Nucleic Acids Research, Vol. 20, No. 7 1785–1791

# Modifications of the *E.coli* Lac repressor for expression in eukaryotic cells: effects of nuclear signal sequences on protein activity and nuclear accumulation

Annabeth Fieck, Denise L.Wyborski and Jay M.Short*
Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037, USA

Received August 27, 1991; Revised and Accepted February 10, 1992

## ABSTRACT

**Eukaryotic expression vectors designed to produce *E. coli* Lac repressor protein targeted to the nucleus of mammalian cells were constructed. These constructions carry the *lac* repressor gene *(lacI)* fused at different positions to a nuclear localization sequence (NLS) from either the SV40 large T antigen or the adenovirus E1a. When the NLS's were fused to the *lacI* gene at the 5′ end, the protein produced exhibited tighter repression of beta-galactosidase expression than the unmodified LacI protein. Localization sequences at the extreme 3′ end of the gene generally diminished induction by IPTG, while introduction of the SV40 NLS nine base pairs upstream of the 3′ end eliminated repressor activity. When either NLS was placed at the 3′ end behind a random nine base pair linker, the activity of the LacI protein depended on the sequence of the linker, and in 9 of 10 linkers tested, activity of the protein was adversely affected. The one exception was the fusion protein from p3′ss, which had the NLS at the 3′ end of *lacI* behind the nine base pair linker, AGC AGC CTG (ser-ser-leu). This protein exhibited efficient nuclear accumulation, strong repressor activity and greater sensitivity to IPTG induction. The functional linker from the p3′ss fusion protein extends the leucine zipper heptad repeat located at the C-terminus of the protein. These data support the role of the leucine zipper in tetramer formation and predict that extension of this zipper will further stabilize the protein. This modified *lacI* gene should be valuable for improved adaptation of the prokaryotic regulatory system to eukaryotic cells.**

## INTRODUCTION

The targeting of cytoplasmic proteins to the nuclei of eukaryotic cells has been accomplished efficiently by the fusion of DNA elements coding for nuclear signal sequences to various regions of heterologous genes (1,2). The fusion of these signal sequences to genes of interest can, however, dramatically affect the expression from those genes or the activity of the expressed fusion products. The effects of different signal sequences placed at various positions within the *lacI* gene on the activity and nuclear localization of the LacI repressor from the *E. coli* lactose operon have been examined. The aim was to obtain a nuclear signal sequence/*lacI* gene fusion which efficiently expressed a repressor having high affinity for operator DNA, but which could be induced with relatively low levels of IPTG, and which was readily transported to the nucleus of mammalian cells.

The Lac repressor binds as a homotetramer to the *lac* operator, blocking transcription of the beta-galactosidase gene (3, review). Inducers such as allolactose (the natural substrate of beta-galactosidase) or isopropyl-beta-D-thiogalactopyranoside (IPTG, a gratuitous synthetic inducer) bind to the repressor and cause a conformational change which effectively decreases the affinity of the repressor for the operator. The approximately 1000-fold reduced affinity of the repressor for operator is similar to the affinity of the repressor for non-specific DNA (the $K_a$ varies depending on the identity of the non-specific DNA), and consequently, the specific binding to operator DNA is competed out by the huge excess of non-specific DNA (3). When the repressor is removed from the operator, transcription from the lactose operon can resume. The elements of the lactose operon have been shown to work in eukaryotic cells (4, 5) and constitute a powerful tool for the control of gene expression in a variety of experimental systems (6, 7).

For increased utility in eukaryotic expression systems, the elements of the lactose operon can be modified in a number of ways. One such modification is the fusion of the Lac repressor to nuclear localization signal sequences. Such fusions allow the protein to be transported and maintained in the nuclei of eukaryotic cells. To study the effects of the nuclear localization signal sequences on Lac repressor function, the SV40 large T antigen signal sequence or the adenovirus E1a signal sequence were fused to the *lacI* gene at various positions. The activities of the resulting proteins were determined *in vitro* and *in vivo* using both eukaryotic and prokaryotic systems. We found that the nuclear localization sequence (NLS)/*lacI* fusions resulted in the production of proteins with significantly different activities than the unmodified protein. Changes in the activities ranged from complete loss to improved repressor function. The changes correlated with the maintenance or destruction of important domains within the Lac repressor protein.

* To whom correspondence should be addressed

## MATERIALS AND METHODS

### Reagents and Enzymes

The restriction and modification enzymes used were obtained from Stratagene Cloning Systems (SCS) and sequenase from U.S. Biochemicals. COS-203 was provided by ATCC, pMJR1560 was purchased from Amersham, and PyF9-1CAT was generously provided by Dr. Heiner Westphal, NIH. All other plasmids, phage stocks, DNA primers, and reagents used in cloning experiments were obtained from SCS. The immunological reagents for indirect immunofluorescence were purchased from Pierce and CJ236 was obtained from Bio-Rad.

