Wobble base-pairing in codon-anticodon interactions: A theoretical modelling study

S Ulen Mangang and R H Duncan Lyngdoh*

Department of Chemistry, North-Eastern Hill University, Bijni Complex, Shillong 793 003, India

The Crick wobble hypothesis attributes the phenomenon of codon degeneracy to a certain impreciseness of pairing between the third base of the codon and the first base of the anticodon. This theoretical study investigates the pairing properties of some wobble bases, including both, observed and unobserved pairs. Some wobble base-pairs are predicted to follow the Watson-Crick pairs in configuration and pairing facility, while others deviate from this norm. The observed U:V pair is unique in that a pairing configuration may be suggested for it wherein the hydrogen-bonding involves the exocyclic 5-carboxymethoxy group of V. By comparing the theoretical data on the configurations of these pairs with the evidence for their existence/non-existence in nature, some guidelines emerge for differentiating between observed and unobserved base pairs on the basis of the pairing configuration.

Introduction

The phenomenon of codon degeneracy has been attributed by Crick[1] to a certain impreciseness (or "wobble") in hydrogen-bonded pairing between the third base of the codon and the first of the anticodon. A good number of such wobble pairs have been experimentally observed. Table 1 lists some observed wobble pairs along with numerous pairs never observed[2]. This theoretical study seeks to differentiate between observed and unobserved wobble pairs purely on the basis of the pairing configuration. Since the first two codon bases pair in approximately Watson-Crick fashion, it is assumed here that the wobble base would accommodate its anticodon counterpart only if the pairing configuration does not deviate too much from the standard Watson-Crick alignment.

This study examines the pairing possibilities from the anticodon side. The following anticodon bases are chosen for study: guanine (G), inosine (I), cytosine (C), N^4-acetylcytosine (ac^4C), uracil (U) and 5-carboxymethoxyuracil (V). Each of these is paired with the four RNA bases from the codon side, viz. A, G, T and U. The pairing energy E_p and the pairing configuration are taken note of to search for determinants which could effectively differentiate between observed and unobserved pairs.

Theoretical methodology

The starting configurations of the base pairs were set up using the PCMODEL package. Molecular orbital calculation and full geometry optimisation were done using the semiempirical PM3 SCF-MO method[3,4] which employs the Davidon-Fletcher-Powell optimisation routine[5,6]. The final energies and geometries thus obtained served as the basis for this study of pairing facility and configuration.

All pairs used a methyl group to represent the sugar moiety. The quantities of interest are the pairing enthalpy E_p in kCal/mol as well as the configuration of the wobble pairs. Each pair is defined by the three configurational markers, viz. the distance R_{nn} between the two would-be glycoside bonded nitrogens N1 and N2 of the two bases involved, the angles θ_1 and θ_2 made by the "glycoside" C-N bonds with the N1-N2 vector, and the dihedral angle ϕ made by C1-N1 on the codon side and N2-C2 on the anticodon side.

*Author for correspondence
Phone: (0364) 25 0041 * 39; E-mail: harku@nehus.ren.nic.in

Table 1—Observed and unobserved wobble pairs involving various anticodon wobble bases

No	Wobble base	Observed pair(s)	Unobserved pair(s)
1	G	C:G, U:G, A:G	G:G
2	I	C:I, U:I, A:I	G:I
3	Q	C:Q	C:Q, U:Q, A:Q
4	C	G:C	U:C, C:C, A:C
5	m^5C	G:m^5C	U:m^5C, C:m^5C, A:m^5C
6	ac^4C	G:ac^4C	U:ac^4C, C:ac^4C, A:ac^4C
7	U	A:U, G:U	U:U, C:U
8	V	A:V, G:V, U:V	C:V
9	S^{mn}	A:S^{mn}, G:S^{mn}	U:S^{mn}, C:S^{mn}
10	S^{mc}	A:S^{mc}, G:S^{mc}	U:S^{mc}, C:S^{mc}

These configurational markers were used in previous work of the author[7,8]. In all the figures, the pointed arrow depicts the glycoside bond.

Results and Discussion

Each wobble base is paired with the four RNA bases A, G, U and C. Table 2 presents the configurational data for pairs arising from G and I. Table 3 gives the configurational data for the pairs arising from C and ac⁴C, and Table 4 for the pairs arising from U and V.

