factor[23]. Thus, compared to the above cases, the heat shock response of Malpighian tubules of *Drosophila* larvae is unique. The regulatory pathways responsible for non-induction of all the common HSPs in Malpighian tubules of *Drosophila* larvae may involve the heat shock transcription factor[24] and/or other auxiliary transcription factors[25] necessary for transcriptional activation of the heat shock genes under conditions of stress. Further studies will help understand the mechanism of this regulation as also the biological significance of this unique situation of non-inducibility of all the common HSPs and induction of a different set in the larger polytenized cells of this particular tissue.

1 Schlesinger, M J., Ashburner, M and Tissieres, A. (eds), *Heat Shock Proteins From Bacteria to Man*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.
2 Lindquist, S , *Annu Rev. Biochem* , 1986, **55**, 1151–1191.
3 Morimoto, R., Tissieres, A. and Georgopoulos, C. (eds), *Stress Proteins in Biology and Medicine*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1990.
4 Ritossa, F. M , *Experientia*, 1962, **18**, 571–573
5 Tissieres, A , Mitchell, H. K. and Tracy, U , *J Mol. Biol* , 1974, **84**, 389–398
6 Lakhotia, S C. and Singh, A K., *J Genet* , 1989, **68**, 129–137
7 DiDomenico, B , Bugaisky, G E. and Lindquist, S , *Cell*, 1982, **31**, 593–603.
8 Palter, B. K , Watanabe, M , Stinson, L., Mahowald, A. P. and Craig, E A , *Mol Cell Biol.*, 1986, **6**, 1187–1203.
9 Lindquist, S. and Craig, E. A , *Annu Rev. Genet.*, 1988, **163**, 631–677
10 Lakhotia, S. C and Mukherjee, T , *Chromosoma*, 1980, **81**, 125–136.
11 Laemmli, U. K , *Nature*, 1970, **227**, 680–685
12. Lakhotia, S. C and Singh, B N., *Indian J Exp Biol* , 1993, **31**, 301–302
13 Kurtz, S , Rossi, J , Petko, L and Lindquist, S , *Science*, 1986, **231**, 1154–1157
14 Patel, N H , *Methods Cell Biol* , 1994, **44**, 445–488
15 Singh, B N and Lakhotia, S C. (in preparation)
16 Mitchell, H K , Petersen, N S and Buzin, C H , *Proc Natl Acad Sci* , 1985, **82**, 4969–4973
17 Rubin, D M , Mehta, A D , Zhu, J , Shoham, S , Chen, X, Wells, Q R and Palter, K. B , *Gene*, 1993, **128**, 155–163
18 Feder, J H , Rossi, J. M , Solomon, J , Solomon, N and Lindquist, S , *Genes Develop* , 1992, **6**, 1402–1413
19 Berridge, M J. and Oschman, J L , *Tissue Cell*, 1969, **1**, 247–272
20. Velazquez, J M. and Lindquist, S , *Cell*, 1984, **36**, 655–662
21 Zimmermann, J L , Petri, W and Meselson, M , *Cell*, 1983, **32**, 1161–1170
22 Joplin, K H and Denlinger, D L , *J Insect Physiol* , 1990, **36**, 239–249
23 Morimoto, R I , Jurivich, D A , Kroeger, P. E , Mathur, S. K , Murphy, S P., Nakai, A , Sarge, K , Abravaya, K. and Sistonem, L. T., in *The Biology of Heat Shock Proteins and Molecular Chaperons* (eds Morimoto, R I , Tissieres, A and Georgopoulos, C), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp 417–455
24. Parker, C S and Topol, J , *Cell*, 1984, **37**, 273–283
25. Lu, Q , Wallrath, L. L , Allan, B D , Glaser, R L , Lis, J T. and Elgin, S C R , *J Mol Biol* , 1992, **222**, 985–998

ACKNOWLEDGEMENTS This work was supported by a research grant from the Council of Scientific & Industrial Research, New Delhi, to SCL and by a research fellowship from Banaras Hindu University to BNS We gratefully acknowledge the kind gift of *7 10* and *7Fb* monoclonal antibodies by Prof S Lindquist

Received 21 April 1995, accepted 23 June 1995

Codon analysis of cyanobacterial genes

S. Krishnaswamy and S. Shanmugasundaram*

Bioinformatics Centre, Department of Biotechnology and *DBT Centre for BGA Biofertilizer, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021, India

