Bi-stranded, multisite replication of a base pair between difluorotoluene and adenine: confirmation by 'inverse' sequencing

Dongyu Liu, Sean Moran and Eric T Kool

Background: The nonpolar nucleoside of difluorotoluene (F) was previously found to behave similarly to thymidine in single-site deoxynucleoside triphosphate (dNTP) insertion experiments with the Klenow fragment (KF) of DNA polymerase I. Further study was needed, first to see whether F–A base pairs could be replicated in more than one sequence context; second to investigate whether specific base pair replication occurs in the presence of four dNTPs; and third to confirm the presence of F in a replicated DNA strand.

Results: A primer bound to a template strand containing eight F residues was extended by KF using the four natural dNTPs at 20 μM. Similarly, the complement (containing eight adenines) was extended using dATP, dGTP, dCTP and dFTP. Comparison of the new strands to authentic strands using standard and 'inverse' chemical sequencing showed identical composition within ± 5%.

Conclusions: The results confirm that F in a template strand encodes the insertion of dATP and that adenine in a template encodes the insertion of dFTP with good specificity in at least six different nearest neighbor contexts. The results confirm that analog F behaves similarly to thymidine despite its poor hydrogen-bonding ability.

Address: Department of Chemistry, and Department of Biochemistry and Biophysics, University of Rochester, Rochester, NY 14627, USA.

Correspondence: Eric T Kool
E-mail: etk@etk.chem.rochester.edu

Key words: hydrogen bonding, Klenow fragment, Maxam–Gilbert chemistry, nonpolar isostere

Received: **1 August 1997**
Revisions requested: **26 August 1997**
Revisions received: **8 October 1997**
Accepted: **13 October 1997**

Chemistry & Biology December 1997, **4**:919–926
http://biomednet.com/elecref/1074552100400919

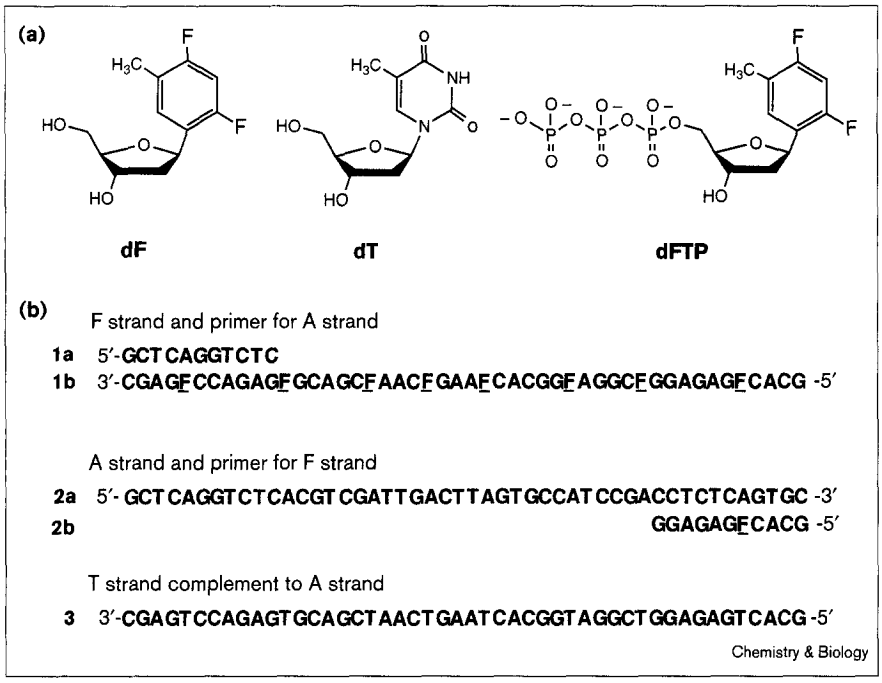
Introduction

Structurally modified nucleoside analogs have proven useful in biophysical studies which aim to pinpoint specific interactions with sugar or base that give rise to a given function or activity. Among such analogs, ones in which the hydrogen-bonding groups on the base have been blocked or deleted have been widely used to examine the importance of such noncovalent interactions. For example, nucleosides lacking one or two of their hydrogen-bonding groups have been used recently to measure the contributions of individual hydrogen bonds to the stability of folded RNA structure [1]. Examples of the bases used in such studies are N7-deazapurines, purine, inosine, and related structures. These or similar analogs have also been used in a number of studies to investigate the involvement of specific interactions in the mechanism of ribozyme catalysis [2–4]. In a different application, DNA nucleoside analogs have been widely useful in probing specific contacts made between proteins and the recognition sites in DNA [5,6]. In addition, a number of such analogs have also been examined as substrates for polymerase enzymes in order to investigate the importance of specific interactions in the mechanism of DNA and RNA synthesis [7–11].

We have proposed a series of nucleoside analogs in which steric shape is conserved as closely as possible, but in which polar functionality is removed [12]. One of these, the deoxynucleoside of difluorotoluene (dF, Figure 1), is a nearly perfect shape mimic of thymidine nucleoside [13]. It has been studied recently to probe the importance of hydrogen bonds in DNA replication [14]. One preliminary study involving single-nucleotide insertions with the Klenow fragment of *Escherichia coli* DNA polymerase I showed that F can serve in a DNA template as an efficient and selective encoder of adenine. Conversely, we have also observed, again in single-nucleotide insertion studies, that the nucleoside triphosphate derivative dFTP is inserted selectively opposite adenine in a template [15]. These early results led us to ask whether multiple substitutions of thymine by F could be tolerated in synthesis of duplex DNAs. This would require that not only insertion but also extension of A–F base pairs (in both directions) be reasonably efficient and selective. It also requires that the replication of this pair proceed well in a number of sequence contexts other than the single context previously studied; this is important because sequence context can play a considerable role in the efficiency and specificity of replication of base pairs [16]. Finally, we wished to develop a method to unambiguously identify the presence of F in a strand of DNA after replication; since it apparently behaves like thymine, standard dideoxynucleotide sequencing procedures would not distinguish the two.

