BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 7547–7555 7547
DOI: 10.1021/bi9006995

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

Gudrun Stengel,[‡] Byron W. Purse,[§] L. Marcus Wilhelmsson,[∥] Milan Urban,[‡] and Robert D. Kuchta*,[‡]

[‡]*Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 30309-0215,* [§]*Department of Chemistry and Biochemistry, University of Denver, Denver, Colorado 80208, and* [∥]*Department of Chemical and Biological Engineering/Physical Chemistry, Chalmers University of Technology, SE-41296 Gothenburg, Sweden*

Received April 23, 2009; Revised Manuscript Received June 2, 2009

ABSTRACT: We studied the incorporation of the fluorescent cytidine analogues 1,3-diaza-2-oxophenothiazine (tC) and 1,3-diaza-2-oxophenoxazine (tCo) by human DNA polymerase α and Klenow fragment of DNA polymerase I (*Escherichia coli*). These tricyclic nucleobases possess the regular hydrogen bonding interface of cytosine but are significantly expanded in size 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 four consecutive nucleotide analogues. Consistent with these results, both DNA polymerases efficiently polymerize dGTP and dATP when tC and tCo are in the template strand. Klenow fragment inserts dGTP with a 4−9-fold higher probability than dATP, while polymerase α favors dGTP over dATP by a factor of 30−65. Overall, the properties of tC(o) as a templating base and as an 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).

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} and 10^{-6} error per nucleotide replicated for proofreading deficient DNA polymerases (*1*). Proofreading and other cellular repair mechanisms reduce the actual error frequency to around 10^{-8}−10^{-11} error 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 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 those of 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 because of the technical challenge of accurately determining the tautomeric equilibrium constant that lies almost completely on the side of the amino (A and C) and keto (T and G) tautomers. X-ray structures of short DNA duplexes frequently capture pyrimidine-purine 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 an aqueous environment, this finding does not exclude the possibility that tautomeric forms occur during nucleotide selection by DNA polymerases. Evidence of 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-6*H*,8*H*-pyrimido[4,5-*c*][1,2]oxazin-2-one (P base) (*20*) (see Chart 1).

tC[1] and tCo [1,3-diaza-2-oxophenothiazine and 1,3-diaza-2-oxophenoxazine, respectively (*21*)] 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

---

[†] This work was supported by grants from the National Institutes of Health (GM54194 and AI59764) to R.D.K.

*To whom correspondence should be addressed. E-mail: kuchta@colorado.edu. Phone: (303) 492-7027. Fax: (303) 492-5894.

[1] Abbreviations: BF pol, *Bacillus stearothermophilus* DNA polymerase; EDTA, ethylenediaminetetraacetic acid; HSV, herpes simplex virus-1; KF, Klenow fragment of DNA polymerase I from *Escherichia coli*; P, 3,4-dihydro-6*H*,8*H*-pyrimido[4,5-*c*][1,2]oxazin-2-one; dPTP, 3,4-dihydro-6*H*,8*H*-pyrimido[4,5-*c*][1,2]oxazin-2-one-2′-deoxyribose 5′-triphosphate; pol α, DNA polymerase α; SD, standard deviation; tC, 1,3-diaza-2-oxophenothiazine-2′-deoxyribose; tCo, 1,3-diaza-2-oxophenoxazine-2′-deoxyribose; Tris-HCl, tris(hydroxymethyl)aminomethane, HCl salt.

Chart 1: N^4-Substituted Cytosine Analogues

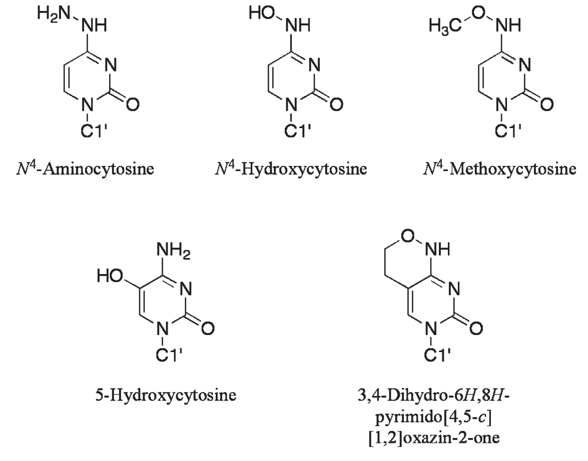

Chart 2: Hydrogen Bonding of the Amino and Imino Tautomer of tC(o) with G and A (X is S for tC, and X is O for tCo)

