Incorporation of reporter-labeled nucleotides by DNA polymerases

Jon P. Anderson[1], Bernhard Angerer[2], and Lawrence A. Loeb[1]

BioTechniques 38:257-264 (February 2005)

The incorporation of fluorescently labeled nucleotides into DNA by DNA polymerases has been used extensively for tagging genes and for labeling DNA. However, we lack studies comparing polymerase efficiencies for incorporating different fluorescently labeled nucleotides. We analyzed the incorporation of fluorescent deoxynucleoside triphosphates by 10 different DNA polymerases, representing a cross-section of DNA polymerases from families A, B, and reverse transcriptase. The substitution of one or more different reporter-labeled nucleotides for the cognate nucleotides was initially investigated by using an in vitro polymerase extension filter-binding assay with natural DNA as a template. Further analysis on longer DNA fragments containing one or more nucleotide analogs was performed using a newly developed extension cut assay. The results indicate that incorporation of fluorescent nucleotides is dependent on the DNA polymerase, fluorophore, linker between the nucleotide and the fluorophore, and position for attachment of the linker and the cognate nucleotide. Of the polymerases tested, Taq *and Vent exo- DNA polymerases were most efficient at incorporating a variety of fluorescently labeled nucleotides. This study suggests that it should be feasible to copy DNA with reactions mixtures that contain all four fluorescently labeled nucleotides allowing for high-density labeling of DNA.*

INTRODUCTION

Besides DNA replication and maintaining the integrity of genomes, DNA polymerases play an essential role in modern molecular biotechnology. Processes such as PCR (1,2), DNA sequencing (3,4), mutation analysis (5,6), and labeling of specific DNA or RNA probes have all depended on the incorporation of nucleoside triphosphates or labeled analogs. These applications have historically relied on radioactive nucleotide substrates, however, there has been increasing emphasis on techniques that use either fluorescent or affinity-based groups such as Biotin or Digoxigenin coupled with enzymatic signal amplification (7). The ability of DNA polymerases to incorporate these modified and labeled nucleotides has been important in monitoring gene expression (8), in situ hybridization (9), and four-color DNA sequencing (10). The methods for introducing nonradioactive detectable groups comprise both enzymatic and chemical additions of 5′ and 3′ markers, as well as introducing internal markers by enzymatic incorporation or post-labeling techniques (11–14). In the case of enzymatic labeling of nucleic acids, the reporter groups are attached to the bases of the deoxynucleoside triphosphates (15,16). Most labeling protocols, such as nick translation and in vitro transcription use a modified nucleotide in combinations with the naturally occurring nucleotides (14,17,18).

In many studies, it is highly desirable to synthesize probes with 100% of one or more of the nucleotide analogs at each position. This facilitates the generation of probes with a high density of fluorophores providing an enhancement in signal in hybridization studies. In some applications, like the single molecule sequencing (19–21), a completely labeled replica of a DNA sequence is obligatory.

In order to fully replace one or more nucleotides with labeled nucleotide analogs, it is desirable to determine which DNA polymerases are most efficacious at incorporating nucleotides labeled with these bulky, hydrophobic sidechains. We investigated 10 different DNA polymerases for their ability to incorporate a variety of different nucleotide analogs. These polymerases represented enzymes from family A, family B, and family reverse transcriptase (RT). In analyzing these polymerases, we developed a new assay to quantitate the production of long fragments of DNA with one or more nucleotides fully replaced with nucleotide analogs. We show that the ability to incorporate nucleotide analogs is not only dependent on the polymerase, but also a function of the nucleotide analog and the position of the linker.

MATERIALS AND METHODS

Polymerases

Commercially available polymerases were used in this study included: *Taq*, Vent, Vent exo-, Deep Vent, and Deep Vent exo- from New England Biolabs (Beverly, MA, USA), *Pfu* from Promega (Madison, WI, USA), and alfalfa mosaic virus reverse trancriptase (AMV-RT), murine leukemia virus reverse transcriptase (MLV-RT), T4, and Pol I-Klenow from Roche Applied Science (Indianapolis, IN, USA). Commercial polymerases were assayed with the buffers provided.

[1]University of Washington, Seattle, WA, USA and [2]Gnothis SA, Lausanne, Switzerland

Modified Nucleotides

Modified nucleotide derivatives used in the analyses included: biotin-16-dUTP, 7-amino-4-methylcoumarin-3-acetic acid (AMCA)-6-dUTP, fluorescein-11-dUTP, tetramethylrhodamine (TMR)-6-dUTP, digoxigenin -11-dUTP, fluorescein-dATP, and IR-770-dATP obtained from Roche Applied Science; Rhodamine Green™-X-dUTP was a gift of K. Muehlegger (Roche Diagnostics GmbH, Mannheim, Germany); and Cy™5-dCTP was obtained from Amersham Biosciences (Piscataway, NJ, USA). (See supplementary material on the *BioTechniques*' web site at http://www.BioTechniques/Feb05/AndersonSupplementary.html.)

