2003

DarA, The Antirestriction Protein of Bacteriophage P1

Priti R. Vemulapalli

DarA, the Antirestriction Protein of Bacteriophage P1

Priti R. Vemulapalli

Submitted in partial fulfillment of the requirements for the
degree of Master of Science in Microbiology from the
Department of Biology of Seton Hall University
May 2003

APPROVAL PAGE

APPROVED BY

MENTOR
Dr. Sulie Chang

COMMITTEE MEMBER
Dr. Allan Blake

GRADUATE ADVISOR
BIOLOGY DEPARTMENT
Dr. Eliot Krause

GRADUATE ADVISOR
BIOLOGY CHAIRPERSON
Dr. Sulie Chang

Acknowledgements

I would like to thank my mentor, Dr. Maria P. MacWilliams, for her guidance and support during and after my research project. I would also like to recognize my family for their support in completing my research and my thesis.

Table of Contents

Tables

Figures

Abstract

Bacteria possess restriction modification (R-M) systems which provide a means of defense against bacteriophage attack. The R-M system comprises of two components: a DNA methyltransferase and a restriction endonuclease. The DNA methyltransferase modifies the host DNA within a specific sequence and the restriction endonuclease cleaves DNA which lacks this modification. There are three classes of R-M systems - types I, II, and III - which differ considerably in their gene and protein structure, functional requirements, and modes of action.

To protect themselves from cleavage, some bacteriophages encode antirestriction systems. We are interested in the Dar antirestriction system of phage P1, a temperate bacteriophage which infects *Escherichia coli*. Bacteriophage P1 is only weakly restricted by many of the type I R-M systems found in *E.coli* due to the actions of two, phage-encoded proteins, DarA and DarB. DarA is functional alone and is known to inhibit the IB subgroup of type I restriction endonucleases. The *darA* gene is expressed late in infection and the resulting protein is packaged in the bacteriophage head along with DNA. Upon infection of a new cell, DarA is coinjected with the DNA and protects the DNA from cleavage by type I R-M systems. As part of the viral particle maturation, the DarA protein is processed by the P1-encoded protease. Proteolytic cleavage releases a small fragment from the amino terminal end.

To investigate the mechanism of DarA antirestriction, we constructed two DarA expression vectors. In each case, the gene of interest was under the control of the arabinose-inducible promoter, p_{BAD}. Plasmid pRR2000 contained the entire *darA* gene while pRR2001 contained a truncated version of the gene encoding the large, carboxyl

fragment of the processed form of the protein. From DNA sequencing and bioinformatics we were able to determine that the cloned DNAs had the correct sequence compared to that of published *darA sequence*. Basic Local Alignment Search Tool (BLAST) and Fast Alignment Search Tools Anything (FASTA) analysis revealed no close DarA relatives. Analysis of *in vivo* protein expression demonstrated that both plasmid systems produced proteins of the predicted sizes and that expression was dependent on the addition of arabinose. We also found that the full-length form and the processed DarA protected incoming phage DNA from restriction. This is the first evidence that DarA can function *in trans* to protect phage from restriction.

The global significance of this research is that bacteriophage P1, equipped with its antirestriction activity, can undermine the defense mechanism of bacteria. Bacteriophage P1 is an efficient transducing phage, and thus facilitates the horizontal transfer of bacterial genes. This means of transfer is a stepping-stone in the evolution of microbes.

Introduction

Bacteria are subject to attack by intracellular parasites, viruses, which are also termed bacteriophages/phages. A single bacterial species can fall prey to a large number of different bacteriophages. In order to survive, bacteria have acquired several different strategies that prevent or inhibit phage infection. Likewise, phage have evolved mechanisms to circumvent the host's defense. We have set out to analyze the complex relationship between bacteriophage P1, and its host organism, *Escherichia coli*.

Bacteria are capable of protecting themselves against bacteriophages at almost all stages of the phage life cycle (26). One highly effective defense is to prevent any productive contact between bacteria and phage. This can be accomplished by the genetic selection of cells that lack the specific phage receptor in the cell wall. Another means to avoid host-phage contact is the secretion of a barrier, such as a capsule or slime layer, that inhibits the approach of the phage. If the bacteriophage DNA has been injected into the cell, the bacterium still has two more survival options. In some cases, one component of the host is not suited to the needs of the bacteriophage and thus the infection will be aborted. For example, the inability of *Salmonella typhimurium* to support a productive phage lambda infection is due in part to the lack of interaction between the *Salmonella* NusA protein and the N protein of lambda (19). Another means of protection is the host cell utilization of restriction endonucleases to destroy unfamiliar DNA. This sort of immunity is referred to as restriction-modification (R-M) (26).

Bacterial-encoded R-M systems provide protection for a microorganism from foreign DNA, foreign DNA implying bacteriophage genomes as well as conjugative plasmids and DNA introduced into the cell via transformation. The restriction component, a site-specific endonuclease, recognizes and cuts unmodified DNA. The linearized DNA is then further degraded by the host cell exonucleases. The modification component of the R-M system is a DNA methyltransferase, which methylates specific adenosine or cytosine residues within the

recognition sequence. This modification allows the bacterium to distinguish its own DNA from foreign DNA, and thus prevents cleavage by endonuclease (27).

There are three classes of R-M systems, types I, II, and III, which differ considerably in protein structure and cleavage mechanisms (27). A type I system consists of a multifunctional enzyme that can catalyze both DNA methylation and restriction activity. The enzyme is composed of three different subunits: HsdR, HsdS, and HsdM, which are the protein products of *hsdR*, *hsdS*, and *hsdM*, respectively. The S subunit is responsible for recognition of the specific DNA sequence targeted by the enzymatic activities of restriction and modification (15, 16, and 17). The R subunit is strictly involved in restriction activity whereas the M that catalyzes the DNA methylation is also necessary for restriction (13).

All natural type I enzymes recognize asymmetric, interrupted DNA sequences (13). These sequences are composed of one, three-base-pair half site and another of four to five base pairs. The two half sites are separated by a nonspecific spacer of defined length, usually six to eight base pairs (Figure 1).

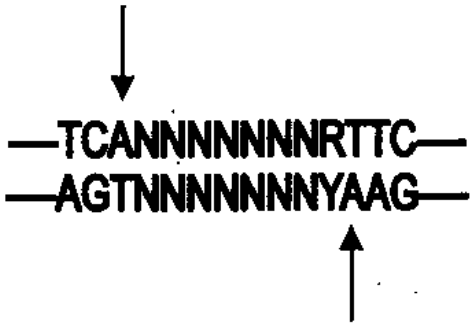

Figure 1. A typical type I R-M recognition site. Type I enzymes recognize asymmetric DNA consisting of two half sites separated by a nonspecific spacer of defined length. The recognition site for *Eco*DXXI is shown. The red arrows correspond to adenine residues that are methylated.

Although methylation is on adenine residues on both strands of the recognition site, DNA restriction occurs randomly *in cis* up to thousands of bases away. The primary function of the methylation activity is to protect the bacterial cell from the corresponding restriction endonuclease. R-M systems require ATP, S-adenosyl methionine (AdoMet) and magnesium (5)

for DNA cleavage. AdoMet is the sole cofactor required for the methylation reaction where it serves as the methyl donor.

Type I systems are further subdivided into families of related enzymes, designated IA, IB, IC, and ID (4). Family members share significant sequence similarity between the structural genes. Mutated genes of one member of the family can be complemented by the wild type alleles from another member. Within a family, the *hsdM* and *hsdR* genes are largely identical throughout their length. On the other hand *hsdS* genes contain two regions of high sequence similarity separated by two extensive, nonconserved regions. It is thought that the regions of high variability code for two protein domains, each of which recognizes one half site of the split recognition sequence characteristic of type I R-M systems (5).

Besides several short amino acid sequence motifs found in HsdM which are common to DNA adenine methyltransferases and a few ATP-binding motifs found in HsdR, there is no sequence homology between structural genes of different families. Thus complementation does not take place between different families.

We utilized bacteriophage P1 in our studies (Figure 2). A generalized transducing phage that can encapsidate up to 100 kb of DNA into its viral particles, P1 also has a broad host range. The wild-type phage has a genome of approximately 90 kb. Its DNA is double-stranded, linear, terminally redundant, and circularly permuted (23). The phage capsid has icosahedral symmetry, with a nonflexible tail. The tail is attached to one of the vertices of the capsid by a head-neck connector and a hollow neck. The tail is slightly thinner than the neck, and consists of a tube and a contractile sheath. The sheath has a cap at the end near the head, at its other end there is a base plate to which there are attached kinked tail fibers. There are two morphological variants of bacteriophage P1. The P1B variant has particles with extended sheaths. The other variant, P1S, produces particles with contracted sheaths (37).

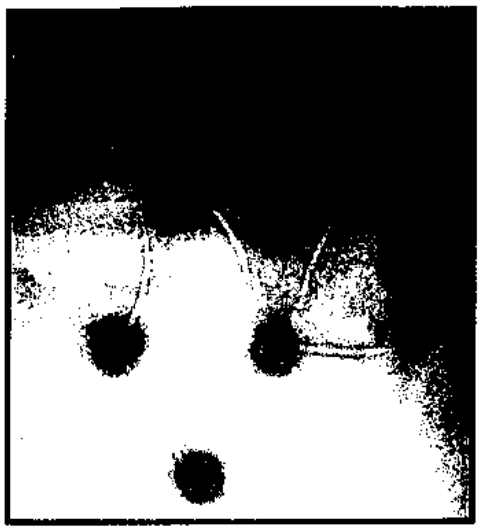

Figure 2. Electron micrograph of negatively-stained coliphage particles. Bacteriophage P1 is a member of the lamboid group of temperate phages which also includes lambda (shown here) P22, and HK022. This group shares a similar particle morphology. In addition, gene order and DNA sequence is somewhat conserved. (micrograph from http://www.biochem.wisc.edu/inman/empics/0022a.jpg)

As a DNA-containing phage that infects *E. coli*, P1 is a potential target for the host restriction endonucleases; however, P1 is only weakly restricted by many of the type I R-M systems found in *E. coli*. Two P1 genes, *darA* and *darB* are responsible for the observed restriction resistance. The *darA* gene product alone protects the phage against the IB family of enzymes exemplified by the enzyme *Eco*AI. The *darB* gene product together with *darA* protects against the IA family including *Eco*KI (22).

The *darA* gene is part of an operon expressed late in the phage lytic cycle. Both DarA and DarB polypeptides are phage head internal proteins and are injected into the cell being infected along with the DNA. Thus the Dar proteins do not provide antirestriction activity in the cells in which they first are synthesized. Instead, these proteins are active in the subsequently infected cells where they protect against type I R-M systems. It is hypothesized that the Dar system only protects the DNA with which it is copackaged (31).

As part of the viral particle maturation, the DarA protein is processed by the P1-encoded

protease. Proteolytic cleavage produces a small, amino-terminal fragment. The fate of this fragment is unknown. Presumably the processed protein is the active form (31).

The present project was designed to determine whether or not the DarA polypeptide is able to protect *in trans* against Type IC R-M systems. As a first step, we cloned the *darA* gene into an expression vector. To investigate whether both the full length DarA polypeptide as well as the shorter, processed version possess antirestriction activity, we constructed two plasmids. The first plasmid, pRR2000, is a pBAD22 derivative that contains the full-length *darA* gene under the control of the regulatable p_{BAD} promoter and a strong, translational initiation signal. The second construct, pRR2001, is virtually identical except that the *darA* sequence begins just after the hypothetical protease cleavage site. The p_{BAD} promoter is tightly regulated with its inducer, arabinose. Hence, in absence of arabinose and with glucose in the medium, the promoter is shut off with no detectable expression. Thus p_{BAD} is superior to the commonly used lactose promoter (p_{lacUV5}) or the synthetic p_{tac} promoter which show low levels of expression under noninducing conditions (18). The pBAD22-derived vectors allow us to regulate *darA* expression. With these vectors it is possible to induce expression of *darA* and monitor the effects of this protein on both the host cell and the resident R-M systems.

