Minireview

Differential codon usage: a safeguard against inappropriate expression of specialized genes?

Milton H. Saier, Jr.*

Department of Biology, University of California at San Diego, La Jolla, CA 92093-0116, USA

Received 2 January 1995; revised version received 14 February 1995

Abstract Recent work has suggested that rare codons are sometimes used for the regulation of specialized gene expression in bacteria. Moreover, the cellular levels of certain tRNAs may fluctuate with growth conditions. Evidence implicating such mechanisms in the control of photosynthesis in *Rhodobacter*, solventogenesis in *Clostridium*, sporulation in *Streptomyces*, and fimbrial phase variation in *E. coli* is summarized. It is suggested that such mechanisms will prove applicable to the control of numerous additional specialized functions, and that the empirical tools for testing this possibility are currently available.

Key words: Codon usage; Bacterium; Translation; Gene expression; Sporulation; Photosynthesis; Solventogenesis; Fimbria

1. Introduction

In 1989, Brinkman et al. noted that eukaryotic proteins such as the human tissue type plasminogen activator, prourokinase, and the gp41 protein of HIV, which have a high content of rare codons in their respective genes, are poorly expressed in *E. coli* [1]. Moreover, induction of the expression of any one of these heterologous, plasmid-encoded genes was found to inhibit cell division and cause plasmid instability. Most remarkably, when the bacteria were simultaneously provided with a plasmid bearing the *dnaY* gene, encoding a rare tRNA ($tRNA^{arg}_{AGA/AGG}$), production of the eukaryotic proteins was increased while plasmid stability and cell viability improved [1].

While these observations were of considerable practical significance to the bioengineer, they foreshadowed observations and experiments that would suggest that the use of rare codons for specialized or differentiation-specific functions in bacteria might provide a general mechanism to ensure proper temporal and spatial expression of the encoding genes. Although this hypothesis is still far from established, work in several laboratories has provided indirect evidence suggesting that rare codon usage is of functional significance in restricting or specifying appropriate gene expression. In this minireview I summarize the evidence concerned with this issue and reiterate the suggestion that the complement of tRNAs found in a particular bacterium under one set of growth conditions may differ from that found under another set of growth conditions.

*Corresponding author. Fax: (1) (619) 534-7108.
E-mail: msaier@ucsd.edu

2. Codon usage and gene expression

All living organisms possess characteristic GC contents and preferred sets of codons used for protein biosynthesis. GC content is a major determinant of codon usage, and codon adaptation indices (CAI values) have proven to provide a reliable, empirically determined estimate of gene expression level for specific groups of organisms [2–4]. tRNA availability during the evolution of an organism may play a significant role in determining its characteristic preferred codon usage. However, genes obtained by horizontal transmission from a phylogenetically divergent organism with GC content and codon usage different from those of the recipient bacterium approach the values characteristic of their newly acquired host only after hundreds of millions of years [5]. This fact suggests, first, that differences in codon usage must have arisen relatively early during prokaryotic evolution, and, second, that the pressure for a newly acquired gene to assume the codon usage of the host organism is minimal. The fact that certain genes exhibiting a relatively high level of specific rare codons can nevertheless be expressed at high levels when cloned behind a strong promoter (see, for example, [6]) has led some investigators to suggest that the use of rare codons does not in fact serve a regulatory function [4]. It should be pointed out in this regard that the inability to demonstrate a regulatory effect with one set of genes expressed under a given set of experimental conditions does not rule out the possibility of an analogous regulatory function for another set of genes expressed preferentially under a different set of conditions. Below I summarize evidence suggesting that various specialized functions, expressed in a variety of bacteria, may be regulated at the translational level by selective use of rare codons in relevant structural genes (see Table 1).