We now report on studies aimed at answering these new questions. We find that the Klenow fragment of DNA

Figure 1



Nucleoside structures and DNA sequences in this study. **(a)** Structure of difluorotoluene nucleoside (dF) next to thymine (T) for comparison, and structure of the triphosphate analog dFTP. **(b)** Sequences of the DNAs used in the replication studies.

polymerase I (KF) can in fact synthesize strands of DNA in which thymine is completely replaced by difluorotoluene and can correctly synthesize the complement of such strands as well. We describe the efficiencies and limitations of replication of the F–A pair with this enzyme, and we also report on a new method developed to sequence DNAs containing F.

Results

Design of sequences

A 48 base pair duplex containing eight F–A pairs was designed to test the ability of DNA polymerase to replicate this non-natural base pair (Figure 1b). It was designed such that in one strand (the 'F' strand, Figure 1b), all thymines are replaced by difluorotoluene. This strand contains eight difluorotoluene (F) bases in six different nearest neighbor contexts. An 11-nucleotide oligodeoxynucleotide (**1a**) complementary to the 3′ end of this 'F' strand was designed to serve as primer to test the insertion of adenines opposite these F residues. Conversely, the complementary 48-mer sequence **2a** (the 'A' strand of the duplex), which contains eight adenines, was designed to test whether the nucleoside triphosphate analog dFTP (Figure 1a) can replace thymidine triphosphate (TTP) for incorporation opposite the eight adenines. In this case, an 11-nucleotide primer (**2b**) containing one F was prepared to be complementary to the 3′ end of sequence **2a**. If successful, polymerase extension would generate sequence **1b** from primer **2b**, and sequence **2a** from primer **1a**. Authentic strands of this duplex **1b–2a** were made by automated solid-phase DNA synthesis to use as polymerase templates and as controls for sequencing reactions. Another 48-nucleotide sequence (**3**) in which all the F residues in **1b** ('F' strand) were replaced by thymine (T) was synthesized for comparison with the 'F' strand in enzymatic DNA synthesis.

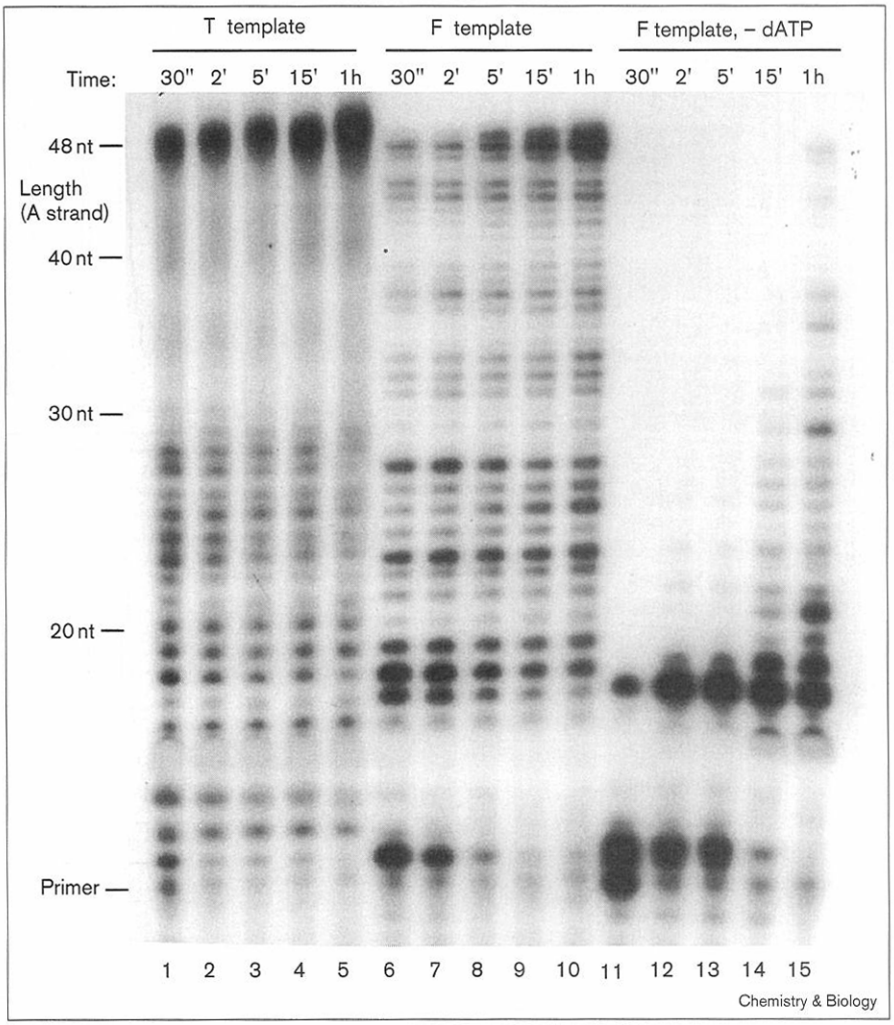
Primer extension on 'F'-containing template strand

We first tested the polymerase reaction with 'F' strand **1b** as template using the KF (exo$^-$ mutant) and primer **1a**. The results are shown in Figure 2 in comparison to synthesis on a template (**3**) containing thymines. The enzymatic DNA synthesis was performed under non-forcing conditions, with 5′-^{32}P-labeled primer and template at 5 μM and with 20 μM each of the four natural nucleotide triphosphates. The time course results show that replacement of T by F in the template has relatively little apparent effect on DNA synthesis (Figure 2, compare lanes 1–5 and 6–10). Some full-length products are seen at 30s in both cases, although accumulation of full-length product is approximately 10-fold slower with the 'F' template. Brief pauses are observed on elongation of putative F–A pairs using the 'F' template. By comparison, control experiments using only dCTP, dGTP and TTP show little or no extension until 1h, serving as indirect evidence that dATP is inserted opposite the F sites in the template.

