Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 1124–1128, February 1988
Evolution

Directional mutation pressure and transfer RNA in choice of the third nucleotide of synonymous two-codon sets

(G+C content of DNA/G·C/A·T pressure/tRNA constraints)

SYOZO OSAWA[†‡], TAKESHI OHAMA[†], FUMIAKI YAMAO[†], AKIRA MUTO[†], THOMAS H. JUKES[‡§], HARUO OZEKI[¶], AND KAZUHIKO UMESONO[¶]

[†]Laboratory of Molecular Genetics, Department of Biology, Nagoya University, Chikusa-ku, Nagoya 464, Japan; [§]Space Sciences Laboratory, University of California, Berkeley, CA 94720; and [¶]Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan

Communicated by Charles Sibley, October 1, 1987 (received for review July 29, 1987)

ABSTRACT **Bacterial species have diverged into a series of families, some with high G+C content in their DNA, and others with high A+T content, resulting, respectively, from G·C- and A·T-directional mutation pressures. Such mutation pressure (G·C/A·T pressure) may be an important determinant for codon usage. It has also been suggested that tRNA acts as a selective constraint for determining codon usage. We have studied the relation between G·C/A·T pressure and tRNA constraints in determining choice of the third nucleotide of eight two-codon sets, using codon usage data obtained from protein genes in four bacterial species, *Mycoplasma capricolum*, *Bacillus subtilis*, *Escherichia coli*, and *Micrococcus luteus*, and in liverwort (*Marchantia polymorpha*) chloroplasts. The genomic G+C contents of these range from 25% to 74%. The results demonstrate that tRNA levels act additively to A·T and G·C pressure in affecting contents of A (pairing with *UNN anticodons, in which *U indicates a 2-thiouridine derivative) and C (pairing with GNN anticodons) or G (pairing with CNN anticodons), respectively, in third nucleotide positions of codons.**

Sueoka (1) proposed a theory to account for the diversification of genomic G+C content, ranging from 25% to 75%, in various species of eubacteria. The G+C content in each case will be determined by the effective base conversion rate of A·T to G·C pairs (G·C pressure) and that of G·C to A·T pairs (A·T pressure). This directional mutation pressure (G·C/A·T pressure) may be an important determinant for codon usage in eubacteria (2–4), as deduced from observations that G+C contents of all three positions of codons show a linear positive relationship with the genomic G+C content of various bacterial species. The strongest correlation is with third positions, largely, but not solely, resulting from G·C/A·T pressure directly exerted on the silent positions of codons. The correlation with the first position may be partly due to silent substitutions between UUR and CUR (both for leucine) and between CGR and AGR (both for arginine) and partly due to nondeleterious replacements between amino acids with differing G+C contents in their codons (2, 3).

Codon choice pattern is also affected by tRNA. Ikemura and Ozeki (5) demonstrated a positive correlation between codon usage and tRNA abundance, and they emphasized that tRNA populations may act as selective constraints for determining the codon usage in *Escherichia coli*, *Salmonella*, and yeast.

The relation between G·C/A·T pressure and tRNA constraints for codon choice has not been studied systematically, and therefore it is examined in this communication.[‖]

MATERIALS AND METHODS

Choice of the third nucleotide of two codon sets was studied by using data for codon usage in proteins produced at very high rates (6), including ribosomal proteins and elongation factors from *Mycoplasma capricolum* [genomic G+C, 25%; G+C of spacer regions (spacer regions are noncoding regions), 20%], *E. coli* (genomic G+C, 50%; spacer, 47%) and *Micrococcus luteus* (genomic G+C, 74%; spacer, 80%), because codon choice is said to be affected strongly by tRNA constraints in these proteins (6). Codon usage data (7) from the genes near the replication origin of *Bacillus subtilis* (genomic G+C, 43%; spacer, 38%), which are presumably for "highly produced" and/or "moderately produced" proteins (6) are included. In addition to the bacteria, data from the chloroplast genome of the liverwort *Marchantia polymorpha* are also included, since the genetic system of chloroplasts may be considered as essentially prokaryotic. The entire sequence of liverwort chloroplast DNA has been recently determined, and the total codon usage in genes and open reading frames as well as the complete set of tRNA genes has been deduced (8, 20). Liverwort chloroplast DNA has 28.8% G+C, which to our knowledge is the highest A+T content so far observed in plant chloroplasts, and the G+C content of total spacer region is 19.3%. The production rate of different proteins in chloroplasts has not been studied. We tentatively assume that in such an economized system, nearly all the protein genes are essential and expressed at about the same rate, corresponding to that of very highly expressed proteins.

