Optimization of non-natural nucleotides for selective incorporation opposite damaged DNA

Diana Vineyard,[a] Xuemei Zhang,[a] Alison Donnelly,[b] Irene Lee[b] and Anthony J. Berdis*[a]

Received 15th August 2007, Accepted 28th September 2007
First published as an Advance Article on the web 12th October 2007
DOI: 10.1039/b712480e

The promutagenic process known as translesion DNA synthesis reflects the ability of a DNA polymerase to misinsert a nucleotide opposite a damaged DNA template. To study the underlying mechanism of nucleotide selection during this process, we quantified the incorporation of various non-natural nucleotide analogs opposite an abasic site, a non-templating DNA lesion. Our kinetic studies using the bacteriophage T4 DNA polymerase reveal that the π-electron surface area of the incoming nucleotide substantially contributes to the efficiency of incorporation opposite an abasic site. A remaining question is whether the selective insertion of these non-hydrogen-bonding analogs can be achieved through optimization of shape and π-electron density. In this report, we describe the synthesis and kinetic characterization of four novel nucleotide analogs, 5-cyanoindolyl-2′-deoxyriboside 5′-triphosphate (5-CyITP), 5-ethyleneindolyl-2′-deoxyriboside 5′-triphosphate (5-EyITP), 5-methylindolyl-2′-deoxyriboside 5′-triphosphate (5-MeITP), and 5-ethylindolyl-2′-deoxyriboside 5′-triphosphate (5-EtITP). Kinetic analyses indicate that the overall catalytic efficiencies of all four nucleotides are related to their base-stacking properties. In fact, the catalytic efficiency for nucleotide incorporation opposite an abasic site displays a parabolic trend in the overall π-electron surface area of the non-natural nucleotide. In addition, each non-natural nucleotide is incorporated opposite templating DNA ~100-fold worse than opposite an abasic site. These data indicate that selectivity for incorporation opposite damaged DNA can be achieved through optimization of the base-stacking properties of the incoming nucleotide.

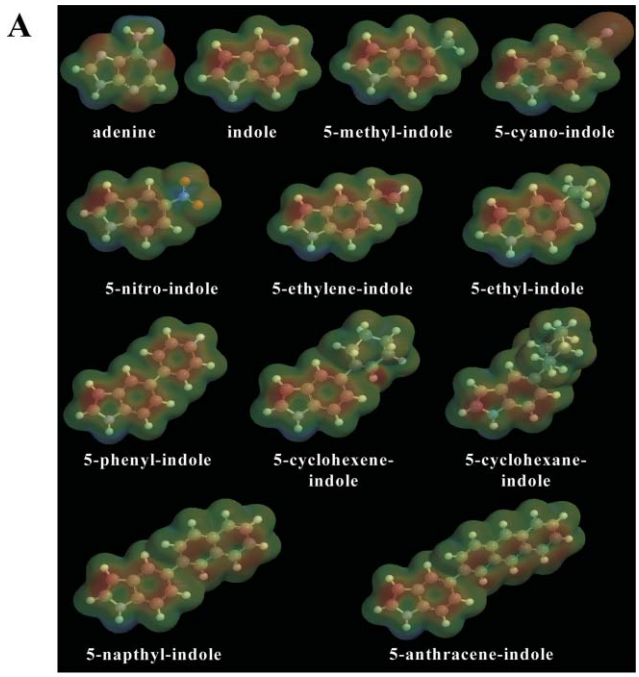
Introduction

DNA polymerases catalyze the incorporation of a nucleoside triphosphate into a growing polymer chain using the template strand as coding information. Mechanistic and structural studies of DNA polymerases have demonstrated that hydrogen-bonding and geometrical constraints play important roles for nucleotide selection during correct DNA synthesis.[1,2] However, the mechanisms by which polymerases mis-replicate DNA still remain undefined, as the rules of base-pairing based solely on hydrogen-bonding patterns and steric fit are violated in many cases.[3–5] The best example of this phenomenon occurs during the process of translesion DNA synthesis, which reflects the ability of a polymerase to incorporate opposite a DNA lesion. One common lesion is an abasic site that lacks coding information in the template strand. Abasic sites can be generated non-enzymatically[6,7] or enzymatically through the action of DNA repair pathways.[8,9] Despite the lack of coding information at an abasic site, we demonstrated that the bacteriophage T4 DNA polymerase preferentially inserts dATP opposite this type of lesion,[10] a phenomenon known as the "A-rule".[11–13] By exploring the mechanism underlying the "A-rule", we developed a strategy to inhibit the propagation of genomic errors by using 5-substituted indolyl triphosphates that are selectively incorporated opposite this non-templating DNA lesion.[14] Although these non-natural nucleotides differ with respect to solvation energy, dipole moment, and shape/size (Fig. 1A), many are incorporated very effectively opposite the DNA lesion. In general, analogs containing extensive π-electron surface areas such as 5-nitroindolyl-2′-deoxyriboside 5′-triphosphate (5-NITP),[15] 5-phenylindolyl-2′-deoxyriboside 5′-triphosphate (5-PhITP),[16] and 5-napthylindolyl-2′-deoxyriboside 5′-triphosphate (5-NapITP)[17] are incorporated opposite the abasic site with incredibly high catalytic efficiencies that result from high binding affinity (low K_d values of <20 μM) coupled with fast incorporation rate constant (k_{pol} values >25 s^{-1}). In addition, most of these non-natural nucleotides are poorly incorporated opposite templating nucleobases.[15–17] Surprisingly, this discrimination occurs through vast reductions in the k_{pol} values during incorporation opposite templating nucleobases rather than through alterations in binding affinity as predicted by a simple steric fit model.

These results lead us to hypothesize that selectivity for incorporation opposite an abasic site could be achieved by balancing the contributions of π-electron density with the relative size of the nucleobase.[18] In this model, increasing the π-electron surface area of the incoming nucleotide should allow for favorable stacking interactions within the void present at an abasic site. In addition, increasing its overall size should sterically perturb interactions with templating bases and thus prevent incorporation opposite undamaged DNA. Indeed, this approach works with a subset of indolyl analogs such as 5-PhITP and 5-cyclohexylindolyl-2′-deoxyriboside 5′-triphosphate (5-CEITP) which are ~1000-fold

[a]*Department of Pharmacology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH, 44106, USA. E-mail: ajb15@case.edu; Fax: +1-216-368-3395; Tel: +1-216-368-4723*
[b]*Department of Chemistry, Case Western Reserve University, 2109 Adelbert Road, Cleveland, OH, 44106, USA*

This journal is © The Royal Society of Chemistry 2007



B 5′-TCGCAGCCGTCCA
3′-AGCGTCGGCAGGTXCCCAAA

X = A,C,G,T, or abasic site

Fig. 1 (A) Structures of non-natural nucleobases used or referred to in this study. Surface ionization potentials for each nucleobase are provided for comparison and were generated using Spartan '04 software. Red indicates the highest electronegative regions, green is neutral, and blue indicates electropositive regions. The partial atomic charges were calculated using the DFT 3-21G** model. (B) Defined DNA substrates used for kinetic analysis. The X in the template strand denotes any of the four natural nucleobases or the presence of 1′,2′-dideoxyribose designed to functionally mimic an abasic site.

more selective for translesion DNA synthesis. However, selectivity diminishes significantly as the π-electron surface area of a nucleobase increases beyond 200 Å^2. The reduced selectivity results primarily from an unexpected increase in the efficiency for incorporation opposite templating DNA. These results imply that nucleotides with large π-electron surface areas can hinder fidelity through an increased propensity for non-selective intercalation during the polymerization process.

