Ambivalent Incorporation of the Fluorescent Cytosine Analogues tC and tCo by Human DNA Polymerase α and Klenow Fragment

ARTICLE *in* BIOCHEMISTRY · AUGUST 2009

Impact Factor: 3.02 · DOI: 10.1021/bi9006995 · Source: PubMed

5 AUTHORS, INCLUDING:

Marcus Wilhelmsson
Chalmers University of Technology
56 PUBLICATIONS **1,622** CITATIONS
SEE PROFILE

Milan Urban
Palacký University of Olomouc
33 PUBLICATIONS **690** CITATIONS
SEE PROFILE

Robert Kuchta
University of Colorado Boulder
91 PUBLICATIONS **2,671** CITATIONS
SEE PROFILE

Published in final edited form as:
Biochemistry. 2009 August 11; 48(31): 7547–7555. doi:10.1021/bi9006995.

Ambivalent Incorporation of the Fluorescent Cytosine Analogues tC and tCo by Human DNA Polymerase α and Klenow Fragment

Gudrun Stengel[1], **Byron W. Purse**[2], **L. Marcus Wilhelmsson**[3], **Milan Urban**[1], and **Robert D. Kuchta**[1,*]

[1] Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 30309-0215

[2] Department of Chemistry and Biochemistry, University of Denver, Denver, CO 80208

[3] Department of Chemical and Biological Engineering/Physical Chemistry, Chalmers University of Technology, SE-41296 Gothenburg, Sweden

Abstract

We studied the incorporation of the fluorescent cytidine analogues 1, 3-diaza-2-oxo-phenothiazine (tC) and 1, 3-diaza-2-oxo-phenoxazine (tCo) by human DNA polymerase α and Klenow fragment of DNA polymerase I (*E. coli*). These tricyclic nucleobases possess the regular hydrogen bonding interface of cytosine but are significantly size expanded toward the major groove. Despite the size alteration both DNA polymerases insert dtCTP and dtCoTP with remarkable catalytic efficiency. Polymerization opposite guanine is comparable to the insertion of dCTP, while the insertion opposite adenine is only ~4-11 times less efficient than the formation of a T-A base pair. Both enzymes readily extend the formed tC(o)-G and tC(o)-A base pairs, and can incorporate at least 4 consecutive nucleotide analogues. Consistent with these results, both DNA polymerases efficiently polymerize dGTP and dATP when tC and tCo are in the template strand. KF inserts dGTP with a 4- to 9-fold higher probability than dATP, while pol α favors dGTP over dATP by a factor of 30-65. Overall, the properties of tC(o) as templating base and as incoming nucleotide are surprisingly symmetrical and may be universal for A and B family DNA polymerases. This finding suggests that the aptitude for ambivalent base pairing is a consequence of the electronic properties of tC(o).

Keywords

Fluorescent Base Analogue; DNA labeling; Kinetics; Nucleotide polymerization; DNA Replication

The ability of DNA polymerases to select and incorporate the correct dNTP helps define the fidelity of DNA replication. Depending on the base mismatch and the local sequence context, error rates vary between 10^{-3}–10^{-6} errors per nucleotide replicated for proof-reading deficient DNA polymerases (1). Proofreading and other cellular repair mechanisms reduce the actual error frequency to around 10^{-8}–10^{-11} errors per nucleotide replicated. While cell survival critically depends upon the maintenance of genome integrity, spontaneous base mutations remain important because they provide the molecular basis for evolution. This suggests nature may have developed a mechanism that ensures a constant, yet low, mutation rate.

Watson and Crick were the first to propose that the minor tautomeric forms of the natural bases may be responsible for the occurrence of base mutations (2). In this model, T and G base pair

#This work was supported by grants from the NIH (GM54194 and AI59764) to RDK.

*Email: kuchta@colorado.edu Phone: 303-492-7027. FAX: 303-492-5894.

via their keto and enol tautomeric forms while C and A form base pairs via their amino and imino tautomeric forms (3,4). The resulting base pairs would have shapes and sizes similar to the classical Watson-Crick base pairs and would not distort the DNA helix as seen for wobble base pairs. The importance of the rare tautomeric forms of the natural bases remains unproven owing to the technical challenge of accurately determining the tautomeric equilibrium constant that lies almost completely on the side of the amino (A, C) and keto (T, G) tautomers. X-ray structures of short DNA duplexes frequently capture pyrimidinepurine mismatches as (sometimes ionic) wobble base pairs despite the larger structural perturbation (5-7). However, NMR has confirmed C-A base pairs in Watson-Crick geometry involving the tautomeric forms of modified cytosine analogues (8-11). While X-ray structures of unmodified DNA emphasize the thermodynamic stability of wobble base pairs in aqueous environment, this finding does not exclude the possibility that tautomeric forms occur during nucleotide selection by DNA polymerases. Evidence for the relevance of the tautomeric equilibrium comes from studies using cytidine analogues with biased tautomeric equilibria as DNA polymerase substrates. Indeed, these compounds are incorporated ambivalently opposite G and A by several bacterial and eukaryotic DNA polymerases (12). Examples include N^4-amino-2′-deoxycytidine (13, 14), N^4-hydroxy-2′-deoxycytidine (15), N^4-methoxy-2′-deoxycytidine (16-18), 5-hydroxy-2′-deoxycytidine (19) and 3,4-dihydro-6H-8H-pyrimido[4,5-c][1,2]oxazin-2-one (P base) (20) (see Chart 1).

tC and tCo (1,3-diaza-2-oxo-phenothiazine and 1,3-diaza-2-oxo-phenoxazine, respectively (21)) [1] are two fluorescent cytosine analogues whose N^4 substituent is locked in the anti conformation relative to O^2 of cytosine by an expanded ring system (Chart 2). The analogues have fluorescence quantum yields of ~ 0.2 both in isolation and in duplex DNA and emission wavelengths of 465 and 505 nm, respectively, making them versatile biophysical probes for studies of DNA-protein interactions (22,23). Substituting a C-G base pair by a tC(o)-G base pair increases the melting temperature of a short oligonucleotide on average by 3 °C depending on the neighboring bases (24). This duplex stabilizing effect likely results from π-π interactions between the aromatic tricycle and the heterocycles of the canonical bases, possibly in conjunction with induced dipole effects. NMR studies indicate that tC-containing DNA duplexes adopt the B conformation with only small local distortions around the tC base (24).

We examined the incorporation of dtCTP and dtCoTP opposite all four natural bases and characterized the templating properties of the analogues using the Klenow fragment of DNA polymerase I (KF) and human DNA polymerase α (pol α). Both enzymes insert dtCTP and dtCoTP opposite G with the same efficiency as dCTP, and insert dtCTP and dtCoTP opposite A only 4-11 times less efficiently than dTTP. When tC(o) is located in the templating DNA strand, KF and pol α polymerize preferentially dGTP across the analogues. Two other representatives of the A and B DNA polymerase family, *Bacillus stearothermophilus* DNA polymerase (BF pol) and herpes simplex virus-1 DNA polymerase (HSV), gave qualitatively similar results. Thus, the base pairing properties of these analogues are intrinsic to the base and are not determined by the differing mechanisms A and B family enzymes use to distinguish between right and wrong dNTPs.

[1]Abbreviations used: BF pol, *Bacillus stearothermophilus* DNA polymerase; EDTA, ethylene diamine tetraacetic acid; HSV, herpes simplex virus-1; KF, Klenow Fragment of DNA polymerase I from *E. coli;* P, 3,4-dihydro-6H-8H-pyrimido[4,5-c][1,2]oxazin-2-one ; dPTP, 3,4-dihydro-6H-8Hpyrimido[4,5-c][1,2]oxazin-2-one-2′-deoxyribose-5′-triphosphate; Pol α, DNA polymerase α; tC, 1,3-diaza-2-oxo-phenothiazine-2′-deoxyribose; tCo, 1,3-diaza-2-oxo-phenoxazine-2′-deoxyribose; TrisHCl, Tris(hydroxymethyl)aminomethane, HCl salt .

