Pergamon

Tetrahedron Letters, Vol. 35, No. 19, pp. 3045-3048, 1994
Elsevier Science Ltd
Printed in Great Britain
0040-4039/94 $6.00+0.00

0040-4039(94)E0510-5

Guanosine Derivatives Bearing an N^2-3-Imidazolepropionic Acid

Norbert V. Heeb and Steven A. Benner*
Department of Chemistry, E.T.H.
CH-8092 Zürich, Switzerland

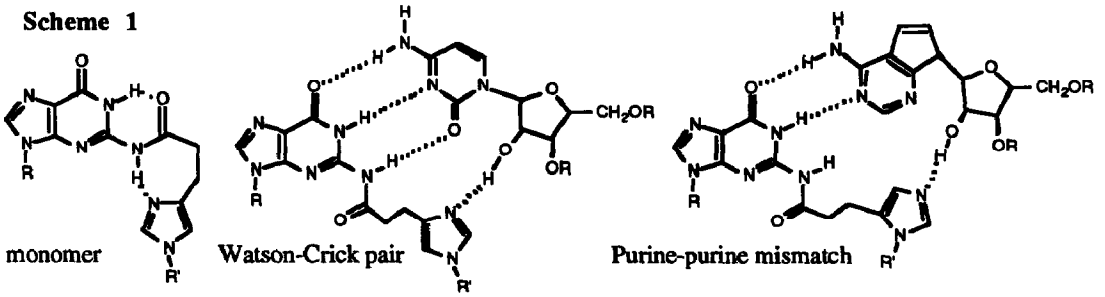
Abstract: Synthesis of a 2'-deoxyguanosine analog tethered through the exocyclic nitrogen via a 3 carbon chain to the 4-position of an imidazole is described. The imidazole forms a hydrogen bond with the 2'-hydroxyl group of a complementary cytosine bound as a Watson-Crick base pair.

Considerable attention has been directed recently towards oligonucleotides bearing functional groups.[1-4] These are interesting because they may catalyze the hydrolysis of complementary oligonucleotides, yielding RNA-targeted "antisense" and DNA-targeted cleaving agents.[5,6] Further, such oligonucleotides might catalyze transesterification reactions at phosphate centers generally,[7] shedding light on the plausibility of the hypothesis that the first form of life on planet earth was an oligonucleotide that catalyzed the template-directed polymerization of oligonucleotides.[8-11] Much experimental work is now being directed towards obtaining such catalytic oligonucleotides in the laboratory.[12]

The enzyme ribonuclease (RNase) provides a well-understood example of how one might catalyze both transesterification reactions at phosphorus and the hydrolysis of RNA.[13] In the RNase mechanism, an imidazole from His 12 acts as a general base abstracting a proton from a ribose 2'-oxygen, which then attacks the 3'-phosphate to form a 2',3'-cyclic phosphate, cleaving the chain. The imidazolium then serves as a general acid to protonate the 2'-oxygen of the cyclic phosphate in a hydrolytic reaction that yields a 3'-phosphate product. Site-directed mutagenesis studies in these laboratories have shown that functional groups from other amino acids, in particular His 119, Lys 41, and Asp 121, are also critical for achieving catalysis.[14] Nevertheless, it is appropriate to ask whether a complementary oligonucleotide can mimic the interaction provided by His 12 in RNase, delivering an imidazole group to an oligoribonucleotide so that it forms a hydrogen bond to the 2'-hydroxyl group.

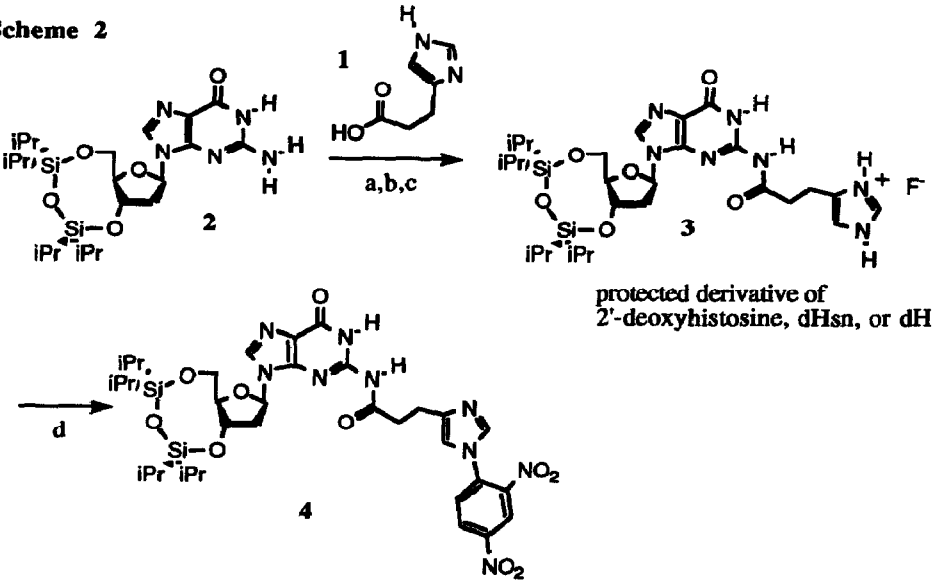
Molecular modelling suggested that an imidazole attached by its 4-position via a 3-atom spacer to the 2-amino group of guanosine should be able to form a hydrogen bond to the 2'-hydroxyl group of a cytosine derivative forming a Watson-Crick base pair with the guanosine (Scheme 1). Attachment of the imidazole at its 4-position (as opposed to the 1 or 2 positions) allows the imidazole to deliver two geometrically distinct nitrogen bases to the vicinity of the 2'-hydroxyl group. This increases the likelihood that the appended imidazole will also form a hydrogen bond to the 2'-hydroxyl group in a purine-purine mismatch (see below).

