# การศึกษาลำดับเบสเริ่มแปลรหัส Non-Shine-Dalgarno/ Shine-Dalgarno ที่มีผลต่อการแสดงออกของยีนรายงานผลในไซยาโนแบคทีเรีย *Synechococcus* PCC7942

# Investigation of Non-Shine-Dalgarno/ Shine-Dalgarno Translational Initiators That Affect Reporter Gene Expression in Cyanobacterium *Synechococcus* PCC7942

**นวพล อุดพ้วย**[1] และวิภา จึงจตุพรชัย[1]

Nawapol Udpuay[1] and Wipa Chungjatupornchai[1]

## บทคัดย่อ

ในการศึกษานี้ได้สำรวจตัวเริ่มแปลรหัสที่เหมาะที่สุด (optimal translational initiators) ในไซยาโนแบคทีเรีย *Synechococcus* PCC7942 โดยใช้ shuttle vector สร้างไลบรารี่สังเคราะห์ที่มีลำดับเบส non-SD/ SD และมียีนรายงาน green fluorescent protein (*gfp*) แล้ว transform เข้า *Synechococcus* ได้ตรวจคัดกรองโดยวัดค่า GFP activity ของ 1,500 โคลนจากไลบรารี่ pSD-TNG2 และ 650 โคลนจากไลบรารี่ pSD-TNG3 แล้วคัดเลือกโคลนมาวิเคราะห์ลำดับเบส non-SD/ SD ผลการทดลองพบว่านิวคลีโอไทด์บน 5'UTR ที่อยู่ upstream ของตัวเริ่มแปลรหัสอาจมีผลกระทบต่อการแสดงออกของยีน *gfp* ใน *Synechococcus* โคลน pSD-TNG2-831, pSD-TNG2-357 และ pSD-TNG2-1459 ซึ่งมีลำดับเบสคล้ายกับ SD คือ AGGAGAATGA, AAGGAACATA และ AAGAACAATG มีค่า GFP activity สูงกว่าพลาสมิดที่ใช้อ้างอิงอยู่ 3.42, 1.97 และ 1.25 เท่าตามลำดับ ดังนั้นผลการทดลองบ่งชี้ว่าลำดับเบสที่คล้าย SD มีประสิทธิภาพในการเริ่มแปลรหัส ส่วนโคลน pSD-TNG3-462, pSD-TNG3-493, pSD-TNG3-552 and pSD-TNG3-556 ซึ่งมีลำดับเบสเป็น non-SD มีค่า GFP activity ต่ำกว่าพลาสมิดที่ใช้อ้างอิง แต่อย่างไรก็ตามโคลน pSD-TNG3-423 ซึ่งลำดับเบสเป็น non-SD คือ AACAGGAAAA กลับมีค่า GFP activity สูงกว่าพลาสมิดที่ใช้อ้างอิงอยู่ 1.42 เท่า ดังนั้นลำดับเบสที่เป็น non-SD น่าจะสามารถเป็นตัวเริ่มแปลรหัสที่มีประสิทธิภาพได้ ขณะนี้กำลังสำรวจและวิเคราะห์ประสิทธิภาพของตัวเริ่มแปลรหัสที่เป็น non-SD ของไลบรารี่ใน *Synechococcus*

**คำสำคัญ** : ลำดับเบส Shine-Dalgarno (SD), ลำดับเบส non-Shine-Dalgarno (non-SD), green fluorescence protein (GFP), ตัวเริ่มแปลรหัส (translational initiator)

## ABSTRACT

In this study, in order to investigate the optimal translational initiators in cyanobacterium *Synechococcus* PCC7942, the synthetic non-SD/ SD sequence libraries containing the green fluorescent protein (*gfp*) gene as a reporter were constructed in a shuttle vector and transformed into *Synechococcus*. A total of 1,500 and 650 clones from libraries pSD-TNG2 and pSD-TNG3, respectively, were screened for GFP activities. The non-SD/ SD sequences in the selected clones were analyzed.

