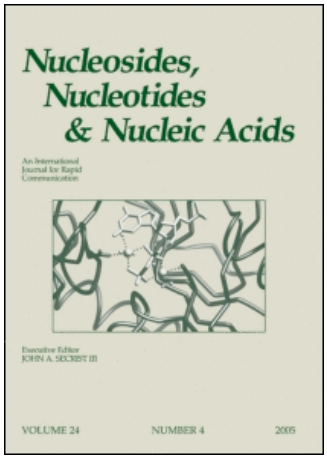

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Incorporation of Multiple Sequential Pseudothymidines by DNA Polymerases and Their Impact on DNA Duplex Structure

Stephanie A. Havemann [a]; Shuichi Hoshika [b]; Daniel Hutter [b]; Steven A. Benner [bc]
[a] Department of Microbiology & Cell Science, Space Life Sciences Laboratory, Kennedy Space Center, University of Florida, Florida, USA
[b] Foundation for Applied Molecular Evolution, Gainesville, Florida, USA
[c] The Westheimer Institute for Science and Technology, Gainesville, Florida, USA

Online Publication Date: 01 March 2008
To cite this Article: Havemann, Stephanie A., Hoshika, Shuichi, Hutter, Daniel and Benner, Steven A. (2008) 'Incorporation of Multiple Sequential Pseudothymidines by DNA Polymerases and Their Impact on DNA Duplex Structure', Nucleosides, Nucleotides and Nucleic Acids, 27:3, 261 - 278
To link to this article: DOI: 10.1080/15257770701853679
URL: http://dx.doi.org/10.1080/15257770701853679

Nucleosides, Nucleotides, and Nucleic Acids, 27:261–278, 2008

ISSN: 1525-7770 print / 1532-2335 online
DOI: 10.1080/15257770701853679

Taylor & Francis
Taylor & Francis Group

INCORPORATION OF MULTIPLE SEQUENTIAL PSEUDOTHYMIDINES BY DNA POLYMERASES AND THEIR IMPACT ON DNA DUPLEX STRUCTURE

Stephanie A. Havemann,[1] Shuichi Hoshika,[2] Daniel Hutter,[2] and Steven A. Benner[2,3]

[1]University of Florida, Department of Microbiology & Cell Science, Space Life Sciences Laboratory, Kennedy Space Center, Florida, USA
[2]Foundation for Applied Molecular Evolution, Gainesville, Florida, USA
[3]The Westheimer Institute for Science and Technology, Gainesville, Florida, USA

□ *Thermal denaturation and circular dichroism studies suggested that multiple (up to 12), sequential pseudothymidines, a representative C-glycoside, do not perturb the structure of a representative DNA duplex. Further, various Family A and B DNA polymerases were found to extend a primer by incorporating four sequential pseudothymidine triphosphates, and then continue the extension to generate full-length product. Detailed studies showed that* Taq *polymerase incorporated up to five sequential C-glycosides, but not more. These results constrain architectures for sequencing, quantitating, and analyzing DNA analogs that exploit C-glycosides, and define better the challenge of creating a synthetic biology using these with natural polymerases.*

Keywords Pseudothymidine; C-glycoside; polymerase; DNA

INTRODUCTION

Various researchers have attempted to increase the number of independently replicable nucleobase pairs in a DNA-like system by changing the arrangement of hydrogen bonding groups on the heterocycle.[1–5] Such attempts immediately encounter a problem that is related to the nature of the bond that joins the heterocycle to its sugar. Each of the four standard nucleotides in DNA joins their heterocycle to their sugar by a carbon-nitrogen bond (N-glycosides). While a nonstandard acceptor-acceptor-donor hydrogen bonding pattern can be implemented on an

Received 13 June 2007; accepted 12 October 2007.
This work was generously supported by a grant from the National Human Genome Research Institute (HG003581).
Address correspondence to Steven A. Benner, Foundation for Applied Molecular Evolution, P.O. Box 13174, Gainesville, FL 32604, USA. E-mail: sbenner@ffame.org

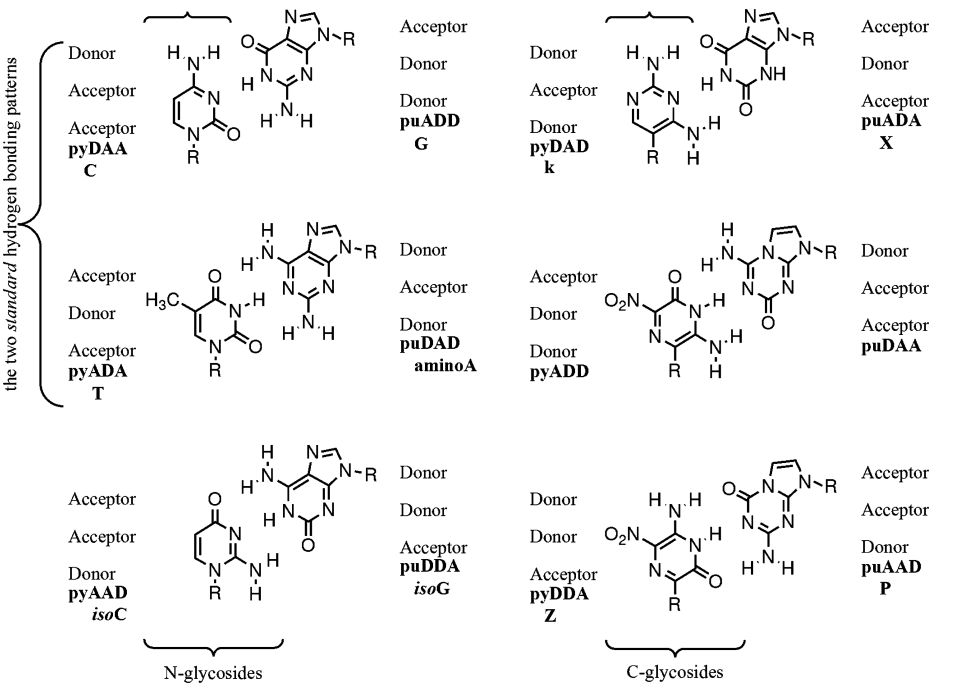

FIGURE 1 C-Glycosides are used for three pyrimidine analogs (pairs to the right) to implement three non-standard hydrogen bond patterns in an artificially expanded genetic information system (AEGIS).[2,11,32,33]

N-glycoside (isocytidine and its analogs, Figure 1), other nonstandard hydrogen bonding patterns must have their six-ring heterocycle attached to their sugar by a carbon-carbon bond. This means that many nonstandard hydrogen-bonding patterns are implemented on C-glycosides (Figure 1).

C-glycosides are robust with respect to chemical degradation, and C-glycosidic analogs of natural nucleotides (e.g., pseudothymidine for thymidine, formycin A for adenosine) could be useful to create more robust DNA structures. Further, C-glycosides are useful for other strategies to expand the genetic alphabet.[6–8] For these and other reasons, C-glycosides are potentially useful in architectures to detect, amplify, quantitate, and sequence DNA and RNA, as well as to support a "synthetic biology."[9] In a synthetic biology, these would be useful first in binding studies, and then in dynamic assay architectures where oligonucleotides incorporating multiple C-glycosides are both copied and synthesized by polymerases.