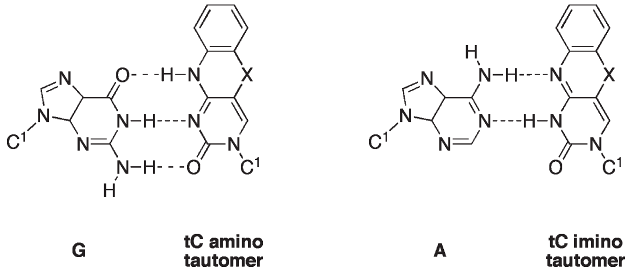

Chart 3: Primer−Template Sequences[a]

DNA$_A$	5'-TCCATATCACAT 3'-AGGTATAGTGTA**A**CTCTTATCATCT
DNA$_T$	5'-TCCATATCACAT 3'-AGGTATAGTGTATCTCTTATCTATCT
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

[a] The letter after DNA designates the template base being replicated.

versatile biophysical probes for studies of DNA−protein interactions (*22*, *23*). Substituting a C-G base pair with 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 $\pi-\pi$ 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 pol), 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 that A and B family enzymes use to distinguish between right and wrong dNTPs.

EXPERIMENTAL PROCEDURES

Materials and Enzymes. All reagents were of the 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-oxophenothiazine and 1,3-diaza-2-oxophenoxazine nucleosides were synthesized according to the methods of Sandin et al. (*25*) and Lin et al. (*21*), respectively. 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, the difference being that the coupling times for both analogues were 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 linear combination of the extinction coefficients of the individual nucleosides at 260 nm (ε_{tC} = 13500 M^{-1} cm^{-1}, and ε_{tCo} = 11000 M^{-1} cm^{-1}) (*22*, *29*). Multiplying the result by a factor of 0.9 accounted for base stacking interactions.

Escherichia coli DNA polymerase I Klenow fragment (exo$^-$) was purchased from New England Biolabs. *B. stearothermophilus* DNA polymerase I (BF pol) was a generous gift from L. Beese (Duke University, Durham, NC). Human pol α (N-His$_6$-p70-p180 complex), expressed in baculovirus-infected Sf9 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 D. Parris (The Ohio State University, Columbus, OH). The herpes simplex virus I DNA polymerase [UL30(exo-)/UL42-His$_6$-C complex (*31*)] was expressed in baculovirus-infected Sf9 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 Primer 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 primer−template sequences that were used 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-labeled primer−template, 10 mM $MgCl_2$, 50 mM Tris-HCl (pH 7.5), 0.05 mg/mL bovine serum albumin, 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 conducted in commercial NEB2 buffer [50 mM NaCl, 10 mM Tris-HCl (pH 7.9), 10 mM $MgCl_2$, and 1 mM dithiothreitol] supplemented with 0.05 mg/mL bovine serum albumin. Reaction mixtures used to derive kinetic parameters contained 350 pM KF and 5 nM pol α. 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 2 volumes of gel loading buffer (90% formamide with 50 mM EDTA) after 1−30 min, depending on the catalytic activity of the enzyme. Extension products were separated by denaturing gel electrophoresis (20% polyacrylamide and 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 the 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, and 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 polymerization of dGTP 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} values 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 ~4−11-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 insertion of 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 prevent 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 C-G base pairs with similar catalytic efficiency, while it elongates the tC-A base pair 30-fold and the tCo-A 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 a T-A base pair.

Table 1: Kinetic Parameters for Incorporation of dNTP Analogues into DNA_N by pol α and KF

