Nonstandard Hydrogen Bonding in Duplex Oligonucleotides. The Base Pair between an Acceptor–Donor–Donor Pyrimidine Analog and a Donor–Acceptor–Acceptor Purine Analog

Johannes J. Voegel and Steven A. Benner*

Department of Chemistry, ETH CH-8092 Zürich, Switzerland

Received March 17, 1994

In its most general form, the Watson–Crick base pair joins a six-membered heterocyclic ring (in natural oligonucleotides, a pyrimidine) with a five–six fused ring system (in natural oligonucleotides, a purine) via three hydrogen bonds, one that joins the two central ring nitrogens of the paired heterocycles, and two that join flanking exocyclic functional groups (Figure 1). To hold the pair together, hydrogen bond donors in one heterocycle must face hydrogen bond acceptors in the other. With three hydrogen bonds, eight (2^3) hydrogen-bonding patterns and 16 independently replicatable bases are conceivable within the Watson–Crick geometry. Six hydrogen-bonding patterns, or 12 independently replicatable bases, are readily accessible using amino and carbonyl functionality (Figure 1).[1,2]

Pyrimidine analogs presenting acceptor–donor–donor and donor–donor–acceptor hydrogen-bonding patterns have proven to be the most difficult to obtain.[1g,3] First, to be aromatic and therefore able to stack, the ring system must be joined to the sugar by a carbon–carbon bond (a "C-nucleoside"). Further, the 6-aminopyrid-2-one structure, which formally presents the correct hydrogen bonding pattern, is evidently readily oxidized.[4] Adding a ring nitrogen to yield the aminopyrimidone pseudocytidine decreases susceptibility to oxidation, but creates an unacceptable tautomeric ambiguity.[5]

We report here that the 6-aminopyrazin-2-one ring system (trivially designated V) as the aglycon in an oligoribonucleotide forms a base pair as an acceptor–donor–donor pyrimidine analog with the donor–acceptor–acceptor purine analog 2′-deoxy-5-aza-7-deazaisoguanosine (trivially designated J) in an RNA–DNA duplex.

The 2′-deoxyriboside of the donor–acceptor–acceptor purine analog[6] was protected and converted into a phosphoramidite suitable for incorporation into a DNA strand by automated synthesis (Figure 2).[7] The riboside 3′,5′-bis(phosphate) (pVp)

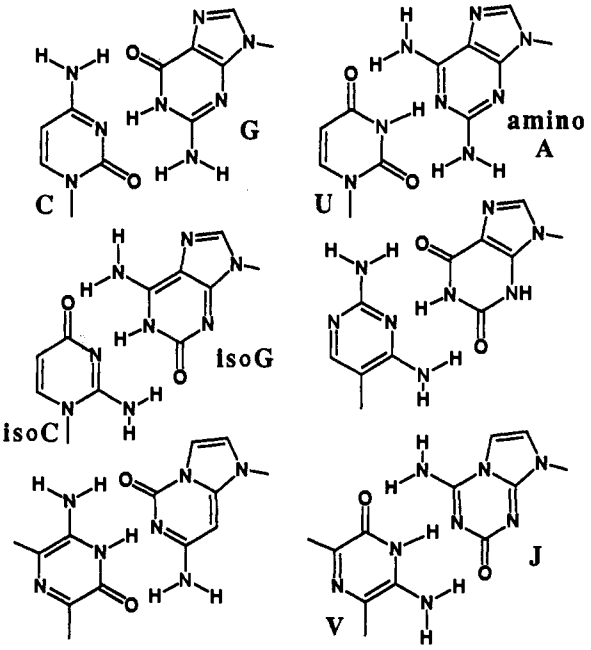

Figure 1. Nucleotide base pairs joined by various combinations of hydrogen bond donor and acceptor groups.

of the acceptor–donor–donor pyrimidine analog was prepared from a known precursor[3] (Figure 2). An oligoribonucleotide containing V was built by sequential addition of pVp and pAp to a starting oligoribonucleotide[8] using T4 RNA ligase and alkaline phosphatase[9] following the procedure of Middleton et al.[10] (Figure 2). All oligonucleotides were purified by HPLC,[11] and duplexes containing central dGC, dAU, and dJV matches were reanalyzed by HPLC after melting curves were obtained.[12]

