By sequencing this product and comparing it to the *E. coli* genome, the exact nature of the disruption could be determined.

Some of the markers have been checked but a thorough check of all markers would be difficult. MG107 should retain the *endA1*, *recA1*, *relA1*, *deoR*, and *lacZ* mutations of its parent strain DH5α and as such should function well in many recombinant DNA procedures. The genotype is as follows:

MG107: *F*$^{-}$, *endA1*, *hsdR17* (r_K^- m_K^+), *glnV44*, *thi-1 deoR*, *recA1*, *relA1*, *supE44*, *Δ(lacZYA-argF)U169*, *λ-*, *(ϕ80dlacΔ(lacZ) M15)*, *pyrFΔ168*, *leuBΔ211*, *trpCΔ137*.

Comparison of KC8, DH5α, and MG107

Table 3 shows a summary of the phenotypic tests conducted for the MG107, DH5α, and KC8 bacterial strains. The product of my project is the bacterial strain MG107, which contains the *pyrF*, *leuB*, and *trpC* mutations and should also retain the characteristics of DH5α. These mutations will be useful for rescuing yeast-*E. coli* shuttle plasmids carrying the yeast *URA3*, *TRP1*, or *LEU2* genes. This makes MG107 ideal for use in conjunction with the yeast two-hybrid system. DNA extracts from a yeast two-hybrid positive can be transformed into MG107 and selected for ampicillin resistant colonies. The yeast two-hybrid plasmid containing the yeast *LEU2* gene and the unknown interacting protein gene can then be identified by replica plating onto M9-leucine medium. In fact, MG107 has recently been used to isolate *LEU2* plasmids from a number of two-hybrid positives previously isolated from a *Grb14* yeast two-hybrid screen (Gietz, 2001).

Since the most important feature of DH5α that we wished to retain in the new strain was its lack of *endA* activity, the *endA* activity of the new strain was tested. The activity is lower than KC8 and compares well with DH5α (data not shown). The comparison of KC8 and MG107 suggests that plasmid DNA degradation by an MG107 miniprep occurs slower than the degradation by a KC8 miniprep. As expected, the endonuclease I acted as a nickase. The super helical circular plasmid DNA (form I) is nicked and therefore becomes nicked circular (form II) or linear (form III) which both run

more slowly. After about 5 hours of incubation, the MG107 sample still contained some super helical circular DNA while the KC8 sample contained smaller amounts (See Figure 9).

Like DH5α, MG107 can be used for blue-white screening to identify mutational disruption of the *lacZ* alpha fragment in plasmid cloning vectors. Finally MG107 shows no antibiotic resistance to tetracycline, kanamycin or ampicillin, making it ideal for recombinant DNA work. MG107 compares well to DH5α and outperforms KC8 in this regard.

Table 3: A Summary of the Phenotypic Test Results.
The various phenotypic tests conducted on the final strain (MG107) and the results of similar tests on KC8 and DH5α.

	Strain and Corresponding Phenotype		
Marker	**KC8**	**DH5α**	**MG107**
Tn5 (tetracycline resistance)	some growth	no growth	no growth
kat (kanamycin resistance)	some growth	no growth	no growth
bla (ampicillin resistance)	no growth	no growth	no growth
hisB (histidine synthesis)	no growth	grows well	grows well
leuB (leucine synthesis)	no growth	grows well	no growth
trpC (tryptophan synthesis)	no growth	grows well	no growth
pyrF (uracil synthesis)	no growth	grows well	no growth
endA (endonuclease activity)	high	low	low
lacZ (ability to convert X-gal)	blue	no change	no change

MG107 can now be used to rescue two-hybrid plasmids from yeast and prepare plasmid DNA which should perform as well as DH5α in a variety of molecular techniques. MG107 performs the functions of both KC8 and DH5α and thus eliminates the need for transformation into a second strain prior to molecular manipulation of

plasmid DNA. The time required for analysis of two-hybrid positives is reduced by at least 2-3 days.

