Chapter 5

Investigations using DGGE of bisulphite sensitivity of DNA from the wild-type, rho and nusG mutant strains
INTRODUCTION

The investigations described in this chapter were done in an attempt to quantify the proportion of R-loops in the rho and nusG mutants of E. coli in comparison with the wild-type strain. Studies described in the previous chapters provided evidence that R-loops occur on the chromosome of E. coli strains compromised for Rho-dependent transcription termination.

Earlier observations in the laboratory had indicated that R-loops occur at increased frequency in the rho and nusG mutants of E. coli owing to defective factor-dependent transcription termination (Harinarayanan and Gowrishankar, 2003). It was argued that one instance of such R-loop occurrence may be the protein coding loci when translation uncoupled transcripts accumulate due to the inability of these mutants to terminate transcription effectively (Gowrishankar and Harinarayanan, 2004). In conformity with this model, it was assumed that R-loops can be detected in the rho and nusG mutants downstream of premature translation termination codons of chain terminating mutant alleles of protein coding loci. Accordingly, a set of three isogenic wild-type, rho and nusG strains containing early nonsense mutations in lacZ and trpE genes were constructed. The transcription of these mutant alleles was rendered constitutive by introducing lacI and trpR mutations into these strains (construction and phenotypic verification covered in detail in Chapter 3, Part A).

It was reasoned that RNA-DNA hybrid formation on the template strand renders the non-template strand unpaired, which can be detected biochemically, providing evidence of occurrence of R-loops. Several biochemical approaches were thus tested to detect single-strandedness at the mutant loci in these isogenic strains. Among the various approaches tested it was observed that in vitro chemical treatment with sodium bisulphite in the absence of denaturation allowed detection of single-stranded DNA in the chromosomal preparations (described in detail in Chapter 3, part B).

Chapter 4 elaborated the systematic approach adopted to standardize the assay for detection of single-strandedness using sodium bisulphite, and the principle involved therein. The studies so far indicated that although bisulphite mutagenesis in the absence of denaturation is the promising approach to detect single-strandedness in the chromosomal preparations tested, the PCR step used in the approach was not sufficiently discriminatory to quantify and compare the extent of single-strandedness in various strains tested. It was frequently observed that making a quantitative distinction between the wild-type and mutant strains was not feasible due to equal intensity of amplification in all the samples tested with primer pairs having either both native primers or at least one non-template-strand-specific (nts) converted primer. Moreover it was also observed that occasionally primer pairs consisting of a template-strand-specific (ts) primer gave amplification in all the samples tested. It was thus important to assess the significance of these observations.

Daskalakis et al. (2002) have reported the use of Denaturing Gradient Gel Electrophoresis (DGGE) after bisulphite treatment to determine the extent of methylation at the p15 and p16 genes in bone marrow mononuclear cells from patients with Myelodysplastic syndrome. The proportion of C-to-T converted amplicons and hence the extent of methylation was determined by extrapolating upon the intensity and the positions of the retarded smears and unretarded bands on the denaturing gradient gel when equivalent amplicons were tested in all the samples. The normal samples were unmethylated and therefore showed
maximal retardation owing to large scale C-to-T conversions after bisulphite treatment. The test samples were inferred to be a mixture of cells having methylated and unmethylated alleles based on the migration pattern on the DGGE gel after bisulphite treatment and PCR amplification which were visualized as smears of varying lengths and intensities on the DGGE gel.

Hence it was decided to include the DGGE approach along with bisulphite treatment of total nucleic acids in the absence of denaturation followed by PCR to determine the proportion of amplicons obtained that contained extensive C-to-T changes. These amplicons were expected to be observed as a retarded species when compared to the unmodified amplicons within the same amplification pool.

**PRINCIPLE**

5.1 Part A: Theory of Denaturing Gradient Gel Electrophoresis

Denaturing Gradient Gel Electrophoresis (DGGE) is highly efficient as a method for detection of DNA sequence differences (Fischer and Lerman, 1979). DGGE separates DNA molecules on the basis of differences in their thermal stabilities caused by differences in base sequence (Abrams and Stanton, 1992). The resolving power of this technique depends on the sudden drop in electrophoretic mobility that occurs when a DNA molecule partially melts, resulting in the formation of a structure that is partly helical and partly random chain (Guldberg et al., 1998).

The sensitivity of DGGE rivals that of sequencing in terms of application. It is widely used in microbial ecology to perform genetic fingerprinting of microbial communities by analyzing the rDNA loci in the populations tested (Muyzer et al., 1993) and also in genotyping studies to identify zygosity of alleles and mutations in diploid organisms occurring as SNPs in two different alleles (Lerman and Silverstein, 1987).

During DGGE, PCR generated DNA molecules encounter increasingly higher concentrations of chemical denaturant as they migrate through a polyacrylamide gel. Upon reaching a threshold denaturant concentration, the weaker melting domains of the dsDNA molecule begin to denature which results in immediate and dramatic slowdown of migration. Different sequences of DNA of same size tend to denature at different denaturant concentrations resulting in a characteristic band pattern on the gel. Each band is expected to represent a unique sequence in the PCR amplicon tested.

By employing DGGE, sequence variants in a given pool of DNA with a fragment size up to 500 bp can be conveniently detected (Myers et al., 1985). Subsequent detection of bands can be done by standard methods such as ethidium bromide staining or silver staining. The range of the denaturant in the denaturing gradient gel can be determined empirically by estimating the expected melting behaviour of the molecules being tested or can be determined after testing several ranges of denaturant.

5.1.1 Use of DGGE in testing for R-loops after bisulphite mutagenesis

In the current set of experiments, to address the problem of the inability to quantify the extent of
single-strandedness in the current set of investigations, the amplicons obtained from various primer pairs in different strains were subjected to DGGE. DGGE analysis allowed quantification of the extent of single-strandedness on each of the complementary strands at various loci by dissecting the contribution made by the unconverted and converted parent DNA template to the final amplification pool. The expectation from the DGGE analysis was that when the amplicons derived after bisulphite treatment in the absence of denaturation and subsequent PCR amplification with nts-specific primers from various samples are analyzed, a fraction of the amplified product in each sample may be retarded to give a smear. The intensity of the retarded smear in comparison to the respective unretarded band would represent the fraction of molecules that were modified by bisulphite and hence were more AT rich when compared to the molecules in the unretarded band.

It was reasoned that increased intensity of retarded bands in the rho and musG mutants when compared to the wild-type, as assessed on DGGE would be taken as evidence for increased occurrence of single-strandedness in the mutants. In conformity with the hypothesis of R-loop formation, it was expected that the intensity of this retarded smear would disappear or reduce in intensity upon RNase H treatment. The retarded smear was expected to be diminished in the amplicons obtained from ts-specific primer. These observations were proposed to be taken as an indirect proof for the occurrence of R-loops in these mutants.