EXPERIMENTAL SECTION

Materials and Enzymes

All reagents were of highest quality commercially available. Unlabeled dNTPs were from Invitrogen and ^{32}P-labeled (d)NTPs from Perkin Elmer Life Sciences. The 1,3-diaza-2-oxo-phenothiazine and 1,3-diaza-2-oxo-phenoxazine nucleosides were synthesized according to Sandin et al. and Matteucci et al., respectively (21,25). The synthesis of the tC and tCo triphosphates was performed following the procedure by Ludwig using phosphorus oxychloride and tetrabutylammonium pyrophosphate (26). DMTr protected phosphoramidites were synthesized using standard methods and the identity of the phosphoramidites was confirmed by ^{31}P NMR (Inova 400 MHz instrument). Established procedures were used for the synthesis of the oligonucleotides on an Applied Biosystems 394 automatic DNA synthesizer with the difference that the coupling times for both analogues was increased to 600 s (27,28). The concentration of the oligonucleotides was determined from their UV absorption at 260 nm. The extinction coefficients of the base analogue-containing oligonucleotides were calculated by combining the extinction coefficients of the individual nucleosides at 260 nm: (ε_{tC} = 13,500 M^{-1} cm^{-1}, ε_{tCo} = 11,000 M^{-1} cm^{-1}) (22,29). Base stacking interactions were accounted for by multiplying the result by a factor of 0.9.

E. coli DNA polymerase I Klenow fragment (exo$^-$) was purchased from New England Biolabs. *Bacillus stearothermophilus* DNA polymerase I (BF pol) was a generous gift from Lorena Beese (Duke University). Human pol α (N-His$_6$-p70-p180 complex), expressed in baculovirus-infected Sf 9 insect cells was grown at the Tissue Culture Core Facility at the University of Colorado Health Sciences Center, and purified by nickel nitrilotriacetic acid column chromatography (30). Pol α and BF pol lack 3′-5′ exonuclease activity. The genes for exonuclease-deficient UL30 and for UL42 were generously provided by Deborah Parris (Ohio State University). The herpes simplex virus I DNA polymerase (UL30(exo-)/UL42-His$_6$-C complex (31)) was expressed in baculovirus-infected Sf 9 insect cells at the Tissue Culture Core Facility at the University of Colorado Health Sciences Center and purified by nickel nitrilotriacetic acid column chromatography.

5′-labeling of primers strands

DNA primers were 5′- ^{32}P-labeled using T4 polynucleotide kinase (New England Biolab) and [γ-^{32}P]ATP. The labeled primer was gel-purified and annealed to the appropriate template strands (32). The sequences of the used primer-templates are given in Chart 3.

Polymerization assays

All kinetic data were determined under steady-state conditions. The pol α, HSV pol, and BF assays contained 0.5 μM 5′- ^{32} P-primer-template, 10 mM $MgCl_2$, 50 mM Tris-HCl pH 7.5, 0.05 mg/ml bovine serum albumine, 1 mM dithiothreitol and various concentrations of natural or analogue dNTPs in a total volume of 5 or 10 μl. The reactions with KF were carried out in commercial NEB2 buffer (50 mM NaCl, 10 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$, 1 mM dithiothreitol) supplemented with 0.05 mg/ml bovine serum albumine. Reactions that served to derive kinetic parameters contained 350 pM KF and 5 nM pol α, respectively. Polymerization was initiated by mixing equal volumes of reaction mixture and enzyme followed by incubation at 37 °C. The reactions were stopped by addition of two volumes gel loading buffer (90 % formamide with 50 mM EDTA) after 1-30 minutes, depending on the catalytic activity of the enzyme. Extension products were separated by denaturing gel electrophoresis (20 % polyacrylamide, 8 M urea) and analyzed by phosphor imaging (Typhoon scanner, Molecular Dynamics). The parameters K_M and V_{max} were obtained by nonlinear curve fitting.

Read through assays

The kinetic parameters for addition of the next correct dNTP onto a just incorporated nucleotide analogue were determined using the "running start" method introduced by Goodman et al. (33). Assays contained 10 μM dtC(o)TP for KF reactions, 20 μM dtC(o)TP for pol α reactions, together with varying amounts of the next correct dNTP (dGTP). They were performed and the reaction products quantified as described above. The ratio between the primer + 1 tC(o) product and the primer + 1 tC(o) + G product yields the kinetic parameters K_M and V_{max} for dGTP polymerization onto the nucleotide analogue.

RESULTS

KF and pol α insert dtCTP and dtCoTP efficiently opposite G and A

We used single nucleotide insertion assays to determine the K_M and V_{max} for the incorporation of dtCTP and dtCoTP into synthetic oligonucleotides of defined sequence. The primer-template sequences were identical in all single nucleotide extension reactions except for the base templating for incorporation of the first dNMP (Chart 3). KF and pol α both catalyze the insertion of the analogue triphosphates opposite G with the same efficiency, V_{max}/K_M, with which they insert dCTP at this position (Table 1). Opposite a template A, both pol α and KF incorporate the analogues only ~ 4- to 11-fold less efficiently than the correct nucleotide dTTP. Both polymerases select against tC(o)-pyrimidine mismatches. KF polymerizes dtCTP and dtCoTP 200- and 1000-fold less effectively opposite T, respectively, than it does dATP, while pol α is even more selective with discrimination factors of 500 and 1100, respectively. The analogous reaction opposite C does not yield any detectable incorporation in the case of pol α and large discrimination factors of 800 and 1600 for dtCTP and dtCoTP using KF, respectively.

Facile polymerization past tC(o)-G and tC(o)-A base pairs

Next we tested the ability of pol α and KF to continue nucleotide polymerization after inserting dtCTP and dtCoTP opposite G and A using the "running start method" introduced by Goodman et al. (33). Assays contained either DNA_G or DNA_A, a fixed concentration of dtC(o)TP and varying concentrations of the next complementary nucleotide, dGTP. We chose C as the second template nucleotide to avoid potential misincorporation of dtC(o)MP opposite A or T, a result that would have complicated data analysis. Pol α elongates the tC(o)-G and CG base pairs with similar catalytic efficiency, while it elongates the tC-A base pair 30-fold and the tCoA base pair 11-fold less efficiently than a T-A base pair (Table 2). Surprisingly, KF elongates the tC(o)-G base pair 30-fold more efficiently than a normal C-G base pair. Additionally, KF also polymerizes dGTP onto a tC(o)-A base pair more efficiently than it elongates either a G-C or T-A base pair.

Kinetic and structural studies indicate that base mismatches and some nucleotide analogues can have long range effects that cause pausing or stalling of polymerases several nucleotides downstream of the mismatch or analogue (34-36). To find out if tC(o)-G and tC(o)-A base pairs have this effect, we examined elongation of DNA_G and DNA_A in assays containing either dtC(o)TP, dATP, dGTP and dTTP (and DNA_G) or dtC(o)TP, dATP, dGTP and dCTP (and DNA_A) (Figure 1). Importantly, the presence of a single tC(o)-G base pair does not cause any downstream problems. More pronounced pausing occurs during the replication of DNA_A, which is probably a cumulative effect of generating several tC(o)-A base pairs, as is required by the template sequence. In contrast, neither enzyme polymerized past a tC(o)-T or tC(o)-C mismatches.

Ambivalent base pairing is typical for A and B family polymerases

To test the generality of these results for other A and B family polymerases, we carried out single nucleotide insertion and read through experiments using BF pol and HSV pol, which belong to the A and B polymerase families, respectively. Figure 1 and 2 show the products of assays containing 25 μM of each dNTP analogue. Ignoring differences in polymerase processivity, the overall pattern repeats for all four polymerases: the polymerase incorporate dtCTP and dtCoTP opposite G and A with efficiencies similar to dCTP and dTTP, respectively. There is some incorporation of the nucleotide analogues opposite T, but very little or no incorporation of dtC(o)TP opposite C. Thus, either these cytosine analogues lack structural features that enable any of the polymerases to prevent polymerization opposite A or the ability to form tC(o)-A mismatches is intrinsic to the base pair itself.

Pol α and KF tackle G and A repeats

To provide a more rigorous challenge for pol α and KF, we conducted primer extension reactions with templates containing 4 consecutive G's followed by 4 A's (DNA_{4G4A}), thus calling for the incorporation of consecutive analogues. In reactions containing only dtCTP, KF inserted 6 instead of the potential 8 nucleotides, whereas it incorporated 8 dtCoTPs (Figure 3). Compared to the incorporation of consecutive dCTPs, the fraction of elongated primer is clearly diminished in reactions employing dtCTP, whereas dtCoTP yields primer elongation comparable to the natural nucleotide (Table 4). Quantification of the extension products reveals that polymerization of the analogues is processive and little pausing occurs, except towards the end of the A repeat. Adding dATP, dGTP and dCTP to assays containing DNA_{4G4A} and either dtCTP or dtCoTP showed that KF can not complete template replication after polymerizing at least 4 consecutive dtCTPs while some complete replication occurs when dtCoTP substitutes for dCTP (Figure 3).