All the pairs studied, whether observed or unob-

Table 2—PM3 data on configuration of pairs arising from guanine and inosine as anticodon wobble bases*

No.	Base pair	E_p	R_{nn}	θ_1	θ_2	ϕ	Status
Pairs from G							
1	A:G	-6.99	11.003	135.06	134.73	-4.54	Obsvd
2	G:G	-10.52	11.240	146.98	147.15	178.57	Unobs
3	U:G	-6.03	9.066	110.97	142.25	13.93	Obsvd
4	C:G	-11.68	9.103	127.65	127.58	-2.40	Obsvd
Pairs from I							
1	A:I	-7.71	11.027	136.73	135.30	-8.51	Obsvd
2	G:I	-9.78	11.236	147.00	146.80	172.23	Unobs
3	U:I	-4.38	9.055	110.68	141.88	4.17	Obsvd
4	C:I	-10.03	9.112	128.16	127.82	4.64	Obsvd

*Units of E_p in kcal mol-1, distances in Å and angles in degrees.

Table 3—PM3 data on configuration of pairs arising from cytosine and N⁴-acetylcytosine as anticodon wobble bases*

No.	Base pair	E_p	R_{nn}	θ_1	θ_2	ϕ	Status
Pairs from C							
1	A:C	-6.02	9.589	156.68	139.52	154.58	Unobs
2	G:C	-11.68	9.103	127.58	127.65	2.40	Obsvd
3	U:C	-4.64	7.247	123.64	124.65	-12.62	Unobs
4	C:C	-6.25	8.439	152.90	152.62	144.26s	Unob
Pairs from ac⁴C							
1	A:ac⁴C	-4.48	9.528	155.03	137.14	136.20	Unobs
2	G:ac⁴C	-7.42	9.080	125.94	127.49	-5.83	Obsvd
3	U:ac⁴C	-1.49	7.128	125.96	121.27	25.68	Unobs
4	C:ac⁴C	-5.16	8.391	152.71	151.16	143.20	Unobs

*See Table 2 for units.

Table 4—PM3 data on configuration for pairs arising from uracil and 5-carboxymethoxyuracil as anticodon wobble bases*

No.	Base-pair	E_p	R_{nn}	θ_1	θ_2	ϕ	Status
Pairs from U							
1	A:U	-5.47	9.170	128.09	131.04	1.74	Obsvd
2	G:U	-6.03	9.066	142.25	110.97	-13.93	Obsvd
3	U:U	-4.37	7.429	99.06	139.90	11.	Unobs
4	C:U	-4.64	7.247	124.65	123.64	-12.62	Unobs
Pairs from V							
1	A:V	-5.93	9.134	127.78	128.16	-6.80	Obsvd
2	G:V	-6.37	9.069	143.04	110.12	-14.63	Obsvd
3	(I)U:V	-4.58	7.430	98.82	139.47	9.61	
	(II)U:V	-5.83	11.452	136.40	115.64	-62.91	Obsvd
	(III)U:V	-8.29	8.532	83.61	80.51	7.12	
4	C:V	-1.91	7.177	124.80	123.41	21.20	Unobs

*See Table 2 for units.

served, are more or less favoured thermodynamically. The hydrogen bonds for all cases are largely linear, with lengths as expected in the range of 1.7 to 2.0 Å.

Base-pairs from G and I

Fig. 1 portrays the eight pairs arising from G and I. The pairing energy E_p ranges from -4.4 to -11.7 kCal/mol (Table 2). It is noteworthy that the largest pairing energy is shown by the standard C:G pair with its three hydrogen bonds, while others have only two. All the pairs are basically planar.

Only the observed C:G and C:I pairs may be characterised as truly double-helical in configuration, not being expected to distort the double-helix during codon-anticodon pairing. The observed pairs A:G, U:G, A:I and U:I are not double-helical in their configuration, as shown by large R_{nn} values of around 11 Å for A:G and A:I, and by atypical θ_1 and θ_2 values for U:G and U:I. Since these pairs are observed, it may be in-

ferred that large R_{nn} distances upto even 11 Å may yet allow for accommodation during codon-anticodon pairing. The A:G mismatch has long been known to be accommodated into the double-helix[9,10]. The R_{nn} distances for U:G and U:I come close to typical values, and their atypical θ_1 and θ_2 values do not disqualify these as wobble pairs.