dNTP	DNA_N	V_{max} (SD) (% extension/min)	K_M (SD) (μM)	V_{max}/K_M (% extension μM^{-1} min^{-1})	discrimination[a]
		pol α			
dCTP	DNA_G	3.63 (0.09)	0.23 (0.02)	15.8	1
dtCTP	DNA_G	4.5 (0.4)	0.18 (0.07)	25	0.6
dtCoTP	DNA_G	3.4 (0.2)	0.13 (0.02)	26.2	0.6
dTTP	DNA_A	4.1 (0.4)	0.6 (0.2)	6.5	1
dCTP	DNA_A	0.21 (0.02)	510 (100)	4×10^{-4}	2×10^4
dtCTP	DNA_A	1.6 (0.2)	2.7 (0.9)	0.6	11
dtCoTP	DNA_A	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_C	2.6 (0.2)	0.33 (0.09)	7.9	1
dCTP	DNA_C			$< 1 \times 10^{-5}$	$> 8 \times 10^{5}$[b]
dtCTP	DNA_C			$< 1 \times 10^{-4}$	$> 8 \times 10^4$
dtCoTP	DNA_C			$< 1 \times 10^{-4}$	$> 8 \times 10^4$
		KF			
dCTP	DNA_G	5.5 (0.2)	0.07 (0.02)	84	1
dtCTP	DNA_G	9.0 (0.2)	0.10 (0.01)	95	0.9
dtCoTPP	DNA_G	9.1 (0.5)	0.09 (0.02)	107	0.8
dTTP	DNA_A	2.3 (0.01)	0.09 (0.04)	26	1
dCTP	DNA_A	4.4 (0.3)	1800 (300)	2×10^{-3}	1.3×10^4
dtCTP	DNA_A	5.3 (0.3)	0.8 (0.2)	6.6	4
dtCoTP	DNA_A	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
dtCoTP	DNA_T	3.1 (0.8)	175 (61)	0.02	950
dGTP	DNA_C	2.3 (0.2)	0.14 (0.07)	16	1
dCTP	DNA_C			$< 1 \times 10^{-5}$	$> 1 \times 10^6$
dtCTP	DNA_C	1.3 (0.1)	80 (16)	0.02	810
dtCoTP	DNA_C	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 the case of dtC(o)TP incorporation, a V_{max}/K_M of $< 1 \times 10^{-4}$ means $< 2\%$ primer extension after incubation with 200 μM dNTP for 20 min. For dCTP incorporation, a V_{max}/K_M of $< 1 \times 10^{-5}$ means $< 5\%$ incorporation after incubation with 8 mM dNTP for 60 min.

Table 2: Next Correct dNTP Insertion on DNA_{NC} by pol α and KF

dNTP	DNA_{NC}	V_{max} (SD) (% extension/min)	K_M (SD) (μM)	V_{max}/K_M (% extension μM^{-1} min^{-1})	discrimination[a]
			pol α		
dCTP, dGTP	DNA_{GC}	11 (1)	6 (1)	1.8	1
dtCTP, dGTP	DNA_{GC}	18 (5)	8 (3)	2.3	0.8
dtCoTP, dGTP	DNA_{GC}	19 (2)	7 (1)	2.7	0.7
dTTP, dGTP	DNA_{AC}	34 (11)	7 (3)	4.9	1
dtCTP, dGTP	DNA_{AC}	0.41 (0.03)	2.6 (0.5)	0.16	30
dtCoTP, dGTP	DNA_{AC}	2.7 (0.6)	6 (3)	0.45	11
			KF		
dCTP, dGTP	DNA_{GC}	5.3 (0.2)	3.2 (0.3)	1.7	1
dtCTP, dGTP	DNA_{GC}	26 (2)	0.5 (0.2)	52	0.03
dtCoTP, dGTP	DNA_{GC}	53 (7)	0.9 (0.4)	59	0.03
dTTP, dGTP	DNA_{AC}	14 (0.6)	1.7 (0.2)	8.2	1
dtCTP, dGTP	DNA_{AC}	25 (2)	1.6 (0.4)	15.6	0.5
dtCoTP, dGTP	DNA_{AC}	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* (i.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*.

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, which is required by the template sequence. In contrast, neither enzyme polymerized past a tC(o)-T or tC(o)-C mismatch.

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 conducted single-nucleotide insertion and read-through experiments using BF pol and HSV pol, which belong to the A and B polymerase families, respectively. Figures 1 and 2 show the products of assays containing each dNTP analogue at 25 μM. If differences in polymerase processivity are ignored, the overall pattern repeats for all four polymerases: the polymerases incorporate dtCTP and dtCoTP opposite G and A with efficiencies similar to those of 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 four consecutive G's followed by four A's (DNA_{4G4A}), thus calling for the incorporation of consecutive analogues. In reaction mixtures containing only dtCTP, KF inserted six instead of the potential eight nucleotides, whereas it incorporated eight 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 that of the natural nucleotide (Table 3). Quantification of the extension products reveals that polymerization of the analogues is processive and little pausing occurs, except toward the end of the A repeat. Adding dATP, dGTP, and dTTP to assays containing DNA_{4G4A} and either dtCTP or dtCoTP showed that KF cannot complete template replication after polymerizing at least four consecutive dtCTPs while some complete replication occurs when dtCoTP substitutes for dCTP (Figure 3).

On the basis of 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 3 and 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.

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 four consecutive C's or tCo's, KF elongated both products similarly well to the primer +8 position, whereas incorporation of four tC's reduced the extent elongated significantly (Table 3).

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 3), whereas KF synthesizes the tCo-G run efficiently but not the tC-G run.

tC and tCo Encode 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 4). 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 A-tC 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 incorporation of dNTP opposite tC(o) suggest that tC(o) is structurally more perturbing in the template than in the primer strand.