### Plasmid Constructions

The prokaryotic Lac repressor expression vectors were produced from a modified pBluescript vector (8), pInt.1, which lacks all elements of the *E. coli* lactose operon. The EcoRI/PstI fragment from pMJR1560 (9) was inserted into pInt.1 to produce pInt.*lacI* (Figure 1), variations of which were cloned by placing adenovirus E1a NLS, 5′ AAG AGG CCT AGA CCT 3′ (10) or the SV40 large T antigen NLS, 5′ CCT AAG AAG AAG AGG AAG GTT 3′ (2) at various positions within the *lacI*$^{q}$ gene (Figure 2) using site-directed mutagenesis of single- stranded DNA (11).

Single-stranded DNA for mutagenesis was produced from a 25 ml culture of phagemid infected CJ236 as per supplier's instructions. Site-directed mutagenesis of the single-stranded phagemid was performed by annealing the DNA to mutagenic primers specific for the changes to be made in the phagemid. Primer extension was carried out essentially as described in the instruction manual for the phagemid, pBluescript II SK(+) (SCS). Following incubation, 10 $\mu$l of the reaction mixture was transformed into 100 $\mu$l of XL1-blue competent cells (SCS). Positive clones were detected by restriction digests of mini-prep DNA from isolated colonies using the unique restriction site, StuI, included in the primer sequence, and then confirmed by sequencing PlasmidQuik (SCS) derived DNA with T7 DNA polymerase.

The eukaryotic Lac repressor expression vectors were produced from pInt.*lacI* by insertion of the polyoma F9-1 promoter from PyF9-1CAT at the BglII site upstream from the *lacI*$^{q}$ gene (12). At the 5′ end of the gene in all vectors, the upstream ATG initiator codons were removed and a modified Kozak consensus ATG sequence (CCACCATG) replaced the *lacI* GTG start codon by site-directed mutagenesis (13). The SV40 polyadenylation site and small intron from pKO-Neo were placed downstream of the *lacI* gene between the PstI and BamHI sites (14). The hygromycin-B resistance gene from COS-203 was inserted downstream of the polyadenylation signal at the BamHI site (15). The PstI, BamHI, and all sites in the upstream polylinker between XbaI and EcoRI were eliminated in the cloning procedures.

### Plate Assay of *lac* Repressor Activity

Lac repressor prokaryotic DNA constructions which have *lacI*-nuclear signal sequence fusions were tested for the ability of the hybrid protein to repress the wild-type *lacZ* gene and to respond to the inducer, IPTG, using a bacterial plate assay. This was performed by transforming the constructs of interest into the DH-9 strain of *E. coli* (generously provided by Dr. Magnus Pfahl) which lacks a functional *lacI* gene but carries an intact *lacZ* gene. Single colonies were selected and grown for 6 hours at 37°C with shaking in 5 ml of LB supplemented with ampicillin. The absorbances of these cultures at 600 nm was measured spectrophotometrically and adjusted to 0.5. 500 $\mu$l of each culture were plated as an overlay onto LB-amp-agar plates and allowed

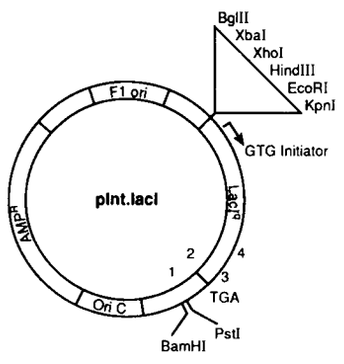



**Figure 1.** Constructions with *lacI*-Nuclear Signal Sequence gene fusions. Vectors used for testing the effects of the NLS on LacI activity were derivatives of the parent vector, pInt.lacI having either the signal sequence of the SV40 large T antigen or the adenovirus E1a early region fused to *lacI* at various positions. Numbers on the diagram indicate points of signal sequence fusions: (1) 3′ end of *lacI* gene with randomly derived nine base pair linkers plus a StuI site between the last codon of the gene and the first base of the signal sequence, (2) 3′ end of gene nine base pairs upstream of stop codon, (3) 3′ end of gene, (4) NarI site of *lacI* ninety base pairs upstream of stop codon.

