Nucleic Acids Research, 1994, Vol. 22, No. 23 4953–4957

Determination of the optimal aligned spacing between the Shine–Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs

Hongyun Chen*, Matthew Bjerknes[1], Ravindra Kumar and Ernest Jay
Department of Chemistry, University of New Brunswick, Fredericton, N.B. E3B 6E2 and
[1]Department of Anatomy and Cell Biology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Received August 21, 1994; Revised and Accepted October 20, 1994

ABSTRACT

The prokaryotic mRNA ribosome binding site (RBS) usually contains part or all of a polypurine domain UAAGGAGGU known as the Shine–Dalgarno (SD) sequence found just 5′ to the translation initiation codon. It is now clear that the SD sequence is important for identification of the translation initiation site on the mRNA by the ribosome, and that as a result, the spacing between the SD and the initiation codon strongly affects translational efficiency (1). It is not as clear, however, whether there is a unique optimal spacing. Complications involving the definition of the spacing as well as secondary structures have obscured matters. We thus undertook a systematic study by inserting two series of synthetic RBSs of varying spacing and SD sequence into a plasmid vector containing the chloramphenicol acetyltransferase gene. Care was taken not to introduce any secondary structure. Measurements of protein expression demonstrated an optimal aligned spacing of 5 nt for both series. Since aligned spacing corresponds naturally to the spacing between the 3′-end of the 16S rRNA and the P-site, we conclude that there is a unique optimal aligned SD–AUG spacing in the absence of other complicating issues.

INTRODUCTION

It is well known that the process of prokaryotic translation initiation involves binding of the 16S rRNA and the initiator tRNA to the mRNA ribosome binding site (RBS) on the mRNAs (1,2). The RBS generally extends 20 nucleotides (nt) on either side of the translation initiation codon (usually AUG) and contains, upstream from the AUG, part or all of a polypurine sequence (UAAGGAGGU) known as the Shine–Dalgarno (SD) sequence (1,3,4). The SD sequence is complementary to a pyridine tract (the anti-SD or ASD region) in the 3′-end of the 16S rRNA and its role in translation initiation is well documented (2; Figure 1). The spacing between the SD sequence and the initiation codon varies considerably in natural messages, with the average being 7 nucleotides (nt) (1,2). Excessively long or short spacing between the SD and the initiation codon may be detrimental to efficient translation initiation (5,6). Most investigators have found an optimal spacing between the SD and the AUG initiation codon in a variety of mRNAs, but confusingly, their estimates have ranged from 5 to 13 nt (Table 1).

Chen (7) and independently, Ringquist *et al.* (8) resurrected Dunn and Studier's (9) notion of what we term an aligned sequence spacing. Aligned spacing is based on the fact that during the formation of a translation initiation complex, ribosomal interaction with the mRNA occurs at two sites, the SD sequence and the initiation codon. The former interaction is mediated by the ASD region, while the latter interaction involves fMet-tRNA in the ribosomal P-site (1). When both interactions occur, a minimal SD–AUG spacing is required, presumably because the 16S rRNA and the fMet-tRNA must be kept a certain distance apart by configurational constraints.

For the sake of discussion, let us assume there is an optimal SD–AUG spacing corresponding to a preferred spacing between the ASD and the P-site of the ribosome. Then the SD–AUG spacing must be defined with respect to a specific nucleotide of the ASD. At this stage, the reference point is arbitrary, so we simply chose the 5′-A of the ASD (Figure 1). The complementary 3′-U of the complete SD 5′-UAAGGAGGU-3′ will be referred to as SD_{ref}.

If the SD of an mRNA contains SD_{ref} (Figure 1A,C), then the SD–AUG spacing will be the number of nucleotides separating the SD from the AUG. In contrast, if the SD does not contain SD_{ref}, then things become more confusing. We must now distinguish between 'spacing' and 'aligned spacing' (Figure 1B). Spacing is the number of nucleotides separating the 3′-end of the partial SD from the AUG, while the aligned spacing is the number of nucleotides separating the nucleotide corresponding to SD_{ref} from the adenine of the AUG (Figure 1B).

Most previous studies have used spacing rather than aligned spacing. Table 1 shows that the discrepancy between studies is reduced if their results are recast in terms of aligned spacing.

