Optimized incorporation of an unnatural fluorescent amino acid affords measurement of conformational dynamics governing high-fidelity DNA replication

Received for publication, August 10, 2020, and in revised form, September 29, 2020 Published, Papers in Press, October 5, 2020, DOI 10.1074/jbc.RA120.015557

Tyler L. Dangerfield and Kenneth A. Johnson*©
From the Institute for Cellular and Molecular Biology, Department of Molecular Biosciences, University of Texas, Austin, Texas, USA

Edited by F. Peter Guengerich

DNA polymerase from bacteriophage T7 undergoes large, substrate-induced conformational changes that are thought to account for high replication fidelity, but prior studies were adversely affected by mutations required to construct a Cys-lite variant needed for site-specific fluorescence labeling. Here we have optimized the direct incorporation of a fluorescent unnatural amino acid, (7-hydroxy-4-coumarin-yl)-ethylglycine, using orthogonal amber suppression machinery in Escherichia coli. MS methods verify that the unnatural amino acid is only incorporated at one position with minimal background. We show that the single fluorophore provides a signal to detect nucleotide-induced conformational changes through equilibrium and stopped-flow kinetic measurements of correct nucleotide binding and incorporation. Pre-steady-state chemical quench methods show that the kinetics and fidelity of DNA replication catalyzed by the labeled enzyme are largely unaffected by the unnatural amino acid. These advances enable rigorous analysis to establish the kinetic and mechanistic basis for high-fidelity DNA replication.

The role of substrate-induced conformational changes in enzyme specificity has been a controversial topic for decades (1). In particular, the extraordinarily high fidelity of DNA replication by DNA polymerases has been the subject of studies in enzymology for many years (2, 3). Paradoxically, high-fidelity polymerases make fewer mistakes and copy DNA at a much faster rate than low-fidelity DNA polymerases (1, 4). Crystal structures of the T7 DNA polymerase (5), a model system for high-fidelity DNA replication, showed large structural changes in the enzyme upon nucleotide binding, as have other polymerases, raising the question of the contribution of this step to the enzyme’s high fidelity. The original T7 DNA polymerase papers were published almost 30 years ago, measuring the correct incorporation (2), misincorporation (3), and exonuclease proofreading reactions (6). These papers provided valuable insights into the kinetic basis for the high DNA replication fidelity of T7 DNA polymerase but were flawed in the analysis of data regarding a prechemistry nucleotide-induced conformational change, based on the interpretation of an observed small thio-elemental effect. The notion that a small observed thio-elemental effect can be interpreted as evidence for the existence of a rate-limiting conformational change preceding chemistry has since been shown to be invalid because the magnitude of the elemental effect is largely dependent on the nature of the transition state and is smaller than previously expected (7, 8). The first direct measurement of a prechemistry conformational change came from studies on the low-fidelity repair enzyme, polymerase β, using a 2-aminopurine fluorescent reporter in the DNA (9–11). These studies showed that the conformational change was faster than chemistry, challenging previous notions that the conformational change had to be rate-limiting to contribute to fidelity. Critics of these studies, however, proposed that the fluorescence signal could be due to changes in the DNA structure and translation, rather than the conformational change in the enzyme. To directly measure the rate of the conformational change for the T7 DNA polymerase, an 8-Cys-lite T7 DNA polymerase variant was labeled with 7-diethylamino-3-(((2-maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC) at a surface-exposed engineered cysteine in the highly mobile fingers domain (4). By labeling the protein instead of the DNA and measuring the relevant rate constants, these studies confirmed that the rate of the conformational change was indeed faster than chemistry for both correct and mismatched nucleotides and discovered that the reverse of the conformational change was much slower than chemistry for the correct nucleotide but much faster than chemistry for the mismatched nucleotide, suggesting a new paradigm for understanding enzyme specificity (4). However, the extensive mutagenesis required to create the 8-Cys-lite variant negatively affected the enzyme. This variant had low solubility and most importantly at least a 50-fold decrease in the apparent Kd for a mismatched nucleotide, reducing the enzyme’s fidelity. Subsequent similar studies using HIV reverse transcriptase, where only a single cysteine had to be removed, confirmed the central role of the nucleotide-induced conformational change in DNA replication fidelity (12). However, HIV reverse transcriptase has only moderate fidelity, so it is important to more accurately assess the role of conformational changes in specificity with a high-fidelity enzyme, which is all the more important given the controversies (13).