The unobserved pairs G:G and G:I are marked by large R_{nn} distances and atypical θ_1 and θ_2 values. It is however, their φ value of about 180° which really characterises the G:G and G:I pairs as being unable to be accommodated into the wobble position.

Base pairs from C and ac⁴C

The pairs arising from C and ac⁴C are portrayed in Fig. 2. The observed G:C and G:ac⁴C pairs display true double-helical configuration. The unobserved pairs U:C and U:ac⁴C are marked by short R_{nn} values (7.13 Å and 7.25 Å respectively) as shown in Table 3,
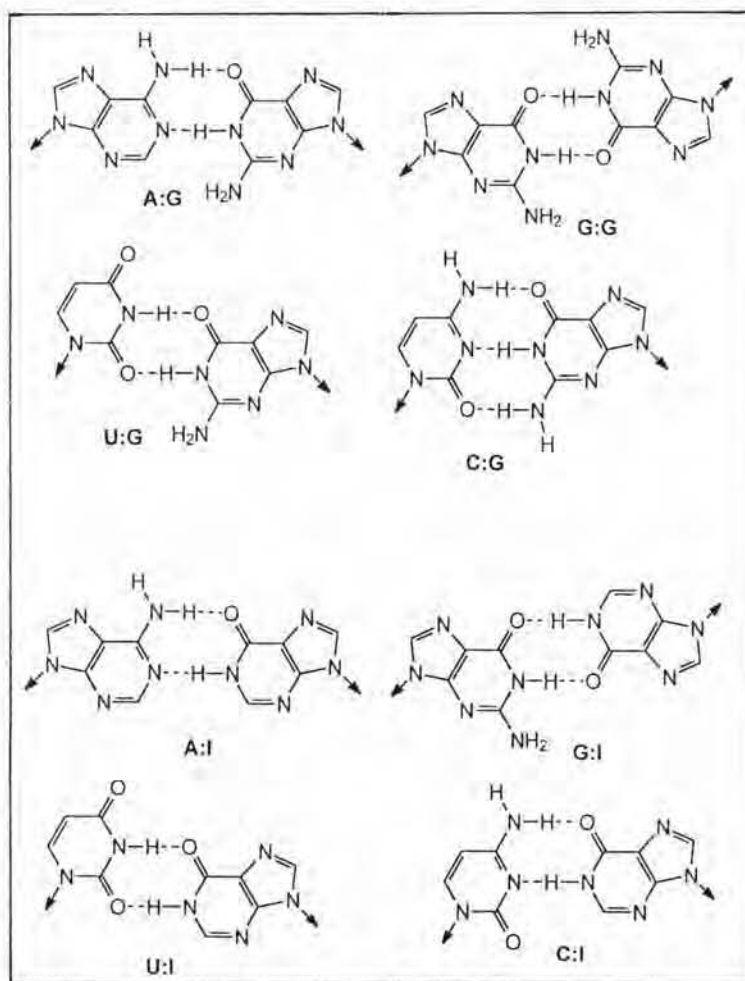