3. Rare codon usage in phototrophic vs. heterotrophic genes in *Rhodobacter*

In 1991, Wu and Saier noted that genes encoding proteins of the photosynthetic apparatuses (reaction center and light harvesting proteins) of the Gram-negative purple bacteria, *Rhodobacter capsulatus* and *R. spheroides*, differed in codon usage from that of genes encoding enzymes of the fructose utilization pathway [7]. While most codons occurred with similar frequencies in these two groups of genes, a few were found to predominate, or be present exclusively in one or the other group (see Table 2 for representative examples). Moreover, other genes, such as those involved in nitrogen utilization or carotenoid biosynthesis, that were expressed under both

SSDI 0014-5793(95)00185-9

Table 1
Selective use of rare codons postulated to control specialized functions in bacteria

Function	Organism	Codon	Amino Acid
Photosynthesis	*Rhodobacter capsulatus*	GCU	Ala
		CUC	Leu
Fructose utilization	*Rhodobacter capsulatus*	AAU	Asn
		UGU	Cys
Solventogenesis	*Clostridium acetobutylicum*	ACG	Thr
Arial mycelium development	*Streptomyces coelicolor*	UUA	Leu
Fimbrial production	*Escherichia coli*	UUG	Leu

heterotrophic and phototrophic conditions, exhibited rare codon usage frequencies that were intermediate between those found in the photosynthetic and fructose-catabolic genes (Table 2). These differences were shown to be statistically significant. It was suggested that different tRNA pools were present under phototrophic vs. heterotrophic growth conditions, and that growth conditions might influence the relative rates of transcription of the tRNA genes and their cognate amino acyl tRNA synthetases. Differences in codon usage might generally allow operation of novel post-transcriptional regulatory mechanisms. It seemed reasonable to suppose that charged tRNA availability and codon usage could provide a safeguard against expression of specialized genes under inappropriate conditions [7].

4. Rare codon usage as a potential regulator of solventogenesis in *Clostridium*

Sauer and Dürre noted in 1992 that a mutational defect preceding the gene *thrA* encoding a rare tRNA, $tRNA^{thr}_{ACG}$ in the low GC Gram-positive bacterium, *Clostridium acetobutylicum*, gave rise to the absence of solventogenesis [8]. This strict anaerobe is a spore-forming bacterium that produces acetone and butanol only during a late stage in the growth cycle. The shift to solventogenesis is accompanied by a series of morphological and physiological changes in motility, shape, and granulose content, culminating in endospore formation. Sauer and Dürre noted that the ACG codon is rarely used and is largely restricted to genes either expressed at the end of exponential growth or involved in the inducible uptake or metabolism of minor carbon and nitrogen sources [8]. Because these investigators did not conduct statistical analyses, it was not possible to state that the observed differences in codon usage reflected a unique characteristic of specific groups of genes encoding specialized functions rather than depressed levels of expressivity [4]. Nevertheless, the potential implications of the observations were clear. As in the case of phototrophic vs. heterotrophic gene expression in *Rhodobacter*, codon usage in *Clostridium* may provide a safeguard to insure proper expression of certain stationary phase vs. log phase genes.

5. Codon usage as a determinant of the differentiated state in *Streptomyces*

Species of the high GC Gram-positive genus *Streptomyces* undergo fungal-like differentiation with the sequential formation of vegetative and aerial mycelia [9,10]. The fact that only the latter structures contain spores reflects the spatial and temporal constraints imposed upon the process of terminal differentiation within this genus. The industrial importance of these organisms is related to their capacity to produce an array of antibiotics and useful secondary metabolites during the post-exponential growth phase. Although these strict aerobes have many of the enzymatic attributes of their low GC Gram-positive cousins, their regulatory mechanisms appear to be remarkably different [11–13].

Leskiw et al. [14] and Fernández-Moreno et al. [15] first observed that, in *Streptomyces coelicolor*, a genetic defect in the gene *bldA*, encoding a rare tRNA, $tRNA^{leu}_{UUA}$ [16,17], blocked aerial mycelium formation and prevented efficient phenotypic expression of several genes containing the rare UUA codon. *bldA* mutations (including deletions) did not interfere with vegetative growth but did prevent aerial mycelium formation and antibiotic production (see [18] for a review). It was suggested that this rare codon occurred preferentially in genes concerned with differentiation and antibiotic production as contrasted with those required for vegetative growth.