[1]สถาบันชีววิทยาศาสตร์โมเลกุล มหาวิทยาลัยมหิดล วิทยาเขตศาลายา อ. พุทธมณฑล จ. นครปฐม 73170

Institute of Molecular Biosciences, Mahidol University, Salaya Campus, Phuttamonthon, Nakhon Pathom 73170

Results showed that the nucleotides in 5'UTR located at upstream of translational initiator may affect the expression of *gfp* gene in *Synechococcus*. The GFP activities of clones pSD-TNG2-831, pSD-TNG2-357 and pSD-TNG2-1459 containing the SD-like decanucleotides AGGAGAATGA, AAGGAACATA and AAGAACAATG were 3.42-, 1.97- and 1.25-fold, respectively, higher than that of corresponding control. Therefore, the results indicated that SD-like sequence is highly efficient for translation initiation. The GFP activities of clones pSD-TNG3-462, pSD-TNG3-493, pSD-TNG3-552 and pSD-TNG3-556 containing non-SD decanucleotides were lower than that of corresponding control. However, the GFP activity of clone pSD-TNG3-423 also containing the non-SD decanucleotides AACAGGAAAA was 1.42-fold higher than that of corresponding control. Thus, the non-SD decanucleotides could be the efficient translational initiator in cyanobacteria. The efficiency of non-SD translational initiators in *Synechococcus* harboring non-SD decanucleotide libraries is currently under investigation.

**Keywords** : Shine-Dalgarno (SD) sequence, non-Shine-Dalgarno (non-SD) sequences, green fluorescence protein (GFP), translational initiator

E-mail : Nawapol_ud@hotmail.com

## INTRODUCTION

The Shine-Dalgarno (SD) sequence consists of three to nine contiguous consensus-purine-rich nucleotides located at 5' untranslated region (5' UTR) of prokaryotic mRNA and has a role in pairing with 16S ribosomal RNA to initiate the translation. The efficiency of translation initiation affects the level of gene expression in prokaryote (De Smit and Van Duin, 1994). The non-Shine-Dalgarno (non-SD) sequences are not consensus-purine-rich nucleotides that involve in translation initiation. However, the exact mechanism of translation initiation of non-SD sequences remains unclear. Recently, the SD-sequence analysis in 162 completed prokaryotic genomes reveals that a considerable number of prokaryotic genes do not have SD sequences, suggesting that non-SD-led genes, including non-SD-led and leaderless transcripts, are as common as SD-led genes (Chang *et al*., 2006). It has been shown that in *E. coli,* the expression of chloramphenicol acetyltransferase gene with some non-SD (AT-rich decanucleotides) translational initiators was higher than that with SD translational initiator (Kolev *et al*., 2003).

Cyanobacteria (blue-green algae) are oxygenic photosynthetic prokaryotes. They have been used as hosts to express several heterologous genes. For example, attempt has been made to express the organophosphorus hydrolase of *Flavabacterium* sp. in cyanobacterium *Synechococcus* PCC7942, in order to detoxify the organophosphates (Chungjatupornchai and Fa-aroonsawat, 2008). However, the level of heterologous gene expression in cyanobacteria is lower when compared with that in *E. coli*.

Current knowledge of the relationship between nucleotide sequences and function of ribosome binding site for translation initiation recognized within cyanobacteria is still limited.

In this study, we have used the green fluorescent protein (GFP) as a reporter to investigate the optimal translational initiators in cyanobacterium *Synechococcus* PCC7942. The synthetic non-SD/ SD sequence libraries were constructed in a shuttle vector and transformed into *Synechococcus*. The GFP activities of *Synechococcus* transformants were screened. The non-SD/ SD sequences in the selected clones were analyzed.

## MATERIALS AND METHODS

### Strains and growth conditions

The *Synechococcus* PCC7942 strain R2-Spc (hereafter, referred to as *Synechococcus*) (Kuhlermeier *et al*., 1983) was grown on BG-11 medium containing 1.5% agar at 30ºC under constant illumination at 4,000 – 4,500 lux. *E. coli* strain DH5$\alpha$ was grown in LB medium.

