Genetic Code-guided Protein Synthesis and Folding in *Escherichia coli*\(^*\)\(^*\)

Shao-liang Hu\(^*\), Mingrong Wang\(^*\), Guoping Cai\(^*\), and Mingyue He\(^*\)

From the \(^*\)School of Life Sciences, Tsinghua University, Beijing 100084, the \(^*\)Shenzhen Graduate School, Tsinghua University, Shenzhen 518055, and \(^*\)Recombinant Protein Therapeutics, Chengdu Institute of Biological Products, Chengdu, Sichuan 610023, China

**Background:** Synonymous codon usage affects protein properties in a given organism.

**Results:** A total of 342 antibody codon variants were identified, differing significantly in solubility and functionality while retaining the identical original amino acid sequence.

**Conclusion:** Genetic codes control protein synthesis and folding. “Codon-preferred” DNA template(s) can be generated by functional screening.

**Significance:** Protein properties can be considerably altered by synonymous codons without substituting amino acids.

Universal genetic codes are degenerated with 61 codons specifying 20 amino acids, thus creating synonymous codons for a single amino acid. Synonymous codons have been shown to affect protein properties in a given organism. To address this issue and explore how *Escherichia coli* selects its “codon-preferred” DNA template(s) for synthesis of proteins with required properties, we have designed synonymous codon libraries based on an antibody (scFv) sequence and carried out bacterial expression and screening for variants with altered properties. As a result, 342 codon variants have been identified, differing significantly in protein solubility and functionality while retaining the identical original amino acid sequence. The soluble expression level varied from completely insoluble aggregates to a soluble yield of \(\sim 2.5\) mg/liter, whereas the antigen-binding activity changed from no binding at all to a binding affinity of \(\sim 10^{-8}\) M. Not only does our work demonstrate the involvement of genetic codes in regulating protein synthesis and folding but it also provides a novel screening strategy for producing improved proteins without the need to substitute amino acids.

Proteins are encoded by genetic codes stored in DNA. The ribosome, the “protein synthesis machinery,” deciphers codons aligned along mRNA to synthesize a specific polypeptide, which then folds into a defined structure/conformation (1). It has long been believed that the amino acid sequence contains all of the essential information required for folding the protein into a specific three-dimensional structure under an appropriate condition despite the detail remains unknown (1). Recent studies have revealed a co-translational protein folding mechanism in which the nascent polypeptide starts to fold immediately after it emerges from the ribosomal tunnel and subsequently folds into a final state with the assistance of molecular chaperones and the ribosome itself (2, 3).

There are 61 universal codons encoding for 20 amino acids, hence offering cells the flexibility to select synonymous codons for translating a polypeptide (4). Interestingly, synonymous codons are non-randomly distributed along genes and the frequency with which an individual synonymous codon is chosen to specify a protein exhibits a “favored” usage bias in a given organism (5). Moreover, codon usage bias varies significantly between different organisms (6) and attempts at producing proteins in heterologous cells often result in a poor synthesis or formation of insoluble aggregates. By prediction, \(\sim 40\%\) of human genes would not express or express very poorly when transformed into *Escherichia coli* (5). However, by mimicking *E. coli* codon usage bias, synonymous substitutions of foreign genes (“codon harmonization”) improved protein synthesis despite a varying success and unpredictability by this approach (7).

Synonymous codons also influence the function of proteins. Naturally occurring silent mutants have been discovered affecting protein folding (8), altering substrate recognitions (9) as well as triggering various diseases (10), suggesting that synonymous codons might control the folding of nascent polypeptides emerging from ribosome by regulating polypeptide elongation rates (4, 11). Indeed, mRNAs composed of different synonymous codons were translated at different rates (12, 13) and the choice of rare codons, the availability of corresponding tRNAs, and adjacent codon pairs was shown to slow down protein synthesis on ribosomes (14). Moreover, identical DNA sequences could generate polypeptides with different secondary structures when the translation speed was altered (11). However, a recent investigation using ribosome density profiling technology suggested that the presence of codons with rare tRNAs did not decrease the translation rates (15). A separate study also indicated that tRNA gene numbers were not completely responsible for the codon usage bias in certain amino acids (16).