Thus it was resolved to incorporate DGGE as a part of the sodium bisulphite analysis to quantify the proportion of molecules having C-to-T conversions in the samples being analyzed. AT richness in the molecules of the retarded smear will be a consequence of bisulphite mutagenesis, and hence when these molecules are eluted, cloned and the individual clones are sequenced, C-to-T transitions were expected to be observed specifically on the non-template strand.

RESULTS

5.2 Part B: Control experiments

5.2.1 Demonstration of difference in mobility of various trp locus clones obtained in previous cloning experiments on a denaturing gradient gel

To check the level of sensitivity of the bisulphite-DGGE approach, PCR amplified inserts from various trpE locus clones obtained in previous experiments were run on a 0-60% denaturing gradient gel. This range of denaturant gradient was empirically chosen to allow optimal separation of various clones tested. As a control, the same samples were also run on a native polyacrylamide gel where the mobility of all the samples was expected to be identical. In the denaturing gradient gel it was expected that the mobility of each of the samples would be unique and would correspond to the individual ATGC contents, with samples being progressively retarded for mobility on the gel with increasing AT content. The bands were subsequently detected by silver staining as per the protocol described in materials and methods section 2.2.2.14 of Chapter 2.
It was observed that as per the expectation all samples exhibited similar migration on a native polyacrylamide gel. Sample 35, which has 35 C-to-T changes was slightly retarded on this gel, conceivably due to secondary structure formation owing to extreme AT-richness. On the denaturing gradient gel, as expected the retardation pattern was concordant with the number of C-to-T changes in the individual amplicons. The mobility shift of the clones in general increased with increasing number of C-to-T changes (The number of C-to-T changes is indicated below each band). Interestingly, sample 5* that contains four C-to-T changes and one T-to-C change from the original sequence (thereby making three net C-to-T changes) is less retarded for mobility than sample 5 that has five C-to-T changes. It was observed that sample 9 that had nine C-to-T changes was actually less retarded for mobility than several other samples with lesser C-to-T changes. Thus it appears that AT content as well as spatial arrangement of the bases in sequence determines the final mobility pattern of a given DNA on a DGGE gel (figure 5.1).

![Figure 5.1: Testing for mobility shift of amplicons derived from individual clones of the trpE2 region having varying number of C-to-T changes. The numbers below each lane indicate the number of C-to-T changes in the sample loaded in the corresponding lane. Sample 5* has four C-to-T changes and one T-to-C change, resulting in three net C-to-T changes. The mobility of sample with zero changes (0) and that of the unmodified genomic DNA control (gDNA) are identical on both PAGE and DGGE analysis. L: 100 bp ladder; PAGE: PolyAcrylamide Gel Electrophoresis; DGGE: Denaturing Gradient Gel Electrophoresis.](image)

5.2.2 Demonstration of difference in mobility of MG1655 DNA upon bisulphite treatment on a denaturing gradient gel

To confirm that in our experiments, the mobility shift was dependent on bisulphite induced C-to-T changes, a control experiment was performed. MG1655 DNA was isolated from a culture grown overnight in LB broth at 37°C using the standard DNA isolation protocol mentioned in section 2.2.2.2 of Chapter 2. The following four preparations were derived in equimolar amounts from the DNA obtained:

i. MG1655 DNA untreated with bisulphite: The DNA preparation thus obtained was quantified and an aliquot was separated to serve as a sample that was untreated with sodium bisulphite thereby representing the unmodified parent DNA.

ii. MG1655 DNA heat denatured untreated with bisulphite: An aliquot of the DNA preparation was heat
denatured but not treated with sodium bisulphite to serve as a control for denaturation.

iii. MG1655 DNA native (non-denatured) bisulphite-treated: The DNA preparation was subjected to bisulphite mutagenesis in the absence of denaturation to assess the magnitude of C-to-T changes and hence the retardation of DNA after such treatment.

iv. MG1655 DNA heat denatured bisulphite treated: The DNA preparation was heat denatured and treated with sodium bisulphite to ensure maximum C-to-T conversions so as to estimate the retardation pattern upon DGGE analysis after such treatment.

The DNA preparations were then used as templates to amplify with trpE2 N–N and Cnts–Cnts primer pairs. (As explained in detail in Chapter 4, the following notations have been used in the primer description: N–N being the native primer pair, Cnts–Csts being the nts-specific converted primer pair and Csts–Cts being the ts-specific converted primer pair. Various combinations of each of these primers as detailed earlier were also employed for amplification.) The expectation was that N–N amplicons would show similar mobility from all preparations since this primer pair is expected to amplify unmodified DNA. The Cnts–Cnts amplicons would show difference in mobility such that the intensity and extent of retardation would be proportional to the number of molecules modified and the number of C’s per molecule modified by bisulphite. Cnts–Cnts amplicons were not expected to be recovered in samples (i) and (ii). N–Cnst and Cnts–N primer pairs were not used in this experiment since the quantity of input template and amplification efficiency with the Cnts–Cnts primer pairs was predictable and it was not intended to quantify the magnitude of C-to-T changes in this experiment. The samples were run on a native polyacrylamide gel and a 0-60% denaturing gradient gel. The bands were detected by silver staining of the gel as described in section 2.2.2.14 of Chapter 2.

Figure 5.2: Testing for mobility shift of amplicons derived from the trpE2 region of various DNA preparations. The mobility of all the amplicons is similar on the native polyacrylamide gel. A few prominent non-specific bands (NS) are also observed on the gel. DGGE analysis reveals that the N–N amplicons do not show retardation of mobility in any of the four samples but the Cnts–Cnts amplicons in (iii) and (iv) show retarded mobility. L: 100 bp ladder; PAGE: PolyAcrylamide Gel Electrophoresis; DGGE: Denaturing Gradient Gel Electrophoresis. (i): MG1655 DNA, non-
denatured, non-bisulphite treated; (ii) MG1655 DNA, denatured, non-bisulphite treated; (iii) MG1655 DNA, non-denatured, bisulphite treated, (iv) MG1655 DNA, denatured, bisulphite treated; N–N: Native primer pair; C_{nts}–C_{nts}: converted primer pair (non-template strand specific). Squares represent unretarded bands and circles represent retarded bands; NS represents non-specific bands.