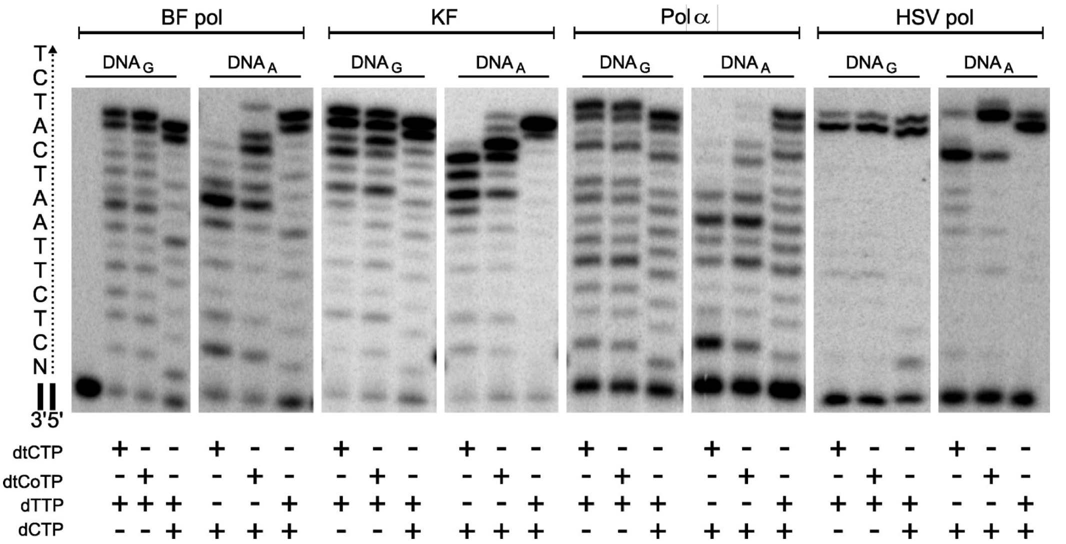

Figure 1: Primer elongation after incorporation of dtCTP and dtCoTP by different DNA polymerases. The template sequence is depicted on the left. N is either G or A, corresponding to DNA$_G$ or DNA$_A$, respectively. Assays contained 0.5 μM primer−template, 25 μM dATP, 25 μM dGTP, and each indicated dNTP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

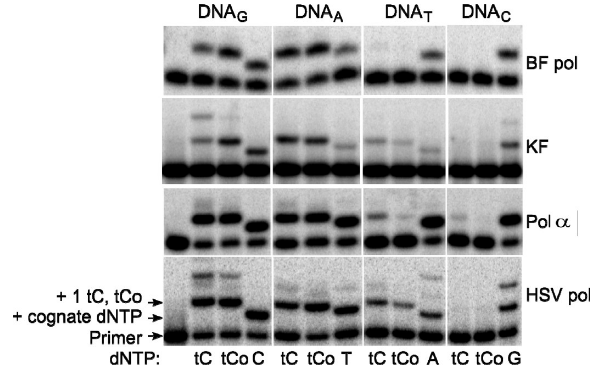

Figure 2: Incorporation of dtCTP and dtCoTP and the natural dNTPs across the four natural bases. All reaction mixtures contained 0.5 μM primer−template and 25 μM dNTP. The enzyme concentrations were as follows: 1 nM BF pol, 350 pM KF, 5 nM pol α, and 5 nM HSV pol.

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 is observed for DNA$_{tC}$, DNA$_{tCo}$, and DNA$_C$, respectively. 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.

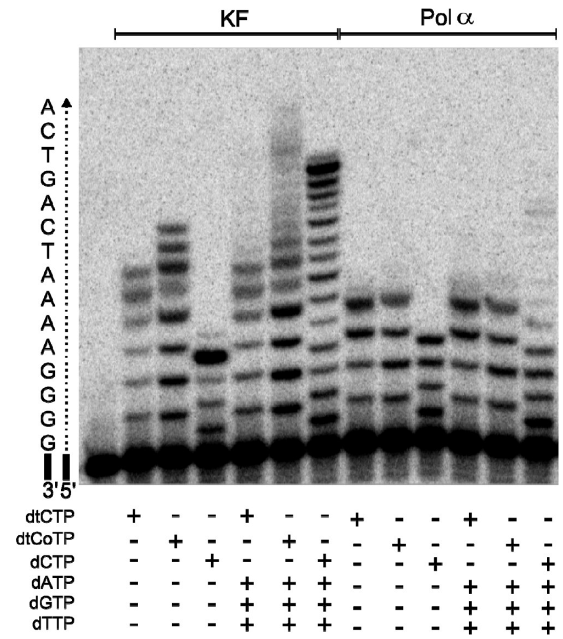

Figure 3: Read-through of DNA$_{4G4A}$ by KF and pol α. Assays contained 0.5 μM primer−template, each designated dNTP at 25 μM, and 350 pM KF or 5 nM pol α.

