J Microbiol Immunol Infect.
2007;40:24-33

Original Article

Studies on synonymous codon and amino acid usages in *Aeromonas hydrophila* phage Aeh1: architecture of protein-coding genes and therapeutic implications

Keya Sau

Department of Biotechnology, Haldia Institute of Technology, Haldia, India

Received: December 1, 2005 Revised: February 28, 2006 Accepted: March 3, 2006

Background and Purpose: Codon and amino acid usage biases determined in numerous organisms have deciphered the architectures of their protein-coding genes to some extent. To understand the architecture of protein-coding genes of *Aeromonas* phages, codon and amino acid usage biases have been investigated in the protein-coding genes of the *Aeromonas hydrophila* phage Aeh1.
Methods: In order to study synonymous codon and amino acid usage biases in Aeh1, all of its protein-coding genes were downloaded and analyzed by standard software programs.
Results: Phage Aeh1 harbors an AT-rich genome. The third position of its synonymous codons carries mostly A or T base and mutational pressure strongly influences the synonymous codon usage bias. Translational selection also influences the codon usage of Aeh1 as its putatively lowly- and highly-expressed genes are influenced by Aeh1-specific tRNAs and by the abundant cellular tRNAs, respectively. Further analysis of amino acid usage shows that amino acid residues are also not randomly utilized in Aeh1 proteins and factors such as hydropathy, aromaticity and cysteine content are mostly responsible for the variation of amino acid usage in Aeh1 proteins.
Conclusions: As Aeh1 does not carry any toxin/antibiotic resistant gene but carries moderately highly expressed genes and relatively few *Ahd*I sites, this study proposes that Aeh1 may be utilized as a therapeutic agent for *A. hydrophila* infections. While codon usage bias in Aeh1 is dictated both by mutational pressure and translational selection, amino acid usage bias in Aeh1 is influenced by hydropathy, aromaticity and cysteine content. Phage Aeh1 may be utilized in phage therapy.

Key words: *Aeromonas hydrophila*; Amino acid metabolism; Bacteriophages; Codon; Gene expression regulation, bacterial

Introduction

Studies on synonymous codon and amino acid usages in living organisms revealed that they vary genome to genome, gene to gene, and even among different parts of a gene. While factors such as mutational pressure [1-3], translational selection [4-9], secondary structure of proteins [10-14] influence the codon usage in various organisms, amino acid usage was shown to be governed by hydrophobicity, aromaticity, cysteine content and mean molecular weight [15-20]. Apart from elucidating the architecture and evolution of genes/genomes, codon usage in particular has a number of practical applications.

Synonymous codon and amino acid usage biases had been studied only in a few genomes of bacteriophages, although they have enriched different fields of biological science [21]. It was reported that codon usage patterns of some coliphages such as T4, T7, lamda (λ), N15, P2, and P4 do not exactly resemble that of their host *Escherichia coli* [22-26]. Synonymous codon usage of highly and lowly expressed genes in T4 were shown to be influenced by the abundant host tRNAs and own tRNAs, respectively, [25]. Codon usage in phage

Corresponding author: Dr Keya Sau, Department of Biotechnology, Haldia Institute of Technology, PO-Hatiberia, Dist-Purba-Medinipur, Haldia 721657, West Bengal, India.
E-mail: keya_sau@yahoo.co.in

T7 was also suggested to be influenced mainly by the abundant host tRNAs [22]. It was described that codon usage in both T4 and T7 is influenced mainly by mutational pressure [26]. In λ, N15, P2, and P4 phages, synonymous codon usage patterns were found nearly similar to that of the lowly expressed genes of *E. coli* [23-26].

In contrast, codon usages of mycobacteriophages and *Staphylococcus aureus* phages were found to be almost identical to their respective bacterial hosts [27-30]. It was shown that mutational pressure and translational selection mainly influence the codon usages in phages of mycobacteria and *S. aureus* [28-30]. While tRNAs of mycobacteriophage Bxz1 were suggested to regulate the expression of both its highly and lowly expressed genes [28], tRNAs present in phages D29 and L5 were suggested to affect their amino acid usage to some extent [27].

Based on codon usage analysis, it was proposed that three *S. aureus* phages — 44AHJD, Phi 68 and K — carrying mostly highly expressed genes, could be utilized in phage therapy [30]. Recently, it was shown that phage K indeed has potential to kill diverse clinical strains of *S. aureus* [31]. In addition, it was shown that in AT-rich *Pseudomonas aeruginosa* phage PhiKZ, codon usage bias is dictated mainly by mutational bias and, to an extent, by translational selection. Analysis also revealed that amino acid usage in PhiKZ proteins is mainly dictated by mean molecular weight, aromaticity and cysteine content [32].

Several phages, which infect different species of *Aeromonas* have been described in the literature [33-37]. Thus far, codon and amino acid usage biases in *Aeromonas* phages have not been studied at length, although such study may reveal the architecture of protein-coding genes and also provide clues as to whether a phage could be used as a therapeutic agent. In this communication, *Aeromonas hydrophila* phage Aeh1 [37] was selected as a representative phage in order to study the synonymous codon and amino acid usage biases in *Aeromonas* phages.