To date, the impact of C-glycosides on DNA, and proteins that interact with DNA, has been tested primarily by incorporating single C-glycosidic units into duplexes. The impact on duplex stability of a *single* C-glycosidic unit appears to be small.[10] For example, Geyer et al. reported a slight (−1.5°C) decrease in the melting temperature of a reference duplex when a single thymidine (paired with adenosine) was replaced by pseudouridine (ψU) (Figure 2).[11] This difference in duplex thermal stability could, of

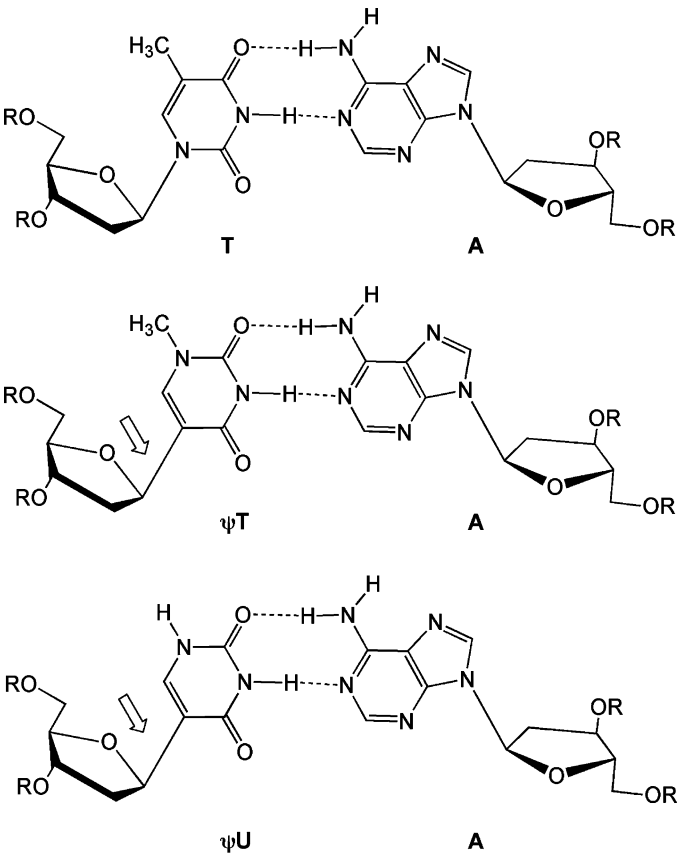

FIGURE 2 Comparison of the structures of thymidine (T), pseudothymidine (ψT), and pseudouridine (ψU) showing the N-glycosidic and C-glycosidic bonds (indicated with the arrow) and the presence of the methyl group in both T and ψT.

course, be attributed to other structural differences between T and ψU, including a methyl group in T that is absent in ψU. Previous studies identified some individual polymerases able to incorporate up to three consecutive C-glycosides,[12–14] but no melting temperatures were reported.

Theoretical considerations suggest that multiple C-glycosides might have an impact not observed with just a few, however. For example, C-glycosidic linkages are believed to change the pucker of the sugar, making it favor the C3′-endo conformation (found regularly in A-DNA and RNA) over the C2′-endo conformation (found regularly in B-DNA).[15] This is just one class of conformational changes that might arise from the "anomeric effect," a theoretical model that has been explored with respect nucleotides.[16] Placing a sufficient number of C-glycosides consecutively in an oligonucleotide might eventually cause a phase transition that drives the duplex to change its conformation from B- to A-DNA. Perhaps consistent with this is classical work; a poly(T)•poly-d(A) duplex was suggested by

TABLE 1 Oligonucleotides used in this study

Name	Sequence (5′ → 3′ Direction)											
P-1	GCG	TAA	TAC	GAC	TCA	CTA	TAG					
T-1	GTT	CCT	GTG	TCG	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-2	TTC	CTG	TGT	CGA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-3	TCC	TGT	GTC	GAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-4	CCT	GTG	TCG	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-5	CTG	TGT	CGA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-6	TGT	GTC	GAA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-7	GTG	TCG	AAA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-8	TGT	CGA	AAA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-9	GTC	GAA	AAA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-10	TCG	AAA	AAA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-11	CGA	AAA	AAA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-12	GAA	AAA	AAA	AAA	ACT	ATA	GTG	AGT	CGT	ATT	ACG	C
T-13	GCG	TAA	TAC	GAC	TCA	CGA	CAG	TCG	GCT	CTG	ACT	GAG
T-14	GCG	TAA	TAC	GAC	TCA	CAA	ACG	TCG	GCT	CTG	ACT	GAG
T-15	GCG	TAA	TAC	GAC	TCC	AAA	AAA	CCG	GCT	CTG	ACT	GAG
T-16	GCG	TAA	TAC	GAC	GAA	AAA	AAA	ACG	GCT	CTG	ACT	GAG
T-17	GCG	TAA	TAC	GAC	AAA	AAA	AAA	AAA	GCT	CTG	ACT	GAG
T-18	CTC	AGT	CAG	AGC	CGA	CTG	TCG	TGA	GTC	GTA	TTA	CGC
T-19	CTC	AGT	CAG	AGC	CGA	CGT	TTG	TGA	GTC	GTA	TTA	CGC
T-20	CTC	AGT	CAG	AGC	CGG	TTT	TTT	GGA	GTC	GTA	TTA	CGC
T-21	CTC	AGT	CAG	AGC	CGT	TTT	TTT	TTC	GTC	GTA	TTA	CGC
T-22	CTC	AGT	CAG	AGC	TTT	TTT	TTT	TTT	GTC	GTA	TTA	CGC
T-23	CTC	AGT	CAG	AGC	CGA	CTG	ΨCG	TGA	GTC	GTA	TTA	CGC
T-24	CTC	AGT	CAG	AGC	CGA	CGΨ	ΨΨG	TGA	GTC	GTA	TTA	CGC
T-25	CTC	AGT	CAG	AGC	CGG	ΨΨΨ	ΨΨΨ	GGA	GTC	GTA	TTA	CGC
T-26	CTC	AGT	CAG	AGC	CGΨ	ΨΨΨ	ΨΨΨ	ΨΨC	GTC	GTA	TTA	CGC
T-27	CTC	AGT	CAG	AGC	ΨΨΨ	ΨΨΨ	ΨΨΨ	ΨΨΨ	GTC	GTA	TTA	CGC

*The Ψ represents the incorporation of a pseudothymidine residue.

circular dichroism (CD) studies to exist in its B-DNA form, while poly-d(U)•poly-d(A) duplex was suggested to exist as A-DNA.[17–20]

To be incorporated into many fully flexible synthetic genetic systems, C-glycosides will need to be incorporated at will into a duplex DNA structure. Given this, it becomes important to determine whether a series of sequential C-glycosidic nucleotides in one strand causes such a conformational change. Accordingly, we prepared the phosphoramidite and triphosphate of pseudothymidine (ψT), the C-glycosidic analog of thymidine that retains the hydrogen bonding pattern complementary to dA as well as the methyl group of T. We then synthesized a series of reference oligonucleotides containing from one to 12 consecutive ψTs. We report here the melting temperatures of duplexes incorporating a dA-ψT base pair, and their CD spectra over a range of temperatures. We also report the ability of eight polymerases to incorporate multiple, consecutive ψTs into a product strand via template-directed synthesis.