Sequencing of the extended products using modified Maxam–Gilbert methods [17] was carried out to identify the nucleotides inserted opposite F sites, and densitometry was used to aid comparison of enzymatically synthesized and authentic strands. As shown in Figure 3, an A-specific

Figure 2

Time course of KF replication of the F-containing template strand **1b** (see Figure 1) (lanes 6–10) using the four natural dNTPs (20 μM each) and primer **1a**, in comparison to naturally substituted template **3** (lanes 1–5). Lanes 11–15 show the effect of withholding dATP from the reaction. Band lengths can be identified by counting upward from the primer band (see length indicators at left). *nt, nucleotides.*

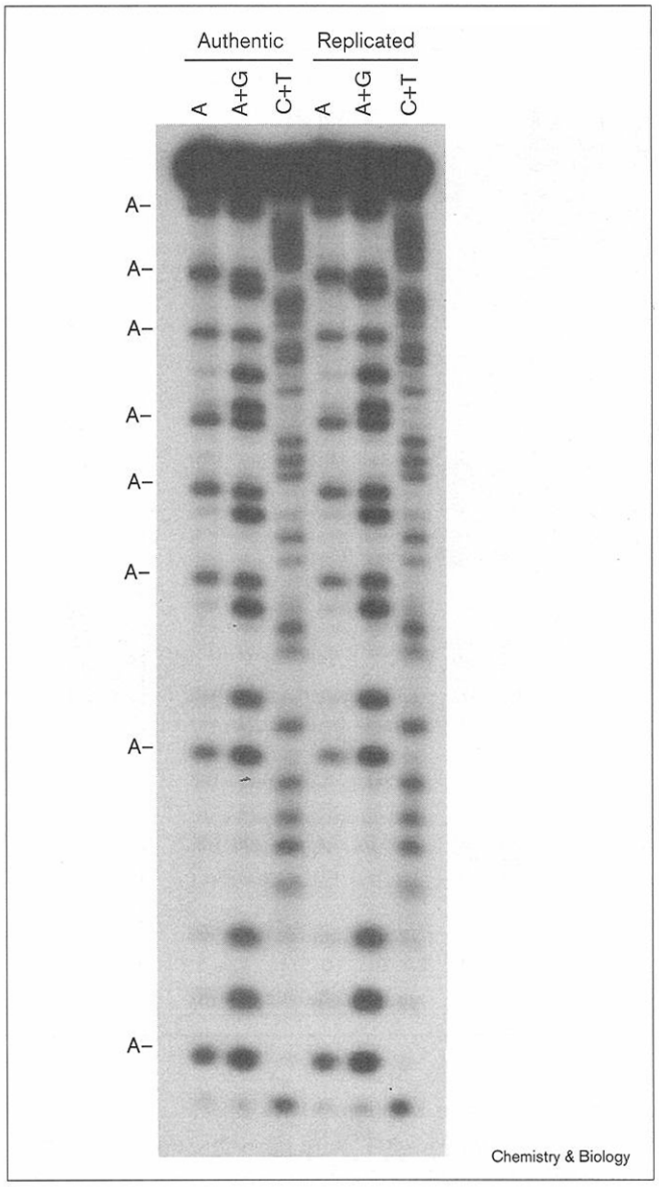


reaction [18] carried out on the primer-extended full-length product excised from a gel confirmed the existence of eight adenines in the newly synthesized strand as predicted, and with no qualitative difference observable with the authentic strand. The sequencing reactions also confirm that both strands are otherwise identical as well. Densitometric analysis of the A reaction lanes shows (Figure 4) that at all eight adenine positions there is no measurable difference between the enzymatically synthesized strand and the authentic strand within ± 5%, which we estimate as the error limits of the experiment. Thus the results indicate that analog F in a template strand encodes adenine with at least good specificity in at least six different nearest neighbor contexts.