The relationship between bacteria cell and phage is a simple model for host-pathogen interactions. Bacteriophages pose a threat to any industry utilizing large scale fermentations (8). In addition, bacteriophages have been proposed as detection agents and therapeutic treatments for bacteria such as *Bacillus anthracis, Staphylococcus aureus,* and *Mycobacterium tuberculosis* (1, 24, 28, and 30). Understanding the complexity of how viruses can survive bacterial defense mechanisms, may lead to the development of innovative methods to inhibit cell lysis.

Phages (and viruses in general) have been proposed to play a pivotal role in evolution. As P1 is a transducing phage, a small number of phage particles released after cell lysis actually

contain host DNA rather than viral DNA. In doing so, these transducing phages have the capability of transferring the encoded genetic attributes to the next host. Such a function may be critical in the shuttling of antibiotic resistance plasmids and has been shown to be a conduit of pathogenicity factors between normal bacterial flora and pathogenic organisms (11 and 33).

One perfect illustration is *Vibrio cholerae*. *Vibrio mimicus* is a reservoir for a filamentous, lysogenic bacteriophage, CTXΦ. Studies have shown that this bacteriophage, and the bacterium may play a significant role in converting *V. cholerae* to a toxigenic strain (7, 9, 12, and 14). Evidence suggests that *V. mimicus* obtained a pathogenicity island from this transducing phage which was then transferred horizontally to *V. cholerae* (7 and 25). In another study, scientists hypothesize the pathogenic strain of *V. cholerae* O139 arose via infection of an O139 $CTX^{ET}\Phi$ lysogen by $CTX^{CALC}\Phi$ (10). This hypothesis is strongly supported by the fact that bacteriophage-mediated transfer of the toxigenic property from a non-pathogenic strain to a pathogenic one has been demonstrated in the laboratory. Antirestriction activity would certainly facilitate horizontal transfer. By understanding antirestriction, novel methods may prevent nonpathogenic bacterial strains from transforming into pathogenic ones.

Materials & Methods

Bacterial Strains, Plasmids, Bacteriophages, and Media

The strains used for this study are listed in Table I. Standard *E. coli* culturing methods were followed and Luria-Bertani (LB) broth and agar plates were prepared as described by Sambrook *et al* (1989). Antibiotics were added to LB broth and agar plates when necessary for selective growth of the desired bacterial strains. Ampicillin (Ap) was used at a concentration of 100 μg/ml, and Chloramphenicol (Cm) was used at 30 μg/ml. Plasmids constructed prior to this study are listed in Table 2. Those plasmids which were constructed during this project are listed in Table 3.

Table 1. Bacterial Strains

Name of Strain	*Genotype*	*Reference*
DH5α (R^-M^+)	F⁻*end*AI *hsd*R17($r_k^-m_k^+$) *sup*E44 *thi*-1 *rec*A1 *gyr*A (Nal^r) *rel*A1 Δ(*lac*ZYA⁻*arg*F)U169*deo*R(ϕ80d*lac*Δ(*lacZ*)M15)	Woodcock *et al.* (1989)
LMG194 (R^+M^+)	F⁻Δ*lac*X74 *gal*E *thi rps*L Δ*pho*A (*pvu*II) Δ*ara*714*leu*::Tn10	Guzman *et al.* (1995)
WA921 ($R_k^-M_k^-$)	*thr leu met*	Wood (1966)
WA2379 ($R_s^+M_s^+$)	*leu*⁻ *met*⁻ *lac*⁻ *tdr hsd*A$^+$	Arber *et al.* (1970)

Table 2. Plasmids

Plasmid	*Relevant Characteristics*	*Reference*
pSHI92	19.3 Kb *Hin*DIII from bacteriophage P1CmSmKm inserted into *Hin*DIII site of pBR322; Ap^R Sm^R Tc^R	Iida *et al.* (1982)
pBAD22	p_{BAD} multiple cloning site; *araC*$^+$; Ap^R	Guzman *et al* (1995)

Table 3. Plasmids Constructed During Research

Plasmid	*Characteristics:* *Plasmid Size, Insert, & Antibiotic Resistance*	*Vector & Cloning Site**
pRR99	6694 bp; *darA* from pSHI92; Ap^R	pBAD22; *Eco*RI-*Sma*I of MCS
pBAD22-Start	5089 bp; PCR fragment including bp 2424 to 2963 of *darA* operon; Ap^R	pBAD22; *Nco*I-*Sma*I of MCS
pBAD22-Protease	4859 bp; PCR fragment including bp 2628 to 2963 of *darA* operon; Ap^R	pBAD22; *Nco*I-*Sma*I of MCS
pRR2000	6562 bp; *Kpn*I-*Hin*DIII from pRR99 (contains the 3'end of *darA*); Ap^R	pBAD22-Start; *Kpn*I (*darA*) - *Hin*DIII (plasmid MCS)
pRR2001	6323 bp, ; *Kpn*I-*Hin*DIII from pRR99 (contains the 3'end of *darA*); Ap^R	pBAD22-Protease; *Kpn*I (*darA*) - *Hin*DIII (plasmid MCS)

* MCS = multiple cloning site

Molecular Biology Reagents

The following reagents were purchased from New England Biolabs: restriction endonucleases, Vent™ thermostable DNA polymerase, Calf Intestinal Phosphatase (CIP) and deoxyribonucleotides. Bacteriophage T4 DNA ligase was obtained from Life Technologies. Amersham Biosciences was the source for deoxyribonucleotides, as well as for reagents for polyacrylamide gel electrophoresis (PAGE) of proteins (protein molecular weight markers, SDS PAGE buffer strips and SDS-PAGE PHAST gels). Protein molecular weight markers from Sigma-Aldrich were also utilized. DNA molecular weight markers were from Life Technologies and the Promega Corporation. The Geneclean kit (Qbiogene) was used for isolating DNA gel fragments and PCR products from enzymatic reactions and agarose gel slices. A variety of DNA midiprep kits were utilized (Life Technologies, Promega, Qiagen) while the Wizard Miniprep System (Promega) was used for all DNA minipreps with one exception: the Qiaprep Spin Miniprep Kit (Qiagen) was used to prepare DNA to be sequenced. All reagents and kits were used in accordance with the manufacturer's instructions. Oligonucleotides used as primers in PCR reactions or DNA sequencing were from Sigma-Genosys.

General DNA Cloning Strategies

DNA ligation mixtures were transformed into competent cells by one of two methods. Calcium Chloride-competent cells were prepared and transformed with DNA according to the protocols in Sambrook *et al.* (1989). Electroporation was also used as a means to introduce DNA into competent cells. The Biorad Gene Pulser was used for all electroporations according to the manufacturer's instructions. The transformed cells were then plated on LB + 100 μg/ml Ap and incubated overnight at 37°C to select for cells containing the plasmid. Isolated colonies were then restreaked onto LB + 100 μg/ml Ap for further colony purification. These plates were

also incubated overnight at 37°C. Liquid cultures were prepared from the resulting clones by inoculating LB + 100 μg/ml Ap with an isolated colony and incubating overnight in a rotating wheel at 37°C. The liquid cultures were used to prepare plasmid DNA minipreps which were subsequently analyzed by restriction site mapping and, in some cases, by DNA sequencing.

Protein and DNA Electrophoretic Techniques

Denaturing SDS-polyacrylamide gels were run on Amersham Biosciences' PHAST Gel System. Protein samples were prepared for electrophoresis and the equipment was used according to the manufacturer's instructions. Agarose gel electrophoresis was carried out in 1X TAE buffer (29). This buffer was utilized as the running buffer as well as for preparing the gels. Varying concentrations of agarose were used to optimize electrophoretic separation of the various sizes of DNA to be analyzed. Ethidium bromide was used in the running buffer and in the gel at a concentration of 0.5 μg/ml.

Subcloning of the darA Gene

In order to obtain a smaller, easier to manipulate plasmid, a *darA*-containing fragment from pSHI92 was subcloned into pBAD22 as described here. The 2.1 Kb fragment containing *darA* was first isolated from pSHI92 by cleaving the plasmid with *Eco*RI and *Eco*RV (Figure 3). The digest was run on a 1.2 % agarose gel. The 2.1 Kb band was excised, Genecleaned, and the isolated DNA was resuspended in TE (10 mM Tris-HCl, pH 8.0; 1 mM EDTA). Plasmid pBAD22 was digested with *Sma*I and *Eco*RI, whose recognition sites are located in the multicloning site (MCS) downstream of the pBAD22 promoter (Figure 3). The pBAD22 vector was first digested with *Sma*I, Genecleaned, resuspended in TE and then digested with *Eco*RI in *Eco*RI buffer. The digests were performed sequentially in order to ensure optimal buffer and

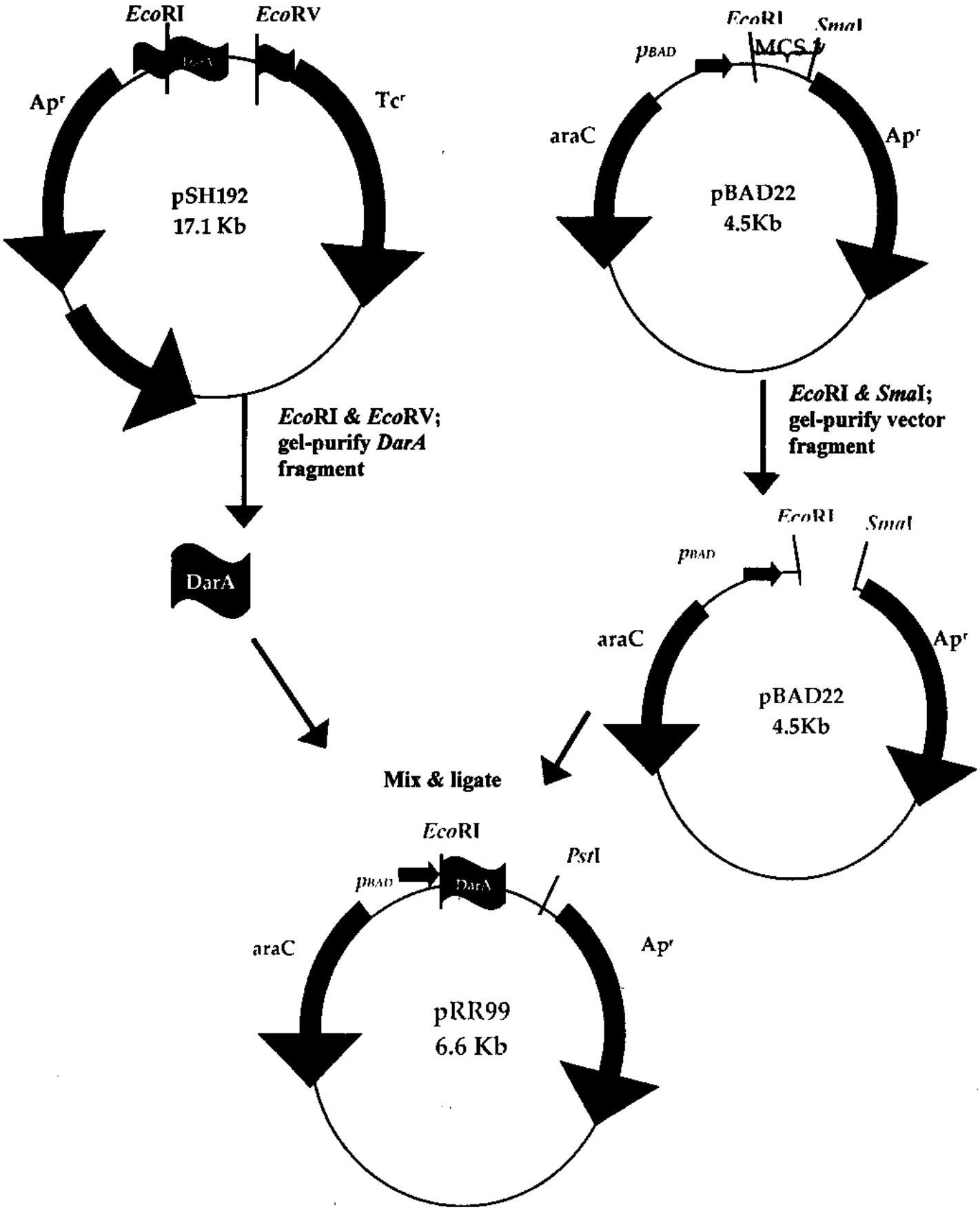

Figure 3. Construction of pRR99. Plasmid pSHI92 was digested with *Eco*RI & *Eco*RV to release a 2.1 kb fragment containing the *darA* gene. Plasmid pBAD22 served as a vector for the *darA* gene, and was digested with *Sma*I & *Eco*RI. The *darA* fragment was gel-purified and ligated to *Eco*RI-*Sma*I-digested pBAD22 to form the recombinant plasmid, pRR99 of 6.6 kb.

and resuspended in TE.

