3700–3706 *Nucleic Acids Research, 1998, Vol. 26, No. 16*

Genetic analysis of prokaryotic and eukaryotic DNA-binding proteins in *Escherichia coli*

Frederick W. Whipple*

Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA

Received May 11, 1998; Revised June 22, 1998; Accepted July 1, 1998

ABSTRACT

This report describes an *Escherichia coli* genetic system that permits bacterial genetic methods to be applied to the study of essentially any prokaryotic or eukaryotic site-specific DNA binding protein. It consists of two parts. The first part is a set of tools that facilitate construction of customized *E.coli* strains bearing single copy *lacZYA* reporters that are repressed by a specific target protein. The second part is a pair of regulatable protein expression vectors that permit *in vivo* production of the target protein at levels appropriate for genetic experiments. When expressed in a properly designed reporter strain, the target protein represses the *lac* genes, resulting in an *E.coli* phenotype that can be quantitatively measured or exploited in large scale genetic screens or selections. As a result, large plasmid-based libraries of protein genes or pools of mutagenized variants of a given gene may be examined in relatively simple genetic experiments. The strain construction technique is also useful for generating *E.coli* strains bearing reporters for other types of genetic systems, including repression-based and activation-based systems in which chimeric proteins are used to examine interactions between foreign protein domains.

INTRODUCTION

Site-specific DNA-binding proteins are central players in the regulation of gene expression and in many other biological processes. In bacteria, these and other proteins are frequently studied genetically via random or localized mutagenesis followed by *in vivo* selection or screening to identify interesting variants. In other organisms, however, this approach is often not practical, particularly in higher eukaryotes. This report describes an *Escherichia coli* genetic system that can facilitate *in vivo* studies of site-specific DNA-binding proteins derived from essentially any source, prokaryotic or eukaryotic. The system provides tools that facilitate the two steps necessary for the development of an indicator system: construction of reporter strains that respond phenotypically to the protein of interest, and expression of the target protein in those strains at appropriate concentrations. Strains constructed using this system bear the *E.coli lacZYA* operon in single copy under the control of an artifical promoter that contains a binding site for the protein under study. One of two IPTG-inducible expression plasmids is used to direct expression of the target protein at an appropriate intracellular concentration, resulting in occupancy of the binding site and repression of the *lac* genes. Relative site occupancy can be quantitated simply by assaying β-galactosidase activity in cultures of growing cells. In addition, the *lac* phenotype that results from DNA binding can be used as the basis of genetic screens or selections. Therefore this system permits powerful bacterial genetic methods to be applied to the study of a wide range of prokaryotic and eukaryotic DNA-binding proteins.

The strain construction strategy takes advantage of several well-established properties of *E.coli*. A customized promoter/operator construct is assembled on a plasmid and then transferred by homologous recombination onto an F′ episome which is subsequently moved via conjugation into a ‘clean’ background strain that lacks the starting plasmid. The recombination and transfer can be performed in one step, making this is a quick and direct method of strain construction. In the finished strain the customized promoter and the entire *lacZYA* operon are located on the F′ episome, which is maintained by *E.coli* permanently and accurately at a copy number of approximately one per cell. This strategy has a number of advantages over many of the existing *lac* fusion techniques. First, because the reporter is not located on a multicopy plasmid or prophage, there is no potential for incompatibility with any application-specific plasmids or phages that may be required. Second, problems such as instability and limited dynamic range that are frequently associated with plasmid-based fusion systems are minimized. Third, the technique produces fusions in which the entire *lacZYA* operon, not just the *lacZ* gene, is placed under control of the artificial promoter/operator construct. This permits full use of the wide range of selective and indicator media available to *lac* geneticists, many of which require coordinate regulation of the *lacZ* and *lacY* genes for optimal performance. Finally, the strain construction method is technically simple to use and is very flexible. It not only accommodates any type of *lac* fusion, but it also permits any desired auxiliary foreign genes or DNA sequences to be incorporated in single copy into the *E.coli* reporter strain.

*Tel: +1 617 432 1922; Fax: +1 617 738 7664; Email: fwhipple@warren.med.harvard.edu

