



## Selective Charging of tRNA Isoacceptors Explains Patterns of Codon Usage

Johan Elf *et al.*

*Science* **300**, 1718 (2003);

DOI: 10.1126/science.1083811

*This copy is for your personal, non-commercial use only.*

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of August 17, 2012):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/300/5626/1718.full.html>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/content/suppl/2003/06/12/300.5626.1718.DC1.html>

This article has been **cited by** 48 article(s) on the ISI Web of Science

This article has been **cited by** 23 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/300/5626/1718.full.html#related-urls>

from careful regulation of the minus end.

The extent of the treadmilling, the creation of microtubule bundles through treadmilling motility, and the absence of other observed mechanisms for polymer repositioning together suggest that treadmilling motility makes an important contribution to the organization of the cortical array.

#### References and Notes

1. C. Lloyd, J. Chan, *Plant Cell* **14**, 2319 (2002).
2. R. J. Cyr, *Annu. Rev. Cell Biol.* **10**, 153 (1994).
3. A. Bichet, T. Desnos, S. Turner, O. Grandjean, H. Hofte, *Plant J.* **25**, 137 (2001).
4. A. T. Whittington *et al.*, *Nature* **411**, 610 (2001).
5. S. Thitamadee, K. Tuchihiro, T. Hashimoto, *Nature* **417**, 193 (2002).
6. T. I. Baskin, *Protoplasma* **215**, 150 (2001).
7. P. B. Green, *J. Cell Biol.* **27**, 343 (1965).
8. J. M. Hush, P. Wadsworth, D. A. Callahan, P. K. Hepler, *J. Cell Sci.* **107**, 775 (1994).
9. C. L. Granger, R. J. Cyr, *Protoplasma* **216**, 201 (2001).
10. M. Yuan, P. J. Shaw, R. M. Warn, C. W. Lloyd, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6050 (1994).
11. G. O. Wasteneys, *J. Cell Sci.* **115**, 1345 (2002).
12. L. Clayton, C. M. Black, C. W. Lloyd, *J. Cell Biol.* **101**, 319 (1985).
13. H. Shibaoka, in *The Cytoskeletal Basis of Plant Growth and Form*, C. W. Lloyd, Ed. (Academic Press, London, 1991), pp. 159–168.
14. S. M. Wick, *Cell Biol. Int. Rep.* **9**, 357 (1985).
15. B. E. Gunning, S. M. Wick, *J. Cell Sci. Suppl.* **2**, 157 (1985).
16. M. Erhardt *et al.*, *J. Cell Sci.* **115**, 2423 (2002).
17. A. R. Hardham, B. E. Gunning, *J. Cell Biol.* **77**, 14 (1978).
18. J. Chan, C. G. Jensen, L. C. Jensen, M. Bush, C. W. Lloyd, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14931 (1999).
19. W. M. Saxton *et al.*, *J. Cell Biol.* **99**, 2175 (1984).
20. See supporting data on Science Online.
21. T. Mitchison, M. Kirschner, *Nature* **312**, 237 (1984).
22. C. M. Waterman-Storer, E. D. Salmon, *Curr. Biol.* **7**, R369 (1997).
23. T. J. Keating, J. G. Peloquin, V. I. Rodionov, D. Momicilovic, G. G. Borisy, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5078 (1997).
24. C. M. Waterman-Storer, E. D. Salmon, *J. Cell Biol.* **139**, 417 (1997).

25. V. I. Rodionov, G. G. Borisy, *Science* **275**, 215 (1997).
26. A. M. Yvon, P. Wadsworth, *J. Cell Sci.* **110**, 2391 (1997).
27. I. A. Vorobjev, V. I. Rodionov, I. V. Maly, G. G. Borisy, *J. Cell Sci.* **112**, 2277 (1999).
28. V. Rodionov, E. Nadezhdina, G. Borisy, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 115 (1999).
29. We thank D. Allen for help in collecting and propagating seed stocks; T. Salmon, C. Somerville, W. Briggs, T. Stearns, J. Theriot, and S. Cutler for useful discussions and comments; and S. Long for her generous support. Supported by the Carnegie Institution of Washington, the Howard Hughes Medical Institute, and U.S. Department of Energy grant DE-FG03-90ER2001.

#### Supporting Online Material

[www.sciencemag.org/cgi/content/full/1083529/DC1](http://www.sciencemag.org/cgi/content/full/1083529/DC1)

SOM Text

Materials and Methods

References

Movies S1 to S5

14 February 2003; accepted 16 April 2003

Published online 24 April 2003;

10.1126/science.1083529

Include this information when citing this paper.

# Selective Charging of tRNA Isoacceptors Explains Patterns of Codon Usage

Johan Elf,<sup>1</sup> Daniel Nilsson,<sup>2</sup> Tanel Tenson,<sup>3</sup> Måns Ehrenberg<sup>1\*</sup>

We modeled how the charged levels of different transfer RNAs (tRNAs) that carry the same amino acid (isoacceptors) respond when this amino acid becomes growth-limiting. The charged levels will approach zero for some isoacceptors (such as tRNA<sup>Leu</sup><sub>2</sub>) and remain high for others (such as tRNA<sup>Leu</sup><sub>4</sub>), as determined by the concentrations of isoacceptors and how often their codons occur in protein synthesis. The theory accounts for (synonymous) codons for the same amino acid that are used in ribosome-mediated transcriptional attenuation, the choices of synonymous codons in trans-translating transfer-messenger RNA, and the overrepresentation of rare codons in messenger RNAs for amino acid biosynthetic enzymes.

