Hydrogen bonding in DNA—a return to the *status quo*

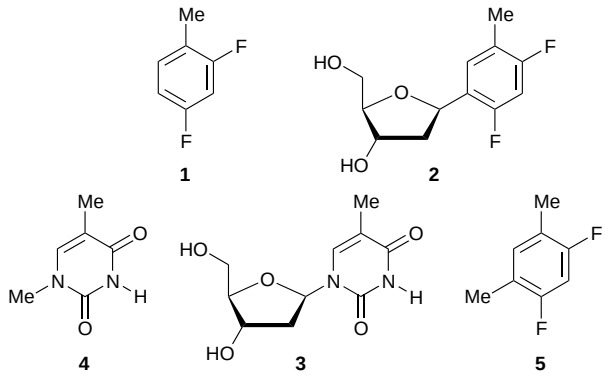
Timothy A. Evans and Kenneth R. Seddon*

School of Chemistry, The Queen's University of Belfast, Stranmillis Road, Belfast, Northern Ireland, UK BT9 5AG

Recent claims that 'conventional hydrogen bonds may not be necessary for high efficiency and fidelity in DNA synthesis' and that 'DNA polymerase can exert high fidelity even when a base pair completely lacks conventional hydrogen bonds' are shown to be in error.

Both Watson and Crick's original *Nature* paper,[1] and Watson's subsequent account of the discovery of the structure of DNA,[2] focus upon the central role of hydrogen bonds between base pairs in defining the formation of the double helix—a view which has become a paradigm, and has remained effectively unchallenged until the appearance this year of a communication (preannounced by Bradley)[3] from Kool,[4] in which the rôle of the hydrogen bond is questioned. Since such doubts as to the importance of hydrogen bonding to molecular recognition undermine not only many of the basic theories of biochemistry, but also the foundations of supramolecular chemistry[5] and crystal engineering,[6–8] it is particularly important that the basis for these radical criticisms be established beyond reasonable doubt. We describe here the reasons why we believe, despite its synthetic elegance, the work of Moran *et al.*[4] is flawed.

At the heart of the communication[4] is the contention that 2,4-difluorotoluene **1** is nonpolar, and incapable of forming hydrogen bonds: if this tenet of their approach is incorrect, the basis for their argument is fallacious. We demonstrate here that they have misunderstood the hydrogen bonding capabilities of such a molecule, that their basic experimental approach is thus flawed, and that there is no reason to change our fundamental understanding of the rôle of hydrogen bonding in the formation of the DNA double helix.



In the title of the paper ('Difluorotoluene, a Nonpolar Isostere for Thymine, Codes Specifically and Efficiently for Adenine in DNA Replication') and in the text it is stated that **1** is nonpolar.[4] However, AM1 calculations† give a dipole moment of 1.86 D (*cf.* AM1 calculations on 1,3-difluorobenzene give a dipole moment of 1.55 D; the experimental value is 1.58 D[9]), a value greater than that for water (1.82 D).[10] Thus, **1** is demonstrably polar, as is the derivative nucleoside **2** which is specifically mentioned[4] as 'a nonpolar shape mimic for natural thymidine' **3**. Furthermore, it is stated[4] that **2** 'would serve as a good test for the importance of thymidine's hydrogen bonding groups on fidelity, because [it] lacks the strongly localized charges but retains nearly the exact steric shape of the natural molecule'. Fig. 1 illustrates the charge distributions of 1-methylthymine **4**, an analogue of **3**, and **5**, an analogue of **2**, as obtained by *ab initio* calculations at the 6-31G** level (essentially similar results were produced by AM1). The pattern and relative magnitude of the localized charges in both molecules are essentially identical, and thus **1** or **2** should not be used for the test described.

The authors state[4] that **2** has a 'non-polar, non-hydrogen-bonding nature', and that their work provides evidence 'that a DNA polymerase can exert high fidelity even when a base pair completely lacks hydrogen bonds'. These statements apparently originate with their belief that the C–F··H–N hydrogen bond is 'unconventional' [*sic*],[4] and that C–H··N hydrogen bonds are not possible upon the observation that 'difluorotoluene is a highly hydrophobic species which does not pair with adenine even in chloroform'.[4] This latter statement is based upon an inconclusive experiment (9-ethyladenine has a very limited solubility in $CHCl_3$), in with 9-ethyladenine was titrated against a dilute (millimolar) solution of **2** in $CDCl_3$, and the N[4]–*H* resonance was monitored on the 9-ethyladenine:[11] the C–*H* resonances of the difluorotoluene do not appear to have been