Incorporation of Fluorescently Labeled Nucleotide Analogs

In routine screening assays we used activated DNA as a template and completely replaced one or more of the nucleotide substrates with a fluorescently labeled analog. Polymerase reactions were carried out for 30 min in 20-µL reaction mixtures containing 1 µg activated calf thymus DNA, 25 µM each dATP, dCTP, and either dTTP or a reporter-labeled dUTP, 5 µM dGTP, 0.25 µCi [α-^{32}P]dGTP (3000 Ci/mmol; PerkinElmer Life and Analytical Sciences, Boston, MA, USA), *Taq* reaction buffer, 5 mM $MgSO_4$, and 0.1–1 U polymerase. Incorporation of modified dNTPs was indirectly traced by quantitating the incorporation of radioactively labeled dGTP. Reactions were incubated for 30 min at 72°C for thermophile polymerases or 37°C for mesophile polymerases and terminated by the addition of 100 µL 0.1 M sodium pyrophosphate, 0.05 M EDTA, and 10 µL calf thymus carrier DNA (5 mg/mL). Natural dNTPs were used as positive controls and served to reference 100% activity for each polymerase tested. Background activity was measured in reactions where dTTP was omitted.

Primer extension reactions were used to determine the length of the newly synthesized DNA product. The template, a 2671-bp fragment derived from the plasmid pACYC184 (New England Biolabs), was generated by PCR using the Expand™ High Fidelity PCR system (Roche Applied Science, according to the manufacturer's instructions). The PCR primers for this fragment were 5′-TGCGCCTTATCCGGTAACTATCGTCTTGA-3′ and 5′-GTGCCCTTAAACGCCTGGTGCTACGCCTG-3′, with 20 pmol of each primer used per 50-µL reaction. Cycling parameters were 1 cycle at 95°C for 2 min, 25 cycles of 95°C for 15 s, 58°C for 30 s, and 72°C for 2 min, then 1 cycle at 72°C for 10 min. After cycling, the reactions were kept at 4°C. DNA fragments were either purified on a QIAquick® Nucleotide Removal kit (Qiagen, Valencia, CA, USA) or on a 50-cm column containing Sephacryl™ S-200 (Amersham Biosciences) to remove dNTPs. The labeling procedure was carried out by briefly heating the 1.2–1.5 pmol DNA for 5 min at 95°C in the presence of excess (20 pmol) of one of the primers. After heating and annealing of the primer at 60°C for 2 min, extension reactions were carried out for 10–90 min at 72°C. The conditions were the same as that used in assays with activated DNA, except that one of the dNTPs was replaced with the corresponding nucleotide fluorophore.

Detection Methods

Incorporation of nucleotide fluorophores. Polymerase activity was determined as previously described (22). Briefly, following the termination of the labeling reaction, 30 µL (about 1/3 of the total reaction) was mixed with 250 µL precipitation solution (1 M HCL, 0.1 M sodium pyrophosphate) in a 96-well silent screen filter plate (Biodyne™ B; Nalge Nunc International, Rochester, NY, USA) mounted on a 96-vacuum manifold (Beckman Coulter, Fullerton, CA, USA). The plates were washed three times with 250 µL precipitation solution and once with 250 µL 95% ethanol to remove the unincorporated radioactivity. The filter was pulled from the plate and dried, and the amount of radioactivity associated with the filter was quantified by Molecular Dynamics Storm™ 840 PhosphorImager™ analysis (Amersham Biosciences). Results from this method were validated by comparing the results to those obtained by the traditional acid precipitation filter assay using liquid scintillation counting (23). The data obtained from the two protocols yielded nearly identical results (data not shown).

All reactions were carried out as duplicates. Two polymerase concentrations for each enzyme (0.1 and 1 U) and two concentrations of the modified nucleotide (25 and 50 µM) were used. The following controls for each enzyme were carried out to calibrate the system: (*i*) positive controls were carried out using all four natural nucleotides; (*ii*) nonspecific binding controls were carried out in the absence of a DNA polymerase; and (*iii*) misincorporation controls used reactions that contained only three of the four natural nucleotides.

Extension cut assay. A new protocol was designed to evaluate the elongation of the DNA strand having one or more nucleotides fully replaced by the labeled nucleotide analogs. The method utilizes a single-stranded DNA template bound to a solid support through a streptavidin-biotin conjugation. A multistep process produced the bound single-stranded template. First, a 1203-bp double-stranded fragment derived from the plasmid pET15b (Novagen, Madison, WI, USA) was produced by PCR. The primers for this fragment were 5′-TTCTCCCATGAAGACGGTACGCGACTG-3′ and 5′-biotin-ACAATCCATGCCAACCCGTTCCATGTGC-3′, with 20 pmol of each primer used per reaction. Cycling parameters were 1 cycle at 95°C for 2 min, 25 cycles of 95°C for 15 s, 58°C for 30 s, and 72°C for 2 min, and then 1 cycle at 72°C for 10 min. The biotin-labeled fragments were separated from the free primers with a PCR purification kit (Qiagen). A total of 2 µg biotin-labeled DNA was diluted to a final volume of 50 µL in a binding buffer containing 1 M NaCl, 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA final concentrations. The DNA solution was added to streptavidin-coated PCR tubes (Roche Applied Science) and allowed to incubate for 20 min at 25°C. The tubes were washed three times with binding buffer and once with distilled water. The double-stranded DNA bound to the streptavidin tubes was denatured by adding 50 µL melting

solution (125 mM NaOH, 0.1 M NaCl) and incubating the tubes for 20 min at 25°C. The tubes were washed three times with binding buffer and once with distilled water, producing single-stranded DNA bound to the PCR tubes. Labeling reactions using the primer 5′-CATGAAGACGGTACGCGACT-GGGCGTGG-3′ were carried out on the single-stranded DNA by allowing for the annealing of the primer (at 58°C for 1 min), followed by an extension reaction for 30 min at 70°C. Reactions were carried out as in the filter binding assays with one or more natural dNTPs replaced with reporter-labeled nucleotide analogs and one of the other nucleotide labeled with α^{32}P.