Future Directions

A mutation of the *hisB* gene could be generated in MG107. The *hisB* gene is equivalent to the yeast marker *HIS3* and this yeast marker is used in some yeast two-hybrid vectors. Although the most commonly used are the *LEU2* and *TRP1* genes, the *HIS3* gene is occasionally used, so the *hisB* gene in DH5α should be considered for disruption to make the strain as versatile as possible for researchers.

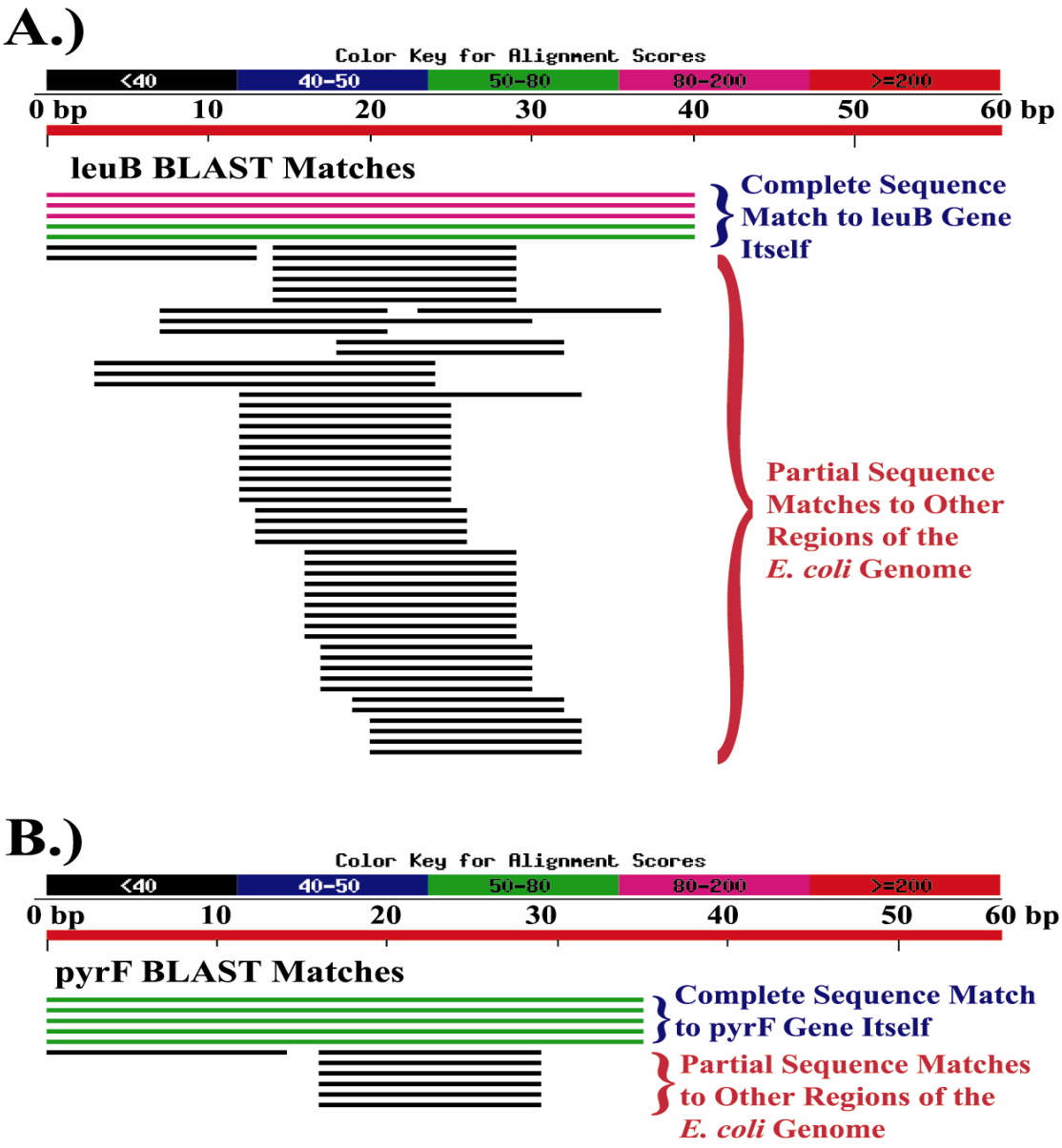
In addition to the assays already conducted to determine the endonuclease activity in MG107 more thorough *endA* assays should be conducted. A more quantitative approach could involve a similar preparation of miniprep samples to which plasmid DNA is added and incubated with magnesium for varying amounts of time. At the end of each time period, a sub-sample could be removed and EDTA used to stop the enzymes activity. This mixture could then be used to electroporate exactly 25 μl of cells and an exact volume of the cells transferred to ampicillin plates. Only cells that received an intact plasmid would be able to grow in the presence of antibiotic, so a sample that produced more colonies would indicate less endonuclease activity. Similarly, samples with high endonuclease activity should have a reduced number of intact plasmids and therefore would produce fewer colonies. The number of colonies produced by each strain would provide a quantitative comparison of the *endA* activity in each strain.