Based on the fraction of elongated primer and the distribution of product bands, pol α processes dCTP and the analogues very much alike within the G quartet (Table 4, Figure 3). No analogue incorporation takes place opposite A in the case of pol α. Judged on the basis of the control experiments with all four natural nucleotides, pol α largely dissociates from the primer-template after the G quartet, consistent with the lower processivity of pol α.

In order to show that bypass of consecutive analogue base pairs is possible we designed a slightly different read through experiment. This time, the assay contained a 20-fold excess of dTTP over dtC(o)TP to drive the formation of correct T-A base pairs subsequent to the four tC(o)-G base pairs (Figure 4). In addition, the KF concentration was increased 10-fold. After incorporation of 4 consecutive C's or tCo's, KF elongated both products similarly well to the primer +8 position, whereas incorporation of 4 tC's reduced the extent elongated significantly (Table 4).

Because pol α did not replicate DNA_{4G4A} completely, even in the presence of the four natural dNTPs, we conducted read through experiments with this template and the fellow B family enzyme, HSV pol instead. HSV pol bypassed the G and A repeat in the presence of both nucleotide analogues at least as well as with only natural dNTPs (Table 4), whereas KF synthesizes the tCo-G run efficiently but not the tC-G run.

tC and tCo encode for the incorporation of dGTP and dATP

Our next objective was to characterize the templating properties of the base analogues. To this end we performed single nucleotide insertion experiments with the four natural dNTPs and DNA_{tC} and DNA_{tCo}, the latter being two DNA constructs that feature a single tC or tCo in the template. Of the natural nucleotides, pol α and KF polymerize dGTP most efficiently across from tC and tCo (Table 3). Both enzymes incorporated dGTP opposite the analogues <10-fold

less efficiently than polymerization of dGTP opposite C, with the reduced catalytic efficiency primarily a consequence of higher K_M values. When it comes to incorporation of dATP opposite tC(o), pol α is clearly more specific than KF. Pol α synthesizes the AtC and A-tCo base pairs 400- and 260-fold less efficiently than a G-C base pair, respectively. KF polymerized dATP opposite tC and tCo 41- and 18-fold less efficiently than it polymerized dGTP across from C. Formation of pyrimidine-tC(o) base pairs was very unfavorable for both polymerases. The slightly larger discrimination factors observed for dNTP incorporation opposite tC(o) suggest that tC(o) is structurally more perturbing in the template than in the primer strand.

We examined the ability of pol α and KF to complete primer extension past a single or consecutive G-tC(o) base pairs. A single G-tC or G-tCo base pair does not pose a major obstacle to continued dNTP polymerization (Figure 5). The total amount of primer extension by KF is 51 % for DNA_{tC}, 41% for DNA_{tCo} and 54% for DNA_C and no significant pausing occurs. For pol α, 35%, 29 % and 48 % primer extension are observed for DNA_{tC}, DNA_{tCo} and DNA_C. In all reactions, 85-90% of the primer +1 product was extended. Thus, these polymerases efficiently extend single tC(o)-G and G-tC(o) base pairs. However, the bypass efficiency decreases in the presence of four consecutive base analogues in the template sequence (Figure 5). Curiously, bypassing consecutive tCo bases constitutes a larger problem than bypassing consecutive tC bases, whereas the polymerization of consecutive dtCoTPs is easier than the formation of a tC repeat in the primer strand.

DISCUSSION

We examined the ability of pol α and KF to polymerize the fluorescent cytidine analogues dtCTP and dtCoTP. Both analogues are derived from cytosine by ring expansion at N^4 and C-5 of C. The resulting bases differ significantly from cytosine at the face of the major groove but retain full H-bonding capability. Remarkably, pol α and KF polymerize the analogue triphosphates as efficiently as dCTP across from G and only 4-11 times less efficiently then dTTP opposite A. When tC or tCo are in the template strand, both DNA polymerases chose to polymerize either dGTP or dATP. Hereby, KF is 4- to 9-fold more likely to insert dGTP than dATP and pol α prefers dGTP over dATP by a factor of 30 to 65. The observed ambivalence of incorporation stands in stark contrast with the ability of both enzymes to discriminate against naturally occurring C-A base mismatches, which display discrimination factors of $>10^3$ (37). The polymerases readily elongated all analogue-purine base pairs, tC(o)-G, G-tC(o), tC(o)-A and A-tC(o), via polymerization of either natural dNTPs or dNTP analogues. Previous work indicated a close correlation exists between successful bypass and the absence of structural perturbation in the primer-template (38). Based on the efficient read through, the base analogues induce surprisingly little structural distortion in either the primer or the template strand.

The mutagenic properties of N^4-substituted cytosine analogues are well documented and have been ascribed to the propensity of the substituted heterocycle to populate both the amino and imino tautomeric forms in solution (4,8-10,12-16,18,20). While the amino form engages in regular C-G base pairs, the imino tautomer is presumed to form C-A base pairs that are isosteric to T-A. Unmodified cytosine exists mainly as the amino tautomer in solution with an amino to imino ratio of $\sim 10^4$, thereby making the incorporation of C opposite A a very unlikely event (39). However, for many N^4-substituted cytosine analogues the equilibrium is shifted in favor of the imino tautomer. For instance, the imino to amino ratio is 10:1 for N^4-hydroxy-cytosine, 30:1 for N^4-aminocytosin and 11:1 for 3,4-dihydro-6H-8Hpyrimido[4,5-c][1,2]oxazin-2-one (P base) (15,20). So far, a univocal correlation between the amino-imino equilibrium and the misincorporation frequency opposite A has only been observed for dPTP. Potentially, the energetically more favorable *syn* conformation (relative to O^2 of cytosine) of the N^4 substituent might interfere with the ability of imino-cytosine to act as hydrogen bond acceptor, thus

destabilizing C-A base pairs (40). Because the N^4-substituent of the P base is locked in the *anti* conformation, unhindered access is provided to N^4, which is probably the origin of the good correlation between tautomeric equilibrium and the frequency of C-A mismatches. NMR studies of P base and N^4-methoxycytosine containing DNA provide experimental evidence for the existence of *syn/anti* conformers and the imino form of modified cytosines (8,11). These studies revealed Watson-Crick geometry for both P-G, N^4-methoxycytosine-G, P-A, and N^4-methoxycytosine-A base pairs that were in slow exchange with the corresponding wobble base pair, especially in the case of the N^4-methoxycytosine-G base pair. Thus, *syn/anti* conversion of the N^4-methyl group destabilizes the Watson-Crick structure.

Although the tautomeric equilibrium of tC and tCo remains to be determined, our results suggest that base pairing of tC and tCo by the action of DNA polymerases is driven by hydrogen bonding of the analogue's amino and imino form with G and A, respectively (Chart 2). The chemical structure of the analogues likely stabilizes the imino tautomer of the analogues, as compared to unmodified C. First, the middle ring of the analogues containing N^4 is aromatic. Thus, N^4 is sp^2 hybridized and bonded to two sp^2 carbons, which are electron-withdrawing. This arrangement helps to stabilize an unshared pair of electrons in a sp^2 orbital on N^4, relative to cytosine. Overall, this inductive effect makes N^4 of the amino tautomer much more acidic than N^4 of cytosine, facilitating the proton shift to N-3. Even more intriguing is that KF and pol α, two polymerases with otherwise fairly divergent substrate specificities, misincorporate the cytidine analogues with almost identical efficiency opposite A and preliminary data suggest that the same applies to HSV pol and BF pol.

If the imino tautomer of the analogues drives their misincorporation opposite A, one would expect that tC(o)-A and A-tC(o) base pairs should form with similar efficiency, provided the polymerase active site accommodates the analogues equally well in the primer and the template strand. To judge the effect of the tautomeric equilibrium on base pairing, it is instructive to compare the catalytic efficiencies, V_{max}/K_M, for the incorporation of dtC(o)TP opposite G and A with those for the incorporation of dGTP and dATP across tC(o) for each enzyme. The data listed in Table 5 reveal that base pair formation is surprisingly symmetrical for pol α and for KF, i.e., base pair processing has similar characteristics whether the analogue is located in the primer or in the template strand. This suggests that the tautomeric equilibrium of the analogues, and hence hydrogen bonding, plays an important role for nucleotide selection by both polymerases.

