Absence of Effect of Varying Thr-Leu Codon Pairs on Protein Synthesis in a T7 System[†]

Li Cheng[‡] and Emanuel Goldman*[,‡,§]

Graduate School of Biomedical Sciences, New Jersey Medical School, Department of Microbiology & Molecular Genetics, University of Medicine and Dentistry of New Jersey, 185 South Orange Avenue, Newark, New Jersey 07103

Received February 5, 2001; Revised Manuscript Received March 20, 2001

ABSTRACT: The over-represented threonine-leucine (Thr-Leu) codon pair ACG CUG has been previously reported to be inhibitory to translation compared to the synonymous under-represented Thr-Leu codon pair ACC CUG, in an *E. coli* system in which the codon pairs were located either 3 and 4, or 6 and 7, or 9 and 10 codons downstream from the initiating codon for the message [Irwin, B., Heck, J. D., and Hatfield, G. W. (1995) *J. Biol. Chem. 270*, 22801−22806]. In the work reported here, these synonymous codon pairs were tested in a T7 system, with the codon pairs located either 14 and 15, or 6 and 7 codons downstream from the AUG start codon. In contrast to the reported findings in the *E. coli* system, there was no difference found in translation between mRNAs containing the respective codon pairs in the T7 system. The reasons for the different findings remain unclear, but presumably are a consequence of differences between the *E. coli* and T7 systems used to assay gene expression. Nevertheless, as a result of this work, it appears that the effect of varying codon pairs reported in the *E. coli* system is not due to a difference in translational step times through the respective codon pairs, as previously proposed.

Regulation at the level of elongation of translation has been increasingly appreciated in recent years, with the discovery of translational "recoding" mechanisms (*1*), and strong correlations of codon bias with efficiency of expression (*2, 3*). Aside from the bias in use of codon synonyms, there has also been observed a bias in codon pair utilization in *E. coli* and other organisms (*4, 5*). This is believed to reflect favored or disfavored interactions of adjacent tRNAs on the translating ribosome (*6*), and provides a plausible explanation for effects of "codon context" on translation efficiency (*7*).

In an experimental test comparing translation efficiencies of mRNAs containing two synonymous pairs of codons, a threonine-leucine (Thr-Leu)[1] sequence encoded by the over-represented codon pair ACG CUG was ∼10-fold lower in expression of a β-galactosidase reporter compared to Thr-Leu encoded by the under-represented codon pair ACC CUG (*8*). These measurements were made with the codon pairs at codons 3 and 4 following the translation start, and the authors indicated that similar results were obtained when the respective codon pairs were placed at codons 6 and 7. Also tested in this work, with similar results, were comparisons of the two Thr-Leu codon pairs at codons 9 and 10 in the leader peptide-encoding mRNA of the *trp* operon attenuator, in a transcriptional fusion to a reporter β-galactosidase (*8*). Neither of the two threonine codons, nor the leucine codon, comprising these codon pairs is individually rare in *E. coli*. Explanations other than varying translational efficiency through the respective codon pairs appeared to be excluded (*8*).

Although the authors interpreted their assays as reflecting secondary interference with translation initiation, the order of magnitude difference in expression levels raised the possibility that the inhibitory codon pair may in fact be rate-limiting for translation, and not merely responsible for transiently obscuring (by paused ribosomes) the translation initiation region. If so, the inhibitory codon pair would be predicted to show inhibition when placed downstream in the mRNA as well, unless this inhibition is also subject to the positional effects of 5′-translational blockage, which we (*9, 10*) and others (*11, 12*) have observed with inhibition of expression by consecutive rare codons.

The 1024-codon gene for wild-type β-galactosidase contains no ACC CUG in-frame codon pairs, and two ACG CUG codon pairs at codons 322−323 and 401−402, which already leads to the expectation that the inhibition is only manifested when this codon pair is located near the 5′ end. Alternatively, expression might be even higher if the downstream over-represented Thr-Leu codon pairs were changed.

We were interested in determining whether inhibition by the over-represented codon pair ACG CUG was subject to 5′-translational blockage, that is, only inhibitory near the translation start and not downstream in the coding sequence. We initially inserted the two Thr-Leu codon pairs into the cloning site near the 5′ end of a T7 vector system which was developed specifically for tests of translation efficiency

[†] This work was supported by a grant from the National Science Foundation (MCB-9981879). L.C. was supported by a graduate fellowship from the UMDNJ Graduate School of Biomedical Sciences.