The G+C contents of spacers from bacteria were taken from Fig. 1 of ref. 3, in which the values are calculated from the 5′- and 3′-flanking sequences of stable rRNA and protein genes. Spacer G+C content of liverwort chloroplasts is the sum of spacers (excluding intron regions), which make up about 20% of the total genome. Since we did not exclude signals such as promoters, terminators, etc. from spacers, G+C contents of spacers used here are only approximate and would deviate somewhat from the content of the "ideal" spacer (see below). The codon usage data and spacer G+C contents are shown in Table 1.

Codons in four-codon sets, those in the two-codon set plus four-codon set for a single amino acid (serine, leucine, and arginine), and codons for isoleucine were not included,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: N is U, C, A, or G; R is A or G; Y is C or U; *U is derivative of 2-thiouridine.

[‡]To whom reprint requests should be addressed.

[‖]Part of this study was presented at the Cold Spring Harbor Symposium held June 30, 1987, on the Evolution of Catalytic Function (4).

Table 1. Codon usage in two-codon sets in liverwort chloroplasts and four species of eubacteria

Amino acid	Codon	Anticodon	Codon usage (%)				
			Chloroplast (19% spacer G+C)	*Mycoplasma capricolum* (20% spacer G+C)	*Bacillus subtilis* (38% spacer G+C)	*Escherichia coli* (47% spacer G+C)	*Micrococcus luteus* (80% spacer G+C)
			NNY codons				
Tyr	UAU	GUA	826 (90)	69 (80)	115 (62)	22 (24)	0 (0)
	UAC		95 (10)	17 (20)	96 (38)	71 (76)	37 (100)
His	CAU	GUG	388 (85)	32 (71)	116 (69)	26 (26)	1 (4)
	CAC		62 (15)	14 (29)	53 (31)	75 (74)	30 (96)
Asn	AAU	GUU	1256 (88)	158 (79)	179 (53)	23 (12)	3 (7)
	AAC		175 (12)	41 (21)	158 (47)	168 (88)	43 (93)
Asp	GAU	GUC	735 (91)	118 (90)	266 (64)	94 (35)	5 (7)
	GAC		72 (9)	13 (10)	151 (36)	177 (65)	71 (93)
Phe	UUU	GAA	1547 (94)	112 (87)	165 (64)	36 (21)	0 (0)
	UUC		97 (6)	17 (13)	93 (36)	133 (79)	43 (100)
Total	NNU	GNN	4752 (90)	490 (83)	841 (60)	168 (26)	9 (4)
	NNC		501 (10)	102 (17)	551 (40)	471 (74)	224 (96)
			NNR codons				
Gln	CAA	*UUG	887 (99)	135 (99)	163 (44)	33 (20)	0 (0)
	CAG	CUG	53 (6)	2 (1)	138 (56)	133 (80)	47 (100)
Lys	AAA	*UUU	1764 (96)	415 (92)	395 (75)	328 (75)	0 (0)
	AAG	CUU (for *M. luteus*)	78 (4)	35 (8)	129 (25)	109 (25)	89 (100)
Glu	GAA	*UUC	1133 (93)	204 (94)	372 (70)	304 (78)	0 (0)
	GAG	CUC (for *M. luteus*)	85 (7)	12 (6)	163 (30)	85 (22)	113 (100)
Total	RAA	*UUY	2897 (95)	619 (93)	767 (73)	632 (77)	0 (0)
	RAG	CUY (for *M. luteus*)	163 (5)	47 (7)	292 (27)	194 (23)	202 (100)
	NAA†	*UNN	3784 (95)	754 (94)	— (—)	— (—)	0 (0)
	NAG†	CUN (for *M. luteus*)	216 (5)	49 (6)	— (—)	— (—)	451 (100)

