Relative Thermodynamic Stacking Constants of Deoxyribo Dinucleotides in the Prebiotic Origin of tRNAs from Bacteriophage T4

NIGEL AYLWARD
Queensland University of Technology
George St., Brisbane, Queensland 4000
AUSTRALIA
n.aylward@alumni.qut.edu.au

Abstract: - Consistent relative stacking interaction energies have been calculated for the sixteen triply ionised deoxyribose dinucleotide structures in the g-g- conformation. These values are lower than for the corresponding ribose dinucleotides and allow the prediction of the composition and complete characterization of a poly deoxyribose nucleotide. When applied to the tRNAs of Bacteriophage T4 they show the anticodons to have the predicted composition and stacking energy. However, the codons have a higher stacking energy enabling them to be preferentially transcribed. The stacking energy for the corresponding ribose anticodon and codon are higher than for the deoxyribose anticodon and codon. The anticodon and codon may be silenced by the synthetic ribose codon and anticodon. The DNA of dinucleotides may be favourably transcribed with Watson-Crick base pairing with a possible 100% accuracy. The prebiotic codons could not be fully translated without a chance of error. The stability of the DNA anticodons has enhanced their preservation.

Key-Words: - Thermodynamic data, base stacking, tRNAs, bacteriophage T4, anticodon, codon.

1 Introduction

The T4 bacteriophage that infects the bacteria Escherichia coli [1] is 18 nm long with a head and tail structure [2] and a molecular weight of 220 megdaltons. It contains 61% DNA consisting of two circular DNA molecules which serve as a template for the synthesis of complimentary viral mRNA molecules in preference to the host-cell mRNA, ultimately leading to the formation of viral coat proteins. T4 contains over a 100 genes and the genetic map is known [2]. The entire process in the assembly of the virus appears to be spontaneous except for one reaction [2].

From a prebiotic perspective it is desirable to show that not only is the virus self assembled, but also its genes may be self assembled. It is known that the nucleic acid bases, uracil, thymine, cytosine, adenine and guanine stack in aqueous solution as free bases [3], nucleosides and nucleotides [4]. The stacking also occurs in single strand polymers [5], and in double and triple helices [6].

This study is to accurately determine the relative stacking energies [7] of the sequences found in the codons and anticodons of the tRNA genes of bacteriophage T4 and compare them to the energy of sequences of the same composition assembled according to their relative stacking free energies calculated by an entirely ab initio method [8].

If correct, this hypothesis predicts that the stability of the stacks formed and subsequently polymerized were assembled according to their thermodynamic stability, and these stable sequences were trapped from ancient time in the genes of living organisms. One suggestion for nucleotides that could have polymerized were the amino acyl derivatives of cyclic-3',5'-nucleotides [9]. The total energy of the polymer may then be calculated according to the one-dimensional Ising model [10].

2 Problem Formulation

The stacking interactions for DNA-DNA, DNA-RNA and RNA-RNA and the Watson-Crick interactions in DNA or RNA are as previously reported [7,11]

2.1 The One-Dimensional Ising Model of a Stacked Polynucleotide

To use the thermodynamic data to characterize a polynucleotide it is necessary to consider the one-dimensional Ising model of a polynucleotide. A nucleotide in solution (T,C,A or G) may interact or stack with another nucleotide (T,C,A or G) to produce a dimer. There are sixteen possible dimers, TT,TC,TA,TG, CT,CC,CA,CG, AT,AC,AA,AG, GT,GC,GA,GG. Each of these dimers has a particular stacking interaction energy. Each may add

a further nucleotide from solution to form a trimer. e.g. TTC. The last interaction also has a specific interaction energy which depends on the attacking base of the dimer being a T and the final base being a C. In this model the single nucleotides are always free in solution, whereas the dimers, trimers, oligomers or polymers are always stacked in a fairly rigid conformation. There are not any conformational degrees of freedom as in the rotational isomeric theory [12][13]. However, the energy of the polynucleotide contains interlocking terms, which depend on the particular bases present in at least two neigbouring monomer units,

$$U(\Phi 1, \Phi 2, \ldots \Phi n) = \Sigma\, U(\Phi k-1, \Phi k) \qquad (1)$$

where, Φk characterizes the interaction of the kth nucleotide unit [12]. Consequently, the polynucleotide represents a statistical system which cannot be divided into elements with mutually independent states, and thus is a cooperative system [12-14]. The interaction partition function (the sum of the statistical weights of the interaction states) of an individual chain, has the form,

$$Z = \sum_{\Phi 1} \sum_{\Phi 2} \ldots\ldots \sum_{\Phi n} e^{-U(\Phi 1, \Phi 2, \ldots \Phi n)/kT} \qquad (2)$$