* Corresponding author. Tel.: 973-972-4367. Fax: 973-972-3644. Email: <egoldman@umdnj.edu>.

[‡] Graduate School of Biomedical Sciences.

[§] New Jersey Medical School.

[1] Abbreviations: Thr-Leu, threonine-leucine codon pair; PCR, polymerase chain reaction; IPTG, isopropyl-β-D-thiogalactopyranoside; MOPS, 3-(*N*-morpholino)propanesulfonic acid; EDTA, ethylenediaminetetraacetic acid; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

10.1021/bi010236v CCC: $20.00
Published on Web 04/26/2001

of rare codons (*9*). To our surprise, the two codon pairs at this site performed equally well in supporting translation. Since the cloning site for this vector inserts at codons 14 and 15 (which are further downstream than positions tested in ref *8*) relative to the initiating AUG, we then constructed variants of the test gene in which the two Thr-Leu codon pairs were now located at codons 6 and 7. Unexpectedly, the two codon pairs again performed equally well in supporting translation of the test gene. Thus, we conclude that, for reasons that are unclear, inhibition by the over-represented ACG CUG codon pair reported in an *E. coli* expression system is not observed in a T7 expression system. However, since the translation apparatus is believed to be the same in both systems, it appears that the results reported here rule out the possibility of different translation step times through the respective codon pairs, as previously proposed (*8*).

EXPERIMENTAL PROCEDURES

Construction of Plasmids Containing the Thr-Leu Codon Pairs. Synthesis of all oligonucleotides, and all DNA sequencing, was performed by the New Jersey Medical School Molecular Resource Facility, under the direction of Robert Donnelly, Ph.D.

To construct pCT1-H14 (over-represented codon pair ACG CUG at codons 14 and 15) and pCT1-L14 (under-represented codon pair ACC CUG at codons 14 and 15), the single-stranded DNA oligonucleotides gatccgacgctgg and gatccgaccctgg were hybridized to their respective complements, aattccagcgtcg and aattccagggtcg, creating *Bam*HI (g/gatcc) and *Eco*RI (g/aattc) sticky ends. These double-stranded oligos were ligated into precut (with *Bam*HI and *Eco*RI) pCT1 (*9*) (see Figure 1, line 3) to create the inserts in the test gene as shown in Figure 1, lines 4 and 5.

The construction of pCT1-H6 (ACG CUG at codons 6 and 7) and pCT1-L6 (ACC CUG at codons 6 and 7) was somewhat more complicated since the cloning site in pCT1 is further downstream, following codon 13. To accomplish these constructions, we took advantage of the modular nature of the pCT1 vector (*9*), which provides an *Nde*I restriction site overlapping the initiation codon in the test gene. The complication is that there is a second *Nde*I site in the vector as well, overlapping the initiation codon of the downstream control gene (see Figure 1, panel 2). Single-stranded DNA oligonucleotides tatggctgaatctaatacgctg and tatggctgaatctaatacctg were hybridized to their respective complements, gatccagcgtattagattcagcca and gatccagggtattagattcagcca, creating *Nde*I (ca/tatg) and *Bam*HI (g/gatcc) sticky ends. These double-stranded oligos were ligated into precut pCT1 (first digested with *Bam*HI, then partially digested with *Nde*I). PCR colony screening was performed on the transformants, with one of the single-stranded oligonucleotides (i.e., used to generate the insert) as the left primer, and an oligo based on the sequence of the complementary strand in the region near the Tϕ element (see Figure 1, panel 2) as the right primer. This facilitated identification of cells with plasmids containing the inserts shown in Figure 1, lines 6 and 7.

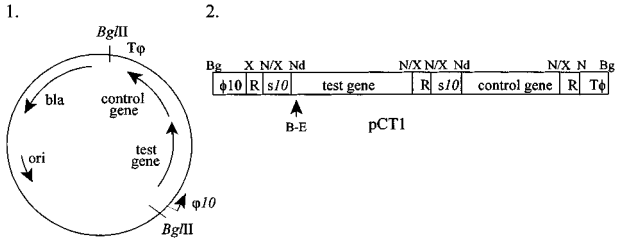
All constructs were verified by DNA sequencing.

