Free Rad. Res. Comms., Vol. 6, No. 2-3, pp. 199-200
Reprints available directly from the publisher
Photocopying permitted by license only

SPECIAL INVITED LECTURE

THE THERMODYNAMIC BASIS OF THE GENETIC CODE

HORST H. KLUMP

Department of Biochemistry, University of Cape Town, Rondebosch, RSA

A systematic investigation of the helix coil transition of synthetic and of naturally occuring DNA sequences by means of microcalorimetry leads to a complete set of thermodynamic parameters as function of the two most important variables of the system,[1] the net base composition (% CG base pairs) and the total counterion concentration. The experimental results are the following, exemplified for the ΔG function:

i) The transition free enthalpy per base pair increases linearly with the net GC content. The slope of the ΔG vs %GC plot is identical for synthetic and for natural DNA sequences, i.e. the stability of an individual base pair inside a long DNA sequence is only marginally dependent of the character of the neighbouring base pair. As a consequence of this the transistion free enthalpy of any known sequence can be calculated from this data-set. This allows a straight forward calculation of the interaction energy for all the triplets of the genetic code.

ii) The entropy per base pair is sequence independent and reflects solely the gain in rotational degrees of freedom of six single bonds per nucleotide upon denaturation.

iii) Since the transition temperature is also linearly dependent of the net sequence composition, the experimentally obtained enthalpy change per base pair and the corresponding transition temperature can serve to calculate the standard free energy change ΔG for an individual base pair.

iv) This value is, as expected, almost independent of the character of the next nearest neighbours.

The comparison of the observed codon usage and the interaction free energy per triplet leads to the conclusion, that it is the interaction free energy ΔG which determines the preference for certain codons in certain genes. The observed codon usage in animal genes was copiled by R. Grantham *et al.*[2]

For a short discussion we can subdivide the triplets into the base preference in position I and II on one hand and in position III on the other hand. A great contrast in the use of C and T as third base exists between the highly or weakly expressed genes (based on quantities of protein synthesized).

The high use of C in quartet codons and duet codons for the highly expressed animal genes reflects the preference of precision over speed, i.e. it is the most tightly bound codon-anticodon complex which is selected for. The base choices are such as to assure the same mean pairing energy with either triplets from quartets or duets.

Contrasted to this clear pattern of choices between pyrimidines, selection between purines for filling the position III is more complex. There is a clear avoidance of the CG doublet by third base use in animals, which disfavors the codons TCG, CCG,

TABLE I

	T		C		A		G		
	1.92	13	3.46	16	−0.04	10	3.46	10	T
T	2.42	28	3.94	18	1.10	23	4.30	13	C
	1.25	2	3.46	9	1.61	stop	3.26	stop	A
	2.42	9	4.30	2	1.13	stop	3.94	12	G
	2.95	9	4.43	14	2.30	10	5.14	8	T
C	3.46	27	4.91	17	3.46	21	5.98	11	C
	2.30	7	4.43	10	2.95	10	5.14	4	A
	3.46	47	5.27	5	3.46	28	5.63	5	G
	0.61	11	3.46	15	1.25	8	3.46	12	T
A	1.13	24	3.94	28	2.42	28	4.30	21	C
	−0.04	4	3.46	11	1.91	19	3.46	8	A
	1.13	16	4.30	6	2.42	49	3.94	10	G
	2.30	9	5.14	28	2.30	16	4.43	22	T
G	3.46	21	5.63	38	3.46	24	5.27	32	C
	2.95	5	5.14	14	2.75	21	4.43	16	A
	3.46	33	5.98	6	3.46	36	4.91	11	G

The standard free enthalpy ΔG for the codon/anticodon interaction per base triplet in kcal/mole triplet (left number), and the observed codon usage in the genes of animals (right number in each column)

ACG, and GCG. As a group these codons are rare in eucaryotes. The potential signal function of the methylated CG in gene control may be the reason for its suppression.

An analysis for the codon usage in bacterial genes would lead to a different picture. (Data not shown) Codon-anticodon pairs of intermediate energy are favored. This preference is favorable for fast and smooth reading by the bacterial polymerase.

Summing up the brief remarks about preferences of codon usage in bacteria and in animals it is tempting to state that the interaction free energy per base triplet is used to modulate genome strategies for gene expression to serve either the need for speed or for precision.

References

1. Klump, H.H. *Biochemical Thermodynamics*, sec. edition, N. Jones, Ed., Elsevier Publ. Amsterdam, (1988).
2. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. and Mercier, R. Codon Catalog Usage is a Genome Strategy Modulated for Gene Expression, *Nucl. Acids Res.*,**9**,p. 43, (1981).