In this, summation is carried out over all 4 interactions of each nucleotide unit. If the free energy of the polynucleotide has the form [12], then,

$$Z = \sum_{\Phi 1} \sum_{\Phi 2} \ldots\ldots \sum_{\Phi n} \prod_{k=1}^{n} g(\Phi k-1, \Phi k) \qquad [3]$$

where, $g(\Phi k-1, \Phi k) = e^{-U(\Phi k-1, \Phi k)/kT}$

In the polynucleotide all the monomer units have identical sets of interactions, $\Phi 1$, $\Phi 2$, $\Phi 3$, $\Phi 4$ for any values of k. Therefore all the quantities $g(\Phi k-1, \Phi k)$ assume identical values of $(g(\Phi(\alpha), \Phi(\beta))$, which are independent of k (where $\alpha, \beta = 1,2,3,4$), and thus may be considered as the elements of a certain matrix,

$$G = \begin{vmatrix} g(\Phi(1),\Phi(1)) & \ldots\ldots & g(\Phi(1),\Phi(4)) \\ . & & . \\ . & & . \\ g(\Phi(4),\Phi(1)) & \ldots\ldots & g(\Phi(4),\Phi(4)) \end{vmatrix} \qquad (4)$$

of order 4 with elements,

$$g(\Phi k\text{-}1(\alpha), \Phi k(\beta)) = g\Phi k\text{-}1(\alpha), \Phi k(\beta) = g\,\alpha,\beta \qquad (5)$$

Introducing cyclic conditions, i.e. setting $\Phi 0 = \Phi n$, we find that Z is equal to the sum of the diagonal elements of matrix G^n :

$$Z = Sp\, G^n = \lambda 1^n + \lambda 2^n + \lambda 3^n + \lambda 4^n \qquad (6)$$

where $\lambda 1$, $\lambda 2$, $\lambda 3$,$\lambda 4$, are the eigenvalues of matrix G. Since all the elements of G are positive, it has a largest eigenvalue [12] $\lambda \equiv \lambda 1 > |\lambda\beta|$ (β=2,3,4), which is real, positive and nondegenerate.

Then for $n >> 1$,

$$Z = \lambda^n \qquad (7)$$

The interaction partition function of the system under consideration gives complete information on its thermodynamic functions and interactions [12]. The interaction energy is completely characterized by a binary distribution function, i.e. by the mean fraction of neighbouring nucleotides which have interactions $\Phi(\alpha)$, $\Phi(\beta)$ (α,β=1,..4).
This fraction,

$$w(\Phi^{(\alpha)}, \Phi^{(\beta)}) = \frac{d \ln \lambda}{d \ln g_{\alpha,\beta}} \qquad (8)$$

and

$$w(\Phi^{(\alpha)}) = \sum_{\nu=1} \frac{d \ln \lambda}{d \ln g_{\nu,\alpha}} \qquad (9)$$

The average number of nucleotide units in a regular α-region of the chain, which consists of nucleotide units in interaction $\Phi^{(\alpha)}$ is equal to the ratio of the total number of nucleotides in that conformation to the number of transitions to monomer units in interaction $\Phi^{(\alpha)}$ from nucleotides in any other interaction,

$$w(\Phi(\alpha)) = \frac{1}{1 - \left(\frac{d \ln \lambda}{d \ln g_{\alpha,\alpha}} \Big/ \sum_{\nu=1} \frac{d \ln \lambda}{d \ln g_{\nu,\alpha}} \right)}. \qquad (10)$$