There are 61 sense codons in the genetic code for only 20 amino acids (1). One amino acid is therefore often encoded by several synonymous codons. These can be read by one transfer RNA (tRNA) or by several isoaccepting tRNAs with sometimes overlapping codon specificity (2). When there are several synonymous codons for an amino acid, these do not occur with equal probability in genes (3) or in ribosome-bound messenger RNAs (mRNAs) (4); there is a distinct codon bias, which can vary from species to species. The concentration of a tRNA isoacceptor is often positively correlated to the frequency of the synonymous codon that it reads (4, 5). In highly expressed genes, one codon in each syn-

onymous set of codons tends to dominate in frequency, and this codon is often read by the most abundant tRNA isoacceptor (6). It has been shown that such an arrangement will lead to more rapid protein synthesis and to a lower level of amino acid substitution errors than when the frequencies of synonymous codons and concentrations of tRNA isoacceptors are more evenly distributed (7). It is therefore likely that codon bias and tRNA isoacceptor concentrations have coevolved, and that the selection pressure for this coevolution is more pronounced for genes with high expression levels than with low expression levels (7, 8). Such observations have been used to define a codon adaptation index (CAI), which is now commonly used to identify highly expressed genes from genome sequence data (9). Until now, codon usage patterns have been interpreted in the context of tRNAs that are always fully charged with their amino acids. However, in many cases, cells are subjected to amino acid starvation, such as when living conditions are harsh or when the environment suddenly deteri-

orates. We investigated whether the same rules for codon adaptation apply when the amino acid supply is scarce as when it is abundant. We suggest that codon adaptation under amino acid limitation must fulfill very different demands, and we identify two main principles that have shaped codon usage under such conditions: (i) When codon reading is part of a control loop that regulates synthesis of the missing amino acid, the translation rate of the selected codon should be as sensitive as possible to starvation. (ii) When, in contrast, de novo synthesis of a protein or a peptide is essential during amino acid starvation, translation of its mRNA should be fast. This means that the rate of codon translation should be as insensitive as possible to amino acid deficiency and also should remain relatively high under amino acid starvation.

We show that when there are several tRNA isoacceptors that read synonymous codons, the rate of translation of some of these codons will be very sensitive to amino acid starvation, whereas the reading of others can be very insensitive. We describe a simple rule of supply and demand that determines whether a codon is sensitive or insensitive to amino acid supply, and this rule is used to predict the choice of codons in transcriptional attenuation for the control of amino acid production in bacteria, the anomalously high frequency of some rare codons in amino acid synthetic operons, and the choice of synonymous codons in the open reading frame of transfer-messenger RNA (tmRNA). Our findings motivate the construction of a starvation CAI (sCAI) that can be used to identify genes that must be expressed under starvation conditions.

**The mechanism behind selective charging of tRNA isoacceptors during amino acid starvation.** The rate of translation of a particular codon in the cell is determined by the level of the charged tRNA isoacceptor (aminoacyl-tRNA) that reads it after entering the mRNA-programmed ribosome in a ternary complex

<sup>1</sup>Department of Cell and Molecular Biology, Uppsala University, Biomedical Center, Box 596, 751 24 Uppsala, Sweden. <sup>2</sup>Center for Genomics and Bioinformatics, Karolinska Institutet, 17177 Stockholm, Sweden. <sup>3</sup>Institute of Molecular and Cell Biology, Tartu University, Riia 23, Tartu, Estonia.

\*To whom correspondence should be addressed. E-mail: ehrenberg@xray.bmc.uu.se

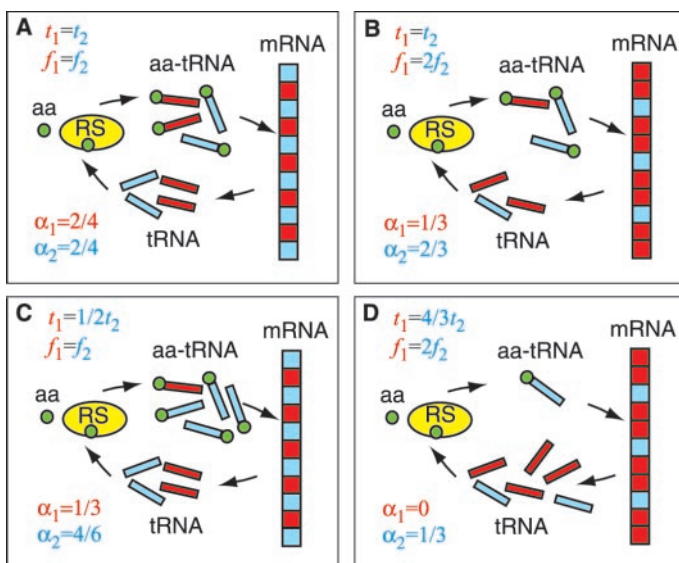
with an elongation factor (EF-Tu in eubacteria or eEF1 $\alpha$  in eukaryotes) and guanosine triphosphate. This is also formally true when the cells are starved for that amino acid, but in this case, the rate of codon reading is only determined by the rates of supply and demand for charged tRNAs, to which the concentration of ternary complexes passively adjusts.