Methods

The genome sequence of bacteriophage Aeh1 was downloaded from GenBank (USA) and its 343 protein coding sequences (carrying 50 or more codons) have been extracted from the genome by the "coderet" (available from: http://bioweb.pasteur.fr/seqanal/interfaces/coderet.html) program. The relative synonymous codon usage (RSCU) in all protein coding sequences was determined, in order to study the overall codon usage variation among the genes. RSCU is defined as the ratio of the observed frequency of codons to the expected frequency if all of the synonymous codons for those amino acids are used equally [38]. RSCU values greater than 1.0 indicate that the corresponding codon is more frequently used than expected, whereas the converse is true for RSCU values less than 1.0. RSCU values in 24 putatively highly expressed genes (e.g., genes encoding outer membrane proteins and those involved in glycolysis, detoxification, etc.) of *A. hydrophila* were also determined for comparison.

GC_{3s} is the frequency of guanine (G) plus cytosine (C) and A_{3s}, T_{3s}, G_{3s}, and C_{3s} are the frequencies of adenine (A), thymine (T), G and C at the synonymous third positions of codons. N_c, the effective number of codons used by a gene, is generally used to measure the bias of synonymous codons and is independent of amino acid composition and gene length [39]. The values of N_c range from 20 (when one codon is used per amino acid) to 61 (when all the codons are used with equal probability). Highly biased genes are generally highly expressed [6] and as there is no information of gene expression level of bacteriophages, the highly biased genes as highly expressed in the Aeh1 phages were considered in this study. Hence the sequences in which the N_c values are less than 30 were considered to be highly expressed and those having N_c values greater than 55 were considered to be lowly expressed genes.

The program CodonW 1.3 (available from: www.molbiol.ox.ac.uk/cu) was used for calculating most of the parameters, including correspondence analysis on the RSCU and amino acid usage. In correspondence analysis, the data are plotted in a multidimensional space of 59 axes (excluding methionine, trpytophan and stop codons) and then it determines the most prominent axes contributing the codon usage variation among the genes. In the present study, RSCU values were used for correspondence analysis in order to minimize the amino acid composition.

Results and Discussion

Overall codon usage analysis

The percent GC content in Aeh1 is 42.57. This indicates that the third position of the synonymous codons of Aeh1 genes should be AT rich. The RSCU values determined in all of the 343 protein coding genes of Aeh1 in fact

Table 1. Overall codon usage analysis in Aeh1

Amino acid	Codon	Aeh1			*Aeromonas hydrophila*	tRNA copy
		Overall RSCU	HEG RSCU	LEG RSCU	HEG RSCU	Aeh1
Phenylalanine	UUU	0.94	0.65	1.07	0.41	
	UUC	1.06	1.35	0.93	1.59	1
Leucine	UUA	0.52	0	0.85	0.03	1
	UUG	1.36	0.16	1.07	0.37	1
	CUU	1.04	0.23	1.28	0.19	
	CUC	0.54	0.31	0.67	1	
	CUA	0.58	0	0.85	0.04	1
	CUG	1.95	5.3	1.28	4.36	
Isoleucine	AUU	1.39	0.67	1.42	0.33	
	AUC	1.38	2.33	1.09	2.35	1
	AUA	0.23	0	0.49	0.31	
Methionine	AUG	1	1	1	1	2
Valine	GUU	2	2.3	1.65	0.72	
	GUC	0.47	0.22	0.63	0.91	
	GUA	0.88	1.07	1.05	0.4	
	GUG	0.66	0.41	0.67	1.97	1
Serine	UCU	1.45	2.91	1.14	0.54	
	UCC	0.77	0.94	0.67	2.22	
	UCA	0.86	0.38	1	0.11	1
	UCG	0.57	0	0.62	0.62	
	AGU	1.26	0.75	1.38	0.23	
	AGC	1.08	1.03	1.19	2.28	1
Proline	CCU	1.01	0.69	1.04	0.3	
	CCC	0.3	0.14	0.49	1.19	
	CCA	1.25	0.83	1.38	0.2	1
	CCG	1.44	2.34	1.09	2.31	
Threonine	ACU	1.59	1.8	1.17	0.56	
	ACC	0.95	2.08	0.85	2.93	
	ACA	1.02	0	1.26	0.13	1
	ACG	0.44	0.11	0.73	0.38	
Alanine	GCU	1.45	2.53	1.1	0.81	
	GCC	0.58	0.69	0.65	2.13	
	GCA	1.22	0.56	1.39	0.3	1
	GCG	0.75	0.21	0.86	0.77	
Tyrosine	UAU	1.33	0.78	1.3	0.59	
	UAC	0.67	1.22	0.7	1.41	
Histidine	CAU	1.08	0.8	1.12	0.49	
	CAC	0.92	1.2	0.88	1.51	1
Glutamine	CAA	1.3	1.45	1.22	0.37	1
	CAG	0.7	0.55	0.78	1.63	
Asparagine	AAU	0.96	0.48	1.11	0.33	
	AAC	1.04	1.52	0.89	1.67	1
Lysine	AAA	1.2	1.14	1.12	0.55	1
	AAG	0.8	0.86	0.88	1.45	
Aspartic acid	GAU	1.25	1.17	1.09	0.83	
	GAC	0.75	0.83	0.91	1.17	1
Glutamic acid	GAA	1.42	1.77	1.36	0.76	2
	GAG	0.58	0.23	0.64	1.24	
Cysteine	UGU	1.16	1.5	1.07	0.29	
	UGC	0.84	0.5	0.93	1.71	1

(Table continued on page 27)

(Table continued from page 26)