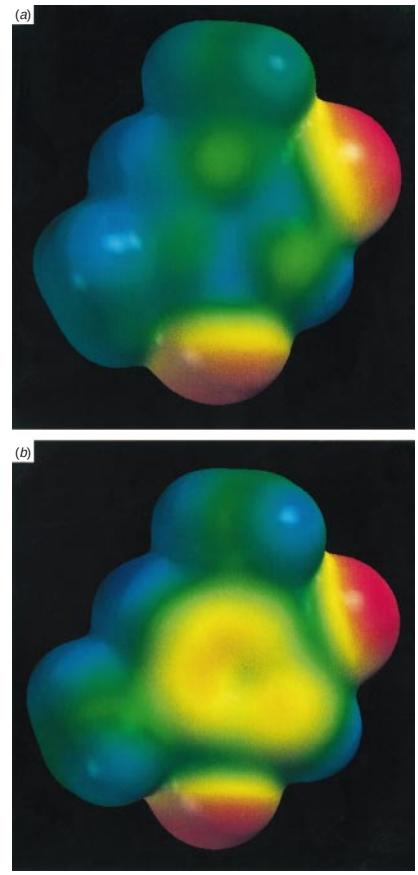


Fig. 1 The total electron density isosurfaces for (*a*) **4** and (*b*) **5**, each being encoded with the value of its electrostatic potential (red = most negative, blue = most positive), calculated at the 6–31G** level. Both show essentially the same relative charge distribution, although the absolute values are smaller for **4** than for **5**. Calculations on **2** and **3** also reveal similar relative charge distributions.

examined. In our hands, even the simple bases pyridine (pK_a 5.18)[12] and 3-chloropyridine (pK_a 2.84),[13] which are respectively stronger and weaker bases than adenine (pK_a 4.20),[14] are capable of forming hydrogen bonds with **1**. Both pyridine and 3-chloropyridine are hydrogen-bond acceptors and, in both cases, the C–*H*^{3} proton of **1** exhibits concentration dependent downfield shifts (see Fig. 2), both with respect to an internal reference ($SiMe_4$), and with respect to the methyl protons of **1**. Indeed, the AA′MXY spin system is converted, in the hydrogen-bonded base pair, into an ABMXY system. This is behaviour characteristic of hydrogen-bond formation in solution,[15–17] and provides hard evidence in direct contradiction of the recent claims.[4] Moreover, the phenomenon of C–H···X hydrogen bonds is well established,[8,18–22] and is no longer considered either controversial or 'unconventional'. In addition, in a study of hydrogen bonds formed by fluorinated hydrocarbons, Howard *et al.*[23] report (on the basis of exhaustive searching of the Cambridge Structural Database and high level *ab initio* calculations) that 'Encouraging for the bio-organic chemist is that substrate/protein interactions may offer an environment for optimal F···H bonding. [...] the F···H–X interaction may contribute to the overall binding energy, up to half of the strength of the original hydrogen bond to oxygen'. Our own gas-phase calculations (AM1, PM3 and 6-31G*) on the stabilization of base pairs between **4** or **5** and 9-methyladenine confirm this view, all resulting in the same fully-optimized co-planar structures **6** and **7**, respectively. Both these base pair structures have hydrogen-bonding motifs which encode identically as $N_1 = R_2^2(8)$ using Etter's topological analytical method for the comparison of hydrogen-bonding networks,[24,25] The calculated NH···F bond in **7** is only about 0.1 Å longer than the calculated NH···O bond in **6**, reflecting the smaller charges in **5** than in **4**, but representing significant hydrogen bonding. The CH···N bond in **7** is, as expected, significantly longer (0.4 Å) than the NH···N bond in **6**, but still represents a significant weak interaction. The difference in binding energies between the two structures represents a 60% decrease in substituting **5** for **4**. This energy difference is entirely in accord with the reported empirical measurements.[11,26,27] The existence of CH···N and NH···F hydrogen bonds in **7** explains the observed surprisingly high efficiency and fidelity in DNA synthesis: the choice of **2** as a replacement for **3** was inspired—it mimicked not only its shape, but also its relative charge distribution and hydrogen-bonding patterns.

In conclusion, the arguments presented elsewhere[4] are flawed: we have demonstrated here that 2,4-difluorotoluene is both polar and capable of acting as a hydrogen-bond donor and acceptor. It has been stated that 'conventional hydrogen bonds may not be necessary for high efficiency and fidelity in DNA synthesis' and that 'DNA polymerase can exert high fidelity even when a base pair completely lacks conventional hydrogen bonds'. These views contradict 'most if not all current models for replication fidelity'.[4] Using the well-established philosophical principle of Occam's Razor,[28] it is clear that these workers[4] have reported no results which disturb the *status quo*. Indeed, in

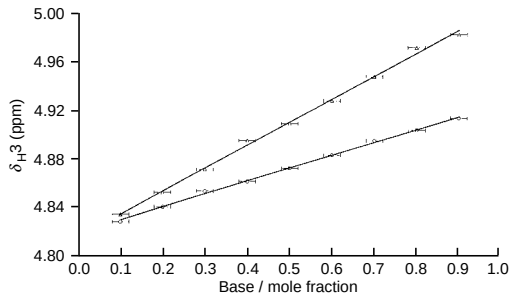
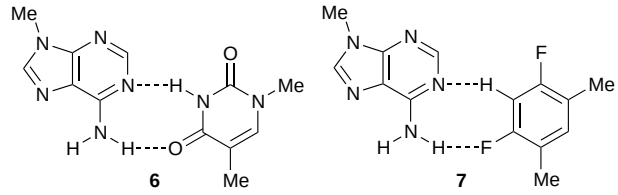


Fig. 2 The chemical shift of the H^3 proton of **1** measured with respect to the methyl resonance, and plotted as a function of the mole fraction of added base, for the addition of (Δ) pyridine and (○) 3-chloropyridine



view of the expected hydrogen bonding ability of **2**, they may well have underlined the validity of the conventional model.