To join the fragments, the excised 2.1 Kb fragment containing the *darA* gene and the digested pBAD22 vector were mixed in a 2:1 ratio (vector: insert). The 10x DNA ligase buffer was added to a final concentration of 1x followed by the addition of 1µl T4 DNA ligase. After an overnight incubation at 15°C, the ligation reaction was transformed into *E.coli* DH5α by the Calcium Chloride method and the samples processed as previously described (See *General Cloning Strategies*). The desired recombinant plasmid, pRR99, was predicted to be 6.6 Kb long (Figure 3). To identify positive clones with the desired construct, miniprep DNAs were prepared and aliquots were digested with *Eco*RI. Miniprep DNA aliquots were also subjected to a *Pst*I-*Eco*RI double digest. Plasmid pBAD22 was used as a negative control for both sets of digests. The digest fragment sizes were analyzed by 0.8% agarose gel electrophoresis (Figure 8).

Introduction of an NcoI Restriction Site by PCR Amplification of the 5' End of DarA

The next step of our cloning strategy involved using PCR to create an *Nco*I recognition site overlapping the translational start site of *darA*. The *Nco*I site would allow placement of the gene downstream of the p_{bad} promoter in the pBAD22 plasmid. Another goal of the project was to create an artificial *darA* construct which encoded a protein lacking the amino terminal region that is normally cleaved by the bacteriophage P1 protease during phage head maturation. We opted to amplify only the 5' end of the gene rather than the entire gene to decrease the chance of picking up a mutation due to polymerase error. The PCR fragment, which encompassed the *dar* sequence from the native initiation codon to just downstream of the *Kpn*I within *darA* (base pair 2424 to 2963 of the *darA*/P1 sequence), was termed "Start". The "Protease" PCR product contained from base pairs 2628 to 2963 of *darA* thus lacked the first 68 codons of the gene.

Plasmid pRR99 was the DNA template for both PCR reactions. The PCR primers used are listed in Figure 4.

Forward Primer for Start site: GAGACGCCCATGGAACAGTTCAATATC

Forward Primer for Protease site: GCTTGACCCATGGCGATTAGTTCTACTC

Reverse Primer for both sites: TAGTGAACCCGGGAGCAACTGTGTGA

Figure 4. PCR primers utilized to amplify the start site and the theoretical protease site of the *darA* gene. Restriction endonuclease recognition sites (*Nco*I, *Sma*I) used for cloning the PCR products are in bold, red type. The ATG triplet intended to serve as the initiation codon in the recombinant gene is underlined.

Three PCR reactions were setup: one for the protease fragment, another for the start fragment, and the third as the negative control. Each reaction was set up with: 10 ng of template DNA, 40 μmoles of each dNTP, 1X Vent polymerase buffer (10 mM KCl; 20 mM Tris-HCl, pH 8.8 at 25^{O}C; 10 mM $(NH_4)_2SO_4$; 2 mM $MgSO_4$; 0.1% Triton X-100), 1X BSA, and 0.5 μl of Vent® DNA polymerase. The protease and start PCR reactions contained 50 μmoles of the reverse primer and 50 μmoles of their respective forward primers (Figure 4). The negative control did not contain any primers. Each reaction was overlaid with mineral oil, and was amplified in the Hybrid Omni Gene thermocycler following the program listed in Table 4. To confirm that the PCR products were the correct sizes, 5 μl of each reaction was run on a 1.2% agarose gel (Figure 9). The start and the protease products are 561 and 360 bp, respectively.

Table 4. Thermocycler Program

Stage	# Cycles	Step	Process	Temp (°C)	Time (min)
1	1	1	Denaturation	92	3.0
2	5	1	Denaturation	92	1.0
		2	Annealing	38	0.5
		3	Synthesis	72	0.5
3	20	1	Denaturation	92	1.0
		2	Annealing	68	0.5
		3	Synthesis	72	0.5
4	1	1	Synthesis	72	5.0

Reconstruction of the darA gene under the control of the pBAD promoter

The PCR fragments were cloned separately into pBAD22. The start fragment, the protease fragment, and pBAD22 all were digested with restriction enzymes *Nco*I and *Sma*I. After digestion, pBAD22 was treated with calf alkaline phosphatase (CIP) to prevent vector recircularization. Prior to the ligation reaction, the inserts and the vectors were Genecleaned. The dephosphorylated vector was then ligated to the start fragment and the protease fragment separately. Both ligation reactions were set up with an insert to vector ratio of 5:1. The ligation mixtures were subsequently transformed into DH5α cells and the resulting Ap^r clones were processed and DNA isolated as described *General DNA Cloning Strategies*. The new recombinant plasmids, pBAD22-START and pBAD22-PROTEASE (Figure 5) had a total size of 5.0 Kb and 4.8 Kb, respectively. For confirmation, plasmid DNA aliquots from the transformants were cleaved with *Hpa*I and *Pvu*I separately and the digested DNAs were analyzed by electrophoresis through a 0.8% agarose gel (Figures 10 and 11, respectively). Plasmid pBAD22 was used as a negative control.

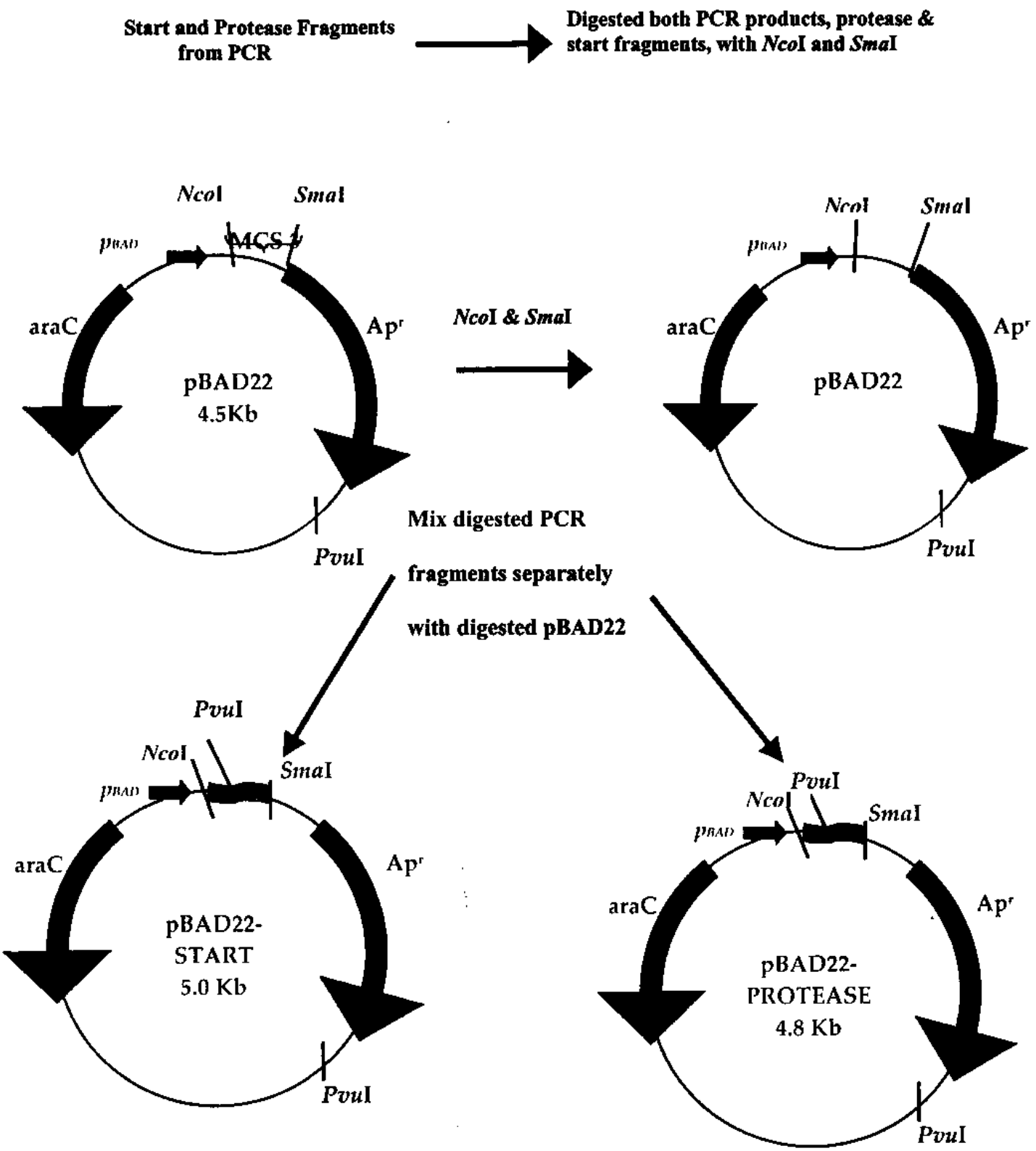

Figure 5. Cloning the PCR products. PCR was used to amplify the protease and start fragments of *darA* while at the same time incorporating an *Nco*I site overlapping the start codons. The PCR products and the pBAD22 cloning vector were digested with *Nco*I and *Sma*I. The digested PCR products were cloned separately into the digested pBAD22. When the amplified start fragment and pBAD22 were mixed and ligated, the resulting recombinant plasmid was pBAD22-START of 5.0 kb. The fusion of the amplified protease fragment with pBAD22 formed pBAD22-PROTEASE of 4.8 kb.

