The difluorotoluene debate—a decade later

Eric T. Kool* and Herman O. Sintim

Received (in Cambridge, UK) 18th April 2006, Accepted 25th May 2006
First published as an Advance Article on the web 7th July 2006
DOI: 10.1039/b605414e

2,4-Difluorotoluene is unusual among hydrofluorocarbons because it is shaped like the DNA base thymine. It was first synthesised as a nucleotide analogue and incorporated into DNA a decade ago. Although it is a nonpolar molecule, it was found to be replicated by DNA polymerase enzymes as if it were thymine. We concluded that replication of DNA base pairs can occur without Watson–Crick hydrogen bonds, and hypothesised that steric effects, rather than these hydrogen bonds, were the main arbiters of DNA replication fidelity. A debate was initiated then, with claims by some that the molecule is polar and forms hydrogen bonds with adenine, thus supporting the hydrogen bonding theory of DNA replication. Here we discuss the evolution of this debate, and reflect on the relevant data that have since come from hundreds of papers and dozens of laboratories. Although discussion on this topic continues, the steric hypothesis for DNA replication is now widely accepted among biochemists, and the changing paradigm has been reflected in textbooks.

Introduction

In 1994 we introduced the concept of ''nonpolar nucleoside isosteres'', which are synthetic nucleoside analogues in which the DNA base is replaced by a less polar structure having a size and shape as close as possible to that of the natural congener.[1] These compounds used substituted benzenes as pyrimidine replacements, and indoles and benzimidazoles as purine replacements (Fig. 1 and 2). Because the polar carbonyl, imino, and amino groups that normally are involved in Watson–Crick hydrogen bonding are entirely missing (replaced by C–F, C–H and CH_3 groups respectively), the compounds were designed to disrupt this hydrogen bonding. The molecules were intended to be used as probes of the importance of the polar functional groups on the biochemistry and biophysics of DNA.

To date, more than ten different nonpolar nucleoside isosteres have come from our laboratory, and some related ones from other laboratories as well. Early on, it was pointed out that some of these compounds were imperfect isosteres, in part because C–H groups were used to replace sp^2-type nitrogens in adenine and cytosine, which adds perhaps 0.5–1.0 Å of steric bulk at N3 in these molecules. At the time we expected that this was a small perturbation, although we now believe that steric effects even on this scale can be large (see below). But as a result of this perturbation, we focused many of our early studies on one of the compounds, difluorotoluene (abbreviated F), which as an isostere is nearly perfect.[2] The deoxyriboside (dF) was prepared and its structure was

Department of Chemistry, Stanford University, Stanford, CA 94305, USA. E-mail: kool@stanford.edu

Eric T. Kool

Eric T. Kool received his PhD from Columbia University in 1988, and was a postdoctoral fellow at the California Institute of Technology before taking an independent position as Assistant Professor at the University of Rochester in 1990. In 1999 he moved to Stanford University, where he is Professor of Chemistry. His research interests include the study of DNA replication and repair, development of DNA-templated chemistry for imaging cellular RNAs, design of new genetic systems, and discovery of new DNA-based fluorophores and sensors.

Herman O. Sintim

Herman O. Sintim received his BSc (first class) from University College London, his DPhil from the University of Oxford and is a postdoctoral fellow in the laboratory of Professor Eric Kool at Stanford. Herman will take up an independent position as Assistant Professor of Chemistry at the University of Maryland, College Park, in the summer of 2006. His research interests include nucleic acids, bacterial quorum sensing and development of new synthetic methods.

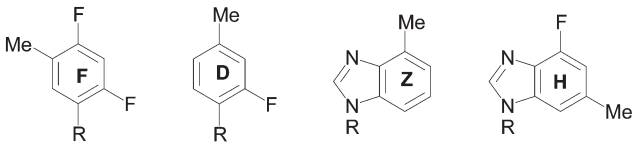

Fig. 1 Nonpolar nucleoside isosteres. R = deoxyribose or ribose.

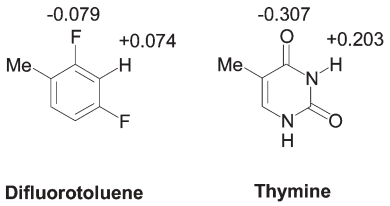

Fig. 2 Difluorotoluene and thymine showing atomic charges at 3 (H) and 4 (F/O).[15]

determined in crystal form and in aqueous solution; in both cases it was almost indistinguishable in conformation from natural thymidine (see Table 1 and Fig. 3).

A surprising observation, and a new hypothesis

In 1995 we began studying the properties of the nucleoside dF, and other isosteres, in DNA.[3] It was immediately apparent that these molecules were strongly destabilising to DNA helices when paired opposite natural bases (see below). In addition, while natural bases clearly showed energetic preferences for pairing with their complements rather than with mismatched bases, the nonnatural isosteres such as difluorotoluene showed little if any such preference (Table 2). Thus, in

Table 1 Comparison of properties of difluorotoluene deoxyriboside with thymidine

	Difluorotoluene	Thymine
Molecular weight	128.12	126.11
Sugar conformation	90% S	70% S
Glycosidic conformation	anti	anti
Log*P* of deoxyriboside.[16]	0.78	−1.1
Dipole moment (base only)/Debye[11]	1.84	4.19
Atomic charge at 3 (H)[15]	0.074	0.203
Atomic charge at 4 (F/O)[15]	−0.079	−0.307

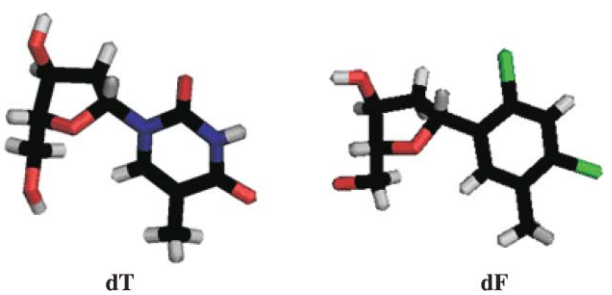

Fig. 3 X-ray crystal structures of dF and dT deoxynucleosides showing the glycosidic *anti* conformation and similar sugar puckers. The Watson–Crick-analogous groups of dF (C–F and C–H) make no hydrogen bonded contacts in the crystal (not shown), whereas those of thymine (C=O and N–H) do.[22c]

Table 2 Pairing stabilities of 12 bp DNA duplexes containing difluorotoluene or thymine paired opposite natural bases

5′-C T T T T C X T T C T T

3′-G A A A A G Y A A G A A

X.Y	T_m/°C	$\Delta G°_{37}$/kcal mol^{-1}	X.Y	T_m/°C	$\Delta G°_{37}$/kcal mol^{-1}
T.A	42.4	−9.7	A.T	42.4	−9.6
T.G	33.3	−7.5	G.T	34.7	−7.9
T.C	29.5	−6.6	C.T	25.7	−6.0
T.T	29.1	−6.8	T.T	29.1	−6.8
F.A	26.2	−6.2	A.F	27.5	−6.6
F.G	23.6	−6.0	G.F	25.1	−6.4
F.C	23.7	−5.8	C.F	25.1	−6.4
F.T	24.0	−5.9	T.F	24.8	−6.4

[a] Conditions: 1 M NaCl, 10 mM phosphate buffer (pH 7.0) with 0.1 mM EDTA, 2.5 μM each strand, monitored at 260 nm. Error in T_m is estimated at ±0.5 °C, and in free energy, ±10%.

DNA alone (without enzymes), difluorotoluene behaved much more like a nonpolar hydrocarbon (such as benzene)[4] than like a DNA base.