The basic principle can be illustrated (Fig. 1) by a case with two different tRNA isoacceptors (red and blue bars) that read two different codons (red and blue squares) for the same amino acid. The tRNAs are charged with amino acid (green) by their common aminoacyl-tRNA synthetase (yellow) with, as assumed here, identical kinetic parameters. The total concentrations of isoacceptors are  $t_1$  and  $t_2$  (total numbers of red or blue bars); their charged concentrations (number of bars with solid green circles) are  $\alpha_1 t_1$  and  $\alpha_2 t_2$ , where  $\alpha_i$  is the fraction of tRNA isoacceptor  $i$  that is charged; and the concentrations of their uncharged forms (number of bars without circles) are  $(1 - \alpha_1)t_1$  and  $(1 - \alpha_2)t_2$ , respectively. The rate of charging of each tRNA isoacceptor is proportional to the concentration of its free form, and the constant of proportionality is assumed to be the same for both isoacceptors. The steady-state flow into ribosomes from each tRNA isoacceptor pool is equal to the average rate of protein synthesis multiplied by its codon frequency ( $f_1$  or  $f_2$ ). Because the steady-state flow of supply must equal the steady-state flow of consumption for each charged isoacceptor, the ratio between their rates of charging and consumption must be equal to the ratio between the frequencies of their respective codons on translating ribosomes. That is, for the supply of the charged tRNAs to meet the demand in pro-

tein synthesis, the following relation must hold

$$\frac{(1 - \alpha_1)t_1}{(1 - \alpha_2)t_2} = \frac{f_1}{f_2} \Leftrightarrow \alpha_2 = 1 - \frac{t_1 f_2}{t_2 f_1} (1 - \alpha_1) = 1 - q(1 - \alpha_1)$$

To see what this means, we first chose indices so that tRNA isoacceptor 1 has the smallest ratio between total tRNA concentration and codon frequency ( $t_1/f_1 < t_2/f_2$  and  $q < 1$ , where  $q = (t_1/f_1)/(t_2/f_2)$ ). When amino acid starvation gets more and more severe, the charged level of tRNA isoacceptor 1 will approach zero, and the charged level of isoacceptor 2 will approach  $1 - q$ . This implies that the isoacceptor with the smallest ratio between total tRNA concentration and codon frequency will totally lose its charging, whereas the charged level of the other isoacceptor may end up at any value between 0 and 100%. The reason for this unexpected behavior is that the rates of supply (proportional to uncharged tRNAs) and demand (proportional to codon frequencies) are always matched at a lower charging level for isoacceptor 1 than for isoacceptor 2. In the special case in which the rate of amino acid supply approaches zero, this means that the charged level of isoacceptor 1 can disappear, whereas isoacceptor 2 can remain almost fully charged. Accordingly, a codon of the type read by isoacceptor 1 should be chosen when the rate of codon reading is used for regulatory purposes to sense amino acid limitation, whereas a codon of the type read by isoacceptor 2 should be used in genes that must be efficiently expressed when the amino acid is in limited supply.



**Fig. 1.** Principle of selective charging of tRNA isoacceptors during amino acid limitation. (A) The number of red and blue squares (codons) and the total number of red and blue bars (isoacceptor tRNAs) are the same; codon usage is balanced to isoacceptor concentrations ( $q = 1$ ), and  $\alpha_i$  is always equal to  $\alpha_2$ . (B) The total numbers of bars are equal, but the number of red squares is two times as large as the number of blue squares ( $q = 0.5$ ). (C) The numbers of red and blue squares are equal, but the number of red bars is half the number of blue bars ( $q = 0.5$ ). (D) There are four red and three blue bars, and there are twice as many red squares as there are blue squares [ $q = (4 \times 1)/(3 \times 2) = 2/3$ ]. When  $q$  is smaller than 1, the proportionality between supply rates and consumption rates is maintained with a larger fraction of charged isoacceptor 2 than of isoacceptor 1 ( $\alpha_2 > \alpha_1$ ). When amino acid starvation becomes increasingly severe,  $\alpha_1$  approaches zero and  $\alpha_2$  approaches  $1 - q$ .

**Selective charging of tRNA isoacceptors and codon usage in *Escherichia coli*.** The total tRNA concentrations and the codon usage frequencies have been estimated for *Escherichia coli* cells growing in different media (4, 5), and the codon specificities of different isoacceptors have been characterized (2). We used these data (Fig. 2 and Table 1) to predict how the charging levels of individual isoacceptors that are cognate

**Table 1.** Sensitivity of the codon translation rate compared with the rate of amino acid supply. The table displays the relative change in codon translation rate normalized to the relative change in the supply of its cognate amino acid (sensitivity amplification) (see SOM text). Codon usage in translation (4) is also shown. Regulatory codons used by *E. coli* in attenuation of transcription are in bold. The growth condition for determination of codon usage frequencies and tRNA concentrations are the same as described in Fig. 2.