### Construction of synthetic non-SD/ SD sequence libraries

The gene cassette, *X-gfp-*(*nos-ter*), was amplified using pUC18-P-GFP-NB as template with primer sets: SD-F2 and Nos-R3; SD-F4 and Nos-R3 (see Table 1). The resulting PCR products were digested with *Bam*HI/*Eco*RI or *Bgl*II/*Eco*RI and cloned into the *Bam*HI/*Eco*RI sites of shuttle vector pKTN-T to obtain libraries, pSD-TNG2 and pSD-TNG3, respectively (Figure 1). The plasmid libraries extracted from pooled *E. coli* transformants were transformed into *Synechococcus* as described (Kuhlermeier *et al*., 1983).

**Table 1** Primers used to construct the non-SD/ SD libraries.

| Primers | Sequence (5'-3') | Restriction site | Target sequence |
|---|---|---|---|
| SD-F2 | CGGGATCCCXXXXXXXXXXTCCCTT**ATG**AGTAAAGGAGAAG | *Bam*HI | *gfp* |
| SD-F4 | CGAGATCTCXXXXXXXXXXTCCCTT**ATG**AGTAAAGGAGAAG | *Bgl*II | *gfp* |
| Nos-R3 | CGGAATTCCCGATCTAGTAACATAGATGACACC | *Eco*RI | *Nos* terminator |

The start codon (ATG) of *gfp* gene is indicated. The randomized nucleotides A, T, C and G are marked by X.