Fig. 1—Base pairs arising out of pairing of guanine (G) and of inosine (I) with the RNA bases A, G, C and U
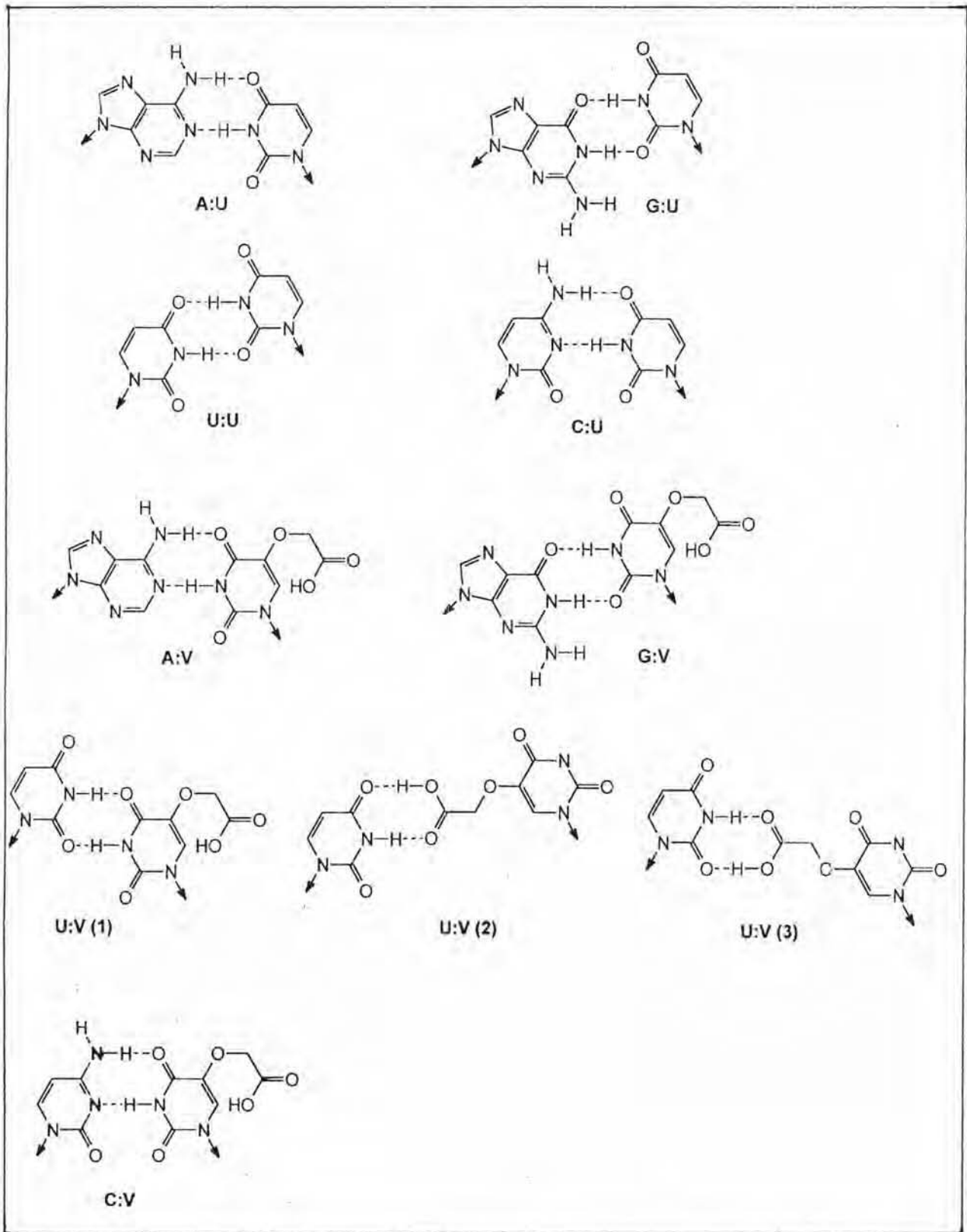
Fig. 3—Base pairs arising out of pairing of uracil (U) and of 5-carboxymethoxyluracil (V) with the RNA bases A, G, C and U (with three options for the U:V pair)
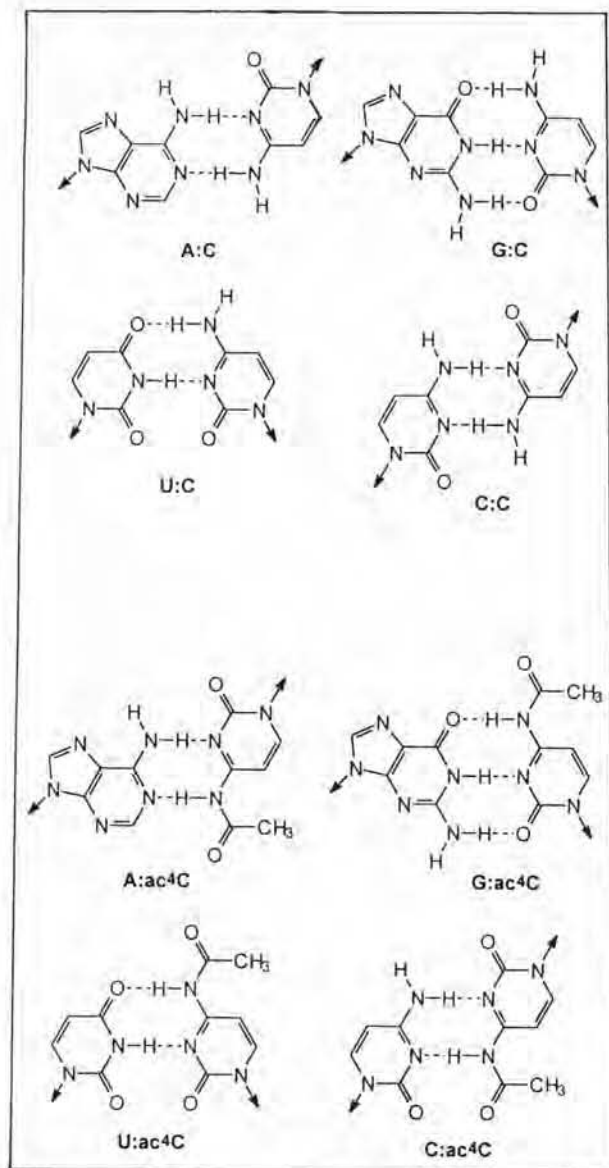
Fig. 2—Base pairs arising out of pairing of cytosine (C) and of N^4-acetylcytosine (ac^{4}C) with the RNA bases A, G, C and U