Amino acid	Codon	Sensitivity	Codon usage ( $\times 10^{-3}$ )
Threonine	ACG	2.4	7.5
	ACA	5.7	3.5
	ACU	6.6	13.9
	<b>ACC</b>	<b>20.9</b>	<b>26.5</b>
Leucine	UUG	0.59	6.3
	UUA	3.0	6.1
	CUG	5.5	60.1
	CUC	24.8	6.2
	CUU	24.8	5.7
	<b>CUA</b>	<b>26.9</b>	<b>2.2</b>
Serine	AGU	3.4	4.0
	AGC	3.4	12.0
	UCG	4.4	6.1
	UCA	7.5	3.9
	UCU	7.9	13.1
	UCC	35.5	11.2
Arginine	AGG	0.0	0.1
	AGA	0.5	1.1
	CGG	1.3	1.8
	CGA	18.0	1.3
	CGU	18.0	31.1
	CGC	18.0	22.3
Alanine	GCG	1.9	30.3
	GCA	1.9	22.1
	GCU	2.0	28.9
	GCC	21.5	19.8
	<b>GUC</b>	<b>19.2</b>	<b>11.3</b>
Valine	GUU	7.7	31.3
	GUA	10	15.9
	<b>GUG</b>	<b>10</b>	<b>21.4</b>
	<b>GUC</b>	<b>19.2</b>	<b>11.3</b>
Glycine	GGG	0.3	4.8
	GGA	0.6	2.7
	GGC	12.5	35.6
	GGU	12.5	38.3
Proline	CCC	2.1	3.3
	CCU	3.7	5.0
	CCG	13.3	29.5
	CCA	32.0	6.5
Isoleucine	AUA	2.0	0.9
	<b>AUU</b>	<b>15.2</b>	<b>21.4</b>
	<b>AUC</b>	<b>15.2</b>	<b>36.7</b>
Glutamine	CAG	14.4	29.2
	CAA	32.2	10.2



## RESEARCH ARTICLES

to the same amino acid vary when the rate of supply of this amino acid approaches zero. Because tRNA isoacceptors in many cases have overlapping codon specificities (2), the steady-state theory that links the degree of amino acid limitation to the charged isoacceptor levels is more complex [see supporting online material (SOM) text] than the simple example above, but the principle remains unaltered. We assumed the kinetic parameters for charging of tRNAs and codon reading by ternary complexes (10) to be the same in each family of isoacceptors. We also neglected the fractions of tRNA molecules that are in a complex with aminoacyl-tRNA synthetases or ribosomes, a simplification that leads to the exact result when synthetases are kinetically neutral to their respective isoacceptors, and the distances between identical codons on mRNAs are statistically unbiased (SOM text). The outcome of this steady-state analysis is shown in Fig. 2. In 8 of the 10 cases in which different isoacceptors read different codons, at least one isoacceptor remained at a high charging level all the way to the limit of zero rate of supply of its cognate amino acid. The two exceptions are Gln and Val, for each of which the charging levels of two isoacceptors decreased uniformly toward zero as the supply of amino acid dwindled. In most cases, it is the charging level of major tRNAs reading abundant codons that approaches

zero when the amino acid supply decreases, but there are notable exceptions. For instance, the minor tRNA<sub>3</sub><sup>Leu</sup> and tRNA<sub>3</sub><sup>Pro</sup> read not only their own rare codons but also an abundant codon together with a major tRNA. When the rate of supply of one of these amino acids approaches zero, the charging level of both a major and a minor tRNA will decrease to very low values.

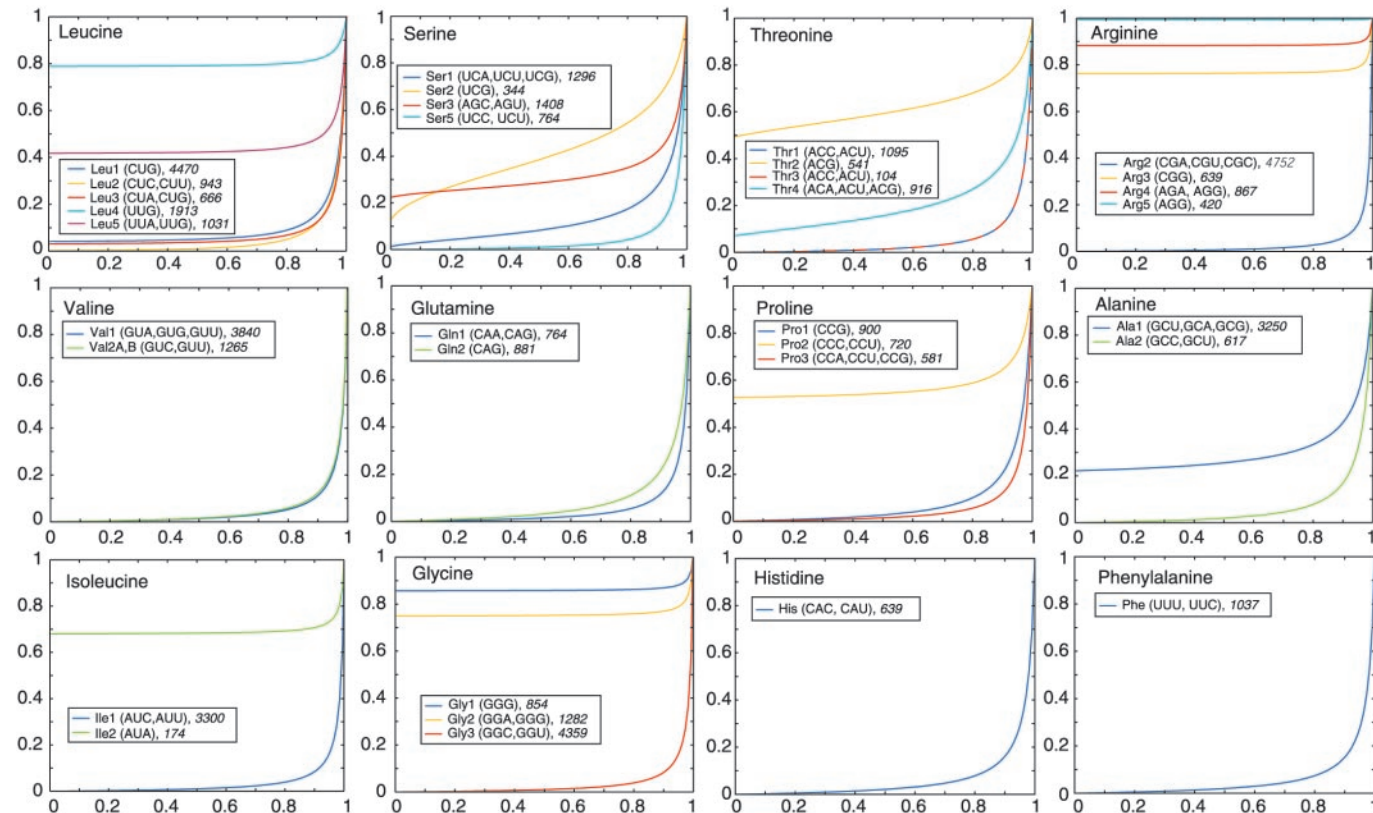
These predictions are consistent with early and until now unexplained experimental data demonstrating that *E. coli* cells subjected to severe amino acid starvation have high residual charging levels for several tRNA isoacceptor families (11–13). Our explanation for these findings is that the total rate of protein synthesis in the cell slows down to match the limited rate of supply of the starved amino acid by very slow translation of some of its synonymous codons. Other codons in the synonymous set, in contrast, continue to be read rapidly by isoacceptors with relatively high charging levels. Our predictions of residual charging levels, summed over all isoacceptors in each family with more than one member, are Gln 0%, Val 0%, Tyr 0%, Ile 3%, Ser 10%, Thr 12%, Pro 17%, Ala 18%, Leu 24%, Arg 25%, and Gly 26% (Fig. 2).

