# Assessment of Site Specific Mutational Effect on Transcription Initiation at *Escherichia coli* Promoter

[1,2,3]S. Kannan, [1]M. Krishnan, [2]S. Sundarraj, [2]C. Meenambigai and [2]M. Vimaladevi
[1]Department of Ecobiotechnology, Bharathidasan University, Tiruchirappalli-620 024 TN, India
[2]Ayya Nadar Janaki Ammal College (Autonomous), Sivakasi-626 124, TN, India
[3]Department of Zoology, School of Life Sciences, Bharathiar University,
Coimbatore-641 046, TN, India

**Abstract: Problem statement:** It is widely accepted thought that the weak promoters control the RNA synthesis and play regulatory role in complex genetic networks in bacterial system. An experiment had been designed to address whether mutations in the -16/-17 region affect the rate of transcription at an activator-independent promoter in *E. coli* or not? **Approach:** The aim of this study was to determine whether mutations in the -16/-17 region affect the rate of expression at an activator-dependent promoter in JM109 strain of *E. coli*. Primers were constructed to amplify the mutant promoter genes through PCR. The amplified PCR product was checked and then inserted into the MCS region of pAA128 plasmid. Further the plasmid vector was transformed into JM109 strain of *E. coli* and then cloned the selected transformats. Finally, the plasmid from each mutant colony was then sequenced using the protocol supplied with the Amersham Pharmacia Biotech T7 sequencing Kit. The JM109 cultures for which the sequences were determined, then assayed for β-galactosidase activity to assess the rate of gene expression from the altered promoters. **Results:** The present investigation revealed that the extended-10 promoter region has a substantial effect on the rate of transcription at weak promoter sequence and also bearing little resemblance to the consensus sequence recognized by RNA. The expression of the genetically engineered plasmid proved that the 2 bps (-16 and -17 base pair) found adjacently upstream of the extended-10 promoter have an effect on the level of transcription. This was achieved by site specific base substitutions into the weak promoter of a modified lac operon lacking any activator or repressor binding sites. The results from gene expression assays of several mutants showed a distinct preference for either GG or TT located adjacently upstream of the extended promoter element. Thus the present study emphasized that extended promoter region also played a key role in regulation transcription initiation in JM109 strain of *E. coli*. **Conclusion:** The present study concluded that the site specific changed in the extended promoter regions, particularly the-17/-16 base pairs had greater influence in the transcription initiation in *E. coli*. Thus the promoter engineering study will definitely pave the way to do both, on or off the genetic switches in bacterial system according to our needs to produce high protein of interest or decrease or block the expression of a particular unwanted protein.

**Key words:** Extended promoters, *E. coli*, Transcription initiation, mutant promoters, JM109 strain

## INTRODUCTION

All cellular organisms use a complex molecular machine, the DNA-dependent RNA Polymerase (RNAP), to execute the first step in gene expression-transcription[1]. RNAP synthesizes the RNA chain complementary to the DNA template strand from Nucleoside Triphospate (NTP) substrates. Transcription is a cyclic process that can be roughly divided into three major steps-promoter DNA binding and RNA chain initiation, progressive RNA chain elongation and termination. The first step of the cycle requires the presence of one of the promoter specific σ subunits that together with the core forms the Holoenzyme[2,3]. The Holoenzyme slides along the DNA[4], locates the promoter sequence, binds to it and then melts the DNA duplex around the transcriptional start-site to form the transcription bubble. This process results in a stable open promoter complex, which initiates RNA synthesis. RNAP then enters a cycle known as abortive initiation,

**Corresponding Author:** S. Kannan, Department of Zoology, School of Life Sciences, Bharathiar University,
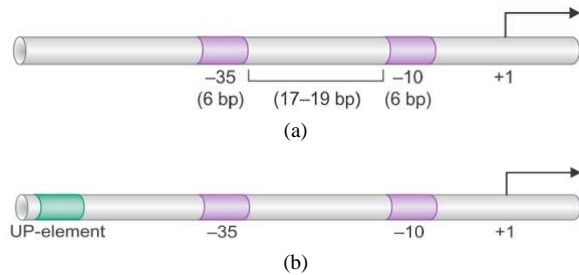Coimbatore-641 046, TN, India

Fig. 1: Showing the Promoter sequence recognized by σ70 of RNAP contains two conserved sequences (-35 and -10 regions/elements) separated by a non-specific stretch of 17-19 nt. Position +1 is the transcription start site. (a): Shows the location of -10 and -35 elements; (b): Pointed out the location of the up element which play a key role in transcription initiation in bacterial system

during which it directs the synthesis of short RNA products that are repeatedly released and re-synthesized[5]. Once, the RNA chain reaches 8-11 nucleotides (nt) the transition to productive elongation usually takes place. During promoter escape, RNAP usually releases σ factor soon after the initiation phase of transcription, relinquishing key contacts with the promoter[6,7].