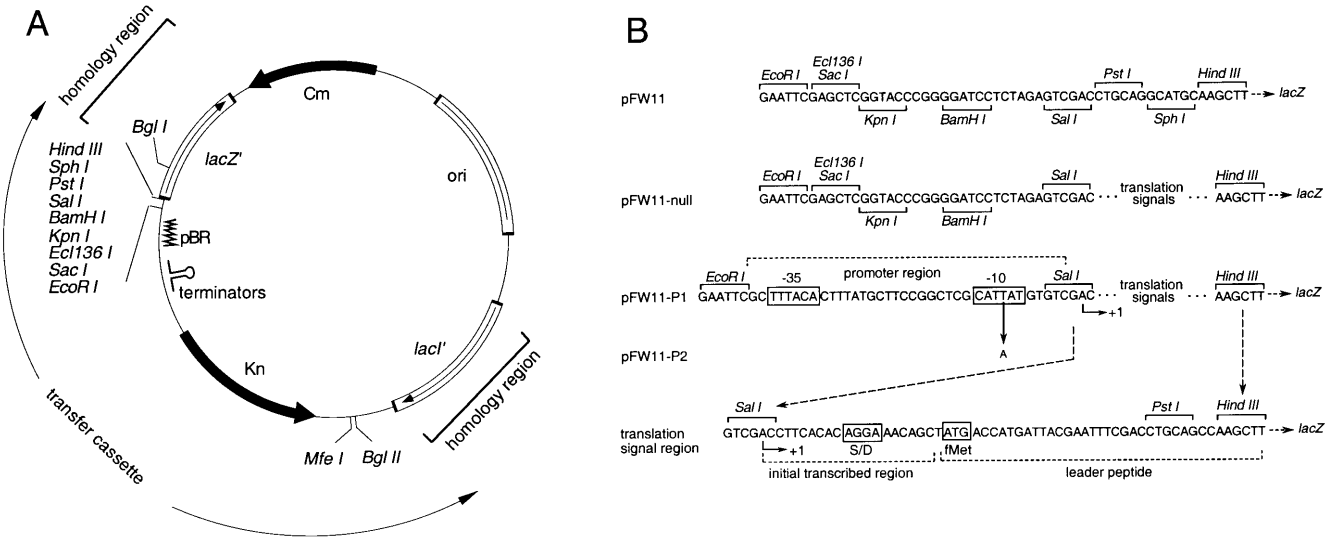

Figure 1. Plasmid pFW11 and its derivatives. (**A**) Map of plasmid pFW11 showing its major features. Selected unique restriction sites are shown. (**B**) Sequences of the *Eco*RI–*Hin*dIII regions of pFW11, pFW11-null and two promoter-containing derivatives.

MATERIALS AND METHODS

Media and standard laboratory procedures

LB liquid media and plates were prepared as described by Miller (1). Antibiotics kanamycin (30 μg/ml), streptomycin (100 μg/ml), chloramphenicol (30 μg/ml) and ampicillin (75 μg/ml) were added as indicated. β-galactosidase assays were performed according to Miller (1) except that strains were grown in LB medium supplemented with appropriate antibiotics and with IPTG as described in the figure legends. MacConkey Lactose medium was purchased in powdered form from Difco Laboratories and prepared as directed. To prepare tetrazolium lactose plates, 50 mg of 2,3,5-triphenyl-2*H*-tetrazolium chloride (Kodak or Merck) and 25.5 g Antibiotic Medium 2 (Difco) were added to 950 ml of water and the mixture was heated to the boiling point prior to autoclaving (the boiling step is essential). After autoclaving and cooling to ~60°C, 50 ml of a filter sterilized 20% solution of lactose (Difco) was added, together with antibiotics, and plates were poured.

Plasmids

Plasmid pFW11, which was constructed in many steps, consists of the following major segments beginning at the *Eco*RI site (Fig. 1). The polylinker and *lacZ*α complementing region from plasmid pUC118 (2) are fused to an additional portion of the *lacZ* gene derived from plasmid pMC903 (3) to yield a *lacZ′* gene that contains codons 8–212 of the natural *lacZ* gene. This region is connected to a 2186 bp *Bst*BI–*Rsa*I segment of plasmid pACYC184 that contains the chloramphenicol acetyltransferase gene (*Cm*) and the origin of replication. Within this segment, the *Eco*RI site in the *Cm* gene and the *Hin*dIII site near the origin of replication have been inactivated by translationally silent directed mutagenesis and by filling in with Klenow fragment, respectively. The next segment, derived from plasmid pAD325 (4), bears DNA encoding the C-terminal 250 amino acids of the *lacI* gene and a segment that corresponds to the 190 bp region between (but not including) the *Eco*RV and *Bam*HI sites of plasmid pACYC184. This is followed by artificial *Bgl*II and *Mfe*I sites, and then by a 1213 bp segment of plasmid pACYC177 which contains a version of the kanamycin resistance gene in which the natural *Hin*dIII site has been removed by translationally silent directed mutagenesis. The final two short segments consist of the *E.coli rrnB* t1 and t2 transcriptional terminators on a 360 bp fragment and the 55 bp region of plasmid pBR322 that immediately precedes its *Eco*RI site. The complete DNA sequence of pFW11 is available upon request.

Plasmid pFW11-null was made by replacing the *Sal*I–*Hin*dIII segment of pFW11 with the translation signal region shown in Figure 1B. Derivatives of pFW11-null were then made by annealing complementary oligonucleotides together to make double stranded DNA fragments bearing *Eco*RI and *Sal*I sticky ends, followed by ligation into the backbone of pFW11-null cleaved with these two enzymes. Annealing was performed at a concentration of 0.4 mg/ml of each oligonucleotide in Sequenase reaction buffer (Amersham). A 1 μl aliquot of a 10- or a 100-fold dilution of the annealing mixture was used in each 10 μl ligation reaction. Ligation mixtures were transformed into either strain MC1061 or strain XL1-Blue and plasmid DNAs purified from cultures grown from single colonies were verified by restriction analysis and DNA sequencing. The sequences of the *Eco*RI–*Sal*I regions of plasmids pFW11-P1 and pFW11-P2 (used to construct strains S11-P1 and S11-P2, respectively) are shown in Figure 1B. The sequences of the 32 bp regions between the *Eco*RI and *Sal*I sites of the plasmids used to construct strains S11-LAM1, S11-LAM2, S11-CRE1 and S11-CRE2 are shown in Figure 4.

