Nucleic Acids Research, 1997, Vol. 25, No. 21 **4209–4218**

# An efficient protocol for linker scanning mutagenesis: analysis of the translational regulation of an *Escherichia coli* RNA polymerase subunit gene

**Derek M. Dykxhoorn, Rebecca St. Pierre, Oded Van Ham and Thomas Linn***

Department of Microbiology and Immunology, Faculty of Medicine, University of Western Ontario, London, Ontario N6A 5C1, Canada

Received August 18, 1997; Accepted September 12, 1997

## ABSTRACT

**A protocol has been developed that is capable of saturating regions hundreds of basepairs in length with linker scanning mutations. The efficacy of this method stems from the design of the linker scanning mutagenesis (LSM) cassette which is composed of a selectable marker flanked by two oligonucleotides, each of which contains a recognition site for a different restriction endonuclease. The cleavage site for one endonuclease is within its recognition site, while the second endonuclease cleaves in the target DNA beyond the end of the cassette. Digestion with these endonucleases and subsequent ligation results in the replacement of 12 bp of the original target sequence with 12 bp of the linker scanning oligonucleotide. We have used this protocol to mutagenize a span of ~400 bp surrounding the start site of the gene for the β subunit (*rpoB*) of *Escherichia coli* RNA polymerase. The translation of the β mRNA has been shown previously to be regulated by the intracellular concentration of either β or β′. Analysis of the linker scanning mutations indicates that sequences extending a considerable distance both upstream and downstream of the start site are required for normal translation. Also a site that appears to be involved in translational repression by excess β′ has been identified.**

## INTRODUCTION

Linker scanning mutagenesis has proven extremely useful in systematically searching for regulatory sites in regions of DNA hundreds of basepairs (bp) in length. The original procedure as developed by McKnight and Kingsbury (1) first involves the creation of a series of 5′ and 3′ sets of deletions onto whose ends are ligated a synthetic oligonucleotide linker. Each mutation is sequenced and then the appropriate 5′ and 3′ deletions are paired so that their endpoints differ by the same number of nucleotides present in the linker. When these ends are ligated together via the linker, the missing nucleotides of the original sequence are replaced by those of the synthetic linker. Though effective, this approach involves a number of time consuming steps. Many deletions have to be generated and sequenced before appropriate 5′ and 3′ combinations can be identified, followed by a separate ligation for each 5′/3′ pair to produce the final linker scanning mutation.

We have developed a more efficient protocol for linker scanning mutagenesis that is capable of generating a library consisting of hundreds of mutations. This protocol makes use of a linker scanning mutagenesis (LSM) cassette which is composed of two synthetic oligonucleotides surrounding the selectable tetracycline resistance gene. The oligonucleotide on one side of the tetracycline resistance gene has the recognition site for *Sma*I, while the other oligonucleotide has a site for *Bpm*I, an endonuclease that cleaves downstream of its recognition site. After the LSM cassette is ligated into a linearized plasmid carrying the DNA sequence to be mutagenized, the DNA is digested with *Sma*I and *Bpm*I, followed by ligation to recircularize the plasmid. This results in a replacement of 12 bp of the original DNA with 12 bp from the LSM cassette.

We have used this protocol to analyze the translation of the *rpoB* gene which encodes the β subunit of *Escherichia coli* RNA polymerase. The *rpoB* gene is part of a complex transcriptional unit that also includes the downstream gene for the β′ subunit of RNA polymerase (*rpoC*) and four upstream ribosomal protein genes (*rplKAJL*) which encode L11, L1, L10 and L7/12 (2). We have shown previously that the translation of both *rpoB* and *rpoC* mRNA is reduced in response to increased intracellular concentrations of either β or β′ (3,4). Preliminary deletion analysis suggested that sequences extending considerably upstream and downstream of the *rpoB* start site were required for normal translation (5) and that sequences extending even further into the structural gene were necessary for feedback regulation (6). Analysis of the mutations generated in the current study more precisely map the sequences important for efficient translation and indicates a potential site on the *rpoB* mRNA required for regulation by β′.

## MATERIALS AND METHODS

### Bacterial strains, plasmids and bacteriophages

*Escherichia coli* strain MG4 (7) is a Δ*lacU169 recA56* derivative of MG1655. The ββ′ expression plasmid, pACTBC, was

*To whom correspondence should be addressed. Tel: +1 519 661 3426; Fax: +1 519 661 3499; Email: tlinn@julian.uwo.ca

The authors wish it to be known that, in their opinion, the first two authors should be considered as joint First Authors

described previously (4). The β′ expression plasmid, pEXT21C1, was constructed by inserting the *rpoC* gene, isolated from pACTC1 (4) as a 5417 bp *Bam*HI–*Hin*dIII fragment, between these same sites in the multiple cloning site (MCS) of the low copy number expression vector pEXT21 (8). The plasmid source of the LSM cassette is pLS2. It was constructed by first inserting a synthetic oligonucleotide with the sequence CTGCACGTGCCCGGGTACGTACTGGAGGCAG into the *Pvu*II site of pUC19. This oligonucleotide contains recognition sequences for *Pml*I, *Sma*I, *Sna*BI and *Bpm*I. When inserted into the *Pvu*II site the sequences at the ends will reconstruct *Pvu*II sites which flank the original oligonucleotide. Next a 1359 bp fragment carrying the tetracycline resistance gene was excised from pBR322 by digestion with *Eco*RI and *Bsm*I, the termini were blunted with T4 DNA polymerase and the fragment was inserted into the *Sna*BI site of the oligonucleotide. The *Eco*RI, *Bsm*I and *Sna*BI recognition sequences are destroyed in this manipulation.

The pLK6 plasmid which carries the *rpoB* target for mutagenesis is based on pGEM7Zf(+) (Promega Corp., Madison, WI). The relevant modifications included insertion into the MCS of a 649 bp *Eco*RI–*Kpn*I *rpoB* fragment that was isolated from pEXBC3 (3). Secondly, the segment of the MCS between the *Sma*I and *Hin*dIII site was removed by digestion with these two enzymes, followed by blunting of the *Hin*dIII terminus with the Klenow enzyme and ligation to recircularize the plasmid. This treatment places the *Bam*HI site downstream of the *Kpn*I site in *rpoB* in the same translational reading frame as the *Bam*HI site in the *lacZ* region of λBC6. Finally, oligonucleotide mutagenesis was used to replace a single nucleotide within the β-lactamase gene in order to destroy the *Bpm*I recognition site normally present while retaining the wild-type amino acid sequence.

λBC6 carries a translational fusion of the wild-type *rpoB* gene to *lacZ* in a fusion vector based on λTLF95, a predecessor of λTLF97 which was described previously (9). The only difference between these vectors is that λTLF95 still has the *Bam*HI sites normally present in the tetracycline resistance and λ *int* genes, and carries four tandem copies of the *rrnBT1* terminator upstream of the MCS instead of just one copy. The *rpoB* region which extends from an engineered *Eco*RI site 245 bp upstream of the *rpoB* start site to the *Kpn*I site 405 bp into the gene was attached in frame to *lacZ* via the *Kpn*I site. Transcription of the *rpoB–lacZ* fusion is driven by a version of the bacteriophage P22 *ant* promoter carrying the 1052 mutation (10). It was inserted into the MCS upstream of the *rpoB* sequence as a 235 bp *Hin*dIII–*Sau*3A fragment isolated from pMS1052 (10).

## Linker scanning mutagenesis protocol

The LSM cassette was released from a CsCl gradient purified preparation of pLS2 by digestion with *Pvu*II. After agarose gel electrophoresis, the 1390 bp band corresponding to the LSM cassette was excised and the DNA extracted by adsorption to powdered glass (11). Most LS mutations presented in this study were produced by first nicking the pLK6 plasmid with pancreatic DNaseI in the presence of $Mg^{2+}$. A total of 870 µg of pLK6 was incubated with DNaseI ($3 \times 10^{-6}$ Kunitz U/µg DNA) for 30 min at 28°C in a buffer composed of 40 mM Tris–HCl (pH 7.5), 6 mM $MgCl_2$ and 0.1 mg/ml BSA. The treated DNA was electrophoresed on several preparative 1.4% agarose gels to resolve the open circular plasmid DNA from the linear and covalently closed circular forms. Approximately 30% of the plasmid was converted to the open circular form during DNaseI treatment. This band was excised from the gel and the DNA extracted by adsorption to powdered glass followed by ethanol precipitation.