The RNAP enzyme and particularly the sigma-core interface, undergoes a significant conformational change[8] that converts the initiation complex into an extremely stable and compact Elongation Complex (EC). The EC probably maintains an almost constant size ($\sim 12\pm 1$ nt) of the transcription bubble and $\sim 8\pm 1$ bp RNA-DNA hybrid until it reaches the termination signal[9,10]. Thus promoters control transcriptional process in all genes. The primitive step in transcription initiation is promoter recognition by RNA polymerase and the different DNA sequence elements. Till date four different elements have been reported. The two principal elements are the -10 hexamer and the -35 hexamer (Fig. 1) which are located in 10 and 35 base pairs upstream from the transcription start point. The other two important promoter elements are the extended -10 elements and the up elements. The extended -10 elements is a 3-4 bp motif located immediately upstream of the -10 hexamer that is recognized by domain 3 of the RNA polymerase σ subunits. An experiment has been designed to find out the influence of mutation at -17 and -16 positions in transcription initiation process. In this context, a genetically engineering a plasmid (pAA182) containing an artificial promoter region was constructed and then transforming

*E. coli* with this plasmid, selecting for the transformed bacterial cell and finally finding the mutations present and assaying for the levels of gene expression from the altered promoters.

**MATERIALS AND METHODS**

The fist step was the amplification of the template DNA into which the -17/-16 bp mutations were introduced by following PCR technique[11]. A constructed promoter sequence was used as a template for the PCR reaction. This template DNA contained the -10 and -35 promoter sequences 5'-TATGGT-3' and 5'-TAGATA-3' respectively, these, while bearing some homology to the consensus sequences were weak promoters. Template sequence used in the present study were: 5'TTC TAG ACA GCT CAT GCA TCT TTG TTA TGG TTA TTT CAT ACC ATA AGC CTA ATG GAG CGA ATT ATG AGA GTT CTG GTT ACC GCC AAG CTT GCC AAC GCC ATT TGG CTA CCC TGC CAC TCA CAC CAT TCA GGC GCC TGG CCG CGT GAA TTT GAT TGG TTA ACA CAC CGA CTA CAA3'.

The DNA primer sequences used were SK6741 and SKAN -17N/-16N; these primers were supplied by Alta Bioscience Pvt. Ltd., UK (Table 1).

The N positions shown in SKAN-17N/-16N represent a random base substitutions. The SKAN-17N and -16N primer had random residues substituted in the positions corresponding to -17 and -16 in an *E. coli* transcription promoter. This created random substitutions at these 2 base positions in the PCR product. The reaction mixture contained a 1x concentration of thermopol buffer, 0.2 mM each of the 4 dNTPs, 1 µL of Template DNA, a 1 µM concentration of both the SKAN-16N and -17N and SK6741 primers, 0.5 µL of Vent DNA pol. (New England Biolabs) and water up to 50 µL. In the PCR program a temperature of 94°C was used for 30 sec to melt the DNA, a temperature of 48°C for 30 sec for the annealing of the primer and a 15 sec period at 74°C for the elongation step. This program was cycled 30 times to amplify the template DNA with random substitution mutations at the positions corresponding to the -17 and -16 bp of a σ[70] recognized *E. coli* promoter region.

Table 1: The nucleotide sequence of selected primers

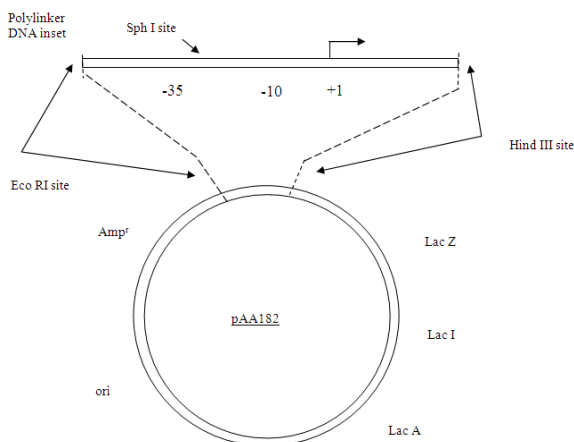| Primer name | Sequence |
|---|---|
| SK6741 | 5'-GTC GGT GTG TTC AC-3' |
| SKAN-17N/-16N | 5'-ACA GCT GCA TGC ATC TTT GTT ATG GTT ATT-3' |