Scheme 1



To prepare such a derivative, 3-imidazol-4-yl-propionic acid **1**, prepared by reduction of urocanic acid, was treated with carbonyldiimidazole in pyridine to yield an activated intermediate (presumably a cyclized derivative), which then reacted with 3',5'-O-(1,1,3,3-tetraisopropyldisilyloxane-1,3-diyl)-2'-deoxyguanosine **2** to yield the amide **3**, whose sugar-base combination is assigned the trivial name "deoxyhistosine" (dHsn or dH), isolated as the hydrofluoride salt (Scheme 2). The pK_a values of **3** were measured to be 7.10 (imidazolium) and 10.08 (N^1H). Treatment of **3** with dinitrofluorobenzene yielded protected derivative **4**.

Scheme 2



(a) 1,1-carbonyldiimidazole, pyridine (b) pyridine, 110°C, 96 h (c) pyridinium fluoride, pyridine-water (d) dinitrofluorobenzene, pyridine.

In $CDCl_3$ at high dilution, the N^2H and N^1H protons of **4** gave signals at 11.28 and 12.05 ppm respectively. The corresponding signals are found at ca. 10.2 and 12.0 ppm in N^2-isobutyrylguanosine derivatives. This suggests that the carbonyl oxygen of the side chain of **4** forms a hydrogen bond with the N^1H (as in isobutyryl-G), while the imidazole nitrogen forms a second hydrogen bond with N^2H (Scheme 1, R' = 2,4-dinitrophenyl). The chemical shift of the N^2H signal varied upon changes in the concentration of **4**, yielding a self-association constant for **4** of 3.5 M^{-1}. The chemical shift of the N^1H is independent of the concentration of **4** up to its solubility limit, suggesting that the self-association of **4** does not occur by intermolecular hydrogen bonding involving N^1H.

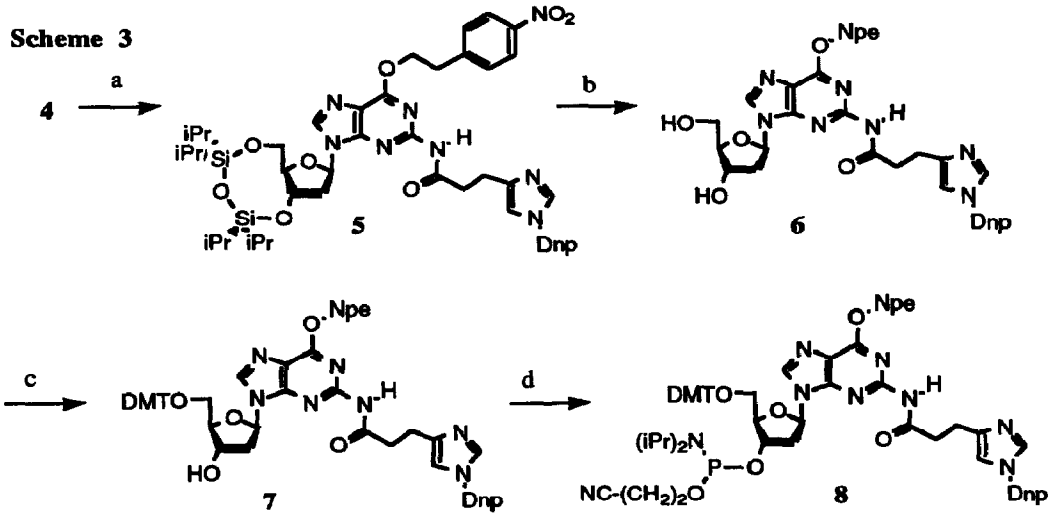
Upon addition to **4** of an equimolar amount of 3',5'-O-(1,1,3,3-tetraisopropyldisilyloxane-1,3-diyl)-2'-deoxycytosine, a complex is formed. Both the N^2H and N^1H signals of **4** and the N^4H signal of the deoxycytosine derivative are shifted downfield in the complex, suggesting that it is held together by intermolecular hydrogen bonding involving N^1H and N^2H of dHsn and N^4H of the dC derivative. The probable structure for the complex is therefore a Watson-Crick base pair (Scheme 1). The complex was then analyzed by ^{13}C nmr spectroscopy using the method of Iwahashi and Kyogoku.[15] Large downfield shifts in the signals arising from C-6 of dHsn and C-2 of the dC derivative were again consistent with the formation of a Watson-Crick base pair between dHsn and dC in $CDCl_3$. From the concentration dependence of the shifts, an association constant of 160 M^{-1} was calculated. The corresponding association constant between 1-cyclohexylthymine and 9-ethyladenine in $CDCl_3$ is 50-70 M^{-1}.[15] This is consistent with the model where the