NMR structural studies of the self-complementary oligonucleotide GAGGTGCA tC CTC showed that it forms a B helix, albeit with a small bend that has been attributed to the presence of TG-AC steps in the sequence (24). However, the structure shows local distortions around the tC base, which mainly originate from contacts between tC and the T adjacent to the G in the complementary strand. In particular, the hydrogen at N^4 of tC forms a bifurcated hydrogen bond with O^6 of G and with O^4 of T. Consequently, the collinear alignment of the hydrogen bond acceptors/donors in a G-tC base pair is slightly perturbed, however, true intercalation of the tC base leading to π stacking interactions with the opposite strand was not observed. This improper alignment of the hydrogen bonding interface may account for the observed incorporation of tC doublets or even triplets via a primer slippage mechanism (Figure 3). In addition, hydrophobic interactions between the incoming dtC(o)TP and the analogue inserted at the 3′-end of the primer may also help drive misincorporation. Given the large hydrophobic cross section of tC(o), the interactions between two adjacent analogues are expected to be unusually strong, possibly changing the binding affinity of the incoming nucleotide. Similar effects may account for the observation that KF polymerizes dGTP ~30 times more efficiently onto a 3′-tC(o) than onto a 3′-C.

Duplex melting experiments provide further evidence for the existence of energetically favorable interactions between tC and the natural bases; the DNA melting temperature increases on average by 3 °C per tC (24). In a 10 base pair duplex, changing the base pair partner with tCo gave T_Ms of tCo-G 46 °C, tCo-A 34 °C, tCo-T 24 °C and tCo -C 20°C. Thus, the tCo-A mismatch is significantly more stable than the other mismatches (41). The fluorescence emission spectrum also mirrors the special character of the tCo-A base pair. It exhibits a fine structure that could be indicative of a stabilized imino tautomer, besides other structural rearrangements (41). Interestingly, the thermodynamic mismatch stabilities correlate with the order of mismatch discrimination by pol α and KF. In contrast, other reported examples question the relevance of thermodynamic stability as a determinant for nucleotide selection (42). For example, polymerases choose to polymerize O^6-methyl-dGTP opposite thymine despite O^6-methyl-G-C being more stable in the DNA duplex (43). Another example is the facile insertion of 2-aminopurine opposite T and C (44), even though the 2-aminopurine-A base pair is more stable in DNA than the 2-aminopurine-C base pair (45). Without knowing the exact nature of the intermolecular forces that stabilize a particular base pair it would be difficult to predict how these forces are amplified in the polymerase active site. While charge based interactions, such as hydrogen bonding and ionic interactions, are likely to be amplified in a water free environment, hydrophobic interactions may be similar in strength to those in water.

The cytosine analogues we present herein constitute a fairly unexplored class of natural base mimics in that they are significantly larger than the canonical bases but retain their hydrogen bonding pattern. Previous work by other groups focused predominantly on separating the effects of base pair size and shape from those of hydrogen bonding potential. Toward this goal, a wealth of nonpolar base mimics were designed in which the hydrogen bonding donors and acceptors were eliminated and replaced by chemical groups of comparable size (reviewed in (46-48)). The facile incorporation of compounds that closely resemble the canonical base pairs in size and shape has given rise to the idea that base pair shape is more important than hydrogen bonding for efficient nucleotide incorporation. However, our own work as well as that from other labs indicate that this kind of size selection functions only for analogues that are closely related in shape to the natural bases (37,38,49-53). Nucleotides of completely unrelated size are often readily but unselectively incorporated by KF and pol α (Ex., 5, 6-dinitrobenzimidazole, 5-nitroindole and 5-trifluoromethyl-1H-benzimidazole). This observation has translated into the idea that at leas some DNA polymerases use a combination of positive and negative selection to polymerize the correct nucleotide. According to this theory DNA polymerases select *for* the correct dNTP and *against* the incorrect dNTP based on the chemical properties of the naturally occurring base mismatches. With regard to dtCTP and dtCoTP incorporation it would mean that the presence of the cytosine-like hydrogen bonding pattern is sufficient to identify the nucleotide as “correct” but the protruding benzene ring in the major groove is an unknown feature that is insufficient to qualify the nucleotide as “wrong”. While the importance of size match at the hydrogen bonding interface appears to be a major prerequisite for correct nucleotide selection, it will be particularly exciting to further explore the effect of groove expansion in conjunction with retained hydrogen bonding patterns.

Nucleosides with ambivalent base pairing properties are expected to be powerful mutagens and therefore they are interesting drug candidates in cancer therapy. The repeating introduction of G-C to A-T transition mutations might corrupt the genome to a degree that triggers apoptosis and thus prevents the division of cancerous cells. In order to propagate base errors through several rounds of cell division the base analogue would ideally possess mutagenic properties as nucleotide and as the templating base. Based on our in vitro experiments, tC and tCo appear to be promising candidates in this regard. The ring structure has the advantage that metabolic enzymes will not be able to deactivate the mutagen by removing N^4. However, at this point it

is unclear whether tC and tCo can be made bioavailable and cell-based experiments are needed to assess their therapeutic potential.

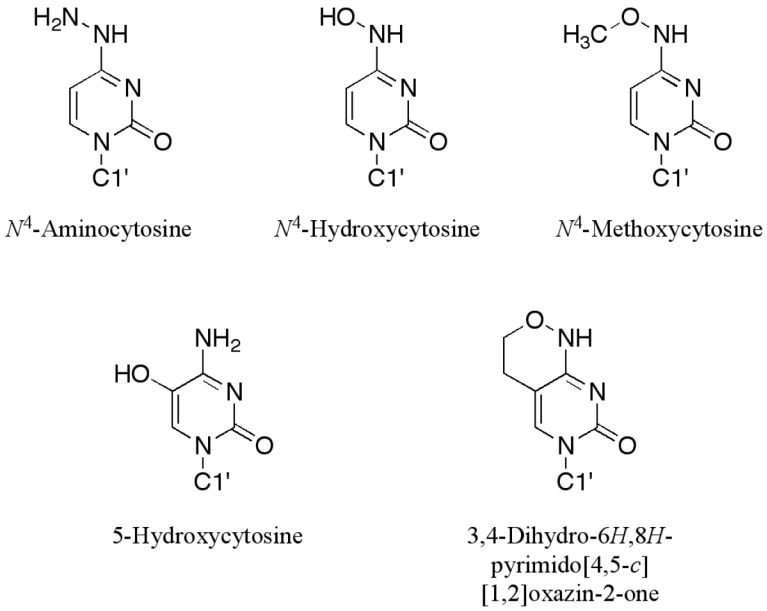
REFERENCES

1. Kunkel TA, Bebenek R. DNA replication fidelity. Annu Rev Biochem 2000;69:497–529. [PubMed: 10966467]
2. Watson JD, Crick FHC. Genetical Implications of the Structure of Deoxyribonucleic Acid. Nature 1953;171:964–967. [PubMed: 13063483]
3. Morgan AR. Base Mismatches and Mutagenesis - How Important Is Tautomerism. Trends Biochem Sci 1993;18:160–163. [PubMed: 8328014]
4. Topal MD, Fresco JR. Complementary Base-Pairing and Origin of Substitution Mutations. Nature 1976;263:285–289. [PubMed: 958482]
5. Johnson SJ, Beese LS. Structures of mismatch replication errors observed in a DNA polymerase. Cell 2004;116:803–816. [PubMed: 15035983]
6. Hunter WN, Brown T, Kennard O. Structural Features and Hydration of a Dodecamer Duplex Containing 2 CA Mispairs. Nucleic Acids Res 1987;15:6589–6606. [PubMed: 3627999]
7. Hunter WN, Brown T, Kneale G, Anand NN, Rabinovich D, Kennard O. The Structure of Guanosine-Thymidine Mismatches in B-DNA at 2.5-a Resolution. J Biol Chem 1987;262:9962–9970. [PubMed: 3611072]
8. Nedderman ANR, Stone MJ, Williams DH, Lin PKT, Brown DM. Molecular-Basis for Methoxyamine-Initiated Mutagenesis - H-1 Nuclear-Magnetic-Resonance Studies of Oligonucleotide Duplexes Containing Base-Modified Cytosine Residues. J Mol Biol 1993;230:1068–1076. [PubMed: 8478918]
9. Fazakerley GV, Gdaniec Z, Sowers LC. Base-Pair Induced Shifts in the Tautomeric Equilibrium of a Modified DNA-Base. J Mol Biol 1993;230:6–10. [PubMed: 8450552]
10. Stone MJ, Nedderman ANR, Williams DH, Lin PKT, Brown DM. Molecular-Basis for Methoxyamine Initiated Mutagenesis - H-1 Nuclear-Magnetic-Resonance Studies of Base-Modified Oligodeoxynucleotides. J Mol Biol 1991;222:711–723. [PubMed: 1660932]
11. Nedderman ANR, Stone MJ, Lin PKT, Brown DM, Williams DH. Base-Pairing of Cytosine Analogs with Adenine and Guanine in Oligonucleotide Duplexes - Evidence for Exchange between Watson-Crick and Wobble Base-Pairs Using H-1-NMR Spectroscopy. J Chem Soc Chem Commun 1991:1357–1359.
12. Negishi K, Bessho T, Hayatsu H. Nucleoside and Nucleobase Analog Mutagens. Mutat Res 1994;318:227–238. [PubMed: 7527486]
13. Nomura A, Negishi K, Hayatsu H. Direct-Acting Mutagenicity of N-4-Aminocytidine Derivatives Bearing Alkyl-Groups at the Hydrazino Nitrogens. Nucleic Acids Res 1985;13:8893–8899. [PubMed: 3909110]
14. Negishi K, Takahashi M, Yamashita Y, Nishizawa M, Hayatsu H. Mutagenesis by N-4-Aminocytidine - Induction of at to GC Transition and Its Molecular Mechanism. Biochemistry 1985;24:7273–7278. [PubMed: 2935188]
15. Brown DM, Hewlins MJE, Schell P. Tautomeric State of N(4)-Hydroxy- and of N(4)-Amino-Cytosine Derivatives. J Chem Soc 1968:1925–9.
16. Reeves ST, Beattie KL. Base-Pairing Properties of N^4-Methoxydeoxycytidine 5′-Triphosphate During DNA-Synthesis on Natural Templates, Catalyzed by DNA Polymerase I of Escherichia Coli. Biochemistry 1985;24:2262–2268. [PubMed: 3888268]
17. Singer B, Abbott LG, Spengler SJ. Assessment of Mutagenic Efficiency of 2 Carcinogen-Modified Nucleosides, 1,N-6-Ethenodeoxyadenosine and O-4-Methyldeoxythymidine, Using Polymerases of Varying Fidelity. Carcinogenesis 1984;5:1165–1171. [PubMed: 6205783]
18. Singer B, Fraenkelconrat H, Abbott LG, Spengler SJ. N-4-Methoxydeoxycytidine Triphosphate Is in the Imino Tautomeric Form and Substitutes for Deoxythymidine Triphosphate in Primed Poly D[A-T] Synthesis with Escherichia-Coli DNA-Polymerase-I. Nucleic Acids Res 1984;12:4609–4619. [PubMed: 6377235]
19. Suen W, Spiro TG, Sowers LC, Fresco JR. Identification by UV resonance Raman spectroscopy of an imino tautomer of 5-hydroxy-2 ′-deoxycytidine, a powerful base analog transition mutagen with