N–N and C_{nts}–C_{nts} amplicons of the \textit{trpE2} region derived from the above mentioned four different DNA preparations when analyzed on a native polyacrylamide gel revealed that a prominent band corresponding to the expected size was observed with the N–N primer pair in all the four lanes corresponding to the four DNA preparations (squared). With the C_{nts}–C_{nts} primer pair, no band corresponding to the band of interest was observed in the lanes corresponding to samples (i) and (ii) that were derived from untreated DNA. A prominent band each of the expected size was observed in lanes (iii) and (iv) which were derived from bisulphite treated templates (squared). Non-specific bands (NS) were also observed in (i), (ii) and (iii) which have most likely arisen due to mispriming of the converted primers (figure 5.2).

On a 0-60% denaturing gradient gel, it was observed that the N–N amplicons derived from all the four templates showed similar mobility corresponding to that of the native unmodified sequence (squared). The C_{nts}–C_{nts} amplicons of the expected size that were observed only in samples (iii) and (iv) showed retarded mobility when compared to the N–N bands (circled). Since it was expected that these amplicons contain extensive C-to-T changes that were a consequence of bisulphite treatment, it was concluded that bisulphite induced C-to-T conversions were the cause of retarded mobility of these amplicons upon DGGE analysis.

5.3 Part C: Testing for mobility shift of various amplicons obtained from wild-type and the \textit{rho} and \textit{nusG} mutant strains and quantification of the extent of R-loop formation in the \textit{rho} and \textit{nusG} mutants

The results obtained from the analysis of previously obtained \textit{trpE} clones with varying C-to-T changes and of DNA preparations with and without bisulphite treatment on DGGE analysis suggested that the retarded mobility of DNA observed on denaturing gradient gels is an adept indicator of bisulphite-induced C-to-T changes on the DNA amplicons being analyzed. It was further reasoned that varying amounts and spatial arrangement of the C-to-T changes in the individual molecules comprising the amplified pool may show retarded mobilities unique to each molecule, thereby resulting in the formation of a retarded smear. The intensity and extent of this smear would provide an estimate of the proportion of molecules involved in R-loop formation as well as the extent of the stretch of DNA in the individual molecules that was amenable to bisulphite modification at the time of exposure.

Thus to quantify the extent of R-loop formation in the \textit{rho} and \textit{nusG} mutants, total nucleic acid preparations were isolated from wild-type (GJ6504), \textit{rho} (GJ6509) and \textit{nusG} (GJ6511) strains as per the protocol described in section 2.2.2.3 of Chapter 2. The nucleic acids obtained were then modified with sodium bisulphite in the absence of denaturation and were subsequently subjected to PCR amplification with the native and converted primer pairs at various test loci. These amplicons were checked on native
polyacrylamide gels as a control to check for equivalent mobility. The PCR products were then subjected to DGGE on a 0-80% denaturing gradient polyacrylamide gel to estimate the amount of AT rich DNA in the amplicons obtained with each primer pair. Silver staining of the gels was performed for detection of bands as described in detail in section 2.2.2.14 of Chapter 2.

The range of the denaturant in the denaturing gradient gel was chosen to be 0-80% since that was the widest range possible, thereby allowing separation of all possible sequence variants from the parent native DNA. Also, a wide range of gradient meant that molecules showing retarded mobility at a particular gradient concentration would form a tight smear owing to the presence of the required denaturant concentration in a relatively smaller space in the gel, than it would be if the range of the gradient was smaller. This was expected to allow greater probability of detection of the retarded smear by allowing more intense staining.

The following PCR conditions were used to amplify 250 bp stretches of bisulphite modified DNA with various primer pairs:

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>95°C</td>
<td>4 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Annealing</td>
<td>50°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>10 min</td>
</tr>
</tbody>
</table>

To test for the occurrence of R-loops downstream of the premature termination codon of the lacZU118 gene in the lac operon and to assess the extent of the region downstream of the stop codon mutation involved in R-loop formation, three regions in the lacZ gene and one negative control region in the lacY gene (designated as lacZ1, lacZ2, lacZ3 and lacY1) had been demarcated and various native and converted primers were designed for each of these regions, as described in detail in chapter 4. Similarly, two regions downstream of the premature stop codon were chosen in the trpE gene and one negative control region in the trpD gene of the trp operon (designated as trpE1, trpE2 and trpD1) to test for the occurrence of R-loops (refer figure 4.3 for a graphical depiction of the regions probed and amplified).

Each of these regions were amplified with the respective N–N, N–Cns and N–Cns primer pairs and the amplified products were analyzed on a pair of gels – native polyacrylamide gel and 0-80% denaturing gradient gel. Assuming that the hypothesis of R-loop formation is valid, the expectation was that the amplicons obtained with various primer pairs for a given region would exhibit similar mobility on a native polyacrylamide gel but show differential mobility on a denaturing gradient gel.

The amplicons obtained from the N–N primer pair, which was ideally expected to amplify only the native unmodified DNA, were expected to migrate to give a band at a region corresponding to the native DNA sequence under interrogation on the denaturing gradient gel. The intensity of the amplicons derived from the N–Cns primer pair was expected to be proportional to the extent of single-strandedness on the non-template strand of the region of the gene being probed and thus provide an indirect estimate of the amount of molecules involved in RNA-DNA hybrid formation on the template strand. It was therefore expected that in accordance
with the R-loop hypothesis, the N–C<sub>ns</sub> amplicons would show retarded mobility in the rho and nusG mutants, which would be significantly higher than that observed in the wild-type strain.

*Apriori* it was assumed that RNA-DNA hybrids do not occur on the non-template strand. Hence amplification from the N–C<sub>is</sub> primer pair was not expected to occur in the three samples tested.

A detailed account of the results obtained using bisulphite-DGGE is provided in the sections below. Finally, a summary of the results obtained using this approach is given in section 5.7 of this chapter at the end of the detailed description.

### 5.3.1 Testing for mobility shift of amplicons obtained from various regions of the lac operon

With the above mentioned expectations, N–N, N–C<sub>ns</sub> and N–C<sub>is</sub> amplicons derived from the three isogenic test strains (wild-type, rho and nusG) at the lacZ1, lacZ2, lacZ3 and lacY regions were subjected to native PAGE and DGGE analysis. The following figures illustrate the observations made from the experiment:

#### 5.3.1.1 lacZ1

As can be noted in figure 5.3, it was observed that at the lacZ1 region, which is immediately downstream of the amber codon mutation of lacZ, the intensity of the amplicons obtained with the control N–N primer pair was fairly high and equivalent in all the three samples tested when analyzed on a native polyacrylamide gel. With the N–C<sub>ns</sub> primer pair, the amplification observed was slightly less intense than with the N–N primer pair and was again observed to be equivalent in all the three samples tested. With the N–C<sub>is</sub> primer pair, no amplification of the expected size was observed in all the three samples.