Tryptophan	UGG	1	1	1	1	1
Arginine	CGU	2.21	4.21	1.63	1.71	
	CGC	1.34	1.43	0.75	2.64	
	CGA	0.95	0.18	1.54	0.13	
	CGG	0.3	0.09	0.62	1.28	
	AGA	1.05	0	1.06	0.1	
	AGG	0.14	0.09	0.4	0.14	
Glycine	GGU	1.77	2.71	1.41	1.45	
	GGC	1.06	1.13	1.21	2.01	
	GGA	1	0.12	1.15	0.1	1
	GGG	0.18	0.04	0.23	0.45	

Abbreviations: RSCU = relative synonymous codon usage ; HEG = Aeh1 putatively highly expressed genes (effective number of codons [N_c] <34); LEG = Aeh1 putatively lowly expressed genes (N_c >55)

show that A- and/or T-ending codons are predominant in this phage (Table 1). As analysis of overall RSCU alone is not sufficient to reveal the heterogeneity of codon usage in Aeh1 genes, N_c and GC_{3s} were also determined.

It was observed that in Aeh1, N_c values range from 27.71 to 61.0 with a mean of 48.65 and standard deviation (SD) of 7.07, whereas, GC_{3s} ranges from 0.256 to 0.586 with a mean of 0.403 and SD of 0.058. Taken together, the results suggest that apart from the mutational bias, other factors might have some influence in the codon usage variation among Aeh1 genes.

Mutational bias versus translational selection in codon usage variation

In order to determine the determinants of codon usage variations in different organisms, N_c plots (a plot of N_c vs GC_{3s}) and correspondence analysis were used widely. It was suggested that the comparison of actual distribution of genes with the expected distribution under no selection might be used if some influences other than mutational bias dictate the codon usage bias in the sequences under study [39]. It was also suggested that if GC_{3s} was the only determinant of N_c, then N_c would fall on the continuous curve between N_c and GC_{3s}.

The N_c plot for the 343 genes of phage Aeh1 shows that a small number of points lie on the expected curve (Fig. 1), which certainly originates from extreme mutational bias. In contrast, most points with low N_c values lie well below the expected curve (Fig. 1), which suggests that these Aeh1 phage genes have an additional codon usage bias which is independent of GC_{3s}.

Correspondence analysis on the RSCU values of the 343 protein-coding genes of Aeh1 phage was done in order to determine the factors influencing the codon usage bias in Aeh1. Fig. 2 shows the distribution of Aeh1 genes on the first two major axes of the correspondence analysis. It is evident from the plot that there is wide variation in codon usage among the genes of Aeh1. However, the first major axis accounted for 12.98% of the total variation and the second major axis accounted for 5.20% of the total variation. The position of the genes along the first major axis is negatively correlated with the A_{3s} ($r = –0.667$, $p<0.01$) and positively correlated with T_{3s} ($r = 0.209$, $p<0.01$) and C_{3s} ($r = 0.621$, $p<0.01$). Interestingly, first major axis does not correlate with G_{3s}. Also, the position of

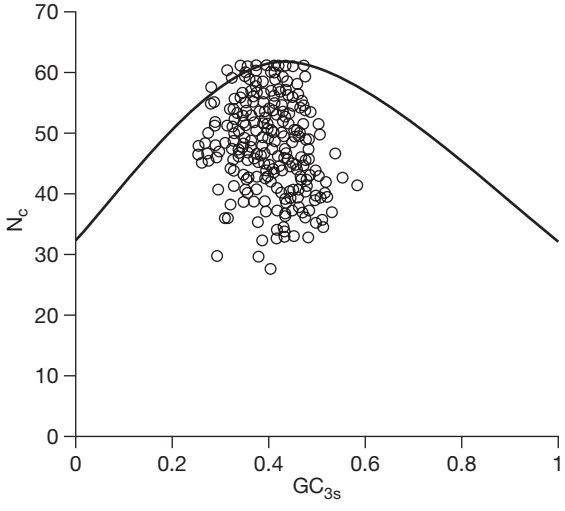

Fig. 1. Plot of effective number of codons (N_c) and guanine (G) plus cytosine (C) frequency at the synonymom third position of codons (GC_{3s}) of phage Aeh1 genes (genes are represented by circles).