In this report, we quantified the selective incorporation of several non-natural nucleotides possessing π-electron surface areas less than 180 Å^2. In general, these analogs are incorporated opposite an abasic site with high catalytic efficiency ($\sim 10^6$ M^{-1} s^{-1}) regardless of π-electron density at the 5-position of the indole ring. These data argue that nucleotide incorporation opposite a non-templating DNA lesion can occur *via* enhanced base-stacking interactions caused by reductions in solvation energies and proper steric arrangements rather than through overall increases in π-electron surface area. All four analogs are poorly incorporated opposite templating nucleobases, as their overall catalytic efficiencies are 100-fold lower than during translesion DNA synthesis. Collectively, these data indicate that selectivity for translesion DNA synthesis can be achieved by judiciously optimizing the base-stacking properties of the incoming nucleotide.

Results and discussion

Synthesis and chemical characterization of 5-cyanoindolyl-2′-deoxyriboside 5′-triphosphate (5-CyITP), 5-ethylindolyl-2′-deoxyriboside 5′-triphosphate (5-EtITP), 5-ethyleneindolyl-2′-deoxyriboside 5′-triphosphate (5-EyITP) and 5-methylindolyl-2′-deoxyriboside 5′-triphosphate (5-MeITP)

We previously synthesized a series of non-natural nucleotide analogs that are incorporated opposite an abasic site with high efficiency and selectivity.[15–19] In general, the overall catalytic efficiency for incorporation is dependent upon the presence of significant π-electron density at the 5-position of the indole ring. The goal of this study is to expand the repertoire of these non-natural nucleotides to include those with smaller π-electron surface areas. The indole triphosphate derivatives reported in this study were synthesized using previously described protocols.[19] In brief, we first generated the 5-substituted indole nucleosides using the protocol of Girgis *et al.*[20] The corresponding triphosphates were then synthesized using the procedure originally established by Smith and coworkers.[21] As previously reported, the yields of the nucleoside triphosphates were limited by the competition of an acid-catalyzed furanoside–pyranoside isomerization step.[19] As such, the proton sponge, 1,8-bis(dimethylamino)naphthalene, was included in the reactions to suppress the competing reactions. The structures of the nucleosides were verified by proton nuclear magnetic resonance (^{1}H-NMR) and mass spectrometry (MS) analyses.

Enzymatic incorporation opposite an abasic site

The kinetic parameters for the incorporation of 5-CyITP, 5-EyITP, 5-EtITP and 5-MeITP (Fig. 1A) opposite an abasic site were measured using a defined DNA substrate (Fig. 1B). Time courses were generated using single turnover conditions in which 1 μM exonuclease-deficient T4 DNA polymerase (gp43 exo$^-$) was incubated with 250 nM DNA and mixed with variable concentrations of non-natural nucleotide (5–1000 μM) and 10 mM Mg^{2+}. Representative time courses for the incorporation of 5-CyITP opposite an abasic site are provided in Fig. 2A. Each time course was fitted using an equation describing a single exponential

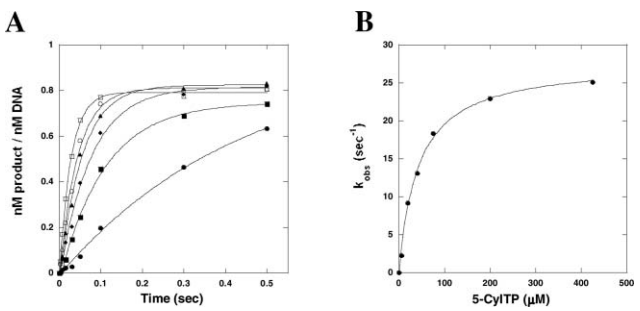


Fig. 2 (A) Rate constants for the incorporation of 5-CyITP opposite an abasic site on as a function of nucleotide concentration. Assays were performed using 1 μM gp43 exo$^-$, 250 nM 13/20 SP, 10 mM $Mg(OAc)_2$, and 5-CyITP in variable concentrations: 5 μM (●), 20 μM (■), 40 μM (◆), 75 mM (▲), 200 μM (○), and 425 μM (□). The solid lines represent the fit of the data to a single exponential. (B) Measured rate constants for incorporation (●) were plotted against 5-CyITP concentration and fitted using the Michaelis–Menten equation to determine values corresponding to K_d and k_{pol}.

This journal is © The Royal Society of Chemistry 2007

Table 1 Summary of kinetic parameters for the incorporation of non-natural nucleotides opposite an abasic site[a]

Nucleotide	k_{pol}/M^{-1} s^{-1}	K_d/μM	(k_{pol}/K_d)/M^{-1} s^{-1}	π-Electron surface area[b]/Å^2	log$P$[c]	Dipole moment[d]	Solvation energy[e]/kJ mol^{-1}	Volume[f]/Å^3
dATP[g]	0.15 ± 0.01	35 ± 5	4.6 × 10^3	145.86	−1.5	2.57	−18.90	124.49
IndTP[h]	0.28 ± 0.07	145 ± 10	1.9 × 10^3	148.11	1.64	2.11	−5.45	132.83
5-MeITP	60.1 ± 3.6	58 ± 10	7.6 × 10^5	148.11	2.12	1.92	−5.29	151.01
5-CyITP	29.0 ± 1.4	79 ± 13	5.0 × 10^5	168.87	1.67	7.18	−7.03	152.29
5-EyITP	94.0 ± 4.7	50 ± 8	1.9 × 10^6	181.45	2.27	2.56	−4.67	165.07
5-EtITP	88.5 ± 4.5	86 ± 12	1.0 × 10^6	148.11	2.54	1.90	−5.07	167.57
5-NITP[i]	126 ± 7	18 ± 3	7.0 × 10^6	174.00	1.67	7.32	−6.92	154.52
5-PhITP[j]	53 ± 4	14 ± 3	3.8 × 10^6	225.16	3.31	2.66	−6.59	215.64
5-CEITP[k]	25 ± 2	5.1 ± 1.7	5.5 × 10^6	238.02	3.15	2.06	−5.11	224.15
5-CHITP[k]	0.70 ± 0.13	44 ± 14	7.4 × 10^4	148.11	3.63	1.97	−4.23	228.93
5-NapITP[l]	27 ± 4	10 ± 5	2.6 × 10^6	274.59	4.31	2.73	−7.21	267.54
5-AnITP[l]	5.3 ± 0.4	25 ± 1.9	2.0 × 10^5	318.87	5.31	2.91	−8.17	321.53