Because the T7 DNA polymerase is an ideal model system for a fast, high-fidelity enzyme, it was essential to improve upon strategies for labeling the enzyme with a fluorophore to report changes in enzyme structure. Here, we present the results of our optimization of the expression, purification, and basic kinetic characterization of a variant containing the fluorescent unnatural amino acid (7-hydroxy-4-coumarin-yl)-ethylglycine...
Unnatural fluorescent amino acid to monitor polymerase dynamics

Figure 1. Conformational changes around the active site of T7 DNA polymerase upon nucleotide binding. Large rearrangements are seen in the fingers domain around the polymerase active site going from the open binary E-D complex (PDB: 2A9Q, gray) to the ternary closed E-D-[MgCl₂-dATP] complex (PDB: 1T7P, blue). We show DNA bases (green) and the incoming dATP (magenta) with two chelated Mg²⁺ ions (red) for the closed state with residue 514 shown with the unnatural amino acid 7-HCou (yellow) modeled in place of glutamic acid (ligand ID: DV7).

(7-HCou) on the surface of the fingers domain. We show that the fluorescence is sensitive to conformational changes of the protein during nucleotide incorporation (Fig. 1). By comparison to WT enzyme, we show that labeling of the enzyme with the fluorescent amino acid has negligible effects on polymerization kinetics and fidelity. By optimizing expression, the proteins could be purified without the use of added tags which also might alter enzyme activity. The methods in this paper overcome the limitations of the cysteine-maleimide labeling strategy and provide the groundwork for full characterization of the role of conformational dynamics of this enzyme on DNA replication fidelity and proofreading. Our optimized methods can be extended to other enzyme systems.

Results

Unnatural amino acid amber codon suppression strategy

Small fluorescence probes with quantum yields that are sensitive to local environment have significant advantages over larger probes commonly used in biochemistry, such as GFP, Cy3, quantum dots, and dual probes for FRET distance measurements. With a probe sensitive to local environment, there is no need to do traditional FRET experiments that require dual labeling with commensurate higher chances to interfere with enzyme function, because comprehensive global data fitting defines the step being measured by the fluorescence signal (14, 15). The optical properties of 7-hydroxycoumarin in various solvents (16) and at different pH values (pKₐ ≈ 7.8 (17)) have been well-characterized and show that this fluorophore can be used to probe local environment.

To circumvent the problems associated with the cysteine labeling strategy used for the 8-Cys-lite variant (18), we tested the amber suppression to insert a site-specific unnatural amino acid, requiring only one residue substitution in the protein to directly incorporate a fluorescent unnatural amino acid into the fingers domain of the T7 DNA polymerase (Fig. 1). This unnatural amino acid system repurposes the amber stop codon (TAG) as the signal for insertion of an unnatural amino acid, using an evolved orthogonal Methanocaldococcus jannaschii tyrosyl-tRNA synthetase variant (aminoacyl tRNA synthetase; AARS) and M. jannaschii tyrosyl-tRNA with CU in the anticodon loop (tRNA_Mo) expressed in E. coli cells. The amber stop codon is the least used of the three stop codons in E. coli (~9%) and rarely terminates essential genes (19). The M. jannaschii AARS is amenable to directed evolution to incorporate a wide variety of unnatural amino acids (4). To implement this system with the T7 DNA polymerase, the codon for residue 514 in T7 gene 5 in pclₘₖ₋₅(T7 G₅ₓ⁻, trxA), described below, was mutated to an amber codon by site-directed mutagenesis to provide the codon for unnatural amino acid incorporation; the resulting plasmid was designated pclₘₖ₋₅(T7 G₅ₓ⁻ E514TAG, trxA). This position was chosen because it was the amino acid position that gave the best signal using the cysteine-MDCC labeling strategy with the 8-Cys-lite T7 DNA polymerase variant (4).