Table 3: Polymerization of Consecutive dtC(o)TPs

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

[a] The assay contained DNA$_{4G4A}$ and the indicated dNTP; quantification of data in Figure 3. [b] The assay contained DNA$_{4G4A}$, dTTP, and the indicated dNTP; quantification of data in Figure 4.

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

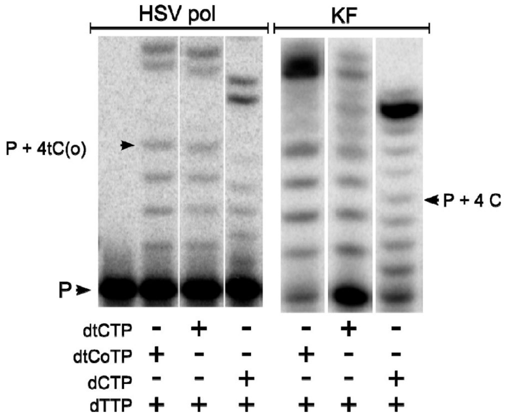

FIGURE 4: Demonstration of bypass of DNA_{4G4A} upon incorporation of dtC(o)TP and an excess of dTTP by A and B family DNA polymerases. HSV pol was used instead of pol α because of its higher processivity on natural substrates. Assays contained 0.5 μM primer−template, 10 μM dCTP or dtC(o)TP, 200 μM dTTP,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

as efficiently as dCTP across from G and only 4−11 times less efficiently then dTTP opposite A. When tC or tCo is in the template strand, both DNA polymerases choose to polymerize either dGTP or dATP. Hereby, KF is 4−9-fold more likely to insert dGTP than dATP, and pol α prefers dGTP over dATP by a factor of 30−65. 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*). On the basis of 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}-hydroxycytosine, 30:1 for *N*^{4}-aminocytosine, and 11:1 for 3,4-dihydro-6*H*,8*H*-pyrimido[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 been observed for only 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 iminocytosine to act as a 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 the tautomeric equilibrium and the frequency of C-A mismatches. NMR studies of P base and *N*^{4}-methoxycytosine-containing DNA provide experimental evidence of the existence of *syn*/*anti* conformers and the imino form of modified cytosines (*8*, *11*). These studies revealed Watson−Crick geometry for 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.

Table 4: Incorporation of Natural dNTPs opposite tC(o) in the Template Strand by pol α and KF

dNTP	DNA_N	V_{max} (SD) (% extension /min)	K_M (SD) (μM)	V_{max}/K_M (% extension μM^{-1} min^{-1})	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	$DNA_{tC(o)}$				nd[b]
dCTP	$DNA_{tC(o)}$				nd[b]
			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	$DNA_{tC(o)}$				nd[b]

[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] No primer extension was observed in the presence of 2 mM dNTP with a reaction time of 30 min.

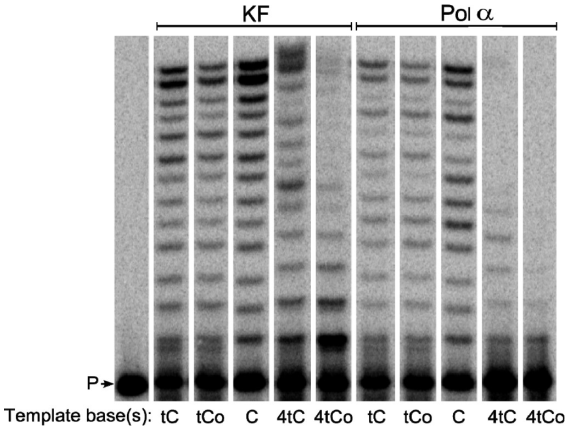

FIGURE 5: Incorporation of natural dNTPs and subsequent primer extension opposite a single tC(o) [$DNA_{tC(o)}$] or across four consecutive tC(o)s [$DNA_{4tC(o)}$]. Assays contained 0.5 μM primer−template, each natural dNTP at 25 μM, and 350 pM KF or 10 nM pol α.

Table 5: Symmetry of Base Pair Formation

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

[a] $R_{\mathrm{dNTP}}=\{V_{\max}/K_{\mathrm{M}}[\mathrm{dtC(o)TP\text{-}G}]\}/\{V_{\max}/K_{\mathrm{M}}[\mathrm{dtC(o)TP\text{-}A}]\}$. [b] $R_{\mathrm{template}}=\{V_{\max}/K_{\mathrm{M}}[\mathrm{dGTP\text{-}tC(o)}]\}/\{V_{\max}/K_{\mathrm{M}}[\mathrm{dATP\text{-}tC(o)}]\}$.