[a] k_{pol}, K_d, and k_{pol}/K_d values were obtained using single turnover reaction conditions (1 μM gp43 exo$^-$, 250 nM DNA substrate, 10 mM Mg^{2+}, and varying concentrations of non-natural nucleotide triphosphate (5–1000 μM). [b] π-electron surface area was calculated using Spartan 2004 software. In all cases, the calculated value represents the surface area of the indole ring in addition to the surface area of the conjugated species of the substituent group. [c] The log$P$ values, the oil-to-water partition coefficients, were calculated using Spartan 2004 software and are used as an indicator of relative hydrophobicity. [d] Dipole moments were calculated using Spartan 2004 software and are used as an indicator of relative electronegativity. [e] Solvation energies were calculated using Spartan 2004 software. [f] Volumes were calculated using Spartan 2004 software are used as an indicator of relative size of the nucleobase. [g] Values taken from Berdis.[10] [h] Values taken from Zhang *et al.*[19] These values are representative of k_{cat}/K_m in which assays were performed using 50 nM gp43, 1000 nM DNA substrate, and variable concentrations of IndTP in the presence of 10 mM Mg^{2+}. [i] Values taken from Reineks and Berdis.[15] [j] Values taken from Zhang *et al.*[16] [k] Values taken from Zhang *et al.*[22] [l] Values taken from Zhang *et al.*[17]

process to define k_{obs}, the rate constant in product formation. The plot of k_{obs} *versus* 5-CyITP concentration is hyperbolic (Fig. 2B), and a fit of the data to the Michaelis–Menten equation yields a k_{pol} value of 29.0 ± 1.4 s^{-1}, a K_d value of 58 ± 10 μM, and a k_{pol}/K_d of 5 × 10^5 M^{-1} s^{-1}. Similar analyses were performed for the incorporation of 5-EyITP, 5-EtITP and 5-MeITP opposite the abasic site (data not shown). Kinetic rate and dissociation constants for the incorporation of all four analogs opposite an abasic site are summarized in Table 1.

Our structure–activity analyses begin by comparing the kinetic parameters for 5-CyITP with those reported for 5-NITP, since both analogs are similar with respect to size, solvation energies, dipole moments, and presence of π-electron density (Table 1). Despite these biophysical similarities, it is clear that the kinetic parameters for 5-CyITP differ from those reported for 5-NITP. Specifically, the k_{pol} of 29 s^{-1} for 5-CyITP is ~4-fold slower than the k_{pol} value of 126 s^{-1} for 5-NITP.[15] In addition, the K_d of 58 μM for 5-CyITP is 3-fold higher than 18 μM for 5-NITP.[15] A possible mechanism to explain this dichotomy is provided by comparing the ionization potentials of both non-natural nucleotides analogs (Fig. 1A). The nitrogen atom of the nitro substituent group of 5-NITP possesses a partial positive charge (indicated by the blue color in the ionization potential map). As illustrated in Fig. 3A, this functional group can interact with the O4 group of thymine at the $n + 1$ position in the template such that the induced dipole–dipole interactions could account for the ability of 5-nitroindole to stack within the void of the abasic site. The resulting stabilization of 5-NITP in an interhelical conformation is manifest by its exquisite catalytic efficiency (~7 × 10^6 M^{-1} s^{-1}),[15] which is unmatched by any other 5-substituted indolyl analog to date.

We next compare the kinetic data for 5-EyITP with that for 5-NITP. Although these analogs are similar with respect to π-electron density, they differ significantly in overall shape, size, solvation energies, and dipole moment. Despite these differences,

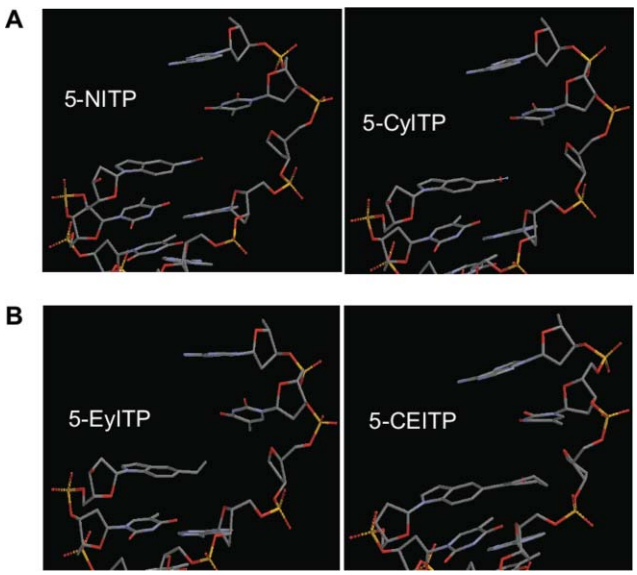


Fig. 3 Theoretical models for the structures of various non-natural nucleotides paired opposite an abasic site. All models were constructed using Spartan '04 and are designed to illustrate the influence of π-electron surface area, shape, and size on the overall catalytic efficiency for incorporation.

the kinetic parameters for 5-EyITP are nearly identical to those for 5-NITP. In particular, the k_{pol} of 94 s^{-1} is only 1.3-fold slower than that of 126 s^{-1} measured for 5-NITP.[15] In addition, the K_d value of 50 μM for 5-EyITP is ~3-fold higher than that for 5-NITP.[15]

These results are somewhat surprising, since 5-EyITP does not possess a partial positive charge (as does 5-NITP) that can interact with the templating base to stabilize the interhelical conformation. Thus, it appears that the π-electron density and solvation energies present on 5-EyITP provide the biophysical forces required for stabilizing the nucleobase at the DNA primer-template. To

emphasize this point, we draw upon comparisons between 5-EyITP and 5-CEITP (Fig. 3B), as both analogs have identical π-electron surface areas as well as comparable dipole moments (Table 1). The most notable differences between these two nucleotides are the larger overall size (volume) and hydrophobicity of 5-CEITP compared to 5-EyITP (refer to Fig. 1A). Indeed, 5-EyITP binds 10-fold more weakly than 5-CEITP[22] (compare 50 μM *versus* 5 μM, respectively). This most likely reflects the reduced size and hydrophobicity of 5-EyITP which limits its ability to adequately displace any water molecules that may occupy the void of the abasic site as effectively as the larger 5-CEITP (Fig. 3B). However, the faster k_{pol} value of 94 s^{-1} for 5-EyITP compared to 25 s^{-1} for 5-CEITP[22] likely reflects the ability of the smaller analog to facilitate the conformational change preceding phosphoryl transfer more effectively than the larger 5-CEITP analog.

Based upon these results, we predicted that the related analog, 5-EtITP, would be poorly incorporated opposite an abasic site since it lacks both the size and π-electron density at the 5-position that are required for efficient stacking interactions. Consistent with this argument is the fact that the K_d of 86 μM for 5-EtITP is higher than that of 50 μM measured for 5-EyITP. However, 5-EtITP is inserted opposite the DNA lesion with a fast k_{pol} value of 88.5 s^{-1} that is esssentially identical to that of 94 s^{-1} measured with 5-EyITP. This result suggests that π-electron density plays a minimal role during nucleotide incorporation opposite an abasic site. Perhaps the most intriguing result, however, is the relatively fast k_{pol} value of 60 s^{-1} measured for 5-MeITP, as this analog also lacks π-electron density at the 5-position. In fact, 5-MeITP is incorporated 2-fold faster than 5-CyITP, an analog containing significant π-electron density. The binding affinity for 5-MeITP is identical, within error, to that of 5-CyITP. This last feature is likely related to their similarity in shape and size (compare volumes of 151.0 *versus* 152.3 Å^3, respectively).