More recently, evidence was presented suggesting that mature $tRNA^{leu}_{UUA}$ accumulates in ever increasing amounts as *S. coelicolor* cultures age, and that the temporally regulated accumulation of this mature tRNA species correlates with an increase in efficiency of UUA-containing messenger RNA transcription and/or translation ([19]; but see also [20]). It seemed to exert regulatory effects on events occurring during late growth, including morphological differentiation and antibiotic production.

6. Rare codon usage and the control of fimbrial production in *E. coli*

A recently noted example of potential rare codon control of

Table 2
Examples of differential codon usage in photosynthetic versus heterotrophic genes in *Rhodobacter*

	Codon	Amino acid	Fractional codon usage for each amino acid			
			fru	*pho*	*nif*	Total
Elevated codon usage in *pho* genes	GCU	Ala	0.02	0.11	0.03	0.05
	CUC	Leu	0.11	0.35	0.16	0.19
Elevated codon usage in *fru* genes	AAU	Asn	0.44	0.05	0.18	0.18
	UGU	Cys	0.25	0.00	0.13	0.12

Data were taken from [7]. Numerical values provide the fractional codon usage for each of the four amino acids (Ala, Asn, Cys, and Leu) in the four catagories indicated with abbreviations as follows: *fru*, fructose utilization (heterotrophic) genes, total of 1724 codons analyzed for *R. capsulatus*; *pho*, photosynthetic genes, total of 1976 codons analyzed for *R. capsulatus*; *nif*, nitrogen utilization genes (mostly concerned with nitrogen fixation), total of 3824 codons analyzed for *R. capsulatus*; Total, all available sequenced genes for *Rhodobacter* species at the time of analysis [7].

specialized gene expression concerns the production of type 1 fimbriae in the Gram-negative enteric bacterium, *Escherichia coli* strain F18, which is able to colonize the mouse colon [21]. Burghoff et al. [22] isolated a 6.5 kb *E. coli* sequence that enhanced the colonizing ability of strain F18 and simultaneously stimulated synthesis of type 1 fimbriae. The gene responsible for this stimulation proved to be the *leuX* gene, encoding a tRNA specific for the rare leucine codon UUG. This gene is in single copy at 97 min on the *E. coli* chromosome, and the encoded tRNA species (LeuX) is apparently dispensable for growth [23]. No effect on growth rate was observed when *leuX* was mutated [24]. Another $tRNA^{leu}$, LeuZ, specific for the UUA leucine codon, presumably recognizes UUG by 'wobble', and can thereby substitute adequately for LeuX, at least with respect to the expression of genes encoding functions required for vegetative growth.

The mechanism by which *leuX* gene expression influences type 1 fimbrial production is probably complex. The *fimA* gene, encoding the principal type 1 fimbrilin, lacks UUG codons altogether [25]. However, synthesis of type 1 fimbriae is subject to phase variation due to inversion of a 314 bp DNA segment that includes the *fimA* promoter ([26]; but see also [27]). The ratio of the products of two *fim* genes, *fimB* and *fimE*, determine the frequencies of inversion in the two opposing directions with high levels of FimB favoring the 'off' to 'on' transition. Since *fimB* has six UUG codons while *fimE* has only two [28], it has been proposed that LeuX influences type 1 fimbrial production by controlling *fimB* expression more stringently than that of *fimE* [29]. In this regard it is interesting to note that *leuX* expression is apparently regulated by two proteins (of 22 and 26 kDa) encoded by genes adjacent to *leuX*. Deletion analyses have suggested that the 22 kDa protein is a transcriptional activator while the 26 kDa protein is a repressor of *leuX* expression. These proteins may therefore be indirect regulators of type 1 fimbrial phase variation, and consequently of net fimbrial production.