The genes from different species of cyanobacteria have been analysed for relationships both between them and with other microbes and organelles. The codon bias indices indicate that the unicellular *Synechococcus* strain shows a preferred codon usage as those of low expressing genes of *E. coli*, *Agrobacterium tumefaciens* and *S. typhimurium*, whereas the filamentous *Anabaena* sp. and *Fremyella diplosiphon* strains show similarity to yeast (*Saccharomyces cerevisiae*), *Bacillus* sp., *B. subtilis* and *Streptococcus* sp. In so far as the codon bias index can be related to expression levels, the data suggest the possibility of suitable cloning hosts. The filamentous and unicellular cyanobacteria can be clearly separated based on the codon third-position bias and also the correlations involving codon usage. In general, correlations in codon usage profiles point to the filamentous strains being more related to gram positive bacteria and eukaryotes, whereas the unicellular strains are more related to gram negative bacteria.

PATTERNS of codon usage vary between organisms[1] and have helped to understand regulation of gene expression[2, 3]. Several organisms have been examined in detail in terms of their biased codon usage[4, 5]. However, only a preliminary analysis has been reported for codon usage in cyanobacterial genes[6] using a limited set of genes. The relationships within cyanobacterial species and with other organisms have been examined using the compilation by Wada *et al.*[7], who have tabulated the codon usage for different organisms. It must be mentioned that a study of codon usage of different genes from an organism, in spite of potential problems of the pattern of usage[8, 9], can lead to acceptable conclusions in certain cases and provide general trends[6, 7].

The codon usage data were taken from the compilation by Wada *et al.*[7]. These data refer to only complete genes and not to introns. The program CODON was used. This was developed by one of us earlier to examine the relationship of codon usage between plants and

Table 1. The percentage of synonymous codons for the unicellular (synechococcus, SYN) and filamentous (Anabaena, ANA), (Fremyella diplosiphon, FDI) strains are shown The preferred codons selected with a cut off limit of sigma are marked with an asterisk

AA	cod	SYN	ANA	FDI
Gly	GGG	13	8	6
Gly	GGA	6	15	15
Gly	GGT	36*	53*	56*
Gly	GGC	45*	24	23
Glu	GAG	38*	21	17
Glu	GAA	62*	79*	83*
Asp	GAT	48*	58*	65*
Asp	GAC	52*	42*	35*
Val	GTG	34*	16	14
Val	GTA	7	32*	30*
Val	GTT	27*	36*	46*
Val	GTC	32*	16	10
Ala	GCG	22*	12	8
Ala	GCA	19*	27*	32
Ala	GCT	28*	42*	51*
Ala	GCC	31*	19	10
Arg	AGG	1	4	2
Arg	AGA	1	14	10
Arg	CGG	20	17	14
Arg	CGA	8	7	7
Arg	CGT	26	32*	34*
Arg	CGC	44*	26*	33*
Ser	AGT	13*	16*	14
Ser	AGC	24*	19*	28*
Ser	TCG	22*	5	3
Ser	TCA	6	15*	8
Ser	TCT	18*	28*	28*
Ser	TCC	18*	17*	19*
Lys	AAG	46*	31	40*
Lys	AAA	54*	69*	60*
Asn	AAT	28	44*	38*
Asn	AAC	72*	56*	62*
Ile	ATA	0	10	5
Ile	ATT	43	52*	51*
Ile	ATC	57*	38*	45*
Thr	ACG	21	5	2
Thr	ACA	6	32*	25
Thr	ACT	21	26*	31*
Thr	ACC	51*	37*	42*
Cys	TGT	33*	52*	38*
Cys	TGC	67*	48*	62*
Tyr	TAT	29	42*	35*
Tyr	TAC	71*	58*	65*
Leu	TTG	27*	26*	27*
Leu	TTA	7	25*	29*
Leu	CTG	36*	16*	13
Leu	CTA	6	15*	16
Leu	CTT	0	6	8
Leu	CTC	24	12	8
Phe	TTT	42*	57*	54*
Phe	TTC	58*	43*	46*
Gln	CAG	45*	27	26
Gln	CAA	55*	73*	74*
His	CAT	30	39*	58*
His	CAC	70*	61*	42*
Pro	CCG	34*	6	7
Pro	CCA	12	28*	21
Pro	CCT	18	37*	45*
Pro	CCC	35*	29*	28*