Plasmid pLX10 (Fig. 3) directs expression of the λcI gene at low levels under the control of a truncated *lacUV5* promoter/operator region. It was constructed from plasmid pLR1 (5) as follows. First, the *Nde*I site near the origin of replication of the pBR322 vector was removed by filling in with Klenow fragment. Second, the region of the ampicillin resistance gene that contains the *Pst*I and *Hin*cII sites was replaced with the corresponding region of plasmid pUC118, which lacks these sites. Third, the *Pst*I site downstream of the λcI gene was converted to an *Xba*I site by cleavage with *Pst*I, trimming with T4 DNA polymerase, and

insertion of a 12 bp *Xba*I linker (New England Biolabs #1082). The *Xba*I–*Bam*HI region of the resulting plasmid was then replaced with the 35 bp *Xba*I–*Bam*HI region of plasmid Litmus28 (New England Biolabs), which contains five unique restriction sites. Finally, a unique *Nde*I site that overlaps the ATG start codon of the λcI gene was introduced by site directed mutagenesis. Plasmid pLX20, which directs expression of higher levels of λcI, was then constructed by introducing an *Nde*I site at the start codon of plasmid pKB280 (6) (which bears a full *lacUV5* promoter/operator region) and then transferring the 101 bp *Eco*RI–*Nde*I promoter-containing fragment into the corresponding position of pLX10.

Plasmids containing the CREB bZIP and full length CREB genes were kindly provided by J. Hoeffler (InVitrogen Corporation). Plasmid pLX20-C-bZIP was constructed by replacing the *Nde*I–*Bam*HI segment of pLX20 with a PCR-amplified fragment encoding the C-terminal bZIP domain of the human CREB protein (beginning at amino acid leucine 274) fused to a short N-terminal leader peptide (MetAlaArgIle). The protein expressed by pLX20-C-bZIP is identical to the CREB bZIP domain previously expressed in *E.coli* by Santiago-Rivera *et al.* (7). Plasmid pLX20-CRB was then constructed by replacing the region between the *Nde*I site and a unique *Xho*I site in the bZIP segment of pLX20-C-bZIP with a PCR-amplified *Nde*I–*Xho*I fragment encoding the remainder of the human CREB protein. The complete coding regions of both pLX20-C-bZIP and pLX20-CRB were verified by DNA sequencing.

Strains

Strains CSH100 [F′*lacproA+,B+(lacI*^{q} *lacPL8)/araD(gpt-lac)5*] and CSH142 [F-/*araD(gpt-lac)5*] (8) were purchased from the Cold Spring Harbor Laboratory. Strain FW102 was constructed by moving a streptomycin-resistant *rpsL* gene into strain CSH142 via P1 transduction. Strains MC1061 and XL1-Blue were obtained from New England Biolabs and Stratagene, respectively.

Customized indicator strains were constructed as follows. Derivatives of plasmid pFW11-null (see above) were introduced into strain CSH100 by transformation using kanamycin as the selective drug. For each strain construction, starter cultures of the appropriate plasmid-containing CSH100 strain and of strain FW102 were grown overnight in LB medium containing and lacking kanamycin, respectively. The next morning, tubes containing 4 ml of plain LB medium (with no antibiotics added) were inoculated with 0.2 ml of each overnight culture, vortexed briefly, and placed in a 37°C water bath for 2 h without agitation. During the final 20 min the FW102 culture (the recipient) was agitated on a culture wheel. The CSH100 strain, however, was not subjected to agitation in order to avoid breaking F pili which are produced by the cells and which are required for conjugation. Upon completion of the 2 h incubation, 0.5 ml aliquots of the two cultures were gently mixed together and returned to the water bath for 1 h without agitation. The mixture was then agitated on a culture wheel very slowly (15 r.p.m.) for 30 min, and then at normal speed for 1 h. 10- and 100-fold dilutions were made and 0.1 ml aliquots of the undiluted and diluted mixtures were spread on LB plates containing streptomycin and kanamycin. As a control, small amounts of the unmixed donor and recipient subcultures were spotted on an additional streptomycin–kanamycin plate. After overnight incubation, isolated colonies were present on at least one of the mixed culture plates, and the control plate showed no growth. Sterile toothpicks were used to patch ~25 colonies first to a plate containing chloramphenicol, and then to a streptomycin–kanamycin plate. Finished strains (identified as clones that were resistant to kanamycin and streptomycin but sensitive to chloramphenicol) were purified by restreaking and then tested as described in Results.

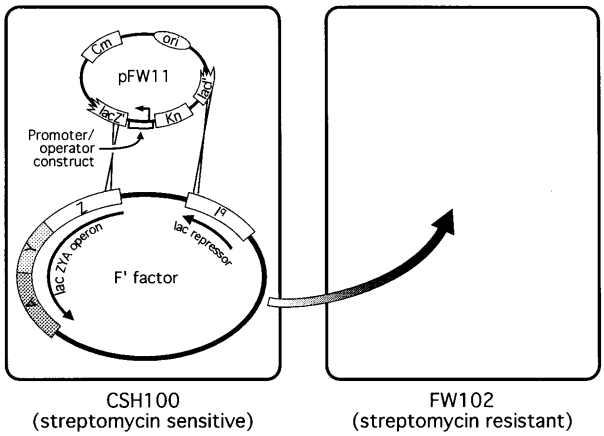

Figure 2. Schematic representation of the strain construction procedure. Within strain CSH100, the transfer cassette of the desired derivative of plasmid pFW11 moves onto the F′ via homologous recombination at the *lacI′* and *lacZ′* regions, respectively (crossed straight lines). The recombinant F′ then moves into strain FW102 via conjugation.