Primer extension on 'A'-containing template

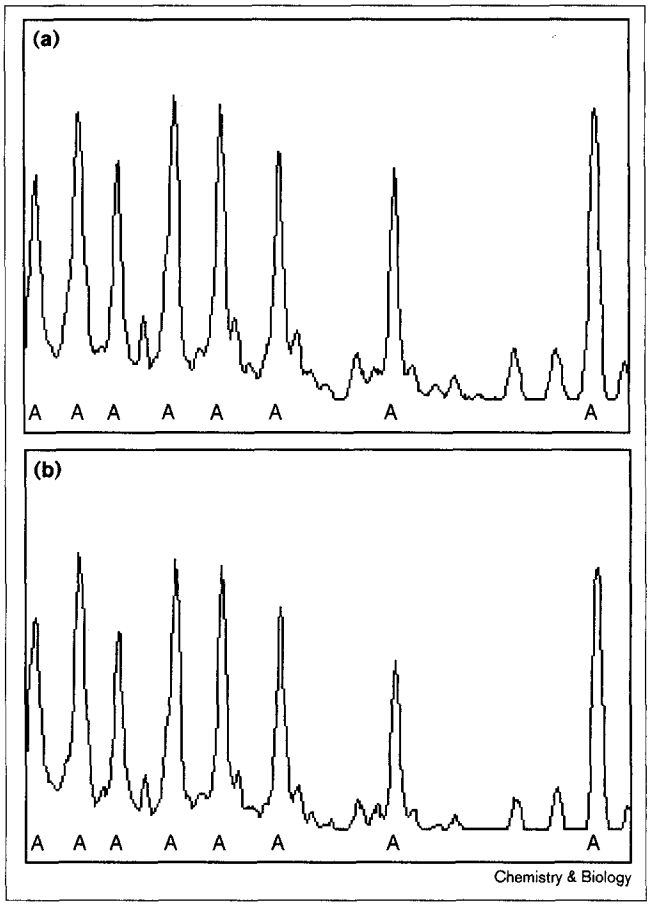
To test whether the deoxynucleoside triphosphate derivative of F (dFTP) can successfully replace TTP in enzymatic full-length DNA synthesis, we carried out analogous experiments using the 'A' strand (**1b**) as template. This was also done using KF (exo$^-$) with dATP, dCTP, dGTP and dFTP at 20 μM each. The reaction time course was monitored from 30 min to 4 h (Figure 5). In this case, although some fully extended products are observed in 30 min, the full-length 48-nucleotide strand (identified by comparison to an authentic marker) does not become the predominant product until 2 h (lanes 6–10). The incorporation of dFTP opposite A appeared to be selective, as similar elongation was not observed in the absence of dFTP (Figure 5, lanes 11–15). Overall, full-length extension was slow relative to when TTP replaces dFTP, in agreement with single-nucleotide insertion experiments showing that efficiency with dFTP is lower than with TTP [15]. On the basis of the positions of paused bands it appears that the slowest steps are the extension (rather than the formation) of F–A pairs, with significant differences depending on context. We also observed somewhat faster gel mobility of the putative F-containing strand relative to the naturally substituted one; this was previously seen in single-nucleotide extension experiments as well.

Figure 3



Chemical sequencing of the strand **2a** replicated from template **1b** (see Figure 1), and comparison with an authentic DNA sequence.

Figure 4



Densitometric analysis showing **(a)** the composition of replicated A strand **2a** (see Figure 1) in comparison to **(b)** an authentic A strand. Data are from A sequencing reactions from lanes 1 and 4 of the autoradiogram in Figure 3.

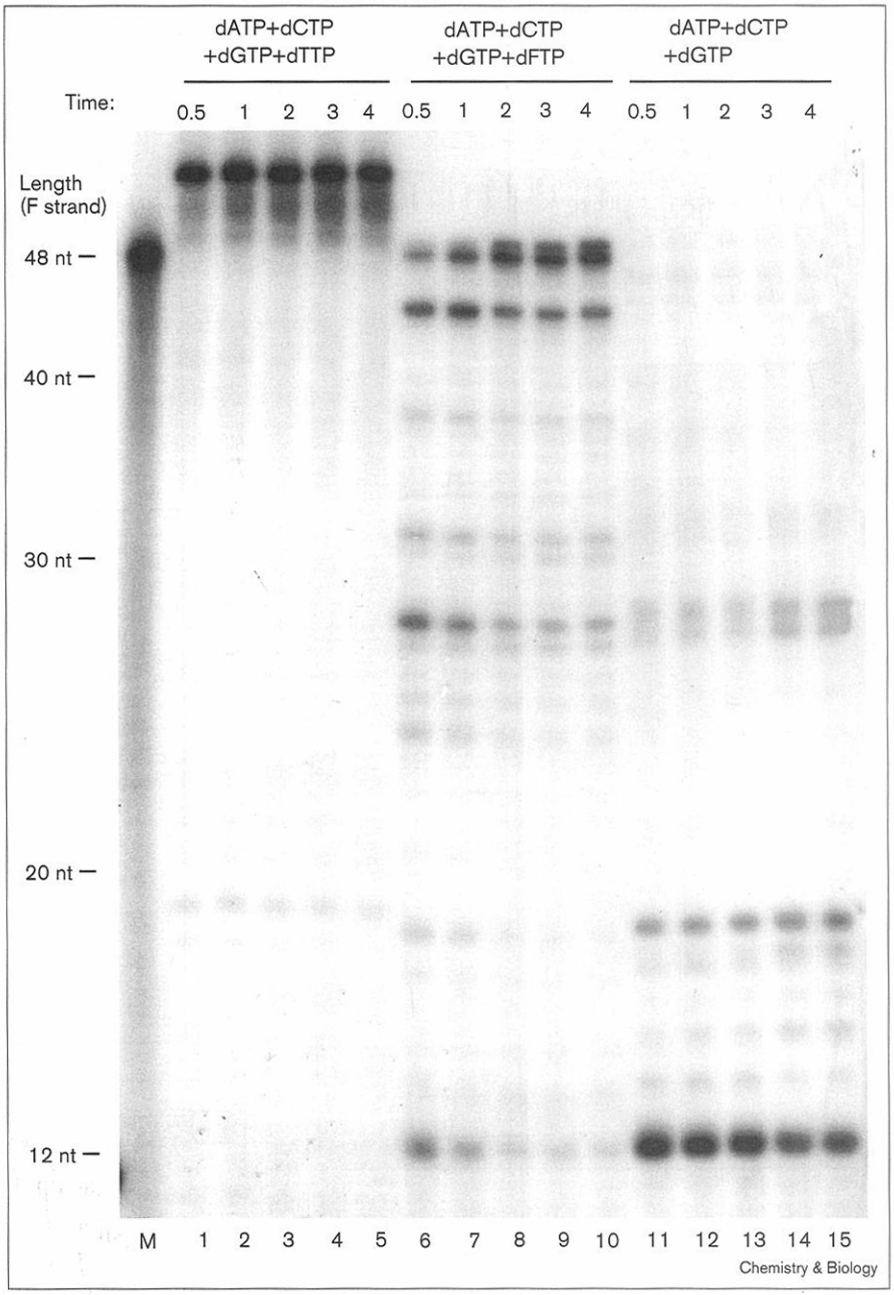
In initial experiments, we also tested a template sequence containing non-isolated adenines. An 81-mer containing an internal segment of sequence -GGTAAGCGAATCT- (along with ten other adenine sites) was used as a template for primer extension with dATP, dCTP, dGTP and dFTP as above (data not shown). Results showed that with these adjacent adenines there was a strong pause on elongation of the first F–A pair to the second at the first -AA- site, and only after extended times (12 h) was a slight amount (~1%) of extension beyond these two -AA- sites seen. At that time point less than 1% of the radiolabel was associated with full-length DNA.