Sources of data are as follows. *Marchantia polymorpha* (liverwort) chloroplasts, all 91 protein genes, including open reading frames (8, 20); *Mycoplasma capricolum*, 20 ribosomal protein genes (9); *Bacillus subtilis*, 20 protein genes near the replication origin (7); *Escherichia coli*, 30 ribosomal protein genes (10) and elongation factors G and Tu genes (11, 12); *Micrococcus luteus*, 2 ribosomal protein genes and elongation factors G and Tu genes (13). The spacer G+C contents given for each species are from ref. 20 for the chloroplasts and ref. 3 for the four eubacteria. *U indicates derivatives of 2-thiouridine (14).

†Since anticodon composition for RAR codons is different from that for CAR codons for *Bacillus subtilis* and *Escherichia coli*, the total values are not given.

because of the complexity of the codon–anticodon recognition pattern. UGY (cysteine) was also omitted because of the limited number of its codons in the sequences examined.

RESULTS AND DISCUSSION

Schematic Representation of Phenomena. Two-codon sets consist of a pair of synonymous codons, NNC and NNU (NNY) or NNG and NNA (NNR). Synonymous pairs of NNG and NNC, or NNA and NNU never exist apart from family sets, in which all four codons are synonymous. Symbols starting with NN refer to codons; those ending with NN refer to anticodons.

We shall postulate, as the first approximation, that the spacer region is free from constraints in the ideal case and its G+C content reflects solely G·C/A·T pressure (see above). If G·C/A·T pressure is the only factor for determining codon choice, the value of NNC/(NNC + NNU) × 100 in the NNY-type two-codon sets or NNG/(NNG + NNA) × 100 in the NNR-type two-codon sets (f) is equal to the G+C percent of the "ideal" spacer (s) and will move from f_1 to f_2 along with s, when s_1 changes to s_2 (Fig. 1*a*). On the other hand, if tRNA constraints are the only factors, the probabilities of the third position being occupied by G/C or A/U are equal. Then f may be expressed as 50 (%) + t (%), where t is the fraction of G+C percent of third nucleotides in codons differing from 50 because of tRNA constraints. The f value may change upon change of t (Fig. 1*b*). Let us consider the case where both G·C/A·T pressure and tRNA constraints affect the codon choice. Here, f is expressed as $s + t$. Upon change of G·C/A·T pressure, $s1$ to $s2$, the G+C content of the third nucleotide, f_1, will shift to f_2 ($= s_2 + t_1$) if tRNA constraints t_1 do not change (Fig. 1*c*). When t_1 changes to t_2, by, for example, addition of a new anticodon, f_1 will shift to f_3 even when s_1 does not change (Fig. 1*d*). The actual examples shown below clearly indicate that the choice of the third nucleotide of codons is affected by both G·C/A·T pressure and tRNA constraints in a combination of the schemes shown in Fig. 1*a*, *c*, and *d*.

NNY-Type Two-Codon Sets. Choice of C or U as the third nucleotide of the two-codon sets was studied for phenylalanine, tyrosine, histidine, asparagine, and aspartic acid. In these cases, only one species of tRNA, with anticodon sequence GNN, is used in translation, so that the amount of tRNA is not an important factor. In Fig. 2*a*, the NNC/(NNC + NNU) × 100 value ($= f$) of two-codon sets (NNY) was plotted against the G+C content of the spacer. The f values of chloroplasts, *Mycoplasma*, and *B. subtilis* are, on average, close to the spacer G+C content (s), indicating that no appreciable tRNA constraints exist in the range between s = 19 and s = 38 as in Fig. 1*a*. In *E. coli*, where G+C content of spacer (s) is 47%, NNC content (f) shifts to 74%, so that the fraction affected by tRNA constraints (t) is 27% ($t = f - s$). The f value reaches the maximum in *Micrococcus* (f = 96; t = 16). These calculations indicate that strong tRNA constraints are exerted to choose C rather than U in *E. coli* and *Micrococcus*.