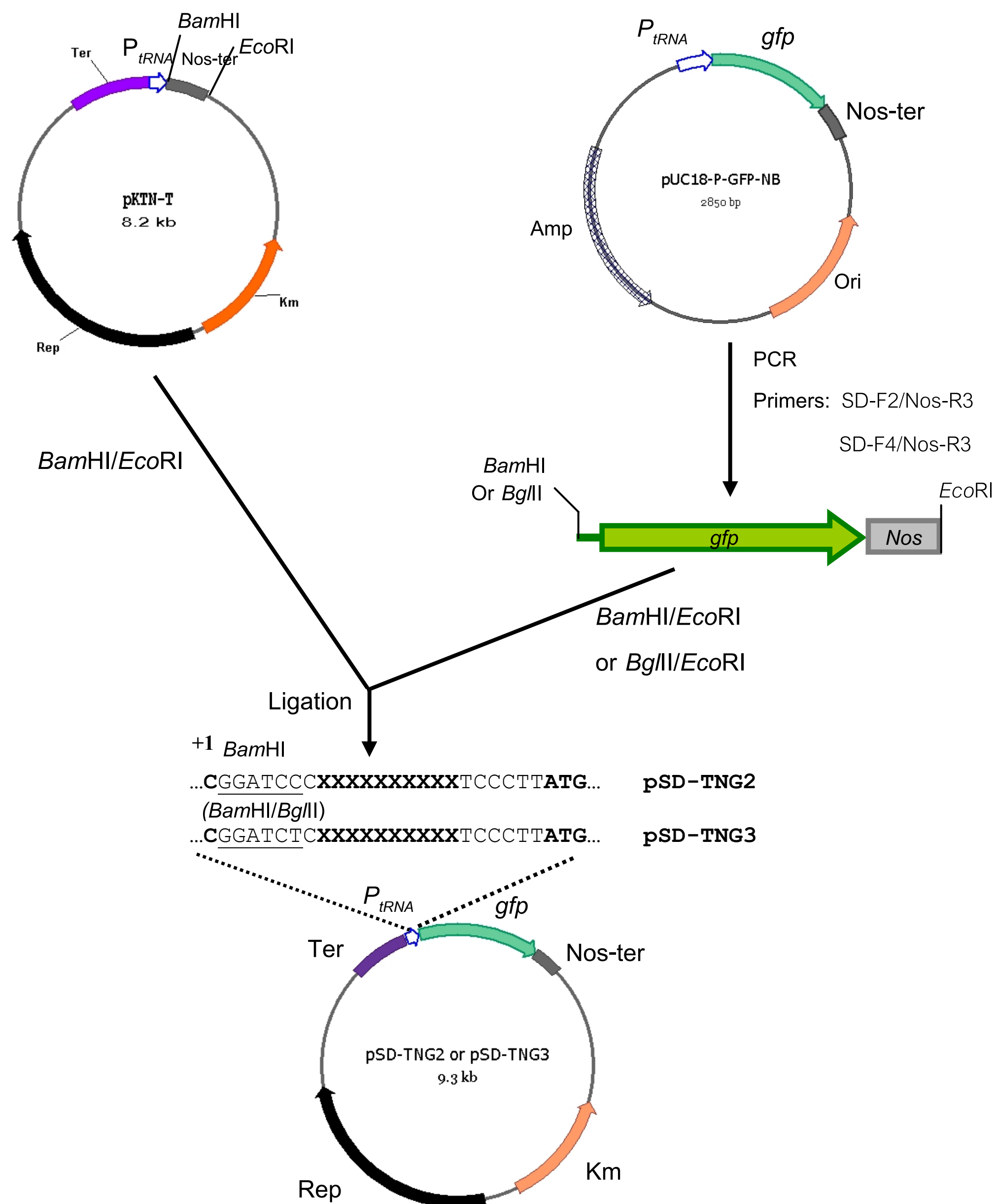


Figure 1 Schematic diagram for construction of pSD-TNG2 and pSD-TNG3 libraries

The PCR products containing gene cassette, X-*gfp*-(*Nos*-ter), were cloned into a shuttle vector pKTN-T. The random nucleotides A, T, C and G are marked by X. The transcription start site (+1) and start codon (ATG) of *gfp* mRNA are indicated.

## GFP fluorescence analysis

The *Synechococcus* cells were harvested and resuspended in phosphate-buffered saline (PBS). The fluorescence intensity of *Synechococcus* whole cells (in 96-well microplates) were measured using spectrofluorometer (Multimode Detector DTX 880, Beckman) with excitation at 485 nm and emission at 535 nm. The specific GFP activity of *Synechococcus* whole cells was defined as the fluorescence intensity divided by the optical density measured at 730 nm.

**Nucleotide sequence analysis**

To rescue the recombinant plasmids, the total DNA of *Synechococcus* was extracted and transformed into *E. coli*. The DNA sequences of the ten randomized nucleotides in the recombinant plasmids extracted from *E. coli* transformants were determined by an automated sequence analyzer.

## RESULTS AND DISCUSSION

**Construction of plasmids harboring synthetic non-SD/ SD sequences**

Plasmids harboring *gfp* gene used in this study were shown in Figure 2. Plasmids pKT-GFP, pSD-TBG2 and pSD-TBG3 were used as controls. The *gfp* gene in all constructs was under the control of cyanobacterial $tRNA^{pro}$ promoter (Chungjatupornchai *et al*, 1999; Chungjatupornchai *et al*, 2002) enabling constitutive expression of GFP (Figure 2).

The *Bam*HI/*Eco*RI-digested or *Bgl*II/*Eco*RI-digested PCR product containing gene cassette *X-gfp*-(*nos-ter*) was cloned into the *Bam*HI/*Eco*RI sites of shuttle vector pKTN-T to obtain plasmid libraries pSD-TNG2 and pSD-TNG3, respectively (Figure 1). Therefore, the *Bam*HI site was retained in plasmid library pSD-TNG2 but lost in pSD-TNG3. The plasmid library pSD-TNG3 extracted from *E. coli* was digested with *Bam*HI in order to get rid of background vector prior to be transformed into *Synechococcus*. The gene cassette *X-gfp*-(*nos-ter*) in plasmid libraries pSD-TNG2 and pSD-TNG3 contained ten randomized nucleotides located at the 7th nucleotide upstream of initiation codon of *gfp* gene which was used as a reporter (Figure 1 and 2).