*To whom correspondence should be addressed at: Ontario Cancer Institute, Princess Margaret Hospital, 500 Sherbourne Street, Toronto, Ontario M4X 1K9, Canada

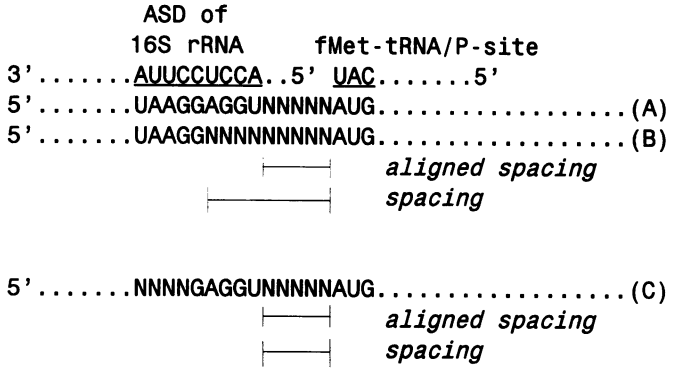

Figure 1. The prokaryotic ribosome binding site. (**A**) The RBS contains part or all of the SD sequence 5'-UAAGGAGGU and the initiation codon AUG. The SD sequence is complementary to the 3'-end of 16S rRNA (the ASD region, underlined). The anticodon UAC (underlined) of the fMet-tRNA is also shown. The 3'-U of the complete SD sequence is chosen as a reference point and named SD_{ref}. (**B**) The RBS with a partial SD UAAGG. The SD−AUG spacing is defined as the number of nucleotides separating the partial SD and the AUG. SD−AUG aligned spacing is defined as the number of the nucleotides separating the nucleotide corresponding to SD_{ref} from the Adenine of the AUG. (**C**) The RBS with a partial SD GAGGU. The SD−AUG spacing and the aligned spacing are the same because the partial SD contains SD_{ref}.

Table 1. [1]Optimal spacing and [2]aligned spacing between the SD and the initiation codon in different systems

SD	Spacing variation	Optimal spacing	Optimal aligned spacing	Reference
.AAGGG...	2−15	7	4	(23)
...GGA...	5−8	8	5	(29)
..AGGA...	8−34	8−10	5−7	(30)
.AAGG....	6−23	8−13	4−9	(31)
.AAGGA...	4−12	7	4	(8)
UAAGGAGG.	4−12	7−8	6−7	(33)
..AGGA...	7−17	7−9	4−6	(22)
.AAGGGU..	2−15	9	7	(20)
UAAGGAGGU*	3−11	5	5	(12)
.AAGGAGGU	3−11	7	7	(26)
..AGGA...	8−20	8−9	5−6	(32)
.AAGGGU..	7−13	8−9	6−7	(34)
.AAGG....	7−14	9	5	(35)

[1]The SD−AUG spacing is the number of nucleotides separating each SD sequence in the given mRNA from the Adenine of the AUG initiation codon. [2]The SD−AUG aligned spacing is the number of nucleotides separating the nucleotide corresponding to SD_{ref} (indicated by *) to the Adenine of the AUG.

Ringquist *et al.* (8) have made a similar observation in terms of natural mRNAs.

In this paper, we re-evaluate the existence of a unique optimal SD−AUG spacing using two series of chemically synthesized RBSs inserted into an expression vector system containing the chloramphenicol acetyltransferase (CAT) gene under the control of a weak promoter. The first series D_n contained the partial SD sequence GAGGU and varied the SD−AUG spacing from 3 to 15 nt. The second series d_n contained the SD sequence UAAGG and varied the spacing from 2 to 17 nt. Measurement of CAT activity was used to test for an optimal SD−AUG spacing in each series. Both series had a unique optimal aligned SD−AUG spacing of 5 nt.