EXPERIMENTAL

Synthesis of Pseudothymidine ψT and ψTTP

Synthesis of the 2′-deoxypseudothymidine (ψT) precursor was performed according to the procedures previously.[7,12,21,22] 2′-Pseudothymidine-5′-triphosphate (ψTTP) was prepared following the standard Ludwig-Eckstein procedure.[23] These syntheses are described in more detail in the Supplementary Material.

Synthesis of Oligonucleotides

The oligonucleotides used for these experiments are listed in Table 1. Those sequences containing only standard nucleotides were obtained commercially from Integrated DNA Technologies (Coralville, IA, USA) as PAGE (Polyacrylamide Gel Electrophoresis) purified (P-1) or desalted (T-1 through T-22) oligonucleotides.

The oligonucleotides containing ψT were synthesized on an Applied Biosystems 394 DNA Synthesizer (Foster City, CA, USA) with a coupling time of 10 minutes; no other adjustments to the protocol were found to be necessary. The oligonucleotides were deprotected under standard conditions (concentrated aqueous ammonia, 55°C, overnight).[24] The solutions containing the pure oligonucleotides were neutralized with acetic acid (5 M) and desalted over Waters Sep-Pak Plus C18-cartridges, following the protocol of the manufacturer, except that the final elution was done with 50% acetonitrile in pure water. The solvents were removed by lyophilization.

All oligonucleotides were then purified by HPLC on a Waters Delta 600 with Waters 2487 Dual wavelength absorbance detector, controlled by Waters Millennium software; Dionex DNAPac PA-100 column, 9 × 250 mm, eluent A = 20 mM NaOH, eluent B = 1 M NaCl in 20 mM NaOH, gradient from 10 to 60% B in 3 minutes, then to 90% B in 30 minutes, flow rate = 2.5 mL/minutes; R_t = 21–24 minutes.

Thermal DNA Duplex Denaturation

Thermal DNA duplex denaturation studies were performed with templates containing up to twelve consecutive dA residues (T-13 through T-17) that were paired with its complement template containing consecutive T (T-18 throughT-22) or ψT residues (T-23 through T-27). Experiments were performed in a buffer (45 mM NaCl, 45 mM sodium citrate, pH 8.1, final vol. 1.5 mL) containing template and its complement (1.5 μM of each). Absorbance (260 nm) was monitored on a Varian Cary 300 Spectrophotometer (Palo Alto, CA) over a range of 25.0 to 90.0°C with a change in temperature of 0.5°C min^{-1} for five heating cycles. The initial heating cycle was discarded and the T_m was determined by averaging the temperatures

of the remaining four cycles. The ΔT_m between similar duplexes was calculated by subtracting the T_m of the duplex containing standard bases from the T_m of the duplex containing C-glycosides. The change in T_m per occurrence of each ψT (ΔT_m/mod) was determined by dividing the ΔT_m by the number of nonstandard bases present in that duplex. Free energies of duplex stability were calculated for 50°C (ΔG°_{50}) according to the methods described by Breslauer,[25,26] since ΔGs are more accurate when given at values near the T_m. The ΔH, ΔS, and ΔG°_{50} values were calculated with the assistance of the Varian Cary WinUV software, which required a user to identify a linear region of the van't Hoff plot.

Circular Dichroism Spectra

Circular dichroism spectra were obtained from duplexes formed from an oligonucleotide containing up to twelve consecutive dA residues (T-13 through T-17) paired with its complement containing consecutive T (T-18 throughT-22) or ψT residues (T-23 through T-27) (Table 1). The complementary oligonucleotides (5 nmol of each in 5 μL of water) were diluted with CD buffer [290 μL, 1 M NaCl, 10 mM Na_2HPO_4, 1 mM Na_2EDTA at pH 7.0]. The mixture was heated for 5 minutes at 96.0°C, and then allowed to cool to room temperature over 2 hours.

CD spectra (200 to 320 nm; wavelength steps of 1 nm) were recorded in a quartz cuvette (0.1 cm path length) using an AVIV Circular Dichroism Spectrometer Model 400 (AVIV Biomedical, Inc., Lakewood, NJ, USA) under a dinitrogen atmosphere at temperatures of 4.0°C, 25.0°C, 40.0°C, 53.5°C, or 58.5°C. All scans were performed in triplicate for each sample mixture. The background was then subtracted using Excel, and the resulting curves were combined to generate averaged spectra.

Polymerases and Triphosphates

Klenow Fragment (3′→5′ exo-), *Bst* DNA Polymerase (Large Fragment), *Taq* DNA Polymerase, $Vent_R$ (exo-) DNA Polymerase, Deep $Vent_R$ (exo-) DNA Polymerase, and Therminator DNA Polymerase were purchased from New England BioLabs. *Tth* DNA Polymerase was purchased from Promega Corporation. *Pfu* (exo-) DNA Polymerase was purchased from Stratagene (La Jolla, CA, USA). Optimal temperatures for polymerase function were 37°C for Klenow (exo-), 65°C for *Bst*, and 72°C for *Taq*, *Tth*, Vent (exo-), Deep Vent (exo-), *Pfu* (exo-), and Therminator.

Buffers used in these experiments were supplied by the manufacturer. Reactions using *Bst*, *Taq*, *Tth*, Vent (exo-), Deep Vent (exo-), and Therminator were performed in 1X ThermoPol Buffer [20 mM Tris-HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, and 0.1% Triton X-100]. Klenow (exo-) reactions were performed in 1 X NEBuffer 2 [10 mM Tris-HCl

(pH 7.9), 50 mM NaCl, 10 mM $MgCl_2$, and 1 mM dithiothreitol]. Reactions using *Pfu* (exo-) were performed in 1 X Cloned *Pfu* Buffer [20 mM Tris-HCl (pH 8.8), 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 0.1 mg/mL nuclease-free bovine serum albumin].

Standard deoxynucleoside triphosphates (dNTPs), comprised of 2′-deoxyadenosine-5′-triphosphate (dATP), 2′-deoxycytidine-5′-triphosphate (dCTP), 2′-deoxyganosine-5′-triphosphate (dGTP), and thymidine-5′-triphosphate (TTP), were purchased from Promega (Madison, Wisconsin). dψTNTP solutions contained dATP, dCTP, dGTP, and ψTTP.

Primer-Extension Assays

Primer-template complexes (15 μL total volume) were prepared by mixing 5′-^{32}P labeled primer (P-1, 25 pmol), non-radiolabeled primer (P-1, 200 pmol), and a template (T-1 through T-16, 300 pmol) in a final volume of 15 μL (Table 1). The mixtures were incubated for 5 minutes at 96°C and allowed to cool to room temperature over 1 hour.

Primer-extension assays contained the primer-template complex (1.5 μL), 1 U of the appropriate polymerase (in variable volume), buffer (1 μL of the appropriate 10 X buffer supplied by the manufacturer), and sterile water (to make 9 μL final volume). Mixtures were incubated at the appropriate temperature for 30 seconds. The reaction was then initiated by adding 1 μL of one of the following: 1 mM TTP, 1 mM ψTTP, 1 mM dNTPs, or 1 mM dψTNTPs. The mixtures were incubated for 2 minutes at the appropriate temperature, and quenched with 5 μL of DNA PAGE Loading Dye mix [98% formamide, 10 mM EDTA, 1 mg/mL xylene cyanol FF, and 1 mg/mL bromophenol blue]. Products (in aliquots of 1 μL) were resolved on denaturing PAGE gels [7 M urea, 20% 40:1 acrylamide: bisacrylamide] and analyzed on a Molecular Imager FX System (Bio-Rad, Hercules, CA, USA).