Bicistronic, heat-induced T7 DNA polymerase expression/purification

We first worked to optimize the expression and purification of WT T7 DNA polymerase. Previous studies showed that coexpression of E. coli thioredoxin and T7 gene product 5 (GP5), the two polypeptides that together constitute the T7 DNA polymerase holoenzyme, greatly enhanced the solubility of a histidine-tagged T7 DNA polymerase variant (20). Although thioredoxin was constitutively expressed in E. coli, significant levels of overexpression of T7 GP5 required overexpression of the thioredoxin cofactor as well. We therefore constructed the plasmid pclₘₖ₋₅(T7 G₅ₓ⁻, trxA) (Fig. 2A), containing a bicistronic operon of T7 gene 5 (without a histidine tag) and trxA cloned into the pclₘₖ backbone (21). The bicistronic operon contains a short linker between the genes for thioredoxin and T7 GP5, consisting of a PmeI restriction site followed by a ribosome binding site (Fig. 2B). The pclₘₖ backbone includes a temperature-sensitive λ clₘₖ repressor encoded on the plasmid that binds tightly to the operator sequence at 30 °C, blocking expression, but becomes inactive at temperatures greater than ~35 °C. Having the repressor encoded by the plasmid allows expression in most E. coli strains, and no chemical inducer is required for high-level expression.

We tested the utility of this system for expressing the T7 DNA polymerase in E. coli harboring the plasmid pclₘₖ₋₅(T7 G₅ₓ⁻, trxA). The cells were grown at 30 °C before heat induction at 42 °C for 20 min, followed by expression at 38 °C for 3 h, and were isolated by centrifugation. Cell pellets were later lysed and processed to obtain samples corresponding to the “Total Cell Protein,” “Soluble Protein,” and “Insoluble Protein” fractions at each time point. The gel in Fig. 2C shows soluble and highly overexpressed T7 DNA polymerase, easily seen above the other E. coli proteins. The procedure for purification of T7
DNA polymerase was based on further optimization of previous protocols (2) to improve purity, increase the capacity of the method to accommodate the large increase in protein expression, and reduce the number of time-consuming dialysis steps. The optimized expression/purification protocol routinely yielded 50–100 mg of WT T7 DNA polymerase/liter of media. The purity and stoichiometry of the final product are shown in Fig. 2D, with T7 GP5 and thioredoxin copurifying in approximately a 1:1 molar ratio.

**Plasmid optimizations for amber suppression machinery**

To change the specificity of the *M. jannaschii* tyrosyl-tRNA synthetase (MjYRS) from tyrosine to 7-HCou, eight mutations were made in the AARS gene of the unnatural amino acid machinery plasmid pSUP-(MjYRS-6TRN), as described by Schultz and co-workers (22). The resulting plasmid containing the mutant AARS was designated pSUP-(CouRS-6TRN) (Fig. 3A). Amber suppression efficiency at position 514 with this system was tested with *E. coli* harboring both pSUP-(CouRS-6TRN) and pcl\[^{\text{I}}\text{ts}\]-(T7 G5x\(^{-}\), trxA) (Fig. 4). As a control, a parallel culture was grown in the absence of 7-HCou to test for background incorporation of naturally occurring amino acids at position 514 in T7 GP5. No band for GP5 was observed in the lanes where 7-HCou was not added to the media (Fig. 4, lane 4), whereas a faint band appeared on the gel for the culture grown in the presence of 7-HCou (Fig. 4, lane 2), indicating low amber suppression efficiency. The literature contains many examples of optimization of various inefficient unnatural amino acid systems (23, 24), so we began testing a number of them to improve the yield of T7 DNA polymerase E514Cou.