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 an 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 the fact 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, $V_{\max}/K_{\mathrm{M}}$, 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 in 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 and 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 (Figures 2 and 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 of 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 bp duplex, changing the base pair partner with tCo gave T_{M} values of 46 °C for the tCo-G base pair, 34 °C for the tCo-A base pair, 24 °C for the tCo-T base pair, and 20 °C for the tCo-C base pair. 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 knowledge of 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 with chemical groups of comparable size (reviewed in refs (*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 laboratories indicates 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 [i.e., 5,6-dinitrobenzimidazole, 5-nitroindole, and 5-(trifluoromethyl)-1*H*-benzimidazole]. This observation has translated into the idea that at least 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 on the basis of 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 for identifying the nucleotide as "correct" but the protruding benzene ring in the major groove is an unknown feature that is insufficient for identifying 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. To propagate base errors through several rounds of cell division the base analogue would ideally possess mutagenic properties as the nucleotide and as the templating base. On the basis of 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, T. A., and Bebenek, R. (2000) DNA replication fidelity. *Annu. Rev. Biochem. 69*, 497–529.
2. Watson, J. D., and Crick, F. H. C. (1953) Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature 171*, 964–967.
3. Morgan, A. R. (1993) Base Mismatches and Mutagenesis: How Important Is Tautomerism. *Trends Biochem. Sci. 18*, 160–163.
4. Topal, M. D., and Fresco, J. R. (1976) Complementary Base-Pairing and Origin of Substitution Mutations. *Nature 263*, 285–289.
5. Johnson, S. J., and Beese, L. S. (2004) Structures of mismatch replication errors observed in a DNA polymerase. *Cell 116*, 803–816.
6. Hunter, W. N., Brown, T., and Kennard, O. (1987) Structural Features and Hydration of a Dodecamer Duplex Containing 2 CA Mispairs. *Nucleic Acids Res. 15*, 6589–6606.
7. Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., and Kennard, O. (1987) The Structure of Guanosine-Thymidine Mismatches in B-DNA at 2.5-Å Resolution. *J. Biol. Chem. 262*, 9962–9970.
8. Nedderman, A. N. R., Stone, M. J., Williams, D. H., Lin, P. K. T., and Brown, D. M. (1993) Molecular-Basis for Methoxyamine-Initiated Mutagenesis: H-1 Nuclear-Magnetic-Resonance Studies of Oligonucleotide Duplexes Containing Base-Modified Cytosine Residues. *J. Mol. Biol. 230*, 1068–1076.
9. Fazakerley, G. V., Gdaniec, Z., and Sowers, L. C. (1993) Base-Pair Induced Shifts in the Tautomeric Equilibrium of a Modified DNA-Base. *J. Mol. Biol. 230*, 6–10.
10. Stone, M. J., Nedderman, A. N. R., Williams, D. H., Lin, P. K. T., and Brown, D. M. (1991) Molecular-Basis for Methoxyamine Initiated Mutagenesis: H-1 Nuclear-Magnetic-Resonance Studies of Base-Modified Oligodeoxynucleotides. *J. Mol. Biol. 222*, 711–723.
11. Nedderman, A. N. R., Stone, M. J., Lin, P. K. T., Brown, D. M., and Williams, D. H. (1991) 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.*, 1357–1359.
12. Negishi, K., Bessho, T., and Hayatsu, H. (1994) Nucleoside and Nucleobase Analog Mutagens. *Mutat. Res. 318*, 227–238.
13. Nomura, A., Negishi, K., and Hayatsu, H. (1985) Direct-Acting Mutagenicity of N-4-Aminocytidine Derivatives Bearing Alkyl-Groups at the Hydrazino Nitrogens. *Nucleic Acids Res. 13*, 8893–8899.
14. Negishi, K., Takahashi, M., Yamashita, Y., Nishizawa, M., and Hayatsu, H. (1985) Mutagenesis by N-4-Aminocytidine: Induction of at to GC Transition and Its Molecular Mechanism. *Biochemistry 24*, 7273–7278.
15. Brown, D. M., Hewlins, M. J. E., and Schell, P. (1968) Tautomeric State of N(4)-Hydroxy- and of N(4)-Amino-Cytosine Derivatives. *J. Chem. Soc.*, 1925–1929.