Thus, it is clear that genetic codes contain additional information beyond amino acid sequences. But how DNA sequences govern the protein synthesis is not clear. To address this issue and investigate how *E. coli* select its codon-preferred DNA...
Effects of Synonymous Codons on Protein Properties in E. coli

**TABLE 1**

| No. | Primers for VH library | Codon usage | Frequency (%) |
|-----|------------------------|-------------|---------------|
| EVHM1 | gcagaagactagctgtgctgagtcgcctg | 3         | 10.1          |
| EVHM2 | cygggcgctgctgtgctgagtcgcctg | 3         | 10.1          |
| EVHM3 | gcagaagactagctgtgctgagtcgcctg | 3         | 10.1          |
| EVHM4 | gcagaagactagctgtgctgagtcgcctg | 3         | 10.1          |
| EVHM5 | gcagaagactagctgtgctgagtcgcctg | 3         | 10.1          |
| EVHM6 | gcagaagactagctgtgctgagtcgcctg | 3         | 10.1          |
| EVHM7 | gcagaagactagctgtgctgagtcgcctg | 3         | 10.1          |

Further 25 cycles (94 °C for 30 s, 55 °C for 30 s, and 68 °C for 5 min) after adding the upstream and downstream flanking primers. A final extension was carried out at 68 °C for 10 min.

**E. coli Cloning, scFv Expression, and Preparation**—PCR products were digested by restriction enzymes NdeI and EcoRI and ligated into the plasmid pET22b (Novagen). E. coli BL21(DE3) (Novagen) was transformed for periplasmic expression of the scFv. For statistical analysis purposes, four duplicated clones from each variant were grown in 2×YT medium containing 50 μg/ml carbenicillin at 37 °C overnight. The overnight culture was then diluted in 1:20 and continued to grow at 37 °C for further 2 h prior to the addition of isopropyl-1-thio-

**UV CD Analysis**—UV CD measurements were performed using the Sigma HRP-coupled monoclonal anti-His6 antibody (1:6000) was added (100 μl/well), and the incubation was continued for 1 h at 37 °C. Finally, the reaction was stopped with 1% BSA. After blocking with 1% BSA for 2 h, the wells were added with either purified or non-purified soluble scFv (100 μl/well in duplicates) and incubated at 30 °C for 1 h. After three washes, Sigma HRP-coupled monoclonal anti-His6, antibody (1:6000) was added (100 μl/well), and the incubation was continued for 1 h at 37 °C. The HRP activity was then developed by addition of 100 μl/well of TMB Liquid Substrate (Sigma) for 5–10 min at 37 °C. Finally, the reaction was stopped with 100 μl of 1 N HCl, and the microtiter wells were read at 450 nm.

The antigen-binding activity of the scFv was examined by a side-by-side comparison using the Sigma HRP-coupled monoclonal anti-His6 antibody (1:6000) was added (100 μl/well), and the incubation was continued for 1 h at 37 °C. The HRP activity was then developed by addition of 100 μl of 3,3′,5,5′-tetramethylbenzidine (TMB) liquid substrate (Sigma) for 5–10 min at 37 °C. Finally, the reaction was stopped with 100 μl of 1 N HCl, and the microtiter wells were read at 450 nm.

The antigen-binding activity of the scFv was examined by a similar procedure as described above except the microtiter wells were coated with purified rabbit polyclonal antibodies against the IgE scFv at 1 μg/well, 4 °C overnight. After blocking with 1% BSA for 2 h, the wells were added with either purified or non-purified soluble scFv (100 μl/well in duplicates) and incubated at 30 °C for 1 h. After three washes, Sigma HRP-coupled monoclonal anti-His6 antibody (1:6000) was added (100 μl/well), and the incubation was continued for 1 h at 37 °C. The HRP activity was then developed by addition of 100 μl of 3,3′,5,5′-tetramethylbenzidine (TMB) liquid substrate (Sigma) for 5–10 min at 37 °C. Finally, the reaction was stopped with 100 μl of 1 N HCl, and the microtiter wells were read at 450 nm.