The first plasmid optimization was designed to test the effect of an amber suppressor tRNA variant of tRNA\[^{\text{Mj}}\] evolved to improve unnatural amino acid incorporation efficiency (25). Cells expressing six copies of tRNA\[^{\text{Mj}}\] from pSUP-(CouRS-6TRN) grew very slowly and did not reach the same stationary phase density as cells expressing the WT enzyme. As previously reported by others (23), one of the tricistronic operons was occasionally lost during cloning steps, presumably from homologous recombination because of the adjacent, identical tRNA operons, resulting in a mixture of cells with different numbers of copies of tRNA\[^{\text{Mj}}\]. Substituting one copy of an optimized tRNA\[^{\text{Mj}}\] designated tRNA\[^{\text{opt}}\] (25), for six copies of tRNA\[^{\text{Mj}}\] has been successful with other unnatural amino acid systems and is reported to reduce the toxicity (25). The optimized tRNA was inserted in place of the six copies of tRNA\[^{\text{Mj}}\] with ligation-independent cloning (26), and the resulting plasmid containing a single copy of tRNA\[^{\text{opt}}\] is designated pSUP-(CouRS-tRNA\[^{\text{opt}}\]) (Fig. 3B). Expression tests of *E. coli* harboring both pcl\[^{\text{I}}\text{ts}\]-(T7 G5x-E514TAG, trxA) and pSUP-(CouRS-tRNA\[^{\text{opt}}\]) show a
large improvement in expression over the pSUP-(CouRS-6TRN) plasmid (Fig. 4, lane 2 versus lane 6). Additionally, no significant band was detected for the T7 DNA polymerase in the cultures without 7-HCou added to the media (Fig. 4, lane 4 versus lane 8), indicating that CouRS appears to be inefficient however specific for 7-HCou over endogenous amino acids. Cells harboring pSUP-(CouRS-tRNAopt) grew much faster and to higher densities than cells harboring pSUP-(CouRS-6TRN) and were not prone to loss of one of the operons. Other studies have found that an inducible copy of the AARS gene, combined with the constitutive copy from the mutant promoter for the *E. coli* glutamine-tRNA ligase gene (*glnS*') (27), can enhance amber suppression efficiency (23, 28). To test the effect of an inducible copy of CouRS in our system, we added a copy of the gene under control of the *l* promoter. Because the T7 DNA polymerase genes are also expressed under this promoter, supplemental expression of CouRS occurs simultaneously upon heat induction. This plasmid, pSUP-(2<p><sub>3</sub>CouRS-tRNAopt) (Fig. 5A), contains a single copy of tRNAopt, p15A origin of replication, CmR gene, and two copies of CouRS. This construct was tested for T7 DNA polymerase E514Cou expression, as above. An enhancement in amber suppression efficiency was observed, so that a significant band of the expected molecular mass appeared in the post-induction samples with 7-HCou added to the expression media (Fig. 5B, lane 2 versus lane 7). One final CouRS optimization was attempted by making the D286R mutation in CouRS (not shown). Some unnatural amino acid systems have shown this mutation to further enhance amber suppression (29); however, we found that this mutation did not significantly enhance the efficiency of our system. The plasmid pSUP-(2<p><sub>3</sub>CouRS-tRNAopt) was therefore used for all further expressions.

Other nonplasmid-based optimizations

Some optimization strategies were tested that did not require any DNA manipulation. Expression tests on the WT enzyme comparing rich media (Terrific Broth) with a supplemented M9 media (see “Experimental procedures”) showed slightly higher specific expression of T7 DNA polymerase with the minimal media (not shown). This has the added advantage of greatly reducing the concentration of naturally occurring amino acids in the media, therefore reducing the likelihood of
Large-scale expression and purification of T7 DNA polymerase E514Cou

Large-scale protein expression was performed by inoculating six 1-liter cultures of supplemented M9 media with E. coli harboring pSUP-(2×CouRS-tRNAopt) and pcl8s-(T7 G5x-E514TAG, trxA). The cultures were grown to an A600 of 0.5, and then the cells were gently pelleted and resuspended in 600 ml of supplemented M9 media containing 1 mM 7-HCou. Protein expression was induced by elevating the culture temperature from 30 to 42 °C, incubating for 30 min, and then dropping the temperature to 38 °C and incubating for 3 h. The T7 DNA polymerase E514Cou variant and the WT enzyme were purified by monitoring absorbance at 335 nm on the HPLC. Purification from a pellet from the 600-ml condensed culture yielded 5–10 mg of the enzyme containing 7-HCou.