Next, 53 µg of this nicked DNA was incubated with S1 nuclease (18 U/µg DNA) for 90 min at 18°C in a solution consisting of 50 mM $CH_3COONa$ (pH 5.7), 200 mM NaCl, 1 mM $ZnSO_4$ and 0.5% glycerol. The reaction was terminated by transfer to ice and the addition of 1/10 vol S1 stop buffer composed of 440 mM Tris–HCl (pH 8.6), 280 mM EDTA and 2.2% SDS. The treated DNA was electrophoresed on a preparative 1.4% agarose gel followed by excision of the linear band and extraction of the DNA by adsorption to powdered glass. Approximately 30% of the starting nicked DNA was converted to the linear form during this incubation. The extracted linear DNA was treated with calf intestinal alkaline phosphatase followed by spermine precipitation (12). A ligation reaction containing 3 µg of the linear pLK6, 0.8 µg of the LSM cassette (molar ratio 1:2) and 6 U of T4 DNA ligase was incubated at 18°C in a 50 µl volume of ligation cocktail I (66 mM Tris–HCl pH 7.5, 1 mM EDTA, 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol and 0.5 mM ATP).

Following ligation, 10 µl of the reaction mix was incubated with *Eco*RI and *Sty*I. The digestion products were resolved by electrophoresis on a 0.8% agarose gel and the 1858 bp band corresponding to the LSM cassette inserted into the original 468 bp *Eco*RI–*Sty*I *rpoB* fragment was excised and the DNA isolated with the Qiaex extraction kit (Qiagen Inc., Chatsworth, CA). This DNA fragment was added to ligation cocktail I containing 1 U of T4 DNA ligase and 2 µg of the *Eco*RI and *Sty*I cleaved plasmid backbone of pLK6 that had been treated with alkaline phosphatase. After incubation at 12°C, the ligation products were recovered by transformation of DH5α cells (10 µl ligation reaction per ml of competent cells) using the procedure of Inoue *et al.* (13). Transformants carrying the LSM cassette in the target region of pLK6 were selected on LB plates containing ampicillin (50 µg/ml) and tetracycline (20 µg/ml). The ~2300 clones recovered were pooled and plasmid DNA was extracted with the Wizard Midiprep DNA purification kit (Promega Corp., Madison, WI). The pooled plasmid DNA (5 µg) was digested with *Sma*I and *Bpm*I, followed by incubation with T4 DNA polymerase and all four deoxynucleoside triphosphates to blunt the termini generated by *Bpm*I. The treated DNA was then added to a 500 µl ligation reaction which was incubated at 18°C and contained 20 U T4 DNA ligase and 8% polyethylene glycol in ligation cocktail II (66 mM Tris–HCl pH 7.5, 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol and 0.1–mM ATP). These conditions were optimized to favor intramolecular blunt end ligation. The ligation products from 300 µl of the reaction were recovered by spermine precipitation and then digested with *Eco*RI and *Bam*HI. The digestion products were resolved by electrophoresis on a 1% agarose gel and the 649 bp band encompassing the start of the *rpoB* gene was excised and the DNA extracted with the Qiaex kit. The purified DNA fragment was ligated between the *Eco*RI generated right arm and *Bam*HI generated left arm of λBC6.

## Recovery and analysis of LS mutations in λBC6

The recombinant versions of λBC6 carrying the LS mutations were recovered from the ligation mixture by *in vitro* packaging (14) and plating on an *E.coli* C600 lawn. Individual plaques were picked and used to make plate lysates as described previously (15). Phage DNA was isolated from 1.7 ml of the plate lysate by

the zinc chloride method (16) followed by spermine precipitation and resuspension of the DNA in 35 µl TE buffer. The approximate position of the LS mutation was determined by digestion of 9 µl phage DNA with *Pml*I and *Bam*HI, and comparison of the resultant fragment's size to standards resolved by polyacrylamide gel electrophoresis. The exact sequence of the mutation was determined using the fmol DNA Sequencing System (Promega Corp., Madison, WI) to analyze 9 µl of the phage DNA.

### Lysogen formation and β-galactosidase assays

Monolysogens of the recombinant phage were recovered in MG4 carrying either pACTBC or pEXT21C1by the protocol described previously (15). Because λBC6 is based on the new λTLF97 vector (9), which contains the tetracycline resistance gene, lysogens were directly selected on LB plates containing 5 µg/ml tetracycline. The plates also contained Xgal to provide an estimate of the fusion's level of β-galactosidase expression. The growth of lysogen cultures and their assay for β-galactosidase levels was as described previously (4).

### Measurement of *rpoB–lacZ* mRNA levels

The amount of *rpoB–lacZ* fusion mRNA was determined with the Lysate Ribonuclease Protection Kit (Amersham Life Sciences Inc., Arlington Heights, IL). The plasmid template (pRNA2F) used for synthesis of the labeled RNA probe was constructed by first isolating a 427 bp DNA fragment originating from λBC6 that extends from a *Sty*I site 221 bp downstream of the start site of *rpoB* to a *Bsu*36I site in *lacZ*. The termini were blunted by treatment with Klenow DNA polymerase and the fragment was inserted into the *Hinc*II site of pGEM3 (Promega Corp., Madison, WI). Cleavage of this plasmid DNA with *Hin*dIII and transcription from the T7 promoter produces an ~490 nt RNA complementary to the *rpoB–lacZ* mRNA made from the λBC6 derivatives. Each monolysogen (λBC6 derivative in MG4 containing pACTBC) to be analyzed was grown at 37°C in LB medium containing 20 µg/ml chloramphenicol. When an $A_{600}$ of 0.4 was reached, 1.4 ml of the culture was centrifuged at 14 000 *g* for 4 min at 4°C and the cell pellet resuspended in 100 µl of lysis solution. Approximately $2 \times 10^6$ c.p.m. of the labeled probe was added to 45 µl of cell lysate, followed by hybridization and processing according to the Lysate Ribonuclease Protection Kit protocol. After polyacrylamide gel electrophoresis the hybrid bands corresponding to the fusion *rpoB–lacZ* mRNA and the cellular *rpoB* mRNA were visualized by exposure of Dupont Cronex 4 film and quantified by analyzing the autoradiographs using a CCD camera and Gel Print Tool Box software version 3.0 (Biophotonics Corp.)

## RESULTS

### Linker scanning mutagenesis protocol

The protocol described is capable of simultaneously producing hundreds of linker scanning mutations. The basic strategy as outlined in Figure 1 makes use of a LSM cassette. The LSM cassette contains a selectable marker, in this case the tetracycline resistance gene isolated from pBR322. It is flanked by two synthetic oligonucleotides that carry the recognition sequences for several restriction endonucleases. The LSM cassette is first inserted into a double stranded DNA break by ligation then subsequently cleaved by digestion with *Sma*I and *Bpm*I. The cleavage sites for *Bpm*I are 16 nt downstream of the recognition

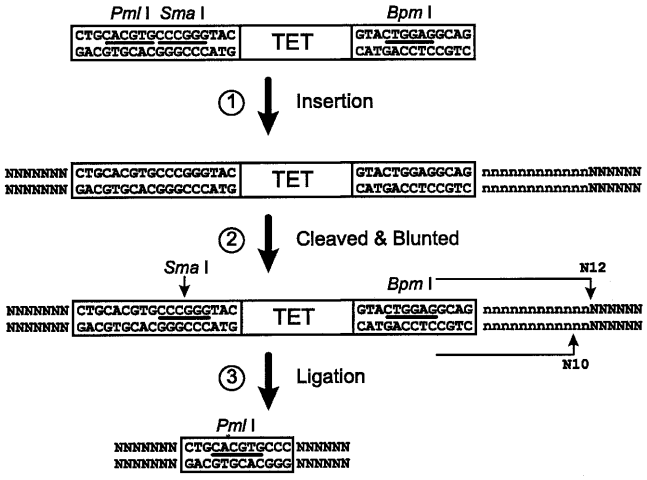



**Figure 1.** Basic strategy for generating linker scanning mutations. 1. The LSM cassette is inserted into a double stranded break in a plasmid by ligation. 2. The resulting product is cleaved with *Sma*I and *Bpm*I, followed by treatment with T4 DNA polymerase to blunt the staggered terminus produced by *Bpm*I. 3. The plasmid is recircularized by ligation, which results in 12 bp of the original sequence (lower case n) being replaced with the 12 bp LS oligonucleotide.

site on the 5′→3′ strand and 14 nt downstream on the 3′→5′ strand. Because the *Bpm*I site is positioned 4 bp from the end of the LSM cassette the cleavages will be located outside of the cassette at 12 and 10 nt, respectively, into the flanking DNA. The 3′→5′ exonuclease activity of T4 DNA polymerase is then used to remove the two 3′ protruding nucleotides and generate a flush terminus 12 nt downstream of the original LSM cassette insertion site. *Sma*I cleaves at a position 12 bp inside the terminus of the oligonucleotide on the opposite side of the tetracycline resistance gene from the *Bpm*I site. Recircularization of the plasmid DNA by ligation results in replacement of the 12 bp of DNA that originally flanked the *Bpm*I end of the LSM cassette by 12 bp from the other end of the cassette. The 12 bp linker that is left behind is marked with a *Pml*I recognition site.