A set of nucleosides that includes dJ and V maintains a complete Watson–Crick base pairing selectivity (Table 1). The reference oligonucleotide containing dGC, dAU, and dJV matches have melting temperatures of 36, 28, and 27 °C, respectively. Incorporating dGU and dAV mismatches lowers the melting temperatures to 17 °C. The first presumably exists as a less favorable wobble base pair, while the second presumably has a Watson–Crick geometry with unfavorable interactions between the exocyclic amino group of V and H-6 of adenine. The dJU pair, presumably joined by two hydrogen bonds but having additional nonbonding repulsion between opposite carbonyl groups, is still weaker (T_m = 13 °C). The dJC and dGV base pairs were too unstable to measure. This specificity implies that, in the expanded basis set, the dJV base pair is joined by three hydrogen bonds.

(1) (a) Benner, S. A.; Allemann, R. K.; Ellington, A. D.; Ge, L.; Glasfeld, A.; Leanz, G. F.; Krauch, T.; MacPherson, L. J.; Moroney, S. E.; Piccirilli, J. A.; Weinhold, E. *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *52*, 53–63. (b) Switzer, C. R.; Moroney, S. E.; Benner S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322–8323. (c) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner S. A. *Nature* **1990**, *343*, 33–37. (d) Piccirilli, J. A.; Moroney, S. E.; Benner S. A. *Biochemistry* **1991**, *30*, 10350–10356. (e) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537–539. (f) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489. (g) Voegel, J. J.; von Krosigk, U.; Benner, S. A. *J. Org. Chem.* **1993**, *58*, 7542–7547.

(2) (a) Rich, A. In *Horizons in Biochemistry;* Kasha, M., Pullman, B., Eds.; Academic Press: New York, 1962; pp 103–126. (b) Zubay, G. In *The Roots of Modern Biochemistry*; Kleinkauf, H., von Doehren, H., Jaenicke, L., Eds.; Walter de Gruyter & Co.: Berlin, 1988; pp 911–916. (c) Strazewski, P.; Tamm, C. *Angew. Chem.* **1990**, *102*, 37–59.

(3) von Krosigk, U. Diss. ETH Nr. 10164, Zürich, 1993.

(4) Seide, O. A.; Titow, A. I. *Chem. Ber.* **1936**, *69*, 1884–1893.

(5) (a) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois J. E. *J. Am. Chem. Soc.* **1976**, *98*, 6338. (b) Katritzky, A. R.; Waring, A. J. *J. Chem. Soc.* **1963**, 3046.

(6) Voegel, J. J.; Altorfer, M. M.; Benner, S. A. *Helv. Chim. Acta* **1993**, *76*, 2061-2069.

(7) The oligonucleotides were synthesized on a "Gene Assembler Plus" of Pharmacia. Except for the coupling step involving dJ, where the reaction time was doubled from 1.5 to 3.0 min, the standard parameters were used. Due to the instability of dJ in concentrated NH_3 (aqueous) at elevated temperatures, "PAC-amidites" of the natural bases were used; these were deprotected under milder conditions (NH_3 (aqueous)/EtOH, 3/1, 3 h at room temperature followed by 1 h at 55 °C).

(8) An RP-C18-HPLC purified *O*-2′-Fpmp- and *O*-5′-DMT-protected 4-mer was purchased from MWG-Biotech, deprotected according to the manufacturers instructions, and isolated by RP-C18-HPLC.

(9) (a) England, T. E.; Uhlenbeck, O. C. *Biochemistry* **1978**, *17*, 2069. (b) Barrio, J. R.; Barrio, M. G.; Leonard, N. J.; England, T. E.; Uhlenbeck, O. C. *Biochemistry* **1978**, *17*, 2077. (c) Romaniuk, P. J.; Uhlenbeck, O. C. *Methods Enzymol.* **1983**, *100*, 52.

(10) Middleton, T.; Herlihy, W. C.; Schimmel, P. R.; Munro, H. N. *Anal. Biochem.* **1985**, *144*, 110.