RESULTS

Thermal DNA Duplex Denaturation

A standard primer-template sequence, used in the past for polymerase studies,[27,28] was used as the starting point for the duplexes examined here. Two nucleotides were added to the sequence and ends were modified to make them non-self complementary. The results of this design were oligonucleotide 36-mers that, with their complements, form ca. 3.3 turns of a standard B duplex (Table 1).

A dA-T pair was present at position 18 of the starting duplex, just left of the midpoint. In the initial experiments, this single dA-T pair was replaced by exactly one dA-ψT pair. The drop in T_m was recorded to be a modest

TABLE 2 Thermal DNA duplex denaturation studies

Duplex	T_m (°C)	ΔT_m (°C)	ΔT_m/mod (°C)	ΔH (kcal/mol)	ΔS (cal/mol/K)	ΔG°_{50} (kcal/mol)	$\Delta\Delta G^\circ_{50}$ (kcal/mol)	$\Delta\Delta G^\circ_{50}$/mod (kcal/mol)	Number of ψTs Present
13/18	77	–	–	−91.8	−236.5	−64.4	–	–	–
13/23	76.8	−0.2	−0.2	−84.3	−213.0	−64.8	−0.4	−0.4	1
14/19	76.5	–	–	−68.6	−167.5	−60.5	–	–	–
14/24	74.5	−2	−0.67	−73.4	−182.8	−59.8	0.6	0.2	3
15/20	75	–	–	−94.5	−243.6	−66.3	–	–	–
15/25	70.1	−4.9	−0.82	−132.7	−359.5	−69.2	−3	−0.5	6
16/21	73	–	–	−101.1	−264.4	−65.6	–	–	–
16/26	65	−8	−0.89	−107.8	−292.5	−55.6	10.1	1.1	9
17/22	69.5	–	–	−103.3	−273.8	−61.1	–	–	–
17/27	58.5	−11	−0.92	−97.0	−139.2	−46.8	14.3	1.2	12

The $\Delta\Delta G^\circ_{50}$ values were calculated with the assistance of the Varian Cary WinUV software, which required the user to identify a linear region of the van't Hoff plot. The modest curvature of that plot generated the largest uncertain in the $\Delta\Delta G^\circ_{50}$ values reported here, and presumably account for the modest negative $\Delta\Delta G^\circ_{50}$ in comparing the 15/20 and 15/25 samples. As the experimental T_m values indicate, the duplex containing six ψTs was nevertheless less stable than the duplex containing no ψTs.

0.2°C (Table 2), within the error of the measurement; the corresponding ΔG°_{50} was −0.4 kcal/mol (−64.4 kcal/mol for the T13:T18 standard duplex, −64.8 kcal/mol for the T13:T23 duplex, nominally indicating that the duplex containing one dA-ψT was slightly more stable, the negative sign for the $\Delta\Delta G^{\circ}_{50}$ providing this inference). This was smaller than the 1.5°C drop reported by Geyer et al. upon the replacement of T with ψU.[11] As the methyl group on T is known to contribute to duplex stability, and as ψT has this methyl group, these data present no inconsistencies.

Additional dA-T or dA-ψT pairs were then added on each side of the central dA-T pair to create oligonucleotide duplexes having more dA-ψT pairs (Table 1). In the design, any pair flanking the growing run of dA-Ts or dA-ψTs that was also a dA-T or T-dA pair was changed to dG-dC or dC-dG to "cap" the ends of the run. The details of any DNA sequence modification are expected to depend on overall nucleotide composition as well as local context, and no small number of experiments can explore these factors comprehensively. Nevertheless, the strategy used here ensured that the weakest portion of the duplex (the run of dA-T or dA-ψT pairs) was flanked by the strongest standard base pair, and gave a mix of contexts that might (one hopes) cause context dependence to be largely averaged out.

These experiments showed a general decrease in the melting temperatures (and $\Delta\Delta G^{\circ}_{50}$s) with increasing numbers of dA-ψT pairs, a decrease that became more apparent as the runs became longer (Table 2). For example, the duplex formed by T-14:T-19, with three central dA-T pairs, melted 2.0°C higher than the analogous T-14:T-24 duplex with three consecutive dA-ψT pairs. This reflects an impact of 0.7°C per pair. The corresponding $\Delta\Delta G^{\circ}_{50}$ was +0.2 kcal/mol.

The T-15:T-20 pair (six dA-T pairs) had a T_m that was 4.9°C higher than the T-15:T-25 pair (six dA-ψT pairs, Table 2). The calculated $\Delta\Delta G^{\circ}_{50}$/mod was −0.5 kcal/mol, with a ΔT_m/mod of 0.82°C. For the duplexes containing nine dA-T /ψT pairs, the ΔT_m/mod was −0.9°C and a $\Delta\Delta G^{\circ}_{50}$/mod was +1.1 kcal/mol (Table 2). The modification of the duplex containing 12 consecutive dA-T/ψT base pairs produced a ΔT_m/mod of −0.9°C and a $\Delta\Delta G^{\circ}_{50}$ /mod of +1.2 kcal/mol (Table 2).

Circular Dichroism

To assess whether the presence of multiple consecutive C-glycosides in double-stranded DNA (dsDNA) cause a transition from B-DNA to A-DNA, CD measurements were recorded for duplexes containing from one to twelve consecutive dA-T or dA-ψT base pairs. While CD spectra can often support multiple interpretations about the details of molecular structure, they should be diagnostic for a global phase transition from B-DNA to A-DNA. Classical work suggests that a conversion of canonical B-DNA to canonical A-DNA should produce a shift of the positive portion of the CD

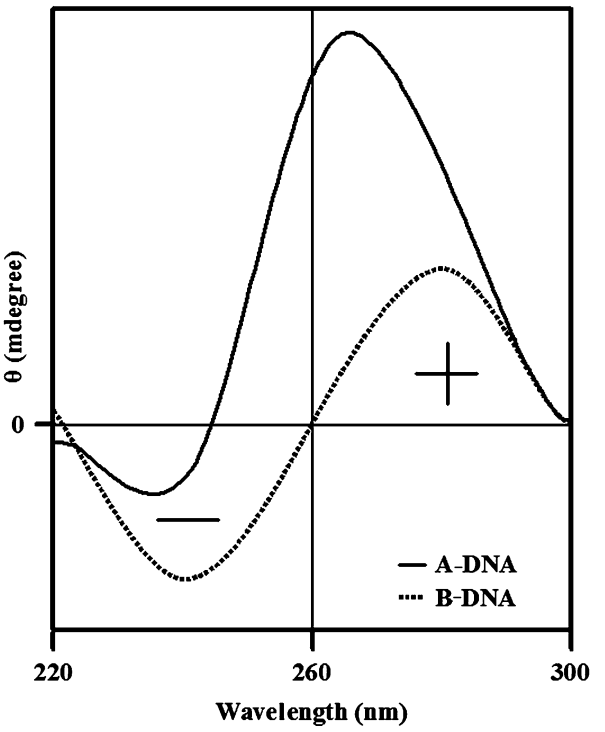

FIGURE 3 Idealized CD spectra of A- and B-DNA forms. Note the shift of the positive potion of the spectrum to shorter wavelengths, from 275 nm in B-DNA to 267 nm in A-DNA.[17] A similar shift with a similar magnitude can be seen in the negative portion, and there is a stronger Cotton effect shown by the A-DNA than the B-DNA.

spectrum to shorter wavelengths, from 275 nm for the B-form to 267 nm for the A-form (Figure 3).[17] A similar shift with a similar magnitude is seen in the negative portion of the spectrum. Further, A-DNA shows a stronger Cotton effect than the B-DNA. Therefore, an increase in the Cotton effect and a shift towards shorter wavelengths should be diagnostic for whether the addition of C-glycosidic units drives the conformation of the duplex from B towards A.