The right-hand and left-hand eigenvectors of matrix G are the columns of matrix V, while the left-hand eigenvectors are the rows of U. Matrices V and U are defined by,

$$GV = V\Lambda$$
$$UG = \Lambda U \qquad (11)$$

where, Λ is a diagonal matrix the elements of which are the eigenvalues, λ_α of matrix G.
It is found,

$$w(\Phi^{(\alpha)}, \Phi^{(\beta)}) = (g_{\alpha,\beta} / \lambda)_{u1\alpha v\beta 1} \quad (12)$$

$$w(\Phi(\alpha)) = u_{1\alpha} v_{\alpha 1} \quad (13)$$

$$w(\Phi(\alpha)) = \lambda / (\lambda - g_{\alpha\alpha}) \quad (14)$$

These quantities allow the polynucleotide to be characterized in terms of its nucleotide fractions (T,C,A,G); its dinucleotide interactions (TT,TC,TA,TG,…..GT,GC,GA,GG), a 4 x 4 matrix, and the average length of sequences (T,C,A,G) throughout its length. Once these percentage quantities are found for a nucleotide interacting with its neighbour, they can be established for any polynucleotide of the same composition by scaling with its number of neighbour interactions.

3 Calculated values from the free energy data for stacked 2-deoxyribose dinucleotides.

To reasonably compare the results with those for nucleotides stacking in solution a number of approximations are necessary. First the entropies of stacking were taken to be those calculated at the HF/6-31G* level of accuracy. With these values the free energies for stacking were calculated [7,11].
However, the bases make strong electrostatic and hydrogen bonds in the gas phase to the maximum values expected. A scaling factor is necessary to make the interactions in keeping with solution data. All of the free energies of stacking as given in Table 7 were scaled by a factor of 1/50. The same as previously used to scale free energy values from ribose dinucleotides [11]
In terms of statistical weights, e-ΔG/RT , the matrix G has the values,

$$\begin{vmatrix} 2.1400 & 1.7251 & 2.2610 & 1.8836 \\ 1.8906 & 2.0369 & 2.0618 & 2.4191 \\ 1.6910 & 2.0712 & 1.7322 & 2.0333 \\ 2.2517 & 3.3713 & 1.7228 & 3.0160 \end{vmatrix}$$

The eigenvalues of G are, 8.6541, 0.5132, 0.1525, -0.3946

The matrix V has the values

$$\begin{vmatrix} 0.4550 & -0.8472 & -0.8001 & -0.1600 \\ 0.4869 & 0.1559 & 0.3250 & -0.6680 \\ 0.4348 & 0.0718 & 0.5012 & 0.0880 \\ 0.6057 & 0.5827 & -0.0551 & 0.7214 \end{vmatrix}$$

and U has the values,

$$\begin{vmatrix} 0.4624 & 0.5399 & 0.4482 & 0.5478 \\ -0.6406 & 0.8760 & -1.4971 & 0.8516 \\ -0.3154 & 0.4017 & 1.6812 & -0.6469 \\ 0.0341 & -1.0944 & 0.7952 & 0.2834 \end{vmatrix}$$

This data allows the weights of the single nucleotides to be calculated as,

w(t) = 0.21039
w(c) = 0.26288
w(a) = 0.19488
w(g) = 0.33180

and the weights of the dinucleotide pairs to be calculated as shown in Table 1.

Table 1. Weights for deoxyribose dinucleotides at 298.15 K.

	T	C	A	G
T	0.05203	0.04488	0.05253	0.06096
C	0.05367	0.06187	0.05593	0.09141
A	0.03985	0.05223	0.03901	0.06378
G	0.06485	0.10391	0.04742	0.11564

The average number of nucleotides in a regular region is calculated as,
$\nu(tt) = 1.33$, $\nu(cc) = 1.31$, $\nu(aa) = 1.25$, $\nu(gg) = 1.54$

These results give,

w(t) + w(a) = 0.40527
w(c) + w(g) = 0.59468

or the ratio, (A+T) / (C+G) = 68.15 %.

This indicates that for a solution that is equimolar in the nucleotides the ratio of (A+T) / (C+G) < 1.0 in the species resulting from stacking. This ratio calculated at 86.1 % is as found in salmon sperm DNA [15].

3.1 A Comparison of the nucleotide frequency quantities for tRNA with the values calculated from the one-dimensional Ising Model of a stacked polynucleotide.

3.1.1 The Serine Anticodon

The serine anticodon depicted here is from the work of [16].
The excised sequence used here, is as indicated:
5'
GGAGGCGTGGCAGAGTGGTTTAATGCACCG
GTCTTGAAAACCGGCAGTCGCTCCGGCGAC
TCATAGGTTCAAATCCTATCGCCTCCGTAA
3'
It has 90 nucleotides and 89 interactions are considered.