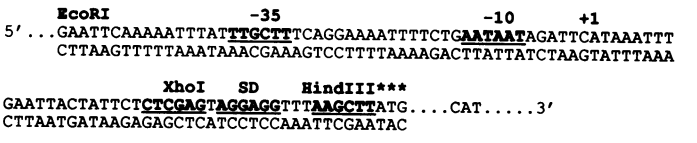

Figure 2. Structure of the expression plasmid vector. The expression vector is a derivative of pBR322 in which the *Eco*RI and the *Hin*dIII region has been replaced by a down-mutant bacteriophage T5P25 promoter P_2 (Rommens *et al.*, 1983) and a strong RBS (Cloney, 1988). The −35 and −10 regions of the P_2 promoter are underlined. The initiating nucleotide of transcription is indicated by (+1). The *Xho*I, *Hin*dIII sites and the SD sequence of the RBS are underlined. The translation initiation codon ATG of the CAT gene is indicated by *.

MATERIALS AND METHODS

Bacterial strain and media

E.coli K12 LE 392 strain, obtained from Dr G. Jay (NIH), was used as a host strain for all the recombinant plasmids. Bacteria were grown in LB broth with 100 μg/ml chloramphenicol (Cm).

Chemicals and enzymes

The four deoxyribonucleotides were purchased from US Biochemical. Restriction endonucleases, T4 DNA ligase, DNA polymerase I (large fragment) and mung-bean nuclease were from Bethesda Research Laboratories. Radiochemicals were from New England Nuclear Corp.

Expression vector

The expression vector (Figure 2) is a pBR322 derivative obtained by replacement of the *Eco*RI−*Hin*dIII fragment with a synthetic DNA fragment that contains: (i) a weak promoter P_2, a down mutant of promoter P_1 (bacteriophage T5 early promoter) in which a mutation converted the Pribnow box TATAAT sequence into a less conserved AATAAT sequence [(10); under the control of P_2, only a small amount of CAT is produced, avoiding the accumulation of high CAT levels which might otherwise distort the measurement of relative translational efficiencies in the variant plasmids and increase the chance of mutation]; (ii) a 43 nt precistronic length which allows optimal expression of the CAT gene [(11); this region is filled with mostly A/T in agreement with Scherer *et al.* (12), and Gold *et al.* (13), who found a preference for A/T in the 5' untranslated region of mRNAs except the SD sequence]; (iii) an *Xho*I site at the 5'-end of the SD and a *Hin*dIII site just prior to the initiation codon ATG which make it easy to replace the SD−ATG region with different synthetic DNA fragments; and (iv) the CAT gene [the CAT gene system was selected for this investigation because it is easy to select bacteria containing an active gene product and a sensitive assay for measuring CAT activity is available, (14)].

Oligonucleotide synthesis

All oligonucleotides were synthesized using an improved modified rapid solid phase phosphite approach and purified as described by Lutze-Wallace (15). The sequence of each synthetic oligonucleotide was confirmed by the Maxam and Gilbert procedure (16). Two series of synthetic RBSs, D_n and d_n, were designed and chemically synthesized to provide variable spacing between the SD sequence and the initiation codon ATG (Figure 3). Each of the RBSs contains the following features: (i) a five base SD sequence (TAAGG or GAGGT); (ii) a tetranucleotide

RBS	Sequence	SD-ATG Spacing	SD-ATG Aligned Spacing
A: SD-ATG spacing series D_n			
	SD		
D_3	TCGAT**GAGGT**CATATG	3	3
D_4	TCGAT**GAGGT**CCATATG	4	4
D_5	TCGAT**GAGGT**CACATATG	5	5
D_6	TCGAT**GAGGT**CTTCATATG	6	6
D_7	TCGAT**GAGGT**CATTCATATG	7	7
D_9	TCGAT**GAGGT**CATTATCATATG	9	9
D_{11}	TCGAT**GAGGT**CTTATTATCATATG	11	11
D_{13}	TCGAT**GAGGT**CTATTATTATCATATG	13	13
D_{15}	TCGAT**GAGGT**CATTATTATTATCATATG	15	15
B: SD-ATG spacing series d_n			
	SD		
d_2	TCGAT**TAAGG**CTATG	2	-2
d_3	TCGAT**TAAGG**CATATG	3	-1
d_5	TCGAT**TAAGG**CACATATG	5	1
d_7	TCGAT**TAAGG**CATTCATATG	7	3
d_8	TCGAT**TAAGG**CATTACATATG	8	4
d_9	TCGAT**TAAGG**CATTATCATATG	9	5
d_{10}	TCGAT**TAAGG**CTATTATCATATG	10	6
d_{11}	TCGAT**TAAGG**CTTATTATCATATG	11	7
d_{13}	TCGAT**TAAGG**CTATTATTATCATATG	13	9
d_{15}	TCGAT**TAAGG**CATTATTATTATCATATG	15	11
d_{17}	TCGAT**TAAGG**CATATTATTATTATCATATG	17	13
The complete SD sequence			
	5′...**TAAGGAGGT***.....ATG		