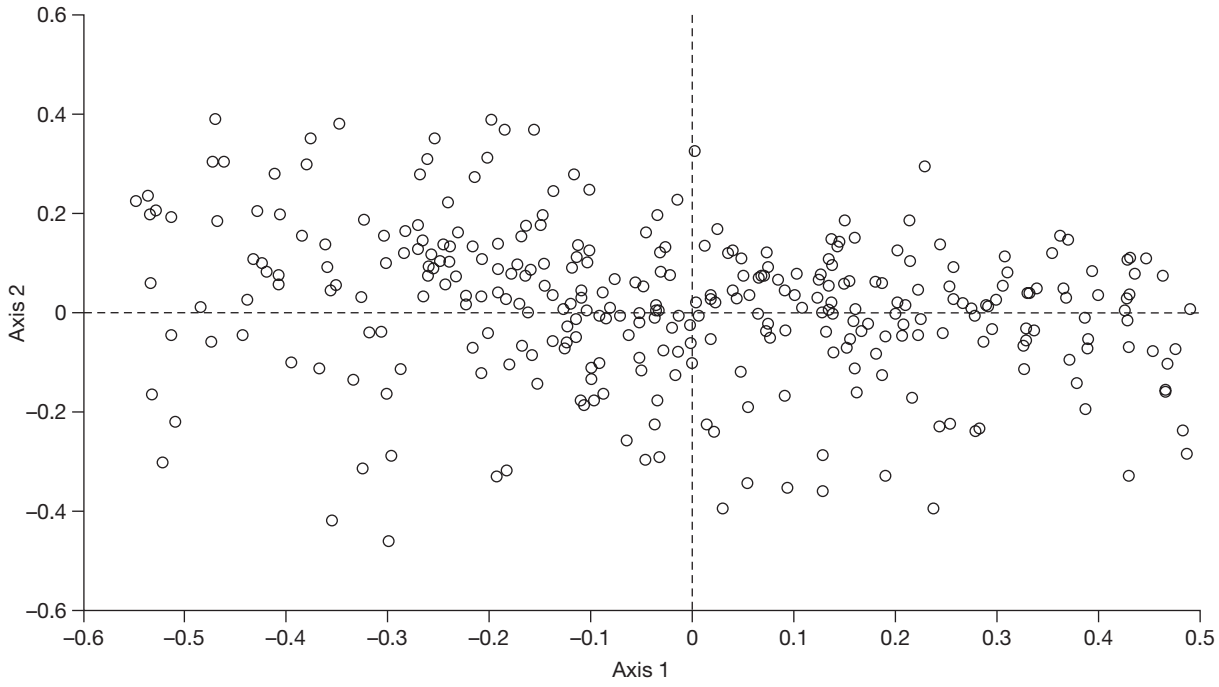

Fig. 2. Positions of Aeh1 genes along the two major axes of variation in the correspondence analysis of relative synonymous codon usage values.

the genes along the second major axis is significantly correlated with C_{3s} (r = –0.172, p<0.01) and G_{3s} (r = 0.125, p<0.05).

As the above results did not clearly demonstrate how A-, T-, G- and C-ending codons are distributed along the first major axis, codon usage in 10% of the genes located at the extreme right of axis 1 has been compared with that of the 10% of the genes located at the extreme left of the axis 1. Chi-squared tests were used to estimate the codon usage variation between these two sets of genes, taking p<0.01 as significance criterion.

Table 2 shows RSCU values for each codon for the two groups of genes. It is important to note that out of 18 codons that are statistically over-represented in genes located on the extreme left side of axis 1, there are 10 C-ending (55.55%), 5 T-ending (27.75%), 2 G-ending (11.1%) and 1 A-ending (5.55%) codon(s). The data in fact support the results of the N_c plot and clearly show that, in addition to mutational pressure, other factors have some influence in the codon usage variation among Aeh1 genes. Interestingly, synonymous codon usages bias in another *Aeromonas* phage 44RR was also found to be influenced by mutational pressure [data not shown]. Mutational pressure was found to play a key role in shaping the codon usage variations in many bacterial and non-bacterial viruses [26,28-30,32,40].

Most bacteriophages do not carry any tRNA encoding gene and depend completely on host tRNAs for their gene expression. However, some phages, such as T4, BxZ1 and PhiKZ, carry some tRNA encoding genes and analyses showed that the phage-specific tRNAs influence the expression of either their lowly- or both the lowly- and the highly-expressed genes of phages [25,28,32]. Phage Aeh1 also harbors 23 tRNA encoding genes within its genome. These tRNAs, if expressed, may incorporate all the amino acids except arginine and tyrosine. As shown in Table 1, in-depth analysis showed that there are 36 over-represented codons in lowly expressed genes of Aeh1 in comparison with its highly expressed genes. Of 36 over-represented codons, 23 are recognized by the anticodons of Aeh1-specific tRNAs. This indicates that nearly 64% of over-represented codons of lowly expressed genes are recognized by Aeh1-specific tRNAs. In contrast, only 48% of over-represented codons of highly expressed genes of Aeh1 are recognized by its own tRNAs. Thus, the data suggest that Aeh1-specific tRNAs preferentially recognize the codons of lowly expressed genes of Aeh1. Function exactly similar to that of phage Aeh1-specific tRNAs has been reported before [25].

The cellular tRNA abundance is positively correlated with over-represented codons of the highly

Table 2. Relative synonymous codon usage (RSCU) values for each codon for the two groups of genes in phage Aeh1. Each group contains 10% of sequences at either extreme of the major axis generated by correspondence analysis