Plasmid pRR99 was digested with *Kpn*I and *Hin*DIII forming two fragments 4.9 kb and 1.7 kb. Plasmid pBAD22-START was also digested with the same enzymes to produce fragments of 4.8 kb and 0.29 kb. Plasmid pBAD22-PROTEASE was cut with the same enzymes as well, and produced the following fragment sizes: 0.29 kb and 4.6 kb. The 1.7 kb pRR99 fragment which contained the 3'end of *darA* was then ligated to the 4.8 kb pBAD22-START vector fragment which contained the 5'end of the gene. The insert to vector molar ratio was 3:1. The new recombinant plasmid, pRR2000, is a total of 6.5 kb (Figure 6). The 4.6 kb pBAD22-PROTEASE fragment which contained the 5'end of the *darA* gene was ligated to the 3' end of the 1.7 kb pRR99 fragment to produce pRR2001. The insert to vector molar ratio was 3:1. This new recombinant fragment is a total of 6.3 kb (Figure 7). To confirm that both vectors contained the insert, the recombinant plasmids were digested with *Nco*I. The fragments from the restriction digest were observed by carrying out electrophoresis using a 0.8% agarose gel (Figure 12: pRR2000, Figure 13: pRR2001).

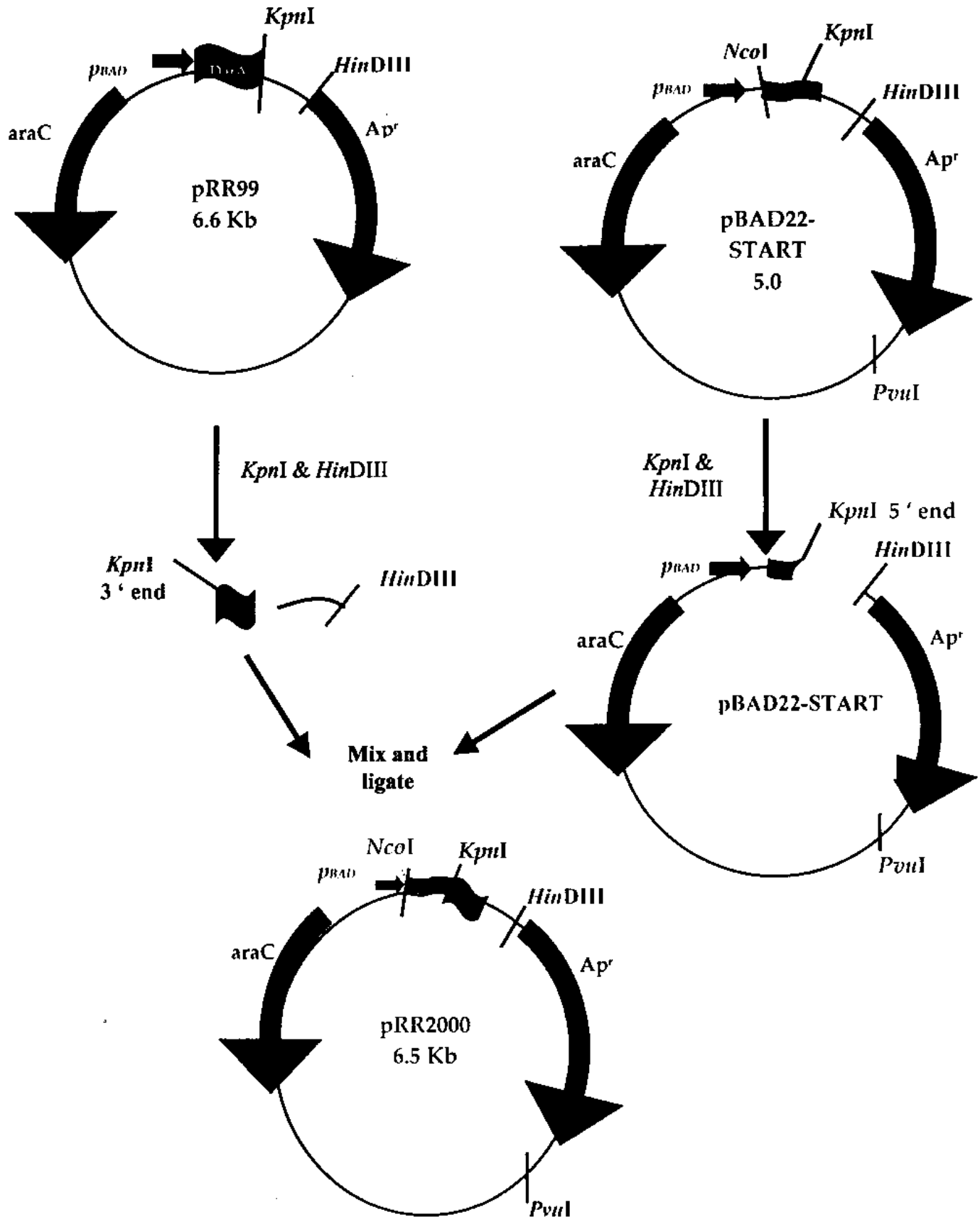

Figure 6. Construction of pRR2000. Plasmids pBAD22-START & pRR99 were digested with *Hin*DIII & *Kpn*I. The smaller fragment of pRR99 was ligated to the larger fragment of pBAD22-START. The new recombinant plasmid was named pRR2000 and was 6.5 kb long.

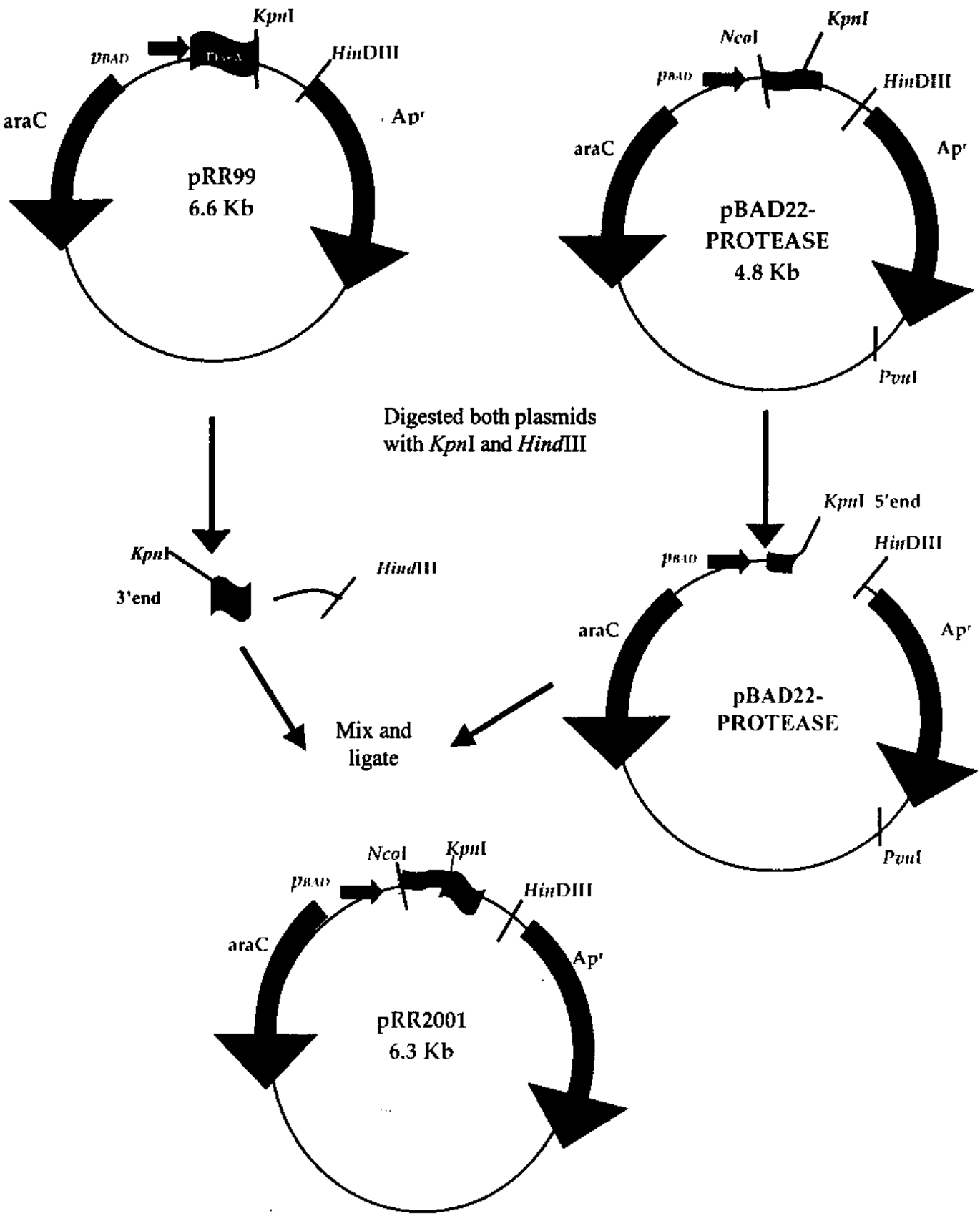

Figure 7. Construction of pRR2001. Both plasmids, pBAD22-PROTEASE & pRR99 were digested with *HinD*III & *Kpn*I. The smaller fragment of pRR99 was ligated to the larger fragment of pBAD22-PROTEASE. The new recombinant plasmid was named pRR2001 for a total of 6.3 Kb.

Sequence Analysis

Before sending samples of pRR2000 and pRR2001 to Arizona State University for sequencing, they were PEG (polyethylene glycol) precipitated. The protocol was suggested by Dr. Scott Bingham from the University. Miniprep DNAs were ethanol precipitated and then resuspended in 50 µl of sterile, deionized water. Thirteen microliters of 4M NaCl were added followed by 63 µl of 13% PEG 8000. The preps were mixed by vortexing and were kept on ice for 20 minutes. The samples were spun at 12,000 RPM for ten minutes. The supernatant was discarded and the pellet was washed with one ml of cold 70% ethanol. After the DNA pellets dried, the samples were then shipped to the Arizona State University sequencing facility where the samples were analysed on an ABI Prism 377 DNA sequencer.

The data obtained allowed us to compare the cloned sequences to that of the published *darA* gene. Using the internet, and going on the National Center for Biotechnology Information (NCBI) website, a Basic Local Alignment Search Tool (BLAST) search was done to compare the cloned sequences of the two plasmids to that of the d*arA* gene. The search was also done to confirm whether the *Nco*I site was indeed cloned into pRR2000 and pRR2001.

Protein Expression

Using LMG194 cells, transformations were performed separately with the following plasmids: pRR2000, pRR2001, and pBAD22. The resulting plasmid-bearing strains were colony-purified and the presence of the correct plasmids verified as described under *General Cloning Strategies*. Since pBAD22 lacks the *darA* gene, it served as a negative control for our protein studies. The parameters used as a starting point for studying protein expression were those suggested by Invitrogen (Invitrogen 1998).

The transformed strains, LMG194/pBAD22, LMG194/pRR2000, and LMG194/pRR2001, were each analysed for arabinose-inducible protein expression as follows. A 1.5 ml aliquot of LB broth + 100 μg/ml Ap was inoculated with an isolated colony and was grown overnight on a rotating wheel at 37°C. The following morning the culture was subcultured at a dilution of 1:100 and incubated on a wheel at 37°C until the culture reached mid-log phase (OD_{600nm} = 0.4; approximately 2.5 hours). At this point, the culture was split in half. Half the culture was induced with arabinose at a final concentration of 0.2% and the other half was not. Induced cultures contained: 4 mls of the 1:100 subculture + 12 mls of LB + 8 μl of 200 mg/ml ampicillin + 160 μl of 20% arabinose. The uninduced cultures had the same components as the induced with the exception of arabinose. A 1.5 ml aliquots were taken of each culture and placed on ice as the zero time point. Cultures were then incubated on a rotating wheel at 37°C. At time points one hour, two hours, four hours, and overnight, 1.5 mls were taken of the induced and uninduced of all three cell lines. Half a milliliter of these aliquots were used to monitor cell growth by OD_{600nm} measurement while the remaining one milliliter was prepared for gel electrophoresis as described below.