a much higher unfavored tautomer frequency than that of the natural residue 2 ′-deoxycytidine. Proc Natl Acad Sci U S A 1999;96:4500–4505. [PubMed: 10200291]

20. Harris VH, Smith CL, Cummins WJ, Hamilton AL, Adams H, Dickman M, Hornby DP, Williams DM. The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: Support for the rare tautomer hypothesis of substitution mutagenesis. J Mol Biol 2003;326:1389–1401. [PubMed: 12595252]
21. Lin KY, Jones RJ, Matteucci M. Tricyclic 2′-Deoxycytidine Analogs - Syntheses and Incorporation into Oligodeoxynucleotides Which Have Enhanced Binding to Complementary RNA. J Am Chem Soc 1995;117:3873–3874.
22. Sandin P, Borjesson K, Li H, Martensson J, Brown T, Wilhelmsson LM, Albinsson B. Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. Nucleic Acids Res 2008;36:157–167. [PubMed: 18003656]
23. Sandin P, Wilhelmsson LM, Lincoln P, Powers VEC, Brown T, Albinsson B. Fluorescent properties of DNA base analogue tC upon incorporation into DNA - negligible influence of neighbouring bases on fluorescence quantum yield. Nucleic Acids Res 2005;33:5019–5025. [PubMed: 16147985]
24. Engman KC, Sandin P, Osborne S, Brown T, Billeter M, Lincoln P, Norden B, Albinsson B, Wilhelmsson LM. DNA adopts normal B-form upon incorporation of highly fluorescent DNA base analogue tC: NMR structure and UV-Vis spectroscopy characterization. Nucleic Acids Res 2004;32:5087–5095. [PubMed: 15452275]
25. Sandin P, Lincoln P, Brown T, Wilhelmsson LM. Synthesis and oligonucleotide incorporation of fluorescent cytosine analogue tC: a promising nucleic acid probe. Nat Protoc 2007;2:615–623. [PubMed: 17406622]
26. Ludwig J. A New Route to Nucleoside 5′-Triphosphates. Acta Biochim Biophys Hung 1981;16:131–133.
27. Beaucage SL, a. C. MH. New class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett 1981:1859–1862.
28. McBride LJ, Caruthers MH. An Investigation of Several Deoxynucleoside Phosphoramidites Useful for Synthesizing Deoxynucleotides. Tetrahedron Lett 1983:245–248.
29. Wilhelmsson LM, Sandin P, Holmen A, Albinsson B, Lincoln P, Norden B. Photophysical characterization of fluorescent DNA base analogue, tC. J Phys Chem B 2003;107:9094–9101.
30. Zerbe LK, Goodman MF, Efrati E, Kuchta RD. Abasic template lesions are strong chain terminators for DNA primase but not for DNA polymerase alpha during the synthesis of new DNA strands. Biochemistry 1999;38:12908–12914. [PubMed: 10504262]
31. Ramirez-Aguilar KA, Low-Nam NA, Kuchta RD. Key role of template sequence for primer synthesis by the herpes simplex virus 1 helicase-primase. Biochemistry 2002;41:14569–14579. [PubMed: 12463757]
32. Maniatis, T.; Fritsch, EF.; Sambrook, J. Molecular Cloning. Cold Spring Harbor Laboratory; Cold Spring Harbor, NY: 1982.
33. Goodman MF, Creighton S, Bloom LB, Petruska J. Biochemical Basis of DNA-Replication Fidelity. Crit Rev Biochem Mol Biol 1993;28:83–126. [PubMed: 8485987]
34. Miller H, Grollman AP. Kinetics of DNA polymerase I (Klenow fragment exo(-)) activity on damaged DNA templates: Effect of proximal and distal template damage on DNA synthesis. Biochemistry 1997;36:15336–15342. [PubMed: 9398262]
35. Carver TE, Hochstrasser RA, Millar DP. Proofreading DNA - Recognition of Aberrant DNA Termini by the Klenow Fragment of DNA Polymerase I. Proc Natl Acad Sci U S A 1994;91:10670–10674. [PubMed: 7938011]
36. Weiss SJ, Fisher PA. Interaction of Drosophila DNA-Polymerase Alpha Holoenzyme with Synthetic Template-Primers Containing Mismatched Primer Bases or Propanodeoxyguanosine Adducts at Various Positions in Template and Primer Regions. J Biol Chem 1992;267:18520–18526. [PubMed: 1526988]
37. Chiaramonte M, Moore CL, Kincaid K, Kuchta RD. Facile polymerization of dNTPs bearing unnatural base analogues by DNA polymerase alpha and Klenow fragment (DNA polymerase I). Biochemistry 2003;42:10472–10481. [PubMed: 12950174]

