NUCLEOTIDE CONFORMATIONS IN CODON-ANTICODON INTERACTIONS

C. J. Alden and Struther Arnott

Department of Biological Sciences
Purdue University
West Lafayette, Indiana 47907, U.S.A.

Received June 7,1973

Summary. The amount of distortion required of the sugar-phosphate backbone in RNA double-helices to accommodate various non-standard base pairs was examined by means of a linked-atom least-squares computer program. Compressive pyrimidine-pyrimidine pairs were found to entail more strain than the extensive adenine-hypoxanthine pair, as predicted by the wobble hypothesis. Purine-pyrimidine pairs involving guanine and uracil also require little torsional strain, indicating the acceptability of such pairings in hairpin stems as well as in codon-anticodon bindings.

Introduction. The accessibility of non-standard base-paring interactions in RNA is of some biological significance, notably in the "wobble" in the codon-anticodon interaction (1) and also in the structure of double-stranded regions of RNA that incorporate G-U pairs, as frequently occurs in stem regions of tRNA. We have examined the amount of torsional strain required to accommodate various base-pairing schemes by means of a linked-atom least-squares computer model building system (2). This system is superior to many in that all assumptions made have to be explicitly specified (and can therefore readily be up-dated, if necessary), and in that all models produced are quantified and thus readily reproducible and comparable.

Procedure. The linked-atom path considered was a closed loop around a double-stranded dinucleotide, as depicted in Fig. 1, and all atoms not directly on the path are included as pendants to atoms on the path. The three initial atoms were repeated at the terminus of the path and the condition that these triangles should coincide was included in the least-squares analysis (by the method of Lagrange's undetermined

Table 1. Base-pair parameters of Figure 1 for various base-pairings. For codon-anticodon binding the first named base of the pair is on the codon (3'-end).

Base-pair	b_1(Å)	b_2(Å)	b_3(Å)	α_1(°)	α_2(°)
Watson-Crick	3.11	7.05	3.11	128.0	128.0
G-U	2.90	7.77	1.56	135.5	119.5
U-G or U-I	1.56	7.77	2.90	119.5	135.5
A-I	3.11	10.02	0.96	129.3	138.0
U-U(1)	4.20	5.45	1.50	117.2	120.5
U-U(2)	1.50	5.45	4.20	120.5	117.2
C-U	4.22	4.31	4.15	121.8	117.0

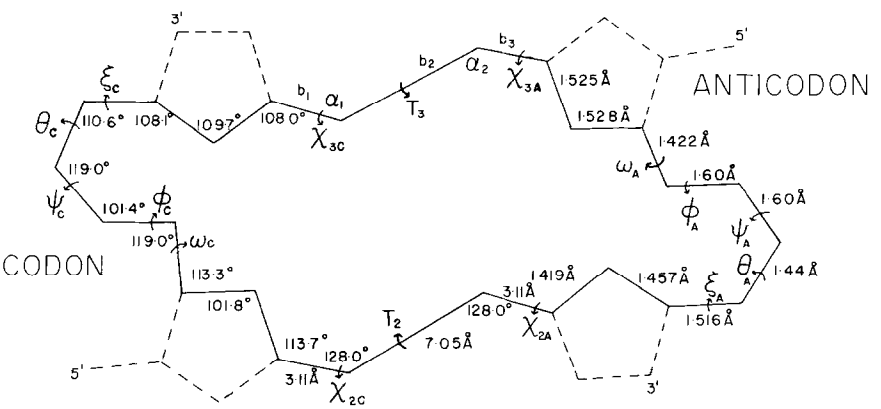

Figure 1. The linked-atom model-building path for a double-stranded dinucleotide in which a non-standard (upper) base-pair succeeds a standard, Watson-Crick base-pair. Bond-lengths and angles are fixed at the values shown and the variable conformation-angles (4) are indicated by arrows.