The pairing of GNN anticodon to NNC codon is generally

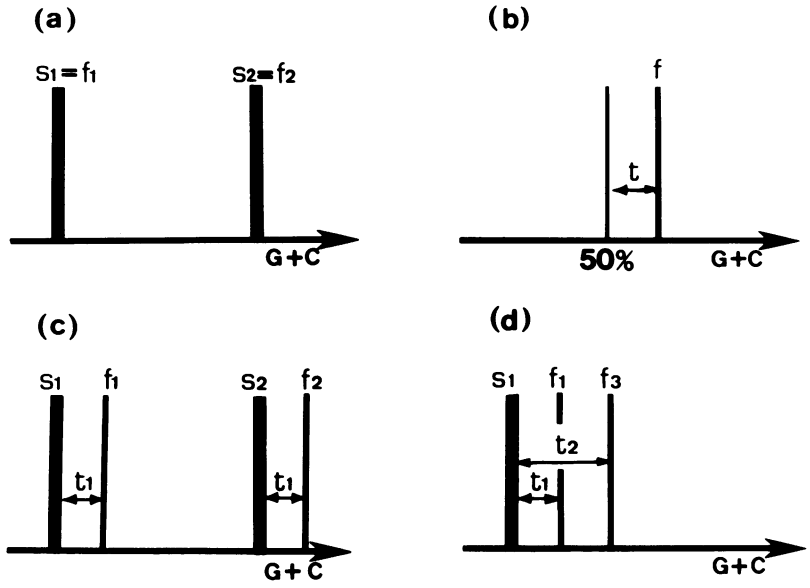

FIG. 1. Schematic representation of the effects of G·C/A·T pressure and tRNA constraints on choice of the third nucleotide of two-codon sets. (*a*) Shift of G+C content of the third nucleotide of codons (f) upon increase in G·C pressure from s_1 to s_2 in the absence of tRNA constraints. (*b*) Shift of f upon increase in tRNA constraints (t) in the absence of G·C pressure. (*c*) Shift of f from f_1 to f_2 upon increase in G·C pressure from s_1 to s_2 in the presence of constant tRNA constraints (t_1). (*d*) Shift of f from f_1 to f_3 upon increase in tRNA constraints from t_1 to t_2 without increase in G·C pressure.

stronger than to NNU codon, because of more stable pairing between G (anticodon) and C (codon) than between G and U (15, 16). However, in a range between about s = 20 (chloroplasts and *Mycoplasma*) and s = 38 (*B. subtilis*), NNU may be used nearly as much as NNC, for NNC would not contribute much to translation efficiency among predominating NNU/A codons in chloroplasts and in high A+T bacteria. This may be the reason for the lack of tRNA constraints. With increase in s, tRNA constraints appear, for NNU codons will become less "favorable" than NNC as numbers of NNC/G codons increase so that NNU codons act as a rate-limiting factor in translation.

The first anticodon nucleotide G is sometimes modified to Q (queuosine) or Gm (2′-*O*-methylguanosine), although its pairing specificity seems to remain the same as that of G (17). We tentatively consider that G, Q, and Gm are equivalent, but it is possible that their pairing preferences differ somewhat from each other. Another possibility is that in *Mycoplasma* and *B. subtilis* (and chloroplasts; see below) anticodons with G, Q, or Gm pairing equally with C and U are responsible for translating both NNC and NNU codons, so that no appreciable tRNA constraints are present, while *E. coli* and *Micrococcus* may have anticodons with another modification of G, pairing more efficiently with NNC than NNU codons, so that strong constraints against NNU by GNN anticodon exist. The G+C contents of spacers of chloroplasts and *Mycoplasma* have been estimated to be about the same in this study, and yet the preference for NNU codons over NNC is somewhat higher in chloroplasts than in *Mycoplasma*. This may be either because of inaccuracy of the estimated spacer G+C contents or because of differences in modification of the first nucleotide G in anticodons.

NNR-Type Two-Codon Sets. Next, choices between G and A of the third nucleotide of codons were studied for glutamine (CAR), lysine (AAR), and glutamic acid (GAR), where, depending on the bacterial species, one species (anticodon *UNN) or two species (anticodons *UNN and CNN) of tRNA are responsible for translating the codons.

tRNAs with *UNN anticodon have been found for glutamic acid, lysine, and glutamine in *E. coli*. It is logical to suppose that *UNN anticodons for these three amino acids are present throughout eubacteria and chloroplasts, for CNN anticodons can read only NNG codons and not NNA codons; NNA in two-codon sets must be translated by *UNN anticodons. Yokoyama and co-workers (14) showed that *U pairs much more readily with A than G, suggesting that NNA codons are read mainly by *UNN.