Each culture aliquot was centrifuged for ten minutes in a microcentrifuge at maximum speed to pellet the cells. The supernatant was removed and discarded. The pellet was resuspended in 100 μl of 1X SDS-PAGE sample buffer (29). The sample was then heated for 95°C for five minutes, vortexed to ensure proper mixing, and centrifuged for ten minutes. Samples were loaded onto an 8-25% gradient SDS PHAST gel. The gel was run on the Amersham Biosciences PHAST gel separation system. The gel was stained in Coomassie Blue Stain for one hour on a platform rocker, and then destained overnight in a destain solution (20% Glacial Acetic Acid, 60% Methanol).

Restriction Assays

For all stages of the assay, the WA921 cultures were grown in LB while WA2379/pBAD22, WA2379/pRR2000, and WA2379/pRR2001 liquid cultures were prepared in LB supplemented with 100 μg/ml Ap to select for plasmid maintenance. Overnight cultures were diluted 1:100 in fresh medium. Arabinose was added to a final concentration of 0.2% when necessary to induction of *darA* expression. These diluted cultures were grown to late log phase (OD_{600nm} ~0.8). Cell culture aliquots were mixed with Luria top agar and plated on LB plates. After the agar had solidified, serial dilutions of a high-titer phage lambda lysate were spotted in 5 μl aliquots. The plates remained at room temperature until the liquid had been absorbed and were subsequently incubated at 37°C. The following morning the plates were observed to ascertain which dilutions produced countable numbers of plaques and the number of plaque forming units/ml (PFU/ml) was calculated for each of the bacterial stains analysed. To evaluate the effectiveness of DarA antirestriction the efficiency of plating (EOP) was calculated. EOP is defined as the PFU for the tester culture divided by the PFU/ml for the nonrestricting strain. The larger the EOP value, the weaker the restriction. In addition, uninduced and induced cultures of the same bacterial strain were also compared to see if any antirestriction activity detected was dependent on *darA* induction.

Results

Subcloning of the darA Gene

Antirestriction systems protect the bacteriophages and conjugative plasmids that encode them from DNA cleavage by bacterial R-M systems. With one recent exception (26), little is known about their mechanism of action. As a first step toward understanding the antirestriction system of phage P1, we set out to clone the P1 *darA* gene and to analyze its ability to inhibit type I restriction endonucleases.

The *darA* gene is part of an operon that is transcribed in the late phase of P1 infection; expression requires specific, phage-encoded transcription factors. In order to study the *darA* gene product apart from the other P1 functions, we needed to construct a plasmid in which *darA* was under the control of a heterologous, inducible promoter. Plasmid pSHI92 was used as the starting point in our cloning procedures. A pBR322 derivative, pSHI92 contains a 19.3 Kb *Hin*DIII fragment from a phage P1derivative (P1CmSmKm) cloned into the analogous site of the vector. Although the *darA* operon is present on pSHI92, the 23.7 Kb size of the plasmid makes its use somewhat problematic for DNA manipulation. The vector, pRR99 (6.6 Kb), was constructed to simplify future cloning strategies. This intermediate plasmid was constructed by insertion of the *darA*-containing 2.1 Kb *Eco*RI- *Eco*RV fragment from pSHI92 into the *Eco*RI-*Sma*I sites of pBAD22.

To confirm that we had the recombinant plasmid with the correct insert and vector backbone, restriction digests were carried out using miniprep DNAs from potential positive clones. Two different restriction digests were performed and analyzed by agarose gel electrophoresis (Figure 8). One was a single enzyme digest with *Eco*RI, and another was a double digest with *Pst*I and *Eco*RI. pBAD22 served as the no-insert

control DNA. Digesting pBAD22 with *Eco*RI produced the linear 4.5 Kb fragment (Figure 8, lane 3) while the double digest produced two fragments of 3.4 Kb and 1.1 Kb (lane 4). *Eco*RI digestion of recombinant plasmids containing the 2.1 Kb *darA* insert resulted in a linear fragment of 6.6 Kb (lanes 7, 9, 11) while three fragments- 3.4 Kb, 2.1 Kb, and 1.1 Kb- were produced by the double digest (lanes 8, 10, 12). Of the six transformants screened, 3 contained the desired construct.

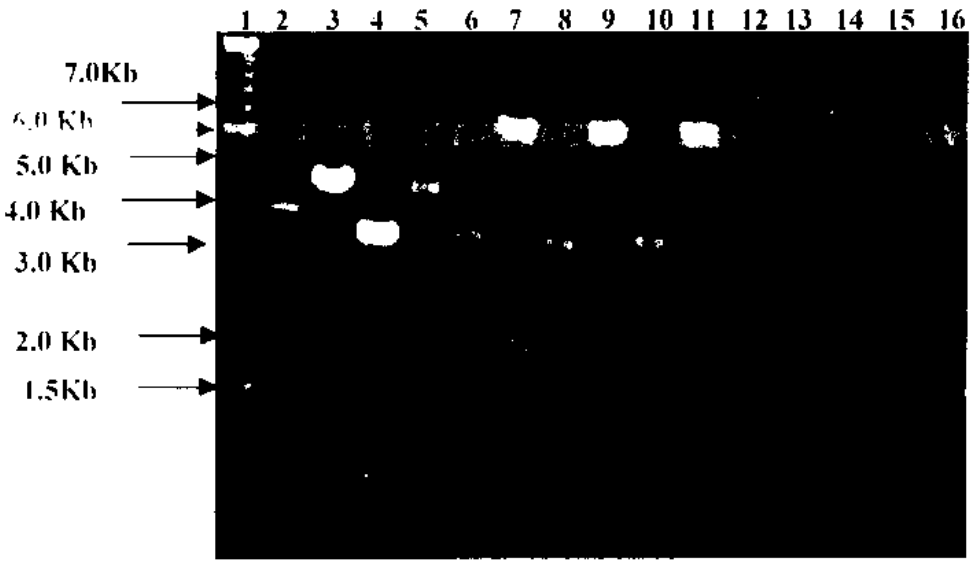

Figure 8. Agarose gel electrophoresis of restriction digests DNAs from possible pRR99 clones. Plasmid DNAs were digested with the enzymes indicated and the digests were electrophoresed on a 0.8% agarose gel as described in Materials and Methods. Plasmid pBAD22 served as the negative control. Lanes 2 - 4 contain pBAD22 DNA digested as follows: lane 2, no enzyme; lane 3, *Eco*RI; lane 4, *Eco*RI + *Pst*I. Six potential positive clones were analyzed (lanes 5-16) with the even lanes containing the *Eco*RI digests and the odd lanes containing the *Eco*RI –*Pst*I double digests of the same miniprep samples. Lanes 7-12 are indicative of the positive pRR99 transformants with one fragment of 6.6 Kb for the *EcoR*I digest and three fragments of 3.4 Kb, 2.1 Kb, and 1.1 Kb for the *EcoR*I + *Pst*I double digest. (The 1.1 Kb fragment is very faint but visible on the original photo.) Lanes 5-6 and 13-16 are the result of transformants without the insert. Lane 1 contains the 1 Kb ladder molecular weight standards. Marker fragment sizes are indicated to the left of the gel.

Introduction of an NcoI Restriction Site by PCR Amplification of the 5' End of the DarA

We were interested in constructing a *darA* expression plasmid in which the levels of *darA* could be tightly regulated. The cloning vector of choice was pBAD22 because it possesses a tightly regulated promoter that is induced to high levels upon addition of arabinose to the growth medium. The p_{BAD} promoter is coupled to a highly efficient Shine-Delgarno (SD) site followed by an *Nco*I site into which the gene of interest can be cloned. To ensure a correctly spaced fusion between *darA* and the SD translation signal on pBAD22, we needed to engineer an *Nco*I site to overlap the *darA* start codon. Toward this end, the 5' end of *darA* was amplified using PCR with a forward primer that altered the region around the start codon appropriately (Figure 4). The reverse primer was complementary to a region downstream of the unique *Kpn*I site. Plasmid pRR99 served as the template in the PCR. The resulting fragment included the *darA* sequence from the initiation site at base pair 2424 to nucleotide 2963 of the published sequence. The complete PCR fragment was predicted to be 561 basepairs long and was termed "Start". An agarose gel was run to confirm the size of the PCR product (Figure 9).

During P1 particle maturation, the *darA* gene product is processed by a phage-encoded protease to produce a short amino terminal fragment and a large carboxyl terminal fragment (31). It is not known whether this processing event is required for antirestriction activity and thus it was of interest to create a truncated *darA* gene, which encoded only the carboxyl terminal fragment. A putative phage processing site was identified in DarA due to its sequence similarity to the bacteriophage lambda protease recognition site (20). Cleavage at this site would theoretically produce a polypeptide similar in size to the mature 68 kDa DarA present in P1 phage particles. The forward

protease PCR primer was designed to contain an ATG initiation codon immediately adjacent to the sequence encoding the beginning of the putative carboxyl terminal *darA* fragment (Figure 4). An *Nco*I site overlapped this ATG to allow for cloning of the resulting PCR product. PCR amplification of pRR99 using the forward protease primer paired with the reverse primer produced the "Protease" DNA fragment of the predicted 360 bp size (Figure 9, lane 4).

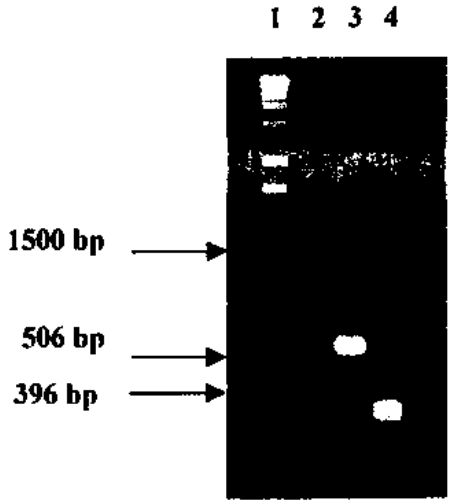

Figure 9. Agarose gel analysis of PCR products. Lane 1 contains the 1 Kb ladder molecular weight standard; relevant fragment sizes are to the left of the gel. Lane 2 corresponds to the negative control PCR that lacked primers. Lane 3 corresponds to the Start fragment of approximately 560 base pairs, and lane 4 contains the 360 bp Protease fragment.

Reconstruction of the darA gene under the control of the pBAD promoter

The Start and Protease PCR fragments were cloned separately into pBAD22 as described in Materials and Methods. Briefly, the PCR products included unique *Nco*I-*Sma*I sites- 5' and 3', respectively- flanking the *darA* sequence (Figure 4). The PCR products were digested with the corresponding enzymes and joined to pBAD22 vector DNA that had been cleaved with the same enzymes. Upon ligation, the intermediate

plasmid with the Start fragment in pBAD22 was referred to as pBAD22-Start. The similarly-named pBAD22-Protease contained the Protease PCR fragment. These new recombinant plasmids had a total size of 5.0 Kb and 4.8 Kb, respectively. To identify positive clones, we performed restriction enzyme digests of miniprep DNA isolated from transformants. Separate digests were carried out using *Hpa*I and *Pvu*I. Plasmid pBAD22 served as the negative control for both digests. Figures 10 and 11 show the agarose gels of the restriction digestion screening for pBAD22-Start and pBAD22-Protease, respectively.