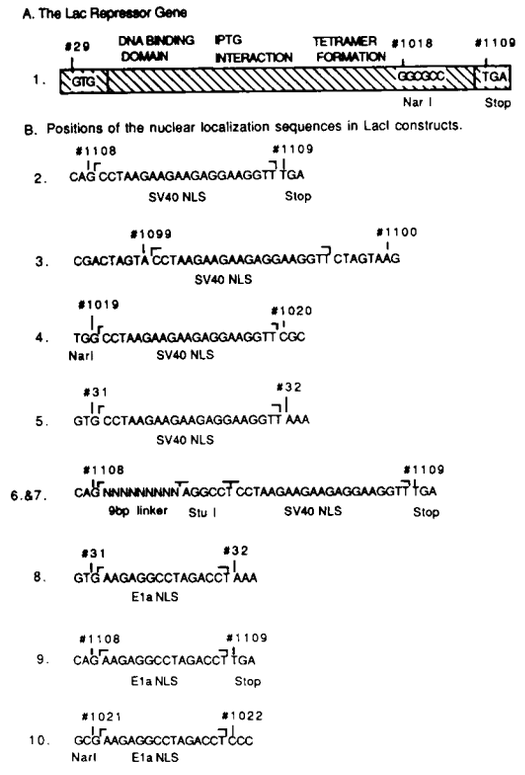



**Figure 2.** Nuclear localization signal sequences (NLS) and positioning in the *lac* repressor gene. A—Diagram of *lac* repressor gene with protein coding domains. B—Positions of the nuclear localization sequences in the *lacI* constructions. All constructions contain the GTG transcriptional start site. The sequence of the 9bp linker is described in the text. The numbers above the nucleotide sequences correspond to the *lacI* sequence of Farabaugh (26). The peptide sequences of the nuclear localization signals are as follows: SV40 large T antigen: 5′ pro lys lys lys arg lys val 3′. Adenovirus E1a early region: 5′ lys arg pro arg pro 3′.

to stabilize for 30 minutes. A grid was drawn on the bottom of each plate delineating 10 squares. Within the boundaries of each square a 2μl drop of 250mg/ml X-gal was placed on the overlay and dried. On top of each of these drops was placed a 10 μl spot of different dilutions of IPTG in descending order from solutions of 1M, 500mM, 50mM, 10mM, 5mM, 2mM, 1mM, 500μM and 100μM. The last square had no IPTG added. After overnight incubation of the plates at 37°C, the level of blue in the squares was quantified by densitometry using the Stratascan 7000 system, to determine three characteristics of the repressor produced from each vector construction. Inducibility was measured by determining the lowest concentration of IPTG required to turn the cells blue, while the extent of induction was judged by the intensity of the blue color after maximum induction (using up to 500 nmols/μl of IPTG). Repressor efficiency was also examined by determining 'leakiness', which was represented by a light blue color in the absence of IPTG.

### Filter Binding Kinetic Experiments on Purified Repressor

Lac repressor protein was purified from NM554 (*lac*−) cells (SCS) as described previously (16). Dissociation constants were determined by the filter binding method of Riggs, *et al.* (17). Briefly, operator-containing bacteriophage (lambda gt11) DNA was labelled with $^{32}$P-dCTP using the prime-it labelling kit (SCS). The labelled DNA was diluted with non-labelled lambda gt11 DNA at 1μg/μl to a specific activity of $2\times10^5$ cpm/ng of total DNA and 1.5 μg of this DNA was mixed with 4.2 ng of purified repressor protein. After 20 minutes at room temperature for equilibration, a 50-fold excess of non-labelled operator-containing DNA was added and samples were taken at time zero and ten minute intervals thereafter. The samples were taken in triplicate and filtered through millipore nitrocellulose filters of 0.25mm pore size. The filters were washed twice, and counted in scintillation vials with 5 ml liquid scintillation fluid. The dissociation constant was determined by the equation; $K_b = -$ (ln cpm′/cpm°)/time′(s), where: cpm′ = counts per minute at time′, cpm° = counts per minute at time zero, time′ = time in seconds after addition of excess operator DNA. Background was determined by including with each set of samples a filter through which labelled operator-containing DNA alone was passed. This filter was washed twice and counted (identical to the repressor samples). The cpm determined for the control filter was subtracted from the cpm determined for each experimental sample. The filtering and binding buffers are as described by Riggs, et al. (17).

| CONSTRUCT 5' Lac I 3' | NLS PLACEMENT | A REPRESSION WITHOUT IPTG | B [IPTG] REQUIRED FOR INDUCTION | C EXTENT OF INDUCTION |
|---|---|---|---|---|
| 1 | NONE | INCOMPLETE | 2 nmols/ul | 1.0 |
| 2 | 3' end | COMPLETE | 2 nmols/ul | 0.5 |
| 3 | 9bp upstream of stop codon | NONE | N/A | N/A |
| 4 | Nar I site | COMPLETE | 1 nmol/ul | 0.6 |
| 5 | 5' end after GTG codon | COMPLETE | 1 nmol/ul | 1.0 |
| 6 | 3' end after 9bp linker #1 | COMPLETE | 0.5 nmol/ul | 1.0 |
| 7 | 3' end after 9bp linker #2 | NONE | N/A | N/A |
| 8 | 5' end after GTG codon | COMPLETE | 1 nmol/ul | 1.0 |
| 9 | 3' end | COMPLETE | 2 nmol/ul | 0.5 |
| 10 | Nar 1 site | COMPLETE | 1 nmol/ul | 0.5 |