**GFP activities of *Synechococcus* harboring plasmid libraries pSD-TNG2 and pSD-TNG3**

A total of 1,500 and 650 clones from *Synechococcus* libraries of pSD-TNG2 and pSD-TNG3, respectively, were screened for GFP activities. The GFP activities of the selected clones were shown in Figure 3. The GFP activities of *Synechococcus* harboring pSD-TNG2-357, pSD-TNG2-831 and pSD-TNG3-423 were significantly higher than those of pKT-GFP, pSD-TBG2 and pSD-TBG3 used as controls (Figure 3). The GFP activity of *Synechococcus* harboring pSD-TBG2 was slightly higher than that of pKT-GFP (Figure 3). The GFP activity of *Synechococcus* harboring pSD-TBG3 was significantly higher than that of pSD-TBG2. The plasmids pSD-TBG2 and pSD-TBG3 contained almost identical 5’UTR, except the nucleotide at position +7 (Table 2). Therefore, the results indicated that nucleotides upstream of the translational initiator may affect the expression of *gfp* gene in *Synechococcus.*

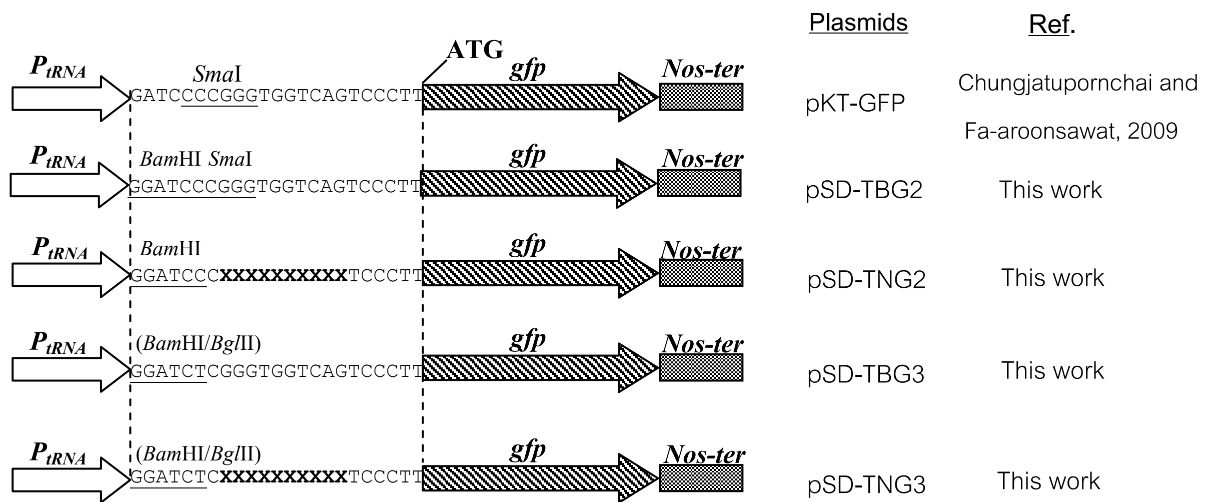


**Figure 2** Plasmids harboring the *gfp* gene.

Plasmids pKT-GFP, pSD-TBG2 and pSD-TBG3 were used as controls. The randomized nucleotides ( A, T, C and G) of plasmid libraries pSD-TNG2 and pSD-TNG3 are marked by X. The start codon (ATG) of *gfp* mRNA are indicated. $P_{tRNA}$, $tRNA^{pro}$ promoter; *gfp*, green fluorescence protein gene; *Nos-ter*, nopaline synthase terminator.