dHsn-dC pair is joined by three hydrogen bonds, while the T-A pair is joined by only two. The association constant for the pairing of dHsn and dC is, however, smaller than that estimated for the pairing of derivatives of dG and dC,[16] suggesting that the disruption of the intramolecular hydrogen bond with N^1H needed to form a Watson-Crick pair has a significant energetic cost.

These studies were then repeated using 3',5'-O-(1,1,3,3-tetraisopropyldisilyloxane-1,3-diyl)-*ribo*cytosine, which carries a 2'-hydroxyl group on the ribose sugar (Scheme 1, R = -Si(iPr)$_2$-O-Si(iPr)$_2$-; R' = 2,4-dinitrophenyl). The imidazole group carried by **4** is expected to form a hydrogen bond to this 2'-hydroxyl group in the Watson-Crick base pair. Consistent with this model, the association constant was found to be 630 M^{-1}, four fold higher than that observed for the pairing of dHsn and dC derivatives. This suggests that the additional hydrogen bond between the imidazole of dHsn and the 2'-hydroxyl group of the derivative of cytosine stabilizes the complex by approximately 0.8 kcal/mol in $CDCl_3$.

The oxygen on dHsn was then protected as the 4-nitrophenylethyl ether and the 5'-hydroxyl group protected as the dimethoxytrityl ether. Short oligonucleotides containing dHsn (HpT, CpH, CpHp, and CpHpT) were prepared in solution by standard methods.[17] Their conformation in aqueous solution was studied by nmr. The C2'-endo conformation was preferred for all deoxynucleotides in these short oligomers, with slightly increasing preference for the C2'-endo conformation in nucleotide units towards the 5'-end. NOE measurements showed a preference for the syn conformation around the glycosidic bond for dHsn.[18] This follows closely the conformational preferences observed with guanosine in short oligonucleotides.[19]

Scheme 3



(a) 2-(*p*-nitrophenyl)-ethanol, triphenylphosphine, diethyl azodicarboxylate (b) Bu_4N^+ F^-, THF, HOAc, pyridine (c) 4,4'-dimethoxytritylchloromethane, Hünig's base, pyridine (d) $Cl(iPr)_2NPOCH_2CH_2CN$, Hünig's base. DMT = dimethoxytrityl; Dpn = 2,4-dinitrophenyl; Npe = 2-(p-nitrophenyl)ethyl.

The corresponding phosphoramidite **8** was then prepared (Scheme 3) and used in a solid phase synthesis of the hexadecanucleotide dT_7HT_8. Mixtures of this hexadecanucleotide with the complementary dA_8CA_7 showed the expected hypochromicity at 260 nm, with the largest effect observed at a 1:1 ratio of the two strands. This suggested that the dT_7HT_8- dA_8CA_7 pair formed a typical DNA duplex. Consistent with the results obtained with monomeric nucleoside derivatives in $CDCl_3$, the dHsn-dC base pair contributed less to overall duplex stability than a dG-dC base pair at the analogous position. At 2 μM concentrations of each

oligonucleotide (in 100 mM KCl in potassium phosphate buffer, 100 mM, pH 6.0), the melting temperature of the duplex between dT_7HT_8 and dA_8CA_7 was 37°C, while the melting temperature of the duplex between dT_7GT_8 and dA_8CA_7 was 43°C. At 1 molar salt, the melting temperatures were 43°C and 51°C respectively.