The pBAD22 vector lacks a *Hpa*I site and thus is not cleaved by the enzyme (See lane 2 in Figures 10 and 11). The inserted *darA* fragment contains a single *Hpa*I site and thus positive clones are linearized by *Hpa*I (See lanes 3-6, Figure 10). In the case of *Pvu*I, pBAD22 contains a single recognition site (lane 7, Figure 10) while insertion of the *darA* fragment introduces a second site (for examples see lanes 8-11, Figure 10). All four putative pBAD-Start DNAs analyzed were the correct construct (Figure 10) while at least eight of the twelve transformants analyzed contained the pBAD-Protease plasmid (Figure 11).

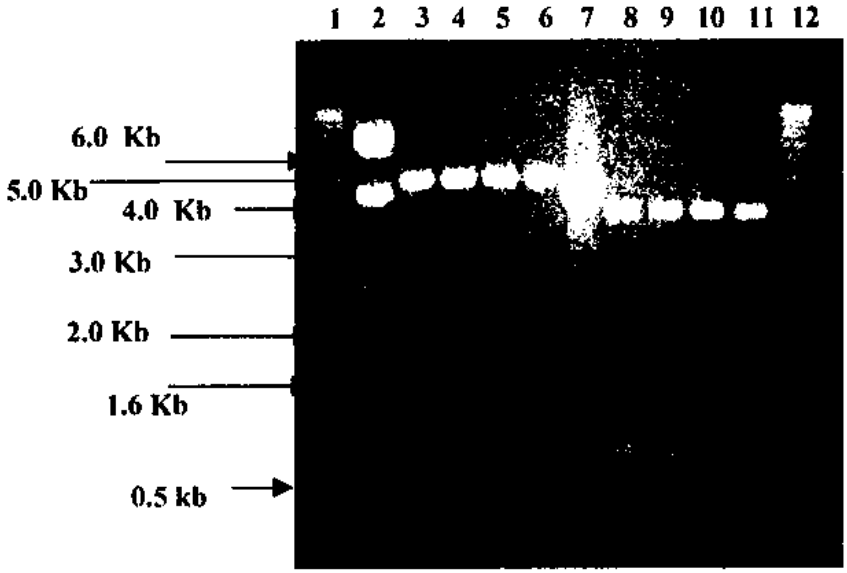

Figure 10. Agarose gel electrophoresis of restriction digests of possible pBAD22-Start clones. Lanes 1 and 12 are the 1 Kb molecular weight standards. Fragment sizes are indicated to the left of the gel. Lanes 2-6 correspond to the *Hpa*I digests, and lanes 7-11 correspond to the *Pvu*I digests. Lane 2 contains the control, pBAD22, which is not cut by *Hpa*I. Lanes 3-6 are indicative of positive transformants where the *Hpa*I digestion produces a linear fragment of 5.0 Kb. Lane 7 corresponds to the pBAD *Pvu*I digest which produces a linear fragment of 4.5 Kb. Lanes 8-11 are indicative of positive transformants where the *Pvu*I digest results in the formation of two fragments: 3.9 Kb and 1.1 Kb.

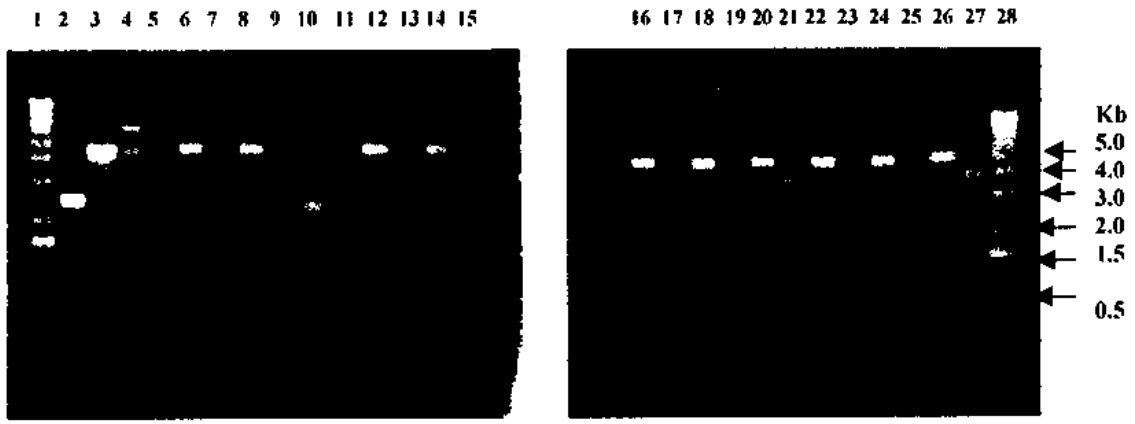

Figure 11. Agarose gel electrophoresis of restriction digests of possible pBAD22-Protease clones. Lanes 1 and 28 contain the 1 Kb molecular weight standards. Sizes are indicated to the right of the gel. The even lanes (2-26) comprise the *Hpa*I digests. The odd numbered lanes (3-27) contain the *Pvu*I digests. pBAD22 is not digested by *Hpa*I (lane 2) while *Pvu*I digestion produces a linear fragment of 4.5 Kb (lane 3). Lanes 12-27 are indicative of positive transformants where the even numbered lanes show a linear

fragment of 4.8 Kb. The odd numbered lanes indicative of the *Pvu*I digest produce two fragments of 3.7 Kb and 1.1 Kb. DNA samples in lanes 4-5 and 10-11 do not contain inserts.

To construct the final recombinant plasmids, pRR99 and the intermediate plasmids, pBAD22-start and pBAD22-protease were digested with *Kpn*I and *Hin*DIII (Figure 6 and 7 respectively). This double digest formed two fragments for pRR99 of 4.8 Kb and 1.7 Kb. For each of the intermediate plasmids two fragments resulted as well (pBAD22-Start: 0.29 Kb and 4.8 Kb; pBAD22-Protease: 0.29 and 4.6 Kb). The smaller fragment of pRR99 containing the 3' end of *darA* was ligated separately to the larger fragments of the intermediate plasmids. The resulting recombinant plasmids were pRR2000 (6.5 Kb), and pRR2001 (6.3 Kb). Plasmid pRR2000 contained the entire *darA* gene with a new *Nco*I site. Plasmid pRR2001 contained the truncated portion of the gene with a new *Nco*I site. A restriction digest using *Nco*I was carried out to identify positive clones. Figures 12 and 13 depict the agarose gels for pRR2000 and pRR2001, respectively.

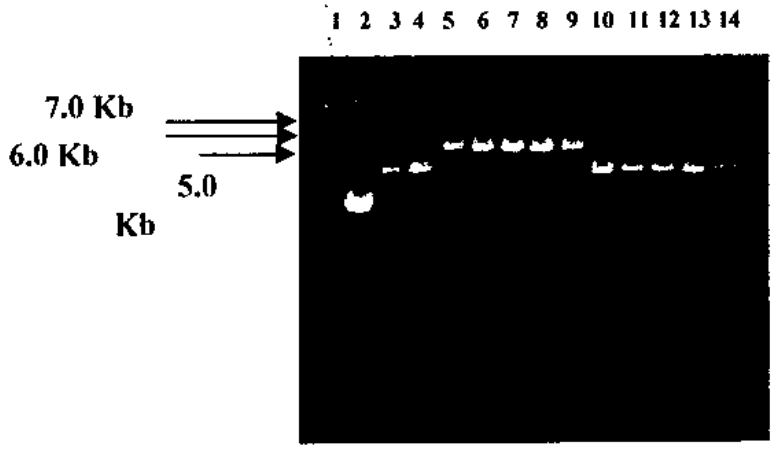

Figure 12. Agarose gel electrophoresis of *Nco*I restriction digests of possible pRR2000 clones. Lane 1 corresponds to the molecular weight standard. The fragment sizes are indicated to the left of the gel. Lane 2 has the negative control, pRR99, which had been digested with *Nco*I. Since it lacks this restriction site, pRR99 was not cut with

the enzyme. Lane 3 has another negative control, pBAD22-Start. When digested with *Nco*I, it will produced a linear fragment of 5.0 Kb. Lanes 4-14 are transformants incubated with the *Nco*I enzyme. Lanes 5-9 have the recombinant plasmid and *Nco*I generates a linear fragment of 6.5 Kb. Lanes 4 and 10-14 do not have the recombinant plasmid and the restriction enzyme produces a linear fragment of 5.0 Kb.

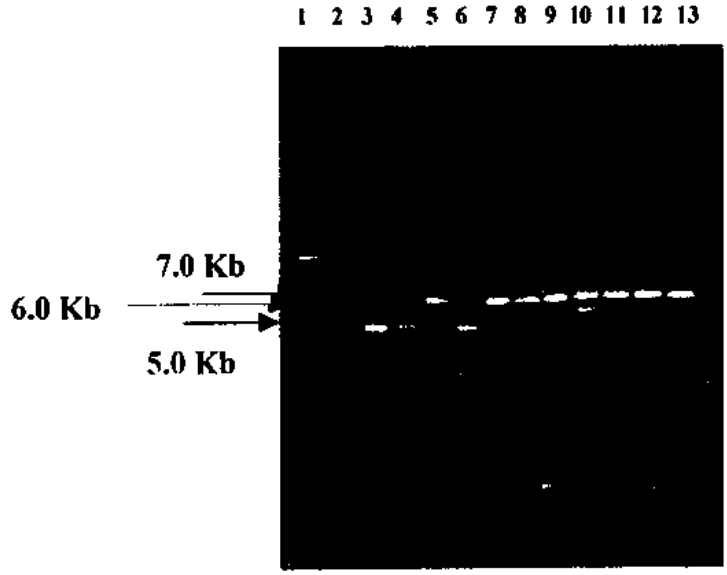

Figure 13. Agarose gel electrophoresis of *Nco*I restriction digests of possible pRR2001 clones. Lane 1 corresponds to the molecular weight standard. The fragment sizes are indicated to the left of the gel. Lane 2 does not contain anything. Lane 3 has a negative control, pBAD22-protease. When digested with *Nco*I, it will produce a linear fragment of approximately 5.0 Kb. Lanes 4-13 are transformants digested with the *Nco*I enzyme. Lanes 5, and 7-13 have the recombinant plasmid and *Nco*I generates a linear fragment of 6.3 Kb. Lanes 4 and 6 do not have the recombinant plasmid and the restriction enzyme produces a linear fragment of approximately 5.0 Kb

Sequence Analysis

To confirm that pRR2000 and pRR2001 contained the desired *darA* constructs, plasmid DNAs from several, independent isolates were sent for DNA sequencing at the Arizona State University DNA Laboratory. The resulting sequences were compared to that of the published *darA* sequence (23) using the BLAST 2 software available on the NCBI web site (http://www.NCBI.gov). Except for the intended changes (I.E. the *Nco*I sites) the newly-generated sequences were 100% identical to the published sequence.

Protein Expression

The recombinant plasmids, pRR2000 and pRR2001, were designed as expression vectors in which *darA* would be under the control of the inducible p_{BAD} promoter. Addition of arabinose was hypothesized to result in high levels of *darA* gene expression and thus protein production.

To test the functionality of these vectors, pRR2000, pRR2001, and pBAD22 were separately introduced into *E. coli* LMG194, and these strains were analyzed for their ability to produce the DarA polypeptide. Overnight cultures of these plasmid-bearing cells were subcultured and grown to mid-log phase. The mid-log cultures were split in two and arabinose was added to one half culture to a final concentration of 0.2%. The cultures were incubated at 37°C. Sample aliquots were removed at varying times and prepared for protein gel analysis as described in Materials and Methods.