When the observed amplicons were analyzed on the denaturing gradient gel, the N–N amplicons gave two bands, conceivably corresponding to the bisulphite converted (circled) and unconverted DNA (squared). Thus it was observed for the first time that the N–N primer pair amplified both native and modified DNA with almost equal intensity. These two species of DNA appeared as one band on a native polyacrylamide gel, but split to give two distinct bands on DGGE. Interestingly, the retarded species formed a compact band on the DGGE gel, indicating that the extent of C-to-T conversions in these bands was limited and probably confined to a small region of the amplicon.

All the N–C<sub>ns</sub> primer pair amplicons showed total retardation in mobility (circled) and the band corresponding to the native unconverted sequence was undetectable in these lanes. The intensity of the band in the wild-type strain was faint when compared to the cognate rho and nusG bands. One likely explanation of the observation that the intensity of the three bands was equivalent on a native gel but different on the denaturing gel is that the extent of C-to-T conversions in the wild-type sample may have been unique to the individual amplicons such that the spread of the retarded smear was more wide in this sample than in the mutant samples. This may have led to the formation of a wide but faint smear that was not detected subsequently. In the rho and nusG samples, the C-to-T conversions may have
been clustered giving amplicons of similar mobility which gave relatively compact smears when compared to the wild-type sample.

Figure 5.3: Testing for mobility shift of amplicons derived from wild-type (W), \( \text{rho} \) (R) and \( \text{nusG} \) (N) strains carrying \( \text{lacZ}_{118} \) mutant locus with various primer pairs from the \( \text{lacZ1} \) region. The proportion of the retarded bands in the N–C\(_{\text{nts}}\) lanes is moderately enhanced in the mutants than in the wild-type strain upon DGGE analysis. There was negligible amplification in the N–C\(_{\text{nts}}\) lanes to make any interpretations from the patterns observed. PAGE: PolyAcrylamide Gel Electrophoresis; DGGE: Denaturing Gradient Gel Electrophoresis; W: wild-type; R: \( \text{rho} \); N: \( \text{nusG} \); N–N: Native primer pair; N–C\(_{\text{nts}}\): Primer pair with non-template strand specific converted reverse primer; N–C\(_{\text{nts}}\): Primer pair with template-strand specific converted reverse primer. Squares represent unretarded bands and circles represent retarded bands.

Thus it could be inferred from these observations that when equimolar amounts of input DNA (equivalent N–N amplicon intensity) are analyzed for the presence of single-strandedness on the non-template strand of transcription (retarded smear in the N–C\(_{\text{nts}}\) lanes) in the \( \text{lacZ1} \) region, there is at best a marginal increase in single-strandedness in the \( \text{rho} \) and \( \text{nusG} \) mutants than in the wild-type strain. No single-strandedness on the template strand (N–C\(_{\text{nts}}\) smears) could be detected at this region in any of the strains tested.

5.3.1.2 \( \text{lacZ2} \)

Figure 5.4 depicts the observations made from PAGE and DGGE analysis of the amplicons derived from the wild-type, \( \text{rho} \) and \( \text{nusG} \) strains at the \( \text{lacZ2} \) region. It was observed that the N–N and N–C\(_{\text{nts}}\) primer pairs gave equal amplification in all the three strains tested as observed on the native gel (squared) but the N–C\(_{\text{nts}}\) primer pair did not give amplification intense enough to be observed on the gel. The N–N primer pair also gave non-specific amplification of a size larger than the expected amplicon size (NS). When these amplicons
were analyzed on a denaturing gradient gel, it was observed that the N–N amplicons gave two sharp bands corresponding to the native sequence (squared) apart from the non-specific band (NS) much higher in the gel. Another band that was slightly spread as a tight smear, probably owing to a few background C-to-T changes as a result of bisulphite exposure also were apparent apart from these two bands (circled). The N–C\textsubscript{nts} amplicons split to give two bands each, one corresponding to the native sequence (squared), the other a smear above the native band (circled). The intensity of all the three bands was equivalent across samples. The N–C\textsubscript{ts} primer pair failed to give any noticeable band in any of the three samples.

**Figure 5.4:** Testing for mobility shift of amplicons derived from wild-type (W), rho (R) and nusG (N) strains carrying \textit{lacZ}\textsubscript{U118} mutant locus with various primer pairs from the \textit{lacZ2} region. The proportion of the retarded bands in the N–C\textsubscript{nts} lanes is equal in all the three strains upon DGGE analysis. There seems to be very small amount of amplification in the N–C\textsubscript{ts} lanes corresponding to the rho and nusG mutants which shows retarded mobility on the denaturing gradient gel. Abbreviations used in the figure are explained in legend to figure 5.3. NS represents non-specific bands.

Thus it was interpreted from these observations that the extent of single-strandness of the non-template strand is similar in the wild-type, rho and nusG strains while the template strand was not unpaired.

**5.3.1.3 lacZ3**

Figure 5.5 depicts the observations made from various amplicons derived from the \textit{lacZ3} region in the wild-type, rho and nusG strains. It was observed that the N–N amplicons derived from the three strains were of equivalent intensity on the native gel (squared). The N–C\textsubscript{nts} amplicon obtained from the wild-type strain was of a significantly lesser intensity than that obtained from the rho and nusG strains.
No band was observed corresponding to the N–Cₜₛ lane in the wild-type strain but faint bands were observed in the rho and nusG samples (squared).

**Figure 5.5:** Testing for mobility shift of amplicons derived from wild-type (W), rho (R) and nusG (N) strains carrying lacZₚₕₚₜ mutagenesis. A significant proportion of the amplicons in the N–Cₜₛ lanes were retarded upon DGGE analysis in all the strain tested. The intensity of the retarded smears in the rho and nusG mutant lanes was approximately three-fold higher than in the wild-type strain. There is considerable amount of amplification in the N–Cₜₛ lanes corresponding to the rho and nusG mutants which shows retarded mobility on the denaturing gradient gel. Abbreviations used in the figure are explained in legend to figure 5.3.

DGGE analysis of these amplicons revealed that in each of the three N–N samples, two bands were observed, one of which corresponded to the native unconverted sequence (squared) while the other was a retarded smear (circled). The intensity of the retarded smear was slightly less intense in the wild-type than in the rho and nusG lanes. The N–Cₜₛ samples also gave two bands each on the DGGE gel, one band corresponding to the native DNA (squared) and another retarded smear (circled) representing bisulphite converted DNA. In the wild-type N–Cₜₛ lane, the retarded band corresponding to the converted sequence was quite prominent, indicating that significant amount of single-strandedness even in this strain. However the intensity of this band was nearly three-fold less than the corresponding band observed in the mutant samples. This indicated that the extent of single-strandedness was nearly three-fold higher in the mutants than in the wild-type strain.