Further 25 cycles (94 °C for 30 s, 55 °C for 30 s, and 68 °C for 5 min) after adding the upstream and downstream flanking primers. A final extension was carried out at 68 °C for 10 min.

**E. coli Cloning, scFv Expression, and Preparation**—PCR products were digested by restriction enzymes NdeI and EcoRI and ligated into the plasmid pET22b (Novagen). E. coli BL21(DE3) (Novagen) was transformed for periplasmic expression of the scFv. For statistical analysis purposes, four duplicated clones from each variant were grown in 2×YT medium containing 50 μg/ml carbenicillin at 37 °C overnight. The overnight culture was then diluted in 1:20 and continued to grow at 37 °C for further 2 h prior to the addition of isopropyl-1-thio-

**UV CD Analysis**—UV CD measurements were performed using the Sigma HRP-coupled monoclonal anti-His6 antibody (1:6000) was added (100 μl/well), and the incubation was continued for 1 h at 37 °C. The HRP activity was then developed by addition of 100 μl of 3,3′,5,5′-tetramethylbenzidine (TMB) liquid substrate (Sigma) for 5–10 min at 37 °C. Finally, the reaction was stopped with 100 μl of 1 N HCl, and the microtiter wells were read at 450 nm.

The antigen-binding activity of the scFv was examined by a side-by-side comparison using the Sigma HRP-coupled monoclonal anti-His6 antibody (1:6000) was added (100 μl/well), and the incubation was continued for 1 h at 37 °C. The HRP activity was then developed by addition of 100 μl of 3,3′,5,5′-tetramethylbenzidine (TMB) liquid substrate (Sigma) for 5–10 min at 37 °C. Finally, the reaction was stopped with 100 μl of 1 N HCl, and the microtiter wells were read at 450 nm.
response time, 8 s; scanning speed, 50 nm/min; measurement range, 260–190 nm; For the signal to noise ratio, each spectrum was the average of four scans. Structure analysis was done by using the CDSSTR software (DichroWeb). Protein concentration for the scans was 0.1 mg/ml in a buffer of 50 mM sodium phosphate (pH 7.2). Near-UV CD was measured as follows: cell length, 1 cm; bandwidth, 1.0 nm; response time, 2 s; scanning speed, 20 nm/min; measurement range, 320–250 nm; each spectrum was the average of four scans. Protein concentration was 0.5 mg/ml in a buffer of 50 mM sodium phosphate (pH 7.2).

**Statistical Analyses**—Statistical analysis was carried out using IBM Statistical Product and Service Solutions (SPSS) 19. Data were collected as means ± S.D. The analysis of variance was used to analyze the variation between samples.

**Bioinformatics Analysis**—Codon Adaptation Index and GC3 analysis was carried out using the software Codon W (SourceForge), and the codon bias usage database was also used. tRNA adaptation index analysis was based on the genomic tRNA database (tRNAscan-SE Genomic tRNA Database). Free energy prediction of 5’ mRNA secondary structure was by the program from the University of Rochester Medical Center.

**RESULTS**

**Design and Construction of Synonymous Codon Libraries**—We have designed the synonymous codon libraries based on a single-chain anti-IgE (scFv) template and substituted every codon with degenerated oligonucleotides at the third position (Table 1). Because the substitution of all codons across the full-length DNA would yield a library size beyond the limitation of *E. coli* transformation efficiency, we constructed two sub-libraries, one for the heavy chain (V\(_{H}\)) and one for the light chain (V\(_{L}\)). To reduce the library size further, each codon was only replaced by a high, medium, and low usage frequency codon, rather than using all possible codons (Table 1). PCR was used to introduce the designed synonymous codon mix into the scFv template. A His\(_{6}\) tag was engineered at the C terminus for scFv detection and purification. To verify the generation of synonymous codon libraries, we directly sequenced the PCR constructs before *E. coli* cloning, showing successful substitutions of individual codons at the third position for the V\(_{H}\) library (Fig. 1).