plant viruses[10]. The codon bias index, which determines the level of usage of preferred codons in a gene, has been defined earlier[4]. In this study, the same definition of the codon bias index is used. However, the selection of preferred codons is not based on the percentage occurrence of codons within a set of synonymous codons. The significance of a percentage occurrence depends on the number of synonymous codons. For example, 50% occurrence for a codon is not significant when there are only two synonymous codons but can be considered preferred if there are three or more synonymous codons in the set. Hence, we have calculated the standard deviation in percentage occurrence within the set of synonymous codons for each amino acid. Those codons in a synonymous set whose percentage occurrence is above a given cut-off limit times the standard deviation are flagged as 'preferred'. These preferred codons of the plant, for example, are used to determine if the codons of the plant virus are biased towards the plant codon usage. The best choice of the cut-off limit is one where the minimum number of codons is chosen as preferred while no amino acid is left without at least one preferred codon (Table 1).

In order to compare two codon usage profiles, for each amino acid a correlation coefficient is calculated using the percentage occurrence of codons within the synonymous set. The match coefficient for the entire set of amino acids is then defined as the sum of the weighted correlation coefficients. The weight for each amino acid is calculated as the ratio of the total number of codons for that amino acid to the total number of codons for all the amino acids of the two tables. This is done to take into account the differences in amino acid composition.

Thus,

$$\text{the match coefficient} = \sum_{I}^{N} w_I c_I,$$

where

N = all amino acids except Met, Trp,

w_I = weight factor,

defined as

$$w_I = \frac{\text{total number of codons for the } i\text{th amino acid}}{\text{total number of codons}},$$

and

c_I = correlation coefficient of synonymous codon usage,

$$c_I = \frac{\sum_{J}^{n_I} (p_{JI} - \bar{p}_I)(q_{JI} - \bar{q}_I)}{\left[\sum_{J}^{n_I} (p_{JI} - \bar{p}_I)^2\right]^{1/2} \left[\sum_{J}^{n_I} (q_{JI} - \bar{q}_I)^2\right]^{1/2}},$$

where P_{ij} is percentage occurrence of the jth codon of the ith amino acid of one table, q_{ij} is the percentage occurrence of the jth codon of the ith amino acid of the other table, $\bar{p}_i, \bar{q}_i$ are the means and n_i is the number of codons of the ith amino acid.

The match coefficient as defined here can vary from –1 (for anticorrelated profiles) to zero (for profiles without correlation) to +1 (for identical profiles). A high match coefficient indicates that amino acids with a higher frequency of occurrence have a similar distribution of codons within their set of synonymous codons.

The codon usage and the set of preferred codons[10] within the cyanobacterial species – the unicellular *Synechococcus* sp. (SYN), the filamentous *Anabaena* sp. (ANA) and *Femyella diplosiphon* (FDI) – are shown in Table 1. It is seen that the unicellular strain has a different usage from the filamentous strain, unlike in the earlier reported analysis[6]. This is also seen from the third-position codon bias (Table 2), where the unicellular strain has less A, T and more C, G as compared to filamentous strains. The match coefficients (Table 2) also show that there is a low correlation of codon occurrence patterns between unicellular and filamentous strains (0.31 in each case), while between the filamentous strains the coefficient is high (0.93). Thus, the match coefficient is able to provide discrimination between the two types of cyanobacterial strains. A comparison of the cyanobacterial codon usage profiles with those of other organisms using the match coefficients (Table 3) shows that the filamentous and unicellular strains of cyanobacteria also have different relationships with other organisms. This is seen from the unicellular strain having a high match coefficient with the gram negative bacteria *S. marcescens*, *Pseudomonas* sp., *A. vinelandii* and the gram positive *Streptomyces* sp. The filamentous strains have a high match coefficient with the gram positive bacteria *B. megaterium*, *Streptococcus* sp. and the eukaryotic yeast (*Schizosaccharomyces pombe* and *S. cerevisiae*), and with chloroplasts of chlamydomonas, pea, maize, spinach and tobacco. A high match coefficient shows a related pattern of codon occurrence. The above-mentioned difference in the codon usage profile implies that the unicellular and filamentous forms could have different evolutionary branches.

The codon bias index refers to how the preferred codons of other organisms are used in cyanobacterial genes, while the match coefficient is an indicator of the correlation in use of all the codons of cyanobacterial genes and the other organism. Thus, the bias and the match coefficient need not relate to each other as has also been shown elsewhere[10].