which could explain their absence in the wobble position. These unobserved pairs also demonstrate lesser or greater values for φ. The dihedral for the C:C and C:ac^{4}C pairs has the large values of 144.3° and 143.2° respectively, which could be the reason for disqualification from being incorporated into codon-anticodon pairing.

Base pairs from U and V

Table 4 shows pairing arising from U and V. The

observed G:U and G:V pairs show atypical θ_1 and θ_2 values as well as some deviation from planarity. However, their typical R_{nn} distances allow them to occur during codon-anticodon pairing.

The pyrimidine-pyrimidine pairs U:U, C:U, U:V and C:V show short R_{nn} values, which may be why U:U, C:U and C:V are not observed. The U:V pair, however, is known to be observed. The reason could be participation of the exocyclic 5-carboxymethoxy group of V during base pairing with V, leading to more typical R_{nn} distances. Three options I, II and III for the U:V pair are given in Fig. 3, of which III is the best choice.

Conclusions

It appears from this limited study that when values of θ_1, θ_2 or R_{nn} markers are close to those of the Watson-Crick base pairs, accommodation into the codon-anticodon pair may be allowed. Values of φ approaching 180° are a prime factor for disallowing the particular base pair. The only observed pyrimidine-pyrimidine pair U:V may be accommodated by invoking the role of the exocyclic 5-carboxymethoxy group of V in hydrogen-bonding with U.

References
1   Crick F H C (1966) *J Mol Biol* 19, 548-556
2   Sodd M A (1967) in *Handbook of Biochemistry and Molecular Biology 3rd edition Nucleic Acids Vol II* (G D Fasman, ed), 423-456, CRC Press, Cleveland, Ohio
3   Stewart J J P (1989) *J Comput Chem* 10, 221-264
4   Stewart J J P (1990) *J Comput Chem* 11, 543- 544
5   Fletcher R & Powell M J D (1963) *Comput J* 6, 163-167
6   Davidon W C (1968) *Comput J* 11, 406-410.
7   Venkateswarlu D, Lyngdoh R H D & Bansal M (1997) *J Chem Soc Perkin Trans 2*, 621-626
8   Venkateswarlu D & Lyngdoh R H D (1995) *J Chem Soc Perkin Trans 2*, 839-846
9   Brown T, Hunter W N, Kneale G & Kennard O (1986) *Proc Natl Acad Sci USA* 83, 2402-2406
10 Webster W D, Sanderson M R, Skelly J V, Neidle S, Swann P F, Li B F & Tickle I J (1990) *Proc Nat Acad Sci USA* 87, 6693-6697