Enzymatic insertion opposite templating nucleobases

The unexpected kinetic behavior for these non-natural nucleotides opposite an abasic site prompted us to test their incorporation opposite templating DNA. Time courses were generated using single turnover conditions as described above. Representative time courses for the incorporation of 5-CyITP opposite a templating thymine are provided in Fig. 4A. The plot of k_{obs} *versus* 5-CyITP concentration is hyperbolic (Fig. 4B), and a fit of the data to the Michaelis–Menten equation yields a k_{pol} value of 0.18 ± 0.02 s^{-1}, a K_d value of 127 ± 55 μM, and a k_{pol}/K_d of 1.4 × 10^3 M^{-1} s^{-1}. Identical analyses were performed with 5-EyITP, 5-EtITP, and 5-MeITP (data not shown). Kinetic parameters for the incorporation of these analogs opposite a templating thymine are summarized in Table 2.

Table 2 Summary of kinetic rate and dissociation constants for the incorporation of 5-CyITP, 5-MeITP, 5-EtITP, and 5-EyITP opposite thymine

Nucleotide	k_{pol}/s^{-1}	$K_d/\mu M$	$(k_{pol}/K_d)/M^{-1}\ s^{-1}$
5-CyITP	0.18 ± 0.02	127 ± 55	1.4 × 10^3
5-MeITP	1.8 ± 0.1	107 ± 22	1.7 × 10^4
5-EtITP	0.61 ± 0.08	157 ± 57	3.9 × 10^3
5-EyITP	2.6 ± 0.1	159 ± 20	1.6 × 10^4

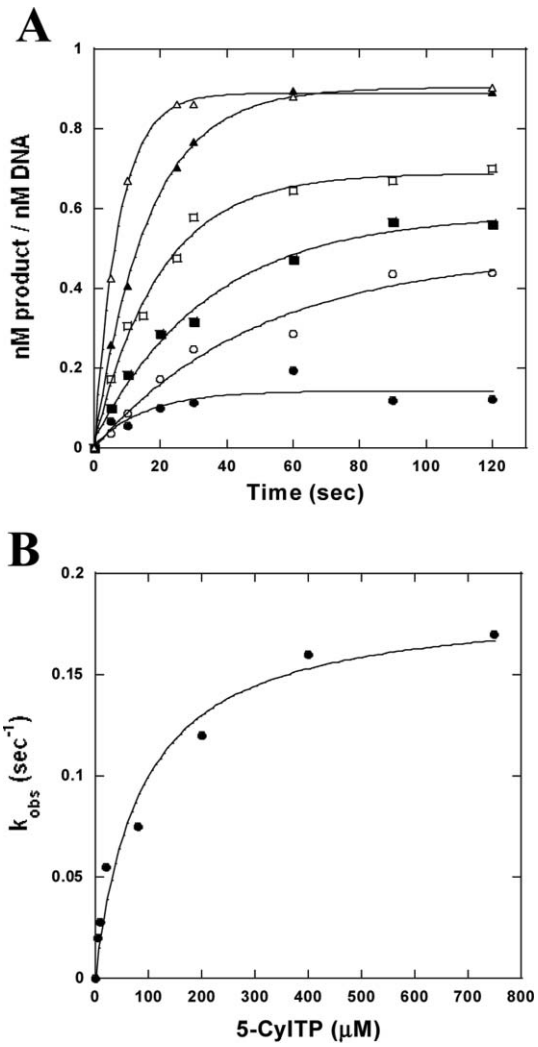


Fig. 4 (A) Rate constants for the incorporation of 5-CyITP opposite a templating thymine as a function of nucleotide concentration. Assays were performed using 1 μM gp43 exo$^-$, 250 nM 13/20$_T$-mer, 10 mM $Mg(OAc)_2$, and 5-CyITP in variable concentrations: 5 μM (●), 10 μM (○), 20 μM (■), 80 μM (□), 400 μM (▲), and 750 μM (△). The solid lines represent the fit of the data to a single exponential. (B) Observed rate constants for incorporation (●) were plotted against 5-CyITP concentration and fitted using the Michaelis–Menten equation to determine values corresponding to K_d and k_{pol}.

In general, the smaller non-natural nucleotides synthesized in this study are poorly incorporated opposite any templating nucleobase. For example, the catalytic efficiency (k_{pol}/K_d) for incorporation opposite a templating thymine is approximately 2 orders of magnitude lower than that for incorporation opposite an abasic site. This lower catalytic efficiency is caused primarily by dramatic reductions in the k_{pol} values (30–60-fold) with minimal effects in binding affinity (~2–3-fold). These results are consistent with our previously proposed model outlined in Fig. 5.[15–19] In this model, the templating base is oriented in an extrahelical position, creating a "void" in the DNA that functionally mimics an abasic site. Since direct hydrogen-bonding interactions between the templating base and the incoming nucleotide are precluded, discrimination is primarily dictated by the conformational change step (k_{pol} effect) rather than initial ground state binding (K_d effect).

This journal is © The Royal Society of Chemistry 2007

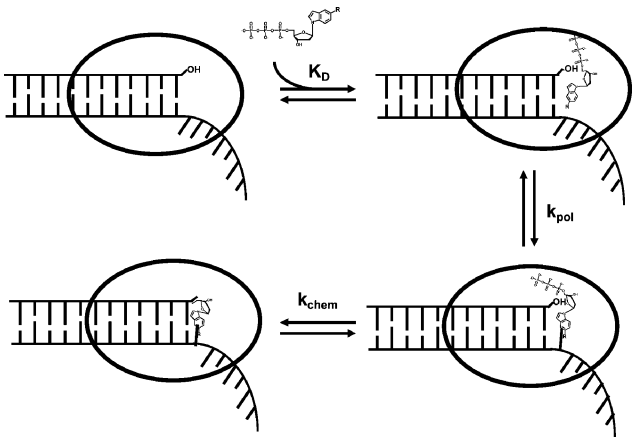


Fig. 5 (A) Proposed model for the enzymatic incorporation of non-natural nucleotides. The first kinetic step represents binding of dXTP to the polymerase–DNA complex (K_d). After nucleotide binding, the polymerase undergoes a conformational change (k_{pol}) that is required to stack the nucleotide into the hydrophobic environment of duplex DNA. The final stage of the catalytic cycle is the phosphoryl transfer step that is required for elongation of the primer strand (k_{chem}).