The usage of preferred codons in organisms has been related to levels of gene expression though it is known that there are other factors that also affect gene expression levels. A comparison of the codon bias indices of

Table 2. Differentiating unicellular from filamentous strain by codon analysis

	SYN	ANA	FDI
	0	–	–
Frequency of base at position 3			
A	0 15	0 29	0 26
T	0 23	0.31	0 35
C	0 36	0 24	0 25
G	0 27	0 17	0 14
Ratio of			
$\frac{XCG}{XCC}$	0 73	0 33	0 26
$\frac{XTA}{XTT}$	0 29	0 78	0 72
Match coefficient relating the codon profile of the different forms of cyanobacteria			
SYN 0	1 0	0 31	0 31
ANA –	0 31	1 0	0 93
FDI –	0 31	0 93	1 0

Table 3. The match coefficient relating the codon profile of the unicellular and filamentous forms with respect to those of the different species

Match coefficient	SYN	ANA	FDT
	0	–	–
Gram negative bacteria			
E coli	0 40	0 24	0 35
Pseudomonas sp.	0 60	–0 25	–0 24
A tumefaciens	0 36	–0 29	–0 19
A vinelandii	0 57	–0 25	–0 23
K pneumoniae	0 55	0 11	–0 09
S. marcescens	0 75	0 04	0 05
S typhimurium	0 43	0 17	0 26
R capsulatus	0 50	–0 40	–0 31
B japonicum	0 53	–0 36	–0 33
Rhizobium	0 54	–0 36	–0 35
Clostridium sp	–0 38	0 47	0 37
Gram positive bacteria			
B subtilis	0 04	0 42	0 47
Bacillus sp	–0 08	0 46	0 47
B megasterium	–0 08	0 66	0 67
B sphaericus	–0 39	0 43	0 42
Streptococcus sp	–0 08	0 62	0 62
S aureus	–0.26	0 52	0 49
B stearothermophilus	0.37	0 16	0 20
Streptomyces sp	0.58	–0 34	–0 35
Eukaryotes			
S. pombe	–0 06	0 61	0 61
S cerevisiae	–0 04	0 62	0 59
Kluyveromyces	–0 29	0 48	0 43
Wheat chloroplast	–0 22	0 60	0 57
Euglena gracilis	–0.28	0 54	0 51
Marchantia polymorpha	–0 25	0 55	0 52
Zea mays	–0 25	0 57	0 55
Pisum sativum	–0 23	0 61	0 60
Spinach	–0 25	0 58	0 57
Nicotiana tabaccum	–0 18	0 62	0 61
Chlamydomonas	0 10	0 73	0 63

Table 4. The codon bias indices (CBI) for the unicellular and filamentous forms with respect to the preferred codons selected for the different species. The cut-off limit used for selection in standard deviation units is given in parantheses

CBI	SYN 0	ANA –	FDI –
Gram negative bacteria			
E. coli (low) (2)*	0.31	0 10	0 14
E coli (high) (2)**	0 17	0 17	0 21
E. coli (2)	0.34	0.20	0 22
Pseudomonas sp (2)	0 22	–0 09	–0 09
A tumefaciens (2)	0 28	–0 06	–0 04
A vinelandii (2)	0 21	–0 06	–0 04
K pneumoniae (2)	0 21	–0 07	–0 09
S marcescens (2)	0 21	–0 05	–0 05
S. typhimurium (2)	0 26	–0 01	–0 07
R capsulatus (2)	0 20	–0 10	–0 12
B japonicum (2)	0 22	–0 13	–0 13
Rhizobium (2)	0 22	–0 14	–0 15
Clostridium sp (2)	–0 18	0 20	0 22
Gram positive bacteria			
B. subtilis (2)	0 09	0 28	0 29
Bacillus sp (2 5)	0 13	0 30	0 29
B megaterium (2 5)	–0 03	0 20	0 17
B sphaericus (2 0)	–0.19	0 22	0 23
Streptococcus sp (2 0)	–0 05	0 27	0.33
S aureus (2 0)	–0 12	0 24	0 26
B stearothermophilus (2 5)	0 16	0.00	–0 04
Streptomyces sp (2 0)	0 17	–0 08	–0 07
Eukaryotes			
S. pombe (2)	–0 03	0 20	0 26
S cerevisiae (2)	0 01	0 28	0.32
Kluyveromyces (2)	–0 12	0 21	0 23
Wheat chloroplast (2)	–0 17	0 25	0.30
M phymorpha (2)	–0 04	0 19	0 22
Z. mays (2)	–0 17	0 26	0 29
P sativum (2)	–0 14	0 24	0.26
Spinach (2)	–0 22	0 24	0 27
N tabaccum (2)	–0 09	0 26	0.29
Chlamydomonas (2)	–0 02	0 27	0 28