CD spectra were recorded on ten duplex DNA molecules containing anywhere from one to twelve consecutive dA-T or dA-ψT base pairs (Table 1). The observed spectra (Figures 4A–4E) were compared to those in Figure 3, the reference spectra for canonical A and B duplexes. In all spectra containing ψT, the maximum of the positive portion of the CD spectrum wavelength shifted at most 2 nm toward longer wavelengths, a magnitude that is negligible, and a direction opposite to that expected if the B-DNA were converting to A-DNA. No trends were observed in the shift no matter how many dA-ψT pairs were present in the duplex.

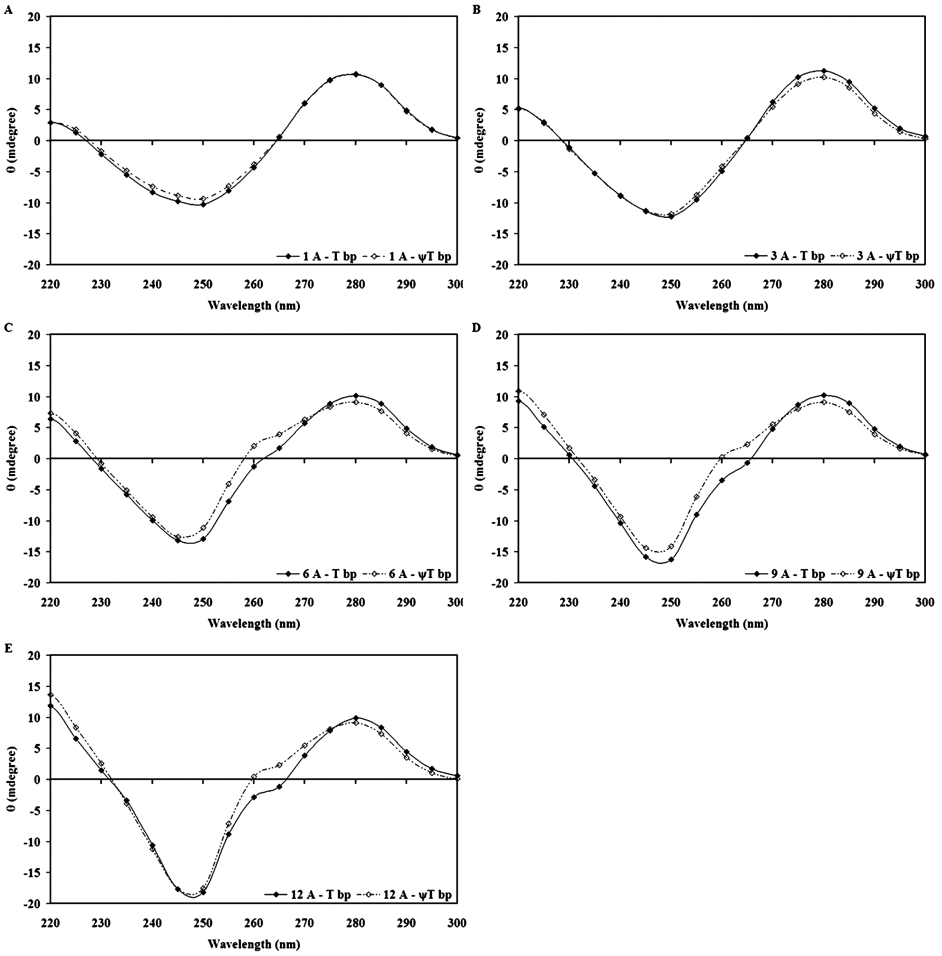

FIGURE 4 Circular dichroism (CD) spectra of double stranded templates with their complements containing different numbers of dA-T or dA-ψT base pairs at 25°C (see Table 1 for sequences). All spectra indicate a B-DNA conformation,[17] and that duplex conformation does not change with increasing numbers of ψT. A) Spectra of duplexes containing one dA-T base pair vs. one dA-ψT base pair. B) Spectra of duplexes containing three dA-T base pairs vs. three dA-ψT base pairs. C) Spectra of duplexes containing six dA-T base pairs vs. six dA-ψT base pairs. D) Spectra of duplexes containing nine dA-T base pairs vs. nine dA-ψT base pairs. E) Spectra of duplexes containing twelve dA-T base pairs vs. twelve dA-ψT base pairs.

The only possible trend is a change in the relative intensity of the positive (at 275 nm) and negative (at 246 nm) band intensities.[17] Here the intensity of the 246 nm band and the 275 nm band increase and decrease (respectively) as more dA-ψT pairs were present in the duplex. As concentrations were carefully controlled, we do not believe that this reflects a change in the concentration of the oligonucleotides. One of these small trends is consistent with, and the other is inconsistent with, the conversion of the duplex structure from canonical B- to canonical A-DNA, but their

small sizes is most consistent with no global change in conformation arising from the introduction of up to 12 C-glycoside units.

Likewise, little evidence was found for a conformational change as the temperature was raised (Figures 5A–5E, and 5G). As the temperature of the sample is increased towards the melting temperature of the duplex containing twelve dA-ψT base pairs (4.0 to 58.5°C), the intensity of the 246 nm band increases slightly and the 275 nm band decreases slightly, again one consistent and the other inconsistent (in the direction but not magnitude) for a B- to A-shift. The wavelengths of the maxima/minima again do not shift consistently. Similar effects are seen within the series simply by raising the temperature, consistent with CD spectra detecting the initial stages of melting of the duplex.

Polymerase Screen Primer-Extension Assays

Four Family A polymerases, Klenow (exo-), *Bst*, *Taq*, and *Tth*, and four Family B polymerases, Deep Vent (exo-), Vent (exo-), *Pfu* (exo-), and Therminator, were screened for the ability to incorporate TTP and dψTTP across from template dA in both 4-base and 13-base primer extension assays (Figures 6A and 6B). In the 4-base extension assay, polymerases were challenged to incorporate four consecutive TTP and ψTTP across from template dA during two-minute incubations at the optimal temperature for each enzyme. The 13-base assay, incubated as described above, took place in the presence of dCTP, dGTP, dATP, and TTP or ψTTP and required incorporation and extension beyond the four consecutive TTP or ψTTP.

Pfu (exo-) was the only polymerase that was not able to generate full-length product (FLP) when challenged to incorporate and extend after incorporation of a run of ψTTPs. All other Family A and Family B polymerases were able to incorporate the four ψTTPs, and then continue the extension of the primer to generate FLP, at least to a detectable amount. Thus, in these experiments, Bst and Therminator polymerases appeared to have consumed virtually all of the primer to generate large amounts of FLP. Klenow (exo-) and Vent (exo-) also incorporated ψTTP well. The other polymerases gave substantial amounts of product that indicated pausing during or after incorporation of the run of ψTTPs.