The sum of the numbers of T,C,A,G nucleotides for the anticodon are, 20,24,20, 26,
The calculated numbers of T,C,A,G are 18.94, 23.66, 17.39, 29.86.
This is depicted in Fig1.
The RMS deviation for the anticodon is 2.40.

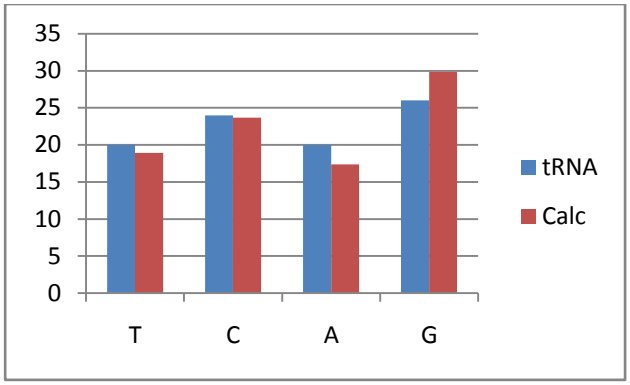

Fig.1. The number of deoxyribose nucleotides T,C,A,G in a stack of 90 for serine anticodon tRNA, compared to the values calculated from the stacking values using the Ising model [10].

Moreover in 8000 randomizations of the anticodon sequence at constant composition the anticodon was within 1% of the mean designating it as an entirely typical polynucleotide formed in accordance with the statistical weights for stacking, although this is a small sample compared to 4^{90} .

For the anticodon the numbers of neighbouring pairs is given as a matrix, Table 2.

	T	C	A	G
T	4	8	4	4
C	5	6	5	8
A	4	3	7	5
G	7	7	4	8

Table.2. Dinucleotide composition of the anticodon of serine tRNA.

The statistical frequencies for the anticodon are depicted in Fig.2 together with the calculated values

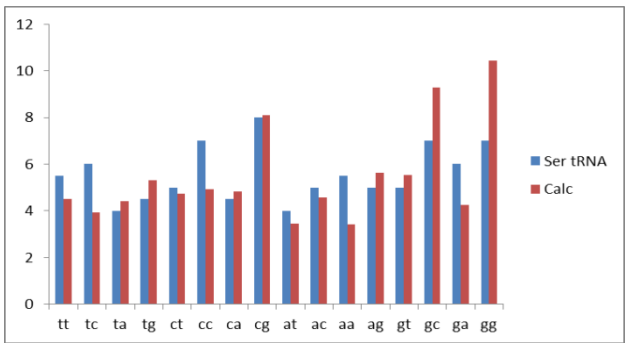

Fig.2. The number of stacked deoxyribose dinucleotides,16, in a stack of 90 for the serine anticodon compared to the values calculated from the stacking values using the Ising model [10].
The RMS deviation of these values is 1.64 .
Although quite large as a percentage, the polynucleotide may have many different sequences of the same energy.

3.1.2 The Serine Codon

The sum of the numbers of T,C,A,G nucleotides for the codon are, 20,26,20, 24,
The RMS deviation for the codon is 3.46
The dinucleotide composition of the serine tRNA codon is shown in Table 3, assumed to be the Watson-Crick complementary strand.

	T	C	A	G
T	7	4	4	5
C	5	8	4	8
A	4	7	4	5
G	3	7	8	6

Table 3. Dinucleotide composition of the codon of serine tRNA.

It is noted that the RMS deviation of the codon from the calculated composition is larger than for the anticodon. The difference is 1.06. This is a general characteristic for the tRNAs as shown in Table 4.

Table 4. A Comparison of the RMS deviations from the calculated values for the anticodon and codon for the tRNAs of T4 for T,C,A,G and the dinucleotides TT---GG..