MS

The 7-HCou–labeled T7 DNA polymerase variant was analyzed by MS performed two ways to address two questions regarding quality control on the purified protein. We reasoned that potential inhomogeneity in our protein preparation due to the unnatural amino acid system could be from either naturally occurring amino acids incorporated at position 514 or, less likely, 7-HCou incorporation at other positions in the protein. Aminoacylation of the orthogonal tRNA with natural amino acids and incorporation into the protein could occur either by promiscuous activity of CouRS or promiscuous activity of the translation machinery. Alternatively, endogenous aminoacylated tRNAs could wobble base pair with the amber codon for insertion. Either situation would lead to a fraction of purified protein that does not contain 7-HCou. To test...
whether naturally occurring amino acids were incorporated at residue 514 in a fraction of our purified protein, samples of WT T7 DNA polymerase and the E514Cou variant were separated by SDS-PAGE and the gel bands containing GP5 were excised and trypsinized. The tryptic peptides were then separated by reverse-phase HPLC and analyzed continuously by MS/MS on a Thermo Fisher Scientific Orbitrap Fusion mass spectrometer, and the data were processed as described under “Experimental procedures.” Approximately 1,700 spectra were matched to T7 gene product 5 peptides for each sample. For the E514Cou sample, 59 spectra were obtained showing the Glu → 7-HCou modification at position 514, and a representative spectrum is shown in Fig. 6. Peptides identified for the E514Cou variant are given in the table in the supporting information. We also separately searched the data against databases with T7 gene product 5 containing other amino acid substitutions at position 514 and found three spectra with glutamine substitution at 514 and one spectrum with tryptophan substitution at 514 and one spectrum with tryptophan insertion at this position, suggesting that ~93% of the protein was labeled with 7-HCou at position 514.

To address whether 7-HCou was incorporated at other positions in the protein, a slightly different mass spectrometry approach was employed. 7-HCou–labeled T7 DNA polymerase was trypsinized in solution, and the peptides were subsequently separated by reverse phase HPLC. Absorbance of 7-HCou was monitored at 335 nm during the separation and two peaks were obtained with an absorbance at this wavelength (not shown). These peaks were analyzed by MALDI-MS, which showed that they were the peptide 509-NQIAAXLPTR518 and the missed cleavage peptide 509-NQIAAXLPTDRN522. Based on these results, there was no significant evidence for incorporation of 7-HCou at any position other than 514 in the labeled protein, and we conservatively estimate the fraction of protein with 7-HCou inserted at position 514 versus natural amino acids to be greater than 90%.

**Fluorescence excitation/emission spectra**

Next, the fluorescence excitation and emission spectra were collected for various T7 DNA polymerase E514Cou complexes (Fig. 7) to determine whether the fluorescence intensity was sensitive to changes in protein conformation upon nucleotide binding. DNA substrates consist of a 27-nt primer annealed to a 45-nt template and are based on the semi-random but defined sequence used in a prior study of HIV reverse transcriptase (30). A dsDNA region of 27 bases has been shown to be long enough for tight binding to the T7 DNA polymerase (2). To prevent chemistry but still allow nucleotide binding and conformational change, a DNA substrate was used containing a 2’,3’-dideoxy modification on the 3’-terminus of the primer strand (4). The excitation spectrum showed an excitation maximum of 361 nm, with a slightly lower intensity peak around 335 nm. A smaller peak of intensity around 280 nm was also observed, resulting from excitation of tryptophan residues and FRET to the 7-HCou. This result gives us the option of exciting the fluorophore directly (ex: 335 nm, not shown) or exciting tryptophan residues (ex: 295 nm) in the protein and monitoring the FRET signal from 7-HCou emission. T7 DNA polymerase E514Cou also showed a small increase in fluorescence upon DNA binding and a further increase in fluorescence (~20%) upon addition of the correct nucleotide to the E + Ddd complex, indicating that the T7 DNA polymerase variant could potentially be used to measure the nucleotide-induced conformational change (Fig. 7).

Both excitation wavelengths gave a signal change upon nucleotide binding; however, the FRET signal was much larger and less sensitive to solution conditions, so it was used in subsequent kinetic experiments and for the fluorescence emission spectra shown here. We cannot identify which Trp residues contribute to the FRET signal. There are 20 total tryptophan residues in the T7 DNA polymerase, two in thioredoxin and 18 in gene product 5. Tallies for the distances of Trp residues to Glu514 are as follows based on the crystal structure: two at 25–30 Å, seven at 30–40 Å, five at 40–50 Å, four at 50–60 Å, and two at 60–70 Å. Most (seven) are in the palm domain, with three in the fingers domain, two in the thumb domain, five in the exonuclease domain, one in the thioredoxin binding domain, and two in thioredoxin. Because of the large number of tryptophan residues in the protein and the added complication of the FRET efficiency effect of the orientation factor for tryptophan residues with similar distances, we are not able to suggest a single candidate Trp as the donor. Nonetheless, kinetic and
Unnatural fluorescent amino acid to monitor polymerase dynamics

The slower phase corresponds to structural changes in the enzyme occurring after the chemistry step, based on measurements of the rate of the chemical reaction under identical conditions, as described below.