Taq Polymerase Primer-Extension Assays

Two of these polymerases, Klenow (exo-) and *Bst*, have lower thermostabilities, and therefore are not useful for standard PCR. According to the manufacturer, Therminator is not recommended for any applications except DNA sequencing and primer-extension reactions. We therefore set these aside in making our selection for more detailed studies, as many

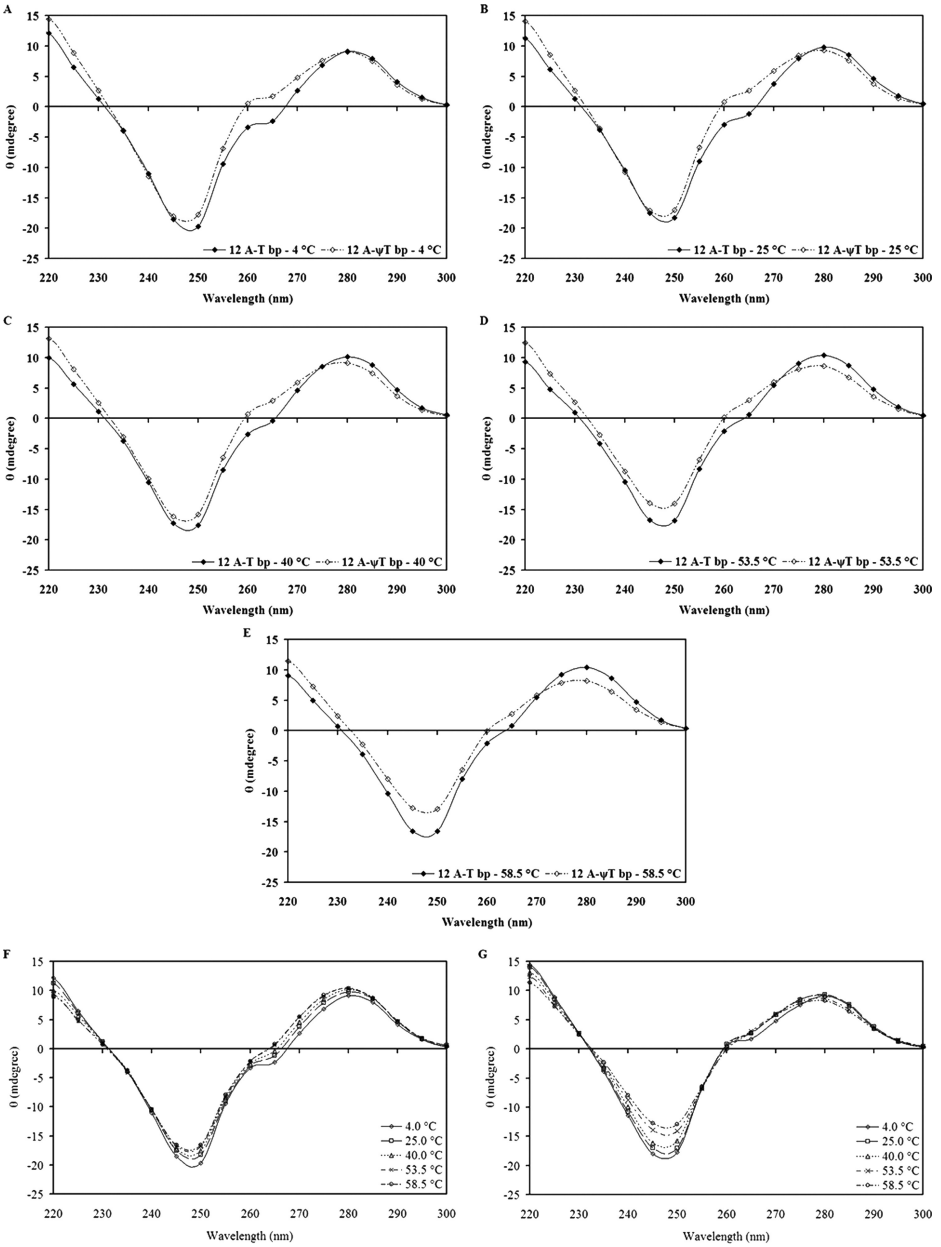

FIGURE 5 Circular dichroism (CD) spectra of double stranded templates with their complements containing twelve dA-T or dA-ψT base pairs at various temperatures (see Table 1 for sequences). All spectra indicate a B-DNA conformation,[17] and that an approach in temperature towards the melting point of the duplex is associated with only a small decrease in the intensity of the negative band (ca. 246 nm) for both helices. Spectra above the line (A–E) show the direct comparison between duplexes containing twelve dA-T base pairs and duplexes containing twelve dA-ψT base pairs at the indicated temperatures. Figures below the line (F-G) directly compare the spectra observed at various temperatures for each helix. F) Spectra of the duplex containing twelve dA-T base pairs at indicated temperatures; G) spectra of the duplex containing twelve dA-ψT base pairs at indicated temperatures.