tRNA	Ser tRNA anti	Ser tRNA codon	Arg tRNA anti	Arg tRNA codon

	codon		codon	
T,C,A,G RMS deviation	2.40	3.46	1.30	2.15
TT,---GG RMS deviation	1.64	2.03	1.79	2.09
	Thr tRNA anticodon	Thr tRNA codon	Pro tRNA anticodon	Pro tRNA codon
T,C,A,G RMS deviation	1.89	2.28	2.00	3.24
TT,---GG RMS deviation	2.20	2.48	2.49	3.05
	Ileu tRNA anticodon	Ileu tRNA codon	Leu tRNA anticodon	Leu tRNA codon
T,C,A,G RMS deviation	2.03	2.37	2.57	2.82
TT,---GG RMS deviation	2.82	3.04	2.55	2.72
	Gly tRNA anticodon	Gly tRNA codon	Gln tRNA anticodon	Gln tRNA codon
T,C,A,G RMS deviation	3.70	3.58	2.39	2.67
TT,---GG RMS deviation	3.57	3.47	2.61	3.20

3.1.3 The Serine Anticodon and Codon Stacking Free Energies.

The calculated values for the stacking free energy values of the anticodon and codon for several of the tRNA from [16-17] are shown in Table 5

Table 5. A Comparison of the stacking energies (h) for the anticodon and codon of tRNAs with the values calculated from the one-dimensional Ising Model of a stacked polynucleotide. T=298.15 K.

tRNA	Stacking free energy, ΔG from deoxyribose dinucleotides	Stacking free energy, ΔG from ribose dinucleotides
Ser tRNA anticodon	-3.22040	-2.92688
Ser tRNA codon	-3.19292	-2.88689
Arg tRNA anticodon	-2.65546	-2.36994
Arg tRNA codon	-2.61303	-2.35691
Thr tRNA anticodon	-2.62747	-2.37119
Thr tRNA codon	-2.54394	-2.35500
Pro tRNA anticodon	-2.54374	-2.25660
Pro tRNA codon	-2.42539	-2.20929
Ileu tRNA anticodon	-2.67985	-2.37114
Ileu tRNA codon	-2.62902	-2.31519
Leu tRNA anticodon	-2.98857	-2.71806
Leu tRNA codon	-2.97090	-2.69992
Gly tRNA anticodon	-2.36667	-2.16657
Gly tRNA codon	-2.38028	-2.18709
Gln tRNA anticodon	-2.54275	-2.23218
Gln tRNA codon	-2.47271	-2.21614

The stacking value for the serine tRNA codon, -3.19292 h, is higher than that of the anticodon, -3.22040. It was also found that the value for the anticodon is within 1 % of the average value calculated for 8100 sequences formed at random, but with the same base composition. This does suggest that at low ionic strength where repulsion between strands is dominant, that the anticodon formed first. However, at high ionic strength the dominance of the strength of GC stacking together with the strength of the G-C Watson Crick base pairing, suggests nucleation of anti-paralleling of the chains would occur and they would propagate together in opposite directions. However, neither the actual type of nucleotide that polymerized or the ionic strength are known [9].

For transcription of the DNA it is apparent that the Watson Crick hydrogen bonding provides both specificity and sufficient energy for an exergonic reaction of either the codon or the anticodon to replicate.

For translation of the codon it is only necessary for the free bases to stack to form the RNA anticodon. This stack being more abundant than the higher energy RNA codon. However, the DNA anticodon also has a favourable exergonic reaction with the RNA codon. However, this reaction appears less likely as the abundance of the RNA codon should be

much less according to the Boltzman Law. The relative values of these stacks are given in Table 5.
In most cases the RMS deviation of the mononucleotide composition is less for the anticodon than for the codon. In most cases RMS deviation of the dinucleotide composition is less for the anticodon than for the codon. The sequence of RMS values for the anticodon approximately follows that of the sequence of tRNA sequences in the genome [16], perhaps suggesting that nucleation sites began near serine and arginine sequences and the environment for stacking became more perturbed as the genome lengthened, as shown,

Genome Sequence

Gln - Leu - Gly - Pro - **Ser** - Thr - iLeu
---------- **Arg**

RMS values
2.61 2.57 3.70 2.00 2.40 1.89 2.03
----------1.30

3.2 The Stacking Energies for the Anticodon and Codon of tRNA for Arginine with the Values Calculated from the One-dimensional Ising Model of a Stacked Polynucleotide

Similar results were found for the arginine anticodon and codon.
The arginine anticodon depicted here is from the work of [16]
The excised sequence used here, is as indicated:

5’
GTCCCGCTGGTGTAATGGATAGCATACGATCCTTC
TAAGTTTGCGGTCCTGGTTCGATCCCAGGGCGGGA
TACCA
3’

It has 75 nucleotides and 74 interactions are considered.