The faint N–Cₜₛ amplicons that were observed in the rho and nusG lanes also showed retarded mobility (circled), thereby indicating that extensive C-to-T changes occurred on these amplicons as well. This indicated that the N–Cₜₛ amplification observed was not a consequence of mispriming or non-specific
amplification. Instead, it represented authentic amplification of the bisulphite modified annotated transcriptional template strand.

These observations indicated that a significant proportion of the non-template strand as well as a considerable proportion of the template strand of the lacZ3 region is single-stranded in the three strains analyzed. Densitometric analysis was done by comparing the intensity of the retarded smear (circled) in the individual lanes for a given strain to the intensity of the native unretarded band (N–N, squared) of the same strain. The analysis revealed that nearly 25-30% of the non-template strand DNA was single-stranded in the rho and nusG mutants. It was also revealed that the extent of single-strandedness in the wild-type strain is approximately 10% of the total DNA analyzed, which is approximately three-fold less than the mutant strains. Approximately 3-5% of the template strand DNA was also estimated to be single-stranded by this analysis at the lacZ3 region in the rho and nusG mutants.

5.3.1.4 lacY1

![Figure 5.6](image)

**Figure 5.6**: Testing for mobility shift of amplicons derived from wild-type, rho and nusG strains carrying lacZU118 and trpE9777 mutant loci with various primer pairs from the lacY1 region. The proportion of the retarded bands in the N–Cnts lanes was equal in all the three strains upon DGGE analysis. Amplification in the N–Cts lanes is undetectable in all the strains tested. Abbreviations used in the figure are explained in legend to figure 5.3.

When N–N, N–Cnts and N–Cts amplicons of the expected negative control locus lacY1 of the lac operon were analyzed on a native polyacrylamide gel, it was observed that the N–N and N–Cnts samples had equivalent intensity in all the three samples tested (squared) whereas no bands were observed in the N–Cts lanes. When DGGE analysis of these samples was performed, it was observed that amplicons obtained with both the N–N and N–Cts primer pairs formed bands corresponding to the native sequence (squared). These
observations indicated that as expected the N–N primer pair amplified the native unmodified DNA. However, the N–C\textsubscript{nts} primer pair also amplified the native unmodified DNA which may have arisen as a consequence of mispriming by the –C\textsubscript{ts} reverse primer.

### 5.3.2 Testing for mobility shift of amplicons obtained from various regions of the \textit{trp} operon

In a manner similar to the \textit{lac} operon, N–N, N–C\textsubscript{nts} and N–C\textsubscript{ts} amplicons derived from the \textit{trpE1}, \textit{trpE2} and \textit{trpD1} regions of the \textit{trp} operon were subjected to native PAGE and DGGE analysis. A broadly similar pattern of single-strandedness as indicated by the intensities of the retarded and unretarded bands was observed even in the \textit{trp} operon as was observed for the \textit{lac} operon. The following figures demonstrate the observations made at each defined region of the \textit{trp} operon:

#### 5.3.2.1 \textit{trpE1}

In the \textit{trpE1} region (figure 5.7), equal intensity of amplification was seen with the N–N primer pair on the native gel (squared). Faint bands of equal intensity were observed in the N–C\textsubscript{nts} lanes in the three samples (squared). Several non-specific bands were also observed in these lanes (NS). Interestingly, in the N–C\textsubscript{ts} lanes, the intensities of the \textit{rho} and \textit{nusG} bands were slightly higher than that of the wild-type band (squared).

![Figure 5.7: Testing for mobility shift of amplicons derived from wild-type, \textit{rho} and \textit{nusG} strains carrying \textit{lacZ}\textsubscript{U118} and \textit{trpE9777} mutant loci with various primer pairs from the \textit{trpE1} region. The proportion of the retarded bands in the N–C\textsubscript{nts} lanes was equal in all the three strains upon DGGE analysis. There was a significant amount of amplification in the N–C\textsubscript{ts} lanes of which a considerable proportion shows retarded mobility on the denaturing gradient gel, the retardation being more in the \textit{rho} and \textit{nusG} mutants than in the wild-type. Abbreviations used in the figure are explained in legend to figure 5.3. NS represents non-specific bands.

Upon DGGE analysis it was observed that the N–N amplicons formed two bands corresponding to the native DNA observed on the native PAGE gel (squared). The N–C\textsubscript{nts} amplicons derived from all the three
strains formed tight retarded smears indicative of C-to-T changes of equal extent in all the strains tested (circled). The N–Cₙₙ lanes also showed tight retarded smears in all the strains, with the intensities of the smears in the lanes corresponding to the rho and nusG strains being greater than that in the wild-type strain (circled). These results indicated that the non-template strand as well as a significant proportion of the template strand at the trpE1 region was single-stranded at the time of bisulphite exposure. The proportion of single-stranded molecules was greater in the rho and nusG mutants than in the wild-type strain. It was estimated that nearly 10% of the template strand DNA was unpaired at the trpE1 region.

**5.3.2.2 trpE2**

When analyzed on a native PAGE gel, the N–N lanes indicated equal intensity of amplicon recovery in the wild-type, rho and nusG strains (squared). The N–Cₙₙ lanes revealed that amplification in the rho and nusG samples was much greater than that observed in the wild-type (squared). With the N–Cₙ₉ primer pair, faint amplification was observed in the three strains.

DGGE analysis of the trpE2 amplicons revealed that the N–N amplicons represent native unmodified DNA (squared), while the N–Cₙₙ and N–Cₙ₉ amplicons, which were observed to be retarded in mobility to form distinct smears (N–Cₙₙ) or bands (N–Cₙ₉) represented DNA that had C-to-T conversions (circled). Therefore it was concluded that a significant proportion of the non-template strand as well as a small proportion of the template strand of the trpE2 region was single-stranded in the parent DNA during bisulphite treatment.

![Figure 5.8: Testing for mobility shift of amplicons derived from wild-type, rho and nusG strains carrying lacZU118 and trpE9777 mutant loci with various primer pairs from the trpE2 region. A significant proportion of the amplicons in the N–Cₙₙ lanes are retarded upon DGGE analysis in all the strain tested. The intensity of the retarded smears in the rho and nusG mutant lanes was approximately three-fold higher than in the wild-type strain. There is a considerable amount of amplification in the N–Cₙ₉ lanes in all the three samples and the lanes corresponding to the rho and nusG mutants show](image)
Densitometric analysis revealed that as in the lacZ3 region, nearly three-fold greater proportion of molecules were single-stranded on the annotated non-template strand of transcription in the rho and nusG mutants than in the wild-type strain. Nearly 30% of the non-template strand in the rho and nusG mutants and 10% in the wild-type strain were determined to be single-stranded. Nearly 5-8% of the template strand DNA was also estimated to be single-stranded in the rho and nusG mutants at the trpE2 region.