The overall protocol that was developed to generate a library of LS mutations is outlined in Figure 2. First, a plasmid containing the segment of DNA to be mutagenized was randomly cleaved with pancreatic DNaseI. For the analysis of *rpoB* translation, a fragment extending from 245 bp upstream to 405 bp downstream of the *rpoB* start site was inserted into a vector based on pGEM7Zf(+) (see Materials and Methods). Two different approaches were used to linearize the plasmid DNA. Initially the plasmid was digested with a limiting amount of DNaseI in the presence of $Mn^{2+}$ to favor double stranded cuts. The reaction was titrated so that ~40% of the covalently closed circular DNA was converted to the linear form which was then isolated by agarose gel electrophoresis. But DNA sequence analysis of the LS mutations that were ultimately generated by this procedure showed that most had suffered an additional deletion in the range of 1–50 bp adjacent to the linker. Previously, a careful analysis of the action of DNaseI in the presence of different divalent metal ions indicated that, although the enzyme produces linear molecules with $Mn^{2+}$, the cleavages are not necessarily in the same position on both strands (17). Moreover, additional cuts are often made in the vicinity of the original cleavage to generate gapped or deleted DNA molecules. These properties of DNaseI with $Mn^{2+}$

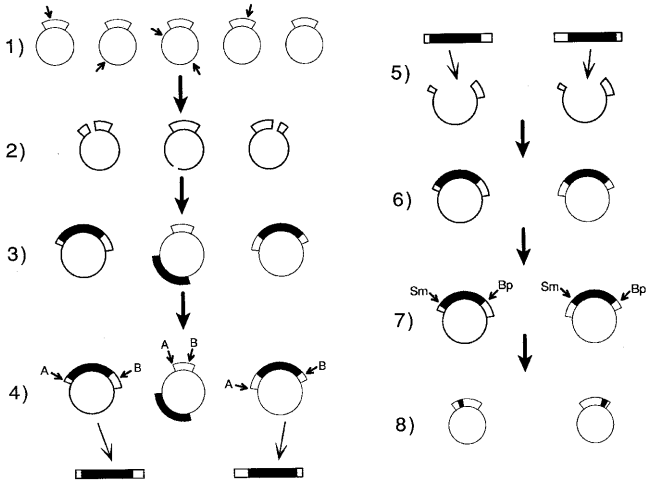



**Figure 2.** Overview of the linker scanning mutagenesis protocol using the LSM cassette. The plasmid vector is indicated with a single line, the insert DNA with an open box, and the LSM cassette with the filled-in box. 1. Production of randomly cleaved linear plasmid DNA by either (i) limited DNaseI digestion in the presence of $Mn^{2+}$, or (ii) partial DNaseI nicking in the presence of $Mg^{2+}$ followed by gel purification of open circular plasmid DNA and limited digestion with S1 nuclease. 2. Purification of linear plasmid DNA by agarose gel electrophoresis. 3. Insertion of the 1390 bp LSM cassette by ligation. 4. Isolation of insertions in the target DNA by digestion with restriction endonucleases A and B, followed by agarose gel purification of the A–B plus 1390 bp DNA fragment. 5. Reinsertion of the target region containing the LSM cassette back into the wild-type plasmid backbone cleaved with A and B. 6. Recovery of a library of LSM cassette insertions confined to the target DNA by transformation with selection of colonies resistant to both ampicillin and tetracycline. 7. Digestion of the pooled library DNA with *Sma*I and *Bpm*I. 8. Recircularization of the plasmid DNA by ligation to generate the final 12 bp LS mutations.

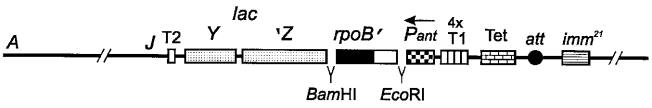



**Figure 3.** Structure of the *rpoB–lacZ* translation fusion phage λBC6. The 649 bp *Eco*RI–*Bam*HI fragment that contains the first 405 bp of *rpoB* was ligated between the *Bam*HI generated left arm of λTLF95 and the *Eco*RI generated right arm of λTLF95 which has been modified with the addition of the P22 *ant*1052 promoter (*Pant*) to drive transcription of the *rpoB–lacZ* fusion. The prime symbols (′) in *lac′Z* and *rpoB′* indicate, respectively, the 5′ and 3′ truncations of these genes. The position of the *rrnBT2* transcriptional terminator (T2), four tandem copies of the *rrnBT1* transcriptional terminator (4× T1) and tetracycline resistance gene (Tet) are indicated. The phage vector attachment site (*att*), immunity region (*imm21*), and representative structural genes in the left arm (*A* and *J*) are also illustrated.

could explain the high proportion of deletions flanking the LS oligonucleotide. Although these deletions marked with the linker can still be useful in a mutational analysis, they do not satisfy the definition of a true linker scanning mutation.

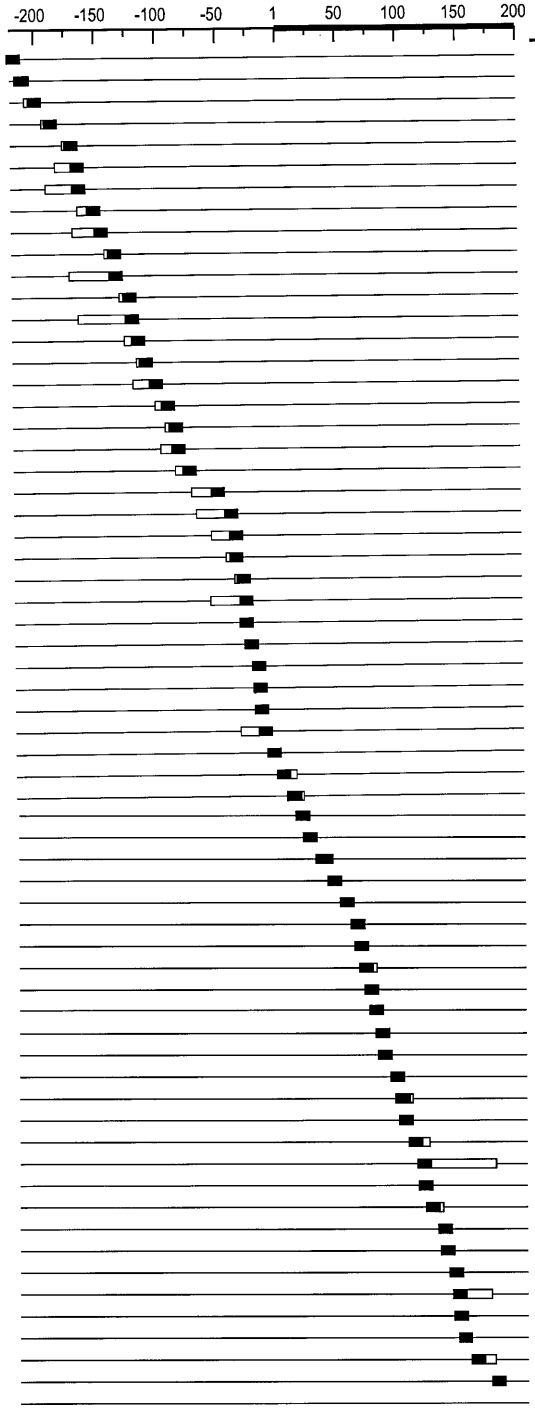
The second approach used to make flush double stranded breaks in the DNA was to first nick one strand with DNaseI in the presence of $Mg^{2+}$, followed by cleavage of the strand opposite the nick with S1 nuclease. In order to increase the probability of just a single nick per DNA molecule, the DNaseI digestion was titrated so that ~30% of the covalently closed circular plasmid DNA was converted to the open circular form. The two populations were resolved by agarose gel electrophoresis. After extraction, the nicked DNA was digested with S1 nuclease so that ~30% was converted to the linear form. The linear molecules were purified by agarose gel electrophoresis and used in the subsequent steps. Sequence analysis of the LS mutations generated by this protocol showed that the majority (>60%) were direct linker replacements with no additional deleted nucleotides. In those cases where deletions flanked the linker, the deletions were generally in the range of 1–6 nt, much smaller than those generated with DNaseI in the presence of $Mn^{2+}$. Presumably the small deletions are the result of S1 nuclease action at the transiently melting ends of the linear DNA molecules.