The number of CNN anticodons used in various bacterial species increases with increasing G+C content of their DNA, suggesting that the development of CNN anticodons has been accelerated by G·C pressure. The reported number of tRNAs having CNN anticodons (not including the CNN anticodons *CAU, CAU, and CCA, which are necessary for translation of isoleucine, methionine, and tryptophan codons) is seven in *E. coli* (G+C, 50%), two in *B. subtilis* (G+C, 43%), one in liverwort chloroplasts (G+C, 29%), and none in *Mycoplasma* group (G+C 25–30%) (4).

Genes for tRNA (CUU) for lysine and tRNA (CUC) for glutamic acid are not present in chloroplasts and have not been reported from *E. coli, B. subtilis*, and *Mycoplasma*. These anticodons may not exist in these three bacteria. On the other hand, tRNAs with anticodons CUC and CUU are major tRNA species in *Thermus thermophilus* (G+C, 69%) (18), and most probably exist in *Micrococcus luteus* (G+C, 74%). tRNA (CUG) for glutamine has been reported from *E. coli* and *Thermus*, and therefore probably exists also in *M. luteus*. *B. subtilis* may have this tRNA, although its presence has not yet been reported. Chloroplasts lack this tRNA, and there is no sign of its presence in *Mycoplasma*.

In Fig. 2*b*, the NNG/(NNG + NNA) × 100 value (f) of respective two-codon sets (NNR) was plotted against the G+C content of the spacer (s), as in the case of NNY codons. Only a very small usage of NNG codons is observed in chloroplasts (f = 5 against s = 19) and in *Mycoplasma* (f = 6 as against s = 20), showing considerable constraints against NNG caused by *UNN anticodons. The percent NNG (f) reaches 26% in *B. subtilis*, nearly in parallel with an increase in s (38%), with a similar t value of −11 as in *Mycoplasma* (t = −14) (Fig. 1*c*). The f value does not change in *E. coli* as compared with *B. subtilis*, so that the t value of *E. coli* decreases to −24 as against t = −11 of *B. subtilis*, indicating that stronger *UNN constraints are acting in this case to favor choice of NNA codons. In contrast,