Figure 3. The nucleotide sequences of the RBS region of the two SD−ATG spacing series (D_n and d_n) between the partial *Xho*I site and the translation initiation codon ATG. The spacing between the SD and the ATG is varied by filling 0−15 A/T. Bold characters indicate the partial SD sequences. The SD−ATG spacing is the number of nucleotides between the partial SD in each series and the ATG while the SD−ATG aligned spacing is the number of nucleotides between SD_{ref} and the ATG. The complete SD sequence (SD_{ref} indicated by *) is also shown for reference.

sequence at its 5′ protruding terminus to direct its ligation at the corresponding *Xho*I site of the expression vector (in order to avoid any secondary structure, the *Xho*I site in the resulting plasmid was modified, **CTCGAG→T**); (iii) a blunt 3′ terminus ending with CA residues, generating a new *Nde*I site: **CATATG**; (iv) 0−15 A/T to vary the spacing between the SD and the ATG, ensuring that no secondary structure results, according to the procedures of Tinoco *et al.* (17).

Plasmid construction

Conditions for all the restriction endonuclease reactions were those recommended by the manufacturers. The expression vector was digested with *Hin*dIII and rendered flush by treatment with mung-bean nuclease. The SD−ATG region was removed from the resulting linear plasmid with *Xho*I. The larger fragment was electroeluted after separation on 1% agarose gel and the fragment was then ligated with the appropriate synthetic RBS. This resulted in two series of plasmids D_n (n = 3, 4, 5, 6, 7, 9, 11, 13, 15 where the n represents the size of the insert separating the SD from the ATG, i.e. the SD−ATG spacing) and d_n (n = 2, 3, 5, 7, 8, 9, 10, 11, 13, 15, 17). All ligation mixtures were used to transform competent *E.coli* LE 392 cells and the candidate clones determined to be Ap^rCm^r were prepared for further characterization by restriction mapping.

Preparation of cell lysates and CAT activity assay

An efficient lysozyme technique developed by Cloney (11) was used in the preparation of cell lysates for CAT activity assay. The harvested cells were resuspended in 1 ml of an ice cold TME buffer solution (38 mM Tris−HCl pH 7.8, 38 μM β-mercaptoethanol, 75 mM EDTA). To this, 10 μl of a freshly prepared lysozyme (10 mg/ml) was added, and the solution thoroughly mixed. After incubation at room temperature for 10 min, Triton X-100 was added to a final concentration of 0.1% and the sample was quickly mixed and put through three cycles of freezing and thawing in powdered dry ice. CAT activity was then assayed spectrophotometrically as described by Shaw (14).

RESULTS AND DISCUSSION

Cells harboring the plasmids were grown in LB broth containing 100 μl/ml Cm to 0.5 OD_{590} and aliquots of cells were harvested 3 h thereafter. The cell lysates were prepared using lysozyme and the amount of CAT directed by D_n and d_n was determined using a spectrophotometric method as described in Materials and Methods. The results are shown in Figure 4. The data are averages of at least three independent trials. Evidently, the optimal SD−ATG spacing in the first series D_n (with SD = GAGGT) was 5 nt while in the second series d_n (with SD = TAAGG), it was 9 nt (Figure 4A). The aligned optimal spacing, however, was 5 nt for both series (Figure 4B). See Introduction for the definition of aligned spacing.