The sum of the numbers of T,C,A,G nucleotides are, 20,19,14, 22,
The calculated numbers of T,C,A,G are, 15.48,19.85,14.55,25.13
The RMS deviation for the anticodon is 1.30. whilst that for the codon is 2.15, Table 4.

For the anticodon the numbers of neighbouring pairs is given as a matrix, Table 6

	T	C	A	G
T	4	6	5	5
C	4	7	3	5
A	6	2	2	3
G	6	4	4	8

Table.6. Dinucleotide composition in the anticodon of arginine tRNA.

Also shown are the numbers for the codon itself, assumed to be the Watson-Crick complementary strand, Table 7.

	T	C	A	G
T	2	4	5	3
C	3	8	5	5
A	6	6	4	4
G	2	4	6	7

Table.7. Dinucleotide composition in the codon of arginine tRNA.

The corresponding calculated values for the arginine tRNA are, Table.8.

	T	C	A	G
T	3.85	3.32	3.89	4.51
C	3.97	4.58	4.14	6.77
A	2.95	3.87	2.89	4.72
G	4.80	7.69	3.51	8.56

Table.8. Calculated dinucleotide composition for the arginine tRNA.

The RMS deviation of these values for the anticodon is 1.79 and for the codon 2.09, Table 4. Although quite large as a percentage, the polynucleotide may have many different sequences of the same energy. The RMS difference between the values for the codon and anticodon is 0.30.
The stacking energies for the bases of the anticodon and codon are given in Table 5.
It was again found that the value for the anticodon is within 1 % of the average value calculated for 8100 sequences formed at random, but with the same base composition, whilst that for the codon was higher.

3.3. The Stacking Energies for the Anticodon and Codon of tRNA for Ileucine with the Values Calculated from the One-dimensional Ising Model of a Stacked Polynucleotide

Similar results were found for the ileucine anticodon and codon.
The ileucine anticodon depicted here is from the work of [16].
The excised sequence used here, is as indicated:

5’
GGCCCTGTAGCTCAATGGTTAGCAGCAGTCCCCTC
ATAAGGGAAAGGTTACCAGTTCAAATCTGGTCTGG
GTCATAT
3’

It has 77 nucleotides and 76 interactions are considered.
The sum of the numbers of T,C,A,G nucleotides are, 20,18,19, 20,
The calculated numbers of T,C,A,G are, 15.90,20.38,14.93,25.80
The RMS deviation for the anticodon is 2.03. whilst that for the codon is 2.37.

For the anticodon the numbers of neighboring pairs is given as a matrix, Table 9.

	T	C	A	G
T	3	7	5	4
C	5	6	7	0
A	5	1	6	7
G	7	4	1	8

Table.9. Dinucleotide composition in the anticodon of ileucine tRNA.

Also shown are the numbers for the codon itself, assumed to be the Watson-Crick complementary strand, Table.10.

	T	C	A	G
T	6	1	5	7
C	7	8	4	0
A	5	7	3	5
G	1	4	7	6

Table.10. Dinucleotide composition in the codon of ileucine tRNA.

The corresponding calculated values for the ileucine tRNA are, Table.11.

	T	C	A	G
T	3.88	3.40	3.82	4.59
C	4.07	4.25	4.17	7.00
A	2.98	3.95	2.96	4.85
G	4.76	8.01	3.66	9.02

Table11. Calculated dinucleotide composition in ileucine tRNA.

The RMS deviation of these values for the anticodon is 2.82 and for the codon 3.04, Table 4. Although quite large as a percentage, the polynucleotide may have many different sequences of the same energy. The RMS difference between the values for the codon and anticodon, RMS = 0.22.
It was again found that the value for the anticodon is within 1 % of the average value calculated for 8100 sequences formed at random, but with the same base composition, but that for the codon is higher..

3.4 The Stacking Energies for the Anticodon and Codon of tRNA for Threonine with the Values Calculated from the One-dimensional Ising Model of a Stacked Polynucleotide

Similar results were found for the threonine anticodon and codon.
The threonine anticodon depicted here is from the work of [16].
The excised sequence used here, is as indicated:

5’
GCTGATTTAGCTCAGTAGGTAGAGCACCTCA
CTTGTAATGAGGATGTCGGCGGTTCGATTCC
GTCAATCAGCACCA
3’

It has 76 nucleotides and 75 interactions are considered.
The sum of the numbers of T,C,A,G nucleotides are, 20,16,17, 20,
The calculated numbers of T,C,A,G are, 15.69,20.11,14.74,25.46
The RMS deviation for the anticodon is 1.89. whilst that for the codon is 2.28.