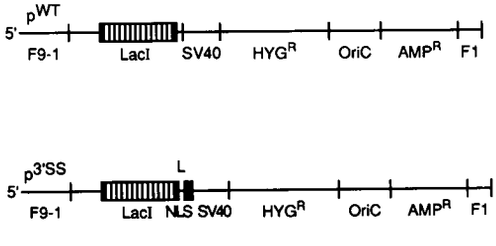
**Figure 3.** Placement of the signal sequences into the *lacI* gene and the resulting effects on repressor activity as measured by the *E. coli* plate assay. Vertical stripes—*lacI* gene, Black box—SV40 large T antigen nuclear signal sequence, Box with diagonal lines—adenovirus E1a nuclear signal sequence, horizontal line between *lacI* and NLS—nine base pair linker. The two linkers designated by the numbers 1 and 2 are two representative examples of the 10 linkers tested. Only one of the linkers had the characteristics of linker #1, three of the linkers had the characteristics of linker #2, and the remaining six linkers adversely affected the activity of the resulting LacI proteins resulting in repression and inductions patterns similar to those of the *lacI*/NLS fusion at the extreme 3′ end of the gene. The extent of induction was measured by determining the blue intensity at high concentrations of IPTG (up to 500nmols/μl) using Stratascan 7000. However, the extent of induction did not increase above concentrations of 50nmols/μl. The constructs are numbered (#1–10) to correlate with those in Figure 2.

### Transfection of NIH 3T3 Cells with Eukaryotic Expression Vectors

Transfection of NIH 3T3 cells was performed using the $CaPO_4$ protocol of the mammalian transfection kit (SCS) as per manufacturers instructions. For each transfection, 20 μg of CsCl purified DNA was used.

### Indirect Immunofluorescence of Repressor in Stably Transfected Cell Lines

The protocol used was described previously (18). Briefly, twenty-four well tissue culture dishes were seeded with varying numbers of stably transformed cells and incubated overnight at 37°C. The cells were washed three times with 1×PBS, fixed in a solution of 1×PBS/4% formaldehyde for 10 minutes, and permeabilized for two minutes in a solution of 1×PBS/0.1% triton X-100. The cells were then washed five times with 1×PBS and incubated for one hour with polyclonal antibody to Lac repressor diluted 1:100 or 1:1000 in 1×PBS/5% normal goat serum at 37°C. After five more washes in 1×PBS/0.5% BSA/0.1% Tween-20, the cells were incubated with fluorescein-conjugated goat-anti-rabbit antibody at 25μg/ml in the same solution at 37°C for 1 hour. Cells were washed five times with 1×PBS/0.1% Tween-20 and fluorescence was visualized and photographed using a Nikon fluorescent microscope. The intensity and localization of the fluorescein labelling was determined by densitometry using the Stratascan system.




| Element | Description | Element | Description |
|---|---|---|---|
| F9-1 | Mutant polyoma promoter | OriC | *E coli* Origin of Replication |
| LacI | *E. coli* Lac repressor gene | AMP$^R$ | Ampicillin Resistance Gene |
| SV40 | SV40 Polyadenylation signal | F1 | Filamentous phage F1 Origin of Replication |
| HYG$^R$ | Hygromycin Resistance Gene with HSV promoter and PolyA | L | Linker: AGC AGC CTG + STU I |
| | | NLS | SV40 Nuclear Localization Signal |

**Figure 4.** Eukaryotic expression vectors transfected into NIH-3T3 cells. pWT = Wild-type repressor; p3′SS = SV40 NLS at 3′ end of gene behind the 9bp linker.

## RESULTS

### Effects of NLS/*LacI* gene fusions on repressor activity

The addition of sequences to the *lacI* gene designed to direct the encoded protein to the nucleus of eukaryotic cells had varied effects on the activities of the resulting repressors. The effects on the *lacI* constructions were determined through the use of a plate assay in *E. coli*. This assay measured the ability of the fusion repressors to completely abolish beta-galactosidase activity, and to be induced upon exposure to IPTG. The fusion proteins acted in a specific, reproducible manner, usually different from that exhibited by the wild-type repressor (Figure 3). The wild-type repressor displays incomplete repression when plated on X-gal in the absence of IPTG (Figure 3, row 1A), suggesting it is slightly 'leaky' (ie...incompletely blocks transcription). The induction threshold for this protein is observed after the addition of 2nmols/$\mu$l of IPTG in a 10 $\mu$l drop. Induction is considered complete if the blue intensity after induction with IPTG equals that of the DH-9 cells plated on X-gal alone (DH-9; *lacZ*$^+$, *lacI*$^-$). Maximum induction with the wild-type repressor was achieved with 50 nmols/$\mu$l of IPTG.