'Inverse' chemical sequencing method

To confirm unambiguously the incorporation of dFTP in the enzymatically synthesized 48-mer strand we developed a chemical sequencing strategy. This approach takes advantage of the fact that hydrazine-mediated cleavage of C and T in the Maxam–Gilbert reaction will not work for difluorotoluene. Thus the base analog F will not be cleaved by the C + T reaction after piperidine treatment and would be expected to give a blank band on the sequencing gel. By running the A + G and C + T reactions in one lane on gel electrophoresis, all natural bases will show up as cleaved bands, leaving blank spots indicating F residues. With this 'inverse' chemical sequencing one can easily identify the positions of the base analog F in the DNA strand. This is shown in Figures 6 and 7, which compare results for the enzymatically synthesized strand with an authentic strand containing eight F residues. The data show that both strands have identical composition within the limits of detection by densitometry (Figure 7),

Figure 5

A time course of KF replication of the A-containing template strand **2a** (see Figure 1) with substitution of dFTP for dTTP. Lanes 1–5 have the four natural dNTPs (20 μM each). Lanes 6–10 have dFTP replacing dTTP. Lanes 11–15 show the effect of withholding dATP from the reaction. The time course scale is in hours. The marker lane (M) shows the mobility of a full-length authentic strand containing eight F residues.



thus independently confirming that the dFTP analog is specifically inserted opposite adenines in the template.

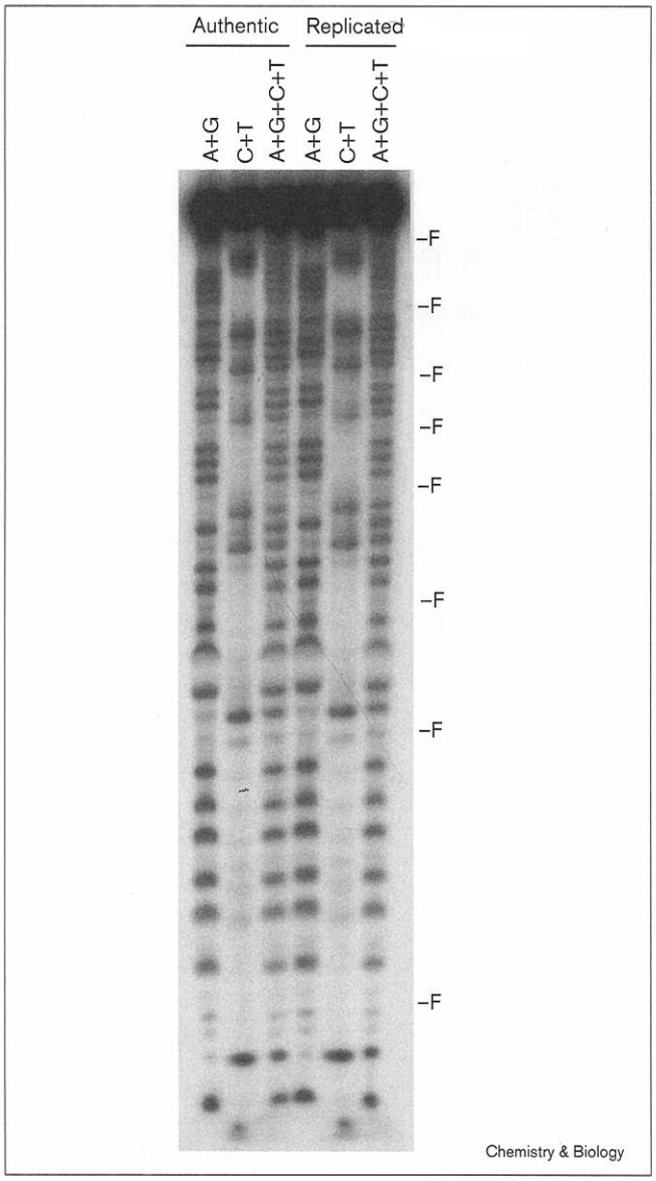
Thermal melting studies

To investigate the effects of the eight F–A pairs on stability of the 48 base pair duplex we carried out thermal denaturation experiments with the modified duplex and with the natural 48 base pair duplex containing T–A pairs at these positions (conditions: pH 7.0, 10 mM Mg^{2+}, 100 mM Na^{+}). The results show that, consistent with previous results, the eight difluorotoluene residues cause significant destabilization. The T_m measured for the modified sequence was 64.9°C, whereas that for the unmodified duplex was 80.9°C.