For the anticodon the numbers of neighboring pairs is given as a matrix, Table.12.

	T	C	A	G
T	5	7	4	4
C	4	3	7	4
A	5	3	2	7
G	6	5	5	4

Table.12. Dinucleotide composition in the anticodon of threonine tRNA.

Also shown are the numbers for the codon itself, assumed to be the Watson-Crick complementary strand, Table.10.

	T	C	A	G

T	2	5	4	7
C	7	4	4	4
A	5	6	5	4
G	3	5	7	3

Table.13. Dinucleotide composition in the codon of threonine tRNA.

The corresponding calculated values are, Table.14.

	T	C	A	G
T	3.90	3.37	3.94	4.57
C	4.03	4.64	4.20	6.86
A	2.99	4.78	4.86	4.78
G	4.86	7.79	3.56	8.67

Table.14. Calculated dinucleotide composition in threonine tRNA.

The RMS deviation of these values for the anticodon is 2.20 and for the codon 2.48. Although quite large as a percentage, the polynucleotide may have many different sequences of the same energy. The RMS difference between the values for the codon and anticodon is 0.28.
It was again found that the value for the anticodon is within 1 % of the average value calculated for 8100 sequences formed at random, but with the same base composition.

3.5 The Combined Stacking and Watson-Crick Interaction Energies for tRNA Duplexes.

The following Table 15 and Table 16, depict the total interaction energies for a tRNA codon with its complimentary RNA strand and compares that energy with that for the tRNA anticodon with its complimentary RNA strand.

tRNA No.of Nucleotides	DNA codon stacking energy	RNA anti codon stacking energy	Watson-Crick Interaction Energy	Total DNA-RNA Duplex Energy
Ileu77	-2.629	-2.371	-3.135	-8.135
Arg75	-2.613	-2.370	-3.127	-8.110
Ser90	-3.193	-2.927	-3.766	-9.886
Thr76	-2.544	-2.371	-2.963	-7.878
Leu87	-2.971	-2.718	-3.525	-9.214
Gly74	-2.380	-2.167	-2.912	-7.459
Pro74	-2.425	-2.257	-3.022	-7.704
Gln73	-2.472	-2.232	-2.917	-7.622

Table 15. Total interaction energies codon DNA with the complimentary RNA.

tRNA No.	DNA anticodon stacking energy	RNA codon stacking energy	Watson-Crick Interaction Energy	Total DNA-RNA Duplex Energy
Ileu77	-2.680	-2.315	-3.135	-8.130
Arg75	-2.656	-2.357	-3.127	-8.139
Ser90	-3.220	-2.887	-3.766	-9.849
Thr76	-2.628	-2.355	-2.963	-7.946
Leu87	-2.989	-2.700	-3.525	-9.214
Gly74	-2.367	-2.187	-2.912	-7.466
Pro74	-2.544	-2.209	-3.022	-7.775
Gln73	-2.543	-2.216	-2.917	-7.676

Table 16. Total interaction energies for the anticodon DNA with the complimentary RNA.

Except for the case of serine tRNA these results tend to favour the transcription of the polynucleotide strand that produces the least stable duplex [18] although the Watson-Crick energies are not well differentiated for the DNA and RNA interactions.

4.Conclusion

The matrix of statistical weights can be diagonalised to produce the largest eigenvalue leading to a prediction of the average composition of polynucleotide with the respective number of nucleotides found in the tRNAs of bacteriophage, T4 [16-17].
The composition and number of nearest neigbours found in the respective tRNAs is within an average of two nucleotides, exact correlation not being expected owing to the enormous number of degenerate sequences with the same composition.
For all the tRNAs studied for serine, threonine, arginine and i-leucine, leucine, proline, glycine and glutamate the free energy of stacking for the anticodon was appreciably lower than for the complimentary codon (presumed Watson-Crick).
The RMS deviation of the codon from the average calculated value was greater than for the anticodon.
The anticodon stacking energies were within 1% of the average for sequences of the same composition, whereas the codon free energies of stacking were higher.