These oligonucleotides are, of course, fully complementary in the Watson-Crick sense, and are likely to form a standard double helical structure. It is important to note, however, that for an imidazole appended to the 2-position amine of a purine to initiate the cleavage of a bound RNA strand, there must be some distortion of the standard double helix, because the 2'-oxygen is not within bonding distance of the 3'-phosphate in this structure. Models suggested that one way to obtain this distortion might be via a purine-purine mismatch (Scheme 1). Thus, an equimolar mixture of dT_7HT_8 and $riboA_{12-18}$ was prepared. This mixture is expected to form a duplex joined by 15 T-A base pairs flanking a distorted H-A base pair, and the expected hypochromicity at 260 nm was observed upon mixing the two oligonucleotides in a 1:1 mixture. No significant amounts of cleavage of the RNA were detected in this complex, possibly because functional groups analogous to other residues in the RNase active site (His 119, Asp 121, and Lys 41) were not present.[14] However, the melting of the duplex between dT_7HT_8 and $riboA_{12-18}$ (35 °C) was *higher* than that between dT_7GT_8 and $riboA_{12-18}$. The increased stability of duplexes with oligo*ribo*nucleotides with oligonucleotides containing Hsn, and the decreased stability of oligo*deoxy*ribonucleotides with oligonucleotides containing G, is again consistent with a model where the imidazole group appended to dHsn reaches across the minor groove of a duplex structure to form a hydrogen bond with a 2'-hydroxyl group on the opposite strand.

These results show for the monomeric species in $CDCl_3$, and suggest for oligomeric species in water, that an imidazole appended to N^2 of guanosine can form a hydrogen bond with the 2'-hydroxyl group of a ribonucleoside on a complementary strand. Work to incorporate other elements of the RNase A active site into an oligonucleotide is in progress.

1. Bashkin, J. K.; McBeath, R. J.; Modak, A. S.; Sample, K. R.; Wise, W. B. *J. Org. Chem.* **1991**, *56*, 3168-3176.
2. Wang, G; Bergstrom, D. E. *Tetrahedron Lett.* **1993**, *34*, 6725-6728.
3 Moser, H. E.; Dervan, P. B. *Science*, **1987**, *238*, 645-650.
4 Sigman, D. S.; Chen, C. B. *Ann. Rev. Biochem.* **1990**, *59*, 207-236.
5. Melton, D., editor, *Antisense RNA and DNA*, **1988**, Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.
6. Huang, Z.; Schneider, K. C.; Benner, S. A. *J. Org. Chem.* **1991**, *56*, 3869-3882.
7 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 Sym. Quant. Biol.* **1987**, *52*, 53-63.
8. Rich, A. in *Horizons in Biochemistry*, **1962**, Eds. M. Kasha and B. Pullman, Academic Press New York, pp 103-126.
9. Woese, C. R. *Origins of the Genetic Code*, **1967**, N.Y., Harper & Row.
10. Orgel, L. E. *J. Mol. Biol.* **1968**, *38*, 381-393.
11. Crick, F. H. *J. Mol. Biol.* **1968**, *38*, 367-379.
12. Bartel, D. P.; Szostak, J. W. *Science* **1993**, *261*, 1411-1418.
13. Blackburn, P.; Moore, S. *The Enzymes*, **1982**, *15*, 317-434.
14. Trautwein, K.; Benner, S. A. *FEBS Lett.* **1991**, *281*, 275-277.
15. Iwahashi, H.; Kyogoku, Y. *J. Am. Chem. Soc.* **1977**, *99*, 7761-7765.
16. The comparison cannot be exact, as the association constant for G and C in $CDCl_3$ was estimated by infra-red spectroscopy. Kyogoku, Y.; Lord, R.C.; Rich, A. *Biochim. Biopys. Acta* **1969**, *179*, 10-17.
17. Pfleiderer, W.; Uhlmann, E. *Helv. Chim. Acta* **1981**, *64*, 1688-1703. Silber, G.; Flockerzi, D.; Varma, R. S.; Uhlmann, E.; Pfleiderer, W.; Charubala, R. *Helv. Chim. Acta* **1981**, *64*, 1704-1716. Pfleiderer, W.; Charubala, R., Trichtinger, T.; Schulz, B. S.; Himmelsbach, F. *Tetrahedron* **1984**, *40*, 59-72.
18. Wüthrich, K. *NMR of Proteins and Nucleic Acids*, 1st ed., **1986**, Wiley, New York, p 208.
19. Cheng, D.M.; Sarma, R.H. *J. Am. Chem. Soc* **1977**, *99*, 7333-7348.

(*Received in Germany* 4 *February* 1994; *accepted* 8 *March* 1994)