16. Reeves, S. T., and Beattie, K. L. (1985) Base-Pairing Properties of N^4-Methoxydeoxycytidine 5′-Triphosphate During DNA-Synthesis on Natural Templates, Catalyzed by DNA Polymerase I of *Escherichia coli*. *Biochemistry 24*, 2262–2268.
17. Singer, B., Abbott, L. G., and Spengler, S. J. (1984) Assessment of Mutagenic Efficiency of 2 Carcinogen-Modified Nucleosides, 1,N-6-Ethenodeoxyadenosine and O-4-Methyldeoxythymidine, Using Polymerases of Varying Fidelity. *Carcinogenesis 5*, 1165–1171.
18. Singer, B., Fraenkelconrat, H., Abbott, L. G., and Spengler, S. J. (1984) 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. 12*, 4609–4619.
19. Suen, W., Spiro, T. G., Sowers, L. C., and Fresco, J. R. (1999) 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. 96*, 4500–4505.
20. Harris, V. H., Smith, C. L., Cummins, W. J., Hamilton, A. L., Adams, H., Dickman, M., Hornby, D. P., and Williams, D. M. (2003) 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. 326*, 1389–1401.
21. Lin, K. Y., Jones, R. J., and Matteucci, M. (1995) Tricyclic 2′-Deoxycytidine Analogs: Syntheses and Incorporation into Oligodeoxynucleotides Which Have Enhanced Binding to Complementary RNA. *J. Am. Chem. Soc. 117*, 3873–3874.
22. Sandin, P., Borjesson, K., Li, H., Martensson, J., Brown, T., Wilhelmsson, L. M., and Albinsson, B. (2008) Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. *Nucleic Acids Res. 36*, 157–167.
23. Sandin, P., Wilhelmsson, L. M., Lincoln, P., Powers, V. E. C., Brown, T., and Albinsson, B. (2005) Fluorescent properties of DNA base analogue tC upon incorporation into DNA: Negligible influence of neighbouring bases on fluorescence quantum yield. *Nucleic Acids Res. 33*, 5019–5025.
24. Engman, K. C., Sandin, P., Osborne, S., Brown, T., Billeter, M., Lincoln, P., Norden, B., Albinsson, B., and Wilhelmsson, L. M. (2004) DNA adopts normal B-form upon incorporation of highly fluorescent DNA base analogue tC: NMR structure and UV-Vis spectroscopy characterization. *Nucleic Acids Res. 32*, 5087–5095.
25. Sandin, P., Lincoln, P., Brown, T., and Wilhelmsson, L. M. (2007) Synthesis and oligonucleotide incorporation of fluorescent cytosine analogue tC: A promising nucleic acid probe. *Nat. Protoc. 2*, 615–623.
26. Ludwig, J. (1981) A New Route to Nucleoside 5′-Triphosphates. *Acta Biochim. Biophys. Hung. 16*, 131–133.
27. Beaucage, S. L., and Caruthers, M. H. (1981) Studies on Nucleotide Chemistry V. Deoxynucleoside Phosphoramidites: A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett. 22*, 1859–1862.
28. McBride, L. J., and Caruthers, M. H. (1983) An Investigation of Several Deoxynucleoside Phosphoramidites Useful for Synthesizing Deoxynucleotides. *Tetrahedron Lett. 24*, 245–248.
29. Wilhelmsson, L. M., Sandin, P., Holmen, A., Albinsson, B., Lincoln, P., and Norden, B. (2003) Photophysical characterization of fluorescent DNA base analogue, tC. *J. Phys. Chem. B 107*, 9094–9101.
30. Zerbe, L. K., Goodman, M. F., Efrati, E., and Kuchta, R. D. (1999) Abasic template lesions are strong chain terminators for DNA primase but not for DNA polymerase α during the synthesis of new DNA strands. *Biochemistry 38*, 12908–12914.
31. Ramirez-Aguilar, K. A., Low-Nam, N. A., and Kuchta, R. D. (2002) Key role of template sequence for primer synthesis by the herpes simplex virus 1 helicase-primase. *Biochemistry 41*, 14569–14579.
32. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Plainview, NY.
33. Goodman, M. F., Creighton, S., Bloom, L. B., and Petruska, J. (1993) Biochemical Basis of DNA-Replication Fidelity. *Crit. Rev. Biochem. Mol. Biol. 28*, 83–126.
34. Miller, H., and Grollman, A. P. (1997) Kinetics of DNA polymerase I (Klenow fragment exo(-)) activity on damaged DNA templates: Effect of proximal and distal template damage on DNA synthesis. *Biochemistry 36*, 15336–15342.
35. Carver, T. E., Hochstrasser, R. A., and Millar, D. P. (1994) Proofreading DNA: Recognition of Aberrant DNA Termini by the Klenow Fragment of DNA Polymerase I. *Proc. Natl. Acad. Sci. U.S.A. 91*, 10670–10674.
36. Weiss, S. J., and Fisher, P. A. (1992) Interaction of *Drosophila* DNA-Polymerase α Holoenzyme with Synthetic Template-Primers Containing Mismatched Primer Bases or Propanodeoxyguanosine