It is plausible that the physical separation between the 3′-end of the 16S rRNA and the P-site of the ribosome determines a preferred spacing between the SD and the AUG of the mRNA. Presumably, this corresponds to the optimal aligned spacing of

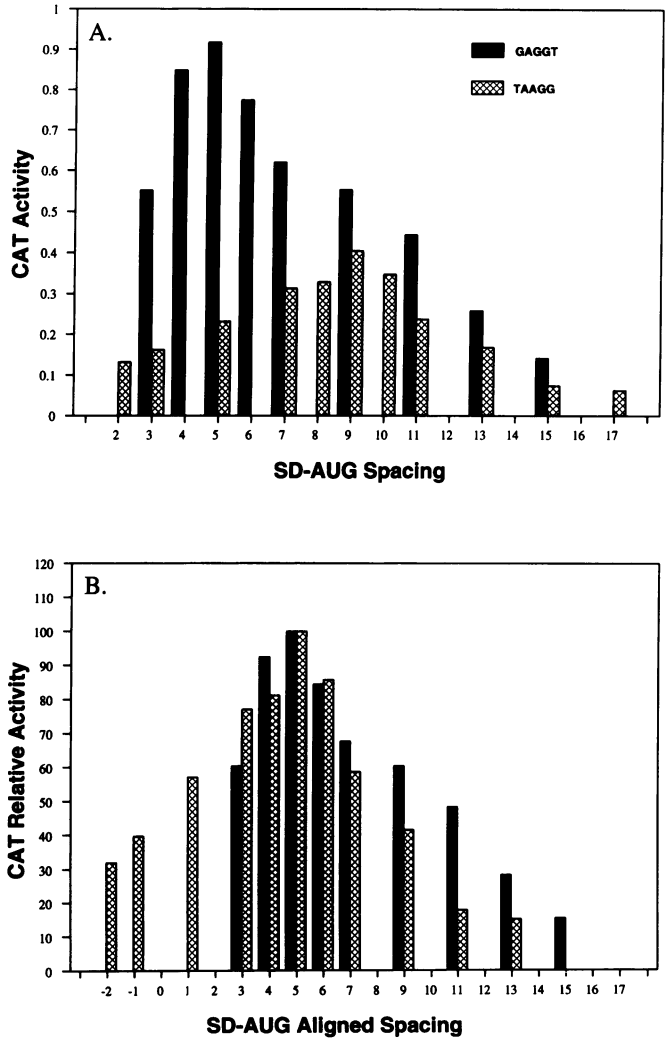

Figure 4. The effect of varying SD−AUG spacing on CAT activity. (**A**) Histogram showing the units of CAT activity (moles of Cm acetylated per minute, per A unit of cells) in cells harboring plasmids D_n and d_n as a function of SD−AUG spacing. The optimal SD−AUG spacing for D_n series is 5 nt while that for d_n series is 9 nt. The data presented in the histogram is the average of three trials. (**B**) Histogram showing the relative CAT activities (percentage of the maximal activity found in each series; data used are the same as in (A) as a function of SD−AUG aligned spacing. Note that in contrast to the presentation in (A), the optimal aligned spacing is the same, 5 nt, for both series.

5 nt we observed. Ringquist *et al.* (8) observed that natural mRNAs of *E.coli* had an aligned spacing of 5 nt. Scherer *et al.* (12) also suggested that 5 nt spacing was optimal with the SD UAAGGAGGU in natural *E.coli* mRNAs. Min *et al.* (18) introduced random mutagenesis within the RBS region and deduced an optimal aligned spacing of 5 nt. Inspecting the other results summarized in Table 1, one can see that the conclusion is similar, although the SD interaction might not strongly constrain the SD−AUG spacing because of other sequence dependent effects.

The presence of secondary structure involving the SD sequence or the AUG has been demonstrated to strongly affect the efficiency of translation initiation (1). When the SD−AUG spacing is manipulated, one of the main complications is the introduction of potential secondary structure involving the SD and/or the AUG due to sequence alterations. Such changes might account for some of the discrepancies in Table 1 (for example, 19). Therefore, we were careful not to introduce secondary structures into the RBS as estimated by the procedures of Tinocol *et al.* (17). The exclusion of secondary structure effects from our study allowed us to measure the effect of SD−AUG spacing in isolation, however, followup experiments should be initiated to investigate the details of potential interaction between SD−AUG spacing and RBS secondary structure as such interactions are presumbly important in many natural mRNAs.