Analysis of Protein Expression. BL27(DE3)pLysS (*9*) cells were used for expression experiments. pLysS is a compatible chloramphenicol-resistant plasmid which supplies T7 lysozyme, an inhibitor of T7 RNA polymerase, to minimize basal expression of T7 promoters in the absence of induction (*13*). Cells carrying the various plasmids shown in Figure 1 were grown in a New Brunswick shaker bath at 37 °C in M9 minimal media supplemented with ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL) to $\sim 2 \times 10^8$ cells/mL (measured as 50 Klett units in a Klett−Summerson colorimeter), and then induced by adding IPTG to 1 mM. In some samples, as indicated, 100 μg/mL rifampicin (from a 10 mg/mL stock dissolved in dimethyl sulfoxide) was added 15 min after induction. For each strain, 50 μL samples were labeled for 2 min at 37 °C with 20 μCi/mL, 50 Ci/mmol [^{35}S]methionine (Amersham) before (0), at 30 and 60 min, or, in some experiments, at 20, 40, and 60 min after induction. Labeling was stopped by mixing each sample with 25 μL of 3× MOPS cracking buffer [3× MOPS cracking buffer contains 60 mM MOPS, 24 mM sodium acetate, 3% SDS, 6 mM EDTA, 30% glycerol, and 0.15% each of bromophenol blue and xylene cyanole FF, and is stored at room temperature; 3% (v/v) β-mercaptoethanol is added to complete the reagent immediately before use]. Samples were placed on ice until completion of the labeling, then boiled in capped tubes for 2 min, stored at −20 °C, boiled again for 1 min, and subjected to sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) in gels containing 10% acrylamide. Dried gels were subjected to autoradiography and/or visualized in a Molecular Dynamics Phosphorimager.

RESULTS

Expression System. The T7 expression system used was originally designed for evaluating the translation-inhibiting effects of low-usage codons (*9*). The system, diagramed in panels 1 and 2 of Figure 1, consists of a plasmid vector containing a 'test' and 'control' gene (derived from T7 gene *9*) in tandem on one transcription unit, controlled by a T7 RNA polymerase promoter (ϕ*10*) and terminator (Tϕ). Both genes are fused to common strong translation start signals (s*10*), and both are flanked (and separated) by single-cut RNaseIII processing sites (R). The control gene has actually the same sequence as the test gene, except for a 45-codon (in-frame) deletion to permit distinguishing by size both test and control protein and mRNA in the same cell. The complete system consists of 3 alternate versions of the test gene in which are inserted 3 alternate cloning sites, respectively, either 13 codons from the AUG start (5′ end), 223 codons downstream from the start (internal), or 307 codons downstream, which is 6 codons from the UAA stop (3′ end). In the experiments reported here, only the version with the cloning site near the 5′ end, pCT1, has been used. The test gene contains a typical high-usage codon composition characteristic of highly expressed genes in *E. coli*, and does not contain any Thr-Leu codon pairs.

The plasmid is inserted into a strain harboring a chromosomal copy of the T7 RNA polymerase gene under control of the *lac*UV5 promoter (*13*); hence, induction with IPTG turns on synthesis of the T7 RNA polymerase, which, in turn, begins transcription of the test and control genes on the vector. Test and control proteins accumulate for well over an hour, and a maximal rate of protein synthesis is obtained by about 20−30 min after induction (*9*, *13*). Insertion of



3. pCT1
ata*cat* *atg* gct gaa tct aat gca gac gta tat gca ct*g* *gat* *ccg* aag ctt *gaa* *ttc* tat gca

4. pCT1-H14
ata*cat* *atg* gct gaa tct aat gca gac gta tat gca ct*g* *gat* *ccg* ACG CTG *gaa* *ttc* tat gca

5. pCT1-L14
ata*cat* *atg* gct gaa tct aat gca gac gta tat gca ct*g* *gat* *ccg* ACC CTG *gaa* *ttc* tat gca

6. pCT1-H6
ata*cat* *atg* gct gaa tct aat ACG CT*G* *gat* *ccg* aag ctt *gaa* *ttc* tat gca

7. pCT1-L6
ata*cat* *atg* gct gaa tct aat ACC CT*G* *gat* *ccg* aag ctt *gaa* *ttc* tat gca