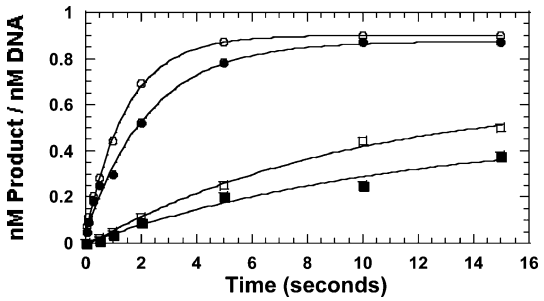


Fig. 6 Representative time courses for to display the dependency of k_{obs} for 1 μM gp43 exo⁻ incorporation of 100 μM 5-MeITP opposite in 250 nM $13/20_X$-mer when the templating base is varied from T (○), A (●), G (□), and C (■). The k_{obs} values were determined by fitting the data with a single exponential equation: k_{obs} opposite T (○) = 0.70 ± 0.01 s^{-1}, k_{obs} opposite A (●) = 0.42 ± 0.01 s^{-1}, k_{obs} opposite G (□) = 0.10 ± 0.04 s^{-1}, k_{obs} opposite C (■) = 0.13 ± 0.07 s^{-1}.

It should be noted that the binding affinities of these smaller non-natural nucleotides are influenced by the presence of a templating nucleobase, since the K_d values measured opposite thymine are ~3-fold higher compared to those measured opposite an abasic site. However, the larger discriminatory factor amongst these molecules is reflected in the k_{pol} step, which we attribute to the conformational change step required to position the templating base into an intrahelical position for proper alignment during the phosphoryl transfer step. Indeed, the kinetic data for 5-CyITP, 5-EyITP, 5-MeITP, and 5-EtITP reveal dramatic reductions in the k_{pol} values (~30–60-fold) when thymine is present in the templating strand.

It is interesting to note that two of the four analogs characterized in this study (5-EyITP and 5-MeITP) are incorporated relatively fast opposite a templating thymine. At face value, the fast k_{pol} values of ~2 s^{-1} are consistent with the shape complementarity model,[23] since these non-natural analogs resemble adenine with respect to shape and size. However, the kinetics of nucleotide incorporation do not always follow the rules set by geometrical constraints. For example, 5-MeITP is incorporated 10-fold faster than 5-CyITP despite obvious similarities in their shape and size. Identical arguments can be made for the 4-fold difference in k_{pol} values between 5-EyITP and 5-EtITP. Collectively, the kinetic data reveal that neither the rate constants for the conformational change step nor their binding affinities increase as the shape and size of the non-natural nucleotide approaches that for the "natural" base-pairing partner.

To ensure that this anomalous behavior is not unique for incorporation opposite thymine, we compared the incorporation of 5-MeITP and 5-CyITP opposite all four templating bases under identical reaction conditions (1 μM gp43 exo⁻, 500 nM $13/20_X$-mer, and 100 μM non-natural nucleotide). Representative data provided in Fig. 6 reveals that 5-MeITP is preferentially incorporated opposite adenine and thymine compared to cytosine and guanine. An identical trend is observed using 5-CyITP as the non-natural nucleotide (data not shown). These results are unusual, since both purine analogs were predicted to be preferentially inserted opposite pyrimidines due to projected similarities in overall shape and size compared to a normal Watson–Crick base pair. One possibility to explain these unusual results is that the non-natural nucleotides used in this study preferentially exist in the *syn* conformation as opposed to the predicted *anti* conformation. Indeed, it has been demonstrated by Hamm and Billig[24] that adenine and guanine preferentially pair opposite 8-oxoguanine and 7-methyl-8-oxoguanine in the *syn* conformation. In both cases, the binding of the modified bases in the *syn* conformation increases the thermal stability of duplex DNA.

Collectively, these mechanistic studies demonstrate that the selectivity for incorporation opposite an abasic site can be modulated through simple alterations to the π-electron surface area of a non-natural nucleotide. This is best illustrated in Fig. 7A, which plots the catalytic efficiency of each 5-substituted indolyl nucleotide as a function of its respective π-electron surface area. The data are best fitted to a parabolic function describing the combined contributions of π-electron density and overall size of the incoming nucleotide toward influencing the optimal catalytic efficiency for nucleotide incorporation in the absence of templating information.[25]

While it is relatively simple to design nucleotides that are selectively incorporated opposite a non-templating DNA lesion, it appears far more challenging to rationally design a nucleotide for selective incorporation opposite a templating nucleobase. As illustrated in Fig. 7B, a defined parabolic trend is non-existent when examining the catalytic efficiency for the incorporation of non-natural nucleotides opposite a templating base. It was previously demonstrated that "large" non-natural nucleotides such as 5-NapITP and 5-AnITP show anomalous kinetic behavior, as they are incorporated with unexpectedly high catalytic efficiencies despite being poor pairing partners for thymine, based on shape/size constraints.[17] In this report, we demonstrate that "smaller" indolyl analogs such as 5-MeITP and 5-EtITP are incorporated poorly opposite templating DNA although they are relatively good shape mimetics of natural purines. All four indolyl analogs are poorly incorporated opposite templating nucleobases, as the overall catalytic efficiencies are ~100-fold lower compared to that for translesion DNA synthesis. The decreased efficiency derives mainly from reductions in k_{pol} rather than through

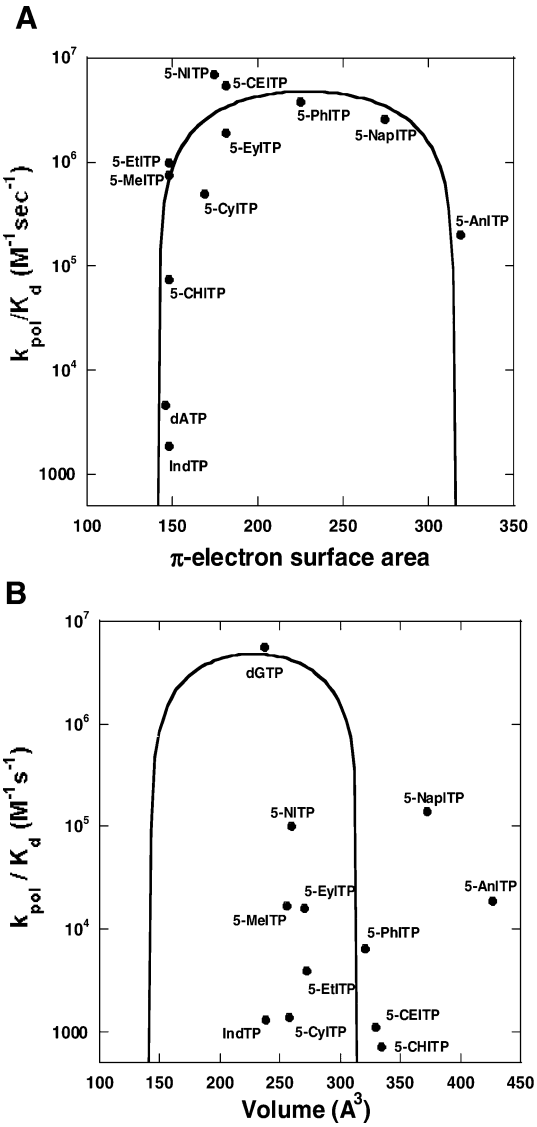


Fig. 7 (A) Dependency of the catalytic efficiency for incorporation opposite an abasic site as a function of π-electron surface area. The solid line represents the fit of the data to a parabolic equation ($R = 0.70$). (B) Dependency of catalytic efficiency for incorporation opposite a templating cytidine as a function of total volume.