The insertion of either the SV40 or E1a NLS to all positions which did not completely inactivate the repressor molecule eliminated the leaky phenotype, as shown in Figure 3, column A. Addition of either NLS at the extreme 3′ end of the *lacI* gene (just prior to the stop codon) produced proteins which were more efficient repressors than wild-type (non-leaky). However, the repression from these 3′ fusion proteins was never completely eliminated (Figure 3, rows 2C and 9C) even after the addition of 10$\mu$l of a 1M solution of IPTG, 500-fold more than what was required with wild-type repressor. Addition of the SV40 NLS at the 3′ end of the *lacI* gene, just nine base pairs upstream of the stop codon abolished all repressor activity (Figure 3, row 3A). Both NLS's placed at the NarI site of the *lacI* gene resulted in a non-leaky protein which required lower concentrations of IPTG for initial induction, but could not be completely induced at the higher IPTG concentrations (Figure 3, rows 4 and 10). The fusion of the NLS's to the 5′ end of the *lacI* gene just after the GTG initiator codon produced proteins in DH-9 cells which exhibited non-leaky repression and inducibility at concentrations of IPTG 50% of those required by the wild-type protein, 10 nmols in a 10 $\mu$l drop (Figure 3, rows 5 and 8). Another set of gene fusions joined the SV40 NLS to the 3′ end of the *lacI* gene following random linkers of nine base pairs and a StuI site. The linkers were produced by site-directed mutagenesis using degenerate synthetic oligonucleotides. Different linkers dramatically changed the activity of the repressors to which they were attached. In nine out of ten cases the activity of the repressor from these constructions was adversely affected (Figure 3, row 7A for example). The one exception, (Figure 3, row 6), was a stronger repressor (non-leaky) than the wild-type, and showed induction at a lower concentration of IPTG than seen with any other construct, 25% of the amount required for unmodified protein or 5 nmols in a 10 $\mu$l drop. The linker between the *lacI* gene and the NLS in this construct was AGC AGC CTG which translates to serine, serine, leucine. This construct was selected for examination in eukaryotic cells. The E1a NLS was initially chosen for its small size in an attempt to minimize distortions

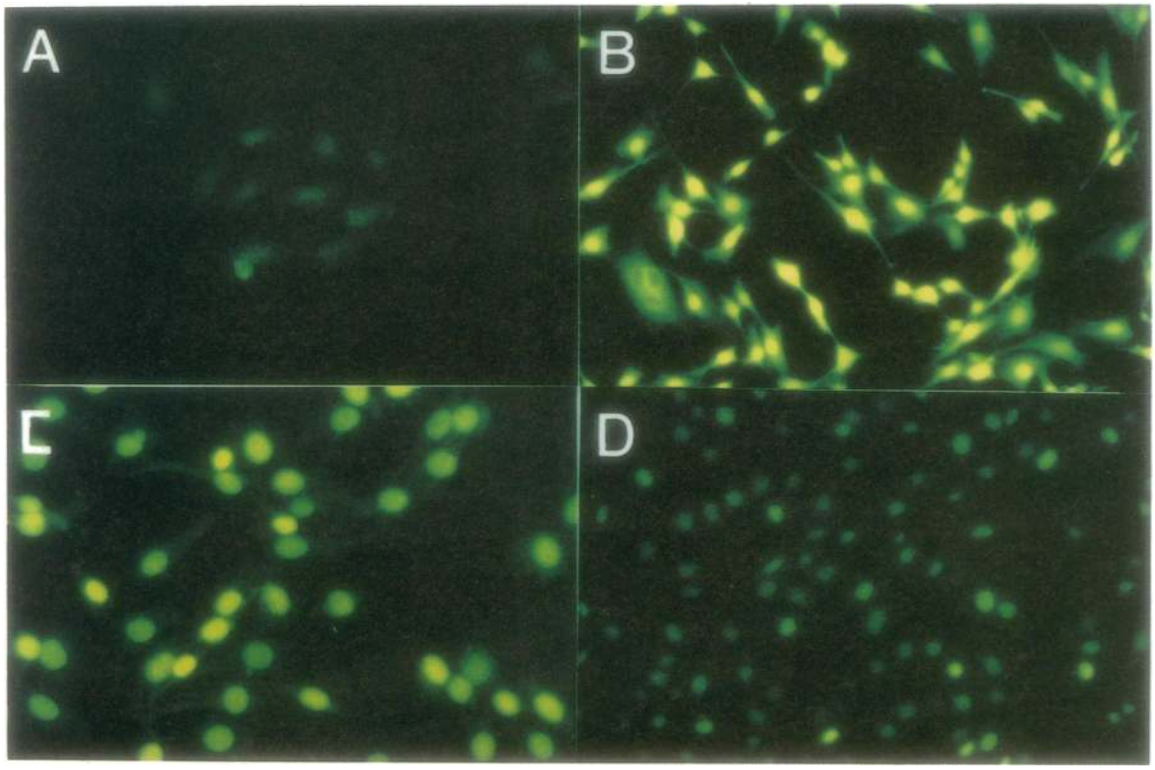



**Figure 5**. Indirect immunofluorescence of stable cell lines expressing LacI. Cells were labelled with an FITC-conjugated goat anti-rabbit antibody after initial binding of an anti-LacI polyclonal antibody produced in rabbits as described in the materials and methods. Panel A—Control 3T3 cells, Panel B—pWT (no NLS), Panel C—p3′ss (3′ LINKER/NLS) higher intensity, Panel D—p3′ss (3′LINKER/NLS).

of the *lacI* fusion proteins. However, its effects on the protein in any given position were the same as the larger SV40 NLS. Consequently, the studies in mammalian cells were performed using only the SV40 NLS/*lacI* fusions in vectors designed specifically for high expression in eukaryotic cells (Figure 4).