A weakness of most of these works, including ours, is that we have implicitly assumed that as the aligned spacing is increased or decreased from the optimum, the measured change in CAT activity results largely from a change in the efficiency of translation initiation. Other potential effects such as changes in transcription efficiency, mRNA stability, etc, could not be ruled out, since we did not measure mRNA levels in our cells. Another important issue is the fact that the manner in which we altered the SD−AUG spacing also changed the precistronic length of the mRNA. The precistronic length increased from 35 to 50 nt as the SD−AUG spacing was altered from 2 to 17 nt. Some authors have pointed out that the precistronic length seems to have an effect on the efficiency of translation initiation (20,21). Cloney (11) also demonstrated an effect of varying the precistronic length from 33 to 49 nt on the level of CAT activity. More work needs to be done to establish the effect of the SD−AUG spacing on translational efficiency without the complication of varying precistronic length or secondary structure.

We also observed that the effect of varying the SD−AUG spacing on translational efficiency was significantly different from what has been obtained previously by Shepard *et al.* (19) and Singer *et al.* (6). These authors found that the insertion of even a single nucleotide into an optimally spaced SD and AUG decreased the level of expression by an order of magnitude. In the present study, however, the plasmids with 4 or 6 nt spacing retained about 80% efficiency (Figure 4). Even when the spacing was shortened to 2 nt or lengthened to 17 nt, about 15% of the maximum was still achieved (Figure 4). The reduced sensitivity of our constructs probably indicates the influence of other elements of the RBS, for example, the adenine-rich 3'-RBS (22). The CAT gene used in this study has a string of 7 A's in the downstream sequence which has been demonstrated to be an important element of the RBS (22). This adenine-rich domain might support efficient expression of the CAT gene thus reducing the need for an optimal SD−AUG spacing. Other elements of the RBS such as 6−10 A residues between the SD and the AUG, certain preferred triplets upstream from the AUG, or long SD sequences have also been demonstrated to allow more flexibility in the SD−AUG spacing (8,23,24).

There is about a two fold maximal variation between the two series of plasmids in the level of CAT activity (see Figure 4A). It is hard to explain the difference because these two series differ only in the SD sequence. One possibility is that the stability of GAGGU base pairing to the ASD may be higher than that of UAAGG, thus allowing a more efficient SD interaction, since the SD interaction energy of GAGGU ($\Delta G = -14.7$ Kcal) is higher than that of UAAGG ($\Delta G = -10.2$ Kcal). However, no direct correlation was found between the stability of the SD interaction and the efficiency of initiation complex formation (25). An alternative explanation is that GAGGU is more in the centre

of the SD interaction than UAAGG. The site of *E.coli* 16S rRNA most frequently involved in SD interaction was noted to be the sequence 5'-CUCC (26). Mutations in SD usually decrease the efficiency of translation initiation, but there is one mutant in which translation was enhanced, apparently by shifting the interaction from a less favourable region of 16S rRNA (5'-AUCAC) to the more favourable site 5'-CUCC (27). In another case in which the SD−ASD complementarity was reduced from 8 to 4 nt without significantly affecting the translational efficiency, the 4 base complementarity was in the most favourable position (GAGG) for interaction with 16S rRNA (28). Thus, as proposed by Kozak (2), the efficiency of translation is likely to be higher when the SD interaction involves the most favourable region of 16S rRNA (i.e., 5'-CUCC) than when it involves off centre sequences.

In conclusion, we confirm that SD−AUG spacing plays a significant role in the process of translation initiation and present evidence that aligned spacing is the most appropriate measure of spacing. An optimal aligned spacing of 5 nt probably corresponds to the preferred spacing between the ASD region of the 16S rRNA and the fMet-tRNA.