Various *E. coli* strains are collectively capable of producing at least six distinct virulence-related fimbriae, each exhibiting specificity for and mediating adhesion to a specific mammalian cell surface macromolecule [30,31]. Expression of these fimbriae is often subject to phase variation in agreement with the belief that successful colonization of the host depends on the timely expression and subsequent silencing of specific virulence-related genes, depending on the stage of infection. A recent analysis has revealed that the *leuX* gene of uropathogenic *E. coli* strain 536 encompasses one of several sites responsible for genetic instability [32]. Internal to *leuX* is one of two 18-nucleotide direct repeats that serve as functional sites for excision of a 190 kb DNA segment. This segment encodes, among other functions, P-related fimbriae. Excision of this DNA segment silences expression of *leuX* (possibly controlling type 1 fimbrial synthesis, as noted above) as well as expression of the genetic apparatus encoding P-related fimbriae. As bacterial cells lacking 'excess DNA baggage' and incapable of making fimbriae divide with increased growth rates, it may be that timely excision provides the bacterium that has already established itself in the host organism with pathogenic advantage [32]. Based on the proposed regulatory role of rare tRNAs in controlling fimbrial production, we suggest that it was not accidental that tRNA loci have come to serve as sites of virulence-associated DNA insertion/deletion phenomena.

7. Conclusions and perspectives

How important are the postulated regulatory mechanisms giving rise to codon-controlled phenotypic gene expression? Are they generally operative for the control of starvation-induced or stress-related vs. vegetative gene expression in *E. coli* and other bacteria [33]? Do they function to safeguard proper temporal expression of sporulation (*spo*)-specific genes at any one stage or during several different stages in the well-defined programs of differentiation of various *Bacillus* species [34]? Do they play a role in the control of growth phase-specific or condition-selective gene expression, e.g. expression of genes concerned with bioluminescence in *Vibrio* species [35,36], bacteriorhodopsin-mediated photosynthesis in archaebacteria [37], or induction of virulence-specific genes in bacterial pathogens of plants, animals and other bacteria [38–41]?

The first step towards answering these important questions would seem to be to analyze functionally related groups of bacterial genes for statistically significant differences in codon usage, as reported by Wu and Saier [7] for the photosynthetic vs. heterotrophic genes of *Rhodobacter*. A second step would be to measure variations in the cellular concentrations of specific rare tRNA species made under relevant but differing physiological conditions. The third step would be to establish a causal relationship between rare codon occurance, tRNA level and gene expressivity. Such studies may lead to recognition of novel codon usage-mediated mechanisms for ensuring the proper expression of temporally and spatially regulated genes in prokaryotic microorganisms. The relevance of such mechanisms to eukaryotic organisms, including protozoa, fungi, plants and animals, could then be ascertained by the application of straight-forward comparative approaches.

Acknowledgements: I thank Lisa Preble for expert assistance in the preparation of this manuscript. Work in the author's laboratory was supported by Public Health Service Grants 5RO1AI21702 and 2RO1AI14176.

References

[1] Brinkman, U., Mattes, R.E. and Burke, P. (1989) Gene 85, 109–114.
[2] Holm., L. (1986) Nucleic Acids Res. 14, 3075–3087.
[3] Sharp, P.M. and Li, W.H. (1987) Nucleic Acids Res. 15, 1281–1295.
[4] Andersson, S.V.E. and Kurland, C.G. (1990) Microbiol. Rev. 54, 198–210.
[5] Groisman, E.A., Sturmoski, M.A, Solomon, F.R., Lin, R. and Ochman, H. (1993) Proc. Natl. Acad. Sci. USA 90, 1033–1037.
[6] Robinson, M., Lilley, R., Little, S., Emtage, J.S., Yarranton, G., Stephens, P., Millican, A., Easton, M. and Humphreys, G. (1994) Nucleic Acids Res. 12, 6663–6667.
[7] Wu, L.-F. and Saier Jr., M.H. (1991) Res. Microbiol. 142, 943–949.
[8] Sauer, U. and Dürre, P. (1992) FEMS Microbiol. Lett. 100, 147–154.
[9] Chater, K.F. (1989) in: Genetics and Molecular Biology of Industrial Microorganisms (Hershberger, C.L., Queener, S.W. and Hegeman, G. eds.) pp. 99–107, American Society for Microbiology, Washington, DC.
[10] Chater, K.F. (1993) Annu. Rev. Microbiol. 47, 668–713.
[11] Kwakman, J.H.J.M. and Postma, P.W. (1994) J. Bacteriol. 176, 2694–2698.
[12] Titgemeyer, F., Walkenhorst, J., Cui, X., Reizer, J. and Saier Jr., M.H. (1994) Res. Microbiol. 145, 89–92.