*Genes with low expression and **genes with high expression

the cyanobacterial genes with respect to other organisms (Table 4) shows that the codon usage of unicellular genes is biased towards *E. coli* (low expression) and *A. tumefaciens*, whereas the filamentous genes are biased towards *S. cerevisiae*, *Bacillus* sp., *B. subtilis* and chloroplasts. In as far as the biased codon usage is a determinant of the levels of gene expression, it is seen through this analysis that *A. tumefaciens* could form an alternate cloning host apart from *E. coli* for unicellular strains and filamentous strains may be better expressed in *S. cerevisiae*, *Bacillus* sp. or *B. subtilis*. These conclusions are being made to form a basis for molecular biology experiments involving cyanobacteria.

1 Andersson, G E and Kurland, C. G, *Microbiol Rev*, 1990, **54**, 198–210

2 Ikemura, T, *J Mol Biol*, 1981, **146**, 1–21

3 de Boer, H A and Kastelein, R A, in *Biased Codon Usage An Exploration of its Role in Optimization of Translation* (eds Reznikoff, W and Gold, L), Buttersworth, Boston, 1986, pp. 225–286

4 Bennetzen, J L and Hall, B D, *J Biol Chem*, 1982, **257**, 3026–3031

5. Sharp, P M and Devine, K. M, *Nucleic Acids Res*, 1989, **17**, 5029–5039.

6 Tandeau de Marsac, N and Houmard, J., in *The Cyanobacteria* (eds Fay, P. and Van Baalen, C), Elsevier, 1987, pp 251–302

7 Wada, K, Wada, Y, Doi, H, Ishibashi, F, Gojobori, T and Ikemura, T., *Nucleic Acids Res*, 1992, **20** (Suppl), 2111–2118

8. Lloyd, A J. and Sharp, P M, *Mol. Gen Genet*, 1991, **230**. 288–294

9. Sharp, P M, Cowe, E., Higgins, D G, Shields, D C, Wolfe, K. H and Wright, F, *Nucleic Acid Res*, 1988, **16**, 8207–8210.

10 Sudha, S N, Krishnaswamy, S and Vaithilingam Sekar, *Curr Sci*, 1992, **63**, 573–575

ACKNOWLEDGEMENTS. We are thankful to DBT, New Delhi, for financial support to Bioinformatics Centre, MKU The work was done under the Technology Development and Demonstration Project on blue–green algae funded by DBT, New Delhi

Received 24 July 1993; revised accepted 23 April 1995

Effects of feeding mulberry (*Morus* sp.) leaves supplemented with different nutrients to silkworm (*Bombyx mori* L.)

A. A. Sarker, M. R. Haque, M. A. Rab and N. Absar*

[1]Bangladesh Sericulture Research and Training Institute, Rajshahi 6100, Bangladesh

*Department of Biochemistry, Rajshahi University, Rajshahi 6205, Bangladesh

The growth of silkworm larvae improved significantly on feeding them mulberry leaves supplemented with different nutrients. The total protein content of silk gland increased, while the water, total lipid and carbohydrate contents decreased. The larval weight, silk gland weight and total protein in silk gland were found to be highest for treatment TS_8 (soymilk + sugar + vitamins + potassium iodide salt), followed by TS_7 (glycine + alanine + sugar + vitamins + potassium iodide salt) and TS_9 (milk powder + sugar + vitamins + potassium iodide salt) treatments. The highest values of single-cocoon weight, single-shell weight and filament length were observed for treatment TS_8 (1.34 g, 0.26 g and 799.42 m/cocoon, respectively), whereas cocoon yield was found to be highest for treatment TS_7 (110.17 g/100 larvae). No significant difference was observed between the treatments TS_7, TS_8 and TS_9, their influence on silk gland and cocoon characteristics being more or less similar. The cocoon yield/100 larvae was increased by 20.61%, 20.38% and 19.06% in treatments TS_7, TS_8 and TS_9, respectively, over the control treatment.