The surprising observation came when difluorotoluene was tested as a substrate for a DNA polymerase enzyme.[5] The compound was synthetically incorporated into a primer-template duplex, in the template strand immediately downstream from the primer 3′ terminus. In this DNA context, a polymerase then has the opportunity to select among nucleoside triphosphates present in solution (typically ~100–500 μM) and add one to the primer terminus, forming the normal phosphodiester bond. This elongated product can easily be distinguished by gel electrophoresis. For example, when the template contains thymine as the next base, a DNA polymerase (such as the well-studied DNA pol I, Klenow fragment, from *E. coli*, abbreviated Kf) will insert deoxyadenosine (dA) opposite this thymine in seconds, whereas it will not readily insert thymidine (dT), deoxyguanosine (dG), or deoxycytidine (dC). When difluorotoluene was tested as the template base, we found that the Kf enzyme efficiently inserted dA opposite this base surrogate, while it did not insert dT, dG, or dC (see Fig. 4). Quantitative kinetics studies of this reaction showed it to be very close to the efficiency of the natural base pair, and the selectivity for insertion of adenine as a partner was also very close to that of a natural pair.[5a] Conversely, dFTP, the nucleoside triphosphate derivative of the dF nucleoside analogue, was also synthesised and tested for incorporation into the elongating primer. Once again, the F–A base pair was enzymatically synthesised with surprisingly high efficiency, and with specificity that was the same as a natural base pair.[5b] It was subsequently shown that all the

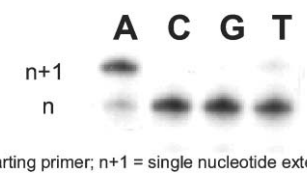

Fig. 4 Thymine-like polymerase activity of difluorotoluene. Gel electrophoretic picture showing polymerase extension products with difluorotoluene as the next nucleobase in the template downstream of the primer.

thymines in a DNA duplex could be replaced enzymatically with difluorotoluene.[5c]

This experimental outcome was surprising because of the expectation that (as taught in biochemistry textbooks at the time) Watson–Crick hydrogen bonds were the primary source of specificity in DNA replication.[6] Here was an example of a molecule that on its own did not show any pairing selectivity and behaved like a nonpolar hydrocarbon, and yet the enzyme was accepting it almost quantitatively as if it were a natural base. Clearly, the specificity of this base pairing was coming from the enzyme's influence and not from the difluorotoluene alone.

As a result of these observations, we began to question the need for Watson–Crick hydrogen bonds in DNA base pair replication at least for this specific enzyme,[5a,b] and began at the same time to develop an hypothesis that relied on steric effects rather than electrostatics to explain the fidelity of replication.[7] The earliest form of the hypothesis was that the enzyme tightly surrounds the base pair being synthesised, enforcing a base pair shape in which only the correct nucleotide would fit opposite the template base shape being addressed. Thus steric clashes would cause rejection of incorrect nucleotides. More details of the hypothesis are given in early reviews on the subject.[7]

It should be noted that the concept of base pair geometry as an influence in DNA replication was long discussed prior to these experiments; indeed, Watson and Crick recognised that the homologous purine–pyrimidine paired structure was significant in forming the double helix.[8] However, it was widely believed that both the specificity of hydrogen bonds and the geometry of pairing were important in replication. What was new in our work was (i) development of an experimental strategy for testing steric and electrostatic effects separately by use of nonpolar isosteres, (ii) the experimental observation that high efficiency and fidelity of replication could be achieved with a nonpolar isostere of a natural base, and (iii) the recognition that a hydrogen bonding explanation for replication was apparently not needed, and that steric effects might be the chief explanation that needed further study.

The debate begins

Clearly, many chemists and biochemists were surprised by these observations (as were we), and a high degree of interest was sparked. Experiments were undertaken in numerous laboratories to address aspects of the data, to repeat the experiments in multiple ways, and to comment on the findings. Not surprisingly, some of the debate was critical of our conclusions. Indeed, one of the authors received personal messages that expressed near-outrage over the claim that the hydrogen bonds of Watson and Crick, long believed to explain not only the self-assembly of the doubly helix, but its replication as well, were not needed.

The first, and sharpest peer-reviewed criticism of our conclusions regarding the replication of difluorotoluene came from Evans and Seddon, who in 1997 published a manuscript describing our conclusions as almost completely "in error".[9] They recognised the importance of the topic, stating that the results "undermine the basic theories of biochemistry". Briefly, the points of contention were these, as outlined in Seddon's claims:

Claim 1: Difluorotoluene is polar

The authors based this conclusion on the fact that it has a dipole moment (1.86 D in their calculations). In addition, the authors performed calculations showing that there are partial charges on fluorine and hydrogen. In a figure they displayed electrostatic potential maps of difluorotoluene and thymine, with the scale magnified for difluorotoluene (the fact that different scales were used was not stated), showing apparently similar charge distributions.

Claim 2: Hydrogen bonding experiments were flawed

The authors claimed that 9-ethyladenine was not highly soluble in chloroform, and implied that this may explain the lack of an apparent complex, despite the fact that such a complex was observed with 9-ethyladenine using a uracil derivative.

Claim 3: Experiments with pyridines show a complex with difluorotoluene

The authors performed NMR titrations of difluorotoluene with pyridine and chloropyridine, and claimed evidence for complexes, which they hypothesised were hydrogen-bonded complexes involving the 3-hydrogen of difluorotoluene.

Claim 4: Fluorocarbons can form hydrogen bonds with amino groups

The authors cited a review which states that "F···H–X interaction may contribute to the overall binding energy, up to half of the strength of the original hydrogen bond to oxygen".[10]

The authors argued that if difluorotoluene forms a hydrogen-bonded complex with adenine, then our hypothesis was "in error", and that there was no reason to change the hydrogen bonding "paradigm" for DNA replication. At the time of Seddon's publication we opted not to respond directly, but rather decided to allow the data to speak for themselves over time. We were not asked by this journal to comment or respond to the manuscript. However, an initial reading of the manuscript found points that could be refuted. For example, the authors used the fact that difluorotoluene has a dipole (which is obvious from its lack of symmetry) to argue that it is polar, but did not mention the fact that its dipole is small as compared with that of thymine (1.84 D *vs.* 4.19 D).[11] Moreover, the authors showed a potentially misleading illustration comparing charge distributions of difluorotoluene and thymine which apparently shows that they are virtually the same; however, Evans and Seddon did not state that they used quite different scales to color the charge maps, and had to magnify the difluorotoluene map considerably to make the two appear similar. In addition, the fact that the authors claimed a hydrogen bonded complex between pyridine and difluorotoluene is curious, considering the experiment was done in chloroform, which itself has a considerably more acidic proton

than the one in difluorotoluene; in addition, the authors apparently did not consider other possible complexes such as ones with stacked geometry, or self-associated structures.

Since the publication of this early critical paper, discussion of these points has continued. For example, there remain a number of more recent claims of hydrogen bonding involving C–F and C–H groups (see below).[12] In addition to hydrogen bonding effects, there is also a continuing debate about our steric hypothesis for DNA replication.[13] This will be discussed below as well.

Debate is essential

At the time that this early discussion of our work began, we recognised that debate about any new hypothesis in science is an important part of progress. New ideas must be tested, and alternative explanations must be considered. Eventually, the preponderance of the data will be most acceptably represented by the best hypothesis, and the majority of the scientific community will come to accept it. The fact that scientists were willing to debate our hypothesis and spend time in the laboratory studying it was a promising sign that the topic was of substantial interest.