Fig. 2: A simplified diagram of the plasmid pAA182 with an exploded view of the polylinker region

The product of the above PCR was then cleaned using phenol/chloroform extraction and ethanol precipitation[11] to remove cell debris and proteins will be a cleaned product which was then digested using the restriction enzymes Sph I and Hind III, using the protocol supplied by New England Biolabs.

The resulting fragments were separated using a 0.9% Agarose gel. The appropriate band was then removed from the gel and the DNA was electro-eluted[11], using a current of 40 mA for 45 min. The electroeluted DNA was then cleaned using phenol/chloroform extraction and ethanol precipitation[11]. And then the Sph I/Hind III fragment of cleaned DNA was then ligated[11] into the promoter region of the plasmid pAA182's modified *E. coli* lac operon (Fig. 2)[12]. The resulting plasmid containing the Sph I / Hind III fragment was then used to transform competent *E. coli*[11] (strain JM109) which were subsequently grown in Lennox broth (20 g Tryptone, 10 g yeast extract, 10 g NaCl L$^{-1}$) containing 80 µg mL$^{-1}$ of Ampicillin. The plasmid pAA182 contained the gene Amp$^r$ for resistance to Ampicillin, allowing selection of transformed JM109 cells with the Ampicillin. The transformed *E. coli* were grown in Lennox broth with Ampicillin and then plated onto Mc Conkey agar containing Ampicillin to the same concentration as in the Lennox broth. McConkey agar plates indicate the ability of bacteria to metabolize lactose; any colonies capable of metabolizing lactose appeared red. Further the appearance red of colony indicated the correct insertion of the Sph I/Hind III fragment into the plasmid pAA182, as the wild type *E. coli* JM109 strain was incapable of metabolizing lactose. Individual colonies were selected and grown in LB + Ampicillin.

The plasmid was then extracted from the selected mutants using the protocol of Maniatis *et al.*[11] with the addition of an RNAse digestion step, digesting with 0.02 mg of RNAse at 37°C for 30 min before the phenol/chloroform extraction step.

The pAA128 plasmid from each mutant colony was then sequenced using the protocol supplied with the Amersham Pharmacia Biotech T7 sequencing Kit. The bacterial cultures for which the sequences were determined and then assayed for β-galactosidase activity by following the method described in Maniatis *et al.*[11]. The *E. coli* cultures containing the sequenced plasmids were grown over night in LB. After this, separate subcultures were established from the overnight cultures and grown aerobically until the A$_{650}$ for each reached a value of 0.4-0.6. At this point the cultures were lysed by the addition of toluene, then removed by evaporation and 100µl of the lysed culture was added to 2.5 mL of Z-buffer (0.375 g KCl, 0.245 g MgSO$_4$.7H$_2$O, 8.53gNa$_2$HPO$_4$, 2H$_2$O, 800 mg Ortho nitro-phenol galactose, 2.70 mL β-mercaptoethanol and H$_2$O to 1 L). These reaction mixtures were then incubated at 37°C until yellow coloration was noticeable. At this juncture, the reaction was stopped by adding 1 mL of NaCO$_3$ (1 M). After stopping the reaction the A$_{420}$ was measured and the rate of ONPG hydrolysis were calculated as per the following equation:

$$\frac{OD_{420}}{OD_{650}} \times \frac{20000}{Reaction\,time\,(min)} = Nmoles\ ONPG\ hydrolyzed$$

$$min^{-1}\ mg^{-1}\ Protein\ dry\ weight$$

## RESULTS

The results from the β-galactosidase assays were shown in Table 2 which explored the effects on activity of the -17/-16 base changes of the mutant lac operon promoter in the plasmid pAA182 in *E. coli* JM109. The standard deviations of the duplicate results are also shown. The assay of the GG mutant showed the highest activity at 692.0 nmoles ONPG/min/mg protein (from herein referred to as units) and also the highest standard deviation, this, assuming a normal distribution, means that 66% of the results will fall between 814.2 and 370.2 units which is a very large variation. The next two most active mutations were the TT and the CC mutants with activities of 595.23 and 578.01 units respectively. These are followed in terms of activity by the CT mutant, which had an activity of 142 units and a standard deviation of 52.9. The two next highest activities were exhibited by the GT and AC mutants,