RESULTS

Construction of unregulated reporter strains

Plasmid pFW11 (Fig. 1A) is a derivative of pACYC184 which contains the 3′ end of the *lacI* gene and the 5′ end of the *lacZ* gene. Between these two gene fragments are located a kanamycin resistance gene, a polylinker cloning site, and several other features. When pFW11 is present in a cell that bears the *lacI* and *lacZ* genes (either on the chromosome or on an episome) the region between the two gene fragments can be transferred by homologous recombination. Thus the region of pFW11 between the two *lac* gene fragments may be considered a transfer cassette.

To make an indicator strain, the desired promoter, operator and translation signals are first assembled on a pFW11 derivative using a simple cloning technique. The resulting plasmid is then introduced into strain CSH100, a streptomycin sensitive strain which bears an F′ episome that carries the complete *lac* operon. Transfer of the cassette from the pFW11 derivative to the F′ occurs via two homologous recombination events that take place at the *lacI′* and *lacZ′* homology regions, respectively (Fig. 2). Singly recombinant intermediates (cointegrant molecules that contain the whole of both episomes) will also be present in the culture. The desired doubly recombinant F′ is distinguishable from the non-recombinant and singly recombinant (cointegrant) forms because it bears the kanamycin resistance gene but lacks the chloramphenicol marker.

Strain construction is completed simply by mixing the above strain with a streptomycin resistant recipient strain (FW102), allowing conjugation to occur, and plating the mixture on medium that contains streptomycin and kanamycin. Only FW102 cells that have acquired either a cointegrant or a doubly recombinant F′ will form colonies on this medium. A clone bearing the desired double recombinant is then identified by screening colonies for sensitivity to chloramphenicol.

Table 1. Identification and testing of promoter–*lacZYA* fusion strains

Starting plasmid	Promoter	Candidates tested	Cm^S clones	Percent Cm^S	β-gal activity
pFW11-null	none	44	6	14	9
pFW11-P1	moderate	56	6	11	307
pFW11-P2	strong	40	8	20	2628

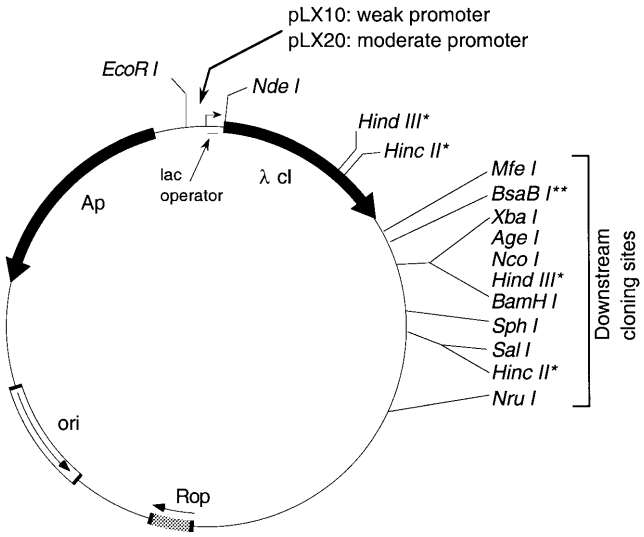

Figure 3. Protein expression plasmids pLX10 and pLX20. The major features of plasmids pLX10 and pLX20 are diagrammed. Unique and semi-unique restriction sites that are useful for cloning purposes are shown. *, the *Hin*dIII and *Hin*cII sites within the λcI gene do not interfere with the use of the downstream *Hin*dIII and *Hin*cII sites for constructing derivatives bearing foreign genes. **, a second *Bsa*BI site is present in the vector, but the latter site is blocked by *dam* methylation and is therefore not cleaved unless the plasmid DNA has been isolated from a *dam*⁻ strain.

To characterize the recombination and mating procedure, test strains were constructed from deriviatives of plasmid pFW11 that bear unregulated promoters. Plasmids pFW11-P1 and pFW11-P2 (Fig. 1B) bear a moderately strong and a strong promoter, respectively, but contain no protein binding sites. Plasmid pFW11-null (Fig. 1B), which has no promoter, was also used as a control. All three plasmids have translation initiation signals (Shine–Dalgarno sequence and ATG start codon) and encode a short leader peptide which is fused in-frame to the *lacZ* gene fragment. Cultures of CSH100 bearing each of these plasmids were mated with cultures of FW102 and dilutions of the mating mixtures were plated on medium containing streptomycin and kanamycin (see Materials and Methods). Colonies were then patch tested on plates containing chloramphenicol. As shown in Table 1, most of the clones were resistant to chloramphenicol, indicating that cointegrant episomes had been transferred. However, the desired chloramphenicol sensitive class was also present at a frequency ranging from 11 to 20%. β-galactosidase activity assays showed that the chloramphenicol-sensitive strains exhibited the expected relative levels of *lacZ* expression, confirming that the promoters and translational signals were functional and that the recombination had properly reconstructed the *lacZ* gene.

Validation with a prokaryotic regulatory system

The artificial promoter/operator strategy was then tested using a well-characterized DNA-binding protein, the bacteriophage λ repressor (λcI) (9). In this experiment, customized indicator strains were provided with controlled steady-state intracellular concentrations of λcI protein. This was done using two IPTG-inducible λcI expression plasmids, pLX10 and pLX20 (Fig. 3; see also Materials and Methods). These plasmids direct expression of low and moderate amounts of λcI, respectively, and the precise degree of expression can be conveniently controlled by adding various concentrations of IPTG to the medium.