**E. coli Expression and Screening for Altered Protein Synthesis**—Both V\(_{H}\) and V\(_{L}\) synonymous codon libraries were cloned
Effects of Synonymous Codons on Protein Properties in E. coli

into E. coli. DNA sequencing was carried out, leading to the identification of 342 variants containing synonymous codons at various positions while maintaining the identical, original amino acid sequence. These clones were expressed, and the synthesized scFv was confirmed by Western blotting (data not shown). Soluble scFv was extracted and subjected to a sandwich ELISA assay in which the soluble scFv was captured on the well by precoated rabbit polyclonal antibodies against the scFv. With the His6 tag at the C terminus, the captured scFv was detected by a HRP-linked anti-His6 antibody. The relative solubility level of scFv was then measured through a side-by-side comparison with the original clone, which generally yielded the scFv at the level of ~1 mg/liter (17). Our results revealed a huge difference in protein solubility among the variants, ranging from completely insoluble aggregates to a soluble yield of ~2.5 mg/liter (Fig. 2 and supplemental Table S1). To examine whether synonymous codons influence total protein expression, the insoluble scFv from the pellet was also monitored by solubilizing the inclusion bodies with 10% sarkosyl (18) followed by sandwich ELISA analysis. This showed that the total scFv expression (insoluble scFv + soluble scFv) also varied between the variants (supplemental Table S1). Western blotting also confirmed the scFv expression with an agreement with the ELISA results (data not shown). However, there was no correlation between the total scFv expression and its soluble production level (supplemental Table S1).

We have also analyzed the mRNA by RT-PCR using randomly selected 20 clones (10 from VH library and 10 from VL library), detecting a very similar amount of mRNA among the clones despite their apparent differences in protein expression (data not shown). Our result suggests that mRNA level was not the major factor affecting the scFv expression.

E. coli Expression and Screening for Altered Functionality—The antigen-binding activity of the scFv was examined by ELISA on wells coated with the antigen IgE. After binding, the bound scFv was detected by the HRP-linked anti-His6 antibody. Similarly, the relative affinity of the scFv was measured by a comparison with the original scFv on the same microtiter plate. This showed that the IgE-binding activity differed from no binding at all to an affinity greater than the original scFv (10^{-8} M) (Fig. 2 and supplemental Table S1) (17). To further characterize the binding activity, four variants, together with the original scFv, were expressed and purified using a His-tagged affinity column. The purified scFvs were then adjusted by dilution to the same level, based on their sandwich ELISA, and used for antigen-binding assays. Fig. 3A shows that while two clones had no binding at all, other two bound to the antigen with affinities greater than the original scFv by ~32- and 8-fold, respectively. This experiment has been repeated and similar results were obtained (data not shown). The sequences of the four variants were aligned with the original scFv, detecting no any positions contributing to the alteration of the binding activity (Fig. 3B).

Relationship between scFv Solubility and Functionality—We noticed that the scFv solubility was not related to its antigen-binding activity (supplemental Table S1). Although variants of highly soluble and active variants were obtained, many clones showed a high solubility but a low affinity or vice versa (Fig. 2 and supplemental Table S1). Notably, most substitutions were deleterious when compared with the original sequence (Fig. 2); however, in general, substitutions at VH region seemed affect the antigen-binding activity (Fig. 2A), whereas VL variants mainly influenced the solubility and decreased the affinity (Fig. 2B). A number of variants with enhanced solubility and antigen-binding affinity over the original scFv were identified from both VH and VL libraries.