Both the codon and anticodon of the tRNAs may be transcribed in exergonic reactions owing to the favourable Watson Crick hydrogen bonding interaction energy. They are thermodynamically accessible in each case due to a combination of the (Watson-Crick free energy of hydrogen bonding plus the free energy for stacking) of the codon base being negative. Seldom is the Watson-Crick interaction sufficient on its own to form the complimentary strand.
Both the codons and anticodons may be naturally translated as the free energy change to form the complimentary strand (Watson-Crick hydrogen bonding + stacking interaction) is negative. However, the concentration of the correct RNA anticodon should be much larger than for the RNA codon.
There is a potential for an error in translation of the codons if the free energy change for the stacking of an added base is more negative than for the sum of the (Watson –Crick hydrogen bonding interaction + the stacking free energy) for the correct complimentary added base, unless steric effects are dominant.
The only biological selection from a prebiotic perspective appears in the subsequent use of the translated modified tRNAs in molecular evolution. This would appear sufficient for these sequences to be preserved since ancient times.

Further work may support the hypothesis that the bases stacked before they polymerized and all living organisms contain an afterglow of biological creation.

Acknowledgements

Appreciation is expressed for the advice and support given to this project by Professor Curt Wentrup of the University of Queensland.
Appreciation is expressed to APAC for facilities at the ANU and QCIF facilities at UQ, and the assistance of Mr.D.Green and H.Hartig.

References

[1] W.G.Walter and R.H.McBee, *General Microbiology*, D.Van Nostrand Comp.Inc. Princeton,1962.
[2] A.L.Lehninger, *Biochemistry*, Worth,New York,1975.
[3] P.O.P.Ts'0, I.S.Melvin and A.C.Olsen, Interaction and association of bases and nucleosides in aqueous solution, *J.Am.Chem.Soc.*, 85, 1963, pp.1289-1296.
[4] P.O.P.Ts'O. Ed, Bases,nucleosides and nucleotides, in *Basic Principles of Nucleic Acid Chemistry*, Academic Press, N.Y.1974, pp.454-584.
[5] D.N.Holcomb and I.Tinoco,Jr., Conformation of polyriboadenylic acid: pH and temperature dependence, *Biopolymers*, 3, 1965, pp.121-133.
[6] R.F.Steiner and R.F.Beers, *Polynucleotides*, Elsevier Publ., Amsterdam, 1961.
[7] N.N.Aylward, Submitted paper for WSEAS, 5th International Conference on Health Science and Biomedical Systems (HSBS '16), Ischia, Italy, June 17-19, 2016.
[8] W.J.Hehre, I.Random, P.V.R. Schleyer, and J.A.Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York., 1986.
[9] N.N.Aylward and N.Bofinger, *J.Theoretical Biology*, 244, 2006, pp.258-274.
[10] E.Ising, *Z.Physik*., 31, 1925, pp.253.
[11] N.Aylward, The relative thermodynamic stability of base stacking in pyrimidine / purine dinucleotides, *WSEAS 14th. International Conference on Computers*, Vol 1, Corfu Greece, 2010, pp.331-337.
[12] T. M. Birshtein and O. B. Ptitsyn *Conformations of Macromolecules* (High Polymers, Vol. XXII), Interscience, New York, 1966.
[13] M.V. Volkenstein, *Configurational Statistics of Polymeric Chains*, Interscience, N.Y.1963.
[14] D.H.Turner, in V.A.Bloomfield, D.M.Crothers, and I.Tinoco,Jr.,(eds.) *Conformational Changes in Nucleic Acids:structures,properties and functions*. University Science Books, Sausalito, Calif.,2000.
[15] D.O. Jordan, *The chemistry of nucleic acids*, Butterworths, 1960.
[16] G.P.Mazzara, G.Plunkett III and W.H.McClain, DNA sequence of the transfer RNA region of bacteriophage T4: Implications for transfer RNA synthesis, *Proc.Natl.Acad.Sci*., USA. 78, 2, 1981, pp.889-892.
[17] European Bioinformatics Institute (EBI) Data Library, http: //www.uni-bayreuth.de /departments/biochemie/trna.
[18] M.Jia and Y.Li, The relationship among gene expression, folding free energy and codon usage bias in Escherichia coli, *FEBS Letters*,579, 2005, pp.5333-5337.

.