ACKNOWLEDGEMENTS

This project was supported in part by grants from the National Cancer Institute of Canada and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Gold, L. (1988) *Ann. Rev. Biochem.* **57**, 199−233.
2. Kozak, M. (1983) *Microbiol. Rev.* **47**, 1−45.
3. Steitz, J. A. (1969) *Nature* **224**, 957−967.
4. Shine, J. and Dalgarno, L. (1974) *Proc. Natl Acad. Sci. USA*. 71, 1342−1346.
5. Roberts, T. M., Bikel, I., Yocum, R.R., Livinstone, D. & Ptashne, M. (1979) *Proc. Natl Acad. Sci. USA* **76**, 5596−5600.
6. Singer, B. S., Gold, L., Shinedling, S. T., Colkitt, M., Hunter, L. R., Pribnow, D. & Nelson, M. A. (1981) *J. Mol. Biol.* **149**, 405−432.
7. Chen, H. (1990) Ph. D. dissertation, University of New Brunswick, Fredericton, New Brunswick, Canada.
8. Ringquist, S., Shinedling, S. Barrick, D., Green, L., Binkley, J., Stormo, G.D. & Gold, L. (1992) *Molecular Microbiology* **6**, 1219−1229.
9. Dunn, J. & Studier, F. (1981) *J. Mol. Biol.* 148, 303−330.
10. Rommens, J., MacKnight, D., Pomeroy-Cloney, L. & Jay, E. (1983) *Nucleic Acids Res.* **11**, 5921−5940.
11. Pomeroy-Cloney, L. (1988) Ph.D dissertation, University of New Brunswick, Fredericton, New Brunswick, Canada.
12. Scherer, G. F. E., Walkinshaw, M. D., Arnott, S. and Morré, D. J. (1980) *Nucleic Acids Res.* **8**, 3895−3907.
13. Gold, L., Pribnow, D., Schneider, T., Shindeling, S., Singer, B. S. & Stormo, G. (1981) *Ann. Rev. Microbiol.* **35**, 365−403.
14. Shaw, W. V. (1975) *Method Enzymol.* (Ed. Hash, J. H.) **43**, 737−755.
15. Lutze-Wallace, C. (1988) Ph.D. Dissertation, University of New Brunswick, Fredericton, N.B., Canada.
16. Maxam, A. M. & Gilbert, W. (1977). *Proc. Natl Acad. Sci. USA* **74**, 560−564.
17. Tinoco, I., Borer, P., Dengler, B. Levine, M. D., Uhlenbech, O. C., Crothers, D. M., and Gralla, J. (1973) *Nature New Biol.* **246**, 40−41−41.
18. Min, K. T., Kim, M. H. & Lee, D.-S. (1989) *Nucleic Acids Res.* **16**, 5075−5088.
19. Shepard, H. M., Yelverton, E. & Goeddel, D. V. (1982) *DNA* **1**, 25−31.
20. Stassens., P., Remaut, E. and Fiers, W. (1985) *Gene* **36**, 211−223.
21. Roberts, T. M., Kachich, R. & Ptashne, M. (1979) *Proc. Natl Acad. Sci. USA* **76**, 760−764.
22. Chen, H., Pomeroy-Cloney, L., Bjerknes, M., Tam, J. and Jay, E. (1994) *J. Mol. Biol.* **240**, 20−27.
23. Curry, K. A. & Tomich, C.-S. C. (1988) *DNA* **7**, 173−179.
24. Dalbøge, H., Carlsen, S., Jensen, E. G., Christensen, T. & Dahl, H.-H. (1988) *DNA* **7**, 399−405.
25. Gren, E. J. (1984) *Biochemie* **66**, 1−29.
26. Stormo, G. D., Schneider, T. D. & Gold, L. M. (1982) *Nucleic Acids Res.* **10**, 2971−2996.
27. Chapon, C. (1982) *EMBO* **1**, 369−374.
28. Gillam, S., Astell, C. R. & Smith, M. (1980) *Gene* **12**, 129−137.
29. Gheyson, D., Iserentant, D., Deron, C. & Fiers, W. (1982) *Gene* **17**, 55−63.
30. Guarente, L., Lauer, G., Roberts, T. M. and Ptashne, M. (1980) *Cell* **20**, 543−553.
31. Itoh, S., Mizukami, T., Matsumoto, T., Nishi, T., Saito, A., Oka, T., Furnya, A., Tak aoka, C. & Taniguichi, T. (1984) *DNA* 3, 157−165.
32. Thummer, C. S., Burgers, T. C. & Tjean, R. (1981) *J. Virol.* **37**, 683−697.
33. Vellanoweth, R. L. and Rabinowitz, J. C. (1992) *Molecular Microbiology* **6**, 1105−1114.
34. Whitehorn, E.A., Livak, K.J. & Petteway, S.R., Jr. (1985) *Gene* **36**, 375−379.
35. Wood, C.R., Boss, M.A., Patel, T.P. and Emtage, J.S. (1984) *Nucleic Acids Res.* **12**, 3937−3948.