FIGURE 1: Test vector and inserted sequences used in this work. Panel 1: overall design of the vector system (*9*), which is based on pBR322. Panel 2: schematic design of the elements used in the expression system. Description of the elements is in the text. Restriction sites: B−E = *Bam*HI−*Eco*RI cloning site; Bg = *Bgl*II; Nd = *Nde*I; N = *Nhe*I; X = *Xba*I; N/X = *Nhe*I/*Xba*I fusion. Line 3: DNA nucleotide sequence from the beginning of the coding region through the cloning site in vector pCT1. Lines 4 and 5: DNA nucleotide sequences from the beginning of the coding region through the cloning sites in pCT1-H14 (line 4) and pCT1-L14 (line 5). Lines 6 and 7: DNA nucleotide sequences from the beginning of the coding region through the cloning sites in pCT1-H6 (line 6) and pCT1-L6 (line 7). The respective Thr-Leu codon pairs (ACG CTG and ACC CTG) are shown in upper case letters. *Nde*I (catatg), *Bam*HI (ggatcc), and *Eco*RI (gaattc) sites are underlined and italicized. The *atg* in the *Nde*I site is the initiating methionine codon. Spaces are placed between codons in the coding region. A *Hind*III site (aagctt) is used as a spacer in the cloning site between the *Bam*HI and *Eco*RI sites in pCT1 (line 3).

alternate sequences into the cloning site in the test gene permits a direct and controlled comparison of translation rates.

There are numerous virtues to this design: (1) The use of a T7 promoter removes the concern of polarity of transcription, or coupling of transcription and translation, since the T7 RNA polymerase is not subject to polarity. (2) Both test and control messages come from a single transcription unit and, following RNaseIII processing, will have identical 5′ and 3′ ends; thus, levels of both mRNAs are expected to be close if not identical. (3) RNaseIII processing of the primary transcript confers a high degree of stability on the resulting mRNAs. (4) The overall common sequence of test and control genes minimizes possible effects of context or other potential variables, and the control gene provides an internal standard between different cultures and experiments. (5) The analysis of the products does not rely on enzymatic function (which may vary), but only on size. (6) The T7 system provides high-level expression with minimal interference from host expression, facilitating easy detection and analysis. (7) Unlike *E. coli* RNA polymerase, T7 RNA polymerase is not subject to inhibition by rifampicin; thus, following a period of induction of the T7 RNA polymerase gene, rifampicin can be added to fully shut off background host expression, and only T7 promoters will remain active, yielding essentially only the T7-directed products.

Expression with Alternate Thr-Leu Codon Pairs Inserted into Vector pCT1. As described in the introduction, we were interested in ascertaining whether inhibition of expression by the over-represented ACG CUG (Thr-Leu) codon pair as reported in an *E. coli* system (*8*) would be subject to 5′-translational blockage, that is, only inhibitory near the translation start and not downstream. Our plan was to first confirm the inhibitory effect when these codon pairs were inserted at the cloning site near the 5′ end of vector pCT1 in our T7 system, then to determine expression levels when these codon pairs were inserted into the downstream cloning site of the test gene in vector pCT2 (*9*). Our first constructions, therefore, inserted the two alternate Thr-Leu codon pairs at codons 14 and 15 of pCT1, as described under Experimental Procedures. When these constructs, called pCT1-H14 [ACG CUG at codons 14 and 15; "H" for "high" (over-represented)] and pCT1-L14 [ACC CUG at codons 14 and 15; "L" for "low" (under-represented)], were tested for incorporation of [^{35}S]methionine into the test and control proteins, we found that there was no effect of varying the codon pairs on the level of expression (data not shown). Hence, there was no need to synthesize constructs with the codon pairs inserted at the downstream cloning site in pCT2.

Because the cloning site we had tested in pCT1 is somewhat further downstream from the translation start compared to the constructs tested in Irwin et al. (*8*) (whose constructs placed the codon pairs at either codons 3 and 4, or 6 and 7, or 9 and 10), we decided to construct alternate versions of the test gene of pCT1, in which the Thr-Leu codon pairs were placed at codons 6 and 7. These constructions, described under Experimental Procedures, created pCT1-H6 (ACG CUG at codons 6 and 7) and pCT1-L6 (ACC CUG at codons 6 and 7).