multipliers) (3). The common good approximation was made that bond-lengths and bond-angles are relatively invariant with respect to conformation-angles and those parameters were fixed at the values shown in Fig. 1. For each pairing scheme the appropriate values for the three lengths b_1, b_2, and b_3 and the two angles α_1 and α_2 and

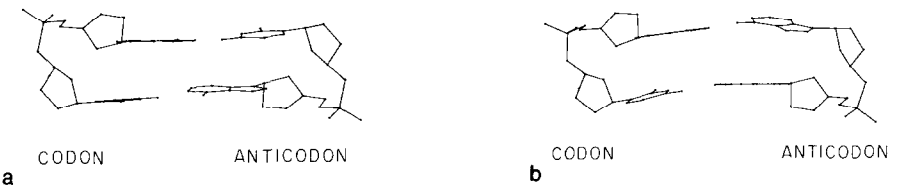

Figure 2. Two successive nucleotide pairs where the upper pair is (a) a standard Watson-Crick pair and (b) an A-I "wobble" pair. Note how very similar are the (C3-endo) sugar-phosphate backbones of both (a) and (b).

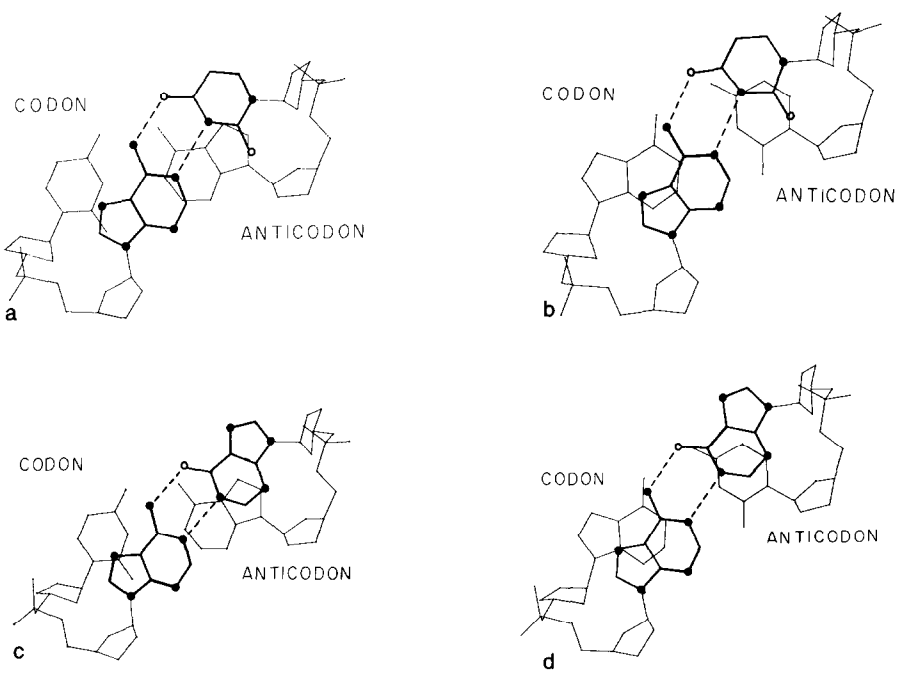

Figure 3. Base overlap for two nucleotide pairs, viewed normal to the plane of the lower codon base, for (a,b) two standard pairs and for (c,d) an A-I "wobble" pair over a standard pair. These also illustrate the small distortion required to accommodate the purine-purine pair and also the similarity in the stackings of A-I and a standard pair over the two kinds of standard pairs.

---

shown in Table 1. Thus in any refinement the only variables are conformation-angles, which are required to be as close as possible (in the least-squares sense) to assigned standard values. These standard angles are based on averages of torsion angle values derived from surveys of X-ray diffraction studies of large numbers of nucleic acid monomers and polymers (5). In least-squares modelling a weighting

factor is assigned to each torsion angle. This factor is 100/(standard deviation)^2, where the standard deviation for each torsion angle is taken from the same survey (5).