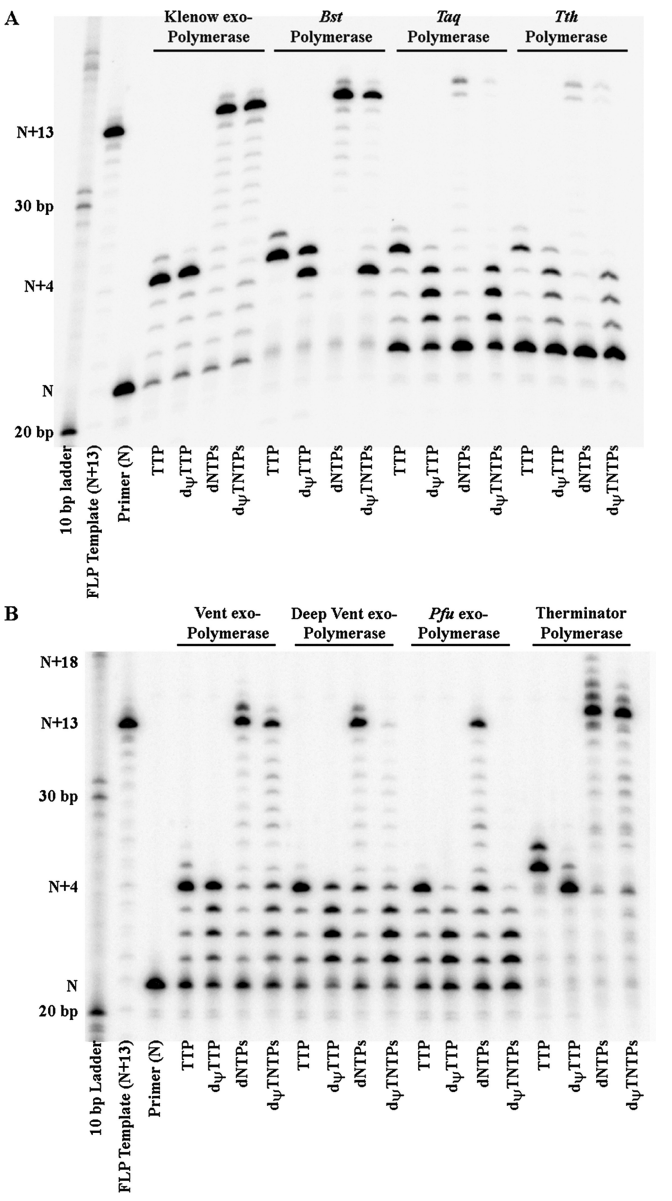

FIGURE 6 Family A and B polymerase screens. Unextended primer is at position N; N+4 is the full-length product (FLP) for the 4-base extension assays; N+13 is the FLP for the 13-base extension assays. Final concentrations: TTP/ψTTP/dNTPs/dψTNTPs (100 μM), radiolabeled P-1 (2.5 pmol), non-radiolabeled P-1 (20 pmol), non-radiolabeled template T-4 (30 pmol), and appropriate polymerase (1 U). A) The incorporation and extension of T and ψT by various Family A polymerases. All polymerases were able to incorporate and extend beyond the four consecutive A-T or A-ψT base pairs to generate some FLP in both the 4-base and 13-base extension assays. Klenow (exo-) and *Bst* most likely generated higher amounts of ψT containing FLP since their optimal temperatures are lower than that of *Taq* and *Tth*.[34,35] B) The incorporation and extension of T and ψT by various Family B polymerases. All polymerases, except *Pfu* (exo-), were able to incorporate and extend beyond the four consecutive A-T or A-ψT base pairs to generate some FLP in both the 4-base and 13-base extension assays. *Pfu* (exo-) was able to generate FLP in the 4-base assay, but not the 13-base assay. Therminator was able to incorporate multiple ψT units, as indicated by the low levels of unextended primer remaining in those lanes.

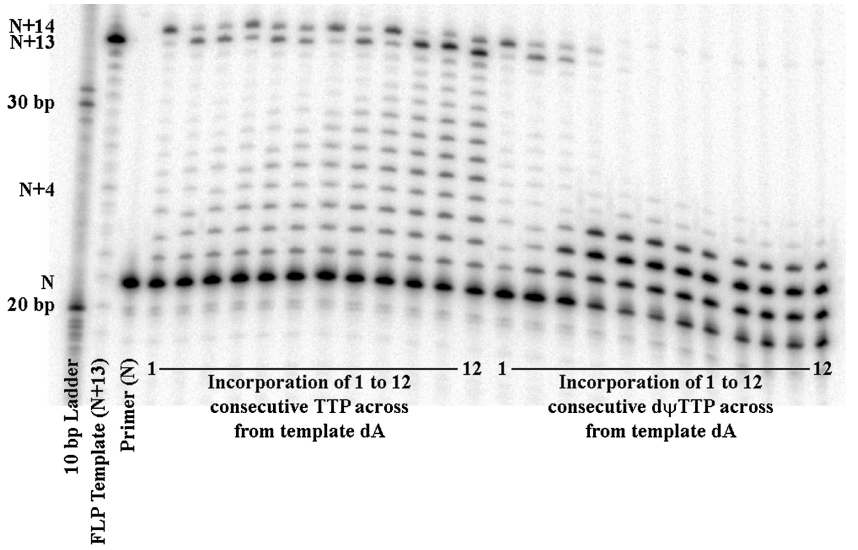

FIGURE 7 Incorporation of one to twelve consecutive T or ψT residues by *Taq* polymerase. Unextended primer is at position N; FLP is denoted by N+13 in all of these assays. Full-length product (N+13) is accompanied by a band at N+14, assigned to a product that has one additional untemplated nucleotide added to the end. While *Taq* easily incorporates one or two ψTs, it does not generate full-length product when challenged to incorporate five or more consecutive ψTs.

architectures for systems biology and synthetic biology that use these C-glycosides also contemplate PCR.

Because of its widespread use in these architectures, *Taq* was selected to be examined further.[12] Primer extension experiments were performed under optimal polymerase conditions using templates T-1 through T-12 (Figure 7). These challenged *Taq* to incorporate from one to 12 consecutive ψTTPs. *Taq* easily incorporates one ψTTP to give full-length product (N+13), together with the N+14 band that is assigned to the product to which has been added a nontemplated nucleotide.[29] Two ψTTPs also appear to be incorporated easily. Decreasing amounts of full-length product (as well as the N+14 product), and increasing amounts of pausing bands, are observed as the number of ψTTPs incorporated increases, with *Taq* being unable to generate detectable amounts of FLP when challenged to incorporate six or more ψTTPs. These data also suggest that *Taq* had difficulty maintaining processivity for long stretches of dA-T base pairs. This has been observed in our laboratory before, and to our knowledge, has no known explanation.

DISCUSSION

The results presented here offer no evidence that adding multiple, consecutive C-glycosides to a strand in a duplex causes the duplex to change

its conformation overall. Further, these results suggest that polymerases such as *Taq*, *Tth*, Vent (exo-), and Deep Vent (exo-) all incorporate multiple consecutive C-glycosides, with *Taq* performing passably, even when challenged to incorporate four consecutive ψTs.

The observation that the addition of multiple ψT (from 1 to 12) is associated with a constant drop in ΔT_m per modification ($-0.9°C$) and a variable $\Delta\Delta G°_{50}$ (from -0.5 to $+1.2$ kcal/mol per modification) is consistent with the absence of a change in conformation.

A variety of hypotheses can be raised to explain why replacing a dA-T pair by a dA-ψT pair causes (on average) a $-0.9°C$ drop in Tm. One hypothesis recognizes that the pair between a thymidine and an adenine uses only one of the two thymidine carbonyl groups, in particular, the one that *is not* between the two heterocyclic ring nitrogens (Figure 2). The pair between a ψT and an adenine uses the *other* carbonyl group, the one that *is* between the two nitrogens. If the two oxygens have different abilities to accept a hydrogen bond, then the one that *is* between the two heterocyclic ring nitrogens must act as an inferior hydrogen-bond acceptor.

Both experiments and quantum mechanical calculations support this hypothesis. In experiments using fast-atom-bombardment tandem mass-spectrometry, Greco et al. suggested that O-4 (the oxygen used in the dA-T pair and not in the dA-ψT pair) has a higher affinity for protons, implying that it is a better hydrogen-bond acceptor.[30] Using theory, density functional theory (DFT) calculations have made a similar implication.[31] Thus, the ΔT_m/mod of $-0.9°C$ may simply reflect the use by ψT of the less basic carbonyl group to form a hydrogen bond with adenine, while T uses the more basic carbonyl group.

Turning to enzymological results, seven polymerases [Klenow (exo-), *Bst*, *Taq*, *Tth*, Vent (exo-), Deep Vent (exo-), and Therminator] were found to create four consecutive dA-ψT base pairs and continue extension. Further testing of *Taq* polymerase demonstrated its ability to incorporate five, but not more, consecutive C-glycoside residues. This number is significant, as the gene for *Taq* has four consecutive dA-T pairs as its longest run. Therefore, *Taq* polymerase should be able to replicate its own gene even if ψT is used instead of T. This can then be utilized as a starting point for directed evolution experiments to improve its performance under this challenge. Experiments to this end are currently under way.