A decade's worth of data

Over the past decade, more than a hundred manuscripts from dozens of laboratories have addressed properties of difluorotoluene deoxyriboside and other nonpolar nucleoside isosteres. In fact, two such isosteres (including difluorotoluene deoxyriboside itself) are now commercially available as phosphoramidite derivatives,[14] making them widely available for study in DNA. Thus we now have much more information about the physical, chemical, biochemical, and biological properties of this compound. This allows one, with much greater confidence, to address many of the early points of contention regarding this molecule. Space does not allow great detail since the number of studies to date is large, and the reader is invited to go to the primary literature to investigate the details and the specific data collected. However, we can summarise the findings and general conclusions to date:

Difluorotoluene is effectively nonpolar

Although difluorotoluene contains two C–F bonds that are somewhat polarized toward F, the base analogue as a whole behaves nonpolar in its properties, and is in some respects even less polar than the hydrocarbon toluene. Calculated atomic charges on fluorines and on the 3-H of difluorotoluene are a small fraction of the charges on the analogous atoms of thymine (see Fig. 2 and Table 1).[15] The calculated dipole moment of difluorotoluene is also much lower than that of thymine.[11] A classic measure of polarity in organic molecules is partitioning between octanol and water (log*P*); the deoxyriboside dT partitions strongly to the aqueous phase, while the dF nucleoside partitions strongly to the octanol phase.[16] Indeed, the log*P* value of dF is greater than that of the toluene nucleoside in which hydrogens replace the fluorines.[16b] This is consistent with the fact that fluorocarbons and fluorine-substituted hydrocarbons are more hydrophobic than pure hydrocarbons.[17]

Difluorotoluene does not measurably form Watson–Crick hydrogen bonds

There are many lines of evidence supporting this, and it would take a full monograph to describe them all in detail. However, some of the main points can be summarised here: (a) The calculated atomic charges of fluorines and hydrogen of F are considerably less than the analogous atoms of thymine, and since electrostatics are a main contributor to hydrogen bonding, this predicts considerably weaker attraction.[15] (b) The nucleoside dF is strongly nonpolar as experimentally measured by log*P*;[16a] this demonstrates the low affinity of the difluorotoluene for water, suggesting that hydrogen bonds to water are not favourable. (c) The nucleoside dF is even less polar in its properties than hydrocarbon analogues,[16b] and it is known that fluorocarbons are more hydrophobic than hydrocarbons.[17] (d) NMR titrations against 9-ethyladenine in $CDCl_3$ show a complex with a uracil derivative but not difluorotoluene;[5a] and yet hydrogen bonds are much stronger in chloroform than in water. Two subsequent NMR studies of difluorotoluene in inter- and intramolecular systems provided new evidence against the Seddon claim of hydrogen bonding with adenine.[18] (e) dF is strongly destabilising in DNA even when paired opposite adenine.[3] This demonstrates that not only is there a lack of attraction between A and F, but in fact suggests net energetic *repulsion* between the two (due to adenine's preference for hydrogen bonding to water). (f) dF shows little or no pairing selectivity for adenine in DNA, in contrast to what is clearly observed with the hydrogen bonding base thymine.[3] (g) Several theoretical studies have concluded that there is little or no hydrogen bonding by F with A in the DNA context. Even in the gas phase, the energy of the F–A complex is only a small fraction of that of T–A.[15,19] (h) Several published analyses of large X-ray crystal structure databases show little or no evidence for any hydrogen bonds in organic molecules involving C–F, particularly so for fluorine on sp^2 carbon.[10,20]

Note that we do not claim that difluorotoluene *cannot* form hydrogen bonds with adenine (particularly in the gas phase, where a weak complex may well be possible), but rather, that such bonds must be so weak in water that they are difficult to measure. Indeed, the evidence suggests that in water, difluorotoluene is energetically repulsive when paired opposite adenine, largely because adenine would much prefer water as a hydrogen-bonding partner.[3]

Difluorotoluene stacks strongly in DNA

A second property that distinguishes difluorotoluene from thymine is its strong stacking propensity in DNA. When one dF nucleoside is placed at the end of a short DNA duplex in a "dangling" position, it stabilises the entire helix by *ca.* -1.3 kcal mol^{-1}, whereas thymidine stabilizes the helix by only -0.6 kcal.[21a] Since the polarisability and size of the two are essentially the same, this cannot be attributed to van der Waals interactions alone; moreover, the added stabilisation by F cannot be attributed to electrostatics since F is much less

polarised than T. We have attributed the strong stacking of F (and other nonpolar nucleobase isosteres) to the hydrophobicity of this compound. The face-to-face geometry of stacking removes considerable energetically unfavourable surface area from water. In general, stacking of aromatic systems in DNA correlates well with size and hydrophobicity.[21b]

Difluorotoluene in DNA does not disrupt the structure of the double helix

Two published structures of DNA containing the dF nucleoside in solution show a lack of distortion of the DNA.[22] One structure (solved by high-resolution 2D-NMR studies) shows difluorotoluene placed opposite adenine in the center of a 12-base-pair helix; little or no difference in the location of adenine and difluorotoluene was found as compared with thymine and adenine in the same duplex (Fig. 5). A second structure shows difluorotoluene paired opposite an adenine isostere, 4-methylbenzimidazole;[22b] again, there is no significant distortion of the double helix. The fact that difluorotoluene does not distort DNA in which it is substituted is important, since it allows scientists to use it as a probe of electrostatic effects in protein–DNA interactions without fear of interference by a non-native structure.

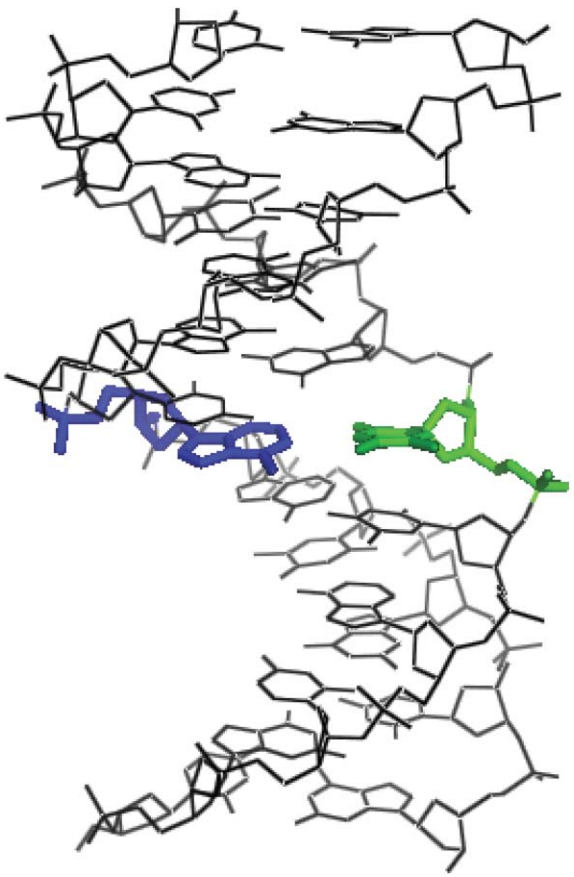

Fig. 5 Structure of a difluorotoluene deoxynucleotide (coloured green) paired opposite adenine (coloured blue) in a 12 bp DNA duplex. Despite the destabilization caused by the nonpolar nucleotide, the double helical structure is unperturbed.[22a]

In the absence of enzymes, difluorotoluene is destabilising and nonselective in its pairing in DNA