alterations in K_d values. Their poor utilization is somewhat surprising, since the overall shape and size of the analogs are very similar to natural purine analogs such as adenine and guanine. Even more intriguing is the fact that none of these modified indolyl nucleotides are preferentially incorporated opposite their predicted "size/shape" partners, *i.e.*, cytosine and thymine. For example, the total surface area of 5-MeITP paired opposite thymine and cytosine is 286.8 Å^2 and 295.0 Å^2, respectively. Both values are close to the total surface area of 283.8 Å^2 for a natural G:C base pair. In addition, the facile incorporation of 5-MeITP and 5-CyITP opposite adenine is counterintuitive, since the predicted surface area of the 5-MeITP:adenine pair of 312.6 Å^2 is significantly larger than that of ~285 Å^2 for a natural base pair. Thus, the kinetic data do not support a model based solely on shape complementarity. However, we hasten to note that other models invoking hydrogen-bonding interactions, hydrophobicity, and π-electron density also cannot completely explain the kinetic behavior of these non-natural nucleotides incorporated opposite templating nucleobases. This deficiency likely reflects the fact that nucleotide incorporation opposite templating DNA is influenced by these interrelated biophysical features. Thus, altering a functional group of an incoming nucleotide undoubtedly influences each of these biophysical features to varying degrees and makes it difficult to unambiguously assess which features are most important. In contrast, the non-instructional nature of an abasic site appears to reduce the overall complexity of these interactions such that the catalytic efficiency for incorporation is primarily influenced by only the size and π-electron density of the incoming nucleotide. These differences suggest that accurately comparing the mechanisms of correct *versus* translesion DNA synthesis is impossible, since the rules governing nucleotide incorporation differ as a result of these different biophysical parameters.

Experimental

Synthesis of non-natural nucleosides and nucleotides

Materials. Tributyl ammonium pyrophosphate was purchased from Sigma. 1-Chloro-1,2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-erythropentofuranose, ethyl acetate, hexane, methanol, dichloromethane, phosphoryl oxychloride, dimethyl formamide, and tributylamine were purchased from ACROS. Trimethyl phosphate and tributylamine were dried over 4 Å molecular sieves. DMF was distilled over ninhydrin and stored in 4 Å molecular sieves. All NMR spectra were recorded in a Gemini-300 FT NMR spectrometer or a Varian 400 MHz NMR spectrometer. Proton chemical shifts are reported in ppm, downfield from tetramethylsilane. Coupling constants (*J*) are reported in hertz (Hz). ^{31}P NMR spectra were taken in D_2O in the presence of 50 mM Tris (pH 7.5) and 2 mM EDTA, and 85% phosphoric acid was used as the external standard. Ultraviolet quantification of triphosphate was performed on a Cary 50 spectrophotometer. High-resolution electrospray mass spectrometery analyses were performed on an Ionspec HiRES ESI-FTICRMS at the University of Cincinnati. 5-Cyanoindole and 5-methylindole were purchased from Acros. 5-Ethylindole was purchased from Wako Chemicals.

The general procedure for preparing the 5-substituted indole triphosphates utilized in this study has been published previously.[17,19,22] In brief, the 5-substituted indole (5 mmol) was reacted with sodium hydride (6 mmol) in 120 mL of anhydrous acetonitrile at room temperature for 30 min. 1-Chloro-1,2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-erythropentofuranose (6 mmol) was then added, and the resulting mixture was stirred at room temperature overnight. Upon evaporation to dryness, the crude reaction mixture was purified by flash column chromatography on silica using dichloromethane and methanol as the solvents (95 : 5). The purified nucleosides were characterized by ^{1}H NMR and mass spectrometry.

To prepare the triphosphates, phosphoryl oxychloride (3 mol equiv, 0.18 mmol) was added dropwise to a pre-chilled (0 °C) solution of the 5-substituted indolyl nucleoside (0.06 mmol) and triethylamine (5 mol equiv., 0.33 mmol) in 0.37 mL of trimethylphosphate. The reaction mixture was stirred for 1.5 h and then treated with a 0.5 M solution of tributylammonium pyrophosphate (0.75 mmol) in DMF and tributylamine (0.75 mmol).

 This journal is © The Royal Society of Chemistry 2007

After stirring at room temperature for 20 min, the reaction mixture was neutralized with 15 mL of 1 M TEAB, followed by an additional 2 h of stirring at room temperature. The crude reaction mixture was evaporated by rota-evaporation under reduced pressure, and then purified by preparative reverse HPLC using a linear gradient of 50% to 80%B within 14 min at a flow rate of 17 mL min^{-1} (mobile phase A: 0.1 M TEAB; B: 50% ACN in 0.1 M TEAB). The desirable nucleotides were characterized by mass spectrometry and ^{31}P NMR.

Synthesis of 5-cyanoindole 2′-deoxyribofuranoside 5′-triphosphate (5-CyITP). 5-Cyanoindole (5-Cy-Ind) was reacted with 1-chloro-2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-erythropentofuranose to yield the nucleoside. ^{1}H-NMR (DMSO, 400 MHz) δ: 2.15–2.21 (m, 1H, 2′-H), 2.37–2.39 (m, 1H, 2′-H), 3.40–3.54 (m, 2H, 5′-H), 3.80 (m, 1H, 4′-H), 4.33 (m, 1H, 3′-H), 4.90 (m, 1H, 5′-OH), 5.30 (br s, 1H, 3′-OH), 6.39 (t, 1H, J = 6.4 Hz, 1′-H), 6.60 (d, 1H, J = 3.2 Hz, 3-H), 7.40 (dd, 1H, J = 1.6 Hz, 8.64 Hz, Ar), 7.71–7.73 (m, 2H, Ar), 8.01–8.02 (m, 1H, Ar). ESI-MS (+), calculated mass spectrum formula $C_{14}H_{14}N_2NaO_3$ for M + Na): 281.0902; experimental mass spectrum: 281.0894. UV (MeOH) λ_{235} (nm), ε = 27 933 cm^{-1} M^{-1}.

The purified 5-cyanoindole 2′-deoxyribofuranoside was then converted to 5-cyanoindole 2′-deoxyribofuranoside 5′-triphosphate (5-CyITP) using the general procedure for synthesizing the triphosphate as described above. The yield of this reaction was 25%. 5-Cyanoindole 2′-deoxyribofuranoside 5′-triphosphate (5-CyITP): ^{31}P NMR (D_2O) δ: −6.5 (γ-P), −10.5 (α-P), −21.9 (β-P). HiRes ESI-MS (−): calculated mass spectrum formula $C_{14}H_{16}N_2O_{12}P_3$ for M − H: 496.9916; experimental mass spectrum: 496.9905.