(11) Both the incorporation of dJ and V into the oligonucleotides and the relative nucleoside composition were verified by digestion of oligonucleotide samples and HPLC analysis of the resulting nucleoside mixtures (Seela, F.; Lampe, S. *Helv. Chim. Acta* **1991**, *74*, 1790–1800). The identity of the peak assigned to the unnatural nucleosides was proven by coinjection with synthetic dJ and V. After each coupling cycle, the products were purified by HPLC to yield product containing V only in the β configuration. The V present in the final oligonucleotide was >90% in the β configuration.

(12) Following the first melting experiment, V was present as a nearly equilibrated mixture of α and β isomers. A second cycle of melting/annealing showed the same melting temperature but proportionately lower hyperchromicity. This suggests that the oligonucleotide containing V in the α configuration does not hybridize.

0002-7863/94/1516-6929$04.50/0

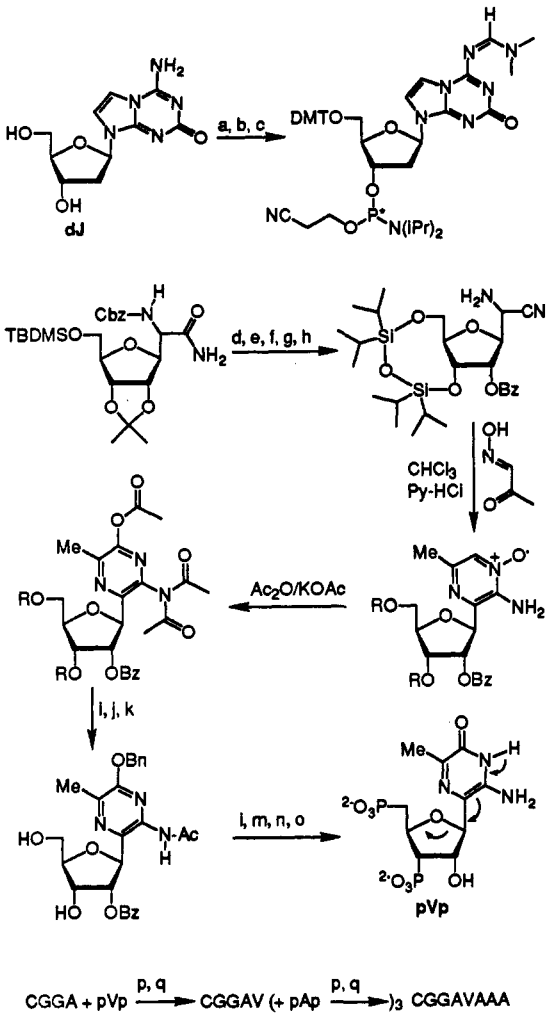

Figure 2. Synthesis of a dJ phosphoramidite, pVp, and an 8-mer RNA oligonucleotide containing V using T4 RNA ligase and alkaline phosphatase. (a) DMF–diethylacetal, DMF; (b) DMTCl, pyridine, DMF; (c) $ClP(N(iPr)_2)OCH_2CH_2CN$, Hünig's base, CH_2Cl_2; (d) TFA/H_2O, 4/1; (e) TFAA/pyridine/dioxane; (f) TipsCl, imidazole, DMF; (g) BzCl, pyridine; (h) Pd–C/H_2/MeOH; (i) EtOH, reflux; (j) $PhCH_2OH$, DEAD, Ph_3P, THF; (k) pyridine–HF, pyridine, 0 °C; (l) $(BnO)_2PN(iPr)_2$, tetrazole, CH_3CN, DMF; (m) N-methylmorpholine *N*-oxide, CH_3CN, DMF; (n) Pd–C/H_2, THF, Et_3NHCO_3 aqueous; (o) $H_2NNH_2 \times H_2O$; (p) T4 RNA ligase; (q) alkaline phosphatase. Arrows in final structure show proposed mechanism for epimerization.