The linear plasmid DNA was treated with alkaline phosphatase then mixed with the 1390 bp LSM cassette and ligated. The ligated DNA is then digested with two restriction endonucleases that cleave uniquely in the insert (indicated as A and B in Figure 2, which in the case of *rpoB* represent *Eco*RI and *Sty*I, whose cleavage sites are 245 bp upstream and 223 bp downstream of the translation start site). The position of these cleavage sites determine the region of insert DNA that will be further analyzed (target DNA). Any target DNA into which the LSM cassette has been ligated will increase in size by 1390 bp. The digestion products are resolved by agarose gel electrophoresis and the target plus LSM cassette band is excised and the DNA extracted. In the case of *rpoB*, the 468 bp fragment containing a LSM cassette increases in size to 1858 bp. Alternatively, the original ligation products can be used to directly transform a competent *E.coli* strain with the selection of ampicillin and tetracycline resistant transformants. The resultant colonies will be a library of clones with the LSM cassette inserted anywhere into the plasmid. DNA prepared from the library is digested with endonucleases A and B and the target plus 1390 bp LSM cassette band purified as described above.

Whether the target plus LSM cassette DNA fragments are isolated directly from the original ligation or from a library of transformants, they are then ligated back into the original wild-type plasmid backbone that has also been digested with the A and B restriction endonucleases. When the ligation products are introduced into *E.coli*, the resulting transformants that are both ampicillin and tetracycline resistant will represent a library of LSM cassette insertions confined to the target DNA. Plasmid DNA prepared from the pooled library is then digested with *Sma*I and *Bpm*I to excise all of the LSM cassette except the 12 bp oligonucleotide adjacent to the *Sma*I site. The plasmid DNA is recircularized by ligation to generate the final linker scanning mutations.

How the phenotype of the resultant LS mutations is analyzed will depend on the particular system. If the plasmid carries all the sequences required for expression of a functional product then the ligated plasmid DNA can be transformed directly into the appropriate strain to generate a library of LS mutants for analysis. Because expression of *rpoB* is subject to gene dosage artifacts, the effects of the LS mutations had to be analyzed in a single copy number vector. Therefore, the recircularized plasmid DNA was digested with *Eco*RI and *Bam*HI to yield a 659 bp fragment which was purified by agarose gel electrophoresis and then ligated between the *Bam*HI-generated left arm and *Eco*RI-generated right arm of λBC6 (Fig. 3). Attachment of the fragment to the left arm produces a translational fusion of *rpoB* to *lacZ*, while the right arm provides a version of the *ant* promoter (10) to drive transcription of the fusion. *In vitro* packaging produces a library

**Table 1.** Effect of linker scanning mutations on translation and translational repression of *rpoB*




| Mutation | Position[1] | Translation (%)[2] | Feedback[3] | β-gal[4] | β-gal[5] (Excess ββ′) |
|---|---|---|---|---|---|
| LS142 | -222 to -211 | 86 | 6.3 | 801.7 ± 6.2 | 127.3 ± 10.0 |
| LSS114 | -216 to -204 | 66 | 5.6 | 609.9 ± 1.0 | 108.4 ± 2.6 |
| LS117 | -208 to -194 | 20 | 7.6 | 183.0 ± 7.8 | 24.1 ± 18.2 |
| LSS62 | -194 to -181 | 42 | 6.0 | 393.0 ± 3.2 | 65.2 ± 4.0 |
| LSS64 | -176 to -164 | 96 | 5.2 | 889.3 ± 4.6 | 171.6 ± 8.0 |
| LSS65 | -183 to -159 | 126 | 5.1 | 1177.2 ± 0.3 | 229.6 ± 4.2 |
| LSS46 | -191 to -158 | 78 | 5.5 | 721.3 ± 2.5 | 131.6 ± 2.5 |
| LSS57 | -165 to -146 | 72 | 5.2 | 666.5 ± 1.5 | 127.3 ± 6.4 |
| LSS53 | -169 to -140 | 97 | 4.9 | 905.9 ± 1.4 | 186.1 ± 3.8 |
| LS90 | -141 to -129 | 96 | 5.3 | 894.3 ± 9.7 | 167.3 ± 2.7 |
| LSS115 | -172 to -128 | 109 | 4.7 | 1015.5 ± 3.1 | 216.0 ± 3.4 |
| LSS113 | -131 to -117 | 80 | 4.7 | 746.3 ± 4.2 | 160.3 ± 24.2 |
| LSS51 | -165 to -115 | 140 | 5.0 | 1298.3 ± 1.1 | 261.4 ± 2.1 |
| LSS48 | -127 to -110 | 79 | 5.4 | 730.9 ± 1.7 | 135.6 ± 4.1 |
| LSS47 | -117 to -104 | 55 | 5.2 | 510.9 ± 1.5 | 98.2 ± 5.7 |
| LSS113 | -120 to -96 | 81 | 5.2 | 749.7 ± 3.0 | 144.1 ± 7.3 |
| LS104 | -102 to -86 | 50 | 6.1 | 463.5 ± 6.4 | 76.4 ± 9.6 |
| LS129 | -94 to -80 | 113 | 5.6 | 1048.8 ± 3.3 | 186.5 ± 1.3 |
| LS100 | -98 to -78 | 74 | 6.1 | 686.2 ± 2.4 | 113.0 ± 7.0 |
| LS160 | -86 to -69 | 17 | 4.0 | 156.3 ± 1.6 | 38.8 ± 1.5 |
| LSS123 | -73 to -46 | 26 | 4.6 | 239.7 ± 2.0 | 51.6 ± 3.7 |
| LSS34 | -69 to -35 | 23 | 4.7 | 211.6 ± 3.6 | 45.2 ± 7.9 |
| LSS54 | -57 to -31 | 15 | 5.2 | 143.8 ± 4.4 | 27.8 ± 1.3 |
| LSS58 | -45 to -31 | 15 | 5.3 | 136.9 ± 2.9 | 26.1 ± 3.4 |
| LSS116 | -37 to -25 | 2 | 5.8 | 18.9 ± 4.5 | 3.3 ± 11.3 |
| LSS31 | -58 to -23 | 7 | 6.7 | 67.5 ± 9.9 | 10.1 ± 2.9 |
| LS122 | -34 to -23 | 1 | 7.9 | 11.9 ± 3.8 | 1.5 ± 14.4 |
| LSS119 | -30 to -19 | 1 | 4.8 | 6.3 ± 16.2 | 1.3 ± 12.6 |
| LS62 | -24 to -13 | 2 | 8.4 | 19.4 ± 12.5 | 2.3 ± 10.6 |
| LSS122 | -23 to -12 | 1 | 5.2 | 5.7 ± 12.4 | 1.1 ± 0.0 |
| LS147 | -22 to -11 | 1 | 7.0 | 10.9 ± 7.3 | 1.6 ± 8.0 |
| LSS32 | -34 to -8 | 0.5 | 4.5 | 0.6 ± 28.1 | 0.1 ± 7.6 |
| LSS702 | -12 to -1 | 12 | 6.8 | 113.7 ± 3.5 | 16.8 ± 13.4 |
| LS135 | -4 to 12 | 16 | 6.5 | 149.4 ± 10.2 | 23.1 ± 13.2 |
| LSS701 | 4 to 18 | 1 | 4.5 | 6.6 ± 10.1 | 1.5 ± 12.9 |
| LSS709 | 10 to 21 | 1 | 5.3 | 13.0 ± 7.5 | 2.5 ± 5.1 |
| LSS843 | 16 to 27 | 0.1 | 5.3 | 0.7 ± 7.5 | 0.1 ± 15.0 |
| LSS44 | 26 to 40 | 9 | 7.2 | 87.9 ± 6.1 | 12.2 ± 17.2 |
| LSS121 | 36 to 47 | 14 | 4.8 | 126.7 ± 2.8 | 19.1 ± 6.1 |
| LSS744 | 46 to 57 | 31 | 6.0 | 289.6 ± 2.3 | 48.3 ± 5.4 |
| LSS718 | 55 to 66 | 37 | 6.4 | 341.2 ± 2.5 | 53.2 ± 2.5 |
| LSS798 | 58 to 69 | 16 | 6.6 | 149.7 ± 1.9 | 22.7 ± 12.7 |
| LSS706 | 62 to 76 | 48 | 6.1 | 446.9 ± 3.3 | 73.1 ± 2.3 |
| LSS809 | 66 to 77 | 3 | 5.1 | 26.9 ± 12.2 | 5.2 ± 20.4 |
| LSS740 | 70 to 81 | 58 | 5.5 | 538.1 ± 3.4 | 97.6 ± 7.0 |
| LS116 | 75 to 86 | 57 | 5.9 | 534.1 ± 7.4 | 90.1 ± 7.0 |
| LSS713 | 77 to 88 | 55 | 6.1 | 513.0 ± 4.8 | 83.9 ± 8.7 |
| LSS719 | 87 to 98 | 61 | 6.2 | 569.1 ± 7.3 | 91.6 ± 10.9 |
| LSS716 | 91 to 105 | 82 | 6.4 | 764.3 ± 5.6 | 119.9 ± 2.0 |
| LS55 | 94 to 105 | 21 | 6.0 | 192.3 ± 3.6 | 32.3 ± 10.6 |
| LSS737 | 102 to 119 | 28 | 6.2 | 256.8 ± 5.6 | 41.4 ± 0.7 |
| LS115 | 109 to 174 | 20 | 6.3 | 186.2 ± 5.1 | 29.4 ± 4.7 |
| LSS736 | 110 to 121 | 43 | 6.9 | 403.9 ± 3.2 | 58.3 ± 1.4 |
| LSS734 | 116 to 130 | 40 | 8.3 | 370.0 ± 7.3 | 44.5 ± 9.0 |
| LSS720 | 126 to 137 | 47 | 6.8 | 435.1 ± 1.3 | 63.6 ± 5.3 |
| LSS735 | 128 to 139 | 68 | 6.1 | 633.4 ± 7.6 | 104.6 ± 10.6 |
| LSS712 | 135 to 146 | 82 | 7.3 | 759.1 ± 5.6 | 104.7 ± 4.0 |
| LS128 | 138 to 170 | 100 | 6.9 | 931.5 ± 2.3 | 134.3 ± 3.2 |
| LSS704 | 139 to 150 | 78 | 7.0 | 724.3 ± 4.0 | 103.8 ± 4.5 |
| LSS714 | 142 to 153 | 84 | 7.2 | 780.6 ± 1.3 | 108.4 ± 5.8 |
| LSS717 | 153 to 173 | 109 | 7.2 | 1012.0 ± 7.0 | 140.0 ± 3.6 |
| LSS722 | 170 to 181 | 109 | 6.6 | 1011.0 ± 6.0 | 152.5 ± 1.0 |
| None | ----- | 100 | 5.0 | 930.7 ± 4.9 | 184.3 ± 6.2 |