Synthesis of 5-ethyleneindole 2′-deoxyribofuranoside 5′-triphosphate (5-EyITP). 5-Ethyleneindole was synthesized from vinyl bromide and indole boronic acid *via* the Suzuki coupling method.[17] 5-Ethyleneindole: ^{1}H NMR ($CDCl_3$, 300 MHz) δ: 5.2 (d, J = 11 Hz, 1H, CH), 5.7 (d, J = 18 Hz, 1H, CH), 6.6 (s, 1H, 3-H), 6.85 (m, 1H, CH), 7.2 (m, 2H, Ar), 7.4 (m, 1H, Ar), 7.65 (m, 1H, Ar), 8.17 (br s, 1H, NH).

The purified 5-ethyleneindole was first converted to 5-ethyleneindole 2′-deoxyribofuranoside and then to 5-ethyleneindole 2′-deoxyribofuranoside 5′-triphosphate (5-EyITP) using the general method outlined above. The yield of 5-EyITP was 44%. 5-Ethyleneindole 2′-deoxyribofuranoside: ^{1}H NMR (DMSO, 400 MHz) δ: 2.20–2.40 (m, 1H, 2′-H), 2.25–2.38 (m, 1H, 2″-H), 3.34–3.45 (m, 2H, 5′-H), 3.72–3.73 (m, 1H, 4′-H), 4.24–4.26 (m, 1H, 3′-H), 4.77–4.82 (m, 1H, 5′-OH), 5.04 (dd, 1H, J = 0.9 Hz, 11 Hz, CH), 5.20 (d, 1H, J = 4.3 Hz, 3′-OH), 5.7 (dd, 1H, J = 1.17 Hz, 17.76 Hz, CH), 6.28 (t, J = 6.93 Hz, 1H, 1′-H), 6.40 (d, 1H, J = 3.51 Hz, 3-H), 6.66–6.74 (m, 1H, CH), 7.24 (dd, 1H, J = 1.66 Hz, 8.68 Hz, Ar), 7.44–7.52 (m, 3H, Ar). ESI-MS (+): calculated mass spectrum formula $C_{15}H_{18}NO_3$ for M + H: 260.13; experimental mass spectrum: 260.08. UV (MeOH) λ_{246} (nm): ε = 19325 cm^{-1} M^{-1}. 5-Ethyleneindole 2′-deoxyribofuranoside 5′-triphosphate (5-EyITP): ^{31}P NMR (D_2O, 300 MHz) δ −6 (γ-P), −11 (α-P), −22.3 (β-P). HiRes ESI-MS (−): calculated mass spectrum formula $C_{15}H_{19}NO_{12}P_3$ for M − H: 498.0126; experimental mass spectrum: 498.0112.

Synthesis of 5-ethylindole 2′-deoxyribofuranoside 5′-triphosphate (5-EtITP). 5-Ethylindole was used as the starting material for synthesizing the 5-ethylindole 2′-deoxyribofuranoside using the general procedure outlined above. ^{1}H NMR (DMSO, 400 MHz) δ: 1.18 (t, J = 7.61 Hz, 3H, CH_3), 2.06–2.14 (m, 1H, 2′-H), 2.34–2.48 (m, 1H, 2′-H); 2.62 (q, 2H, J = 7.61 Hz, CH_2), 3.40–3.56 (m, 2H, 5′-H), 3.77–3.78 (m, 1H, 4′-H), 4.27–4.33 (m, 1H, 3′-H), 4.89 (t, J = 5.37 Hz, 1H, 5′-OH), 5.20 (d, J = 4.1 Hz, 1H, 3′-OH), 6.29–6.31 (m, 1H, 1′-H), 6.40 (d, J = 3.52 Hz, 1H, 3-H), 6.97 (dd, J = 1.56 Hz, 8.39 Hz, 1H, Ar), 7.30–7.32 (m, 1H, Ar), 7.44 (d, J = 8.4 Hz, 1H, Ar), 7.49 (d, J = 3.2 Hz, 1H, Ar). HiRes FAB-MS (+): calculated mass spectrum formula $C_{15}H_{20}NO_3$ for M + H: 262.14432; experimental mass spectrum: 262.14270. UV (MeOH) λ_{266} (nm): ε = 4774 cm^{-1} M^{-1}.

The purified 5-ethylindole 2′-deoxyribofuranoside was then converted to 5-ethylindole 2′-deoxyribofuranoside 5′-triphosphate (5-EtITP) using the general procedure described above. The yield of this reaction was 10%. 5-Ethylindole 2′-deoxyribofuranoside 5′-triphosphate (5-EtITP): ^{31}P NMR (D_2O/Tris) δ −8.2 (γ-P), −10.9 (α-P), −22.9 (β-P). HiRes ESI-MS (−): calculated mass spectrum formula $C_{15}H_{21}NO_{12}P_3$ for M − H: 500.0277; experimental mass spectrum: 500.0259.

Synthesis of 5-methylindole 2′-deoxyribofuranoside 5′-triphosphate (5-MeITP). 5-Methylindole was used as the starting material for synthesizing the 5-methylindole 2′-deoxyribofuranoside using the general procedure outlined above. ^{1}H NMR (DMSO, 400 MHz) δ: 2.07–2.123 (m, 1H, 2′-H), 2.293–2.396 (m, 1H, 2′-H), 2.27 (s, 3H, CH_3), 3.35–3.47 (m, 2H, 5′-H), 3.69–3.72 (m, 1H, 4′-H), 4.22–4.26 (m, 1H, 3′-H), 4.78 (t, J = 5.47 Hz, 1H, 5′-OH), 5.185 (d, J = 4.29 Hz, 1H, 3′-OH), 6.22–6.26 (m, 1H, 1′-H), 6.31 (d, J = 3.20 Hz, 1H, Ar), 6.88 (dd, J = 1.56 Hz, 8.39 Hz, 1H, Ar), 7.23 (s, 1H, Ar), 7.35 (d, J = 8.40 Hz, 1H, Ar), 7.43 (d, J = 3.20 Hz, 1H, Ar). HiRes ESI-MS (+): calculated mass spectrum formula $C_{14}H_{17}NNaO_3$ for M + H: 270.1106; experimental mass spectrum: 270.1104. UV (MeOH) λ_{269} (nm): 2563 cm^{-1} M^{-1}. The purified 5-methylindole 2′-deoxyribofuranoside was then converted to 5-methylindole 2′-deoxyribofuranoside 5′-triphosphate (5-MeITP) using the general procedure described above. The yield of this reaction was 33%. 5-Methylindole 2′-deoxyribofuranoside 5′-triphosphate (5-MeITP): ^{31}P NMR (D_2O) δ −6 (γ-P), −10.4 (α-P),−21.7 (β-P). HiRes ESI-MS (−): calculated mass spectrum formula $C_{14}H_{19}NO_{12}P_3$ for M − H: 486.0120; experimental mass spectrum: 486.0272.