### Indirect immunofluorescence of *lacI* transformed cells

To determine the abundance and location of the repressor protein in the stable cell lines producing *lacI* mRNA, the cells were examined using an indirect immunofluorescence technique (see materials and methods). The cells are viewed under a fluorescence microscope and the presence of LacI protein is seen as green staining in the regions of the cells harboring the protein. To determine a background staining level of cells with no repressor, NIH 3T3 cells were used as controls yielding low stain intensity (Figure 5A). The cell lines with the F9-1 promoter driving the *lacI* gene without a NLS (pWT) showed strong fluorescence throughout the cells (Figure 5B). The stable cell lines from clones of the p3'ss construct which has the SV40 nuclear signal sequence at the 3' end of the *lacI* gene behind the optimum nine base pair linker displayed approximately 98% of LacI staining in the nucleus (Figure 5C and 5D). The immunofluorescence of the cells with the 3' signal sequence behind the nine base pair linker is shown at two levels of intensity to ensure that any cytoplasmic expression would be detected. These results were quantitated by densitometry and are tabulated in Table 1.

**Table 1.** Indirect Immunofluorescence of mammalian cells expressing LacI repressor from constructs with and without the SV40 nuclear signal sequence.

| Construct | Description | Fluorescence Intensity | Fluorescence Localization |
|---|---|---|---|
| NIH 3T3 (5A) | No repressor | 0.0 | Non-specific |
| pWT (5B) | No NLS | 0.8–1.0 | 10% nuclear |
| p3'ss (5D) | NLS at 3' of gene behind 9bp linker | 0.2–0.8 | 98% nuclear |

All constructs were tested in NIH-3T3 cells at three different cell densities. Quantification of the immunofluorescence was performed with the Stratascan 7000 densitometry system. The numbers are determined by measuring the maximum intensity of fluorescence in several different cells, and subtracting out background. Intensity measurements were normalized with intensity from cells expressing unmodified repressor being set equal to 1.0, and the intensity from the other cell lines represented as a fraction of 1.0.

**Table 2.** Results of LacI-filter binding experiments to determine values of dissociation constants for wild-type and modified LacI proteins.

| PROTEIN | $K_b$ (dissociation constant) | Relative amount of protein bound at equilibrium |
|---|---|---|
| LacI Previous* | $6.2\times10^{-4}$ sec$^{-1}$ +/− $1.3\times10^{-4}$ | 1.0 |
| LacI unmodified | $5.8\times10^{-4}$ sec$^{-1}$ +/− $2.0\times10^{-4}$ | 1.0 |
| LacI 3'ss w/linker | $3.9\times10^{-4}$ sec$^{-1}$ +/− $1.1\times10^{-4}$** | 1.2 |

*This is the value obtained in previous studies (19) included for comparison. The *lacI* unmodified is the protein produced for this study from our vector constructions.
**The standard deviation was determined using multiple time-points from several different experiments.

### Determination of the dissociation constant of the mutant and wild-type LacI proteins

The data from plate assays of the *lacI* fusion constructs suggested that proteins were being produced with activities differing from the wild-type protein due to the presence and position of the NLS. The *lacI* fusion in p3'ss which placed the NLS at the 3' end of the *lacI* gene downstream of a specific three amino acid linker produced a protein which exhibited more desirable repression and induction properties than the wild-type protein and was efficiently transported to the nuclei of mammalian cells. To investigate this protein further, it was purified from *E. coli* cells and compared to the purified wild-type protein in a filter binding assay previously used to determine the dissociation constant of different LacI proteins (19). The previously determined dissociation constant of wild-type Lac repressor (at an ionic strength of 0.025) is $6.2\times10^{-4}$ sec$^{-1}$ +/−$1.3\times10^{-4}$. We obtained a value of $5.8\times10^{-4}$ sec$^{-1}$ +/−2.0 x$10^{-4}$ for the wild-type protein (Table 2). The dissociation constant for the NLS/*lacI* fusion protein from p3'ss was $3.9\times10^{-4}$ sec$^{-1}$ +/−$1.1\times10^{-4}$. The dissociation constants obtained for the wild-type and NLS fusion proteins suggest that the signal sequence does not significantly affect normal dissociation kinetics of the protein in the absence of inducer. However, at equilibrium conditions, the amount of repressor bound to DNA is approximately 20% greater with the signal sequence fusion protein than with the wild-type protein. This may have been reflected in the data from the plate assay, as it could explain the loss of the 'leaky' phenotype with the modified repressor compared to that of the wild-type repressor in the absence of