[1]Position in the wild-type sequence that has been replaced by the LS oligonucleotide (and deleted in those mutations that have an accompanying deletion). Numbering relative to the *rpoB* start site (+1). The LS oligonucleotide is represented as a filled-in box, and the deletions with an open box.

[2]Translation in percent was calculated as 100 × (β-gal units for mutation/β-gal units for wild-type *rpoB*–lacZ fusion).

[3]Translational feedback was calculated as (β-gal units in the absence of IPTG/β-gal units in the presence of 1 mM IPTG).

[4]β-galactosidase activities were determined in the absence of IPTG as described in the Materials and Methods; all values are the averages of at least six assays performed on three or more independent cultures. The percent standard deviation is presented with each average.

[5]β-galactosidase activities as described in footnote 4, except that they were determined in the presence of 1 mM IPTG to induce β and β′ production from pACTBC.

of phage clones each containing an LS mutation in the *rpoB* segment of the *rpoB–lacZ* fusion. After plating on a bacterial lawn, individual plaques were used to produce a plate lysate of each phage clone (see Materials and Methods). A DNA preparation made from each plate lysate was used for restriction site analysis and DNA sequencing.

The approximate position of the LS mutation can be readily mapped by digestion with *Pml*I, which cleaves uniquely in the LS oligonucleotide, and another restriction endonuclease that cleaves in the flanking DNA (e.g. enzymes A or B in Fig. 2). The size of the resulting DNA fragment determined on polyacrylamide gels provides a measure of the distance of the LS mutation from the restriction site reference point. Such restriction site mapping is useful for rapidly screening clones to identify mutations in a particular region of the target DNA. The exact position of the mutations was then determined by thermal DNA cycle sequencing of the *rpoB* target region.

## Analysis of sequences important for efficient translation of *rpoB*

Sixty two mutations which provide continuous coverage of 400 bp extending from ~220 bp upstream to 180 bp downstream of the *rpoB* start site were chosen for further analysis. The positions of these mutations, which include both true linker scanning mutations and LS oligonucleotide insertions with accompanying deletions, are summarized in Table 1. Lysogens of the recombinant λ phage carrying a mutation in the *rpoB–lacZ* fusion were made in *E.coli* strain MG4 which also contained pACTBC. β-galactosidase assays on exponentially growing cultures of these lysogens provide a measure of *rpoB* translation in each mutant (Table 1). Many mutations were found to affect translation of the *rpoB–lacZ* fusion, but an overall profile emerges from the map of translational efficiency versus mutation position that is presented in Figure 4. Starting from the 5′ end there appears to be a significant drop in translation when sequences centered around –200 to –190 are altered. Mutating sequences between approximately –180 and –120 has a variable but lesser effect with translation levels ranging from an increase of 40% to a decrease of 30%. Most mutations in the range of –120 to –80 decrease translation from a modest 20 to 50%, with the exception of LS129 which increases it slightly. However, as the 3′ end of the mutation gets closer than 70 nt upstream of *rpoB*, the translation level drops dramatically. Five mutations that have the 3′ end of the mutation in the –70 to –30 interval have translation levels in the range of 26–15% of the wild-type. The region spanning –25 to +16 appears to be critical, since 10 out of 13 mutations in this area reduce translation to 2% or less. Then as the 5′ end of the mutations move 16–91 nt downstream of the *rpoB* start site there is a general increase in translation ranging from 9 to 82% of the wild-type (with the exception of LSS809). This is followed by a drop in translation for three mutations that have 5′ and/or 3′ endpoints between +94 and +109. As the 5′ end of the mutations move from +110 to +138, there again is a general increase in translation to normal levels. Finally six mutations with a 5′ end mapping in the interval from +135 to +170 have translation levels which vary from the wild-type by only 22% or less.

## Stability of the mutated *rpoB–lacZ* mRNA

Although the assumption is that the primary effect of these mutations is on the translation efficiency of *rpoB*, some could have their main influence on mRNA stability which in turn would affect the level of the β subunit–β-galactosidase fusion protein. The RNA level of 14 selected mutants, which had translation levels ranging from 113 to 1%, were directly measured by the RNase protection assay (see Materials and Methods). The 490 nt RNA probe extends from a *Sty*I site 221 nt downstream of the *rpoB* start, through the *rpoB–lacZ* junction into *lacZ*. This probe sequence begins just downstream of the target region for LS mutagenesis, so an identical hybrid will be formed with all mutated RNAs. The RNA from each cellular extract will form two hybrids with this probe. One, with the *rpoB–lacZ* junction RNA from the fusion will be 427 bp, while a smaller hybrid of 182 bp will form with the *rpoB* RNA produced from both the chromosomal copy of *rpoB* and any leaky expression from the repressed copy of *rpoB* on pACTBC. This smaller hybrid provides a convenient internal standard which was used to normalize the levels of *rpoB–lacZ* RNA measured in the mutants.

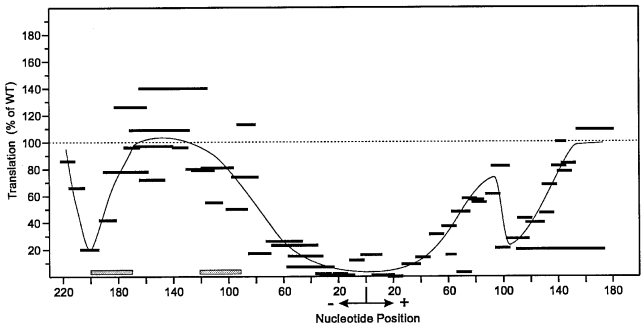



**Figure 4.** Map of sequences important for translation of *rpoB*. For each mutation (indicated with a thick black line) the level of translation of the *rpoB–lacZ* fusion listed in Table 1 is plotted versus the mutation position. The nucleotide position is with respect to the A (+1) of the start codon. The curve drawn through the points is based on a 9th degree parabolic regression conducted by the Fig. P software (Biosoft). The shaded boxes indicate the position of the two regions that form the proposed stem of the RNaseIII processing site.