Nucleotide incorporation analyses

Materials. [γ-^{32}P]ATP was purchased from Perkin Elmer Life and Analytical Sciences (Boston, MA). Unlabeled dNTPs (ultrapure) were obtained from Pharmacia. Magnesium acetate and Trizma base were from Sigma. Urea, acrylamide, and bisacrylamide were from National Diagnostics (Rochester, NY). Oligonucleotides, including those containing a tetrahydrofuran moiety mimicking an abasic site, were synthesized by Operon Technologies (Alameda, CA). Single-stranded and duplex DNA were purified and quantified as described.[26] All other materials were obtained from commercial sources and were of the highest quality available. The exonuclease-deficient mutant of gp43 (Asp-219 to Ala mutation) was purified and quantified as previously described.[27,28]

Enzyme assay methods. The assay buffer used in all kinetic studies consisted of 25 mM Tris-OAc (pH 7.5), 150 mM KOAc, and 10 mM 2-mercaptoethanol. All assays were performed at 25 °C. Polymerization reactions were monitored by analysis of the products on 20% sequencing gels as previously described.[10] Gel images were obtained with a Packard PhosphorImager using the OptiQuant software supplied by the manufacturer. Product formation was quantified by measuring the ratio of ^{32}P-labeled extended and non-extended primer. The ratios of product formation are corrected for the substrate in the absence of polymerase (zero point). Corrected ratios are then multiplied by the concentration of primer/template used in each assay to yield the total product. All concentrations are listed as final solution concentrations.

Pre-steady-state nucleotide incorporation assays. Kinetic parameters k_{pol}, K_d, and k_{pol}/K_d for each non-natural nucleotide were determined by monitoring the rate constants in product formation using a fixed amount of gp43 (1 μM) and DNA substrate (500 nM) at varying concentrations of nucleoside triphosphate (0.01–1 mM). Aliquots of the reactions performed by manual quenching techniques were quenched in 350 mM EDTA, pH 8, at times ranging from 5–300 s.

A rapid quench instrument (KinTek Corporation, Clarence, PA) was also used to monitor the time course in non-natural nucleotide incorporation. Experiments were performed using single turnover reaction conditions in which 250 nM DNA was incubated with 10 mM $Mg(OAc)_2$ and variable concentrations of the nucleotide analog (5–500 μM). This solution was mixed with 500 nM gp43 exo$^-$ and 10 mM $Mg(OAc)_2$. The reactions were quenched with 350 mM EDTA, pH 8, at variable times (0.005–5 s).

In all cases, samples were diluted 1 : 1 with sequencing gel load buffer and products were analyzed for product formation by denaturing gel electrophoresis. Data obtained for the pre-steady-state rates in DNA polymerization measured under single turnover reaction conditions were fitted with eqn (1):

$$y = A(1 - e^{-kt}) + C, \quad (1)$$

where A is the burst amplitude, k is the observed rate constant (k_{obs}) in product formation, t is time, and C is a defined constant. Data for the dependency of k_{obs} as a function of dNTP concentration was fitted with the Michaelis–Menten equation (eqn (2)) to provide values corresponding to k_{pol} and K_d:

$$k_{obs} = k_{pol}[dXTP]/(K_d + [dXTP]), \quad (2)$$

where k_{obs} is the observed rate constant of the reaction, k_{pol} is the maximal polymerization rate constant, K_d is the dissociation constant for dXTP, and [dXT]P is the concentration of non-natural nucleotide substrate.

Acknowledgements

This research was supported through funding from the National Institutes of Health (CA118408) to AJB.

References and notes

1 W. A. Beard and S. H. Wilson, *Structure*, 2003, **11**, 489–496.
2 C. M. Joyce and S. J. Benkovic, *Biochemistry*, 2004, **43**, 14317–14324.
3 R. A. Perlow-Poehnelt, I. Likhterov, D. A. Scicchitano, N. E. Geacintov and S. Broyde, *J. Biol. Chem.*, 2004, **279**, 36951–36961.
4 M. A. Graziewicz, J. M. Sayer, D. M. Jerina and W. C. Copeland, *Nucleic Acids Res.*, 2004, **32**, 397–405.
5 R. L. Eoff, K. C. Angel, M. Egli and F. P. Guengerich, *J. Biol. Chem.*, 2007, **282**, 13573–13584.
6 T. Lindahl, *Nature*, 1993, **362**, 709–715.
7 K. S. Gates, T. Nooner and S. Dutta, *Chem. Res. Toxicol.*, 2004, **17**, 839–856.
8 M. L. Dodson and R. S. Lloyd, *Free Radical Biol. Med.*, 2002, **32**, 678–682.
9 J. S. Sung and B. Demple, *FEBS J.*, 2006, **273**, 1620–1629.
10 A. J. Berdis, *Biochemistry*, 2001, **40**, 7180–7191.
11 B. S. Strauss, *Bioessays*, 1991, **13**, 79–84.
12 J. S. Taylor, *Mutat. Res.*, 2002, **510**, 55–70.
13 S. Boiteux and M. Guillet, *DNA Repair*, 2004, **3**, 1–12.
14 X. Zhang, I. Lee and A. J. Berdis, *Biochemistry*, 2005, **44**, 13111–13121.
15 E. Z. Reineks and A. J. Berdis, *Biochemistry*, 2004, **43**, 393–404.
16 X. Zhang, I. Lee and A. J. Berdis, *Biochemistry*, 2005, **44**, 13101–13110.
17 X. Zhang, A. Donnelly, I. Lee and A. J. Berdis, *Biochemistry*, 2006, **45**, 13293–13303.
18 B. Devadoss, I. Lee and A. J. Berdis, *Biochemistry*, 2007, **46**, 4486–4498.
19 X. Zhang, I. Lee and A. J. Berdis, *Org. Biomol. Chem.*, 2004, **2**, 1703–1711.
20 N. S. Girgis, H. B. Cottam, S. B. Larsons and R. K. Robins, *Nucleic Acids Res.*, 1987, **15**, 1217–1226.
21 C. L. Smith, A. C. Simmonds, I. R. Felix, A. L. Hamilton, S. Kumar, S. Nampali, D. Loakes and D. M. Brown, *Nucleosides Nucleotides*, 1998, **17**, 541–554.
22 X. Zhang, I. Lee, X. Zhou and A. J. Berdis, *J. Am. Chem. Soc.*, 2006, **128**, 143–149.
23 E. T. Kool, *Annu. Rev. Biochem.*, 2002, **71**, 191–219.
24 Although the size of the nucleotide is important, it cannot entirely account for facile incorporation. This is best exemplified by the large difference in k_{pol}/K_d values for 5-PhITP and 5-CH-ITP (compare 3.8×10^6 M^{-1} s^{-1} *versus* 7.4×10^4 M^{-1} s^{-1}, respectively) despite similarities in their overall surface area (compare 223.2 Å^2 *versus* 240.9 Å^2, respectively). In addition, defined corollaries are not evident between the catalytic efficiency in incorporation opposite the abasic site with the overall size, solvation energy, or dipole moment of the incoming nucleotide.
25 M. L. Hamm and K. Billig, *Org. Biomol. Chem.*, 2006, **4**, 4068–4070.
26 T. L. Capson, J. A. Peliska, B. F. Kaboord, M. W. Frey, C. Lively, M. Dahlberg and S. J. Benkovic, *Biochemistry*, 1992, **31**, 10984–10994.
27 M. W. Frey, N. G. Nossal, T. L. Capson and S. J. Benkovic, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 2579–2583.
28 J. Rush and W. H. Konigsberg, *Prep. Biochem.*, 1989, **19**, 329–340.

This journal is © The Royal Society of Chemistry 2007