These two constructs were examined for incorporation of [^{35}S]methionine into the test and control proteins. Cells were grown to mid-log phase in minimal media, as described under Experimental Procedures, and T7 transcription was induced by the addition of IPTG. [^{35}S]Methionine was added to portions of the cultures either before induction of the T7 RNA polymerase (i.e., no IPTG), in lanes marked "0" in Figure 2, or 30 or 60 min after induction, as indicated in the figure. Since radiolabeling was performed as 2 min pulses of [^{35}S]methionine at the various times, the protein bands reflect the rates of incorporation at the indicated times, and not the accumulation of label in the proteins following induction. As shown in Figure 2, both test and control proteins were synthesized comparably in cells containing either pCT1-H6 (marked "H6" in Figure 2) or pCT1-L6 (marked "L6" in Figure 2), both in the absence and in the presence of rifampicin. Quantitation of these bands in a phosphorimager indicates that the ratios of synthesis of test to control proteins in the H6 samples were, if anything, even slightly higher that the ratios obtained with the L6 samples (data not shown). Because this result was so unexpected, we extracted plasmid DNAs using a Qiagen mini-prep kit from the very same cultures used in this experiment, and had these DNAs sequenced again, which confirmed the sequences as shown in Figure 1, lines 6 and 7. Thus, we were forced to conclude that the inhibition by the over-represented Thr-Leu codon pair observed in an *E. coli* expression system did not occur in our T7 system.

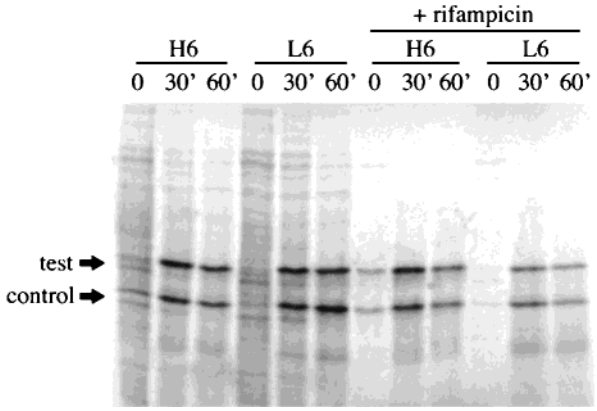


FIGURE 2: Synthesis of test and control proteins in pCT1-H6 and pCT1-L6. Cells were grown to mid-log phase, induced with IPTG at 0 time, pulse-labeled for 2 min with [^{35}S]methionine at 30 and 60 min post-induction in the absence or presence of rifampicin, and subjected to SDS−PAGE, as described under Experimental Procedures. The "0" time samples for pCT1-H6 (H6 in the figure) and pCT1-L6 (L6 in the figure) were labeled prior to induction. The uninduced rifampicin-treated samples ("0" time) were exposed to rifampicin for 10 min prior to addition of [^{35}S]methionine. Induced rifampicin-treated samples were exposed to rifampicin starting at 15 min post-induction. Following SDS−PAGE, the gel was dried under vacuum, exposed to X-ray film for 7 h, and imaged in a Molecular Dynamics Phosphorimager following a 2 h exposure, as shown in the figure.

DISCUSSION

In this work, we have demonstrated that inhibition by the over-represented Thr-Leu codon pair ACG CUG observed in an *E. coli* system does not occur in a T7 system. There are a number of differences between the two systems, even including the strains used, and, presumably, one or more of these differences is responsible for the different results.

A major difference is the RNA polymerase used for transcription of the genes containing the alternate codon pairs. The T7 RNA polymerase is highly processive, is considerably faster than the *E. coli* RNA polymerase (*14*), and is not subject to polarity or to coupling of transcription to translation. Perhaps the inhibition observed in the *E. coli* system depends after all on intrinsic polarity or other properties of the *E. coli* RNA polymerase.

Another major difference is in the stability of the mRNAs in the two systems. mRNA for β-galactosidase, known to be unstable, was not directly measured in Irwin et al. (*8*); their conclusions relied on measurements of the activity of a downstream gene (the transacetylase) transcribed from the same promoter, as a control for transcription. While this is a suitable control for transcription starts and for effects of polarity, it is not suitable for assessing the possibility of differential mRNA stability near the 5′ end of the message, encoding the β-galactosidase reporter. By contrast, the mRNA in our vector system is designed to be highly stable, with RNaseIII processing sites at the 3′ ends of both the test and control genes, and mRNAs have been measured by Northerns for this vector (*9*, *10*). Thus, differential mRNA stability with the different codon pairs could also account for the results of Irwin et al. (*8*).