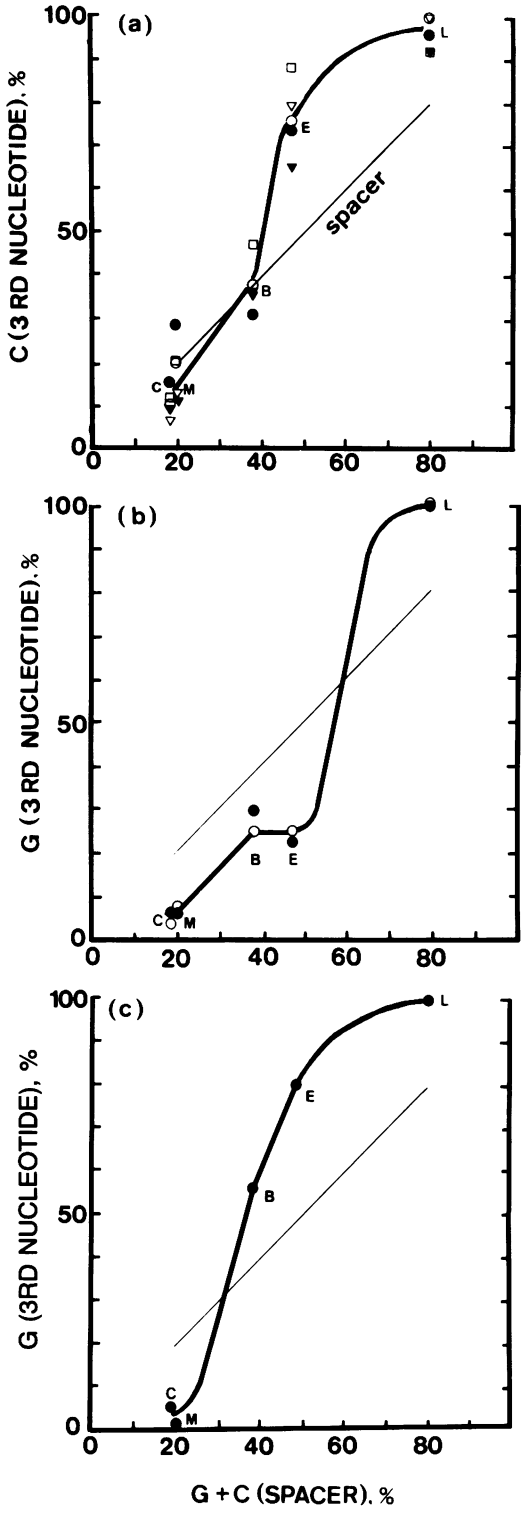

FIG. 2. Relation between G+C content of spacer and C (*a*) or G (*b* and *c*) content of the third nucleotide of two-codon sets. (*a*) Codons for tyrosine (○), histidine (●), asparagine (□), aspartic acid (▼), and phenylalanine (▽). (*b*) Codons for lysine (○) and glutamic acid (●). (*c*) Codon for glutamine (●). C, *Marchantia polymorpha* chloroplast; M, *Mycoplasma capricolum*; B, *Bacillus subtilis*; E, *Escherichia coli*; L, *Micrococcus luteus*. The straight lines indicate % spacer G+C levels from Table 1, plotted for "ideal" values (45° slope).

f value of *Micrococcus* reaches nearly 100 with *t* value of 20, indicating that strong constraints by CNN anticodons exist to choose NNG codons. In the case of glutamine, CUG anticodon is present in *E. coli* and is presumed to exist in *Micrococcus* and probably in *B. subtilis*, but it is absent from chloroplasts and *Mycoplasma* (see above). CAG usage is almost absent from chloroplasts and *Mycoplasma* and appears in *B. subtilis* ($f = 56$; $t = 18$). The high value of *t* is maintained in *E. coli* ($t = 25$) and in *Micrococcus* ($t = 20$).

The above observations suggest that, without CNN anticodons, NNG is a very "bad" codon because of only poor pairing ability of *UNN anticodons with NNG codons, and the constraints by *UNN are strong so as to keep NNG usage down to around 25%. The appearance of CNN anticodons thus seems to be adaptive, for an increase in *s* value tends to increase NNG codons that must be translated. *UNN constraints cannot accommodate any more of the increasing NNG codons by an increased *s* value. This would be the reason why CNN anticodons are generated, presumably by G·C pressure, in high G+C bacteria.

The above view is supported by the codon choice pattern for glutamine, lysine, and glutamic acid in *E. coli* (6). Usage of CAG for glutamine is highest in very highly produced proteins (89%) and decreases to 60% with the decreasing production rate of proteins. This pattern can be explained by the constraints of CNN anticodons, which are exerted at a maximum on very highly produced proteins and decrease in parallel with the decreasing production rate. In contrast, usage of NNG codons for lysine and glutamic acid is low (about 25%) and nearly unchanged regardless of the magnitude of the protein production, suggesting that, without CNN anticodons, *t* value (by *UNN constraints) has been saturated before *s* has reached 47% (= *s* of *E. coli*), not only in very highly produced proteins but also in others, including rarely produced proteins. This suggestion is further strengthened by the fact, as noted above, that usage of NNG codons in *B. subtilis* is almost the same (27%) as that in *E. coli* (23%).

Tryptophan codon usage in *Mycoplasma capricolum* supports the conclusion presented here. In *Mycoplasma capricolum*, UGA, a termination codon in the universal code, and UGG, the universal tryptophan codon, are both used for tryptophan (19). To translate these two codons, *UCA anticodon is used, while anticodon CCA, although still present, seems to be inactive and to be in the process of being discarded (ref. 19 and unpublished data). Usage of UGA codons much predominates over UGG; 24 tryptophan codons are UGA (92.3%) and 2 are UGG (7.7%). Although the number of tryptophan codons examined is not large, the value for UGA usage is similar to that for NNA usage in other two-codon sets (94% average), suggesting considerable constraints against codon UGG by *UCA anticodon, as in the case of other NNA/G two-codon sets.