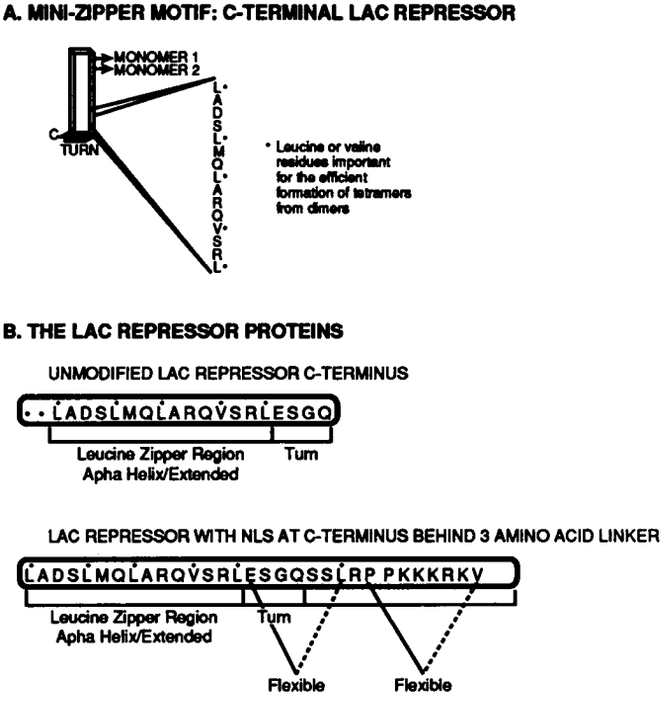



**Figure 6.** Predicted carboxy-terminal structure of the LacI repressor. A—Leucine mini-zipper domain near C-terminal end of LacI protein dimer showing residues important for tetramerization on each LacI monomer, B—Predicted secondary structure of the C-termini of the unmodified and modified Lac repressor proteins showing the seven amino acid heptad repeats and the residues which may act to extend the leucine heptad repeating structure.

IPTG. These data indicate that the Lac p3'ss fusion protein has a higher affinity for DNA, and that this increase may be mediated through a shift in the association constant.

## DISCUSSION

The fusions of Lac repressor to either the SV40 or E1a nuclear localization signal sequence resulted in the production of two general categories of proteins: those which lost all activity, and those which retained some level of repressor function. Of the ones retaining activity, all exhibited the ability to more tightly repress and reduce the production of beta-galactosidase as compared to the unmodified protein in *E. coli* cells. NLS's are commonly basic in nature as is the case with the two tested here (20). The fusion of these basic peptides to the Lac repressor adds a region of positive charge to the protein's charge profile. Previous studies of the interaction between the Lac repressor and operator show an extremely fast reaction rate of association which is attributed to an electrostatic attraction between a positively charged site on the repressor and the negatively charged phosphate backbone of the operator DNA (21). The addition of a positively charged sequence to the Lac repressor may add to the natural attractive forces between the repressor and operator. This interpretation is consistent with the kinetic data for the NLS/*lacI* fusion protein from p3'ss which exhibits an approximately 20% increase in the level of repressor-operator complex at equilibrium.

The indirect immunofluorescence studies in mammalian cells compared pWT to p3'SS only, based on this fusion protein's strong repression and increased sensitivity to IPTG induction. Approximately 98% of the repressor protein was transported to the nucleus of the NIH 3T3 cells. This represents an approximately 10-fold improvement in the transport of the repressor to the nucleus. Previously Labow et al. reported similar nuclear accumulation of Lac repressor when fused to the transcriptional activation domain of the herpes simplex virus type 1 virion protein 16 which had an SV40 NLS at the 5' end (18). The repressor in those studies was fused to a transcriptional activating domain that is normally transported to the cell nucleus without the aid of the SV40 NLS. Although it is likely that the 5' positioning of the NLS in the *lacI* gene is also effective at inducing nuclear transport, the contribution of the transcription factor sequences in this fusion protein for nuclear localization has not been determined.

The amino acid composition of the sixty C-terminal residues of NLS/*lacI* fusions having the SV40 NLS either at the C-terminus (amino acid #360), three residues in from the C-terminus, or behind a three amino acid linker at the C-terminus, were analyzed by computer modeling using the Intelligenetics PC/GENE software package. The NLS adds flexibility to the LacI protein, changing the 'flexibility profile' of the protein in the region of the fusion. Based on 'Flexpro' program analysis the greatest amount of flexibility is imparted to the LacI protein with the fusion of the NLS to the C-terminal end with or without the linker (22). In both cases the highest flexibility peak corresponds to the point of NLS fusion. The fusion of the NLS three amino acids internal to the C-terminus imparted some measure of flexibility to the LacI protein, but considerably less than the other fusions. Since this construction lacked repressor activity, increased flexibility may be an important factor in the ability of the repressor to retain activity after the addition of an NLS. The nuclear signal sequences are positively charged and, with the formation of tetramers, are brought into close proximity. The ability of these sequences to 'flex' away from each other may play a role in the ability of the tetramers to maintain stability with an increase in electrostatic repulsion from the positive charges.