REFERENCES

1. Kool, E.T. Replacing the nucleobases in DNA with designer molecules. *Acc. Chem. Res.* **2002**, 35, 936–943.
2. Benner, S.A. Understanding nucleic acids using synthetic chemistry. *Acc. Chem. Res.* **2004**, 37, 784–797.
3. Tor, Y.; Dervan, P.B. Site-specific enzymatic incorporation of an unnatural base, N(6)-(6-aminohexyl)isoguanosine, into RNA. *J. Am. Chem. Soc.* **1993**, 115, 4461–4467.

4. Marx, A.; Detmer, I.; Gaster, J.; Summerer, D. Probing DNA polymerase function with synthetic nucleotides. *Synthesis-Stuttgart* **2004**, 1–14.
5. Hwang, G.T.; Romesberg, F.E. Substituent effects on the pairing and polymerase recognition of simple unnatural base pairs. *Nucleic Acids Res.* **2006**, 34, 2037–2045.
6. Hirao, I. Unnatural base pair systems for DNA/RNA-based biotechnology. *Curr. Opin. Chem. Biol.* **2006**, 10, 622–627.
7. Wellington, K.W.; Benner, S.A. A review: Synthesis of aryl C-glycosides via the Heck coupling reaction. *Nucleosides, Nucleotides Nucleic Acids* **2006**, 25, 1309–1333.
8. Lee, A.H.F.; Kool, E.T. Novel benzopyrimidines as widened analogues of DNA bases. *J. Org. Chem.* **2005**, 70, 132–140.
9. Sismour, A.M.; Benner, S.A. Synthetic biology. *Exp. Opin. Biol.l Ther.* **2005**, 5, 1409–1414.
10. Charette, M.; Gray, M.W. Pseudouridine in RNA: What, where, how, and why. *International Union of Biochemistry and Molecular Biology Life* **2000**, 49, 341–351.
11. Geyer, C.R.; Battersby, T.R.; Benner, S.A. Nucleobase pairing in Watson-Crick-like genetic expanded information systems. *Structure* **2003**, 11, 1485–1498.
12. Lutz, S.; Burgstaller, P.; Benner, S.A. An in vitro screening technique for DNA polymerases that can incorporate modified nucleotides. Pseudo-thymidine as a substrate for thermostable polymerases. *Nucleic Acids Res.* **1999**, 27, 2792–2798.
13. Sismour, A.M.; Lutz, S.; Park, J.H.; Lutz, M.J.; Boyer, P.L.; Hughes, S.H.; Benner, S.A. PCR amplification of DNA containing non-standard base pairs by variants of reverse transcriptase from human immunodeficiency virus-1. *Nucleic Acids Res.* **2004**, 32, 728–735.
14. Piccirilli, J.A.; Moroney, S.E.; Benner, S.A. A C-nucleotide base pair: methylpseudouridine-directed incorporation of formycin triphosphate into RNA catalyzed by T7 RNA polymerase. *Biochem.* **1991**, 30, 10350–10356.
15. Davis, D.R. Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res.* **1995**, 23, 5020–5026.
16. Plavec, J.; Tong, W.M.; Chattopadhyaya, J. How Do the Gauche and anomeric effects drive the pseudorotational equilibrium of the pentofuranose moiety of nucleosides. *J. Am. Chem. Soc.* **1993**, 115, 9734–9746.
17. Ivanov, V.I.; Minchenk, L.E.; Schyolki, A.K.; Poletaye, A.I. Different conformations of double-stranded nucleic-acid in solution as revealed by circular-dichroism. *Biopolymers* **1973**, 12, 89–110.
18. Saenger, W. *Principles of Nucleic Acid Structure*. Springer-Verlag: New York; 1984.
19. Chandrasekaran, R.; Radha, A. Structure of poly D(a). Poly D(T). *J. Biomol. Struct. Dyn.* **1992**, 10, 153–168.
20. Chandrasekaran, R.; Radha, A.; Park, H.S.; Arnott, S. Structure of the beta-form of poly-D(a).poly-D(U). *J. Biomol.Struct. Dyn.* **1989**, 6, 1203–1215.
21. Bhattacharya, B.K.; Devivar, R.V.; Revankar, G.R. A practical synthesis of N1-methyl-2′-deoxy-psi-uridine (psi-thymidine) and its incorporation into G-rich triple-helix forming oligonucleotides. *Nucleosides Nucleotides* **1995**, 14, 1269–1287.
22. Zhang, H.C.; Daves, G.D. Syntheses of 2′-deoxypseudouridine, 2′-deoxyformycin-B, and 2′,3′-dideoxyformycin-B by palladium-mediated glycal aglycon coupling. *J. Org. Chem.* **1992**, 57, 4690–4696.
23. Ludwig, J.; Eckstein, F. Rapid and efficient synthesis of nucleoside 5′-O-(1-thiotriphosphates), 5′-triphosphates and 2′,3′-cyclophosphorothioates using 2-chloro-4h-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.* **1989**, 54, 631–635.
24. Sambrook, J.; Fritsch, E.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, **1989**.
25. Marky, L.A.; Breslauer, K.J. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **1987**, 26, 1601–1620.
26. Breslauer, K.J. Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions, in *Methods in Enzymology: Energetics of Biological Macromolecules*, eds. M.L. Johnson, G.K. Akers, Academic Press, San Diego, CA, 1995, pp. 221–242.
27. Sismour, A.M.; Benner, S.A. The use of thymidine analogs to improve the replication of an extra DNA base pair: a synthetic biological system. *Nucleic Acids Res.* **2005**, 33, 5640–5646.
28. Hendrickson, C.L.; Devine, K.G.; Benner, S.A. Probing minor groove recognition contacts by DNA polymerases and reverse transcriptases using 3-deaza-2′-deoxyadenosine. *Nucleic Acids Res.* **2004**, 32, 2241–2250.

29. Hu, G.X. DNA Polymerase-catalyzed addition of nontemplated extra nucleotides to the 3′ end of a DNA fragment. *DNA Cell Biol.* **1993**, 12, 763–770.
30. Greco, F.; Liguori, A.; Sindona, G.; Uccella, N. Gas-phase proton affinity of deoxyribonucleosides and related nucleobases by fast-atom-bombardment tandem mass-spectrometry. *J. Am. Chem. Soc.* **1990**, 112, 9092–9096.
31. Zhu, W.L.; Luo, X.M.; Puah, C.M.; Tan, X.J.; Shen, J.H.; Gu, J.D.; Chen, K.X.; Jiang, H.L. The multiplicity, strength, and nature of the interaction of nucleobases with alkaline and alkaline earth metal cations: A density functional theory investigation. *J. Phys.l Chem.y A* **2004**, 108, 4008–4018.
32. Switzer, C.; Moroney, S.E.; Benner, S.A. Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **1989**, 111, 8322–8323.
33. Piccirilli, J.A.; Krauch, T.; Moroney, S.E.; Benner, S.A. Enzymatic Incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **1990**, 343, 33–37.
34. Rappaport, H.P. The fidelity of replication of the three-base-pair set adenine/thymine, hypoxanthine/cytosine and 6-thiopurine/5-methyl-2-pyrimidinone with T7 DNA polymerase. *Biochem. J.* **2004**, 381, 709–717.
35. Horlacher, J.; Hottiger, M.; Podust, V.N.; Hubscher, U.; Benner, S.A. Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. *Proc. Nat. Acad. Sci. USA* **1995**, 92, 6329–6333.