An interesting extension of the work on MG107 could involve creating a new mutation of the *endA* gene. I have found references stating that the current *endA1* mutation (a point mutation) is leaky and allows residual endonuclease activity

(Cherepanov and Wackernagel, 1995; and Durwald and Hoffmann-Berling 1968). Although this *endA* activity is much lower than the wild type level of activity, and easily controlled, it might still be worthwhile to disrupt the gene further. Since the gene has no compulsory biological function for the bacteria the entire sequence could be removed and the new strain compared quantitatively to DH5α for residual endonuclease I activity (Durwald and Hoffmann-Berling, 1968).

Finally, the ultimate goal of my project was to create a strain useful for the rescue of yeast two-hybrid plasmids. MG107 does perform this function and hopefully it will be used successfully in the next two-hybrid screen conducted in the Gietz lab. The true test will be the quality of the sequence generated using MG107 miniprep DNA. If it performs as expected the strain will be made available to other research groups via CGSC.

Appendix A: Summary of BLAST Results for *pyrF* and *leuB* Primers



A.) The first five bars represent sequence matches to different citations of the *leuB* gene itself. The next 44 bars are matches to different regions of the *E .coli* chromosome. These do not represent 44 different regions of homology as many are repeats of the same sequence from different citations. The longest exact match was 15 bp.
B.) The first five bars represent matches to the *pyrF* gene itself. The next six are matches to other regions of the *E. coli* chromosome (longest exact match was 15 bp).
Note: In both A and B no matches for the last 20 bp (the priming sequence) were found.

References

Berger, I. and Cohen, A. 1989. Suppression of recA deficiency in plasmid recombination by bacteriophage lambda beta protein in RecBCD- ExoI- *Escherichia coli* cells. *Journal of Bacteriology* 171:3523-3529.

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 7:1513-123.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., and Shao Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453-174.

Chen, Y., Narendra, U., Iype, E.L., Cox, M.M. and Rice, A.P. 2000. Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. *Molecular Cell* 6:885-897.

Cherepanov, P.P. and Wackernagel, W. 1995. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158:9-14.

Dabert, P. and Smith, G.R. 1997. Gene replacement with linear DNA fragments in wild-type *Escherichia coli*: enhancement by Chi sites. *Genetics* 145:877-89.

Datsenko, K.A. and Wanner, B.L. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc.Natl.Acad.Sci.U.S.A* 97:6640-665.

Dower, W.J., Miller, J.F. and Ragsdale, C.W. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Research* 16:6127-645.

Durwald, H. and Hoffmann-Berling, H. 1968. Endonuclease I-deficient and Ribonuclease I-deficient Escherichia coli Mutants. *J.Mol.Biol.* 34:331-346.

Fields, S. and Song, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature.* 340:245-26.

Fong, W.G., Liston, P., Rajcan-Separovic, E., St Jean, M., Craig, C. and Korneluk, R.G. 2000. Expression and genetic analysis of XIAP-associated factor 1 (XAF1) in cancer cell lines. *Genomics* 70:113-22.

Gietz, R.D. *Personal Communication*, 2001. (UnPub)

Gietz, R.D. and Sugino, A. 1988. New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* 74:527-34.