In this communication, we have dealt only with two-codon sets for eight amino acids as occurring in abundantly expressed genes. Other studies will be needed to establish whether our findings are generally applicable to other codon sets and to genes that are less frequently expressed.

CONCLUSION

In a given bacterium, the G+C content of spacer (noncoding) DNA (*s*) can be regarded as being determined solely by G·C/A·T pressure. The G+C content of spacer DNA may differ from the G+C content of third nucleotides in coding regions. This difference (*t*) results from the effect of anticodons on choice of third nucleotides in codons. As shown in Fig. 2, tRNA constraints generally act additively to G·C/A·T pressure.

In five NNY two-codon sets (Fig. 2*a*), the percentage of C in third nucleotide codon positions is increased by constraints exerted by GNN anticodons in the presence of high G·C pressure. No difference from G+C content of spacer DNA occurs when A·T pressure predominates.

In NNR two-codon sets with A·T pressure, the percentage of A in the third nucleotide positions of codons is increased by constraints exerted by *UNN anticodons when these anticodons occur in the absence of CNN anticodons (Fig. 2*b*). In contrast, with G·C pressure (Fig. 2*c*), the percentage of G in third positions of codons is increased by constraints exerted by CNN anticodons.

This work was supported by grants from the Ministry of Education, Science and Culture of Japan, by the Naito Science Foundation, and by National Aeronautics and Space Administration Grant NGR 05 003 460 to the University of California, Berkeley.

1. Sueoka, N. (1962) *Proc. Natl. Acad. Sci. USA* **48,** 582–592.
2. Jukes, T. H. & Bhushan, V. (1986) *J. Mol. Evol.* **24,** 39–44.
3. Muto, A. & Osawa, S. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 166–169.
4. Osawa, S., Jukes, T. H., Muto, A., Yamao, F., Ohama, T. & Andachi, Y. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** in press.
5. Ikemura, T. & Ozeki, H. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47,** 1087–1097.
6. Sharp, P. M. & Li, W.-H. (1986) *J. Mol. Evol.* **24,** 28–38.
7. Ogasawara, N. (1985) *Gene* **40,** 145–150.
8. Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-I., Inokuchi, H. & Ozeki, H. (1986) *Nature (London)* **322,** 572–574.
9. Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F. & Osawa, S. (1987) *Mol. Gen. Genet.* **210,** 314–322.
10. Schnier, J., Kitakawa, M. & Isono, K. (1986) *Mol. Gen. Genet.* **204,** 126–132.
11. Yokota, T., Sugisaki, H., Takanami, M. & Kaziro, Y. (1980) *Gene* **12,** 25–31.
12. Zengel, J. M., Arch, R. H. & Lindahl, L. (1984) *Nucleic Acids Res.* **12,** 2181–2192.
13. Ohama, T., Yamao, F., Muto, A. & Osawa, S. (1987) *J. Bacteriol.* **169,** 4770–4777.
14. Yokoyama, S., Watanabe, T., Murao, K., Ishikura, H., Yamaizumi, Z., Nishimura, S. & Miyazawa, T. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 4905–4909.
15. Grosjean, H., Sankoff, D., Min Jou, W. & Cedergren, R. J. (1978) *J. Mol. Evol.* **12,** 113–119.
16. Ikemura, T. (1981) *J. Mol. Biol.* **151,** 389–409.
17. Nishimura, S. (1979) in *Transfer RNA: Structure, Properties and Recognition*, eds. Schimmel, P. R., Soll, D. & Abelson, J. M. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 59–79.
18. Hara-Yokoyama, M., Yokoyama, S., Watanabe, T., Watanabe, K., Kitazume, Y. & Miyazawa, T. (1986) *FEBS Lett.* **201,** 149–152.
19. Yamao, F., Muto, A., Kawauchi, Y., Iwami, M., Iwagami, S., Azumi, Y. & Osawa, S. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 2306–2309.
20. Ozeki, H., Ohyama, K., Inokuchi, H., Fukuzawa, H., Kohchi, T., Sano, T., Nakahigashi, K. & Umesono, K. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** in press.