Two strains, S11-LAM1 and S11-LAM2, were constructed using the method described above. Each bears a 17 bp consensus λ operator site between the –35 and –10 hexamers of the promoter that controls the *lac* genes (Fig. 4A). The strains are identical except for substitutions at two positions in the –10 promoter elements that dramatically alter promoter strength. The strains were transformed with pLX10, pLX20, and also a control plasmid that does not express λcI (pLR1ΔcI) (5), and β-galactosidase activities of cultures grown at various IPTG concentrations were measured. As shown in Figure 4A, repression of *lacZ* was a direct function of IPTG concentration regardless of the strength of the promoter in the indicator construct and also regardless of which λcI expression plasmid was used. Substantial repression was observed even at zero IPTG because pLX10 and pLX20 direct the synthesis of significant amounts of λcI protein even under repressing conditions.

It should be noted that in this system IPTG results in lowered *lac* expression, which is the reverse of the effect seen with the wild type *lac* operon. The *lac* repressor protein, whose gene (*lacI*) is located on the F′, controls promoters located *in trans* on pLX10 and pLX20. It has no *cis* effect on the *lacZYA* operon because the normal *lac* control region on the F′ has been replaced by the artificial indicator construct. The F′ bears the overexpressing $lacI^q$ allele to ensure that the dose of Lac repressor protein present in the cell is sufficient to control promoters located on multicopy plasmids.

A mixed plating experiment was then performed to determine if this system is sensitive enough to permit the visual identification of colonies bearing subtly altered mutant derivatives of a DNA-binding protein. The λcI mutant E34K is known to bind operator sites with elevated affinity (10,11). Strain S11-LAM1 was transformed with each of two IPTG-inducible expression plasmids that bear either the wild type or the mutant E34K λcI gene. β-galactosidase assays performed over a range of IPTG concentrations yielded curves similar to those of Figure 4A (not shown); activities at an intermediate IPTG concentration (10 μM) were 124 and 78 units, respectively. Cultures of these two transformants were mixed in a ratio of 30:1 (wild type:mutant), and plated on LB media containing antibiotics, X-gal and 10 μM IPTG. After incubation for ~20 h, colonies bearing the mutant were pale blue and could be reliably distinguished from the slightly darker blue colonies that contained the wild type plasmid. When the plates were stored