38. Beckman J, Kincaid K, Hocek M, Spratt T, Engels J, Cosstick R, Kuchta RD. Human DNA polymerase alpha uses a combination of positive and negative selectivity to polymerize purine dNTPs with high fidelity. Biochemistry 2007;46:448–460. [PubMed: 17209555]
39. Katritzky AR, Waring AJ. Tautomeric Azines .1. Tautomerism of 1-Methyluracil and 5-Bromo-1-Methyluracil. J Chem Soc 1962:1540.
40. Morozov YV, Savin FA, Chekhov VO, Budowsky EI, Yakovlev DY. Photochemistry of N-6-Methoxyadenosine and of N-4-Hydroxycytidine and Its Methyl-Derivatives .1. Spectroscopic and Quantum Chemical Investigation of Ionic and Tautomeric Forms - Syn-Anti Isomerization. J Photochem 1982;20:229–252.
41. Borjesson K, Sandin P, Wilhelmsson LM. Nucleic acid structure and sequence probing using fluorescent base analogue tC(O). Biophys Chem 2009;139:24–28. [PubMed: 18963381]
42. Petruska J, Goodman MF, Boosalis MS, Sowers LC, Cheong C, Tinoco I. Comparison between DNA Melting Thermodynamics and DNA-Polymerase Fidelity. Proc Natl Acad Sci U S A 1988;85:6252–6256. [PubMed: 3413095]
43. Snow ET, Foote RS, Mitra S. Kinetics of Incorporation of O6-Methyldeoxyguanosine Monophosphate During Invitro DNA-Synthesis. Biochemistry 1984;23:4289–4294. [PubMed: 6386047]
44. Patro JN, Urban M, Kuchta RD. Role of the 2-Amino Group of Purines during dNTP Polymerization by Human DNA Polymerase alpha. Biochemistry 2009;48:180–189. [PubMed: 19072331]
45. Eritja R, Kaplan BE, Mhaskar D, Sowers LC, Petruska J, Goodman MF. Synthesis and Properties of Defined DNA Oligomers Containing Base Mispairs Involving 2-Aminopurine. Nucleic Acids Res 1986;14:5869–5884. [PubMed: 3737416]
46. Henry AA, Romesberg FE. The evolution of DNA polymerases with novel activities. Curr Opin Biotech 2005;16:370–377. [PubMed: 16006114]
47. Henry AA, Romesberg FE. Beyond A, C, G and T: augmenting nature's alphabet. Curr Opin Chem Biol 2003;7:727–733. [PubMed: 14644182]
48. Kool ET. Active site tightness and substrate fit in DNA replication. Annu Rev Biochem 2002;71:191–219. [PubMed: 12045095]
49. Kincaid K, Beckman J, Zivkovic A, Halcomb RL, Engels JW, Kuchta RD. Exploration of factors driving incorporation of unnatural dNTPS into DNA by Klenow fragment (DNA polymerase I) and DNA polymerase alpha. Nucleic Acids Res 2005;33:2620–2628. [PubMed: 15879351]
50. Ogawa AK, Wu YQ, Berger M, Schultz PG, Romesberg FE. Rational design of an unnatural base pair with increased kinetic selectivity. J Am Chem Soc 2000;122:8803–8804.
51. Wu YQ, Ogawa AK, Berger M, McMinn DL, Schultz PG, Romesberg FE. Efforts toward expansion of the genetic alphabet: Optimization of interbase hydrophobic interactions. J Am Chem Soc 2000;122:7621–7632.
52. Berger M, Wu YQ, Ogawa AK, McMinn DL, Schultz PG, Romesberg FE. Universal bases for hybridization, replication and chain termination. Nucleic Acids Res 2000;28:2911–2914. [PubMed: 10908353]
53. Ogawa AK, Wu YQ, McMinn DL, Liu JQ, Schultz PG, Romesberg FE. Efforts toward the expansion of the genetic alphabet: Information storage and replication with unnatural hydrophobic base pairs. J Am Chem Soc 2000;122:3274–3287.



N^4-Aminocytosine

N^4-Hydroxycytosine

N^4-Methoxycytosine

5-Hydroxycytosine

3,4-Dihydro-6*H*,8*H*-pyrimido[4,5-*c*][1,2]oxazin-2-one

Chart 1.
N^4-substituted cytosine analogues.

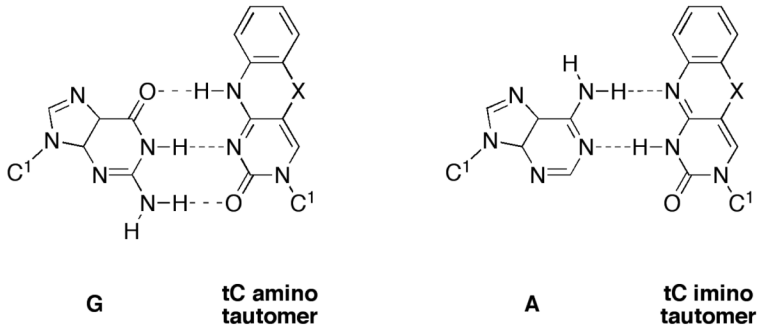


Chart 2.
Hydrogen bonding of the amino and imino tautomer of $tC^{(O)}$ with G and A (X = S for tC, X = O for tCo).

DNA_A	5’–TCCATATCACAT 3’–AGGTATAGTGTA**A**CTCTTATCATCT
DNA_T	5’–TCCATATCACAT 3’–AGGTATAGTGTA**T**CTCTTATCTATCT
DNA_C	5’–TCCATATCACAT 3’–AGGTATAGTGTA**C**TTCTTATCTATCT
DNA_G	5’–TCCATATCACAT 3’–AGGTATAGTGTA**G**CTCTTATCTATCT
DNA_{4G4A}	5’–TCCATATCACAT 3’–AGGTATAGTGTA**GGGG**AAAATCAGTCA
DNA_{tC}	5’–TCCATATCACAT 3’–AGGTATAGTGTA(**tC**)ATCTTATCTATCT
DNA_{tCo}	5’–TCCATATCACAT 3’–AGGTATAGTGTA(**tCo**)ATCTTATCTATCT
DNA_{4tC}	5’–TCCATATCACAT 3’–AGGTATAGTGTA(**tC**)(**tC**)(**tC**)(**tC**)ATATCATCT
DNA_{4tCo}	5’–TCCATATCACAT 3’–AGGTATAGTGTA(**tCo**)(**tCo**)(**tCo**)(**tCo**)ATATCATCT

Chart 3.
Sequences of Primer-Templates. The letter after “DNA” designates the template base being replicated.

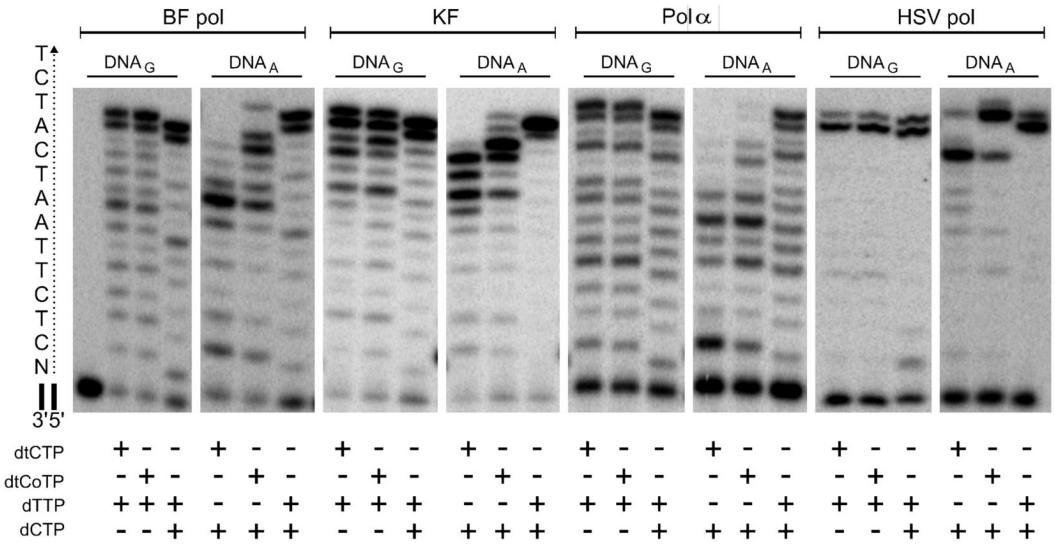


Figure 1.

Primer elongation after incorporation of dtCTP and dtCoTP by different DNA polymerases. The template sequence is depicted on the left. N is either G or A, corresponding to DNA_G or DNA_A, respectively. Assays contained 0.5 μM primer-template, 25 μM dATP, 25 μM dGTP and 25 μM of each indicated dNTP. The enzyme concentrations were 1 nM BF pol, 350 pM KF, 5 nM pol α and 5 nM HSV pol.

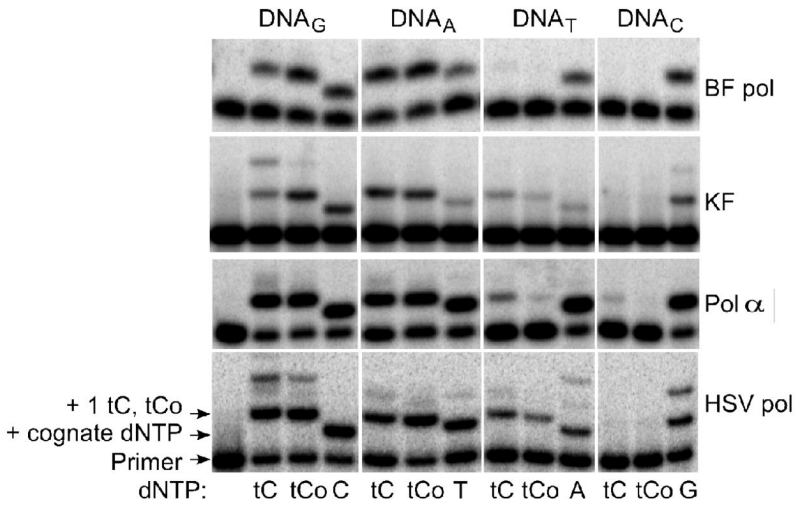


Figure 2.