Following the extension reactions, the tubes were again washed three times with binding buffer and once with distilled water to remove any unincorporated nucleotides. The newly produced double-stranded DNA was then cleaved with restriction enzymes to release a portion of the DNA into solution. Restriction enzymes were chosen that cleaved at a series of GC sites to reduce the interference by modified nucleotides. The restriction reactions were carried out in the streptavidin-coated PCR tubes in a total volume of 50 µL according to the manufacturer's instructions. The released DNA from the restriction reactions was then either analyzed on the filter plates or displayed on a 6% nondenaturing acrylamide gel.

RESULTS

Initial Screening of the DNA Polymerases

To investigate the ability of DNA polymerases to incorporate reporter-labeled nucleotide analogs, we screened 10 different polymerases against a panel of nucleotide analogs. These analogs included dUTP-labeled biotin, AMCA, Rhodamine Green, fluorescein, and TMR. In each test, we compared the ability of the polymerase to incorporate the natural nucleotide to that of each analog. Two concentrations of polymerase (0.1 and 1 U) and two concentrations of nucleotide fluorophore (25 and 50 µM) were assayed, with each comparison repeated two times, for a total of eight comparisons for each dye-polymerase combination. An activated calf thymus DNA primer template was used for the initial screening of the polymerases, and the assays were performed in a 96-well format (22).

The results of the polymerase screen showed that most of the polymerases tested that contained a 3′ to 5′ exonuclease activity (proofreading ability) were inefficient at incorporating the nucleotide analogs. These polymerases included Vent, Deep Vent, T4, and *Pfu*. This result is consistent with a previous study on incorporating 2′-fluoro-modified nucleic acids using enzymes with or without 3′ to 5′ exonuclease activity (24). Therefore, we focused on the polymerases that lacked a 3′ to 5′ exonuclease activity and Klenow, which exhibits a less active 3′ to 5′ exonuclease and did incur a similar reduction of analog incorporation as seen in the other proofreading polymerases. We analyzed two polymerases from family A (*Taq* and Klenow), family B (Vent exo$^-$ and Deep Vent exo$^-$), and family RT (MLV and AMV). For each of these polymerases, the relative ability to incorporate the nucleotide analogs as compared to natural nucleotides was determined (Figure 1). The results show that no single polymerase was more efficient at incorporating all of the different analogs. *Taq* polymerase from family A and Vent exo$^-$ polymerase from family B were further analyzed.

Multiple Nucleotide Replacements

To further probe incorporation by *Taq* and Vent exo$^-$, we replaced two of the natural nucleotides with nucleotide analogs and utilized activated calf thymus DNA as a template primer. We tested each of the dU derivatives used in previous experiments alone and in combination with Cy5 dCTP, fluorescein dATP, and IR-770-dATP (Figure 2). As expected, the results show that complete replacement of a natural nucleotide with a single nucleoside analog results in more extensive synthesis than replacement of two natural nucleotides with two nucleoside analogs. These experiments show that *Taq* and Vent exo$^-$ inefficiently incorporate the dA and dC derivatives. When these derivatives are used in combination with the dU derivatives, the amount of labeled product was further reduced. Furthermore, the incorporation of two nucleoside analogs often reduces the incorporation efficiency to a point that it is below the background signal of incorporating only three of the four natural nucleotides.

Primer Extension Reactions

Although activated calf thymus DNA was used effectively as a template for our initial polymerase screening, the randomly nicked template can incorporated single nucleotide additions without further extension and thus produces a high background signal. To minimize terminal addition reactions, we performed reactions on a more defined primer-template system. On a 2.6-kb linear fragment, we tested the ability of *Taq* and Vent exo$^-$ to incorporate fluorescein dATP, and each of the dU derivatives alone or in combination with Cy5 dCTP. For each of the derivatives, the relative ability to incorporate the nucleotide analogs as compared to natural nucleotides was determined. Using a two-sample T-test for equal means with $\alpha = 0.01$, the results show that *Taq* is significantly better than Vent exo$^-$ at incorporating the dU analogs AMCA, TMR, and fluorescein, while Vent exo$^-$ is significantly better at incorporating Rhodamine Green dUTP and Cy5 dCTP (Figure 3). The results also show that fluorescein dATP is incorporated minimally and produces a signal not significantly different than the background (Figure 3).

Directly Observing the Incorporation of Nucleotide Analogs

The previous experiments indicate that DNA polymerases can incorporate a variety of nucleotide fluorophores, but they utilized a secondary α^{32}P reporter for quantitations. To directly measure the incorporation of the nucleotide fluorophore, we used a phosphoimager to detect the fluorescence signal associated with the nucleotide analogs. After the removal of the labeled DNA from residual non-incorporated-labeled nucleotides with two different methods

(binding to glass surfaces/Qiagen and size-exclusion chromatography), the DNA was precipitated on the nylon membrane and assayed both for radioactivity and fluorescence. The results of the fluorescent detection were in accord with those obtained by measuring incorporations of complementary nucleotide dye analogs into the growing DNA strand (Figure 4).