Amino acid	Codon	RSCU[b]	N[b]	RSCU[c]	N[c]	Amino acid	Codon	RSCU[b]	N[b]	RSCU[c]	N[c]
Phenylalanine	UUU	0.63	74	1.07	144	Serine	UCU[a]	2.49	118	1.17	92
	UUC[a]	1.37	161	0.93	124		UCC[a]	1.12	53	0.62	49
Leucine	UUA	0.00	0	1.48	112		UCA	0.30	14	1.37	108
	UUG	0.40	24	1.56	118		UCG	0.08	4	0.75	59
	CUU	0.34	20	1.04	79	Proline	CCU	0.94	31	1.20	77
	CUC	0.35	21	0.22	17		CCC	0.24	8	0.41	26
	CUA	0.02	1	0.99	75		CCA	0.36	12	1.39	89
	CUG[a]	4.89	290	0.70	53		CCG[a]	2.45	81	1.00	64
Isoleucine	AUU	0.92	96	1.45	211	Threonine	ACU[a]	2.02	150	1.21	143
	AUC[a]	2.07	217	0.92	134		ACC[a]	1.80	134	0.47	55
	AUA	0.01	1	0.64	93		ACA	0.12	9	1.66	195
Methionine	AUG	1.00	182	1.00	168		ACG	0.05	4	0.66	78
Valine	GUU	2.14	245	1.96	222	Alanine	GCU[a]	2.03	253	1.13	130
	GUC	0.35	40	0.42	48		GCC[a]	0.78	97	0.45	52
	GUA	0.88	101	1.05	119		GCA	0.84	105	1.49	172
	GUG	0.62	71	0.57	65		GCG	0.35	43	0.93	107
Tyrosine	UAU	0.98	78	1.47	217	Cysteine	UGU	1.00	25	1.35	50
	UAC[a]	1.02	82	0.53	79		UGC	1.00	25	0.65	24
Tryptophan	UGG	1.00	51	1.00	118						
Histidine	CAU	0.76	35	1.28	62	Arginine	CGU[a]	3.57	151	0.80	36
	CAC[a]	1.24	57	0.72	35		CGC[a]	2.06	87	0.44	20
Glutamine	CAA[a]	1.41	173	1.15	127		CGA	0.28	12	1.33	60
	CAG	0.59	73	0.85	93		CGG	0.05	2	0.24	11
Asparagine	AAU	0.60	68	1.19	230	Serine	AGU	0.85	40	1.26	99
	AAC[a]	1.40	158	0.81	157		AGC	1.16	55	0.84	66
Lysine	AAA	1.14	259	1.26	257	Arginine	AGA	0.05	2	2.75	124
	AAG	0.86	196	0.74	150		AGG	0.00	0	0.44	20
Aspartic acid	GAU	1.14	171	1.29	219	Glycine	GGU[a]	2.42	212	1.32	157
	GAC	0.86	129	0.71	120		GGC[a]	1.30	114	0.70	83
Glutamic acid	GAA	1.51	321	1.36	221		GGA	0.18	16	1.71	203
	GAG	0.49	104	0.64	103		GGG	0.09	8	0.27	32

Abbreviation: N = number of codons

[a]Codons whose occurrences are significantly ($p<0.01$) higher in the extreme left side of axis 1 than the genes present on the extreme right of the first major axis.

[b]Genes on the extreme left of axis 1.

[c]Genes on the extreme right of axis 1.

expressed genes in several organisms [4,22,23,41-44]. In coliphage T4, synonymous codon usage of the highly expressed genes was also shown to be influenced by the abundant cellular tRNAs [25]. This may also be true for the highly expressed genes of Aeh1. As the status of tRNA copy number is not known in *A. hydrophila* at present, the RSCU values of the putatively highly expressed genes of *A. hydrophila* were determined and the resulting data compared with that of highly expressed genes of Aeh1 (Table 1). Of 18 over-represented codons of highly expressed genes of Aeh1, 8 codons are also over-represented in the highly expressed genes of *A. hydrophila*. In addition, codons CGU, CGC, GGU, GGC, GCU, and GCC are also abundant in the highly expressed genes of both Aeh1 and its host *A. hydrophila*. The data indeed suggest that abundant tRNAs of *A. hydrophila* influence the synonymous codon usage of the highly expressed genes of Aeh1. It is interesting to note that the translational selection also influences the codon usage bias of *Aeromonas salmonicidia* phage 44RR to some extent.

Amino acid usage

Correspondence analysis on the relative amino acid usage of the 343 proteins showed that the first and second major axes of correspondence analysis accounted for 20.46% and 9.89% of the total variation of the amino acid composition of Aeh1 proteins, respectively. Further

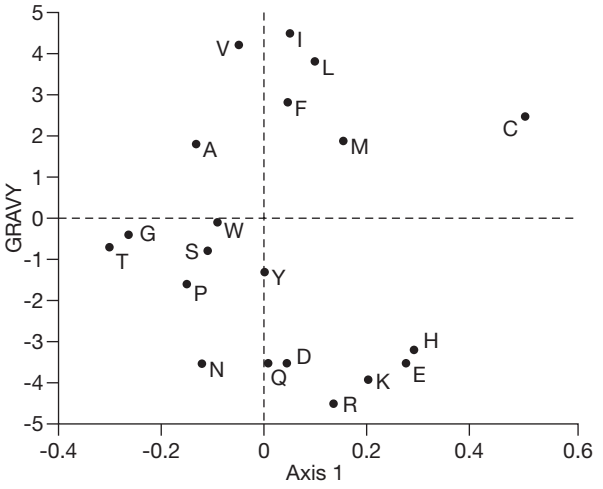

Fig. 3. Correlation between grand average hydropathy (GRAVY) scores of amino acid residue vs the axis 1 values of correspondence analysis. Single letter codes (A = alanine; C = cysteine; D = aspartic acid; E = glutamic acid; F = phenylalanine; G = glycine; H = histidine; I = isoleucine; K = lysine; L = leucine; M = methionine; N = asparagine; P = proline; Q = glutamine; R = arginine; S = serine; T = threonine; V = valine; W = tryptophan; Y = tyrosine) are used to show the positions of 20 amino acid residues (amino acid residues are represented by black dots).