**Table 2.** *rpoB–lacZ* RNA levels

| Mutation | Translation | mRNA amount[a] |
|---|---|---|
| none (BC6) | 100 | 100 |
| LS129 | 113 | 85 |
| LS128 | 100 | 101 |
| LS90 | 96 | 98 |
| LS142 | 86 | 93 |
| LS100 | 74 | 86 |
| LS116 | 57 | 92 |
| LS104 | 50 | 101 |
| LS55 | 21 | 58 |
| LS117 | 20 | 70 |
| LS115 | 20 | 55 |
| LS160 | 17 | 50 |
| LS135 | 16 | 71 |
| LS62 | 2 | 37 |
| LS122 | 1 | 35 |

[a]Average of two determinations as described in Materials and Methods.

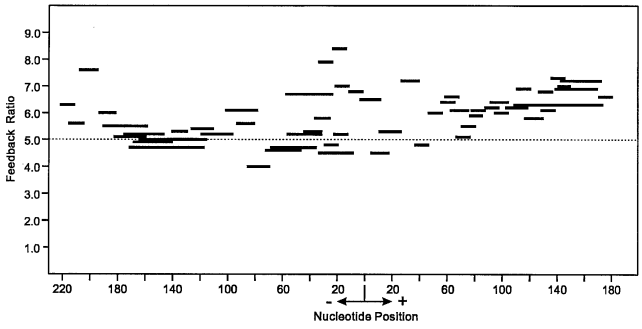



**Figure 5.** Effect of mutations on the translational repression exerted on *rpoB* by increased concentrations of both β and β′. For each mutation (indicated with a thick black line) the feedback ratio listed in Table 1 is plotted versus the mutation position. The dotted line indicates the feedback ratio for the wild-type *rpoB* sequence (5.0) after induction of β and β′ production from pACTBC.

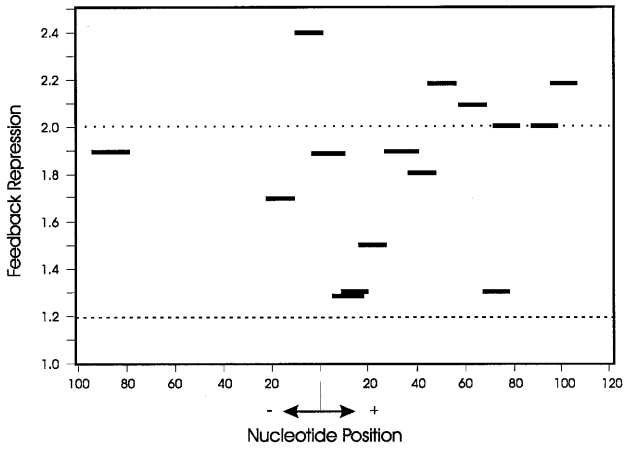



**Figure 6.** Effect of mutations on the translational repression exerted on *rpoB* by increased concentrations of β′. For each mutation (indicated with a thick black line) the feedback ratio listed in Table 3 is plotted versus the mutation position. The upper dotted line indicates the feedback ratio for the wild-type *rpoB* sequence (2.0) after induction of β′ production from pEXT21C1. The lower broken line indicates the feedback ratio for a control fusion of an intact *lacZ* gene attached downstream of the *ant*1052 promoter.

The 14 mutations analyzed fall into three groups with respect to translation level and mRNA amounts (Table 2). Seven of the mutations in which translation levels ranged from 113 to 50% had mRNA amounts that did not vary significantly from the wild-type *rpoB–lacZ* fusion. The next group of five mutations that had translation levels in the range of 21–16% showed mRNA levels that decreased to 71–58% of the wild-type. Finally, two mutations that drastically decreased translation to 1–2% of the wild-type showed a drop in mRNA levels to about one third (35 and 37%) the normal level. So both groups that showed a substantial reduction in translation of the β subunit–β-galactosidase fusion protein also showed a reduction in the level of the *rpoB–lacZ* junction mRNA. However, the levels of RNA were still considerably higher than the level of translation (3–4-fold higher for the intermediate group, and 19- and 35-fold higher for the most severe mutations).

### Analysis of sequences required for autogenous regulation of *rpoB* translation

The current collection of mutations was analyzed in an attempt to more precisely define the sequences essential for feedback regulation of *rpoB* translation. As described above, lysogens of each recombinant phage were made in an *E.coli* strain which also contains the ββ′ expression plasmid pACTBC (4). The *rpoBC* genes in this plasmid are transcribed from the *tac* promoter which is regulated by the *lac* repressor. In the absence of IPTG the β-galactosidase activity produced from the *rpoB–lacZ* fusion will reflect the level of translation with normal concentrations of β and β′, while in the presence of inducer the activity will represent the translation level of *rpoB–lacZ* in the presence of an excess of both β and β′. The ratio of these two levels of *rpoB* translation is presented as the feedback ratio in Table 1. Under these conditions the wild-type fusion (λBC6) shows a 5-fold reduction in translation (feedback ratio of 5.0). The feedback ratio is plotted versus mutation position in Figure 5, but, surprisingly, none of the mutations show a significant decrease in feedback under these conditions.

We demonstrated previously that not only was the combination of β and β′ an effective regulator of *rpoB* translation, but that excess levels of each individual subunit could also cause repression of *rpoB* translation (4). Therefore, if each subunit interacted with a different region on the *rpoB* mRNA it would be possible that even if a given mutation eliminated the repression site for one subunit, the other subunit could still interact with its site and the feedback ratio would not decrease appreciably. To directly test this, a series of recombinant phage carrying selected LS mutations were used to lysogenize MG4 containing pEXT21C1. This low copy number expression plasmid overproduces only the β′ subunit under the control of the *tac* promoter and *lac* repressor. Although higher copy number plasmids that express β′ alone can achieve feedback ratios as high as 15, this high a level of expression is lethal to the cell, presumably as the result of limiting expression of the chromosomal copy of *rpoB* (4). The level of β′ overproduction from pEXT21C1 with 10 μM IPTG gives a lower feedback ratio of 2.0 for the wild-type *rpoB–lacZ* fusion but is not detrimental to cell growth (Table 3, Fig. 6). There is a significant loss of repression by two partially overlapping mutations, LSS701 (+4 to +18) and LSS709 (+10 to +21), which both have a feedback ratio of 1.3. The partially overlapping LSS843 (+16 to +27) has an intermediate feedback ratio of 1.5. Seven of the eight LS mutations that map downstream of LSS843 had near normal repression with the exception of LSS809 (+66 to +77) which had a feedback ratio of 1.3. The overproduction of β′ from an expression plasmid has been shown previously to exert a small general reduction on the expression of other genes (4). To quantify this non-specific effect under these conditions a fusion of the intact *lacZ* gene downstream of the *ant*1052 promoter was assayed. The expression of this control fusion, which has the same reporter gene and promoter as the *rpoB–lacZ* fusions but without the specific *rpoB* sequences, was reduced 1.2-fold in the presence of β′. Therefore, the ratio of 1.3 seen with some of the mutations represents almost a complete loss of repression of *rpoB* translation.

## DISCUSSION

The analysis of cloned genes and their regulatory sequences by *in vitro* mutagenesis generally proceeds from a primary analysis with simple deletions to finally generating single base substitutions. A variety of methods can be used to produce deletions that extend into the region of interest from both the upstream and downstream

directions (18). This initial deletion analysis can define the borders of important regulatory elements and point the way to more precise mapping of critical sequences. Ultimately, the contribution of individual nucleotides can be examined by constructing single base substitutions with techniques such as oligonucleotide directed mutagenesis (19). But in most cases the region delimited by the deletion mutants is generally too large to analyze efficiently by single base substitutions. Surveying this region with a collection of 'clustered point mutations', that is, changes in the range of 8–12 adjacent base pairs at a time, is an effective intermediate step. This can be achieved by generating linker-scanning mutations. The original procedure (1) and even a later streamlined version (20) devised to construct such mutants are laborious. However using the protocol described here, regions of DNA on the order of 100–500 bp can be readily saturated with LS mutations.

The most problematic step of the overall protocol is the efficient conversion of circular plasmid DNA to full length linear molecules with blunt termini. Digestion with DNaseI in the presence of $Mn^{2+}$ can produce such linear molecules (21). However we found in agreement with earlier work (17) that, even when the limited digestion is carefully titrated, a significant fraction of the molecules undergo secondary cleavage with the loss of nucleotides adjacent to the initial cleavage site, and as a result the majority of mutations had a deletion flanking the LS oligonucleotide. We found a much higher proportion of full length linear molecules was produced by first nicking the circular plasmid DNA with DNaseI in the presence of $Mg^{2+}$ and then linearizing the nicked DNA with S1 nuclease, although the S1 nuclease digestion must be carefully controlled to limit nibbling at the termini of the linear DNA.