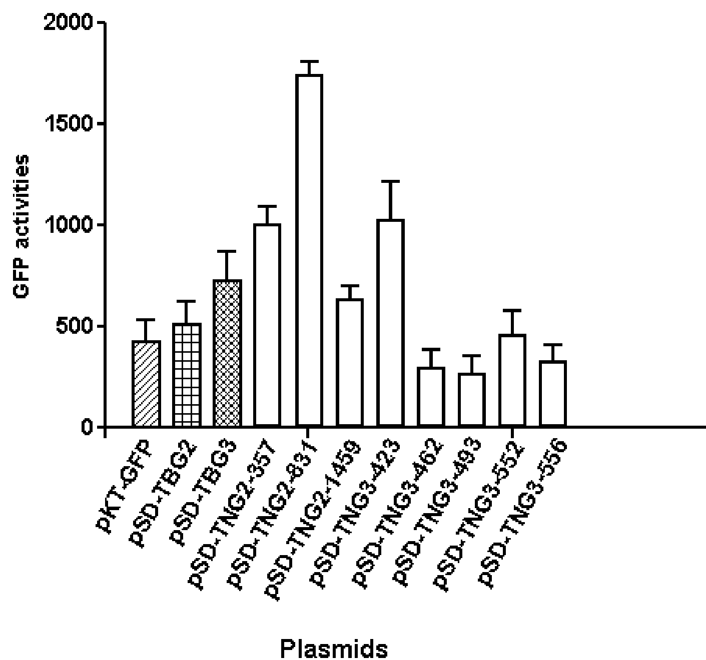


**Figure 3** GFP activities of selected *Synechococcus* clones.

GFP activities of recombinant *Synechococcus* were determined. *Synechococcus* harboring plasmids pKT-GFP, pSD-TBG2 and pSD-TBG3 were used as controls. The selected clones from libraries pSD-TNG2 and pSD-TNG3 are indicated. Each value and error bar represents the means of three independent experiments and its standard deviation.

Table 2 showed that the GFP activities of clones pSD-TNG2-831, pSD-TNG2-357 and pSD-TNG2-1459 containing the SD-like decanucleotides AGGAGAATGA, AAGGAACATA and AAGAACAATG were 3.42-, 1.97- and 1.25-fold, respectively, higher than that of corresponding control. Therefore, the results indicated that SD-like sequence is highly efficient for translation initiation. The GFP activities of clones pSD-TNG3-462, pSD-TNG3-493, pSD-TNG3-552 and pSD-TNG3-556 containing non-SD decanucleotides were lower than that of corresponding control. However, the GFP activity of clone pSD-TNG3-423 also containing the non-SD decanucleotides AACAGGAAAA was 1.42-fold higher than corresponding control. Thus, the non-SD decanucleotides could be the efficient translational initiator in cyanobacteria. According to Kolev *et al.*, 2003, in *E. coli*, some non-SD (AT-rich decanucleotides) translational initiators are more efficient than SD translational initiator.

**Table 2** Expression of *gfp* gene in *Synechococcus* under the control of random (non-SD/ SD) translational initiators.

| Plasmids[a] | Sequence[b] | GFP activities (%)[c] | | |
|---|---|---|---|---|
| Consensus SD sequence | **AAGGAGGTTT** | | | |
| pKT-GFP | cgatccccGG**G**TG**G**TCAGtcccttATG | 100 | | |
| pSD-TBG2 | cggatcccGG**G**TG**G**TCAGtcccttATG | 118 | 100 | |
| pSD-TBG3 | cggatctcGG**G**TG**G**TCAGtcccttATG | 168 | 142 | 100 |
| pSD-TNG2-357 | cggatccc**AAGGA**ACATAtcccttATG | 233 | 197 | 139 |
| pSD-TNG2-831 | cggatccc**A**G**G**AGAA**T**GAtcccttATG | 404 | 342 | 240 |
| pSD-TNG2-1459 | cggatccc**AAG**A**A**CAATGtcccttATG | 148 | 125 | 88 |
| pSD-TNG3-423 | cggatctc**AA**CAG**G**AAAAtcccttATG | 240 | 203 | 142 |
| pSD-TNG3-462 | cggatctcT**A**AC**A**AAAAAtcccttATG | 70 | 60 | 42 |
| pSD-TNG3-493 | cggatctc**AA**AA**A**ACAAAtcccttATG | 62 | 53 | 37 |
| pSD-TNG3-552 | cggatctc**AAGG**CCCAAAtcccttATG | 107 | 91 | 64 |
| pSD-TNG3-556 | cggatctcGG**G**A**A**AAGCAtcccttATG | 77 | 65 | 46 |

[a:] Plasmids pSD-TBG2 and pSD-TBG3 are controls of libraries pSD-TNG2 and pSD-TNG3, respectively.