Gietz, R.D., Triggs-Raine, B., Robbins, A., Graham, K.C. and Woods, R.A. 1997. Identification of proteins that interact with a protein of interest: Applications of the yeast two-hybrid system. *Molecular and Cellular Biochemistry* 172:67-79.

Girvitz, S.C., Bacchetti, S., Rainbow, A.J. and Graham, F.L. 1980. A rapid and efficient procedure for the purification of DNA from agarose gels. *Analytical Biochemistry* 106:492-46.

Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology* 177:4121-4130.

Hamilton, C.M., Aldea, M., Washburn, B.K., Babitzke, P. and Kushner, S.R. 1989. New method for generating deletions and gene replacements in *Escherichia coli*. *Journal of Bacteriology* 171:4617-422.

Hanahan, D., Jessee, J. and Bloom, F.R. 1991. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods in Enzymology* 204:63-113.

Hemming, R., Agatep, R., Badiani, K., Wyant K., Arthur G., Gietz R.D. and Triggs-Raine B. 2001. Human growth factor receptor bound 14 binds the activated insulin receptor and alters the insulin-stimulated tyrosine phosphorylation levels of multiple proteins. *Biochem.Cell.Biol.* 79:21-32.

Kalchman, M.A., Graham, R.K., Xia, G., Koide H.B., Hodgson J.G., Graham K.C., Goldberg Y.P., Gietz R.D., Pickart C.M., and Hayden M.R. 1996. Huntingtin is ubiquitinated and interacts with a specific ubiquitin-conjugating enzyme. *Journal of Biological Chemistry* 271:19385-1994.

Kato, C., Ohmiya, R. and Mizuno, T. 1998. A rapid method for disrupting genes in the Escherichia coli genome. *Biosci.Biotechnol.Biochem.* 62:1826-189.

Maniatis, T., Fritsch, E.F. and Sambrook, J. *Molecular Cloning, A Laboratory Manual (2nd ed.)*, New York: Cold Spring Harbor Laboratory, 1989.

Murphy, K.C. 1998. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *Journal of Bacteriology* 180:2063-271.

Ratzkin, B. and Carbon, J. 1977. Functional expression of cloned yeast DNA in Escherichia coli. *Proc.Natl.Acad.Sci.U.S.A* 74:487-91.

Sasarman, A. and Horodniceanu, T. 1967. Locus Determining Normal Colony Formation on the Chromosome of *Escherichia coli* K-12. *Journal of Bacteriology* 94:1268-1269.

Schoenfeld, T., Mendez, J., and Storts, D.R., 1995. Effects of Bacterial Strains Carrying *endA1* Genotype on DNA Quality Isolated with Wizard(TM) Plasmid Purification Systems. *Promega Notes Magazine* 53:12-15.

Struhl, K. *Personal Communication*, 2000. (UnPub).

Sumida, K., Saito, K., Ooe, N., Isobe, N., Kaneko, H. and Nakatsuka, I. 2001. Evaluation of in vitro methods for detecting the effects of various chemicals on the human progesterone receptor, with a focus on pyrethroid insecticides. *Toxicology* 118:147-55.

Ullsperger, C.J. and Cox, M.M. 1995. Quantitative RecA protein binding to the hybrid duplex product of DNA strand exchange. *Biochemistry* 34:10859-1066.

Waite, L.L. and Cox, M.M. 1995. A protein dissociation step limits turnover in FLP recombinase-mediated site-specific recombination. *Journal of Biological Chemistry* 270:23409-23414.

Watt, V.M., Ingles, C.J., Urdea, M.S. and Rutter, W.J. 1985. Homology requirements for recombination in *Escherichia coli*. *Proc.Natl.Acad.Sci.U.S.A* 82:4768-472.

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z., and Graham M.W. 1989. Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Research* 17:3469-3478.

Yanisch-Perron, C., Vieira, J. and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-19.

Zhang, Y., Buchholz, F., Muyrers, J.P. and Stewart, A.F. 1998. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genetics* 20:123-18.