To analyze cellular protein expression, the samples were run on denaturing 8-25% gradient polyacrylamide gels. Figures 14 and 15 depict *darA* gene expression in LMG194/pRR2000 and LMG194/p2001, respectively. For these experiments, equivalent amounts of cell culture were loaded in each lane. It is evident from the time point samples that the total amount of protein increased with incubation time (For example, see Figure 14, lanes 3-6). LMG194/pBAD22 lacks the *darA* gene and thus cannot produce the corresponding protein (lane 2 in Figures 14 and 15; Figure 16, lane 7). A comparison of the LMG194/pBAD22 sample with LMG194/2000 induced samples revealed a polypeptide band running just below the 97 kDa marker that is absent in the pBAD22 lane (Figure 14, lane 2 versus lanes 3 - 6). Synthesis of this peptide is dependent upon the addition of arabinose as shown by the fact that this band is

undetectable in the uninduced LMG194/pRR2000 cultures (lanes 7, 8). It is reasonable to hypothesize that this band is indeed the DarA polypeptide based on the fact that it is only present in cells containing the pRR2000 *darA* expression plasmid which had been incubated in the presence of arabinose. The band migration is also consistent with the 77 kDa predicted molecular weight of the DarA protein (31).

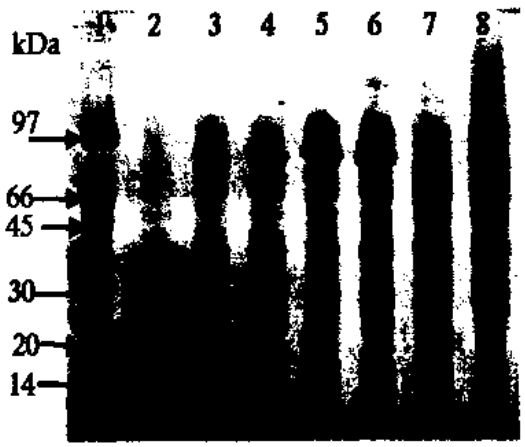

Figure 14. Detection of DarA expression by SDS-PAGE analysis of cell lysates. Lane 1: molecular weight markers with sizes listed to the left of the gel. Lane 2: LMG194/pBAD22 induced with arabinose for four hours. Lanes 3-6: LMG194/pRR2000 induced with arabinose for one, two, four hours, and overnight, respectively. Lane 7 and 8: LMG194/pRR2000 grown in the absence of arabinose for four hours and overnight, respectively.

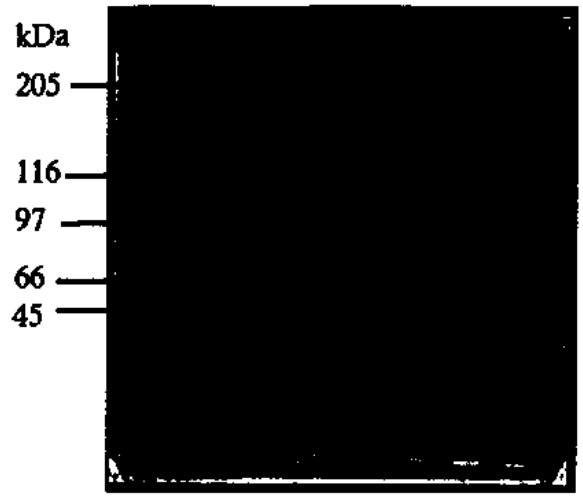

Figure 15. Detection of truncated DarA expression by SDS PAGE analysis of cell lysates. Lane 1: molecular weight markers with sizes listed to the left of the gel. Lane 2: LMG194/pBAD22 induced with arabinose for four hours. Lanes 3-6: LMG194/pRR2001 induced with arabinose for one, two, four hours, and overnight, respectively. Lane 7 and 8: LMG/pRR2001 grown in the absence of arabinose for four hours and overnight, respectively.

From the gel in Figure 15 it is apparent that a darkly-staining polypeptide band of approximately 66 kDa was present in the induced LMG194/pRR2001 samples but was not apparent in either the LMG194/pBAD22 or the uninduced LMG194/pRR2001 lysates. The molecular weight of this over-expressed polypeptide was quite similar to that of the processed, mature DarA. From these results we propose that the synthetic *darA* construct was able to produce a stable, truncated DarA polypeptide. In both over-expression systems, the total amount of DarA protein increased with increasing incubation time.

Another SDS PAGE gel was run to directly compare the sizes of the proteins produced via the new recombinant plasmids. From Figure 16 it is evident that pRR2000 produced a protein of a size between 66 and 97 kDa which is larger than the protein that

pRR2001 produces at approximately 66 kDa. This result is in agreement with the protein sizes predicted by the cloned *darA* constructs.

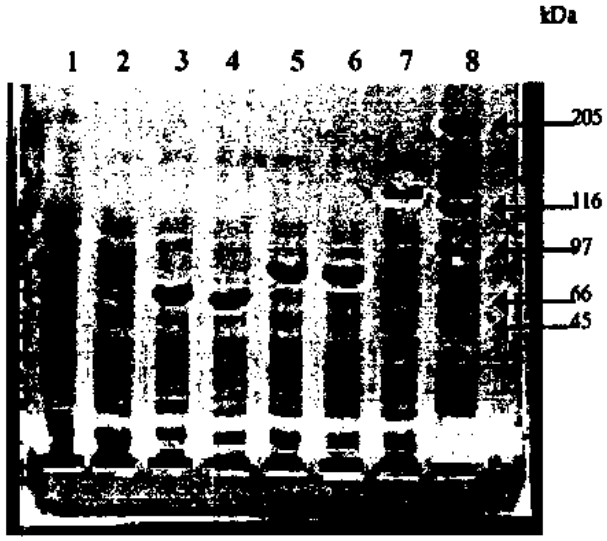

Figure 16. SDS-PAGE analysis of full-length and truncated DarA polypeptides. Lanes 1 and 2: LMG194/pRR2000 and LMG194/pRR2001, respectively, uninduced overnight. Lanes 3 and 4: LMG194/pRR2001, 4 hour arabinose induction. Lanes 5 and 6: LMG194/pRR2000, 4 hour arabinose induction. Lane 7: LMG194/pBAD22, 4 hour arabinose induction. Listed to the right of the gel are the molecular weights of the markers in lane 8.

Restriction Assays

The abilities of the recombinant DarA proteins to inhibit *Eco* AI restriction activity *in vivo* were analyzed via phage lambda restriction assays. Restriction activity is assessed by comparing the number of lambda phage plaques formed on the test strain lawn to the number of plaques formed on the non-restricting control:

$$\frac{\text{Tester strain PFU/ml}}{\text{Nonrestricting strain PFU/ml}} = \text{Efficiency of Plating (EOP)}$$

An EOP of 1.0 indicates the lack of restriction activity while cell expressing R-M systems would have lower EOP values. EOP values also allow us to detect antirestriction activity. The two recombinant plasmids, pRR2000 and pRR2001, were tested to see if

they encode protection against the *Eco*AI restriction system (type IB family) constituitively expressed in WA2379 cells. The negative control used in these experiments was pBAD22. If DarA could inhibit *Eco*AI, WA2379 cells expressing a functional *darA* gene would produce a greater number of PFU/ml (and likewise a higher EOP) than cells lacking a DarA-producing plasmid. Table 5 summarizes the results.

Table 5. Summary of the Restriction Assays

Strain	+/- ara	PFU/ml	+ara/-ara Ratio	EOP
WA921 (R^+M^+)	+	8×10^{10}	1.1	1.0
WA921 (R^+M^+)	-	7×10^{10}		0.88
WA2379 / pBAD22	+	3.35×10^{8}	1.1	0.42
WA2379 / pBAD22	-	3.05×10^{8}		0.38
WA2379 / pRR2000	+	6.45×10^{8}	1.8	0.81
WA2379 / pRR2000	-	3.55×10^{8}		0.44
WA2379 / pRR2001	+	5.7×10^{8}	1.4	0.71
WA2379 / pRR2001	-	4.1×10^{8}		0.51

+/- ara indicates whether cells were grown in the presence or absence of arabinose.

WA921, the nonrestricting control strain, had, by definition, an EOP of 1 (Table 5). WA2379/pBAD22 encoded a functional *Eco*AI R-M system as detected by an EOP value of less than half the control. This level of *Eco*AI restriction activity is typical for lambda which contains only a single *Eco*AI recognition site.

When the full-length *darA* gene was expressed in WA2379/pRR2000 cells, restriction activity was inhibited as shown by an increase in the EOP value (Table 5). This value was dependent on arabinose induction (+ara/-ara = 1.8). The truncated *darA* protein, expressed from pRR2001 in WA2379 also inhibited *Eco*AI restriction activity,

albeit, at a lower level. These results demonstrate that both the full-length and the truncated DarA proteins expressed from the newly constructed plasmids act to inhibit the host cell type IB restriction activity and, in so doing, make the cells more susceptible to phage infection.

Discussion & Conclusion

Our main goal was to determine whether the full length DarA or the shorter, processed form of the protein were capable of inhibiting type I R-M systems "*in trans*". We successfully constructed two DarA-expressing plasmids, pRR2000 and pRR2001. Plasmid pRR2000 is 6.5 kb recombinant plasmid which possesses ampicillin resistance, and consists of the pBAD promoter, the full-length *darA* gene, and an *Nco*I site. pRR2001 is 6.3 kb and possesses the same characteristics with the exception that only the *darA* sequence just after the hypothetical protease cleavage site is present on the plasmid. These plasmids were shown to be correct by restriction site analysis and DNA sequencing.

Once the plasmid constructs were verified, we wanted to determine whether or not the full-length and truncated forms of DarA protein could be overexpressed using our newly constructed plasmid systems. The use of the pBAD22 promoter allowed us to tightly regulate gene expression. The unprocessed DarA protein is approximately 77 kDa (31). Plasmid pRR2000 produced the full-length protein at a size between 97 and 66 kDa. Since we did not determine the molecular weight of the DarA bands, we are estimating that the full length may be 77 kDa. We are also estimating that the pRR2001 produced a smaller, approximately 66 kDa protein that is of similar size to the mature, processed DarA polypeptide found in the phage head (20 and 31). Both the full length and truncated protein appear to be fairly stable as suggested by their intense polypeptide bands present in lysates from overnight induced cultures as visualized by polyacrylamide gel analysis (Figures 14 and 15). The lack of obvious degradation products in the protein extracts from cells induced for prolonged periods is also evidence of protein stability.

We know that the plasmid has the *darA* gene and makes the DarA protein, but is it functional in protecting against type IB R-M systems? According to our results in Table 5, the EOP values for the induced WA2379/pRR2000 and WA2379/pRR2001 were both higher than for WA2379/pBAD22. This showed us that the incoming phages were less restricted and therefore both plasmids encoded a functional antirestriction activity. It was also apparent that the induced strains produced more phage compared to the uninduced strains. Thus the antirestriction activity was dependent on the *darA* sequence itself.

Previous work suggested that the DarA polypeptide only protected DNA with which it had been copackaged in a bacteriophage P1 capsid (22 and 31). We have shown this not to be the case. DarA can protect the heterologous phage lambda and the antirestriction protein does not need to be simultaneously introduced into the cell along with the DNA. The discrepancy between the two experiments is likely due to the means of expression and the amounts of DarA polypeptide present. In our system, the DarA polypeptide was under the control of the heterologous *pBAD* promoter and the gene could be induced to high levels of expression as detected by SDS-PAGE analysis. The protein levels were sufficient to inhibit the *Eco*AI restriction endonuclease which was coexpressed in the same cell. Iida and coworkers (22) analyzed the ability of $darA^+$ phage particles to protect the DNA of a coinfecting dar^- phage particles. The amount of DarA protein packaged in the phage head apparently is only sufficient to protect the copackaged DNA and does not extend protection to any coinfecting phage.