If the resultant library of LS mutations does contain a substantial fraction with accompanying deletions many of those can be eliminated by excision of the target sequence from the pooled plasmid DNA with the appropriate restriction enzymes followed by separation on high resolution polyacrylamide gels. Extraction of the full length target band and ligation into the plasmid backbone would produce a second library containing a higher proportion of LS mutations with no deletions. The efficiency of this additional step will depend on the size of the target sequence and the resolving power of the gel system. Although larger deletions should be eliminated, smaller deletions of several nucleotides would probably not be resolved from the full length fragment and thus would still be present in the final library.

Any other procedure that productively converts closed circular DNA to full length linear molecules would greatly increase the ease and efficiency of the overall protocol. One such approach is to linearize the DNA by partial digestion using a variety of restriction enzymes that cleave DNA frequently and generate blunt ends. Although this may produce useful mutations, it would not saturate the target sequence since the distribution of the LS mutations would be dictated by the location of the restriction sites.

**Table 3.** Effect of linker scanning mutations on the translational repression of *rpoB* by β′

| Mutation[a] | Position of mutation[b] | Feedback[c] | β-gal[d] | β-gal (excess β′)[e] |
|---|---|---|---|---|
| BC6 | none | 2.0 | 1185 ± 3.0 | 597 ± 7.8 |
| LS129 | –94 to –80 | 1.9 | 1189 ± 4.5 | 627 ± 3.5 |
| LSS122 | –23 to –12 | 1.7 | 19 ± 8.4 | 11 ± 3.7 |
| LSS702 | –12 to –1 | 2.4 | 176 ± 2.3 | 73 ± 5.6 |
| LS135 | –4 to 12 | 1.9 | 156 ± 4.5 | 82 ± 2.9 |
| LSS701 | 4 to 18 | 1.3 | 14 ± 5.5 | 11 ± 3.8 |
| LSS709 | 10 to 21 | 1.3 | 18 ± 6.1 | 14 ± 9.2 |
| LSS843 | 16 to 27 | 1.5 | 0.9 ± 2.3 | 0.6 ± 4.9 |
| LSS44 | 26 to 40 | 1.9 | 105 ± 4.0 | 55 ± 1.3 |
| LSS121 | 36 to 47 | 1.8 | 182 ± 4.9 | 102 ± 1.0 |
| LSS744 | 46 to 57 | 2.2 | 357 ± 4.8 | 160 ± 4.9 |
| LSS798 | 57 to 68 | 2.1 | 196 ± 5.0 | 92 ± 1.5 |
| LSS809 | 66 to 77 | 1.3 | 30 ± 2.5 | 24 ± 1.5 |
| LSS740 | 70 to 81 | 2.0 | 680 ± 4.3 | 342 ± 4.2 |
| LSS719 | 87 to 98 | 2.0 | 693 ± 4.9 | 339 ± 5.3 |
| LS55 | 94 to 105 | 2.2 | 287 ± 4.0 | 132 ± 5.6 |
| pANT1052 | none | 1.2 | 4670 ± 1.7 | 3816 ± 4.9 |

[a]The LS mutations present in the *rpoB* gene fused in frame to *lacZ* and carried on a lambda vector. pANT1052 contains an intact copy of *lacZ* downstream of the *ant*1052 promoter.

[b]Position in the wild-type sequence that has been replaced by the LS oligonucleotide (and deleted in those mutations that have an accompanying deletion). Numbering relative to the *rpoB* start site (+1).

[c]Translational feedback was calculated as (β-gal units in the strain without a subunit expression plasmid/β-gal units in the strain expressing β′ from pEXT21C1 in the presence of IPTG).

[d]β-galactosidase activities were determined in the absence of a subunit expression plasmid as described in Materials and Methods; all values are the averages of at least six assays performed on three or more independent cultures. The percent standard deviation is presented after each average.

[e]β-galactosidase activities as described in footnote 4, except they were determined in the presence of excess β′ expressed from pEXT21C1 induced with 10 μM IPTG.

The size of the LSM cassette and the selectable marker it carries are key features of this protocol and once linear DNA molecules are generated, the remaining protocol is direct and efficient. Although the LSM cassette can be inserted anywhere into the plasmid when the DNA is randomly cleaved with DNaseI, the increase in the target DNA fragment size provides a direct physical selection for only those insertions which are present in that sequence. Following ligation of these larger fragments back into the plasmid backbone, the isolation of tetracycline resistant colonies provides a phenotypic selection for only those recombinant plasmids that contain the LSM cassette.

The current LSM cassette was designed with recognition sites for *Sma*I, *Bpm*I and *Pml*I because of the absence of these sites in the *rpoB* target fragment and pLK6 plasmid. This cassette can be used directly to mutagenize any other recombinant plasmid that does not have recognition sites for these enzymes. If the plasmid does contain these sites, either they would first have to be eliminated by oligonucleotide mutagenesis or a new LSM cassette would have to be designed with recognition sites for other enzymes that do not cut the plasmid. In addition to providing a convenient marker for the position of the LS mutation, the unique *Pml*I site present on the LS oligonucleotide can facilitate the construction of deletions spanning the distance between any pair of LS insertions. Following digestion with *Pml*I, the ligation of any upstream *Pml*I terminus to any other downstream *Pml*I terminus results in the deletion of the intervening sequence.

Analysis of the series of LS mutations that were isolated in *rpoB* indicates that an extensive region surrounding the start site is important for efficient translation. All 23 mutations that alter sequences between –69 and +58 reduce translation to a third or less as compared to the wild-type sequence. In particular, 10 out of 13 mutations that alter sequences in the region of –25 to +16 have translational levels of 2% or less. For most *E.coli* genes, the minimal translational initiation region (TIR) is defined as the start codon and the upstream Shine–Dalgarno region composed of 3–8 nt that match the sequence AAGGAGGU (22). The distance between these two sites can vary between 4 and 12 nt. In *rpoB* there are 5 nt between the AGGA Shine–Dalgarno sequence and the AUG start codon. Surprisingly, two LS mutations that directly overlap these sites affect translation to a lesser extent than the flanking mutations (LSS702, –12 to –1, 12% translation; LS135, –4 to 12, 16% translation). However, the altered sequence in LSS702 has a potential Shine–Dalgarno sequence (AGGGG) located 9 nt upstream of the start codon. Although the LS135 mutation eliminates the normal start codon, it positions an alternative GTG start codon 5 nt downstream of an improved Shine–Dalgarno sequence (AGGAAG). Sequences flanking the Shine–Dalgarno sequence and start codon appear to be critical because six mutations that map immediately upstream of the Shine–Dalgarno region and three mutations that map just downstream of the start codon all reduce translation to 2% or less.

Two general models might explain why extensive sequences 5′ and 3′ to the start codon are required for efficient translation of *rpoB* mRNA. One proposal is that positioned within this region there are elements in addition to the Shine–Dalgarno sequence and start codon that are important for the mRNA interaction with the 30S subunit. A variety of additional recognition elements have been suggested in different transcripts (reviewed in ref. 23). The sequence surrounding the *rpoB* start site was analyzed to identify any regions of complementarity with 16S rRNA using the Best Fit program of the Wisconsin GCG software package (version 9). Interestingly, several regions of the *rpoB* mRNA in the vicinity of the start site showed a high degree of complementarity to sequences in 16S rRNA. A stretch of 8 nt (+5 to +12) has 100% complementarity to the 16S rRNA sequence from 453 to 460. It is noteworthy that this region overlaps the sequence (458–466) which has been suggested to enhance the level of translation of the *atpE* gene as well as translation of the bacteriophage T7 gene 10 (23 and references therein). An adjacent rRNA sequence (463–472) has also been proposed to provide an extra site of interaction with the *infC* mRNA to increase its level of translation (22,23). In addition, the interval of *rpoB* mRNA from –16 to +3 shows ~80% complementarity with the 16S rRNA sequence from 1053 to 1071. This region of 16S rRNA may be important in interactions with mRNA leading to translation initiation since nucleotide 1052 of 16S rRNA can be cross linked to the +6 nucleotide of mRNA (23).

A second model is that the RNA secondary structure of this overall region may be important to keep the TIR accessible to the 30S subunit. Changes upstream or downstream of the TIR may lead to alterations in secondary structure that then sequester the TIR. Preliminary computer analysis of possible mRNA secondary structures using the MFold program (24) of the Wisconsin GCG software package (version 9) suggests that at least some of the mutants have the potential to form secondary structures different from the wild-type sequence. However, we have not found a compelling correlation between the accessibility of the TIR in the putative secondary structures and the measured level of translation for the different mutants. The MFold program does not predict possible psuedoknot or tertiary structure that would also influence the accessibility of the TIR. Therefore direct structure mapping of this region will be required to assess its importance in the translation of *rpoB*.

