APPENDIX 1: Structures of Base Pairs Involving at Least Two Hydrogen Bonds

Provided by Mark E. Burkard and Douglas H. Turner
Department of Chemistry, University of Rochester
Rochester, New York 14627-0216

Ignacio Tinoco, Jr.
Department of Chemistry, University of California, Berkeley
Structural Biology Division, Lawrence Berkeley National Laboratory
Berkeley, California 94720-1460

The structures of 29 possible base pairs that involve at least two hydrogen bonds are given in Figures 1–5 (for further descriptions, see Saenger, in Principles of nucleic acid structure, p. 120. Springer-Verlag [1984]). A base pair that is not a Watson-Crick pair or a G·U wobble pair is called a base-base mismatch, or an internal loop of two nucleotides. All the base pairs can be divided into two classes: normal and flipped. The normal class is defined by the arrangement of the Watson-Crick base pairs. The hydrogen bonding occurs for nucleotides with antiparallel strands and anti orientation of the bases relative to the ribose rings. The 11 base pairs that can be made with this same arrangement of nucleotides are called normal; they are shown in Figures 1 and 2. The remaining 18 base pairs require that one of the bases be flipped (inverted) by either reversing the direction of the strand or by switching the base from anti to syn. (Figs. 3–5). Normal base-base mismatches are found more often than flipped mismatches.
Figure 1 Five possible normal purine-pyrimidine base pairs. The Watson-Crick A·U, Watson-Crick G·C, and G·U wobble pairs fit into a double helix with very little distortion; A·U and A·C reverse Hoogsteen are mismatches. The plus and minus signs represent the direction of the strands (antiparallel) for anti nucleotides. The same orientation of each base can be obtained by reversing the direction of the strand and rotating the base around the glycosidic bond to syn.
Figure 2 Six possible normal purine-purine and pyrimidine-pyrimidine base pairs. The plus and minus signs represent the direction of the strands for anti nucleotides. The significance of the pluses and minuses is clear from the fact that no rotation of the bases in the plane of the figure can superimpose, for example, a $+\text{A}$ base on a $-\text{A}$ base. To superimpose two bases, either one strand must be reversed or the base must be changed from anti to syn. Sheared and imino G-A mismatches are found often in RNA structures.
Figure 3  Six possible flipped purine-pyrimidine mismatches. Note that all the nucleotides are labeled +. Each base pair could as well have been rotated 180° around an axis in the plane of the figure and labeled −. These base pairs can be formed from parallel strands with anti bases, or from antiparallel strands with one base changed to syn. Syn bases are higher energy conformations than anti bases.
Figure 4  Seven possible flipped purine-purine mismatches. Note that all the nucleotides are labeled +. Each base pair could as well have been rotated 180° around an axis in the plane of the figure and labeled \( - \). These base pairs can be formed from parallel strands with \( \text{anti} \) bases, or from antiparallel strands with one of the purine bases changed to \( \text{syn} \). \( \text{Syn} \) purines are higher energy than \( \text{anti} \) purines, but they have been found in several RNA molecules.
Five possible flipped pyrimidine-pyrimidine mismatches. Note that all the nucleotides are labeled $-$. Each base pair could as well have been rotated 180° around an axis in the plane of the figure and labeled $+$. These base pairs can be formed from parallel strands with *anti* bases, or from antiparallel strands with one of the pyrimidine bases changed to *syn*. *Syn* pyrimidines are high-energy conformations that have been very rarely identified, but they do occur.