5.3.2.3 trpD1

The amplicons derived from the expected negative control trpD1 for the trp operon, when analyzed on native PAGE gave equal intensity of amplification in the three strains tested with the N–N and N–Cts primer pairs (squared). The amplicons generated by the N–Cts primer pair were of slightly higher intensity when derived from the rho and nusG templates than from the wild-type (squared).

Upon DGGE analysis it was observed that a significant proportion of the N–N amplicons and the entire N–Cts and N–Cts amplicons derived from the three strains showed retarded mobility (circled). These results indicated that a significant proportion of the DNA on the non-template strand as well as the template strand is in single-stranded form in the rho and nusG mutants as well as the wild-type strain. Interestingly, it was noted that the intensity of the retarded smear in the lanes corresponding to the template as well as the non-template strand was similar in all the three strains (circled).

**Figure 5.9:** Testing for mobility shift of amplicons derived from wild-type, rho and nusG strains carrying lacZ\textsubscript{U118} and trpE\textsubscript{9777} mutant loci with various primer pairs from the trp\textsubscript{D1} region. The intensity of the N–Cts amplicons is equal in all the strains while the intensity of the N–Cts amplicons is higher in the lanes corresponding to the rho and nusG mutants than that corresponding to the wild-type strain. The entire N–C\textsubscript{cts} and N–C\textsubscript{cts} amplicons show retarded mobility upon DGGE analysis. Abbreviations used in the figure are explained in legend to figure 5.3.
5.3.3 Sequencing of the retarded and unretarded bands from the \textit{rho lacZ3} and \textit{rho trpE2} amplicons obtained on the DGGE gel

According to the hypothesis under investigation, bisulphite treatment of total nucleic acids followed by PCR amplification of modified DNA and subsequent DGGE analysis will result in appearance of smears of retarded DNA bands on the gel. The retarded bands will have extensive C-to-T changes and are expected to represent the non-template strands of the DNA involved in R-loop formation in the parent DNA pool.

\textbf{Figure 5.10:} Sequencing of clones derived from retarded and unretarded bands obtained from the N–C\textsubscript{nts} amplicons of the \textit{rho} mutant strain carrying \textit{lacZ118} and \textit{trpE9777} mutant loci at the \textit{lacZ3} (top panel) and \textit{trpE2} (bottom panel) regions. The arrows emerging from the gel show the positions of the bands from where the DNA was eluted, cloned and sequenced. Annotations in the figure are explained in legend to figure 4.6. It was observed that extensive C-to-T changes occurred in the sequences derived from the clones containing the DNA from the retarded smears in both loci tested while the unretarded bands had considerably fewer C-to-T changes.
To test the above-mentioned expectation, the retarded and unretarded bands from the N–C
amplicons of the rho mutant derived from the lacZ3 and trpE2 regions were electro eluted from the DGGE gel separately (protocol described in Chapter 2, section 2.2.2.9.1) and the DNA from the individual bands was cloned into the Invitrogen TA cloning vector. The individual clones thus obtained were screened for the presence of the insert and the insert-containing clones were sequenced. The sequences so obtained were analyzed for C-to-T changes and are graphically depicted in the two panels of figure 5.10.

As can be observed from the depictions in the figure, in the sequences derived from the unretarded band of the N–C
amplicons at both the lacZ3 and trpE2 regions, more than 80% of the clones show 4% or less C-to-T changes. Greater than 80% of the clones obtained from the retarded smear showed extensive (40% or more) C-to-T changes. Upon further analysis it was observed that all the C-to-T changes in these amplicons occurred exclusively on the non-template strand of transcription. Thus the sequencing results establish the expectation that the retarded bands represent the AT-rich fraction of the amplicon generated, corresponding to C-to-T changes on the non-template strand transcription at the loci being tested.

5.4 Part D: Control experiments

5.4.1 RNase H treatment reduces the intensity of the retarded smear observed at the lacZ3 and trpE2 regions

The results so far generated provided evidence in favour of the R-loop hypothesis, i.e., that there is increased occurrence of single-strandedness of the non-template strand of transcription in the rho and nusG mutants, downstream of the premature translation termination codon in the lacZ118 and trpE9777 genes. Appearance of a more intense retarded smear in these samples than that observed in the wild-type sample when equimolar amounts of bisulphite treated preparations are amplified with N–C
primer pairs was observed at these loci. Cloning of the unretarded and retarded bands and sequencing of the individual clones established that only the retarded smears represented DNA having extensive C-to-T conversions specifically on the non-template strand of transcription at these loci.

To further determine that the observed single-strandedness of the non-template strand at these mutant loci is due to RNA-DNA hybrid formation on the template strand, it was decided to perform native PAGE and DGGE analysis of N–N, N–C
and N–C
amplicons derived from the bisulphite treated wild-type, rho and nusG mutant samples following in vitro RNase H treatment of the total nucleic acids. The expectation was that if the observed retarded smear was indeed due to R-loop formation at these loci, then RNase H treatment would totally abolish or greatly reduce the intensity of the retarded smear that was observed on DGGE analyses. This analysis was done with lacZ3 and trpE2 since these represent the terminal regions of untranslated transcription in the two mutant loci where the extent of single-strandedness on the non-template strand was observed to be maximal in previous analyses.

The panels in the following figure (figure 5.11; top panel - lacZ3 and bottom panel - trpE2) indicate
the observations made from the native PAGE and DGGE analysis of the RNase H treated total nucleic acids derived from the wild-type, "rho and nusG" mutant strains followed by bisulphite treatment in the absence of denaturation and subsequent PCR amplification with the indicated primer pairs. As can be observed from these figures, at both the lacZ3 and trpE2 regions, when equimolar amounts of input DNA, as evidenced by equal intensity of amplification with the N–N primer pair (squared), were subjected to amplification by the N–C
ts and N–C
ts primer pairs, the intensity of the retarded smears from these amplicons was much reduced (circled) when compared to the same amplicons derived from bisulphite modified preparations not treated with RNase H. Furthermore, as expected, when the band intensities for the three strain samples after RNase H treatment was compared, the intensities of the mutants were not intense as compared to the wild-type, and in fact was somewhat weaker than the band intensity of the wild-type strain.