When substituted into DNA opposite adenine near the center of DNA duplexes, F destabilises the entire duplex by 3–4 kcal mol^{-1} as compared to the case where thymine is placed at the same position.[3] Moreover, difluorotoluene yields nearly the same thermal melting temperature (T_m) values and free energies for these duplexes regardless of whether F is paired opposite A, C, T, or G. By contrast, thymine (in the same context) shows a strong preference for pairing with adenine, giving 3–4 kcal mol^{-1} greater thermodynamic stability than the mismatched pairings (Table 2). The lack of selectivity of difluorotoluene is attributed to its lack of significant hydrogen bonding ability, and the destabilisation by this molecule is attributed to the cost of desolvation of the polar DNA bases near it. Other nonpolar nucleoside isosteres (whether they contain fluorine or not) give virtually the same results.[23]

Difluorotoluene is replicated efficiently by several high-fidelity DNA polymerases

Despite the destabilisation and lack of selectivity shown by difluorotoluene in DNA alone, replicative DNA polymerase enzymes have been shown to process this nucleoside with high efficiency and fidelity.[5,24] Among the enzymes shown to have this activity are DNA pol I (Klenow fragment), Taq polymerase, and T7 DNA polymerase. As a template nucleoside, F is processed with quantitative efficiency near that of natural thymidine, and the specificity of deoxyadenosine insertion (rather than the other nucleotides) is nearly that of thymidine itself.[5a] As a deoxynucleoside triphosphate analogue, dFTP is processed with somewhat lower efficiency than dTTP, but its selectivity for insertion opposite adenine remains near the levels of dTTP.[5b]

Difluorotoluene is replicated as thymine in living cells

Single difluorotoluene residues have been placed into a bacteriophage genome, which was transfected into *E. coli* cells and replicated.[25c,d] The efficiency of replication was relatively high (at 10–15% that of a natural base), and importantly, the difluorotoluene was accurately replaced with thymine in the progeny phage, with mispairings occurring only once per $\sim$1000 replications (see Fig. 6). This occurred in the presence of the five known DNA polymerases that operate in *E. coli*, along with the intact DNA repair machinery. Importantly, a benzimidazole-based adenine isostere was also replicated successfully (albeit at lower efficiency), demonstrating that fluorine is not needed for a nonpolar isostere to have biological activity.

New analogues of difluorotoluene add support to the steric hypothesis

Because it is virtually the same size and shape as thymine, difluorotoluene serves as a nearly ideal molecular probe of the importance of electrostatic effects in natural processes involving thymine. The above findings led to our ruling out of a purely hydrogen bond-based explanation for DNA replication,

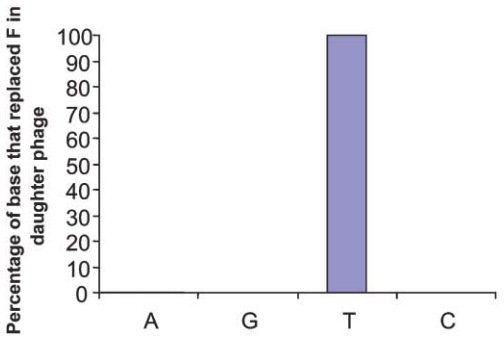

Fig. 6 Thymine-like coding behaviour of difluorotoluene in living *E. coli*. Shown is a histogram of recovered DNA after replication, showing which bases replaced F in the replicated progeny. Actual percentages of A, G, C are 0.1%, 0.1% and 0% respectively.[25d]

and led us to propose a steric-based explanation instead. Thus we needed new molecules to test this steric argument. As described below, a series of thymidine nonpolar analogues that have varied size, and a series that have varied shapes have been studied recently. Both series revealed remarkable sensitivity to steric effects in DNA replication.

New nonpolar nucleoside variants: analogues of difluorotoluene

To explicitly test the effects of steric changes in DNA without interference from hydrogen bonding, we first prepared a series of thymidine shaped mimics in which size was gradually varied over a 1.0 Å range (Fig. 7).[25] These were varied by replacing the fluorine groups at the 2 and 4 positions in difluorotoluene with smaller (H) and larger (Cl, Br, I) groups. Thus overall, the approximate shape of thymine was retained, while size was varied.

A set of thymidine shaped variants was also prepared. These compounds maintained the approximate size of thymidine, but varied shape by altering the location of groups along the base pairing edge of the molecule.[26] In combination with the size-varied compounds, these two molecule probe sets were prepared to test steric effects in DNA at a level of resolution (*i.e.*, in sub-Angstrom increments) that was previously untested.

New data on the steric hypothesis for DNA replication

The thymidine analogues having varied size were studied recently as substrates for DNA polymerase I (Klenow fragment), and in living *E. coli* cells as well. In the enzymatic studies, the compounds showed remarkable differences in activity associated with small changes in size. Changes of as little as 0.5 Å in size resulted in more than three orders-of-magnitude change in activity.[25c] The largest analogues were rejected as substrates, presumably as a result of steric clashes. Importantly, analogues that were too small were as inefficient as too-large ones; this effect is thought to be due to the energetic cost of voids left in the active site. Another interesting observation was that the best thymidine analogue among these in DNA replication was dichlorotoluene, which is slightly larger than the natural base. We hypothesised that the enzyme has evolved a slightly too-large active site to allow for evolutionarily beneficial mutations.

In living bacterial cells, a similar level of sensitivity to size was seen.[25c] The smallest analogue (toluene) was much less efficiently replicated than the best (dichlorotoluene), and larger analogues were poorly replicated, apparently due to steric clashes in the replicative polymerases. The fidelity was greatest with dichlorotoluene, which was processed quite efficiently, at *ca.* 20% the efficiency of a natural base. Aside from the importance in the basic study of DNA replication, this is a significant step in the long-term goal of designing nonnatural replacements for the components of living systems: dichlorotoluene is the first efficient human-designed base that functions well in a living cell.

Varied shapes (Fig. 8) have also recently been found to have large influences on DNA replication by DNA polymerase I (Klenow fragment). Mono- and di-chlorinated toluenes with varying substitution patterns varied in activity by as much as 3500-fold with this enzyme.[26] The most efficient analogues (for replication opposite adenine) were the ones that closely resembled the shape of thymine.

Note that important work on varied nucleobase shapes (also using substituted benzenes) has come recently from the Romesberg and Shultz laboratories, and has given useful evidence of the importance of close "hydrophobic packing".[26b,c]

Continuing points of discussion, and unanswered questions

There remain a number of issues that are still under study and debate in the literature. The following are three examples:

1. Not all polymerases are the same

The earliest studies of DNA polymerase activities using nonpolar nucleoside analogues involved DNA polymerase I (in particular, the Klenow fragment), and T7 DNA polymerase. These enzymes fall into the "A" family of polymerases, a homologous group of enzymes that typically exhibit high fidelity

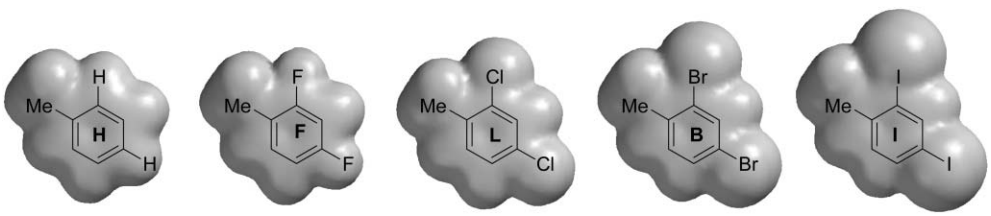

Fig. 7 New thymine-mimicking nonpolar isosteres having size varied over a 1.0 Å range.[25a–c]