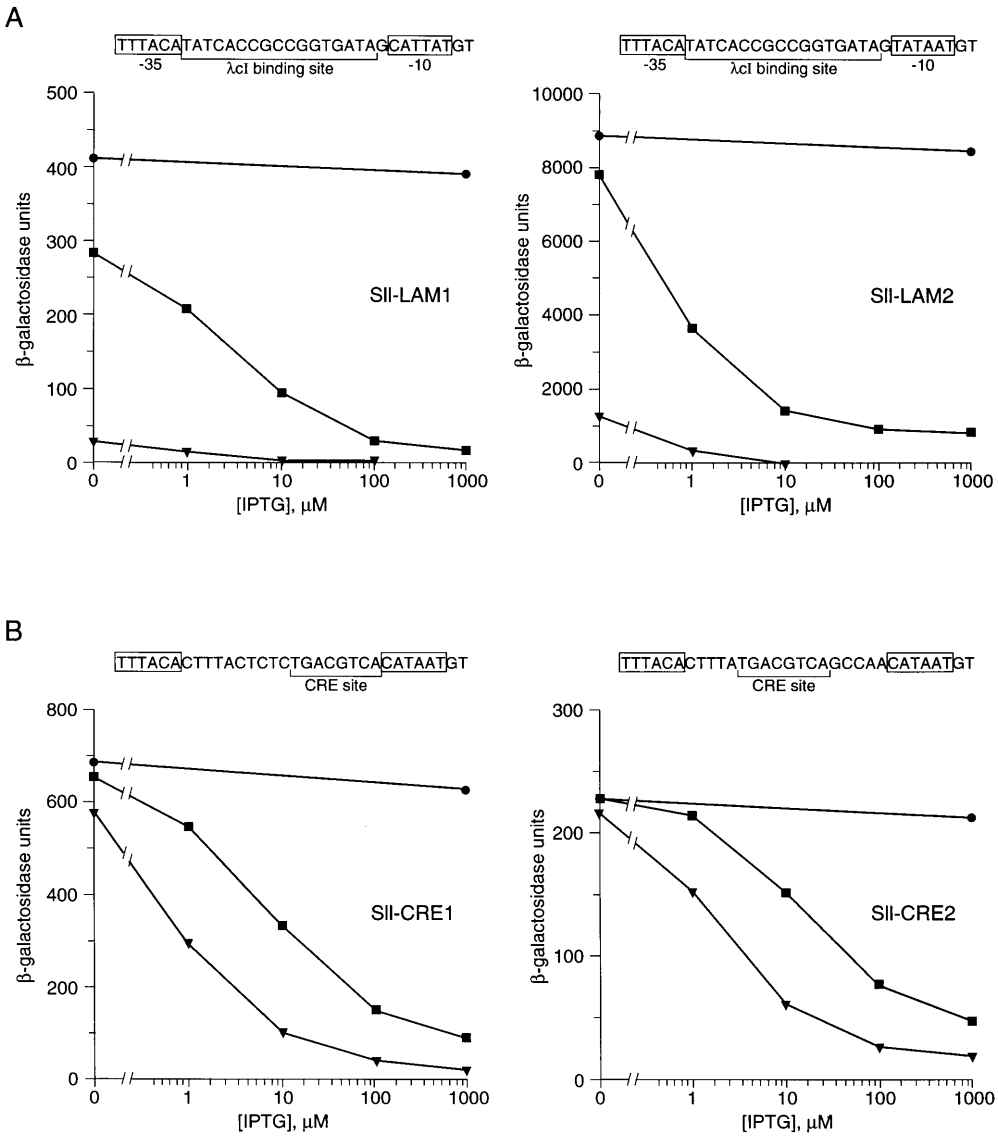

Figure 4. Repression of the *lacZ* gene in customized reporter strains by the corresponding site-specific DNA binding proteins. The sequence of the artificial promoter/operator present in each strain is indicated. Each strain was transformed with plasmids encoding the appropriate DNA binding protein and β-galactosidase activity was measured over a range of IPTG concentrations. (**A**) Strains S11-LAM1 and S11-LAM2 containing plasmids pLR1ΔcI (circles), pLX10 (squares) and pLX20 (inverted triangles). (**B**) Strains S11-CRE1 and S11-CRE2 containing plasmids pLR1ΔcI (circles), pLX20-CRB (squares) and pLX20-C-bZIP (inverted triangles).

at 4°C for several days, the color difference between the two classes of colonies became much more pronounced.

Validation with a eukaryotic regulatory system

To test the system with a eukaryotic DNA-binding protein, the human transcription factor CREB was chosen. CREB binds via a basic leucine zipper motif (bZIP) to 8 bp DNA recognition sites known as CRE sites (for review, see 12). Two *E.coli* indicator strains were constructed which each bear a consensus CRE site within the promoter spacer region and differ only in that the CRE sites are located on opposite sides of the DNA helix (Fig. 4B). These strains (named S11-CRE1 and S11-CRE2) were transformed with a control plasmid (pLR1ΔcI) or with derivatives of pLX20 in which the λcI gene has been replaced by either the complete CREB protein or the isolated CREB bZIP domain. As shown in Figure 4B, repression increased with increasing concentrations of IPTG in all cases. The precise location of the CREB binding site within the spacer region had only a minor effect on the efficiency of repression, although the resulting differences in promoter sequence did cause an ~3-fold difference in absolute β-galactosidase activity levels. Western blot analysis confirmed that the intracellular concentration of the full-length CREB protein was IPTG-dependent (data not shown). Strain S11-CRE1 was then tested phenotypically on lactose tetrazolium indicator plates, which yield white colonies when *lac* is expressed and red colonies when it is not. As shown in Figure 5, colonies containing the control plasmid were white on this medium, whereas those containing the bZIP plasmid were bright red.

DISCUSSION

The basic elements of the system described here are plasmid pFW11 and strains CSH100 and FW102. These elements provide

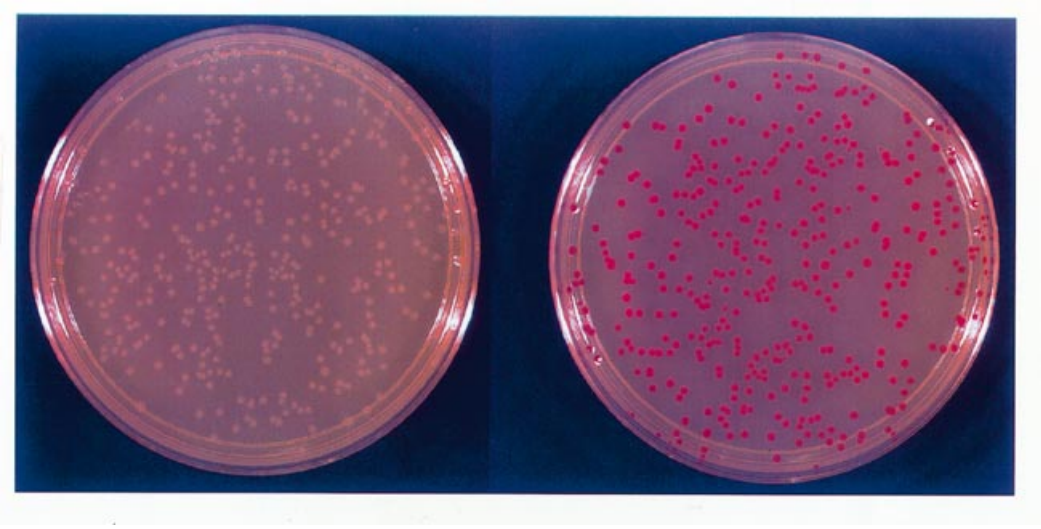

Figure 5. Effect of expression of the CREB bZIP domain in an indicator strain. *Escherichia coli* strain S11-CRE1 was transformed with either the control plasmid pLR1ΔcI (**left**) or the CREB bZIP expression plasmid pLX20-C-bZIP (**right**), plated on tetrazolium lactose plates containing ampicillin and kanamycin, and incubated overnight at 37°C.

a convenient and flexible method of constructing stable *E.coli* strains in which the *lacZYA* operon is placed under the control of any desired natural or artificial control region. Plasmid pFW11-null, a specialized derivative of pFW11, simplifies the assembly of artificial promoters that contain sites to which specific DNA binding proteins can bind, thereby repressing expression of the *lac* genes. Finally, plasmids pLX10 and pLX20 provide the means to express the corresponding proteins in *E.coli* strains at concentrations that are appropriate for genetic experiments. With this system essentially any site-specific DNA-binding protein can be made to exhibit a phenotype in *E.coli* that can be measured quantitatively and that can also be exploited in large scale genetic screens or selections.