Bioinformatics Analysis—Sequence alignment of the 342 codon variants was carried out to detect any possible associations between synonymous codons and their corresponding properties. Unfortunately, this did not identify any regions or individual positions that contributed to the altered solubility and functionality of the scFv (Fig. 3B; data not shown). Substitutions at either VH or VL region could increase or decrease the solubility and affinity of the scFv (supplemental Table S1). We also analyzed the codon adaptation indices such as codon adaptation index, tRNA adaptation index, and GC3, which have been developed to define the relative “adaptiveness” of a codon.
and used to predict protein expression from DNA sequences (19). Again, we did not detect their correlation with the protein properties, although there may be a weak association between GC3 and expression level in the VL library (data not shown).

**CD Spectroscopy**—To assess the conformation/folding of variants, CD analysis was performed by comparing the original scFv with affinity-altered variants (Fig. 4). Far-UV CD analysis has showed that all the selected scFvs exhibited distinct CD spectra with a negative band between 216–220 nm and a positive band between 195–200 nm, which are characteristic of H-strand contents. However, they varied in intensity at the peak 220 nm and/or 195 nm, indicating the protein secondary structure was slightly altered by the synonymous mutations (Fig. 4A). With CDSSTR software available, we were able to calculate the fraction of changes in the structure for each individual variant (Fig. 4A).

**FIGURE 3. Comparison of scFv variants with the original scFv.** A, ELISA of antigen-binding activity. An equal amount of purified scFvs was used. IgEscFv V\(_H\)-H (red line) and IgEscFv V\(_L\)-L (blue line), variants from V\(_H\) library; IgEscFv V\(_H\)-H (pink line) and IgEscFv V\(_L\)-L (green line), variants from V\(_L\) library; IgEscFv WT (black line), the original scFv. B, DNA sequencing alignment. The sequences of scFv variants and the original scFv are aligned, and colors indicate the substituted synonymous codons. Red, high frequency codons; yellow, medium frequency codons; blue, low frequency codons. The nucleotide substitutions are underlined. AA, amino acid.

**FIGURE 4. CD spectra analysis of scFv synonymous variants.** A, the far-UV CD spectra. The table shows the relative affinity and the secondary structure fraction of variants. B, the near-UV CD spectra. 1, IgEscFv WT, the original scFv (black line); 2, IgEscFv V\(_H\)-H (red line); 3, IgE scFv VH-M1 (blue line); 4, IgEscFv V\(_L\)-M2 (pink line).
Effects of Synonymous Codons on Protein Properties in E. coli
carried out to analyze the environmental change of aromatic residues on the protein surface, such as tryptophan (Trp) at 288–293 nm, tyrosine (Tyr) at 275–282 nm, and phenylalanine (Phe) at 255–278 nm. Fig. 4B shows the near-UV CD spectra, revealing clearly the variation at ∼284 and 291 nm, which was most likely from the contribution of Tyr and Trp located within or adjacent to CDRs (Figs. 3B and 4B). Our results indeed provide the structural explanation for the altered affinity by the synonymous variants.

Interestingly, variants with no antigen-binding activity or a very low affinity could not be purified by protein L-agarose (data not shown). Perhaps these variants were unstable and formed aggregates on the column during the purification or they contained an altered conformation not recognized by protein L.

DISCUSSION
We have shown that a scFv encoded by different synonymous codons could be synthesized in E. coli with considerably changed solubility and antigen-binding activity while retaining the identical, original amino acid sequence. The significant difference in the antibody affinity (Figs. 2 and 3 and supplemental Table S1) clearly suggests a structural/conformational alteration caused by synonymous codons. Our CD analysis from affinity-improved variants also detected the structural/conformational changes (Fig. 4), directly demonstrating that synthesis and folding of proteins in E. coli is indeed controlled by genetic codes. Our results also suggest that production of heterologous proteins in E. coli for structural and functional studies should be carried out with caution because foreign proteins may not be faithfully regenerated in bacteria. Because the scFv was designed to be secreted into the periplasm of E. coli in a structurally loosely folded state (20), our results favor the previous proposal that synonymous codons only affect the encoded protein at the level of secondary structure (11).