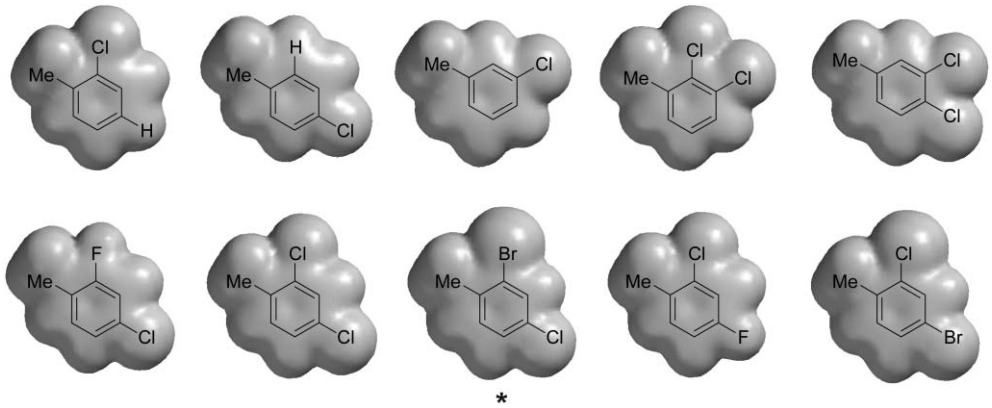

Fig. 8 Some shape variants of difluorotoluene. Despite being structurally related, they vary by 3500-fold in activity with DNA polymerase, demonstrating the importance of nucleobase shape in DNA replication.[26] The compound marked with an asterisk is replicated with fidelity higher than thymine itself.

and are involved in genomic replication.[27] The A-family enzymes commonly accept nonpolar isosteres (difluorotoluene in particular) as good substrates.[5,24] However, there are at least six classes of DNA polymerases now known (classified by sequence homology): A, B, C, X, Y, and reverse transcriptases.[27] Nonpolar isosteres have begun to be tested with a wider set of enzymes from these different classes. To date, it appears that these analogues are often good substrates for the A, B, and reverse transcriptase classes of polymerases.[28] They have not been studied as yet with C-family enzymes. With the X family there is only early data; it appears that difluorotoluene is not a good substrate for pol beta from this class,[28] but it is a substrate for the enzyme TdT.[29] Finally, there is the set of Y-family polymerases; these enzymes generally play the role of repair enzymes, assisting DNA replication in the presence of damaged or mispaired bases. To date, nonpolar isosteres have been tested with a few examples from this class (including dinB, Dpo4, pol kappa, and pol eta),[30] and these analogues are generally very poor substrates. Kinetics studies have shown them to be processed with efficiencies near those of mismatched based pairs (or less), and with considerably lower fidelity than natural hydrogen-bonded pairs are.

The finding that Y family enzymes accept nonpolar nucleoside isosteres poorly suggests that this class of enzymes uses a different mechanism for selecting and incorporating nucleotides into DNA. The current consensus hypothesis is that these repair enzymes (in marked contrast to A family replicative enzymes) may require Watson–Crick hydrogen bonding to incorporate a nucleotide.[30] The mechanism by which this might works is still unclear; however, these enzymes appear to be structurally more ''open'' around their active sites,[31] and recent steric studies suggest that they are quite flexible and nonselective with regard to base pair size.[30*d*] It has been hypothesised that in the absence of strong steric constraints, hydrogen bonding may aid in positioning of the incoming nucleotide so that bond formation may occur efficiently.

2. Some discussion of ''weak'' hydrogen bonds in DNA continues

There remains a good deal of literature discussion concerning hydrogen bonds involving weaker hydrogen donors such as C–H, and weaker hydrogen bond acceptors such as C–F.[12] In nonpolar environments, such as the gas phase, there is significant evidence that these bonding interactions are energetically real. In crystal structures, the presence of short atomic distances and/or implied bond directionality could provide some support for such interactions,[10] but two recent reviews argue against any C–F hydrogen bond acceptors, especially involving sp^2 carbon.[10,20*a*] In the aqueous environment with DNA or RNA, it is very difficult to find any experimental evidence for positive energetic contributions for such interactions. The water solvent makes any hydrogen bonds much weaker than they are in nonpolar environments, by providing a high dielectric that shields much of the electrostatic component, and by providing direct competition for the hypothesised donors and acceptors. Nevertheless, such bonds in RNA have been hypothesised in a recent published study.[12*b*,*c*] Interestingly, despite this claim, base pairing by difluorobenzene showed no selectivity for adenine over other bases. We remain unconvinced of any significant energetic contributions from C–H donors or C–F acceptors in DNA or RNA in water (see our discussion above).

3. Debate on the influence of steric effects continues

A few recent studies have raised new questions about the effects of sterics with the A and B family DNA polymerases.[13,32] Berdis studied a series of 5-substituted indoles with T4 DNA polymerase, and found low selectivity and activity; it was concluded that steric effects were not an influence with this enzyme.[13*a*] Similarly, Engels and Kuchta studied DNA Pol I (Klenow) using 5- and 6-substituted benzimidazoles, and concluded that steric effects were not important after observing low selectivity with different analogue substitutions.[13*b*] We have pointed out that these results can be explained by steric reasoning, and we suggested that all of these analogues were too large to be accepted as part of a base pair in the standard *anti* conformation.[26] We hypothesised that instead they are flipped into the *syn* conformation, where they are likely to have some shape similarity to C. Interestingly, in most cases these analogues showed low activity and weak G selectivity, as might be

expected if this were the case.[13a,b] New approaches to evaluating such *syn*/*anti* effects would be useful in shedding more light on this issue.

Finally, Davisson and Bergstrom have studied pyrrole DNA base surrogates as possible "universal" nucleotide analogues.[32] They recently claimed that their observation of selectivity between G and A with these molecules is not consistent with a steric argument. We do not agree with this interpretation, since the bases used had low activity, were conformationally ambiguous, and had an unknown level of selectivity; nevertheless, the debate on steric effects is still continuing, and there remain some open questions among some workers in the research community.

Untested questions

The existence of these debated points adds justification for continuing study of difluorotoluene and other nonpolar nucleobase analogues. In addition to these debates, there remain other, as yet largely unexplored, issues regarding DNA structure and replication that might well be addressed by such molecular probes. Among these remaining questions are the following:

1. How do as-yet-unstudied DNA polymerases respond to nonpolar nucleoside isosteres? To date, nonpolar nucleoside isosteres have been studied primarily with A- and Y-family DNA polymerases. Although one or two preliminary studies have qualitatively examined difluorotoluene as a substrate for B and X family polymerases and reverse transcriptases, to date there is little if any kinetic data available for any of these classes of enzymes, and several examples of each class are known. In addition, the C-family of polymerases remains completely unstudied. Thus the influence of hydrogen bonding and steric effects are essentially unknown for most known DNA polymerases. In this regard, the field is still in its early stages.