analysis revealed that first major axis is negatively correlated with grand average hydropathy (GRAVY) [r = –0.313, p<0.01] of each Aeh1 protein. GRAVY or hydropathy denoting the measure of hydrophobic character present in an amino acid is used during analysis of the structure of a polypeptide or protein sequence. However, a plot of the GRAVY score of each amino

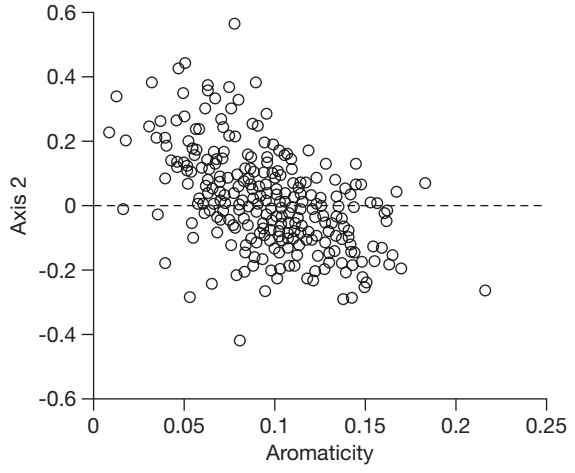

Fig. 4. Correlation between the aromaticity of Aeh1 proteins and axis 2 values of correspondence analysis. Aeh1 proteins are represented by circles.

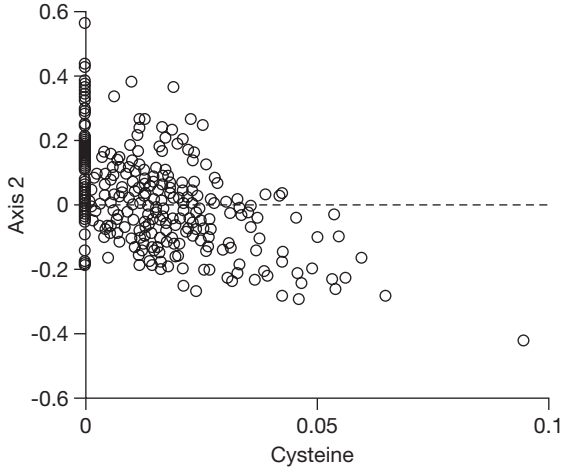

Fig. 5. Correlation between cysteine content of Aeh1 proteins and axis 2 values of correspondence analysis. Aeh1 proteins are represented by circles.

acid residue vs the first major axis shows that charged residues such as lysine, arginine, histidine, aspartic acid and glutamic acid are located on the positive side of the first axis (Fig. 3). In contrast, the second major axis is significantly correlated (r = –0.517, p<0.01) with aromaticity of each Aeh1 protein (Fig. 4). Amino acid frequency analysis shows that all of the aromatic amino acids are rare in Aeh1 proteins (data not shown). Aromatic amino acids were also rare in *E. coli, Thermogota maritama, Giardia lamblia* and phage PhiKZ proteins, and it was suggested that these amino acids are not incorporated preferentially in proteins as their biosynthesis is energetically expensive for organisms [16-18,32].

Further analysis has shown that the second major axis is also significantly correlated (r = –0.500, p<0.01) with the cystine content of Aeh1 proteins (Fig. 5). Interestingly, among the 343 Aeh1 proteins, 63 proteins do not carry any cystine residue, whereas 9 proteins which are located at the extreme right side in Fig. 5 are found to harbor more than 5% cystine residues. It would be interesting to explore the contribution of these 9 cystine-rich proteins towards temporal gene regulation as well as the development of Aeh1 in its host. It was noticed that amino acid usage bias of *Aeromonas* phage 44RR is also influenced by the above three factors [data not shown].

Therapeutic potential of Aeh1

A. hydrophila causes several diseases in fish, amphibians, reptiles and humans and present antibiotic therapy could not eliminate it completely from earth

[45]. Several reports showed that lytic phages could be utilized successfully to cure diverse bacterial infections [46-48]. No phage capable of restricting the diseases caused by *A. hydrophila* has been reported so far. As described above, lowly expressed genes of Aeh1 are expressed preferentially by its own tRNAs, whereas its highly expressed genes are expressed preferentially by the abundant host tRNAs. Furthermore, from correspondence and related analysis on RSCU values of Aeh1 genes, it is found that N_c is significantly correlated ($r= –0.616$, $p<0.01$) with the position of Aeh1 genes distributed along the first major axis. Together, these data suggest that genes of Aeh1 phage may be expressed rapidly by the translational machinery of *A. hydrophila*. If this is true, then phage Aeh1 must be substantially virulent in nature. Analysis also revealed that Aeh1 does not carry any toxin-encoding/antibiotic resistance genes and harbors only 3 recognition sites for the restriction endonuclease, *Ahd*I. Moreover, it resembles the T-even phages both morphologically and genomically [37]. Several groups showed that phage T4 and T4-like phages such as JS4, JSD.1, JSL.6, JS94.1, e11/2 and pp01 could be utilized to kill pathogenic *E. coli* strains which cause several human diseases such as diarrhea, hemorrhagic colitis and hemolytic-uremic syndrome. [49-51]. In addition to T4-like phages, several other phages of the *Myoviridae* family, such as K, ENB6, PPpW-1 and PPpW-3, were also shown to efficiently kill the pathogenic strains of *S. aureus* [31], *Enterococcus faecium* [52], and *Pseudomonas plecoglossicida* [53], respectively. Incidentally, it was shown by codon usage analysis that phage T4 mostly harbors highly expressed genes [25,30]. Based on these findings, this study suggests that Aeh1 might be utilized in curing *A. hydrophila* infections if it were to pass clinical and pharmaceutical development steps successfully.