It was observed that codons with a high-frequency usage were mainly located at structural regions, whereas rare codons were more likely for β-strand, random coil and domain boundaries (21). Studies on rare codon distribution have also suggested the conservation between structurally related proteins from different organisms (8). However, our sequence analysis from 342 codon variants did not detect any individual regions or positions contributing to the altered protein properties (Fig. 3B and supplemental Table S1), suggesting multiple codons across the entire sequence may act in a synergistic manner. Recently, GFP synthesis was significantly improved by synonymous substitutions of its N-terminal 40 nucleotides that were proposed to form a specific 5′-mRNA secondary structure favoring the translation initiation (22). However, our results clearly showed that an identical N-terminal sequence could produce the protein either at a high level or a low yield, depending on the codon distribution at the downstream sequence (Fig. 3B and supplemental Table S1, the Vλ library), suggesting that the N-terminal sequence is not the sole or indispensable factor affecting protein expression as observed by others (23). Moreover, by using the same method as described in the GFP work (22), we could not detect any ΔG association of 5′-mRNA with the scFv expression and function (Fig. 5). Also, our codon adaptation index, tRNA adaptation index, and GC3 analyses failed to show any clear correlation between the codon adaptation indices and the scFv properties. Taken together, our work suggests that the effects of synonymous codon on protein properties cannot be accurately predicted, and attempts at expressing heterologous proteins in E. coli using a computational design remain a challenge.

Our strategy of screening synonymous codon libraries offers a powerful novel tool for identification of “codon-modified” DNA template(s) for synthesis of required variants. In this approach, various synonymous codons are combined along the DNA template to allow E. coli to choose its “preferred” template(s) for translating the protein with enhanced properties. The desirable variants can then be identified through screening of the required function/properties. Our successful discovery of enhanced codon variants, in particular the increase in scFv affinity by 32-fold (Fig. 3A), verifies the feasibility of engineering proteins without the necessity of substituting amino acids. We have successfully applied this approach to a number of other proteins, highlighting its general applicability (data not shown). It remains to be tested whether synonymous codons

![Image](313x383 to 563x733)
can affect antibody specificity and stability. We envisage that the power of this method can be further enhanced if a high-throughput screening technique (e.g. a display method or a reporter gene) is combined or the library diversity is further enlarged to contain all of the possible synonymous codons.

Modulating protein synthesis by genetic codes may have biological importance such as providing an additional control over protein synthesis. Recently, codon usage bias has been observed to influence cell cycle development (24), the responses of the cell to stress-specific conditions (25, 26), and protein phosphorylation profile and stability (27). In addition, it was discovered that silent mutations could cause serious diseases (10) or affect frame-shift in a given organism (28). Our strategy may be applied to address these issues and study disease genes.

Acknowledgment—We thank YongZhi He for help with the bacterial screening.