Adducts at Various Positions in Template and Primer Regions. *J. Biol. Chem. 267*, 18520–18526.

37. Chiaramonte, M., Moore, C. L., Kincaid, K., and Kuchta, R. D. (2003) Facile polymerization of dNTPs bearing unnatural base analogues by DNA polymerase α and Klenow fragment (DNA polymerase I). *Biochemistry 42*, 10472–10481.
38. Beckman, J., Kincaid, K., Hocek, M., Spratt, T., Engels, J., Cosstick, R., and Kuchta, R. D. (2007) Human DNA polymerase α uses a combination of positive and negative selectivity to polymerize purine dNTPs with high fidelity. *Biochemistry 46*, 448–460.
39. Katritzky, A. R., and Waring, A. J. (1962) Tautomeric Azines. 1. Tautomerism of 1-Methyluracil and 5-Bromo-1-Methyluracil. *J. Chem. Soc.,* 1540.
40. Morozov, Y. V., Savin, F. A., Chekhov, V. O., Budowsky, E. I., and Yakovlev, D. Y. (1982) 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. 20*, 229–252.
41. Borjesson, K., Sandin, P., and Wilhelmsson, L. M. (2009) Nucleic acid structure and sequence probing using fluorescent base analogue tC-(O). *Biophys. Chem. 139*, 24–28.
42. Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., and Tinoco, I. (1988) Comparison between DNA Melting Thermodynamics and DNA-Polymerase Fidelity. *Proc. Natl. Acad. Sci. U.S.A. 85*, 6252–6256.
43. Snow, E. T., Foote, R. S., and Mitra, S. (1984) Kinetics of Incorporation of O6-Methyldeoxyguanosine Monophosphate During Invitro DNA-Synthesis. *Biochemistry 23*, 4289–4294.
44. Patro, J. N., Urban, M., and Kuchta, R. D. (2009) Role of the 2-Amino Group of Purines during dNTP Polymerization by Human DNA Polymerase α. *Biochemistry 48*, 180–189.
45. Eritja, R., Kaplan, B. E., Mhaskar, D., Sowers, L. C., Petruska, J., and Goodman, M. F. (1986) Synthesis and Properties of Defined DNA Oligomers Containing Base Mispairs Involving 2-Aminopurine. *Nucleic Acids Res. 14*, 5869–5884.
46. Henry, A. A., and Romesberg, F. E. (2005) The evolution of DNA polymerases with novel activities. *Curr. Opin. Biotechnol. 16*, 370–377.
47. Henry, A. A., and Romesberg, F. E. (2003) Beyond A, C, G and T: Augmenting nature's alphabet. *Curr. Opin. Chem. Biol. 7*, 727–733.
48. Kool, E. T. (2002) Active site tightness and substrate fit in DNA replication. *Annu. Rev. Biochem. 71*, 191–219.
49. Kincaid, K., Beckman, J., Zivkovic, A., Halcomb, R. L., Engels, J. W., and Kuchta, R. D. (2005) Exploration of factors driving incorporation of unnatural dNTPS into DNA by Klenow fragment (DNA polymerase I) and DNA polymerase α. *Nucleic Acids Res. 33*, 2620–2628.
50. Ogawa, A. K., Wu, Y. Q., Berger, M., Schultz, P. G., and Romesberg, F. E. (2000) Rational design of an unnatural base pair with increased kinetic selectivity. *J. Am. Chem. Soc. 122*, 8803–8804.
51. Wu, Y. Q., Ogawa, A. K., Berger, M., McMinn, D. L., Schultz, P. G., and Romesberg, F. E. (2000) Efforts toward expansion of the genetic alphabet: Optimization of interbase hydrophobic interactions. *J. Am. Chem. Soc. 122*, 7621–7632.
52. Berger, M., Wu, Y. Q., Ogawa, A. K., McMinn, D. L., Schultz, P. G., and Romesberg, F. E. (2000) Universal bases for hybridization, replication and chain termination. *Nucleic Acids Res. 28*, 2911–2914.
53. Ogawa, A. K., Wu, Y. Q., McMinn, D. L., Liu, J. Q., Schultz, P. G., and Romesberg, F. E. (2000) Efforts toward the expansion of the genetic alphabet: Information storage and replication with unnatural hydrophobic base pairs. *J. Am. Chem. Soc. 122*, 3274–3287.