We are indebted to the EPSRC and Royal Academy of Engineering for the award of a Clean Technology Fellowship (to K. R. S.), to the EPSRC for the award of a ROPA grant, and to Richard Murphy, Mark Nieuwenhuyzen, and Drs John Holbrey, Martyn Earle, Yasmin Patell and John Wilson for their assistance and advice. We would particularly like to acknowledge a suggestion from a referee which prompted our use of 3-chloropyridine.

Footnotes and References

* E-mail: k.seddon@qub.ac.uk

† All the calculations reported here were performed using the Spartan 4.1.1 suite of software. The results cited correspond to the Hamiltonians quoted in the text, and all results represent fully optimized geometries with no applied constraints. In all cases, calculations were performed at the AM1, PM3, STO-3G, 6-31G* and 6-31G** levels, and full details of these calculations will be reported in a future manuscript. Inclusion of an aqueous solvation sheath in $AM1_{aq}$ calculations does not significantly perturb the conclusions.

1 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737.
2 J. D. Watson, *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*, Atheneum, New York, 1968.
3 D. Bradley, *New Scientist*, 22nd February 1977, **153**, 19.
4 S. Moran, R. X.-F. Ren, S. Rumney IV and E. T. Kool, *J. Am. Chem. Soc.*, 1997, **119**, 2056.
5 J.-M. Lehn, *Supramolecular Chemistry: Concepts and Perspectives*, VCH, Weinheim, 1995.
6 G. R. Desiraju, *Crystal Engineering: The Design of Organic Solids*, Elsevier, Amsterdam, 1989.
7 *The Crystal as a Supramolecular Entity*, ed. G. R. Desiraju, Wiley, Chichester, 1996, vol. 2.
8 C. B. Askeröy and K. R. Seddon, *Chem. Soc. Rev.*, 1993, **22**, 397.
9 E. M. Moore and M. E. Hobbs, *J. Am. Chem. Soc.*, 1949, **71**, 411.
10 A. L. McClellan, *Tables of Experimental Dipole Moments*, W. H. Freeman and Company, San Francisco, 1963.
11 B. A. Schweitzer and E. T. Kool, *J. Am. Chem. Soc.*, 1995, **117**, 1863.
12 R. K. Murmann and F. Basolo, *J. Am. Chem. Soc.*, 1955, **77**, 3484.
13 H. C. Brown and D. H. McDaniel, *J. Am. Chem. Soc.*, 1955, **77**, 3752.
14 R. M. Izatt and J. J. Christensen, *J. Phys. Chem.*, 1962, **66**, 359.
15 M. Nakano, N. I. Nakano and T. Higuchi, *J. Phys. Chem.*, 1967, **71**, 3954.
16 J. W. Akitt, *N.M.R. and Chemistry: An Introduction to the Fourier Transform–Multinuclear Era*, Chapman and Hall, London, 2nd edn., 1983.
17 A. G. Avent, P. A. Chaloner, M. P. Day, K. R. Seddon and T. Welton, *J. Chem. Soc., Dalton Trans.*, 1994, 3405.
18 G. R. Desiraju, *J. Chem. Soc., Chem. Commun.*, 1989, 179.
19 G. R. Desiraju, *Acc. Chem. Res.*, 1991, **24**, 290.
20 G. R. Desiraju, *Mol. Cryst. Liq. Cryst.*, 1992, **211**, 63.
21 G. R. Desiraju, *Angew. Chem., Int. Ed. En.*, 1995, **34**, 2311.
22 C. B. Aakeröy and K. R. Seddon, *Z. Naturforsch., Teil B*, 1993, **48**, 1023.
23 J. A. K. Howard, V. J. Hoy, D. O'Hagan and G. T. Smith, *Tetrahedron*, 1996, **52**, 12613.
24 M. C. Etter, *Acc. Chem. Res.*, 1990, **23**, 120.
25 M. C. Etter, J. C. Macdonald and J. Bernstein, *Acta Crystallogr., Sect. B*, 1990, **46**, 256.
26 J. Pranata, S. G. Wierschke and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1991, **113**, 2810.
27 I. K. Yanson, A. B. Teplitsky and L. F. Sukhodub, *Biopolymers*, 1979, **18**, 1149.
28 B. Russell, *History of Western Philosophy*, George Allen & Unwin, London, 2nd edn., 1961.

Received in Cambridge, UK, 27th March 1997; Revised version received, 31st July 1997; 7/05239A