Several testable predictions can be made from our theory. For instance, when cells are starved for an amino acid, the missense error rates at codons read by isoacceptors for which the charg-

ing levels go to very low levels will increase drastically in wild-type cells and increase even more in *relA* mutants lacking the stringent response (14, 15). The *relA* effect is predicted because the stringent response reduces the demand for a limiting amino acid in protein synthesis, thereby allowing the charging levels of its isoacceptors to increase (14). In contrast, the missense error rates at codons read by isoacceptors with high residual charging levels are expected to increase little when amino acid limitation rises and are not expected to depend on whether the genetic background is wild type or *relA*.

Sörensen (16) measured individual charging levels for the major isoacceptors tRNA<sub>2</sub><sup>Arg</sup>, tRNA<sub>1</sub><sup>Thr</sup>, and tRNA<sub>1</sub><sup>Leu</sup> in *E. coli* strains starved for Arg, Thr, or Leu, respectively. He found that each charging level was very low for wild-type bacteria and even lower for *relA* bacteria, in agreement with the predictions in Fig. 1, where the charging levels of these particular isoacceptors go to very low values when the supply rates of their cognate amino acids approach zero.

There is a narrow selection of regulatory synonymous codons in the leaders of mRNAs for those amino acid biosynthetic operons that are controlled by ribosome-mediated attenuation of transcription (17, 18). However, the principle underlying this selection of codons remains unclear. These regulatory codons encode amino



**Fig. 2.** Predicted tRNA isoacceptor charged levels ( $\alpha_i$ ) at varying rates of supply of amino acids. The x axis represents the total rate of protein synthesis during starvation for one amino acid normalized to the maximal rate. The value 0 corresponds to zero rate of supply and the value 1 corresponds to a saturating supply rate. The y axis represents the

percentages of charging of individual isoacceptors as functions of the amino acid supply rate. The numbers in the legends are the total number of the different isoacceptor tRNA molecules in an *E. coli* cell grown in acetate at 0.4 doubling per hour (4). The ratios [Ile1]/[Ile2] and [Gly1]/[Gly2] are from (5).

acids that are synthesized by the enzymes expressed from the controlled operon. When ribosomes are slow in translating such regulatory codons, such as when there is a deficient supply of their cognate amino acid, this signals to the RNA polymerase to continue transcription from the leader sequence into the structural genes of the operon. When, in contrast, there is a sufficient amino acid supply, allowing the ribosomes to translate the regulatory codons quickly, the expression of the coding sequences of the operon is prevented by the formation of a stem loop structure in the leader of the mRNA, which signals termination of transcription.

In *E. coli*, transcriptional attenuation is used to regulate expression from the *leu*, *thr*, *ilvGMEDA*, *ilvBN*, *trp*, and *his* operons, as well as the *pheA* gene (17). In the first four of these leader sequences, there are several isoacceptors for each synthesized amino acid. It has been suggested (17) that rare codons read by minor isoacceptors would be optimal as regulatory codons. However, this cannot be true in general, because it is only in the case of the *leu* and *ilvGMEDA* operons that a rare codon (CUA encoding Leu) is used in *E. coli* leader sequences.