Initial assembly of a customized promoter–operator construct can be achieved simply by annealing a pair of custom oligonucleotides together and ligating them directly into the backbone of pFW11-null (see Materials and Methods). Optionally, a white/blue colony screen may be used at this stage to distinguish correctly made derivatives of pFW11-null from any improperly ligated or parental plasmids that may be present in the ligation mix. In an appropriate cloning strain, a promoter-containing derivative of pFW11-null will yield blue colonies on X-gal/IPTG indicator plates, whereas incorrect constructs (which have no promoter in front of the *lacZ* gene fragment) will yield white colonies. This screen is analogous to the blue/white screen frequently used with popular cloning vectors such as pUC18 and its relatives (13), but in this case the colony colors are reversed. Any cloning strain that expresses the LacZ α-complementing fragment (lacZΔM15) may be used. Examples are JM105, DH5-α, XL1-Blue and Top10. With most of these strains, IPTG (1 mM) must be added to the medium for full expression of the α-complementing fragment.

Several other features of pFW11 should be noted. A strong transcriptional terminator located upstream of the polylinker region attenuates any read-through transcription coming from upstream of the artificial promoter/operator which might otherwise mask protein-mediated repression of the *lac* genes. The DNA just upstream of the *Eco*RI site is derived from pBR322, allowing sequencing primers that anneal to this region to be used. A secondary cloning region located between the *Kn* and *lacI′* genes contains restriction sites for *Mfe*I and *Bgl*II. These enzymes produce cohesive ends that are compatible with those produced by several frequently used enzymes, including *Eco*RI, *Bam*HI, *Bcl*I, *Apo*I, *Bst*YI and *Sau*3AI. Any genes or DNA sequences placed in this region will be transferred to the F′ and will be present permanently in the finished indicator strain. This can be especially useful in cases where the DNA binding protein of interest requires auxiliary foreign proteins for its proper functioning. Finally, the *Bgl*I site, which is unique in pFW11 and present naturally in the *lacZ* gene is very useful for transferring existing *lac* fusions from other vectors or strains into the pFW11 system.

Many existing protein expression vectors are not suitable for use in genetic experiments because they are designed for maximum protein production. With such vectors, cells are grown under repressing conditions and the culture is induced only when it has reached an optimal density. Large amounts of protein are then produced, but the cells usually stop growing and often die. Genetic studies, on the other hand, require that the target protein be expressed continuously at levels that are high enough to elicit a phenotype, but not so high that they become toxic to the cells. Plasmids pLX10 and pLX20 provide appropriate low and moderate expression levels that can be fine-tuned with IPTG. The unique *Nde*I site located at the start codon of both of these plasmids facilitates insertion of foreign genes. An *Nde*I site may be placed at the start codon of any desired target gene and used, together with any convenient downstream site, to move the foreign gene into pLX10 or pLX20, replacing the λcI gene. It should be noted that foreign proteins will become toxic to *E.coli* at various expression levels. Conditions should be chosen such that the strain produces colonies of uniform size on plates and grows in liquid culture at approximately the same rate as control cultures that lack the expression plasmid.

One of the advantages of working with the *lac* system is the wide range of analytical tools that can be used to monitor *lac* expression levels (8,14). For example, there are several types of colorimetric indicator media. On LB plates containing X-gal, colonies change from white to shades of blue over a range of *lac* expression that corresponds to ~20–100 β-galactosidase units. MacConkey lactose and tetrazolium lactose media, on the other hand, have a higher useful range (~100–1000 U) and are most sensitive at the low and high end of the range, respectively. Most indicator media other than

X-gal detect by-products of the fermentation of lactose, a process which requires the products of the *lacZ* and *lacY* genes. Lactose minimal medium, which can be used for true genetic selections, also requires both *lacZ* and *lacY.* Many existing *lac* fusion techniques produce fusions to only the *lacZ* gene. These systems work well with X-gal and in liquid enzyme assays, but with lactose fermentation media they either do not work at all or (if *lacY* is provided *in trans*) they work, but with greatly lowered sensitivity. Fusions that include the entire *lac* operon, such as those described here, result in strains that have maximum sensitivity and flexibility. It should be noted that protein expression from pLX10 or pLX20 will be fully induced whenever lactose-containing medium is used. IPTG has no effect in this case.

In designing promoter/operator constructs, the inherent promoter strength can be changed at will by modifying the sequence and/or spacing of the –35 and –10 hexamer elements. In general, changes that increase or decrease conformity to the consensus pattern TTGACA-N_{17}-TATAAT will correspondingly strengthen or weaken the promoter (for review, see 15). However, because overall promoter strength can usually be predicted only approximately, some applications may require the construction and testing of several strains that bear promoters with slightly altered –10 or –35 hexamers. The ability to easily construct strains with modified promoters and the availability of indicator media that are sensitive at different *lac* expression ranges are therefore distinct advantages of this system.