Discussion

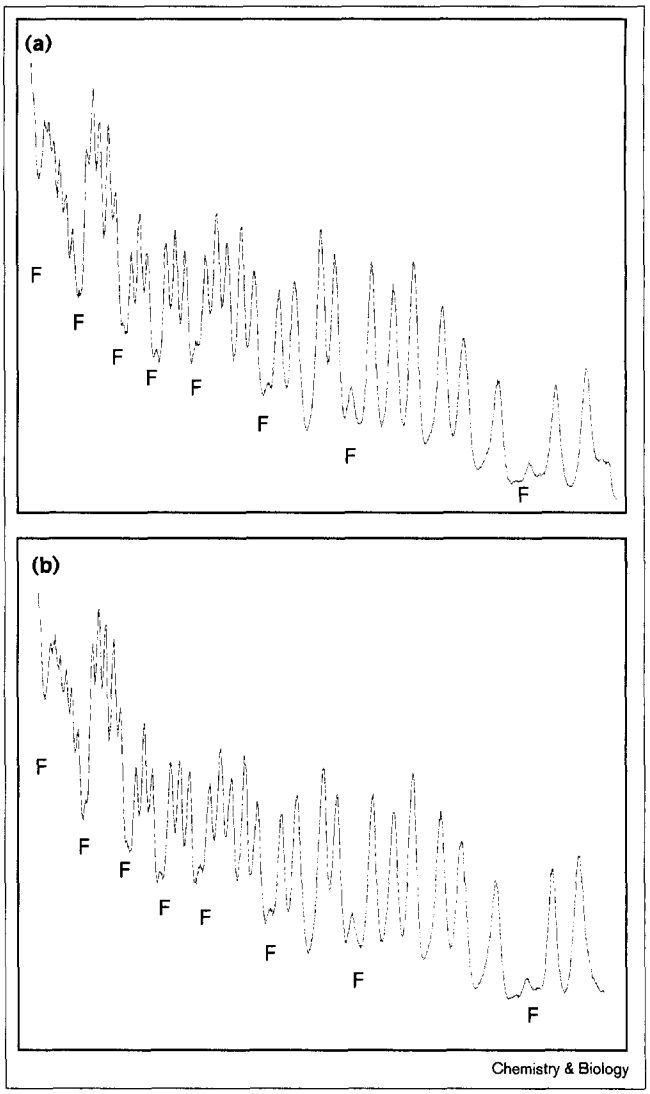
Our results establish that the F–A base pair can, to a reasonable extent, replace the T–A pair in replication of duplex DNA by the KF polymerase. Within error limits of the chemical sequencing methods, we find that the fidelity of the A–F pairing is essentially as good as that of other natural base pairs. Of course, accurate measures of fidelity require much more sensitivity than chemical sequencing because of the small error rates for most polymerases [19].

Figure 6



'Inverse' Maxam–Gilbert sequencing of the F-containing strand **1b** (see Figure 1) replicated from template **2a**, and comparison with an authentic DNA sequence.

Figure 7



The analysis of F-containing strands by phosphorimaging densitometry. **(a)** The strand synthesized by KF polymerase containing eight F residues, showing 'inverse' sequencing, wherein F residues are denoted by missing peaks. **(b)** An authentic DNA strand of the same sequence produced by a DNA synthesizer for comparison.

Our previous quantitative studies indicate that that fidelity of F–A replication is only lower than T–A by a small degree (about twofold) [14]. For most of our envisioned applications the fidelity of F–A pairing is for practical purposes the same as T–A pairing. The main limitation appears to be the poor enzymatic insertion of F opposite two consecutive A nucleotides, which appears to be unfavorable enough to cause a relatively strong stall, at least for this polymerase and under these conditions.

It remains to be seen how general the ability to process the F analog is for various polymerases. In the context of the templates described here we have also tested the KF enzyme that does not lack a 3′ exonuclease, and we found similar results but with somewhat slower synthesis both for natural and non-natural nucleotides, as expected (data not shown). We also examined a T7 DNA polymerase mutant lacking 3′ exonuclease activity and found similar processing of the F–A pair.

Our early findings of high fidelity for replication of the F–A pair [14] were surprising in face of the widely held belief that Watson–Crick complementarity and the specificity of hydrogen bonding are the chief mechanisms responsible for fidelity. Our quantitative evidence supported the conclusion that this hydrogen bonding may not be nearly as

important as once believed, and our current hypothesis is that shape complementarity may play a more direct role. The present results provide unambiguous evidence that dFTP is inserted and that the F–A pair is viable in replication. It is interesting to note the apparent efficiency of replication of F–A pairs (especially for insertion of adenine opposite F) despite the fact that the DNA being synthesized is considerably less stable than DNA containing analogous T–A pairs. We have shown that the pairing of F in short synthetic DNAs is inherently nonselective and is energetically destabilizing [15,20], consistent with the present thermal denaturation data. It is clear that the destabilization of the F–A pair must be realized after the transition state for the rate-limiting step for synthesis of the pair; otherwise replication of the pair would be practically impossible. This of course does not imply that the rate-limiting steps are the same for T–A pairs and F–A pairs, but merely that they are not greatly different in activation energy.

We have also demonstrated a potentially useful method for identification of F in the context of the four natural bases. Common polymerase-based sequencing approaches such as the Sanger method would not be able to distinguish thymine and difluorotoluene, since KF (and T7 DNA polymerase as well; S.M., D.L. and E.T.K., unpublished data) inserts adenine opposite either one. Thus, the development of a chemical method for sequencing DNAs containing the analog F increases the utility of F in a number of possible applications.

In one possible application we envisage the use of F as a probe for hydrophobic and hydrophilic contacts between proteins and DNAs. It could be inserted into shorter strands synthetically, or could potentially be incorporated enzymatically into DNA strands too long to be readily constructed by automated synthesizer. Additionally, mixtures of F and T could be incorporated at given positions in a sequence (for example, a protein-binding site) to construct combinatorial libraries which might be useful for probing hydrophobic or hydrophilic contacts. Our inverse sequencing approach could be used to identify which positions preferentially contain F or T (or some fraction of both). Finally, F may be generally useful for investigating the importance of the polar hydrogen-bonding groups of thymine in many biomolecular complexes involving DNA or nucleotides.