Another difference is related to the translation start signals. Our T7 system employs a translational enhancer element (s*10* in Figure 1), derived from upstream sequences native to T7 gene *10*. This element facilitates efficient translation starts (*15*), and perhaps programs the ribosome in some unknown fashion to be impervious to certain inhibitory signals such as the over-represented Thr-Leu codon pair. This would be analogous to anti-termination mechanisms for *E. coli* RNA polymerase, in which transcription of certain sequences causes the RNA polymerase to be re-programmed to ignore transcription termination signals (*16*).

The reading contexts around the codon pairs in our experiments are also different from the contexts in which the codon pairs were tested in Irwin et al. (*8*). We note, however, that the immediate sequence context surrounding the codon pairs inserted at codons 6 and 7 in this work is different from the sequence context surrounding our insertions at codons 14 and 15 (see Figure 1, lines 4 and 5 vs lines 6 and 7). Because Irwin et al. (*8*) also obtained similar results in different reading contexts, it seems unlikely that this could be the explanation for the different results.

Still another parameter to consider in these experiments is the relative level of ribosomes versus mRNA. In the T7 system, a lot of mRNA is generated quickly, which may lead to a condition of limiting ribosomes, whereas in the *E. coli* RNA polymerase system, especially in rich media [i.e., apparently as used in Irwin et al. (*8*)], one might expect conditions of ribosome excess. Transient effects on initiation by inhibition of elongation near the translation start may be exacerbated when there is ribosome excess, but are not expected when ribosomes are limiting. However, it should be noted that we observed no inhibition by the ACG CUG codon pair on apparent basal expression in uninduced samples in the absence of rifampicin, when ribosomes would be expected to be in excess.

Whatever the basis for the different results in the two systems, the reported inhibition by the over-represented ACG CUG codon pair seems unlikely to reflect a straightforward inhibition of translational step times, as concluded by Irwin et al. (*8*). Since both the *E. coli* and T7 systems use the same translational apparatus, and since there was no effect of varying codon pairs in the T7 system, we conclude that, as a first approximation, there are no significant translational step time differences between the two Thr-Leu codon pairs in *E. coli*.

REFERENCES

1. Gesteland, R. F., and Atkins, J. F. (1996) *Annu. Rev. Biochem. 65*, 741−768.
2. Sharp, P. M., and Li, W.-H. (1987) *Nucleic Acids Res. 15*, 1281−1295.
3. Zhang, S., Zubay, G., and Goldman, E. (1991) *Gene 105*, 61−72.
4. Gutman, G. A., and Hatfield, G. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3699−3703.
5. Hatfield, G. W., and Gutman, G. A. (1992) in *Transfer RNA in Protein Synthesis* (Hatfield, D. L., Lee, B. J., and Pirtle, R. M., Eds.) pp 157−190, CRC Press, Boca Raton, FL.
6. Smith, D., and Yarus, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4397−4401.
7. Buckingham, R. H. (1994) *Biochimie 76*, 351−354.
8. Irwin, B., Heck, J. D., and Hatfield, G. W. (1995) *J. Biol. Chem. 270*, 22801−22806.
9. Rosenberg, A. H., Goldman, E., Dunn, J. J., Studier, F. W., and Zubay, G. (1993) *J. Bacteriol. 175*, 716−722.
10. Goldman, E., Rosenberg, A. H., Zubay, G., and Studier, F. W. (1995) *J. Mol. Biol. 245*, 467−473.
11. Chen, G. F. T., and Inouye, M. (1990) *Nucleic Acids Res. 18*, 1465−1473.

12. Chen, G. F. T., and Inouye, M. (1994) *Genes Dev. 8*, 2641−2652.
13. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.
14. Iost, I., Guillerez, J., and Dreyfus, M. (1992) *J. Bacteriol. 174*, 619−622.
15. Olins, P. O., and Rangwala, S. H. (1989) *J. Biol. Chem. 264*, 16973−16976.
16. Greenblatt, J., Mah, T. F., Legault, P., Mogridge, J., Li, J., and Kay, L. E. (1998) *Cold Spring Harbor Symp. Quant. Biol. 63*, 327−336.

BI010236V