2. Can steric effects in the DNA be complemented by steric effects in mutant DNA polymerases? We and others have hypothesised that high-fidelity DNA polymerases tightly surround the incipient base pair to enforce the size and shape of the incoming nucleobase.[33] A close examination of X-ray crystal structures of DNA polymerases bound to DNA reveals sidechains that interact closely with the sugar and base of the incipient base pair. In principle, one could test the steric influence of a given enzyme on the DNA not only by varying the size and shape of the bases, but also by varying the sizes and shapes of amino acid sidechains near these points of contact. It is possible, for example, that a larger base pair might be more readily accepted if smaller amino acid sidechains were substituted in the active site. Screening of polymerase mutants with modified DNA nucleoside analogues has recently begun to be pursued in a few laboratories.[34]

3. Can structural studies with DNA polymerases complement what we currently know about the effects of hydrogen bonds and steric effects? Now that it is possible to co-crystallise a number of different DNA polymerases with DNA template, primer and nucleotide bound, it may be of interest to obtain structures of one or more enzymes with nonpolar nucleoside analogues as part of the incipient pair. Questions to be answered include whether they occupy the same orientation and geometry as their natural counterparts. It would also be of interest to observe how altered sizes and shapes of DNA bases and pairs affect the local active site and the global structure.[35]

The combination of several ongoing debates and a number of unanswered questions leaves much justification for continuing to study difluorotoluene and other nucleoside analogues as substrates for polymerases. We hope to study many of these issues in the future.

Other applications of difluorotoluene and nonpolar nucleoside isosteres

Of course, nucleosides and nucleotides play a much broader role in biology than simply forming a double helix and encoding the copying of that helix. As a result, difluorotoluene deoxyriboside and other nonpolar nucleoside isosteres have recently begun to be employed much more widely in other biochemical and biological studies. Below is a brief summary of some of those studies.

DNA proofreading

One important mechanism by which DNA polymerases increase fidelity of replication is in proofreading, whereby an already-added nucleotide is removed again. This happens more frequently when the nucleotide is incorrect, which leads to an increase in total fidelity. Difluorotoluene has been used in quantitative studies of proofreading,[36] which led to the hypothesis that hydrogen bonds in the terminal base pair are important in determining editing rates.

DNA curvature

It has been recognized for some time that certain sequences of DNA are inherently curved, leading to bending of the helix overall. The mechanisms that cause this curvature are under debate, and some of the leading hypotheses involve localised electrostatic interactions.[37] Difluorotoluene substitution in curved DNA sequences has been observed to affect bending only at certain locations, which suggests that local electrostatic interactions (such as with metal ions) may be a chief cause of this curvature.[37b,c]

DNA repair processes

A number of DNA repair enzymes are known to bind and excise bases involved in damaged or mismatched pairs in DNA. Of interest is the mechanism by which these base excision repair (BER) enzymes identify such pairs as damaged. Difluorotoluene and other nonpolar nucleoside isosteres have been used in multiple studies of these mechanisms.[38] Difluorotoluene deoxyriboside is of special interest because its C–C glycosidic bond cannot be cleaved by glycosidases.

Nucleoside kinase studies

The high activity of difluorotoluene with replicative polymerase enzymes has led to interest as to whether such nucleosides might be incorporated into native DNA when incubated with cells. If such nucleosides were taken up into cells and

phosphorylated by kinase enzymes, this could lead to useful applications in antiviral and anticancer studies. Nonpolar nucleosides such as difluorotoluene deoxyriboside have been studied with a number of kinases, and some have been shown to exhibit high levels of activity.[39]

DNA helix stabilisation

The strong stacking properties of the nonpolar nucleosides have been employed in stabilisation of helical and folded DNAs.[40] For example, they have been substituted into loops of hairpin DNAs, where they stabilise the folded structure significantly.[40a] Difluorotoluene nucleoside has been examined for its effects not only in double-stranded DNAs, but in triple helices as well.[40b]

Transcription

Nonpolar nucleoside analogues have been shown to have a useful application for *in vitro* transcription to produce RNAs in the test tube.[41] A common problem in this so-called "run-off" transcription is addition of one extra ribonucleotide beyond the end of the DNA template. This results in inhomogeneous product RNAs having different lengths. It has been observed that if the DNA template contains one nonpolar nucleoside analogue at its end, the extra nucleotide addition is suppressed.

RNA helix studies

The concept of nonpolar nucleoside isosteres has been expanded into the RNA world in multiple laboratories. Ribonucleoside isosteres have been reported; in the case of RNA, the analogue of difluorotoluene is difluorobenzene, isosteric with uracil.[42] Studies by Engels have used such compounds to explore the effects of hydrogen bonding on pairing in RNA helices.[42b,c]

DNA ligation

The mechanism by which ligase enzymes, which join duplex DNAs together, has recently been investigated by use of difluorotoluene near the ligation junction.[43] Minor groove hydrogen bonding interactions was important for ligation reactions by both Tth and T4 DNA ligases.

RNA interference mechanisms

RNA interference is a widely useful tool in biology, for knocking down the activity of specific genes. The mechanism of mRNA recognition by short interfering RNAs was studied recently using nonpolar ribonucleoside isosteres.[44] Remarkably, at some positions in the RNA, difluorobenzene or difluorotoluene could replace uracil with nearly full retention of cellular activity, and with sequence specificity at least as high as that of uracil.

Nonnatural base pair design

The strong stacking and hydrophobic nature of nonpolar nucleoside isosteres has been involved in several studies of potential nonnatural DNA base pairs, wherein both bases in a pair are nonnatural.[3,25b,45] The goal is to find pairs that are stable and are selective, with each base preferring to pair with the other rather than with the natural DNA bases. Hirao reported an efficient designed partner for difluorotoluene,[45] and some of the larger nonpolar isosteres of thymine also pair well with one another.[25b] Several labs have recently been exploring the possibility of hydrophobic base pair design.[46]

Charge transfer in DNA

The effects of oxidation potential of DNA bases on cation radical migration through DNA was studied with difluorotoluene.[47] Interestingly, the molecule caused no measurable change in charge migration relative to natural thymine.

Analogues including altered backbone

There has recently been some interest in combining DNA base analogues with synthetically altered analogues of the DNA backbone.[48] One recent example combined the well-known and high-affinity peptide nucleic acid (PNA) backbone structure with the base analogue difluorotoluene.[48a] Other nonpolar base surrogates have been investigated with PNA as well.[48b] A different example using a peptide backbone was also reported.[47c]

The changing of a scientific paradigm

As T. S. Kuhn pointed out,[49] the changing of a long-standing scientific hypothesis (paradigm) requires much work and repeated debate. While disruptive and facing resistance at first, a successful new hypothesis eventually becomes widely accepted. In the present case, some discussion of hydrogen bonding and steric effects in DNA and in DNA replication remains in the literature, and should continue as long as some scientists remain unconvinced by the existing data. However, many scientists have come to accept that Watson–Crick hydrogen bonds are not the chief force in selection of nucleotides during DNA replication, and that steric effects are the most important factor in this fidelity. Recent reviews of DNA replication fidelity now reflect the idea that replicative DNA polymerases are governed largely by steric effects in making DNA copies with high fidelity.[33] Textbooks have changed as well; for example, the classic biochemistry text of Stryer in its earlier editions cited the specificity of hydrogen bonding as the source of replication fidelity,[6a] whereas more recent editions have removed this explanation and have instead cited the complementary fit of the bases.[50]

Acknowledgements

We are grateful for the dedicated efforts of our coworkers on this project over the past decade. We also thank the US National Institutes of Health (GM072705) for support of this work.