Many studies have established the feasibility of using artificial *E.coli* genetic systems to study both native and foreign proteins (16–25). Often a protein or protein domain that mediates dimerization or some other specific protein–protein interaction is fused to the DNA binding domain of a prokaryotic repressor. When expressed in *E.coli*, the resulting protein complex can bind to a strategically placed DNA site and repress a reporter gene, thus creating a phenotype that can be exploited. The system described here is suitable for experiments of this type. Two strategies have also been developed that result in activation of a reporter gene. In one case two promoters are positioned facing each other such that transcription initiating at one of them interferes with transcription from the other. When a foreign protein represses the first promoter, transcription from the other promoter (which is fused to an antibiotic resistance gene) is no longer impeded, and the cells become resistant to the antibiotic (25). In the second case, two foreign protein domains are fused separately to a DNA-binding protein and to a subunit of RNA polymerase, respectively. An interaction between the foreign domains activates transcription from an artificial construct in which a DNA binding site positions the first chimeric protein properly with respect to a promoter that controls a reporter gene (21). The strain construction technique described here has been used for the construction of single copy reporters for use with the latter assay (S. Dove, personal communication).

Activation-based and repression-based systems have both been used successfully to search through large plasmid pools to find rare variants that encode proteins having a specific activity. For example, an activation-based system reliably identified a plasmid encoding *trp* repressor from a mixture of plasmids containing only 1 part in 100 000 of the desired plasmid (25). A repression-based system picked out a plasmid encoding an HMG-box protein from a large human cDNA library based on the protein's ability to heterodimerize with the c-Myc oncoprotein, inactivating an artificial chimeric repressor that dimerized via its c-Myc domain (22). Artificial systems have also been used to screen large numbers of mutants of a given gene to identify functional variants (16). Both activation-based and repression-based systems can be exploited not only for genetic screens but also for genetic selections. The choice of an appropriate system will depend upon the protein–protein and protein–DNA interactions involved in each specific case. Protein domains derived from bacteria, yeast and mammals have all been made to exhibit usable phenotypes in *E.coli*. Based on the success of these systems, the use of *E.coli* genetic systems for the characterization and analysis of foreign proteins can be expected to continue to increase in the future.

ACKNOWLEDGEMENTS

I wish to thank James Hoeffler for providing plasmids containing the CREB gene. I am very grateful to Michael Malamy and Ann Hochschild for helpful discussions, and to Simon Dove and Ann Hochschild for providing critical comments on the manuscript. I am also very grateful to Ana Pis-Lopez and Vladimir Podolny for assistance in the construction of plasmids and of strains. Finally, I wish to thank Vivek Bajaj and C. Lee, Simon Dove, Robert Langdon, Jennifer Leeds and Jon Beckwith, and Vladimir Podolny for testing the system in various experimental situations. This work was supported by a Postdoctoral Cancer Fellowship from the American Cancer Society, Massachusetts Division, Inc. and by National Institutes of Health grant GM44025 (to Dr Ann Hochschild).

REFERENCES

1 Miller,J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2 Vieira,J. and Messing,J. (1987) *Methods Enzymol.*, **153**, 3–11.
3 Casadaban,M.J., Chou,J. and Cohen,S.N. (1980) *J. Bacteriol.*, **143**, 971–980.
4 Derman,A.I., Puziss,J.W., Bassford,P.J., Jr and Beckwith,J. (1993) *EMBO J.*, **12**, 879–888.
5 Whipple,F.W., Kuldell,N.H., Cheatham,L.A. and Hochschild,A. (1994) *Genes Dev.*, **8**, 1212–1223.
6 Backman,K. and Ptashne,M. (1978) *Cell*, **13**, 65–71.
7 Santiago-Rivera,Z.I., Williams,J.S., Gorenstein,D.G. and Andrisani,O.M. (1993) *Protein Sci.*, **2**, 1461–1471.
8 Miller,J.H. (1992) *A Short Course in Bacterial Genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
9 Ptashne,M. (1992) *A Genetic Switch.* Cell Press and Blackwell Scientific Publications, Cambridge, MA.
10 Nelson,H.C.M. and Sauer,R.T. (1985) *Cell*, **42**, 549–558.
11 Whipple,F.W., Ptashne,M. and Hochschild,A.H. (1997) *J. Mol. Biol.*, **265**, 261–265.
12 Lee,K.A.W. and Masson,N. (1993) *Biochim. Biophys. Acta*, **1174**, 221–233.
13 Vieira,J. and Messing,J. (1982) *Gene*, **19**, 259–268.
14 Silhavy,T.J., Berman,M.L. and Enquist,L.W. (1984) *Experiments with Gene Fusions.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15 Harley,C.B. and Reynolds,R.P. (1987) *Nucleic Acids Res.*, **15**, 2343–2361.
16 Hu,J.C., O'Shea,E.K., Kim,P.S and Sauer,R.T. (1990) *Science*, **250**, 1400–1403.
17 Marchetti,A., Abri-Marti,M., Illi,B., Cesareni,G. and Nasi,S. (1995) *J. Mol. Biol.*, **248**, 541–550.
18 Jappelli,R. and Brenner,S. (1996) *J. Mol. Biol.*, **259**, 575–578.
19 Zeng,X. and Hu,J.C. (1997) *Gene*, **185**, 245–249.
20 Ma,J. (1992) *Gene Exp.*, **2**, 139–146.
21 Dove,S.L., Joung,J.K. and Hochschild,A.H. (1997) *Nature*, **386**, 627–630.
22 Bunker,C.A. and Kingston,R.E. (1995) *Nucleic Acids Res.*, **23**, 269–276.
23 Castagnoli,L., Vetriana,C. and Cesareni,G. (1994) *J. Mol. Biol.*, **237**, 378–687.
24 Cairns,M.T., Green,A.J., White,P.M., Johnston,P.G. and Brenner,S. (1997) *Gene*, **185**, 5–9.
25 Elledge,S.J., Sugiono,P., Guarente,L. and Davis,R.W. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 3689–3693.