Figure 5.11: Testing for mobility shift of amplicons derived from RNase H treated nucleic acid samples of wild-type,
rho and nusG strains carrying lacZU118 and trpE9777 mutant loci with various primer pairs from the lacZ3 (top panel) and trpE2 (bottom panel) regions. The intensities of the retarded smears was considerably reduced in each of these samples when compared to the RNase H untreated counterparts. Abbreviations used in the figure are explained in legend to figure 5.3.

Interestingly the smear in the N–Cts lanes also disappeared after RNase H treatment. This indicated that the signal observed as a retarded smear corresponding to single-strandedness on the canonical template strand was also a consequence of RNA-DNA hybrid formation on the canonical non-template strand. The implied interpretation of this is that R-loops occur at the loci tested due to a low level of antisense transcription occurring as well. This may be kept under check by the Rho-NusG surveillance system in a strain proficient for transcription termination. However no signal corresponding to the unpaired template strand was observed in the wild-type strain within the limits of sensitivity of the DGGE assays.

5.4.2 Testing for single-strandedness on the template and non-template strands at the trpE2 region in the isogenic wild-type, rho and nusG strains that carry wild-type trpE gene

As per the hypothesis under investigation, R-loops are expected to occur at increased frequency downstream of the premature translation termination codons of mutated structural genes as a consequence of production of translation-uncoupled transcripts in strains mutated for rho or nusG. It was thus expected that in these mutants R-loops will not occur in wild-type genes, or may do so at much lower frequencies, since transcription is coupled to translation in such genes. It was therefore hypothesized that the signal corresponding to single-strandedness on the non-template strand of the trpE+ gene would not occur on DGGE analysis in the rho and nusG mutants.

Among the various loci tested the trpE2 N–Cns amplicons gave a clear differential signal in the rho and nusG mutants carrying the trpE9777 allele when compared to the isogenic wild-type strain upon DGGE analysis. Therefore N–N, N–Cns and N–Cts amplicons of the trpE2 region derived from isogenic wild-type (GJ6525), rho (GJ6531) and nusG (GJ6532) strains carrying the trpE+ gene were subjected to native PAGE and DGGE analysis. It was expected that the signal corresponding to single-strandedness on the non-template strand would be diminished in these strains.

It was observed that when amplicons with various primer pairs derived from equimolar amounts of input DNA were subjected to native PAGE analysis, the bands obtained were of equivalent intensity in all the strains tested (squared). DGGE analysis revealed that the intensity of the retarded smears in all the three strains was considerably lower for the trpE+ allele (circled) when compared to those observed with the trpE9777 allele (circled, figure 5.8). Moreover the intensities of the retarded smears in the rho and nusG lanes with both N–Cns and N–Cts primer pairs were marginally higher than in the wild-type strain. Since R-loops were not expected to form when simultaneous translation of a transcript occurs, it was reasoned that the observed smears of equal intensity in all the strains tested corresponded to the active transcription bubble or the lagging strand of the replication fork or both. It may also be possible that low level of R-loops occur even when transcription and translation are coupled,
contributing to the marginal difference observed between the wild-type and mutant strains.

**Figure 5.12:** Testing for mobility shift of amplicons derived from wild-type, *rho* and *nusG* strains carrying *lacZ* and *trpE* genes with various primer pairs from the *trpE2* region. A significant proportion of the amplicons in the N–C<sub>nts</sub> lanes was retarded upon DGGE analysis in all the strain tested. There was minimal amplification in the N–C<sub>is</sub> lanes as well which shows retarded mobility on the denaturing gradient gel. The intensities of the N–C<sub>nts</sub> and N–C<sub>is</sub> amplicons were marginally higher in the *rho* and *nusG* mutants than in the wild-type strain. Abbreviations used in the figure are explained in legend to figure 5.3.

### 5.5 Part E: Summary of results of the bisulphite-DGGE approach

The following is a general summary of the observations made from the bisulphite-DGGE experiments at the *lac* and *trp* operons of the wild-type, *rho* and *nusG* strains:

1. DGGE results indicate that there is a difference in the proportion of non-template strand molecules that were single-stranded as compared to the template strand within each strain. The proportion of non-template strand molecules in the single-stranded form was observed to be greater than the template strand at all the regions of the *lac* and *trp* operons tested. For example at the *lacZ3* region the intensity of the retarded smear in the *rho* and *nusG* mutants when amplified with N–C<sub>nts</sub> primer pair was greater than that observed in the wild-type strain. Similar observations were made with the N–C<sub>is</sub> primer pair even at the *trpE2* region.

2. The non-template strand specific retardation signal was higher in the *rho* and *nusG* mutants than the wild-type strain with the following trend:
   i. There is no appreciable difference in the intensity of the retarded band across the three strains immediately downstream of the premature stop codons.
   ii. There is a very subtle difference in the intensity of the retarded band between the wild-type and the *rho* and *nusG* mutants further downstream in the ORF.
   iii. There is a marked difference in the intensity of the retarded band between the wild-type and
the rho and nusG mutants at the 3’ end of the gene close to the natural translation stop signal. The intensities of the retarded bands indicated that nearly 10% of the non-template strand DNA in the wild-type and nearly 30% in the rho and nusG mutants at the loci tested was unpaired and susceptible to bisulphite treatment.

iv. There is no difference in the mobility shift of the amplicons in the three strains in the ORF downstream of the gene with premature termination codon; i.e., in the lacY and trpD gene of the respective operons when non-template strand specific signal is analyzed.

The intensity of the retarded bands with template strand specific primers is faint when compared to the non-template strand. As a general trend, the intensity of the template-strand specific signal was greater in the rho and nusG mutants than in the wild-type strain. This observation is true for lacZ, trpE as well as trpD but was not observed in lacY. Moreover, the single-strandedness at the template strand of trpE in the rho and nusG mutants showed 5’–3’ funneling effect as was observed for the non-template strand of lacZ and trpE in these strains. Also, it seemed that the amount of annotated template strand DNA in single-stranded form in the trpD gene is higher than that observed in the lacY gene.

4. Sequencing of the non-template strand specific retarded and unretarded bands derived from the N–C_mn amplicons of the rho mutant strain establish that the retarded bands represent the AT-rich molecules corresponding to C-to-T changes exclusively on the non-template strand.

5. The intensity of the retarded non-template strand as well as the template strand band was a direct correlate of the number of molecules that were in R-loop formation since the signal weakens considerably when analyzed after RNase H treatment. Hence the results seem to suggest that the rho and nusG mutants have increased R-loops on the chromosome at the loci tested when compared to the wild-type. The propensity of R-loop occurrence does not seem to be confined to the annotated template strand of transcription. Low level of transcription seems to occur even in the antisense orientation as evidenced by appearance of a faint smear with template strand-specific primer pairs and the disappearance of the smear upon RNase H treatment.