REFERENCES

1. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. Science 181, 223–230
2. Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis. Nature 475, 324–332
3. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I. Jr., and Bustamante, C. (2011) The ribosome modulates nascent protein folding. Science 334, 1723–1727
4. Zhang, G., Hubalewska, M., and Ignatova, Z. (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nat. Struct. Mol. Biol. 16, 274–280
5. Plotkin, J. B., and Kudla, G. (2011) Synonymous but not the same: the causes and consequences of codon bias. Nat. Rev. Genet. 12, 32–42
6. Sharp, P. M., Cowe, E., Higgins, D. G., Shields, D. C., Wolfe, K. H., and Shields, D. C. (1987) The Codon Adaptation Index: a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281–1295
7. Sørensen, M. A., and Pedersen, S. (1991) Absolute translation rates. In The proteins: Properties, Structure, and Genesis (C. B. Anfinsen, J. D. Bell, E. J. Orme, and D. M. Richardson, Eds.), pp. 147–161. Academic Press, San Diego
8. Komar, A. A. (2009) A pause for thought along the co-translational folding pathway. Trends Biochem. Sci. 34, 1–6
9. Kimchi-Sarfaty, C., Oh, J. M., Kim, I. W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V., and Gottesman, M. M. (2007) A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science 315, 525–528
10. Chen, R., Davydov, E. V., Sirota, M., and Butte, A. J. (2010) Non-synonymous and synonymous coding SNPs show similar likelihood and effect size of human disease association. PLoS One 5, e13574
11. Saunders, R., and Deane, C. M. (2010) Synonymous codon usage influences the local protein structure observed. Nucleic Acids Res. 38, 6719–6728
12. Sørensen, M. A., and Pedersen, S. (1991) Absolute in vivo translation rates of individual codons in Escherichia coli: The two glutamic acid codons GAA and GAG are translated with a three-fold difference in rate. J. Mol. Biol. 222, 265–280
13. Spencer, P. S., Siller, E., Anderson, J. F., and Barral, J. M. (2012) Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. J. Mol. Biol. 422, 328–335
14. Carnarozzi, G., Schraudolph, N. N., Faty, M., von Rohr, P., Friberg, M. T., Roth, A. C., Gonnet, P., Gonnet, G., and Barral, Y. (2010) A role for codon order in translation dynamics. Cell 141, 355–367
15. Li, G. W., Oh, E., and Weissman, J. S. (2012) The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature 484, 538–541
16. Satapathy, S. S., Dutta, M., Buragohain, A. K., and Ray, S. K. (2012) Transfer RNA gene numbers may not be completely responsible for the codon usage bias in asparagine, isoleucine, phenylalanine, and tyrosine in the high expression genes in bacteria. J. Mol. Evol. 75, 34–42
17. Wang, M., Zhang, Y., Du, T., Dai, Y., He, Y., Yu, X., Cong, C., and He, M. (2011) Bacterial expression and characterization of a novel human anti- IgE scFv fragment. MAbs. 3, 495–499
18. Tao, H., Liu, W., Simmons, B. N., Harris, H. K., Cox, T. C., and Massiah, M. A. (2010) Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. BioTechniques 48, 61–64
19. Sharp, P. M., and Li, W. H. (1987) The Codon Adaptation Index: a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281–1295
20. Driessen, A. J., and Nouwen, N. (2008) Protein translocation across the bacterial cytoplasmic membrane. Annu. Rev. Biochem. 77, 643–667
21. Thanaraj, T. A., and Argos, P. (1996) Protein secondary structural types are differentially coded on messenger RNA. Protein Sci. 5, 1973–1983
22. Kudla, G., Murray, A. W., Tollervey, D., and Plotkin, J. B. (2009) Coding-sequence determinants of gene expression in Escherichia coli. Science 324, 255–258
23. Welch, M., Govindarajan, S., Ness, J. E., Villalobos, A., Gurney, A., Minshall, J., and Gustafsson, C. (2009) Design parameters to control synthetic gene expression in Escherichia coli. PLoS One 4, e7002
24. Frenkel-Morgenstern, M., Danon, T., Christian, T., Igarashi, T., Cohen, L., Hou, Y. M., and Jensen, L. I. (2012) Genes adopt non-optimal codon usage to generate cell cycle-dependent oscillations in protein levels. Mol. Syst. Biol. 8, 1–10
25. Chan, C. T., Pang, Y. L., Deng, W., Babu, I. R., Dyavaiah, M., Begley, T. J., and Dedon, P. C. (2012) Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. Nat Commun. 3, 937–945
26. Begley, U., Dyavaiah, M., Patil, A., Rooney, I. P., DiRenzo, D., Young, C. M., Conklin, D. S., Zitomer, R. S., and Begley, T. J. (2007) Trm9-catalyzed tRNA modifications link translation to the DNA damage response. Mol. Cell 28, 860–870
27. Zhou, M., Guo, J., Chai, J., Chae, M., Chen, S., Barral, J. M., Sachs, M. S., and Liu, Y. (2013) Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature 495, 111–115
28. Huang, Y., Koonin, E. V., Lipman, D. J., and Przytycka, T. M. (2009) Selection for minimization of translational frame shifting errors as a factor in the evolution of codon usage. Nucleic Acids Res. 37, 6799–6810