Thus there are two criteria for the goodness of any model. If the initial and final triangles of atoms cannot be induced to coincide (as is characteristic of overly compressed models) the model is clearly unacceptable. Second, the weighted sum of the differences of the variable torsion angles from their standard values provides a measure of the strain of the backbone; this sum may be expressed as $R = \sum \omega_n (\tau_n - \tau_{ns})^2/100N$, where ω_n is the weighting factor, τ_n is the refined torsion angle, τ_{ns} is the standard value of the torsion angle, and N is the number of variable torsion angles.

Results. The conditions we specified for the refinements of dinucleotide loops like Fig. 1 are as follows:

a) The torsion angles for all the riboses were fixed at values corresponding to a standard C3-endo conformation (5).

b) The standard and initial values chosen for the variable torsion angles were from A-RNA (6).

c) The initial and standard value chosen for T_3 (the angle between the planes of the upper pair of bases was 0^o with an assumed standard deviation of 5^o. T_2, χ_{2A}, and χ_{2C} were fixed at the values (-13.8^o, 74.9^o, 74.9^o) in A-RNA (6).

The three pyrimidine-pyrimidine pairs entail the greatest amount of torsional strain: the average root mean squared difference, i.e. $\langle R^{\frac{1}{2}} \rangle$, in a torsion angle from its standard value is more than one standard deviation. In contrast the purine-pyrimidine and purine-purine pairs can be accommodated with torsion angle differences only half that of the strained cases. In Table 2 the values of the torsion angles for all pairs examined are presented. Figures 2 and 3 emphasize the relatively small changes in a standard RNA backbone that are required to accommodate

Table 2. Conformation-angles for backbones accommodating various base-pairs in the upper position of Fig. 1. The mean values and standard deviations from these mean values in monomers (6) is given for comparison.

Angle	Monomer Mean and e.s.d.	Watson-Crick	U-G	G-U	A-I	C-U	U-U(1)	U-U(2)
χ_{3A}	81(11)	80	80	83	80	98	105	91
T_3	0(5)	1	1	1	0	3	4	2
χ_{3C}	81(11)	74	79	62	95	51	49	54
ω_C	-123(24)	-152	-158	-153	-145	-138	-138	-137
φ_C	-60(14)	-72	-69	-73	-86	-64	-69	-59
ψ_C	-60(14)	-69	-68	-69	-68	-71	-71	-70
θ_C	180(15)	171	164	-175	151	-168	-161	-176
ξ_C	52(8)	46	47	43	47	40	39	43
ω_A	-123(24)	-145	-149	-156	-159	-139	-133	-146
φ_A	-60(14)	-68	-64	-77	-64	-91	-102	-79
ψ_A	-60(14)	-64	-65	-61	-72	-58	-62	-56
θ_A	180(15)	169	169	164	-179	157	156	159
ξ_A	52(8)	48	47	51	41	41	40	43

the A-I wobble pair. Finally, the relative ease with which non-standard pairs involving G and U are accommodated indicates that they not only can occur in codon-anticodon interactions but would also fit readily into double-helical stems, rather than forming bulges or loops as is sometimes envisaged.

Acknowledgements. This research was supported by grants from NIH and NSF.

References.

1. Crick, F. H. C., J. Mol. Biol. 19, 548 (1966).

2. Arnott, S., Prog. in Biophys. and Mol. Biol. 21, 267 (1970).
3. Hancock, H., "The Theory of Maxima and Minima" Ch. 6 (Dover Publications, Inc., N.Y., 1960).
4. Arnott, S., and Hukins, D. W. L., Nature 224, 886 (1969).
5. Arnott, S., and Hukins, D. W. L., Biochem. J. 130, 453 (1972).
6. Arnott, S., Hukins, D. W. L., and Dover, S. D., Biochem. Biophys. Res. Comm. 48, 1392 (1972).