[b:] The first nucleotide is the transcription start site (+1). The ten randomized nucleotides identical to consensus SD sequence are in bold. The start codon (ATG) of *gfp* gene is indicated. The *Bam*HI site is underlined.

[c:] The GFP activities were normalized to that of control (pKT-GFP, pSD-TBG2 and pSD-TBG3).

## CONCLUSIONS

The nucleotides in 5'UTR located at upstream of translational initiator may affect the expression of *gfp* gene in *Synechococcus.* The SD-like sequence is highly efficient for translation initiation in *Synechococcus*. However, the non-SD decanucleotides could be the efficient translational initiator in cyanobacteria. According to Kolev *et al.*, 2003, in *E. coli*, some non-SD (AT-rich decanucleotides) translational initiators are more efficient than SD translational initiator. The efficiency of non-SD

translational initiators in *Synechococcus* harboring non-SD decanucleotide libraries is currently under investigation.

## ACKNOWLEDGMENTS

We thank Sirirat Fa-aroonsawat for technical assistance. This work was supported by The Thailand Research Fund (TRF). Nawapol Udpuay is supported by The 60$^{th}$ Year Supreme Reign of his Majesty King Bhumibol Adulydej Scholarship of the Faculty of Graduate Studies, Mahidol University. Wipa Chungjatupornchai is a recipient of TRF.

## REFERENCES

Chang, B., S. Halgamuge, and S-L. Tang. 2006. Analysis of SD sequences in completed genomes: Non-SD-led genes are as common as SD-led gene. *Gene*. 373: 90-99.

Chungjatupornchai, W., and S. Fa-aroonsawat. 2009. Translocation of green fluorescent protein to cyanobacterial periplasm using ice nucleation protein. *The Journal of Microbiology.* 47(2): 187-192.

Chungjatupornchai, W., and S. Fa-aroonsawat. 2008. Biodegradation of organophosphate pesticide using recombinant cyanobacteria with surface- and intracellular-expressed organophosphorus hydrolase. *Journal of Microbiology and Biotechnology.* 18(5): 46-51.

Chungjatupornchai, W., S. Fa-aroonsawat, and S. Panyim. 2002. Characterization of regions of the cyanobacterial tRNA$^{pro}$ gene that affect the expression of a beta-glucuronidase reporter gene. *FEMS Microbiology Letters*. 211: 57-64.

Chungjatupornchai, W., T. Senawong, and S. Panyim. 1999. Isolation and characterization of *Synechocccus* PCC 7942 promoters: tRNA$^{pro}$ gene function as a promoter. *Current Microbiology*. 38: 210-216.

De Smit, M.H., and J. Van Duin. 1994. Translational initiation on structure massagers: Another role for the Shine-Dalgardo interaction. *Journal of Molecular Biology*. 235: 173-184.

Kolev, V., I. Ivanov, A. Berzal-Herranz, and I.G. Ivanov. 2003. Non-Shine-Dalgarno initiators of translation selected from combinatorial DNA libraries. *Journal of Molecular Microbiology and Biotechnology*. 5: 154–160.

Kuhlemeier, C., A.A. Thomas, A. van der Ende, R.W. van Leen, W.E. Borrias, C.A. van den Hondel and G.A. van Arkel. 1983. A host-vector system for gene cloning in the cyanobacterium *Anacystis nidulans* R2. *Plasmid.* 10: 156-163.