Extension Cut Assay

In order to measure the length of the newly synthesized product of the reactions, we developed an extension cut assay to test for the extension of the primer along a single-stranded template that is bound to a solid support. As the DNA strand is extended, restriction sites are created that can be used to monitor the size of the extended product. We used the extension cut assay to test *Taq* and Vent exo$^-$ at incorporating Rhodamine Green-labeled dUTP. The results of this assay show that the polymerases can produce DNA strands of greater than 900 bp that have dTTP fully replaced by Rhodamine Green dUTP (Figure 5). For an extended product, the restriction enzyme cleaves the double-stranded DNA at one or more sites, creating bands on a gel that can be observed. If the extension fails to proceed past the site of cleavage, then the single-stranded DNA at that position will not be cut by the restriction enzyme and will not produce a band on the gel. Extended products that contain dTTP replaced with Rhodamine Green dUTP migrate more slowly on the gel and will appear larger than the same sized DNA fragments containing only natural DNA.

DISCUSSION

Nucleotide analogs are routinely used to label, isolate, study, and manipulate DNA in a wide variety of applications (8–10,25). These nonradioactive nucleotide analogs are introduced into a DNA strand by chemical and enzymatic 5′ and 3′ end labeling and through internal enzymatic labeling or post-labeling methods (11–14). Most methods however, replace only a small fraction of the natural nucleotides with their respected analog counterpart. Complete replacement of a natural nucleotide with a modified nucleotide analog has been difficult to achieve. PCR labeling methods that completely replaced a nucleotide often produced an incorrect product (26). Partial nucleotide replacement is inadequate for applications such as single molecule sequencing, which requires complete replacement of two or more nucleotides. We were interested in identifying polymerases that can efficiently incorporate a variety of nucleotide analogs, including large hydrophobic fluorescent dye analogs. We screened 10 wild-type DNA polymerases from families A, B, and RT for the ability to incorporate several nucleotide analogs. Our analysis indicates that a strong 3′ to 5′ exonuclease activity hinders the ability of the polymerase to incorporate and extend nucleotide analogs in DNA. Presumably, the proofreading ability of the enzymes results in removal of the added nucleotide fluorophores (27).

We compared two polymerases each from families A, B, and RT that lack or had little 3′ to 5′ exonuclease activity. An analysis of these six polymerases showed that all could incorporate the nucleotide analogs. However, trends in the ability of the different polymerase family types to incorporate the different nucleotide derivatives were observed, with the family A polymerases appearing more efficient at incorporating dUTP labeled with AMCA (Figure 1). We chose to further analyze *Taq* from family A and Vent exo$^-$ from family B. Family RT polymerase was not chosen for further analysis, because even though the relative rates of analog incorporation were comparable to the family A and B polymerases, the rates of natural and analog nucleotide incorporation for the family RT polymerases was significantly lower than *Taq* and Vent exo$^-$ (data not shown).

Both an activated calf thymus DNA and a 2.6-kb linear template were used to test the ability to incorporate single and multiple nucleotide analogs using either *Taq* or Vent exo$^-$ polymerase

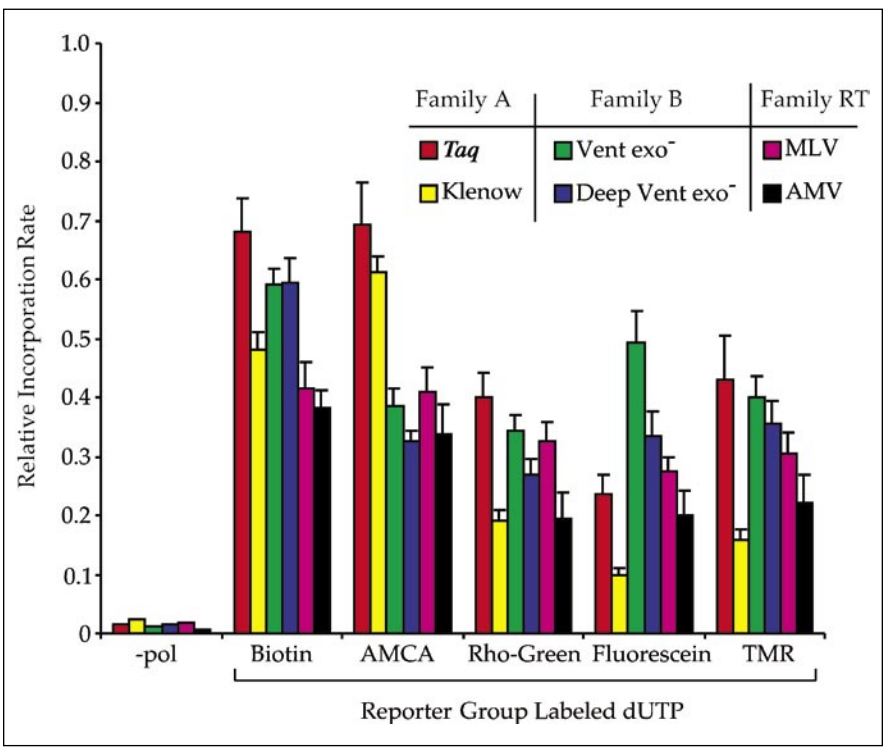

Figure 1. Relative ability to incorporate nucleotide analogs. The relative incorporation rate of each dUTP nucleoside analog as compared to natural nucleotides was determined for two family A, B, and RT polymerases. An activated calf thymus template and an [α-^{32}P]dGTP secondary reporter were used to determine the amount of nucleotide incorporation. Error bars indicate one standard error higher than the average. AMCA, 7-amino-4-methylcoumarin-3-acetic acid; Rho-Green, Rhodamine Green; TMR, tetramethylrhodamine.