Acknowledgments

I thank Dr. T.C. Ghosh and Dr. S. Sau, both from Bose Institute, Calcutta, for critically reading and modifying this manuscript. I also thank Dr. N. Bhattacharyya, Head, Department of Biotechnology, Haldia Institute of Technology for his help and encouragement during the entire work.

References

1. Levin DB, Whiteome B. Codon usage in nucleopolyhedroviruses. J Gen Virol. 2000;81:2313-25.
2. Jenkins GM, Pagel M, Gould EA, de A Zanotto PM, Holmes EC. Evolution of base composition and codon usage bias in the genus *Flavivirus.* J Mol Evol. 2001;52:383-90.
3. Jenkins GM, Holmes EC. The extent of codon usage bias in human RNA viruses and its evolutionary origin. Virus Res. 2003;92:1-7.
4. Grantham R, Gautier C, Gouy M, Jacobzone M, Mercier R. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. 1981;9:r43-74.
5. Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol. 1985;2:13-34.
6. Sharp PM, Cowe E. Synonymous codon usage in *Saccharomyces cerevisiae*. Yeast. 1991;7:657-78.
7. Lesnik T, Solomovici J, Deana A, Ehrlich R, Reiss C. Ribosome traffic in *E. coli* and regulation of gene expression. J Theor Biol. 2000, 202:175-85.
8. Ghosh TC, Gupta SK, Majumdar S. Studies on codon usage in *Entamoeba histolytica.* Int J Parasitol. 2000;30:715-22.
9. Gupta SK, Ghosh TC. Gene expressivity is the main factor in dictating the codon usage variation among the genes in *Pseudomonas aeruginosa*. Gene. 2001;273:63-70.
10. Oresic M, Shalloway D. Specific correlations between relative synonymous codon usage and protein secondary structure. J Mol Biol. 1998;281:31-48.
11. Xie T, Ding D. The relationship between synonymous codon usage and protein structure. FEBS Lett. 1998;434:93-6.
12. Chiusano ML, Alvarez-valin F, Di Giulio M, D'Onofrio G, Ammirato G, Colonna G, et al. Second codon positions of genes and the secondary structures of proteins. Relationships and implications for the origin of the genetic code. Gene. 2000; 261:63-9.
13. Gupta SK, Majumdar S, Bhattacharya TK, Ghosh TC. Studies on the relationships between the synonymous codon usage and protein secondary structural units. Biochem Biophys Res Commun. 2000;269:692-6.
14. D'Onofrio G, Ghosh TC, Bernardi G. The base composition of the genes is correlated with the secondary structures of the encoded proteins. Gene. 2002;300:179-87.
15. Basak S, Banerjee T, Gupta SK, Ghosh TC. Investigation on the causes of codon and amino acid usages variation between thermophilic *Aquifex aeolicus* and mesophilic *Bacillus subtilis*. J Biomol Struct Dyn. 2004;22:205-14.
16. Lobry JR, Gautier C. Hydrophobicity, expressivity and aromaticity are the major trends of amino-acid usage in 999 *Escherichia coli* chromosome-encoded genes. Nucleic Acids Res. 1994;22:3174-80.
17. Garat B, Musto H. Trends of amino acid usage in the proteins from the unicellular parasite *Giardia lamblia*. Biochem Biophys Res Commun. 2000;279:996-1000.
18. Zavala A, Naya H, Romero H, Musto H. Trends in codon and