These behaviors are the minimal that are required for a coding system. A detailed analysis of the behavior of the expanded basis set raises, however, general questions about the structural origin of duplex stability in DNA–RNA hybrids. In particular, the duplex containing the nonstandard dJV base pair is less stable (T_m = 27 °C) than the GC base pair (T_m = 36 °C), even though both are presumably joined by three hydrogen bonds. Together with recent results with DNA duplexes containing 1-methylpseudouridine,[13] the enzymatic synthesis of DNA duplexes containing this C-glycoside,[1d] and experimental melting temperatures with C-glycosides bearing 2,6-diaminopyrimidine as aglycon,[1c] it is possible to suggest that C-nucleoside analogs of pyrimidines will generally be found to form weaker base pairs than the corresponding N-nucleosides. As context and other effects also influence the contribution of a base pair to duplex stability, more data must be collected with more analogs embedded in more sequences to test this generalization.

(13) Rosenberg, I.; Soler, J. F.; Tocik, Z.; Ren, W.-Y.; Ciszewski, L. A.; Kois, P.; Pankiewicz, K. W.; Spassova, M.; Watanabe, K. A. *Nucleosides Nucleotides* **1993**, *12*, 381.

Table 1. Melting Temperatures for Duplexes Formed with Ribo-GAACXAAA and Deoxyribo-TTTYGTTC[a]

	T_m		
Y	X = C	X = U	X = V
dG	36	17	<5
dA	5	28	17
dJ	<5	13	27

[a] Values in °C. In aqueous buffer containing 10 mM NaH_2PO_4 (pH 7.0), 1 M NaCl, 0.1 mM EDTA. Oligonucleotide concentrations ca. 1.3 μM. Heating rate 1 °C/min. Ultraviolet absorbance measured at 254 nm. Duplicate experiments (except for the dAV mismatch) yielded identical T_m values to within ±1 °C.

Slowly at room temperature and more rapidly at high temperature, the oligonucleotide containing V undergoes a clean reaction to yield a second oligonucleotide product. This was assigned as the oligonucleotide containing the α-V-riboside, by analogy with the epimerization observed with the free riboside.[14] HPLC analysis provided an equilibrium constant (66 ± 5% to 34 ± 5%, β:α). At 50.5 ± 0.5 °C, the pseudo-first-order rate constant for this epimerization (α to β) was 0.070 min^{-1} ($t_{1/2(Pseudo)}$ = 10 min; 0.1 M Et_3NH^+ ^-OAc, pH 7.1, containing 10% MeCN). The same equilibrium ratio of products was obtained starting from oligonucleotides containing either anomer.

While epimerization does not preclude laboratory experiments such as those described here, it is clearly undesirable in a DNA molecule intended for the storage of genetic information. It presumably occurs by the mechanism shown in Figure 2, analogous to that for the slower epimerization of pseudouridine, which is found in natural oligonucleotides in noncoding roles. Thus, such epimerization might be expected with any nucleoside where electrons from a ring N–H bond can flow via a π system to break the C1–O4 bond in the sugar. It is difficult to imagine an acceptor–donor–donor pyrimidine analog where this electron flow would not be possible. This may explain why this hydrogen-bonding scheme is not found in oligonucleotides from organisms that reside on planet earth and is not expected in any DNA-like molecules that might be found extraterrestrially. It is worth noting, however, that analogs with the sugar ring oxygen replaced by a methylene group would not suffer this epimerization.

Acknowledgment. The authors are indebted to Ms. B. Brandenberg, Mr. M. Sperl, and Dr. B. Jaun for NMR measurements, to R. Häfliger, O. Greter, and Dr. W. Amrein for MS data, and to Dr. E. Zass for literature searches. J.J.V. thanks Dr. Rosenkranz for a scholarship. Support from the Swiss National Science Foundation and Sandoz AG is gratefully acknowledged.

(14) For the free ribonucleoside V, an acid- and base-catalyzed isomerization leads to two furanosyl and two pyranosyl anomers analogous to the one observed with pseudouridine (Cohn, W. E. *J. Biol. Chem.* **1960**, *235*, 1488. Chambers, R. W. *Prog. Nucleic Acid Res. Mol. Biol.* **1966**, *5*, 349), although at significantly faster rates. The two furanosyl epimers are expected without further degradation if the 5′-OH functionality is blocked, as in the 5′-phosphate derivative (Chambers, R. W.; Kurkov, V.; Shapiro, R. *Biochemistry* **1963**, *2*, 1192) or the oligonucleotide described here.