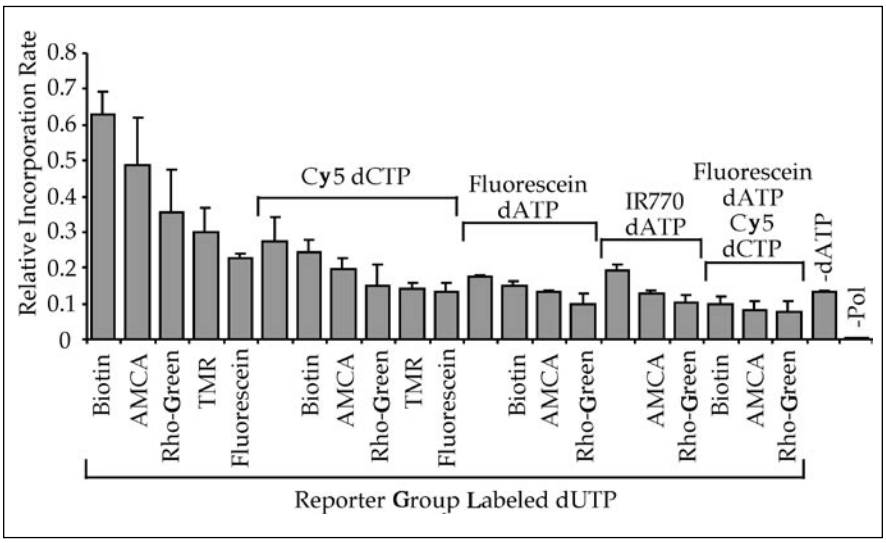

Figure 2. Multiple nucleotide analog incorporation by *Taq* and Vent exo⁻ polymerases. An activated calf thymus template and an [α-^{32}P]dGTP secondary reporter were used to determine the amount of nucleotide incorporation produced by *Taq* and Vent exo⁻ DNA polymerases. Natural nucleotides were replaced by one or more nucleotide analogs. The height of the bars indicate the average relative incorporation rate of *Taq* and Vent exo⁻ polymerases. Controls were included that contained either polymerase and no dATP or no polymerase. Error bars indicate one standard error higher than the average relative incorporation. AMCA, 7-amino-4-methylcoumarin-3-acetic acid; Rho-Green, Rhodamine Green; TMR, tetramethylrhodamine.

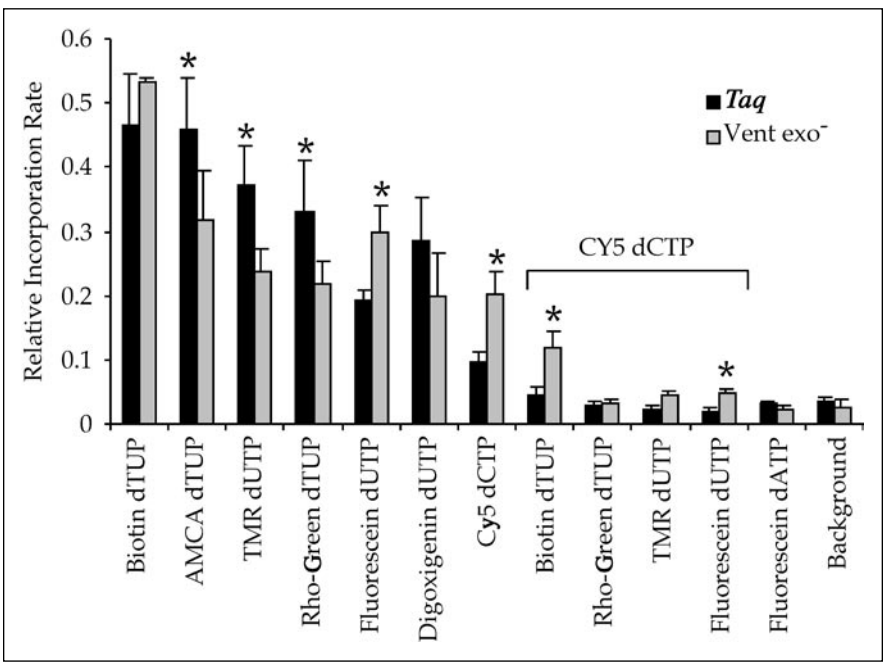

Figure 3. Analog incorporation of *Taq* and Vent exo⁻. The relative incorporation rates of various nucleotide analogs as compared to natural nucleotides using *Taq* and Vent exo⁻ polymerases are shown. A defined primer/template system and an [α-^{32}P]dGTP secondary reporter were was used to determine the amount of nucleotide incorporation. Significant differences in the abilities of the two polymerases to incorporate a given nucleotide analog are indicated by an asterisk (*). A two-sample T-test for equal means with $\alpha = 0.01$ was used to determine significance. Error bars indicate one standard error higher than the average. Including no dATP in the reaction produced the background incorporation rate. AMCA, 7-amino-4-methylcoumarin-3-acetic acid; Rho-Green, Rhodamine Green; TMR, tetramethylrhodamine.