Two other regions considerably farther from the TIR have a significant impact on translation levels. Upstream of the TIR, mutation LS117 (–208 to –194) and to a lesser extent the flanking mutation LSS62 (–194 to –181) showed reduced translation. This region has been shown to contain an RNaseIII processing site with the 3′ end of the cleaved transcript mapping in the vicinity of –194 to –191 and the 5′ end in the vicinity of –183 to –180 (25). It has been proposed that sequences from –200 to –170 and –121 to –91 form a stem that is important for RNaseIII processing (26). Therefore, the sequence alteration in LS117 and LSS62 may directly affect cleavage by RNaseIII or interfere with formation of the secondary structure that may be necessary for efficient RNaseIII recognition. We have not directly assayed processing of the mutant transcript, but if it is reduced it would suggest that cleavage of the polycistronic *rpoKAJLrpoBC* mRNA at this site is important for normal translation of *rpoB*.

Alteration of another region located ~105–126 nt downstream of the *rpoB* start also reduces translation to ~20–40% of the wild-type, although it is not clear how sequences >100 nt into the structural gene could have such an effect on translation. With a transcription rate of ~50 nt/s in *E.coli* (27) it would seem unlikely that sequences this far downstream could affect the secondary structure of the RNA in the vicinity of the start site before the first 30S particle would bind to the TIR.

Analysis of *rpoB–lacZ* mRNA levels for a selected set of LS mutations indicated that those mutations which showed significantly decreased translation also had reduced amounts of RNA. However, in all these cases the decrease in the level of the β subunit–β-galactosidase translation product was much greater

that the drop in RNA levels. The four upstream protein genes (*rplKAJL*) are also subject to autogenous translational regulation (28), with two independent translational feedback circuits, one involving *rplKA* and the other for *rplJL*. It has been demonstrated that the stability of the *rplKA* mRNA is decreased as an indirect effect of reduced translation, presumably as a result of decreased ribosome density on the message (29). The finding that the LS mutations reduce RNA levels much less than they decrease the translation product, also suggests that the primary effect of these mutations is to decrease translation which in turn reduces the stability of the mRNA. Although it is possible that some LS mutations could have their main effect on RNA stability, we assume that the overall profile seen in Figure 4 reflects sequences which have their primary effect on translation.

Previously, using a collection of *rpoB–lacZ* translational fusions with differing amounts of *rpoB* deleted, we found that a construct which retained only 70 bp of the *rpoB* sequence downstream of the start site exhibited very little repression in response to increased concentrations of β and β′ (6). The simplest interpretation of this result is that sequences important for autogenous regulation of β translation extend >70 bp into the structural gene. However, none of the LS mutations examined in this study showed a significant reduction in the feedback ratio when the intracellular level of both β and β′ was increased after induction of their expression from pACTBC. It is possible that the individual β and β′ polypeptides each interact with different sites on the *rpoB* mRNA to reduce translation. If none of the mutations in this study covers the sites for both subunits then, although the binding site for one subunit may be mutated, the feedback may not be reduced significantly due to the binding of the other subunit at its site. To circumvent this potential problem, a select group of mutations was analyzed using a plasmid that overproduces only β′. The partially overlapping LSS701 (+4 to +18) and LSS709 (+10 to +21) mutations had the lowest levels of repression in response to excess β′ while LSS843 (+16 to +27) had an intermediate level. Because the flanking mutations LS135 (–4 to +12) and LSS44 (–26 to +40) showed normal repression, it suggests the sequences important for regulation by β′ extend no further than +13 to +25. Since all three affected mutations have some overlap in the region from +13 to +18, the sequence required for repression may be contained within this interval.

Two general models of translation regulation have emerged from the analysis of a variety of phage and *E.coli* genes (22,28). In one model the repressor binds directly over or immediately adjacent to the TIR and sterically prevents interaction with the 30S subunit. In the second model, binding of the repressor stabilizes an inhibitory secondary structure that sequesters the TIR. It is possible that the large β′ polypeptide (155 kDa) binding in the vicinity of +13 to +25 could directly block access to the TIR. We have analyzed the RNA sequence in the region surrounding the *rpoB* start site (–200 to +200) using the Mfold program to determine if there is any correlation between the predicted RNA secondary structures and the ability of β′ to repress *rpoB* translation. The most energetically favorable structure of this overall region contains the *rpoB* start site at the base of a weak stem–loop structure which has a 4 nt stem (+2 to +5 paired with +16 to +19) and a 10 nt loop (+6 to +15). Due to the inherent weakness of this structure (–2.1 kcal), this region which overlaps the interval important for repression by β′ (+13 to +18 or +25) may often be single stranded *in vivo* thereby allowing direct binding of β′ to the primary sequence. The reduced level of repression seen with the downstream LSS809 (+66 to +77) mutation might be explained by the observation that this mutation causes drastic changes in the predicted RNA structure and incorporates the putative β′ binding site into a long, stable stem–loop structure. The predicted RNA secondary structure of the other LS mutations that do not reduce repression consistently show the same weak stem–loop structure seen with the wild-type sequence. A change in the RNA secondary structure in this region may also explain the previous observation that a deletion which retained only 70 bp of *rpoB* had reduced repression (6). The predicted folding of this deleted RNA incorporates the +13 to +25 region into a duplex structure which may be inaccessible to β′. More direct techniques such as RNA footprinting will be required to confirm the prediction that β′ binds in the vicinity of +13 to +25 and that binding is reduced in the mutated sequences which lower repression.

## ACKNOWLEDGEMENT

This study was supported by the Medical Research Council of Canada.

## REFERENCES

1 McKnight,S.L. and Kingsbury,R. (1982) *Science*, **217**, 316–324.
2 Steward,K.L. and Linn,T. (1991) *J. Mol. Biol.*, **218**, 23–31.
3 Passador,L. and Linn,T. (1989) *J. Bacteriol.*, **171**, 6234–6242.
4 Dykxhoorn,D.M., St. Pierre,R. and Linn,T. (1996) *Mol. Microbiol.*, **19**, 483–493.
5 Passador,L. and Linn,T. (1997) *Can. J. Microbiol.*, **47**, 819–826.
6 Passador,L. and Linn,T. (1992) *J. Bacteriol.*, **174**, 7174–7179.
7 Linn,T. and Ralling,G. (1985) *Plasmid*, **14**, 134–142.
8 Dykxhoorn,D.M., St. Pierre,R. and Linn,T. (1996) *Gene*, **177**, 133–136.
9 St. Pierre,R. and Linn,T. (1996) *Gene*, **169**, 65–68.
10 Moyle,H., Waldburger,C. and Susskind,M.M. (1991) *J. Bacteriol.*, **173**, 1944–1950.
11 Vogelstein,B. and Gillespie,D. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 615–619.
12 Hoopes,B.C. and McClure,W.R. (1981) *Nucleic Acids Res.*, **9**, 5493–5504.
13 Inoue,H., Nojima,H. and Okayama,H. (1990) *Gene*, **96**, 23–28.
14 Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
15 Linn,T. and St. Pierre,R. (1990) *J. Bacteriol.*, **172**, 1077–1084.
16 Santos,M.A. (1991) *Nucleic Acids Res.*, **19**, 5442.
17 Campbell,V.W. and Jackson,D.A. (1980) *J. Biol. Chem.*, **255**, 3726–3735.
18 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
19 Smith,M. (1985) *Annu. Rev. Genet.*, **19**, 423–462.
20 Haltiner,M., Kempe,T. and Tjian,R. (1985) *Nucleic Acids Res.*, **13**, 1015–1025.
21 Heffron,F., So,M. and McCarthy,B.J. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 6012–6016.
22 Gold,L. (1988) *Annu. Rev. Biochem.*, **57**, 199–233.
23 McCarthy,J.E.G. and Brimacombe,R. (1994) *Trends Genet.*, **10**, 402–407.
24 Zuker,M. (1989) *Science*, **244**, 48–52.
25 Downing,W.L. and Dennis,P.P. (1987) *J. Mol. Biol.*, **194**, 609–620.
26 Barry,G., Squires,C. and Squires,C.L. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3331–3335.
27 Yager,T. and von Hippel,P.H. (1987) In Neidhardt,F.C., Ingraham,J.C., Brooks-Low,K., Magasanik,B., Schaecter,M. and Umbarger,H.E. (Eds) Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology*. American Society for Microbiology, Washington, DC, USA.
28 Zengel,J.M. and Lindahl,L. (1994) *Prog. Nucleic Acid Res. Mol. Biol.*, **47**, 331–370.
29 Cole,J.R. and Nomura,M. (1986) *J. Mol. Biol.*, **188**, 383–392.