[13] Titgemeyer, F., Walkenhorst, J., Reizer, J., Stuiver, M.H., Cui, X. and Saier Jr., M.H. (1995) Microbiology 141, 51–58.
[14] Leskiw, B.K., Lawlor, E.J., Fernandez-Abalos, J.M. and Chater, K.F. (1991) Proc. Natl. Acad. Sci. USA 88, 2461–2465.
[15] Fernández-Moreno, M.A., Caballero, J.L., Hopwood, D.A. and Malpartida, F. (1991) Cell 66, 769–780.
[16] Bibb, M.J., Findlay, P.R., and Johnson, M.W. (1984) Gene 30, 157–166.
[17] Wright, F. and Bibb, M.J. (1992) Gene 113, 55–65.
[18] Leskiw., B.K., Bibb, M.J. and Chater, K.F. (1991) Mol. Microbiol. 5, 2861–2867.
[19] Leskiw, B.K., Mah, R., Lawlor, E.J. and Chater, K.F. (1993) J. Bacteriol. 175, 1995–2005.
[20] Gramajo, H.C., Takano, E. and Bibb, M.J. (1993) Mol. Microbiol. 7, 837–845.
[21] Cohen, P.S., Rossoll, R., Cabelli, V.J., Yang, S.L. and Laux, D.C. (1983) Infect Immun. 40, 62–69.
[22] Burghoff, R.L., Pallesen, L., Krogfelt, K.A., Newman, J.V., Richardson, M., Bliss, J.L., Laux, D.C. and Cohen, P.C. (1993) Infect. Immun. 61, 1293–1300.
[23] Yoshimura, M., Inokuchi, H. and Oxeki, H. (1984) J. Mol. Biol. 177, 627–644.
[24] Komine, Y., Adachi, T., Inokuchi, H. and. Ozeki, H. (1990) J. Mol. Biol. 212, 579–598.
[25] Klemm, P. (1984) Eur. J. Biochem. 143, 395–399.
[26] Abraham, J.M., Freitag, C.S., Clements, J.R. and Eisenstein, B.I. (1989) Proc. Natl. Acad. Sci. USA 82, 5724–5727.
[27] McClain, M., Blomfield, I.C., Eberhardt, K.J. and Eisenstein, B.I. (1993) J. Bacteriol. 175, 4335–4344.
[28] Klemm, P. (1986) EMBO J. 5, 1389–1393.
[29] Newman, J.V., Burghoff, R.L., Pallesen, L., Krogfelt, K.A., Kristensen, C.S., Laux, D.C. and Cohen, P.S. (1994) FEMS Microbiol. Lett. 122, 281–287.
[30] Van Rosmalen, M. and Saier Jr., M.H. (1993) Res. Microbiol. 144, 507–527.
[31] Klemm, P. (1994) Fimbriae: Adhesion, Genetics, Biogenesis, and Vaccines, CRC Press, Boca Raton, FL.
[32] Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschäpe, H. and Hacker, J. (1994) Infect. Immun. 62, 606–614.
[33] Matin, A. (1990) FEMS Microbiol. Ecol. 74, 185–196.
[34] Losick, R. and Stragier, P. (1992) Nature 355, 601–604.
[35] O'kane, D.J. and Prasher, D.C. (1992) Mol. Microbiol. 6, 443–449.
[36] Meighen, E.A. (1993) FASEB J. 7, 1016–1022.
[37] Kuan, G. and Saier Jr., M.H. (1994) Res. Microbiol. 145, 273–285.
[38] Thomashow, M.F. and Cotter, T.W. (1992) J. Bacteriol. 174, 5767–5771.
[39] Gross, R. (1993) FEMS Microbiol. Rev. 104, 301–326.
[40] Briggs, S.P. and Johal, G.S. (1994) Trends Genet. 10, 12–16.
[41] Saier Jr., M.H. (1994) FEBS Lett. 337, 14–17.