(Figures 2 and 3). These analyses show that the DNA polymerase can have a profound effect on the ability to incorporate different nucleotide analogs, with no single polymerase appearing superior at incorporating all of the analogs. *Taq* polymerase was significantly better than Vent exo⁻ at incorporating AMCA, TMR, and Rhodamine Green-labeled dUTP, while Vent exo⁻ was significantly better at incorporating fluorescein dUTP and Cy5 dCTP. Furthermore, these experiments clearly demonstrate the influence of the position in the heterocycle of the nucleobase on which the linker-group and the reporter are attached. Modifications in the N7 position of the deazapurines (15) are incorporated fairly well, because the linker can protrude into the solvent accessible major groove of the DNA. The dA derivatives used here, however, are exclusively modified in the C8 position. The linker of the dA derivatives is therefore situated close to the sugar-phosphate backbone and may not allow efficient incorporation. Modifications in the C8 position of the dA derivatives may lead to a change/twist in the backbone structure of the DNA, diminishing further elongation due to steric interactions. The low incorporation rates of all the dA derivatives, even when they are used alone, may reflect these detrimental interactions. In all the cases where we combined these dA derivatives with other modified nucleotides, the incorporation efficiency is further reduced.

Reduced incorporation efficiency of Cy5 dCTP may be explained by the attachment of the reporter group to the position 4 of the 2′-deoxycytidine, which is involved as an electron donor in the base pairing with the C-6 carbonyl group of 2′-deoxyguanosine. Although this position is also pointing towards the major groove of the DNA, steric hindrance may be responsible for the reduced substrate activity, compared with dC derivatives, where the linker/reporter moiety is attached at the 5-position of the heterocycle (21). The data presented here show that the position for attaching reporter groups on 2′-deoxynucleotides must be carefully considered if a high label density is required, as the dA and dC derivatives used here, which work

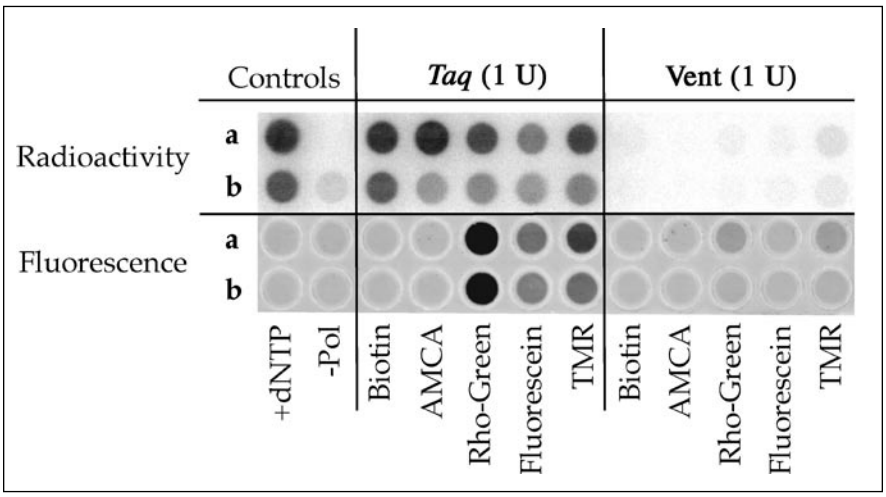

Figure 4. Direct observation of nucleotide analog incorporation. Phosphorimager results showing the incorporation of the radioactive secondary reporter [α-^{32}P]dGTP and the corresponding fluorescence of the nucleotide analogs. The labeled DNA was purified away from the unincorporated nucleotide analogs by either (a) glass bead DNA binding columns or (b) Sephacryl S-200 size exclusion columns. Rhodamine Green (Rho-Green), tetramethylrhodamine (TMR), and fluorescein were the only analogs tested that produce excitation/emission spectra capable of being detected with our instrumentation. Vent polymerase was used as a control of the DNA purification procedures, producing little signal by both radioactive and fluorescence detection methods. AMCA, 7-amino-4-methylcoumarin-3-acetic acid.

quite well, if used together with natural dNTPs. As has been shown by others, a more preferred position on the purine heterocycles is the 7-deaza position and for pyrimidines the heterocycle position 5 (15).

The ability of *Taq* and Vent exo$^-$ polymerases to incorporate nucleotide analogs was directly observed by visualizing the fluorescent signal produced by several of the analogs (Figure 4). The detection system was limited to detecting only the fluorescent analogs that were excited at a wavelength of 450 nm and had emissions greater than 520 nm. Rhodamine Green, TMR, and fluorescein fulfilled these requirements and could be detected. Even though these fluorescent dyes could be qualitatively detected, quantitation of the fluorescent signal was not possible due to quenching of these highly labeled DNA strands, as well as a high background and saturation of the fluorescent signal (9).