For attenuation mechanisms to work well, the rates of translation of regulatory codons, and therefore the charging levels of the tRNAs that read them, must be sensitive to the rates of supply of their cognate amino acids. Accordingly, we suggest that the regulatory codon in a set of synonymous triplets should be the one for which the response in translation rate is most sensitive [highest sensitivity amplification (19)] to variation in the rate of supply of its cognate amino acid. Sensitivity amplification parameters were calculated (SOM text) from the data in Fig. 2 and are shown in Table 1. From the table, we predicted that the *leu* attenuator should have CUA; the *thr* attenuator ACC; and the *ilvGMEDA* and *ilvBN* attenuators should have CUA for Leu, AUU or AUC for Ile, and GUC for Val as regulatory codons. These predictions score

extremely well (Table 1) when compared to sequence data (17) showing that the *leu* attenuator uses four rare CUA codons and the *thr* attenuator uses eight major ACC codons. Furthermore, the *ilvGMEDA* and *ilvBN* attenuators use GUC or GUG for Val, the *ilvGMEDA* uses AUU or AUC for Ile, and *ilvBN* uses CUA and CUC for Leu.

The attenuator-controlled *his* and *pheA* operons are also interesting in this context. In these cases, a single tRNA isoacceptor reads both codons that encode the amino acid (UUU or UUC for Phe and CAC or CAU for His). Because we predicted that the sensitivity in these cases would be the same for the two codons, the evolutionary choice of regulatory codon should be neutral (Fig. 2). This is confirmed by sequence data, showing that the attenuator leader for *pheA* uses three UUU codons interchangeably with four UUC codons, and the leader for the *his* operon uses three CAC and four CAU codons.

The present analysis suggests that rare codons, read by minor tRNAs that retain high charging levels during starvation for their cognate amino acids (Fig. 2), can be the most rapidly translated codons during amino acid limitation. This means that these codons, which are read comparatively slowly under conditions of balanced growth, may become the most rapidly translated and the least error-prone in situations of starvation.

When bacterial cells are subjected to downshifts from a medium containing all amino acids to a medium lacking one or several of them, rapid production of the enzymes that synthesize these amino acids must be accomplished in their absence. It is then expected that codons read by tRNA isoacceptors that retain high charging levels during starvation for their cognate amino acid should be overrepresented in mRNAs for the enzymes that synthesize it. At the same time, rare codons read by tRNAs that lose their charging at the onset of amino acid starvation should be

underrepresented. These predictions were tested for the amino acid biosynthetic *arg*, *thr*, *leu*, and *ilv* genes, for which expression is upregulated more than 30 times when their respective amino acids become rate-limiting for protein synthesis (20). We observed (Table 2) that the usage of insensitive codons, read by tRNAs that are predicted to retain high charging levels under amino acid starvation, was significantly higher than the average usage (4) in protein synthesis at the low growth rate of 0.4 doubling per hour. The infrequent UUG and UUA codons for Leu, the rare ACG codon for Thr, the rare CGG codon for Arg, and the rare AUA codon for Ile are overrepresented by a factor of at least 2. At the same time, the rare Leu codon CUA, read by a minor tRNA that loses its charge when the Leu supply becomes rate-limiting (Fig. 2), is absent in the *leuA*, *leuB*, *leuC*, and *leuD* genes, although an unbiased estimate would predict the presence of three such codons.

Another case in which rapid translation of codons that are cognate to a limiting amino acid is essential concerns tmRNA, known to rescue ribosomes stalled at starved codons or at the end of truncated mRNAs (21, 22). During trans-translation, tmRNA adds a decapeptide to the C terminus of nascent proteins on stalled ribosomes, which tags the peptide chains for subsequent degradation by Clp and other proteases. The recent finding that the bacterial protein RelE truncates mRNA codon-specifically in the ribosomal A site suggested that RelE and tmRNA are part of a bacterial stress response that leads to rapid adaptation after nutritional downshifts (23). Two of the amino acids in the decapeptide encoded by tmRNA have several codons and isoacceptors; there are four Ala residues, encoded by three GCU and one GCA, and a single Leu residue encoded by UUA. All these codons are read by tRNA isoacceptors that we predict will have high charging levels during starvation for their cognate amino acids (Fig. 1). This suggests that these choices of codons have evolved to ensure rapid trans-translation during amino acid limitation, possibly as a way to optimize a RelE-dependent mechanism for fast recovery after the onset of amino acid starvation (23).

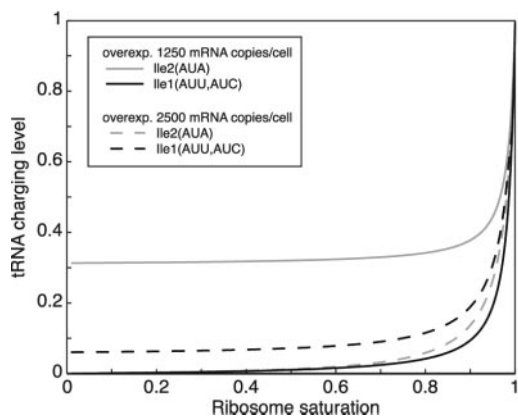
**Selective charging of tRNA isoacceptors modulated by overexpression.** There is one experimental result obtained for *E. coli* cells lacking the stringent response (*relA*) that superficially contradicts one of our predictions. Barak *et al.* (24) designed a  $\beta$ -galactosidase assay to detect frame shifting resulting from inefficient reading of A-site codons. They compared the frame-shifting propensity with either AUA or AUC in the ribosomal A site as a function of increasingly severe inhibition of the rate of charging of both tRNA<sub>1</sub><sup>Ile</sup>, reading AUC, and tRNA<sub>2</sub><sup>Ile</sup>, reading AUA. They found that the relative increase in frame shifting was much higher for AUA than for AUC codons, in contrast to our expectation that