One interesting reflection is that expression of DarA within a normally restricting bacterial cell actually makes that cell more sensitive to phage infection and lysis. Thus

such antirestriction activity could be exploited in the future to facilitate new antibacterial therapies.

Future work will involve repeating the restriction assay several times to replicate the results and to extend our analysis to determine whether or not DarA is active against the untested Types IC and ID systems. Other unanswered questions that can be the basis of future experiments include: How does DarA work in conjunction with DarB to protect from Type IA systems? Are both DarA and DarB required to protect against Type IC and ID systems?

The biochemical mechanism of DarA antirestriction has yet to be determined. One possibility is that DarA may work like the Ocr (overcome classical restriction) protein of bacteriophage T7. Although unrelated to phage P1, T7 also falls prey to type I DNA restriction enzymes when it infects *E.coli* (6). Ocr has the size and shape to effectively block the DNA binding site of a type I enzyme (32). This highly acidic protein functions as a competitive inhibitor by preventing a restriction endonuclease from binding its DNA target sequence (3). Thus the Ocr acts as a DNA mimic and the R-M enzyme binds to it. The resulting Ocr-Type I R-M enzyme complex can neither restrict nor modify DNA; therefore, the phage is able to proliferate (34).

Although they differ greatly in size, there are a few similarities between these two antirestriction proteins. Both Ocr and DarA are fairly acidic. Also, DarA has a sequence motif that Ocr uses as a DNA mimic. Ocr has ED(X)3E(X)2E (reference 34), while DarA has the similar amino acid fragment, EE(X)3E(X)2E (personal observation). Site-directed mutagenesis can be used to determine if this acidic motif is required for DarA

antirestriction function. Biochemical and biophysical analyses can be used to further address the DNA mimic hypothesis.

With its antirestriction activity, bacteriophage P1 can protect itself from the R-M systems of bacteria. This bacteria-phage system is an ideal model for host-pathogen interactions and exquisitely illustrates the coevolution of such partners. Bacterial cells which normally fall prey to phage develop R-M systems to screen incoming DNA. Antirestriction activities then evolved as a means to thwart the protective action of the R-M systems. Thus the interplay between phage and bacterium illustrates how organisms will evolve ways to defend themselves from attack, thereby promoting their survival.

From a mutualistic sense, Bacteriophage P1 is an efficient transducing phage, and thus facilitates the horizontal transfer of bacterial genes. This means of transfer may explain the wide array of characteristics obtained horizontally by bacteria, and the rapid spread of traits amongst bacteria. This constant development can prove to be deleterious for humans. One example is the emergence of *V. cholerae* to a toxigenic strain as discussed in the Introduction section. An understanding of how antirestriction works can lead to the development of methods to prevent non-pathogenic strains from being converted to pathogenic ones, or even to prevent viral proliferation in the food industry. Finally, as mentioned, an understanding of antirestriction activity may lead to the development of novel methodologies to target life-threatening bacterial infections.

Literature Cited

1. **Alisky, J., K. Iczkowski, A. Rapoport, and N. Troitsky**. 1998. Bacteriophages show promise as antimicrobial agents. Journal of Infection **36**: 5-15.
2. **Arber, W. and D. Wauters-Willems**. 1970. Host-Specificity of DNA Produced by *Escherichia coli:XII*. The two restriction and modification systems of strain 15T⁻. Molecular and General Genetics **108**: 203-217.
3. **Atanasiu, C., O. Byron, H. McMiken, S. S. Sturrock, and D. T. F. Dryden.** 2001. Characterization of the structure of Ocr, the gene 0.3 protein of bacteriophage T7. Nucleic Acids Research **29**:3059-3068.
4. **Bickle, T.** 1987. DNA restriction and modification systems, p. 692-696. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaecter, and H. E. Umbarger (ed.), *Esherichia coli* and *Salmonella*: Cellular and Molecular Biology, 2nd ed. American Society for Microbiology, Washington, D.C.
5. **Bickle, T. A. and D. H. Kruger**. 1993. Biology of DNA restriction. Microbiological Review **57**: 434-450.
6. **Blackstock, J. J., S. U. Egelhaaf, C. Atanasiu, D. T. F. Dryden, and W. C. K. Poon**. 2001. Shape of Ocr, the gene 0.3 protein of bacteriophage T7: Modeling based on light scattering experiments. Biochemistry **40**: 9944-9949.
7. **Boyd, E. F., K. E. Moyer, L. She, and M. K. Waldor**. 2000. Infectious CTXΦ and the Vibrio pathogenicity island prophage in *Vibrio mimicus*: Evidence for recent horizontal transfer between *V. mimicus* and *V. cholera*. Infection and Immunity **68**:1507-1513.

8. **Brussow, H., 2001.** Phages of dairy bacteria. Annual Review of Microbiology **55**: 283-303.

9. **Dalsgaard, A., O. Serichantalergs, A. Forslund, W. Lin, J. Mekalanos, E. Mintz, T. Shimada, and J. G. Wells.** 2001. Clinical and environmental isolates of *Vibrio cholerae* serogroup O141 carry the CTX phage and the genes encoding the toxin-coregulated pili. Journal of Clinical Microbiology **39**:4086-4092.

10. **Davis, B. M. , H. H. Kimsey, A. V. Kane, and M. K. Waldor.** 2002. A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. EMBO Journal **21**: 4240-4249.

11. **Davis, B. M., H. H. Kimsey, W. Chang, and M. K. Waldor.** 1999. The *Vibrio cholerae* O139 Calcutta bacteriophage CTXΦ is infectious and encodes a novel repressor. Journal of Bacteriology **181**: 6779-6787.

12. **Davis, Brigid, M. K. E. Moyer, E. F. Boyd, and M. K. Waldor.** 2000. CTX prophages in classical biotype *Vibrio cholerae*: Functional phage genes but dysfunctional phage genomes. Journal of Bacteriology **182**: 6992-6998.

13. **Dryden, DT, N. E. Murray, and D. N. Rao.** 2001. Nucleoside triphosphate-dependent restriction enzymes. Nuceic Acids Research **29**: 3728-3741.

14. **Faruque, S. M., M. Asadulghani, M. Rahman, M. K. Waldor, and D. A. Sack.** 2000. Sunlight-induced propagation of the lysogenic phage encoding the cholera toxin. Infection and Immunity **68**: 4795-4801.

15. **Gann, A. A., A. J. Campbell, J. F. Collins, A. F. Coulson, and N. E. Murray NE.** 1987. Reassortment of DNA recognition domains and the evolution of new specificities. Molecular Microbiology 1:13-22.

16. **Gubler, M., D. Braguglia, J. Meyer, A. Piekarowicz , and T. A. Bickle.** 1992. Recombination of constant and variable modules alters DNA sequence recognition by type IC restriction-modification enzymes. EMBO Journal **11**: 233-240.

17. **Gubler, M. and T. A. Bickle.** 1991. Increased protein flexibility leads to promiscuous protein—DNA interactions in type IC restriction-modification systems. EMBO Journal **10**: 951-957.

18. **Guzman, Luz-Maria, D. Belin, M. J. Carson, and Jon Beckwith.** 1995. Tight regulation, modulation, and high level expression by vectors containing the arabinose P_{BAD} promoter. Journal of Bacteriology **177**: 4121-4130.

19. **Harkki, A. and E. T. Palva.** 1984. Application of phage lambda technology to *Salmonella typhimurium*. Construction of a lambda-sensitive *Salmonella* strain. Molecular and General Genetics **195**: 256-259.

20. **Hendrix, R. W. , J. W. Roberts, F. W., Stahl, and R. A. Weisberd** (eds). 1983. Lambda II. Cold Spring Harbor Press, Cold Spring Harbor, NY

21. **Iida, S., J. Meyer, P. Linder, N. Goto, R. Nataya, H. J. Reif, and W. Arber.** 1982. The Kanamycin resistance transposon Tn 2680 derived from the R plasmid Rts1 and carried by phage P1Km has flanking 0.8-kb-long direct repeats. Plasmid **8**: 187-198.

22. **Iida, S., M. B. Streiff, T. A. Bickle, and W. Arber.** 1987. Two DNA Antirestriction systems of bacteriophage P1, *darA* and *darB*: Characterization of *darA*$^-$ Phages. Virology **157**:156-166.

23. **Iida, S., R. Hiestand –Nauer, H. Sandmeier, H. Lehnherr, and W. Arber.** 1998. Accessory genes in the *darA* operon of bacteriophage P1 affect

antirestriction function, generalized transduction, head morphogenesis, and Host Cell Lysis. Virology **251**: 49-58.

24. **Jacobs, W. R. Jr, R. G. Barletta, R. Udani, J. Chan, G. Kalkut, G. Sosne, T. Kieser, G. J. Arkis, G. F. Hatfull, and B. R. Bloom.** 1993. Rapid assessment of drug susceptibilities of *Mycobacterium tuberculosis* by means of luciferase reporter phages. Science **260**: 819-222.
25. **Karaolis, D. K. R., S. Somara, D. R. Maneval Jr., J. A. Johnson, and J. B. Kaper.** 1999. A bacteriophage encoding a pathogenicity island, and Type-IV pilus and a phage receptor in Cholera bacteria. Nature **399**: 375-379.
26. **Kruger, D. and T. A. Bickle.** 1983. Bacteriophage survival: Multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. Microbiological Review **47**: 345-360.
27. **Redashci, N. and T. A. Bickle.** 1996. DNA restriction and modification systems, p. 773-781. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaecter, and H. E. Umbarger (ed.), *Esherichia coli* and *Salmonella*: Cellular and Molecular Biology, 2nd ed. American Society for Microbiology, Washington, D.C.
28. **Riska, P. F. and W. R. Jacobs Jr.** 1998. The use of luciferase-reporter phage for antibiotic-susceptibility testing of mycobacteria. Methods in Molecular Biology **101**: 431-455.
29. **Sambrook, J. and D. W. Russel.** 1989. A Laboratory Manual, 2nd Ed. Cold Spring Harbor Press, Cold Spring Harbor, NY.

30. **Schuch, R., D. Nelson D, and V. A. Fischetti.** 2002. A bacteriolytic agent that detects and kills *Bacillus anthracis*. Nature **22**:884-889.

31. **Streiff, M. B., Shigeru, I. and T. Bickle.** 1987. Expression and proteolytic processing of the *darA* antirestriction gene product of bacteriophage P1. Virology **157**:167-171.

32. **Sturrock, S. S., D. T. F. Dryden, C. Atanasiu, J. Dornan, S. Bruce, A. Cronshaw, P. Taylor, and M. D. Walkinshaw.** 2001. Crystallization and preliminary S-ray analysis of Ocr, the product of gene 0.3 of bacteriophage T7. Acta Crystallographica **D57**:1652-1654.

33. **Wagner, P. L. and M. K. Waldor.** 2002. Bacteriophage control of bacterial virulence. Infection and Immunity **70**: 3985-3993.

34. **Walkinshaw, M. D., P. Taylor, S. S. Sturrock, C. Atanasiu, T. Berge, R. M. Herderson, J. M. Edwardson, and D. T. F. Dryden.** 2002. Structure of Ocr from bacteriophage T7, a protein that mimics B-form DNA. Molecular Cell **9**: 187-194.

35. **Wood, W. B.**, 1966. Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA. Journal of Molecular Biology **16**: 118-133.

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

37. **Yarmolinsky, M. B. and N. Sternberg.** 1988. Bacteriophage P1, p. 291-332. *In* H. Fraenkel-Conrat and R. R. Wagner (ed), The Bacteriophages, Vol. 1. Plenum Press, New York, New York.