Incorporation of dtCTP and dtCoTP and the natural dNTPs across the four natural bases. All reactions contained 0.5 μM primer-template and 25 μM dNTP. The enzyme concentrations were 1 nM BF pol, 350 pM KF, 5 nM pol α and 5 nM HSV pol.

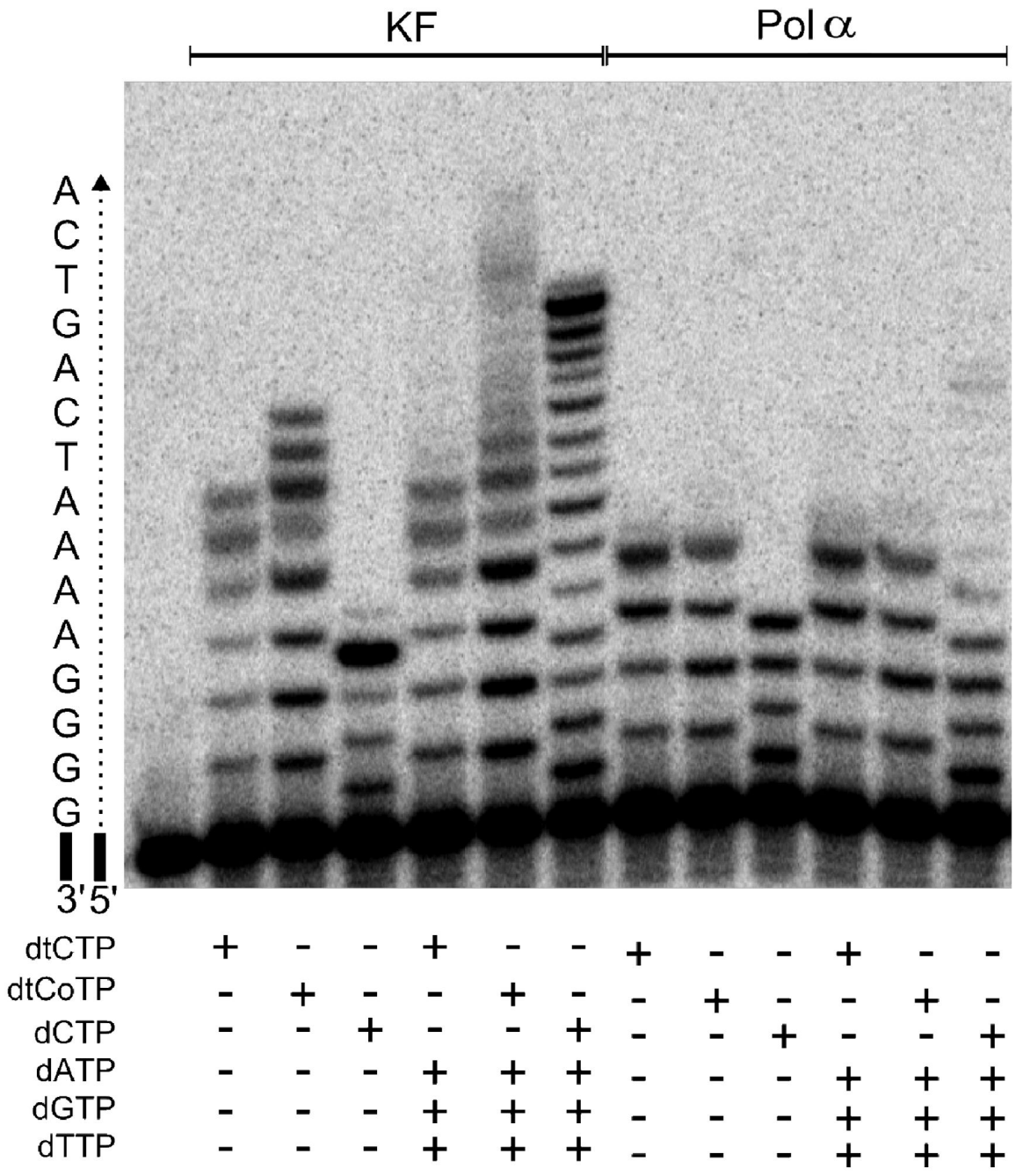


Figure 3.
Read through of DNA_{4G4A} by KF and pol α. Assays contained 0.5 μM primer-template, 25 μM of each designated dNTP and 350 pM KF or 5 nM pol α.

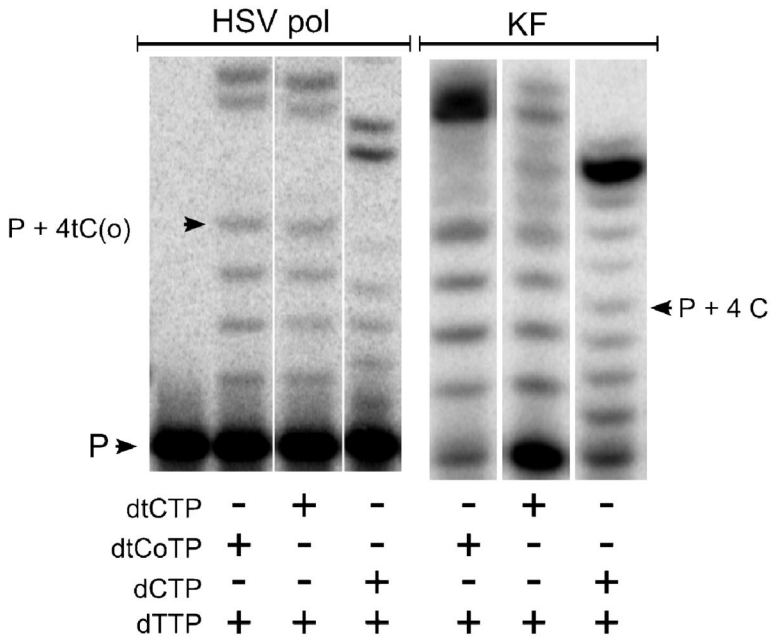


Figure 4.

Demonstration of bypass of DNA_{4G4A} upon incorporation of $dtC^{(O)}TP$ and an excess of dTTP by A and B family DNA polymerases. HSV pol was used instead of pol α due to its higher processivity on natural substrates. Assays contained 0.5 μM primer-template, 10 μM of dCTP or $dtC^{(O)}TP$, 200 μM dTTP and 5 nM HSV pol or 3.5 nM KF, respectively.

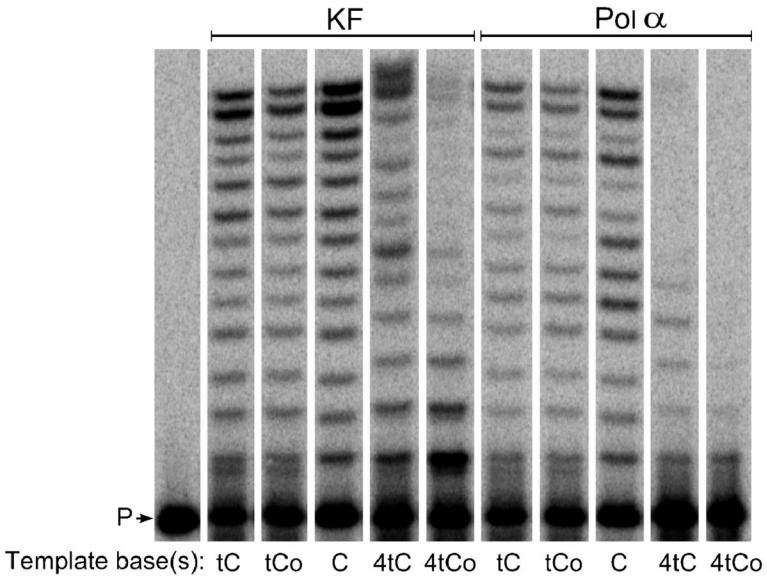


Figure 5.
Incorporation of natural dNTPs and subsequent primer extension opposite a single $tC^{(O)}$ (DNAtC(O)) or across four consecutive $tC^{(O)}$s (DNA4tC(O)). Assays contained 0.5 μM primer-template, 25 μM of each natural dNTP and 350 pM KF or 10 nM pol α, respectively.