**Table 2.** Codon usage for Leu, Thr, Arg, and Ile in genes for their respective biosynthetic enzymes. The codons are ordered in groups depending on their predicted sensitivity to starvation (Table 1) as follows: insensitive (sensitivity  $\leq 3$ ), intermediate ( $3 < \text{sensitivity} \leq 10$ ), and sensitive ( $10 < \text{sensitivity}$ ). The number of occurrences of the codons in the genes (28) is reported, together with the 95% confidence interval (see SOM text) for what is expected from codon usage in proteins (Table 1). The insensitive codons are used significantly more than the level expected in the average protein and the sensitive codons are used less.

Genes	Codon groups	Occurrences (n)	Expected
<i>leuABCD</i>	Insensitive (UUG,UUA)	28	8–24
	Intermediate (CUG)	87	73–92
	Sensitive (CUC,CUU,CUA)	4	12–27
<i>argIFEHB</i>	Insensitive (AGG,AGA,CGG)	14	0–7
	Sensitive (CGA,CGU,CGC)	62	69–76
<i>ilvBN, ilvGMEDA</i>	Insensitive (AUA)	8	0–5
	Sensitive (AUU,AUC)	142	145–150
<i>thrBC</i>	Insensitive (ACG)	9	1–8
	Intermediate (ACA,ACU)	9	6–16
	Sensitive (ACC)	14	11–22



**Fig. 3.** Effect of overexpression of a reporter gene with unusual Ile codon usage. The predicted charging levels for tRNA<sup>Ile</sup><sub>1</sub> (Ile1) and tRNA<sup>Ile</sup><sub>2</sub> (Ile2) are shown for the varying Ile supply and for two different expression levels of the reporter construct (1250 or 2500 mRNA copies per bacterial cell). The overexpressed mRNA has one AUU and one AUA codon. At about 2200 mRNA copies per cell, the model predicts a shift in which one of the charging levels goes to zero.



the charged level of tRNA<sup>Ile</sup><sub>2</sub> is little and the charged level of tRNA<sup>Ile</sup><sub>1</sub> is substantially reduced by strong Ile starvation (Fig. 2). However, it must be taken into account that in the AUA experiment the reporter gene had one AUU and one AUA codon in the nonshifted (+1) frame. Accordingly, overexpression of this construct changed the codon usage for Ile considerably, given that the rare AUA codon is normally used 50 times less frequently than other Ile codons. In fact, our theory predicts, in this case, weak frame shifting at AUA codons when the mRNA is expressed from low-copy-number plasmids and strong frame shifting when the same mRNA is expressed from high-copy-number plasmids, as found by Barak *et al.* (24). The expected change in the charging level pattern when the plasmid copy number is increased is illustrated in Fig. 3. For 1250 mRNAs per cell (low-copy-number plasmid), the charged level of tRNA<sup>Ile</sup><sub>2</sub> stays high and the charged level of tRNA<sup>Ile</sup><sub>1</sub> dwindles, but for 2500 mRNAs per cell (high-copy-number plasmid), the charged level of tRNA<sup>Ile</sup><sub>1</sub> stays high and the charged level of tRNA<sup>Ile</sup><sub>2</sub> approaches zero, as the charging efficiency of isoleucyl-tRNA synthetase decreases.

The predictions of the theory (Fig. 2) rely on the assumptions of kinetic equivalence [identical  $k_{cat}/K_m$  values, where  $k_{cat}/K_m$  is the effective association rate constant for the binding of a tRNA isoacceptor to its cognate synthetase (see SOM text)] of the tRNA species within each family of isoacceptors as they are charged by their aminoacyl-tRNA synthetase and, if there is overlap in codon reading by isoacceptors, when they interact with mRNA-programmed ribosomes. These kinetic features are largely unknown, and future results that challenge the assumption of kinetic equivalence will have a direct impact on the predicted charging levels. Furthermore, there may be some future changes in the way that synonymous codons are assigned to individual isoacceptors (2), and this could influence some of the curves in Fig. 2.

Application of the theory also depends on experimental estimates of tRNA isoacceptor concentrations and the frequencies by which codons occur on translating ribosomes. Two

groups (4, 5) have measured isoacceptor concentrations under similar conditions, and in some cases there are clear differences between their results, giving an indication of the precision of these estimates. From this, we foresee that adjustments of the numbers that enter the theory will be necessary, and these are likely to affect some of the results in Fig. 2.

#### A starvation codon adaptation index.

This theory, in spite of these assumptions and the uncertainty of the experimental data, has thrown light on codon usage in cases where no predictions could be made before. Its basic feature, that charging levels of isoacceptors will respond very differently to amino acid limitation, is robust to changes in parameter values and adds a dynamic aspect to the genetic code. That is, codons that are translated slowly under some conditions will be read most rapidly in other circumstances. In fact, the expectation that the reading of certain codons can occur with a high rate when there is virtually no supply of their cognate amino acids suggests that the prediction of genes for which expression is essential during amino acid limitation could be based on a suitably defined sCAI, in analogy with the CAI that has been successfully used to predict highly expressed genes (9).