Notes and references

1 B. A. Schweitzer and E. T. Kool, *J. Org. Chem.*, 1994, **59**, 7238–7242.

2 K. M. Guckian and E. T. Kool, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 2825–2828.
3 B. A. Schweitzer and E. T. Kool, *J. Am. Chem. Soc.*, 1995, **117**, 1863–1872.
4 T. A. Millican, G. A. Mock, M. A. Chauncey, T. P. Patel, M. A. Eaton, J. Gunning, S. D. Cutbush, S. Neidle and J. Mann, *Nucleic Acids Res.*, 1984, **12**, 7435–7453.
5 (*a*) S. Moran, R. X.-F. Ren, S. Rumney and E. T. Kool, *J. Am. Chem. Soc.*, 1997, **119**, 2056–2057; (*b*) S. Moran, R. X.-F. Ren and E. T. Kool, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 10506–10511; (*c*) D. Liu, Sean Moran and E. T. Kool, *Chem. Biol.*, 1997, **4**, 919–926.
6 (*a*) L. Stryer, in *Biochemistry*, W. H. Freeman, New York, 4th edn., 1995, p. 800; (*b*) J. D. Watson, N. H. Hopkins, J. W. Roberts, J. A. Steitz and A. M. Weiner, in *Molecular Biology of the Gene*, Benjamin Cummings, Menlo Park, 4th edn., 1987, p. 283.
7 (*a*) E. T. Kool, *Biopolymers*, 1998, **48**, 3–17; (*b*) E. T. Kool, J. C. Morales and K. M. Guckian, *Angew. Chem., Int. Ed.*, 2000, **39**, 990–1009; (*c*) E. T. Kool, *Annu. Rev. Biophys. Biomol. Struct.*, 2001, **30**, 1–22.
8 (*a*) J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737–738; (*b*) H. Echols and M. F. Goodman, *Annu. Rev. Biochem.*, 1991, **60**, 477–511; (*c*) A. Kornberg, in *DNA replication*, Freeman & Co., New York, 2nd edn., 1992.
9 T. A. Evans and K. R. Seddon, *Chem. Commun.*, 1997, **21**, 2023–2024.
10 J. A. K. Howard, V. J. Hoy, D. O'Hagan and G. T. Smith, *Tetrahedron*, 1996, **52**, 12613–12622.
11 B. D. Silverman, M. C. Pitman and D. E. Platt, *J. Biomol. Struct. Dyn.*, 1999, **16**, 1169–1175.
12 (*a*) V. R. Thalladi, H. -C. Weiss, D. Blaser, R. Boese, A. Nangia and G. R. Desiraju, *J. Am. Chem. Soc.*, 1998, **120**, 8702–8710; (*b*) J. W. Bats, J. Parsch and J. W. Engels, *Acta Crystallogr., Sect. C*, 2000, **56**, 201–205; (*c*) J. Parsch and J. W. Engels, *Helv. Chim. Acta*, 2000, **83**, 1791–1808; (*d*) J. A. Frey, R. Leist and S. L. Leutwyler, *J. Phys. Chem. A*, 2006, **110**, 4180–4187; (*e*) J. A. Frey, R. Leist and S. L. Leutwyler, *J. Phys. Chem. A*, 2006, **110**, 4188–4195.
13 (*a*) X. Zhang, I. Lee and A. Berdis, *J. Biochem.*, 2005, **44**, 13101–13110; (*b*) K. Kincaid, J. Beckman, A. Zivkovic, R. L. Halcomb, J. W. Engels and R. D. Kuchta, *Nucleic Acids Res.*, 2005, **33**, 2620–2628.
14 Glen research catalogue No. 10-1095-95.
15 C. F. Guerra and F. M. Bickelhaupt, *J. Chem. Phys.*, 2003, **119**, 4262–4273.
16 (*a*) $\log P_{exp.} = 0.78$ for difluorotoluene (dF) nucleoside; (*b*) $\log P_{exp.} = 0.71$ for toluene nucleoside (dH).
17 L. Barthel-Rosa and J. A. Gladysz, *Coord. Chem. Rev.*, 1999, **190–192**, 587–605.
18 (*a*) K. S. Schmidt, R. K. O. Sigel, D. V. Filippov, G. A. van der Marel, B. Lippert and J. Reedijk, *New J. Chem.*, 2000, **24**, 195–197; (*b*) N. Shibata, B. K. Das, K. Harada, Y. Takeuchi and M. Bando, *Synlett*, 2001, 1755–1758.
19 (*a*) R. Meyer and J. Suhnel, *J. Biomol. Struct. Dyn.*, 1997, **15**, 619–624; (*b*) X. Wang and K. N. Houk, *Chem. Commun.*, 1998, **23**, 2631–2632; (*c*) D. Barsky, E. T. Kool and M. E. Colvin, *J. Biomol. Struct. Dyn.*, 1999, **16**, 1119–1134; (*d*) E. Cubero, C. A. Laughton, F. J. Luque and M. Orozco, *J. Am. Chem. Soc.*, 2000, **122**, 6891–6899; (*e*) J. Florian, M. F. Goodman and A. Warshel, *J. Phys. Chem. B*, 2000, **104**, 10092–10099; (*f*) E. C. Sherer, S. J. Bono and G. C. Shields, *J. Phy. Chem. B*, 2001, **105**, 8445–8451; (*g*) J. Sponer, P. Jurecka and P. Hobza, *J. Am. Chem. Soc.*, 2004, **126**, 10142–10151.
20 (*a*) J. D. Dunitz, *ChemBioChem*, 2004, **5**, 614–621; (*b*) J. D. Dunitz and R. Taylor, *Chem.–Eur. J.*, 1997, **3**, 89–98; (*c*) L. Shimoni and J. P. Glusker, *Struct. Chem.*, 1994, **5**, 383–397.
21 (*a*) K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, P. L. Paris and E. T. Kool, *J. Am. Chem. Soc.*, 1996, **118**, 8182–8183; (*b*) K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, D. C. Tahmassebi and E. T. Kool, *J. Am. Chem. Soc.*, 2000, **122**, 2213–2222.
22 (*a*) K. M. Guckian, T. R. Krugh and E. T. Kool, *Nat. Struct. Biol.*, 1998, **5**, 954–959; (*b*) K. M. Guckian, T. R. Krugh and E. T. Kool, *J. Am. Chem. Soc.*, 2000, **122**, 6841–6847; (*c*) D. W. Young, P. Tollin and H. R. Wilson, *Acta Cryst.*, 1969, **B25**, 1423–1432.
23 B. A. Schweitzer and E. T. Kool, *J. Am. Chem. Soc.*, 1995, **117**, 1863–1872.
24 (*a*) T. Kim, L. G. Brieba, T. Ellenberger and E. T. Kool, *J. Biol. Chem.*, 2006, **281**, 2289–2295; (*b*) J. C. Morales and E. T. Kool, *J. Am. Chem. Soc.*, 2000, **122**, 1001–1007.
25 (*a*) T. Kim and E. T. Kool, *Org. Lett.*, 2004, **6**, 3949–3952; (*b*) T. Kim and E. T. Kool, *J. Org. Chem.*, 2005, **70**, 2048–2053; (*c*) T. Kim, J. C. Delaney, J. M. Essigmann and E. T. Kool, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15803–15808; (*d*) J. C. Delaney, P. T. Henderson, S. A. Helquist, J. C. Morales, J. M. Essigmann and E. T. Kool, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 4460–4473.
26 (*a*) H. O. Sintim and E. T. Kool, *Angew. Chem., Int. Ed.*, 2006, **45**, 1974–1979; (*b*) A. M. Leconte, S. Matsuda and F. E. Romesberg, *J. Am. Chem. Soc.*, 2006, **128**, 6780–6781; (*c*) S. Matsuda, A. A. Henry and F. E. Romesberg, *J. Am. Chem. Soc.*, 2006, **128**, 6369–6375.
27 K. Bebenek and T. A. Kunkel, *Adv. Protein Chem.*, 2004, **69**, 137–165.
28 J. C. Morales and E. T. Kool, *Biochemistry*, 2000, **39**, 12979–12988.
29 Y. Cho and E. T. Kool, unpublished results.
30 (*a*) O. Potapova, C. Chan, A. M. DeLucia, S. A. Helquist, E. T. Kool, N. D. F. Grindley and C. M. Joyce, *Biochemistry*, 2006, **45**, 890–898; (*b*) W. T. Wolfle, M. T. Washington, E. T. Kool, T. E. Spratt, S. A. Helquist, L. Prakash and S. Prakash, *Mol. Cell. Biol.*, 2005, **25**, 7137–7143; (*c*) H. Hwang and J.-S. Taylor, *Biochemistry*, 2004, **43**, 14612–14623; (*d*) S. Mizukami, T. Kim, S. A. Helquist and E. T. Kool, *Biochemistry*, 2006, **45**, 2772–2778.
31 W. Yang, *FEBS Lett.*, 2005, **579**, 868–872.
32 N. Paul, V. C. Nashine, G. Hoops, P. Zhang, J. Zhou, D. E. Bergstrom and V. J. Davisson, *Chem. Biol.*, 2003, **10**, 815–825.
33 (*a*) M. F. Goodman, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 10493–10495; (*b*) T. A. Kunkel and K. Bebenek, *Annu. Rev. Biochem.*, 2000, **69**, 497–529; (*c*) E. T. Kool, *Annu. Rev. Biochem.*, 2002, **71**, 191–219; (*d*) M. Strerath, J. Cramer, T. Restle and A. Marx, *J. Am. Chem. Soc.*, 2002, **124**, 11230–11231.
34 (*a*) A. M. Leconte, L. Chen and F. E. Romesberg, *J. Am. Chem. Soc.*, 2005, **127**, 12470–12471; (*b*) F. J. Ghadessy, N. Ramsay, F. Boudsocq, D. Loakes, A. Brown, S. Iwai, A. Vaisman, R. Woodgate and P. Holliger, *Nat. Biotechnol.*, 2004, **22**, 755–759.
35 S. J. Johnson and L. S. Beese, *Cell*, 2004, **116**, 803–816.
36 J. C. Morales and E. T. Kool, *Biochemistry*, 2000, **39**, 2626–2632.
37 (*a*) N. V. Hud and J. Plavec, *Biopolymers*, 2003, **69**, 144–159; (*b*) A. Maki, F. E. Brownewell, D. Liu and E. T. Kool, *Nucleic Acids Res.*, 2003, **31**, 1059–1066; (*c*) A. S. Maki, T. Kim and E. T. Kool, *Biochemistry*, 2004, **43**, 1102–1110.
38 (*a*) T. W. Kirby, E. F. DeRose, W. A. Beard, S. H. Wilson and R. E. London, *Biochemistry*, 2005, **44**, 15230–15237; (*b*) C. L. Chepanoske, C. R. Langelier, N. H. Chmiel and S. S. David, *Org. Lett.*, 2000, **2**, 1341–1344; (*c*) A. W. Francis, S. A. Helquist, E. T. Kool and S. S. David, *J. Am. Chem. Soc.*, 2003, **125**, 16235–16242; (*d*) K. R. Fox, S. L. Allinson, H. Sahagun-Krause and T. Brown, *Nucleic Acids Res.*, 2000, **28**, 2535–2540; (*e*) A. C. Vallur, J. A. Feller, C. W. Abner, R. K. Tran and L. B. Bloom, *J. Biol. Chem.*, 2002, **277**, 31673–31678; (*f*) M. J. Schofield, F. E. Brownewell, S. Nayak, C. Du, E. T. Kool and P. Hsieh, *J. Biol. Chem.*, 2001, **276**, 45505–45508; (*g*) K. Drotschmann, W. Yang, F. E. Brownewell, E. T. Kool and T. A. Kunkel, *J. Biol. Chem.*, 2001, **276**, 46625–46629; (*h*) T. J. Begley, B. J. Haas, J. C. Morales, E. T. Kool and R. P. Cunningham, *DNA Repair*, 2002, **2**, 107–120.
39 A. S. Al-Madhoun, S. Eriksson, Z.-X. Wang, E. Naimi, E. E. Knaus and L. I. Wiebe, *Nucleosides Nucleotides Nucleic Acids*, 2004, **23**, 1865–1874.
40 (*a*) X.-F. Ren, B. A. Schweitzer, C. J. Sheils and E. T. Kool, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 743–746; (*b*) S. Rumney, IV and E. T. Kool, *J. Am. Chem. Soc.*, 1995, **117**, 5635–5646.
41 S. Moran, R. X.-F. Ren, C. J. Sheils, S. Rumney, IV and E. T. Kool, *Nucleic Acids Res.*, 1996, **24**, 2044–2052.
42 (*a*) K. M. Guckian, PhD Thesis, 1999, University of Rochester; (*b*) M. Zacharias and J. W. Engels, *Nucleic Acids Res.*, 2004, **32**, 6304–6311; (*c*) J. Parsch and J. W. Engels, *Helv. Chim. Acta*, 2000, **83**, 1791–1808.