amino acid usage in *Thermotoga maritima*. J Mol Evol. 2002; 54:563-8.
19. Banerjee T, Basak S, Gupta SK, Ghosh TC. Evolutionary forces in shaping the codon and amino acid usages in *Blochmannia floridanus*. J Biomol Struct Dyn. 2004;22:13-23.
20. Naya H, Zavala A, Romero H, Rodriguez-Maseda H, Musto H. Correspondence analysis of amino acid usage within the family *Bacillaceae*. Biochem Biophys Res Commun. 2004; 325:1252-7.
21. McGrath S, Fitzgerald GF, van Sinderen D. The impact of bacteriophage genomics. Curr Opin Biotechnol. 2004;15: 94-9.
22. Sharp PM, Rogers MS, McConnell DJ. Selection pressures on codon usage in the complete genome of bacteriophage T7. J Mol Evol. 1984-85;21:150-60.
23. Gouy M. Codon contexts in enterobacterial and coliphage genes. Mol Biol Evol. 1987;4:426-44.
24. Holm L. Codon usage and gene expression. Nucleic Acids Res. 1986;14:3075-87.
25. Kunisawa T. Synonymous codon preferences in bacteriophage T4: a distinctive use of transfer RNAs from T4 and from its host *Escherichia coli*. J Theor Biol. 1992;159:287-98.
26. Kunisawa T, Kanaya S, Kutter E. Comparison of synonymous codon distribution patterns of bacteriophage and host genomes. DNA Res. 1998;5:319-26.
27. Kunisawa T. Functional role of mycobacteriophage transfer RNAs. J Theor Biol. 2000;205:167-70.
28. Sahu K, Gupta SK, Ghosh TC, Sau S. Synonymous codon usage analysis of the mycobacteriophage Bxz1 and its plating bacteria *M. smegmatis*: Identification of highly and lowly expressed genes of Bxz1 and the possible function of its tRNA species. J Biochem Mol Biol. 2004;37:487-92.
29. Sahu K, Gupta SK, Sau S, Ghosh TC. Comparative analysis of the base composition and codon usages in fourteen mycobacteriophage genomes. J Biomol Struct Dyn. 2005;23: 63-71.
30. Sau K, Gupta SK, Sau S, Ghosh TC. Synonymous codon usage bias in 16 *Staphylococcus aureus* phages: implication in phage therapy. Virus Res. 2005;113:123-31.
31. O'Flaherty S, Ross RP, Meaney W, Fitzgerald GF, Elbreki MF, Coffey A. Potential of the polyvalent anti-*Staphylococcus* bacteriophage K for control of antibiotic-resistant staphylococci from hospitals. Appl Environ Microbiol. 2005;71: 1836-42.
32. Sau K, Sau S, Mandal SC, Ghosh TC. Factors influencing the synonymous codon and amino acid usage bias in AT-rich *Pseudomonas aeruginosa* phage PhiKZ. Acta Biochim Biophys Sin (Shanghai). 2005;37:625-33.
33. Merino S, Camprubi S, Tomas JM. Isolation and characterization of bacteriophage PM3 from *Aeromonas hydrophila* the bacterial receptor for which is the monopolar flagellum. FEMS Microbiol Lett. 1990;57:277-82.
34. Merino S, Camprubi S, Tomas JM. Isolation and characterization of bacteriophage PM2 from *Aeromonas hydrophila*. FEMS Microbiol Lett. 1990;56:239-44.
35. Merino S, Camprubi S, Tomas JM. Identification of the cell surface receptor for bacteriophage 18 from *Aeromonas hydrophila*. Res Microbiol. 1990;141:173-80.
36. Chow MS, Rouf MA. Isolation and partial characterization of two *Aeromonas hydrophila* bacteriophages. Appl Environ Microbiol. 1983;45:1670-6.
37. Tetart F, Desplats C, Kutateladze M, Monod C, Ackermann HW, Krisch HM. Phylogeny of the major head and tail genes of the wide-ranging T4-type bacteriophages. J Bacteriol. 2001; 183:358-66.
38. Sharp PM, Li WH. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 1987;15:1281-95.
39. Wright F. The 'effective number of codons' used in a gene. Gene. 1990;87:23-9.
40. Gu W, Zhou T, Ma J, Sun X, Lu Z. Analysis of synonymous codon usage in SARS *Coronavirus* and other viruses in the *Nidovirales*. Virus Res. 2004;101:155-61.
41. Ikemura T. Correlation between codon usage and tRNA content in microorganisms. In: Hatfield DL, Lee BJ, Pirtle RM, eds. Transfer RNA in protein synthesis. Boca Raton, Ann Arbor, London, Tokyo: CRC Press; 1992.
42. Zhou J, Liu WJ, Peng SW, Sun XY, Frazer I. Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. J Virol. 1999;73: 4972-82.
43. Kanaya S, Yamada Y, Kinouchi M, Kudo Y, Ikemura T. Codon usage and tRNA genes in eukaryotes: correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis. J Mol Evol. 2001; 53:290-8.
44. Kanaya S, Yamada Y, Kudo Y, Ikemura T. Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus subtilis* tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. Gene. 1999;238:143-55.
45. Kao HT, Huang YC, Lin TY. Fatal bacteremic pneumonia caused by *Aeromonas hydrophila* in a previously healthy child. J Microbiol Immunol Infect. 2003;36:209-11.
46. Sulakvelidze A, Alavidze Z, Morris JG Jr. Bacteriophage therapy. Antimicrob Agents Chemother. 2001;45:649-59.
47. Nakai T, Park SC. Bacteriophage therapy of infectious diseases in aquaculture. Res Microbiol. 2002;153:13-8.
48. Matsuzaki S, Rashel M, Uchiyama J, Sakurai S, Ujihara T, Kuroda M, et al. Bacteriophage therapy: a revitalized therapy

against bacterial infectious diseases. J Infect Chemother. 2005; 11:211-9.
49. Bruttin A, Brüssow, H. Human volunteers receiving *Escherichia coli* phage T4 orally: a safety test of phage therapy. Antimicrob Agents Chemother. 2005:49:2874-8.
50. Chibani-Chennoufi S, Sidoti J, Bruttin A, Kutter E, Sarker S, Brüssow H. In vitro and in vivo bacteriolytic activities of *Escherichia coli* phages: implications for phage therapy. Antimicrob Agents Chemother. 2004;48:2558-69.
51. O'Flynn G, Ross RP, Fitzgerald GF, Coffey A. Evaluation of a cocktail of three bacteriophages for biocontrol of *Escherichia coli* O157:H7. Appl Environ Microbiol. 2004;70:3417-24.
52. Biswas B, Adhya S, Washart P, Paul B, Trostel AN, Powell B, et al. Bacteriophage therapy rescues mice bacteremic from a clinical isolate of vancomycin-resistant *Enterococcus faecium.* Infect Immun. 2002;70:204-10.
53. Park SC, Shimamura I, Fukunaga M, Mori KI, Nakai T. Isolation of bacteriophages specific to a fish pathogen, *Pseudomonas plecoglossicida*, as a candidate for disease control. Appl Environ Microbiol. 2000;66:1416-22.