Table 2: Showing the β-Galactosidase activity for the possible mutants of -17 and -16 N of lac operon promoter region in pAA182 plasmid of *E. coli* (JM109)

| -17/-16 mutation | β-Galactosidase activity (nmol ONPG hydrolyzed min$^{-1}$ mg$^{-1}$ protein) | Standard deviation |
|---|---|---|
| AA | 2.00 | 2.10 |
| CT | 142.04 | 52.90 |
| CC | 578.01 | 175.87 |
| AC | 101.07 | 49.00 |
| GG | 692.00 | 220.00 |
| GT | 120.00 | 49.02 |
| TT | 595.23 | 191.62 |

120 and 101 units, the standard deviations for these results were very similar and quite low at 49.02 and 49.0 respectively which indicates much better grouping of results than the GG mutant. The least active mutant was the AA mutant, which exhibited virtually no β-galactosidase activity with an average activity of just 2 units. This result should be interpreted very carefully due to the fact that the standard deviation is larger than the result, which statistically should result in a substantial percentage of any repeat assays showing no activity.

## DISCUSSION

The results clearly show that mutations at positions -17/-16 did have an effect on the level of transcription from the weak *E. coli* promoter sequence that used in the present investigation. This correlates well with research already published about the -17/-16 promoter element. This region was found to have a substantial effect on the level of expression of α-amylase in *Bacillus subtilis*, base substitutions in this region led to in one instance a 94-fold reduction in the level of expression. The α-amylase promoter exhibits little homology to the consensus[13]. The same authors concluded that the ideal sequence for the tetramer located 1 base pair upstream of the-10 promoter element was 5'-TGTG-3'. There are a few suggestions in the tabulated data that suggest the same may be true of *E. coli* promoters, but as we did not manage to assay all possible nucleotide combinations that exhibited the highest levels of expression in the TT mutant. This may suggest that the base T is important at one or both of the mutated positions. Expanding on this, the mutants GT and CT exhibited relatively low levels of expression at this type of promoter. This in turn suggests that T located at -17 is responsible for elevating the level of transcription. This deduction could also be applied to the G located at -17. As the GG mutant exhibits the next highest level of transcription at the promoter used in this experiment and the GT mutant exhibits low level

of expression it could be concluded that G located at position -16 increases the level of expression. Besides, as the experimenter did not assay all the possible combinations of -17/-16 mutations, unable to confirm this deduction. Also, unlike in the research paper mentioned above, we had no wild type promoter with which to compare our results and hence, synthetic promoter is used which does not exist in nature. To draw any further conclusions it would have been necessary to assay all or most of the possible -17/-16 mutations, but It could be better to speculate that the TG mutant would exhibit the greatest levels of transcription at weak σ$^{70}$ recognized activator-independent promoters. This is based partly on the similarities between the σ$^A$ subunit of *Bacillus subtilis* and the σ$^{70}$ subunit of *E. coli*, however, this is far from certain as whilst σ$^{70}$ does recognize the same consensus -10 and -35 promoters as σ$^A$, σ$^A$ will often not translate *E. coli genes*[13]. This is thought to be due to differences in the patterns of minor promoters recognized and it is possible that the recognition of the extended-10 promoter may be one of the differences between σ$^{70}$ and σ$^A$. The results, while showing large standard deviations, did appear to follow the same trend in all the repeats of the assays. The results from the AA mutant virtually eliminate any expression at the promoter used. This makes it impossible to speculate which position occupied by A would cause the greatest suppression or indeed if any reasonable level of transcription can be achieved with A occupying either position.

This result suggests that it could be that the region of the σ$^{70}$ gene located directly after region 2.5 is responsible for the recognition of the -17/-16 promoter element. This result gains support from the observations of Barne *et al.*[14] that region 2.5 is responsible for the recognition of the extended-10 promoter (i.e., -16/-15). It could also be possible that the -17/-16 promoter element is recognized by the same helix-turn-helix motif of σ$^{70}$ that binds the extended -10 promoter. This would be challenging to determine though, as the interaction would probably not be sufficient to protect from foot-printing reagents. The X-ray crystallographic structure of σ$^{70}$ bound to DNA would however aid in determining this concept, nevertheless, the complexity and size of the enzyme, the availability of this soon is quite unlikely. The next step in classifying the-17/-16 promoter element should be to attempt to assay all 7 possible combinations of nucleotides and to do this for a variety of promoter, ranging from strong to weak. This would allow a more definite interpretation of the effects of the -17/-16 region on transcription initiation at *E. coli* promoters.