Table 1

Kinetic parameters for incorporation of dNTP analogues into DNA_N by pol α and KF.

dNTP	DNA	V_{max} (SD) $\left[\frac{\%\ extension}{\text{min}}\right]$	K_M (SD) [μM]	V_{max}/K_M $\left[\frac{\%\ extension}{\mu M \cdot \text{min}}\right]$	Discrimination [a]
Pol α					
dCTP	DNA	3.63 (0.09)	0.23 (0.02)	15.8	1
dtCTP	DNA	4.5 (0.4)	0.18 (0.07)	25	0.6
dtCoTP	DNA	3.4 (0.2)	0.13 (0.02)	26.2	0.6
dTTP	DNA	4.1 (0.4)	0.6 (0.2)	6.5	1
dCTP	DNA	0.21 (0.02)	510 (100)	4×10^{-4}	2×10^{4}
dtCTP	DNA	1.6 (0.2)	2.7 (0.9)	0.6	11
dtCoTP	DNA	2.2 (0.1)	2.4 (0.5)	0.9	7
dATP	DNA_T	4.8 (0.4)	0.5 (0.1)	10.7	1
dCTP	DNA_T			$< 1 \times 10^{-5}$	$> 1 \times 10^{6}$
dtCTP	DNA_T	0.66 (0.006)	40 (9)	0.02	540
dtCoTP	DNA_T	0.9 (0.2)	88 (39)	0.01	1070
dGTP	DNA	2.6 (0.2)	0.33 (0.09)	7.9	1
dCTP	DNA			$< 1 \times 10^{-5}$	$> 8 \times 10^{5}$ [b]
dtCTP	DNA			$< 1 \times 10^{-4}$	$> 8 \times 10^{4}$
dtCoTP	DNA			$< 1 \times 10^{-4}$	$> 8 \times 10^{4}$
KF					
dCTP	DNA	5.5 (0.2)	0.07 (0.02)	84	1
dtCTP	DNA	9.0 (0.2)	0.10 (0.01)	95	0.9
dtCoTP	DNA	9.1 (0.5)	0.09 (0.02)	107	0.8
dTTP	DNA	2.3 (0.01)	0.09 (0.04)	26	1
dCTP	DNA	4.4 (0.3)	1800 (300)	2×10^{-3}	1.3×10^{4}
dtCTP	DNA	5.3 (0.3)	0.8 (0.2)	6.6	4
dtCoTP	DNA	6.1 (0.5)	1.3 (0.3)	4.7	6
dATP	DNA_T	1.08 (0.03)	0.06 (0.02)	19	1
dCTP	DNA_T	0.60 (0.06)	1300 (400)	5×10^{-4}	4×10^{4}
dtCTP	DNA_T	1.4 (0.06)	13 (2)	0.11	170

dNTP	DNA	V_{max} (SD) $\left[\frac{\%\ extension}{min}\right]$	K_M (SD) [μM]	V_{max}/K_M $\left[\frac{\%\ extension}{\mu M \cdot min}\right]$	Discrimination [a]
dtCoTP	DNA_T	3.1 (0.8)	175 (61)	0.02	950
dGTP	DNA	2.3 (0.2)	0.14 (0.07)	16	1
dCTP	DNA			$< 1 \times 10^{-5}$	$> 1 \times 10^{6}$
dtCTP	DNA	1.3 (0.1)	80 (16)	0.02	810
dtCoTP	DNA	0.56 (0.08)	100 (29)	0.01	1620

[a] Discrimination is defined as V_{max}/K_M for the incorporation of the correct nucleotide (i.e. dATP opposite T) divided by V_{max}/K_M for analogue incorporation opposite the same template.

[b] Some mismatches were incorporated at levels too low for accurate determination. In case of dtC(o)TP incorporation, $V_{max}/K_M < 1 \times 10^{-4}$ means less than 2 % primer extension after incubation with 200 μM dNTP for 20min. For dCTP incorporation, $V_{max}/KM < 1 \times 10^{-5}$ means less than 5 % incorporation after incubation with 8 mM dNTP for 60min.

Table 2

Next correct dNTP insertion on DNA_{NC} by pol α and KF.

dNTP	DNA_N	V_{max} (SD) $\left[\frac{\%\ extension}{min}\right]$	K_M [μM]	V_{max}/K_M $\left[\frac{\%\ extension}{\mu M \cdot min}\right]$	Discrimination [a]
			Pol α		
dCTP, dGTP	DNA_G	11 (1)	6 (1)	1.8	1
dtCTP, dGTP	DNA_G	18 (5)	8 (3)	2.3	0.8
dtCoTP, dGTP	DNA_G	19 (2)	7 (1)	2.7	0.7
dTTP, dGTP	DNA_A	34 (11)	7 (3)	4.9	1
dtCTP, dGTP	DNA_A	0.41 (0.03)	2.6 (0.5)	0.16	30
dtCoTP, dGTP	DNA_A	2.7 (0.6)	6 (3)	0.45	11
			KF		
dCTP, dGTP	DNA_G	5.3 (0.2)	3.2 (0.3)	1.7	1
dtCTP, dGTP	DNA_G	26 (2)	0.5 (0.2)	52	0.03
dtCoTP, dGTP	DNA_G	53 (7)	0.9 (0.4)	59	0.03
dTTP, dGTP	DNA_A	14 (0.6)	1.7 (0.2)	8.2	1
dtCTP, dGTP	DNA_A	25 (2)	1.6 (0.4)	15.6	0.5
dtCoTP, dGTP	DNA_A	33 (1)	0.4 (0.1)	82.5	0.1

[a] Discrimination is defined as V_{max}/K_M for the incorporation of dGTP opposite C following the incorporation of a matching nucleotide at position N (f. e. dGTP after dATP has been polymerized across T) divided by V_{max}/K_M for dGTP polymerization subsequent to insertion of a nucleotide analogue at position N.

Table 3

Incorporation of natural dNTPs opposite tC(o) in the template strand by pol α and KF.

dNTP	DNA_N	V_{max} (SD) $\left[\frac{\%\ extension}{min}\right]$	K_M (SD) [μM]	V_{max}/K_M $\left[\frac{\%\ extension}{\mu M \cdot min}\right]$	Discrimination [a]
		Pol α			
dGTP	DNA_C	2.6 (0.2)	0.33 (0.09)	7.9	1
dGTP	DNA_{tC}	1.8 (0.6)	2.3 (0.2)	1.3	6.1
dGTP	DNA_{tCo}	2.2 (0.4)	2.0 (0.09)	0.9	8.8
dATP	DNA_{tC}	54 (19)	0.89 (0.06)	0.02	395
dATP	DNA_{tCo}	35 (4)	1.02 (0.02)	0.03	263
dTTP	$DNA_{tC(o}$				n.d. [b]
dCTP	$DNA_{tC(o}$				n.d.
		KF			
dGTP	DNA_C	2.3 (0.2)	0.14 (0.07)	16.2	1
dGTP	DNA_{tC}	2.3 (0.9)	7.8 (0.8)	3.4	4.8
dGTP	DNA_{tCo}	1.1 (0.8)	3.7 (0.9)	3.4	4.8
dATP	DNA_{tC}	19 (13)	8 (2)	0.4	41
dATP	DNA_{tCo}	4 (2)	3.6 (0.5)	0.9	18
dTTP	DNA_{tC}	190 (99)	0.59 (0.09)	0.003	5400
dTTP	DNA_{tCo}	288 (110)	0.72 (0.09)	0.003	5400
dCTP	$DNA_{tC(o}$				n.d.

[a] Discrimination is defined as V_{max}/K_M for the polymerization of dGTP across C divided by V_{max}/K_M for the incorporation of a natural dNTP across tC(o) in the template strand.

[b] n. d. means no primer extension was observed in the presence of 2 mM dNTP and a reaction time of 30 min.

Table 4

Polymerization of consecutive dtC(o)TPs

dNTP	Total fraction of elongated primer (%)[a]		Fraction of primer elongated past A repeat (%) [b]	
	KF	Pol α	KF	HSV pol
dtCTP	11	17	23	23
dtCoT	31	17	61	23
dCTP	25	21	75	20

[a] Assay contained DNA_{4G4A} and the indicated dNTP; quantification of data in Figure 3.

[b] Assay contained DNA_{4G4A}, dTTP and the indicated dNTP; quantification of data in Figure 4.

Table 5

Symmetry of base pair formation.

Analogue	KF		Pol α	
	R_{dNTP} [a]	$R_{template}$	R_{dNTP} [a]	$R_{template}$ [b]
tC	14.3	9	42	65
tCo	22.8	4	29	30

[a] $R_{dNTP} = \frac{V_{max}/K_M\ (dtC(O)TP - G)}{V_{max}/K_M\ (dtC(O)TP - A)}$

[b] $R_{template} = \frac{V_{max}/K_M\ (dGTP - tC(O))}{V_{max}/K_M\ (dATP - tC(O))}$