This direct observation of fluorescent signal demonstrates that the nucleotide analogs are being incorporated by the DNA polymerases, but does not indicate the lengths of the labeled products. Therefore, we developed an extension cut assay, using *Taq* or Vent exo$^-$ to produce a DNA strand that has dTTP completely replaced with a Rhodamine Green dUTP nucleotide analog. The extension cut assay uses a single-stranded DNA template bound to a solid support by a streptavidin-biotin conjugation. As the primer is elongated, the resulting double-stranded product can be cleaved by a series of restriction enzymes, indicating the length of the double-stranded product and thus the length of the extended primer. This method of determining the length of a labeled DNA strand has unique advantages over other methods. The hydrophobic nature of many of the nucleotide analogs makes size determination using an agarose or acrylamide gel difficult. Analogs incorporated into a DNA fragment can alter the physical properties of the DNA molecule, causing the DNA to migrate in a gel at a slower rate than DNA fragments containing only natural nucleotides. Furthermore, as the proportion of hydrophobic analogs increase, the DNA fragments begin to produce smears in the gel and may fail to enter the gel matrix itself.

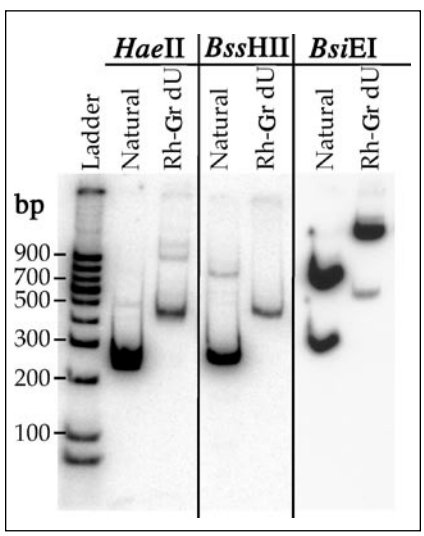

Figure 5. Extension cut assay. Products produced by the extension cut assay and run on a 6% nondenaturing acrylamide gel are shown, along with a labeled 100-bp ladder (Biotools, B&M Labs, S.A., Madrid, Spain). Restriction enzymes used to cleave the extended product that contains either natural nucleotides or dTTP replaced by Rhodamine Green dUTP (Rh-Gr dU) are indicated. The apparent shift in the bands produced when dTTP is replaced by the nucleotide analog is due to the hydrophobicity of the analog molecule.

The extension cut assay requires only that the product be cleaved by restriction enzymes. Since restriction enzymes may be unable to bind and cleave nucleotides that contain large hydrophobic analogs, we chose restriction enzymes that cleave only at GC sites, reducing any affects of replacing dTTP with Rhodamine Green dUTP. This analysis demonstrates that fragments >900 bp can be produced with dTTP completely replaced with a nucleotide analog (Figure 5).

One strategy for single molecule sequencing requires that two or more nucleotides be completely replaced with fluorescent nucleotide analogs (19–21). With current post-labeling methods unable to label 100% of the available positions, enzymatic incorporation of these nucleotide analogs may be the only method capable of complete labeling of a base at all positions. The overall results indicate the feasibility of completely labeling DNA with one or more fluorescent nucleotide analogs, synthesizing a product of 900 nucleotides in length. These steps provide the first requirements for an approach for copying single nucleotides of DNA and delineating nucleotide sequence by detection of released fluorescently labeled bases.

Polymerases both lacking 3′ to 5′ exonuclease activity and selected to

incorporate nucleotide analogs may possess low base incorporation fidelity. Examples of polymerase mutants that incorporate various modified nucleotides have been shown to have error rates as high as 0.1% (28,29). Such high error rates however should not preclude an enzyme from being used in single molecule sequencing. In most situations, random errors can be identified and eliminated by sequencing a region multiple times, allowing for the identification of the correct sequence.