## CONCLUSION

In conclusion the present investigation suggests that the site specific changes in the extended promoter regions, particularly the-17/-16 base pairs have greater influence in the transcription initiation in *E. coli*. The result we obtained hint that the GG and CC mutants would exhibit the greatest levels of transcription at weak $\sigma^{70}$ recognized activator-independent promoters. Thus the promoter engineering study will definitely pave the way to do both up and down ward regulation of transcription in prokaryotic system according to our needs to produce high protein of interest or decrease or block the expression of a particular protein.

## REFERENCES

1. Browning, D.F. and S.J.W. Busby, 2004. The regulation of bacterial transcription initiation. Nat. Rev. Microbiol., 2: 57-65. http://www.nature.com/nrmicro/journal/v2/n1/abs/nrmicro787.html

2. Burgess, R.R. and L. Anthony, 2001. How sigma docks to RNA polymerase and what sigma does. Curr. Opin. Microbiol., 4: 126-131. DOI: 10.1016/S1369-5274(00)00177-6

3. Boreikhov, S. and E. Nudler, 2008. RNA polymerase: The vehicle of transcription. Trends Microbiol., 16: 126-134. http://www.ncbi.nlm.nih.gov/pubmed/18280161

4. Sakata-Sogawa, K. and N. Shimamoto, 2004. RNA polymerase can track a DNA groove during promoter search. Proc. Natl. Acad. Sci. USA., 101: 14731-14735. http://www.pnas.org/cgi/content/abstract/101/41/14731?ck=nck

5. Ishihama, A., 1993. Protion-protein communication within the transcription apparatus. J. Bacteriol., 175: 2483-2489. http://jb.asm.org/cgi/content/citation/175/9/2483

6. Bar-Nahum, G. and E. Nudler, 2001. Isolation and characterization of sigma 70-retaining transcription elongation complexes from *E. coli*. Cell, 106: 443-451. http://linkinghub.elsevier.com/retrieve/pii/S0092867401004615

7. Reppas, N.B., J.T. Wade, G. Church and K. Struhl, 2006. The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. Mol. Cell, 24: 747-757. http://www.cell.com/cell-stem-cell/abstract/S1097-2765(05)01681-3

8. Murakami, K.S. and S.A. Darst, 2003. Bacterial RNA polymerase: The whole story. Curr. Opin. Struct. Biol., 13: 31-39. DOI: 10.1016/S0959-440X(02)00005-2

9. Nudler, E., S. Boreikhov, D. Hurd and K. Struhl, 1997. The RNA/DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. Cell, 89: 33-41. http://www.ncbi.nlm.nih.gov/pubmed/9094712

10. Kireeva, M.L. *et al*., 2000. The 8-nucleotide-long RNA: DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. J. Biol. Chem., 275: 6530-6536. http://www.jbc.org/cgi/content/abstract/275/9/6530

11. Maniatis, T., E. Fritsch and J. Sambrook, 1989. Molecular Cloning-A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor. USA.

12. Rodrigue, S., J. Brodeur, P.E. Jacques, A.L. Gervais, R. Brzezonski and L. Gaudreau, 2007. Identification of mycobacterial and sigma factor finding sites by chromatin immuno precepitation assays. J. Bacteriol., 189: 1505-1513. http://jb.asm.org/cgi/content/abstract/189/5/1505

13. Grainger, D.C., H. Aiba, D. Hurd, D.F. Browning and S.J. Busby, 2007. Transcription factor distribution in *Escherichia coli* studies with FNR protein. Nucleic Acid Res., 35: 269-278. http://www.ncbi.nlm.nih.gov/pubmed/17164287

14. Wade, J.D., D.C. Roa, D.C. Grainger, D. Hurd, S.J. Busby, K. Struhl and E. Nudler, 2006. Extensive functional overlap between sigma factor in *Escherichia coli*. Nat. Struct. Mol. Biol., 13: 806-814. http://www.nature.com/nsmb/journal/v13/n9/abs/nsmb1130.html