43 P. Liu, A. Burdzy and L. C. Sowers, *Nucleic Acids Res.*, 2004, **32**, 4503–4511.
44 (*a*) J. Xia, A. Noronha, I. Toudjarska, F. Li, A. Akinc, R. Braich, M. Frank-Kamenetsky, K. G. Rajeev, M. Egli and M. Manoharan, *ACS Chem. Biol.*, 2006, **1**, 176–183; (*b*) A. Somoza, J. Chelliserrykattil and E. T. Kool, *Angew. Chem., Int. Ed.*, 2006 in press.
45 T. Mitsui, A. Kitamura, M. Kimoto, T. To, A. Sato, I. Hirao and S. Yokoyama, *J. Am. Chem. Soc.*, 2003, **125**, 5298–5307.
46 (*a*) T. Mitsui, M. Kimoto, A. Sato, S. Yokoyama and I. Hirao, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 4515–4518; (*b*) S. Matsuda and F. E. Romesberg, *J. Am. Chem. Soc.*, 2004, **126**, 14419–14427; (*c*) L. Zhang and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 74–75; (*d*) G. Mathis and J. Hunziker, *Angew. Chem., Int. Ed.*, 2002, **41**, 3203–3205; (*e*) A. Zahn, C. Brotschi and C. J. Leumann, *Chem.–Eur. J.*, 2005, **11**, 2125–2129; (*f*) J. S. Lai and E. T. Kool, *J. Am. Chem. Soc.*, 2004, **126**, 3040–3041.
47 A. Joy, G. Guler, S. Ahmed, L. W. McLaughlin and G. B. Schuster, *Faraday Discuss.*, 2006, **131**, 357–365.
48 (*a*) N. Shibata, B. K. Das, H. Honjo and Y. Takeuchi, *J. Chem. Soc., Perkin Trans. 1*, 2001, **14**, 1605–1611; (*b*) H. Challa and S. A. Woski, *Tetrahedron Lett.*, 1999, **40**, 419–422; (*c*) B. K. Das, N. Shibata and Y. Takeuchi, *J. Chem. Soc., Perkin Trans. 1*, 2002, 197–206.
49 T. S. Kuhn, in *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, 3rd edn., 1996.
50 J. M. Berg, J. L. Tymoczko and L. Stryer, in *Biochemistry*, W. H. Freeman, New York, 5th edn., 2002, p. 751.