ACKNOWLEDGMENTS

The authors thank Dr. Klaus Muehlegger for generously providing the Rhodamin-green-X-dUTP nucleotide analog. This work was supported by grants from the National Institutes of Health; CA78885, ES-07032-22, and the University of Washington (UW) National Institute of Environmental Health Sciences (NIEHS) sponsored Center for Ecogenetics and Environmental Health, grant NIEHS P30ES07033.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. **Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich, and N. Arnheim.** 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science *230*:1350-1354.
2. **Higuchi, R., B. Krummel, and R.K. Saiki.** 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res *16*:7351-7367.
3. **Reeve, M.A. and C.W. Fuller.** 1995. A novel thermostable polymerase for DNA sequencing. Nature *376*:796-797.
4. **Tabor, S. and C.C. Richardson.** 1995. A single residue in DNA polymerases of the *Escherichia coli* DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides. Proc. Natl. Acad. Sci. USA *92*:6339-6343.
5. **Chen, X. and P.Y. Kwok.** 1997. Template-directed dye-terminator incorporation (TDI) assay: a homogeneous DNA diagnostic method based on fluorescence resonance energy transfer. Nucleic Acids Res. *25*:347-353.
6. **Charmley, P., D. Nickerson, and L. Hood.** 1994. Polymorphism detection and sequence analysis of human T-cell receptor V alpha-chain-encoding gene segments. Immunogenetics *39*:138-145.
7. **Kessler, C., H.J. Holtke, R. Seibl, J. Burg, and K. Muhlegger.** 1990. Non-radioactive labeling and detection of nucleic acids. I. A novel DNA labeling and detection system based on digoxigenin: anti-digoxigenin ELISA principle (digoxigenin system). Biol. Chem. Hoppe Seyler *371*:917-927.
8. **Schena, M., D. Shalon, R.W. Davis, and P.O. Brown.** 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*:467-470.
9. **Randolph, J.B. and A.S. Waggoner.** 1997. Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes. Nucleic Acids Res. *25*:2923-2929.
10. **Kaiser, R.J., S.L. MacKellar, R.S. Vinayak, J.Z. Sanders, R.A. Saavedra, and L.E. Hood.** 1989. Specific-primer-directed DNA sequencing using automated fluorescence detection. Nucleic Acids Res. *17*:6087-6102.
11. **Smith, L.M., J.Z. Sanders, R.J. Kaiser, P. Hughes, C. Dodd, C.R. Connell, C. Heiner, S.B. Kent, and L.E. Hood.** 1986. Fluorescence detection in automated DNA sequence analysis. Nature *321*:674-679.
12. **Trevisiol, E., A. Renard, E. Defrancq, and J. Lhomme.** 2000. Fluorescent labelling of oligodeoxyribonucleotides by the oxyaminoaldehyde coupling reaction. Nucleosides Nucleotides Nucleic Acids *19*:1427-1439.
13. **Oyama, F., R. Kikuchi, A. Omori, and T. Uchida.** 1988. 3′-End labeling of DNA fragments by AMV-reverse transcriptase. Nucleic Acids Symp. Ser. *19*:115-116.
14. **Starke, H.R., J.Y. Yan, J.Z. Zhang, K. Muhlegger, K. Effgen, and N.J. Dovichi.** 1994. Internal fluorescence labeling with fluorescent deoxynucleotides in two-label peak-height encoded DNA sequencing by capillary electrophoresis. Nucleic Acids Res. *22*:3997-4001.
15. **Seela, F., E. Feiling, J. Gross, F. Hillenkamp, N. Ramzaeva, H. Rosemeyer, and M. Zulauf.** 2001. Fluorescent DNA: the development of 7-deazapurine nucleoside triphosphates applicable for sequencing at the single molecule level. J. Biotechnol. *86*:269-279.
16. **Muhlegger, K., E. Huber, H. von der Eltz, R. Ruger, and C. Kessler.** 1990. Non-radioactive labeling and detection of nucleic acids. IV. Synthesis and properties of digoxigenin-modified 2′-deoxyuridine-5′-triphosphates and a photoactivatable analog of digoxigenin (photodigoxigenin). Biol. Chem. Hoppe Seyler *371*:953-965.
17. **Yu, H., J. Chao, D. Patek, R. Mujumdar, S. Mujumdar, and A.S. Waggoner.** 1994. Cyanine dye dUTP analogs for enzymatic labeling of DNA probes. Nucleic Acids Res. *22*:3226-3232.
18. **Zhu, Z., J. Chao, H. Yu, and A.S. Waggoner.** 1994. Directly labeled DNA probes using fluorescent nucleotides with different length linkers. Nucleic Acids Res. *22*:3418-3422.
19. **Jett, J.H., R.A. Keller, J.C. Martin, B.L. Marrone, R.K. Moyzis, R.L. Ratliff, N.K. Seitzinger, E.B. Shera, and C.C. Stewart.** 1989. High-speed DNA sequencing: an approach based upon fluorescence detection of single molecules. J. Biomol. Struct. Dyn. *7*:301-309.
20. **Stephan, J., K. Dorre, S. Brakmann, T. Winkler, T. Wetzel, M. Lapczyna, M. Stuke, B. Angerer, et al.** 2001. Towards a general procedure for sequencing single DNA molecules. J. Biotechnol. *86*:255-267.
21. **Augustin, M.A., W. Ankenbauer, and B. Angerer.** 2001. Progress towards single-molecule sequencing: enzymatic synthesis of nucleotide-specifically labeled DNA. J. Biotechnol. *86*:289-301.
22. **Glick, E., J.P. Anderson, and L.A. Loeb.** 2002. In vitro production and screening of DNA polymerase eta mutants for catalytic diversity. BioTechniques *33*:1136-1144.
23. **Kornberg, A. and T.A. Baker.** 1992. DNA Replication. W.H. Freeman and Co., New York.
24. **Ono, T., M. Scalf, and L.M. Smith.** 1997. 2′-Fluoro modified nucleic acids: polymerase-directed synthesis, properties and stability to analysis by matrix-assisted laser desorption/ionization mass spectrometry. Nucleic Acids Res. *25*:4581-4588.
25. **Foldes-Papp, Z., B. Angerer, W. Ankenbauer, and R. Rigler.** 2001. Fluorescent high-density labeling of DNA: error-free substitution for a normal nucleotide. J. Biotechnol. *86*:237-253.
26. **Zhu, Z. and A.S. Waggoner.** 1997. Molecular mechanism controlling the incorporation of fluorescent nucleotides into DNA by PCR. Cytometry *28*:206-211.
27. **Perler, F.B., S. Kumar, and H. Kong.** 1996. Thermostable DNA polymerases. Adv. Protein Chem. *48*:377-435.
28. **Patel, P.H., H. Kawate, E. Adman, M. Ashbach, and L.A. Loeb.** 2001. A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. J. Biol. Chem. *276*:5044-5051.
29. **Gardner, A.F. and W.E. Jack.** 1999. Determinants of nucleotide sugar recognition in an archaeon DNA polymerase. Nucleic Acids Res. *27*:2545-2553.

Received 23 August 2004; accepted 17 September 2004.

Address correspondence to:

Jon P. Anderson
LI-COR, Inc.
4308 Progressive Avenue
P.O. Box 4000
Lincoln, NE 68504, USA
e-mail: janderson@licor.com