

*Letters to the Editor***Beyond the Wobble: The Rule of Conjugates****Bhyravabhotla Jayaram**

Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi-110016, India

Received: 29 November 1996 / Accepted: 8 June 1997

**Abstract.** A reexamination of the genetic code suggests a rule of conjugates which captures the observed quartet degeneracies without exception. *Adenine is the conjugate of cytosine and uracil is the conjugate of guanine.* Further analysis reveals that the rule of conjugates is a macrolevel manifestation of the molecular-level hydrogen-bonding and base-stacking interactions at the decoding site. This new perspective is of significance to evolutionary discussions of nucleic acid bases, genetic code, and interactions involving RNAs.

**Key words:** Translation — Wobble — Degeneracies — RNA-RNA interactions — Genetic code

Ever since the genetic code has been deciphered, a mystery has surrounded the degeneracies observed in the code (Watson et al. 1987). The number of amino acids coded for with a three-letter code is not 64, not 48 with the equivalence of the pyrimidines (cytosine, C and uracil, U) in the third position, not even 32 (if the purines adenine, A and guanine, G were also to be equivalent to each other), but 20. While the wobble hypothesis (Crick 1966) captures the general spirit of the molecular events at the recognition site—the dynamics of the third base of the codon on m-RNA in the presence of the anticodon and in the milieu of ribosomal machinery in particular—

not just the degeneracies but the wobble itself appears to be idiosyncratic. A further analysis of the code reveals some beautiful, unexceptionable, inherent interrelationships between the nucleic acid bases, notably in the second position and to an extent in the first position of the codon, which partly explain the degeneracies, providing a molecular basis for the nature of the observed wobble.

The wobble hypothesis allows for G.U base pairing at the third position of the codon in addition to the usual Watson-Crick G.C and A.U pairing, laying out a mechanism for the realization of the degeneracies. An alternative explanation for some of the observed degeneracies comes from the “two out of three proposal” (Lagerkvist 1978), which maintains that only the first two bases of the codon are recognized by the anticodon. Consider the quartet degeneracies in the genetic code (given in Table 1 for ready reference), i.e., codons wherein the third base is inconsequential to the specification of the amino acid coded for. It is not apparent how the first two bases of the codon would cause these degeneracies and whether any prediction could be made based on the first two bases.

An analysis of the code suggests a rule of conjugates. (The word “conjugate” is used almost synonymously with “antonym” and “opposite” here.) *A is the conjugate of C and U is the conjugate of G.*  $A_1$  and  $A_2$  (adenine in the first and second positions of the codon) act as the conjugates of  $C_1$  and  $C_2$ . Similarly  $G_1$  and  $G_2$  act as the conjugates of  $U_1$  and  $U_2$ . If  $A_1A_2$  forms a non-quartet,  $C_1C_2$  forms a quartet. Similarly  $A_1U_2$ ,  $U_1A_2$ ,  $U_1U_2$ ,  $A_1G_2$ ,  $G_1A_2$ ,  $C_1A_2$ ,  $U_1G_2$ , act as the conjugates of  $C_1G_2$ ,  $G_1C_2$ ,  $G_1G_2$ ,  $C_1U_2$ ,  $U_1C_2$ ,  $A_1C_2$ ,  $G_1U_2$ , respec-

**Table 1.** The genetic code<sup>a</sup>

First base ↓	Second base				Third base ↓
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	End	End	A
	Leu	Ser	End	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

<sup>a</sup> Format adapted from Schulz and Schirmer (1979)

tively. Thus, simply by noting the behavior of one-half of the codons, the wobble behavior of the other half can be predicted.

A closer examination helps to unravel the chemical principles behind this prediction scheme. While the demands on the second base are stringent with regard to the above rule of conjugates, if the predictions are to have no exception, those on the first base are a bit more flexible. For A<sub>1</sub>, even G<sub>1</sub> serves as an antonym, and similarly, C<sub>1</sub> can serve as the antonym of U<sub>1</sub>. This gives a parallel series with an interchange of G<sub>1</sub> for C<sub>1</sub> and vice versa, and similarly A<sub>1</sub> for U<sub>1</sub>. (A<sub>1</sub>A<sub>2</sub>, A<sub>1</sub>U<sub>2</sub>, U<sub>1</sub>A<sub>2</sub>, U<sub>1</sub>U<sub>2</sub>, A<sub>1</sub>G<sub>2</sub>, U<sub>1</sub>G<sub>2</sub>, C<sub>1</sub>A<sub>2</sub>, G<sub>1</sub>A<sub>2</sub>, act as the conjugates of G<sub>1</sub>C<sub>2</sub>, C<sub>1</sub>G<sub>2</sub>, C<sub>1</sub>C<sub>2</sub>, C<sub>1</sub>G<sub>2</sub>, G<sub>1</sub>U<sub>2</sub>, C<sub>1</sub>U<sub>2</sub>, U<sub>1</sub>C<sub>2</sub>, A<sub>1</sub>C<sub>2</sub> respec-

tively.) This establishes that while it is only the number of hydrogen bonds formed by the first base that is of significance to wobble, it is both the size (pyrimidine vs purine) and the number of hydrogen bonds (two vs three) formed by the second base of the codon which determine the wobble.

The rule of conjugates carries with it a stereochemical interpretation of the wobble. With six hydrogen bonds formed by the first two bases (G<sub>1</sub>G<sub>2</sub>, G<sub>1</sub>C<sub>2</sub>, C<sub>1</sub>G<sub>2</sub>, C<sub>1</sub>C<sub>2</sub>), the identity of the third base is irrelevant. With four hydrogen hydrogen bonds formed by the first two bases (U<sub>1</sub>U<sub>2</sub>, U<sub>1</sub>A<sub>2</sub>, A<sub>1</sub>U<sub>2</sub>, A<sub>1</sub>A<sub>2</sub>), the third base has to be specified. With five hydrogen bonds formed by the first two bases, a pyrimidine in the second position (A<sub>1</sub>C<sub>2</sub>, U<sub>1</sub>C<sub>2</sub>, G<sub>1</sub>U<sub>2</sub>, C<sub>1</sub>U<sub>2</sub>) renders the third base inconsequential but a purine (C<sub>1</sub>A<sub>2</sub>, G<sub>1</sub>A<sub>2</sub>, U<sub>1</sub>G<sub>2</sub>, A<sub>1</sub>G<sub>2</sub>) necessitates the specification of the third base. Viewed in this context, wobble is a manifestation of the conformation and dynamics of the third base of the codon as facilitated by the first two bases.

That the behavior of the codons can be condensed into a simple rule is too fascinating and thought provoking to ignore. From an evolutionary standpoint, the Watson-Crick base pairing (rule of complementarity) and the rule of conjugates together sum up the nature of mRNA-tRNA recognition in general terms. Whether the intrinsic dynamics conveyed by the rule of conjugates is specific to codon-anticodon recognition or is more general to interactions involving single-stranded RNAs remains to be seen. For now, the role of the preceding bases in determining the wobble stands accentuated.

## References

- Crick FHC (1966) Codon-anticodon pairing: the wobble hypothesis. *J Mol Biol* 19:548–555
- Lagerkvist ULF (1978) “Two out of three”: an alternative method for codon reading. *Proc Natl Acad Sci USA* 75:1759–1762
- Schulz GE, Schirmer RH (1979) Principles of protein structure. Springer-Verlag, New York, p 7
- Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM (1987) Molecular biology of the gene, vol 1, chap 15. 4th ed. (The Benjamin/Cummings, CA)