Recent studies on the C-terminal end of the Lac repressor protein describe cooperative protein-protein interactions between dimers of Lac repressor (23 and 24). These reports describe leucine heptad repeats designated, 'mini-zipper motifs', at the C-termini which are responsible for the formation of tetramers from dimers (Figure 6A). The Lac repressor protein from p3'ss with the NLS fused at the C-terminus to the three amino acid linker plus a StuI site, adds a leucine residue seven amino acids from the C-terminal leucine of the 'mini-zipper', extending the heptad repeat (Figure 6B). It is probable that the added leucine residue from the linker at this position is further stabilizing the leucine zipper, overcoming potential repulsive forces of the NLS, and increasing the energetic favorability of tetramer formation. In all of the constructions described in this paper the 'mini-zipper' motif is preserved, however, the fusion protein from the p3'ss construction is the only one where the leucine zipper is extended. These data support the role of the leucine zipper in tetramer formation and predict that further extension of this zipper will further stabilize the protein. Such modifications may allow the design of other Lac repressor fusions to a variety of proteins without adversely affecting Lac repressor function.

The Lac repressor variants were produced and subsequently characterized in an attempt to identify an inducible negative control element capable of operating efficiently in mammalian cells and transgenic animals. With this goal in mind, other elements of the system were recently characterized and reported by this lab (25). The phenotype of the Lac repressor fused to the SV40 NLS at its C-terminus behind a three amino acid linker from p3'ss, which exhibits tighter repression and greater sensitivity to Lac inducers, should improve the adaptation and functionality of the Lac repressor system to eukaryotic cells, as well as prokaryotic cells, for control of gene expression.

## ACKNOWLEDGEMENTS

We wish to thank Dr. Mario Bourdon for assistance with the indirect immunofluorescence, Mark Bergseid and Matt Petri for protein purification, Tom Tomzinski and Tim Sanchez for supplying the lambda gt11 DNA, John Fox for his Stratascan expertise, and Dr. Kathleen Matthews for critical reading of the manuscript. This research was supported in part by NIH grant #2R44GM42291-02.

## REFERENCES

1. Dingwell, C., Sharnick, S.V., and Laskey, R.A. (1982) Cell 30, 449–458.
2. Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984) Cell 39, 499–509.
3. Bourgeois, S., and Pfahl, M. (1976) Adv. Protein Chem. 30, 1–99.
4. Hu, M.C.-T., and Davidson, N. (1987) Cell 48, 555–566.
5. Brown, M., Figge, J., Hansen, U., Wright, C., Jeang, K-T., Khoury, G., Livingston, D.M., and Roberts, T.M. (1987) Cell 49, 603–612.
6. Tepper, R.I., Levinson, D.A., Stanger, B.Z., Compos-Torres, J., Abbas, A.K., and Leder, P. (1990) Cell 62, 457–467.
7. Giordano, T.J., Deuschle, U., Bujard, H., and McAllister, W.T. (1989) Gene 84, 209–219.
8. Short, J.M., Fernandez, J.M., Sorge, J.A., and Huse, W.D. (1988) Nucleic Acids Research 16:15, 7583–7599.

9. Stark, M.J.R. (1987) Gene 51, 255–267.
10. Lyons, R.H., Ferguson, B.Q., and Rosenberg, M. (1987) Mol. and Cell Biol 7:7, 2451–2456.
11. Kunkel, T.A. (1985) Proc. Natl. Acad. Sci. USA 82, 488–493.
12. Bohnlein, E., Chowdhury, K., and Gruss, P. (1985) Nucleic Acids Research 13:13, 4789–4809.
13. Kozak, M. (1986) Cell 44, 283–292.
14. Van Doren, Hanahan, and Gluzman. (1984) J. Virology 50:2, 606–614.
15. Kioussis, D., Wilson, F., Daniels, C., Leveton, C., Taverne, J., and Playfair, J.H.L. (1987) EMBO Journal 6:2, 355–361.
16. Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Springs Harbor University Press, Cold Spring Harbor, New York.
17. Riggs, A.D., Suzuki, H., and Bourgeois, S. (1970) J. Mol. Biol. 48, 67–83.
18. Labow, M.A., Baim, S.B., Shenk, T., and Levine, A.J. (1990) Mol. and Cell Biol. 10:7, 3343–3356.
19. Riggs, A.D., Bourgeois, S., and Cohn, M. (1970) J. Mol Biol. 53, 401–417.
20. Silver, P. (1991) Cell 64, 489–497. Review.
21. Hamada, F., Ohshima, Y., and Horiuchi, T. (1973) J. Biochem 73, 1299.
22. Karplus, P.A., and Schulz, G.E. (1985) Naturwissenschaften 72, 212–213.
23. Alberti, S., Oehler, S., Wilcken-Bergmenn, B., Kramer, H., and Muller-Hill, B. (1991) The New Biologist 3:1, 57–62.
24. Chakerian, A.E., Tesmer, V.M., Manly, S.P., Brackett, J.K., Lynch, M.J., Hoh, J.T., and Matthews, K.S., (1991) J. Biol. Chem. 266:3, 1371–1374.
25. Wyborski, D.L., and Short, J.M. (1991) Nucleic Acids Research 19:17, 4647–4653.
26. Farabaugh, P.J. (1978) Nature 274:24, 765–769.