Our theory has been tested and discussed in relation to experimental data from *E. coli*. However, aminoacylation of several tRNAs by the same enzyme and their deacylation in stoichiometric proportions during protein synthesis is a kinetic motif common to all kingdoms (25). Uneven charging of isoacceptors in response to amino acid limitation is therefore expected to be widespread, and its effects on codon bias in organisms other than *E. coli* constitute an interesting field of future research.

#### References and Notes

1. F. Crick, L. Barnett, S. Brenner, R. Watts-Tobin, *Nature* **192**, 1227 (1961).
2. G. Björk, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, F. Neidhardt *et al.*, Eds. (American Society for Microbiology Press, Washington, DC, 1996), pp. 861–886.

3. Y. Nakamura, *Nucleic Acids Res.* **24**, 214 (1996).
4. H. Dong, L. Nilsson, C. Kurland, *J. Mol. Biol.* **260**, 649 (1996).
5. T. Ikemura, *J. Mol. Biol.* **146**, 1 (1981).
6. L. E. Post *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1697 (1979).
7. M. Ehrenberg, C. G. Kurland, *Q. Rev. Biophys.* **17**, 45 (1984).
8. M. Bulmer, *Nature* **325**, 728 (1987).
9. P. M. Sharp, W. H. Li, *Nucleic Acids Res.* **15**, 1281 (1987).
10. For the predictions in Fig. 2,  $k_{ik} = 0$  for interactions between ternary complexes containing an isoacceptor *i* that cannot read a codon *k*. The parameter  $k_{ik}$  is  $k_{cat}/K_m$  for the binding of a ternary complex containing EF-Tu, guanosine triphosphate, and a charged tRNA isoacceptor *i* to a ribosome with a codon of type *k* in the A site (see SOM text). For isoacceptors that read the same codon, the  $k_{ik}$  parameters have been set as equal. The probability that the ribosome accepts a cognate isoacceptor after ternary complex association is very close to 100% (26). Therefore, only large kinetic differences between isoacceptors will lead to substantial differences in the  $k_{cat}/K_m$  values ( $k_{ik}$  parameters) for their interactions with the ribosome. Substantial deviations from uniform  $k_{cat}/K_m$  values within isoacceptor families would primarily affect residual charging levels during amino acid starvation and could also alter the predictions of which tRNA levels go to zero and which stay high. The  $k_{cat}/K_m$  values of charging ( $c_i$ ) of isoacceptors that are cognate to the same amino acid have been set as equal. In general, these parameters are not known, but similar aminoacylation kinetics have been found in our laboratory for all Leu isoacceptors (27).
11. C. D. Yegian, G. S. Stent, *J. Mol. Biol.* **39**, 45 (1969).
12. A. Bock, L. E. Faiman, F. C. Neidhardt, *J. Bacteriol.* **92**, 1076 (1966).
13. D. W. Morris, J. A. DeMoss, *J. Bacteriol.* **90**, 1624 (1965).
14. P. O'Farrell, *Cell* **14**, 545 (1978).
15. M. Cashel, D. R. Gentry, V. J. Hernandez, D. Vinella, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, F. Neidhardt *et al.*, Eds. (American Society for Microbiology Press, Washington, DC, 1996), pp. 1458–1496.
16. M. A. Sorensen, *J. Mol. Biol.* **307**, 785 (2001).
17. R. Landick, C. Yanofsky, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, F. Neidhardt *et al.*, Eds. (American Society for Microbiology Press, Washington, DC, 1987), pp. 1276–1301.
18. C. Yanofsky, *Nature* **289**, 751 (1981).
19. D. E. Koshland Jr., A. Goldbeter, J. B. Stock, *Science* **217**, 220 (1982).
20. F. Neidhardt *et al.*, Eds., *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (American Society for Microbiology Press, Washington, DC, 1996).
21. E. D. Roche, R. T. Sauer, *EMBO J.* **18**, 4579 (1999).
22. J. H. Withey, D. I. Friedman, *Curr. Opin. Microbiol.* **5**, 154 (2002).
23. K. Pedersen *et al.*, *Cell* **112**, 131 (2003).
24. Z. Barak, D. Lindsley, J. Gallant, *J. Mol. Biol.* **256**, 676 (1996).
25. M. Ibbá, D. Söll, *Annu. Rev. Biochem.* **69**, 617 (2000).
26. M. V. Rodnina, W. Wintermeyer, *Annu. Rev. Biochem.* **70**, 415 (2001).
27. M. Ehrenberg, unpublished data.
28. F. R. Blattner *et al.*, *Science* **277**, 1453 (1997).
29. We thank O. G. Berg, G. Björk, H. Bremer, R. H. Buckingham, J. Gallant, C. G. Kurland, J. Paulsson, D. Söll, M. Sorensen, and C. Yanofsky for helpful comments on the manuscript, and M. Lovmar for pointing out the tmRNA connection. Supported by the National Graduate School of Scientific Computing, the Wenner-Grenska Samfundet Foundation, and the Estonian Science Foundation.

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/300/5626/1718/DC1  
SOM Text

24 February 2003; accepted 25 April 2003