DISCUSSION

Previous chapters (Chapters 3 and 4) described the construction of isogenic test strains GJ5604 (Wild-type), GJ6509 (rho) and GJ6511 (nusG) and the selection and customization of the bisulphite mutagenesis approach to test the hypothesis of R-loop occurrence in E. coli. A major problem with the approach so far employed was that the proportion of single-strandedness in the DNA region being probed was not quantifiable. Hence DGGE in conjunction with bisulphite treatment in the absence of denaturation was employed to quantitate the proportion of molecules that were converted in the final amplicon pool generated after bisulphite mutagenesis and PCR. These molecules gave an indirect estimate of the proportion of DNA templates that were single-stranded in the original DNA preparation.
at the loci tested.

The results of the PAGE and DGGE analyses of the amplicons derived with different primer pairs from the lacZ1, lacZ2, lacZ3 and lacY1 regions of the lac operon and the trpE1, trpE2 and trpD1 regions of the trp operon indicated that when equimolar amounts of input DNA (as indicated by the equivalent intensities of amplicons obtained with the N–N primer pairs) are analyzed for C-to-T conversions, the non-template strand is unpaired immediately downstream of the premature termination codon in both lacZ and trpE genes.

a. There is a funneling effect of the observed single-strandedness in the rho and nusG mutants

The intensity of single-strandedness immediately downstream of the premature translation stop codon is comparable in the wild-type and the rho and nusG mutants. This intensity was higher when obtained from a mutant allele (for e.g., trpE9777) than when obtained from the wild-type allele (for e.g., trpE'). The intensity gradually increases only in the rho and nusG mutants as the region being probed moves further away from the premature stop codon. This may indicate that transcription continues up to a certain point even after translation is terminated, thereby contributing to equal intensity of the retarded smear in the three strains immediately downstream of the premature termination codon owing to active transcription. Thereafter, Rho and NusG bring about termination of translation in the wild-type strain, whereas transcription continues well beyond the stop codon in the rho and nusG mutants. The observation that the difference in intensity of the retarded smear corresponding to single-stranded non-template strand DNA is least at the proximal end and greatest at the distal end of the lacZ and trpE genes implies that there is a funneling effect of defective transcription termination activity in the rho and nusG mutants, such that there is cumulative occurrence of single-strandedness at the 3' end of the genes being probed.

b. The signal corresponding to single-strandedness correlates inversely with translation

At the lacY1 and trpD1 regions that are immediately downstream of the lacZ and trpE genes respectively, the intensity of the amplicons obtained from the N–N and N–C
\text{nts} primer pairs are equivalent as also the intensities of the retarded smears on denaturing gradient gels, indicating that the extent of single-strandedness in these regions is the same in all the three strains. Two alternative explanations are likely for this observation; the simplest explanation is that that in the downstream ORF, ribosomes are loaded and translation occurs, coupled along with transcription. Occupancy of the mRNA by the ribosomes in the mutant strains may reduce the propensity of R-loop occurrence and therefore single-strandedness on the non-template strand, thereby equalizing the intensity of the signal observed in the three strains. De novo transcription of the lacY and trpD genes also occurs in all the three strains. Thus the transcription bubble may cause the non-template strand to be unpaired and sensitive to bisulphite, which is subsequently picked up as a signal in DGGE analysis. However, this explanation does not account for polarity, i.e., lacY is not transcribed in the wild-type strain and hence
the signal in the wild-type strain is not explained. Therefore bisulphite sensitivity in \textit{lacY} may represent a signal corresponding to the single-strandedness of the lagging strand template of replication which is also the non-template strand of transcription at this locus.

c. \textit{Low level of RNA-DNA hybrid occurrence was observed even on the annotated non-template strand of transcription}

Interestingly, signal corresponding to single-strandedness of the annotated template strand of transcription was also observed in all the three strains, \textit{rho} and \textit{nusG} being greater than the wild-type. This observation was made at the \textit{lacZ}, \textit{trpE} and \textit{trpD} genes but not in the \textit{lacY} gene. Of greater interest was the observation that in the \textit{trpE} gene, the \textit{trpE2 N}–\textit{C}ts amplicon gave a differential signal in the wild-type and mutant strains that was intensified in the \textit{trpE1 N}–\textit{C}ts lanes; i.e., when moved from 5’ to 3’ assuming transcription of the annotated non-template strand of the \textit{trpE} gene. This is similar to the funneling effect observed in the annotated ORF’s when analyzed from the 5’ end to 3’ end. This signal disappeared when the preparations were treated with RNase H prior to bisulphite treatment.

Interestingly, the observation that the annotated transcriptional template strand of the \textit{lacZ}, \textit{trpE} and \textit{trpD} but not the \textit{lacY} gene was single-stranded was substantiated when, as described in detail in Chapter 6, whole-genome sequencing analysis of the bisulphite treated DNA from wild-type and \textit{nusG} strains was performed and it was observed that high level of single-strandedness occurred on the template strand of \textit{lacZ}, \textit{trpE} and \textit{trpD} but lower level of single-strandedness occurred on the template strand of \textit{lacY}.

In conclusion, the results of the DGGE analyses suggested that R-loops occur at increased frequency in the \textit{rho} and \textit{nusG} mutants of \textit{E. coli}, at the \textit{lacZ} and \textit{trpE} loci bearing premature stop-codon mutations. Next, in order to assess the susceptibility of the entire chromosome of these mutants to R-loop formation, whole genome analysis of the bisulphite treated preparations was undertaken. Conclusive evidence was thus obtained for occurrence of R-loops all over the genome in a strain defective for Rho-dependent transcription termination. It was also observed for the first time that low levels of R-loops occur in the genome of even the wild-type strain that is proficient for Rho-dependent transcription termination. Interestingly it was observed that the \textit{rac} genes were deleted in more than 90% of cells in a growing culture of a mutant that is defective for Rho-dependent transcription termination. Clusters of heightened bisulphite sensitivity were observed in the strains indicating increased single-strandedness on one or both DNA strands. These clusters overlapped with nearly fifty genes identified to be targets of Rho-dependent transcription termination by an alternate approach (Peters \textit{et. al} 2009). Likewise, several clusters were also in or near the horizontally acquired non-essential gene regions (Posfai \textit{et. al}., 2006) that are silenced by Rho-dependent termination (Cardinale \textit{et. al}., 2008). The results of these experiments are described in detail in Chapter 6.