Significance

Difluorotoluene (F), a nonpolar isostere of thymidine (T), is useful for evaluating the importance of hydrogen bonding in biomolecular interactions. We have shown that a 48-mer strand containing eight F nucleosides can be replicated efficiently in the presence of KF (exo-) and the four natural dNTPs. Standard Maxam–Gilbert sequencing has established that the product strand is identical to an authentic version and thus contains only adenine at the eight sites encoded by F. Conversely, the complementary adenine-containing strand was used as a template for the uptake of dFTP (the nucleoside triphosphate derivative of F) at the eight possible sites. An 'inverse' Maxam–Gilbert chemical sequencing method was used to establish that F is indeed present in the new strand at the eight sites opposite adenines. Some brief pausing after single dFTP incorporation is observed, and sequential incorporation of two dFTPs causes a strong pause. The results establish: first, that replication involving F succeeds in at least six different sequence contexts; second, that chemical methods can be used to sequence highly unreactive nucleotides; and finally, that hydrogen bonding between bases may be less important in DNA replication than previously believed.

Materials and methods

Oligonucleotides

Oligodeoxynucleotides were synthesized on an Applied Biosystems 392 DNA synthesizer using standard β-cyanoethyl phosphoramidite chemistry. Non-natural nucleoside 1′–2,4-difluorotoluene deoxyriboside phosphoramidite was synthesized as described previously [21] and was incorporated using the standard ABI coupling cycle; stepwise coupling yields were all greater than 95% as determined by trityl cation monitoring. Oligonucleotides were deprotected in concentrated NH_4OH (55°C, 12 h). After lyophilization the DNA was purified by preparative denaturing polyacrylamide gel electrophoresis, and isolated by excision, crushing and eluting into 0.2 M NaCl. The salts were removed by dialysis against distilled deionized water, and the DNA was quantitated by absorbance at 260 nm, using the nearest neighbor method to calculate molar extinction coefficients.

Synthesis of 1′,2′-dideoxy-1′-(2,4-difluorotolyl)-(-D-2′-deoxyribofuranose 5′-triphosphate (dFTP)

Ten milligrams (50 mmol) of the dF nucleoside [21] was dissolved in 250 ml of trimethylphosphate (Aldrich). The solution was cooled to 0°C. Proton Sponge (16 mg, Aldrich) and $POCl_3$ were added and the solution was stirred at 0°C for 2 h. Tributylamine (750 ml) and tributylammonium pyrophosphate (Sigma; 42.5 mg in 500 ml dimethylformamide (DMF) were added. After 1 min, triethylammonium bicarbonate (5 ml of 1 M stock at pH 7.5) was added and stirred at room temperature for 20 min. The reaction mixture was purified by liquid chromatography (Sephadex A25) using a gradient of triethylammonium bicarbonate (pH 7.5) from 0.1–1 M. The yield was estimated by spectrophotometric analysis to be 7.5 mmol (15%) using an extinction coefficient of $\varepsilon = 1200$ at 270 nm. ^{31}P NMR (400 MHz; D_2O referenced to phosphoric acid): δ −5.5(doublet); −10.5(doublet); −21.5 (triplet). UV(H_2O): $\lambda_{max} = 270, 280$ nm.

Template-directed primer extension

Primers **1a** and **2b** were used in combination with template strands **1b** and **2a**, respectively (Figure 1). Primer was (1 ml, 50 μM) end-labeled in a 20 μl volume by adding 2 μl 10× polynucleotide kinase buffer (700 mM Tris-HCl (pH 7.0), 100 mM $MgCl_2$, 50 mM dithiothreitol (DTT), 5 U T4 polynucleotide kinase, 5 μl 6000 Ci/mmol [γ-^{32}P]ATP, and 11 μl distilled deionized H_2O. This mixture was incubated at 37°C for 45 min, heated at 68°C for 5 min, allowed to cool to room temperature, microcentrifuged briefly, and then used directly in the annealing step. In separate reactions, 2 μl of a 50 μM stock of template strand **1b** or **2a**, 2 μl of 50 μM stock of primer **1a** or **2b**, and 2 μl of the 5′-^{32}P-end-labeled primer, were combined with 2μl of 10× Klenow fill-in buffer (0.5M Tris-HCl, pH 7.5, 0.1M $MgCl_2$, 10 mM DTT 0.5 mg/ml bovine serum albumin (BSA), and 9 μl of distilled deionized H_2O was added. The mixture was heated to 70°C for 5 min and then slowly cooled to room temperature over a period of 1 h. The desired dNTPs (2 μl of 2 mM stock) were added, and 1 μl (2.5 units) of KF (exo– enzyme;

Amersham) was added to the annealing mixture, and incubated at 37°C. Aliquots (3 µl) were removed at desired time points and added to 3 µl gel loading buffer (80% formamide, 10 mM EDTA pH 8.0, 1 mg/ml xylene cyanol, 1 mg/ml bromphenol blue). The samples were heated for 5 min at 90°C, and aliquots (3–5 µl) were loaded onto a 10% polyacrylamide gel containing 8 M urea. Following electrophoresis (1500–1800 V), the gels were visualized by autoradiography.

Chemical sequencing

Prior to sequencing, authentic DNA strands **1b** and **2a** were 5′ end-labeled with ^{32}P, purified by 10% denaturing polyacrylamide gel electrophoresis and excised from the gel using a razor blade. The gel slices were crushed, vortex-mixed overnight in a Spin-X 0.22 µm cellulose acetate centrifuge tube (Costar) in 400 µl sodium acetate (pH 5.2), and precipitated at −70°C by the addition of 3 vol of 100% ethanol. The pellet obtained after microcentrifugation at 14,000 rpm for 10 min at 4°C was washed with 70% ethanol and dried by lyophilization. Then the samples were redissolved in 100 µl of 1 M aqueous piperidine. This solution was heated at 90°C for 30 min, cooled to room temperature, and microcentrifuged briefly. Samples were lyophilized, twice resuspended in 20 µl H_2O and relyophilized, and then redissolved in 20 µl H_2O for A + G or C + T sequencing reactions according to the standard protocol [22]. DNA strands **1b** and **2a** prepared from ^{32}P-labeled primer extension by Klenow enzyme (exo$^-$) were excised from the gel and treated the same way as above except without piperidine treatment prior to Maxam–Gilbert sequencing. The adenine-specific sequencing reaction was carried out according to published methods [18].

Acknowledgements

We are grateful to the National Institutes of Health (GM52956) for support of this work.

References

1. SantaLucia, J., Kierzek, R. & Turner, D.H. (1992). Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin. *Science* **256**, 217-219.
2. Slim, G., Pritchard, C., Biala, E., Asseline, U. & Gait, M.J. (1991). Synthesis of site-specifically modified oligoribonucleotides for studies of the recognition of TAR RNA by HIV-1 tat protein and studies of hammerhead ribozymes. *Nucleic Acids Symp. Ser.* **24**, 55-58.
3. Strobel, S.A. & Cech, T.R. (1996). Exocyclic amine of the conserved G.U pair at the cleavage site of the *Tetrahymena* ribozyme contributes to 5′-splice site selection and transition state stabilization. *Biochemistry* **35**, 1201-1211.
4. Tuschl, T., Ng, M.M., Pieken, W., Benseler, F. & Eckstein, F. (1993). Importance of exocyclic base functional groups of central core guanosines for hammerhead ribozyme activity. *Biochemistry* **32**, 11658-11668.
5. Lesser, D.R., Kurpiewski, M.R. & Jen-Jacobson, L. (1990). The energetic basis of specificity in the EcoRI endonuclease–DNA interaction. *Science* **250**, 776-786.
6. Lever, C., Li, X. & Cosstick, R. (1993). Thermodynamic stability and drug-binding properties of oligodeoxyribonucleotide duplexes containing 3-deazaadenine:thymine base pairs. *Nucleic Acids Res.* **21**, 1743-1746.
7. Singer, B., Spengler, S.J., Fraenkel-Conrat, H., & Kusmierek, J.T. (1986). O4-Methyl, -ethyl, or -isopropyl substituents on thymidine in poly(dA-dT) all lead to transitions upon replication. *Proc. Natl Acad. Sci. USA* **83**, 28-32.
8. Strazewski, P. & Tamm, C. (1990). Replication experiments with nucleotide base analogues. *Angew. Chem. Int. Ed. Engl.* **29**, 36-57.
9. Freese, E. (1959). The difference between spontaneous and base analogue induced mutations of phage T4. *Proc. Natl Acad. Sci. USA* **45**, 622-633.
10. Yu, H., Eritja, R., Bloom, L.B. & Goodman, M.F. (1993). Ionization of bromouracil and fluorouracil stimulates base mispairing frequencies with guanine. *J. Biol. Chem.* **268**, 15935-15943.
11. Horlacher, J., Hottiger, M., Podust, V.N., Hubscher, U. & Benner, S.A. (1995). Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. *Proc. Natl Acad. Sci. USA* **92**, 6329-6333.
12. Schweitzer, B.A. & Kool, E.T. (1994). Nonpolar aromatic nucleosides as hydrophobic isosteres of DNA nucleosides. *J. Org. Chem.* **59**, 7238-7242.
13. Guckian, K.M. & Kool, E.T. (1997). Highly precise shape mimicry by difluorotoluene deoxynucleoside, a replication-competent substitute for thymidine. *Angew. Chem.*, in press.
14. Moran, S., Ren, R.X.F., Rumney, S. & Kool, E.T. (1997). Difluorotoluene, a nonpolar isostere of thymine, codes specifically and efficiently for adenine in DNA replication. *J. Am. Chem. Soc.* **119**, 2056-2057.
15. Moran, S., Ren, R.X.F. & Kool, E.T. (1997). A thymidine triphosphate shape mimic lacking Watson–Crick pairing ability is replicated with high specificity. *Proc. Natl Acad. Sci. USA*, **94**, 10506-10511.
16. Petruska, J. & Goodman, M.F. (1985). Influence of neighboring bases on DNA polymerase insertion and proofreading fidelity. *J. Biol. Chem.* **260**, 7533-7539.
17. Maxam, A.M. & Gilbert, W. (1977). A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA* **74**, 560-564.
18. Iverson, B.L. & Dervan, P.B. (1987). Adenine specific DNA sequencing reaction. *Nucleic Acids Res.* **15**, 7823-7830.
19. Kornberg, A & Baker, T.A. (1992). *DNA Replication.* (2nd edn.), p. 113, W.H. Freeman, New York.
20. Schweitzer, B.A. & Kool, E.T. (1995). Hydrophobic, non-hydrogen-bonding bases and base pairs in DNA. *J. Am. Chem. Soc.* **117**, 1863-1872.
21. Ren, X.-F., Schweitzer, B.A., Sheils, C.J. & Kool, E.T. (1996). Stable DNA loop formation by hydrophobic nucleoside isosteres. *Angew. Chem.* **108**, 834-837; *Angew. Chem. Int. Ed. Engl.* **35**, 743-746.
22. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). In *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbour Laboratory Press, Cold Spring Harbour Laboratory, NY.