Experimental strategy to measure fidelity/correct nucleotide burst experiments

The MDCC-labeled 8-Cys-lite T7 DNA polymerase (4, 18) has reduced fidelity because of the extensive mutagenesis required to achieve site-specific labeling at an engineered cysteine residue. We therefore tested whether the 7-HCou substitution altered the fidelity of our T7 DNA polymerase E514Cou variant. To address this question, pre-steady-state kinetic experiments were performed at various nucleotide concentrations for both the correct and mismatched nucleotide to estimate the parameters $k_{pol}$ and $K_{d,app}$ for each nucleotide (30). These parameters are important because the ratio $k_{pol}/K_{d,app}$ determines enzyme specificity and the ratio of $(k_{pol}/K_{d,app})_{Mismatch}$ versus $(k_{pol}/K_{d,app})_{Correct}$ defines the fidelity of the enzyme. To measure $k_{pol}$ and $K_{d,app}$ for the correct nucleotide, rapid-quench pre-steady-state burst experiments were performed (Fig. 9), which allowed reaction times down to 2.5 ms to be accurately measured. The temperature of the reaction was maintained at 4 °C with a circulating water bath. T7 DNA polymerase (60 nM WT and 75 nM E514Cou) with thioredoxin (2 μM), BSA (0.1 mg/ml), and 6-carboxyfluorescein (FAM)-labeled-27/45-18T (with T as the templating base) DNA (125 nM WT and 150 nM E514Cou) were mixed with various concentrations of dATP (2.5–110 μM) and Mg$^{2+}$ (12.5 mM) to start the reaction. EDTA from the quench syringe was used to stop the reaction, and products were analyzed by denaturing PAGE or capillary electrophoresis. The exponential “burst” phase arises from the fast rate of incorporation of nucleotide by the DNA-bound enzyme, relative to the slow rate of DNA dissociation, before binding a new substrate DNA molecule to allow further nucleotide incorporation. Data were fit using KinTek Explorer with the model in Scheme 2 to extract the parameters $k_{pol}$ and $K_{d,app}$. For comparison, the results of equation-based data fitting are shown in Fig. S5. Global data fitting returned values with standard errors based on confidence contour analysis to define $k_{pol} = 181 \pm 25$ s$^{-1}$ and $K_{d,app} = 28 \pm 8$ μM for WT, and $k_{pol} = 202 \pm 14$ s$^{-1}$ and $K_{d,app} = 15 \pm 3$ μM E514Cou T7 DNA polymerase. A $\chi^2$ threshold of 0.96 was used to define confidence limits based on the number of parameters and the number of data points in the fitting (14, 31). Both enzymes have almost identical values of $k_{pol}$, whereas the apparent $K_d$ for the nucleotide appears slightly lower for the E514Cou variant. Nonetheless, these data indicate that incorporation of 7-HCou has a minimal effect on the kinetics of correct nucleotide incorporation.

Misincorporation reactions

The misincorporation reaction was much slower and could therefore be assayed by handmixing. Reactions were performed using a modified tube rack, designed to sit inside a refrigerated water bath to provide a constant temperature of 4 °C for the

Figure 7. Fluorescence excitation and emission spectra for various T7 DNA polymerase E514Cou complexes. The excitation spectrum (red) has a maximum at 361 nm but also contains a smaller peak with a local maximum around 280 nm. Emission spectra were collected with excitation of the sample at 295 nm, first for the free enzyme (300 nM) with thioredoxin (6 μM) and Mg$^{2+}$ ($E$, black). An excess of DNA$_{dd}$ (400 nM) was added, resulting in a small increase in fluorescence for the binary complex ($E$ + DNA$_{dd}$ blue). dATP (10 μM) was then added and resulted in a further increase in fluorescence for the ternary complex ($E$ + DNA$_{dd}$ + dATP, green). The emission maximum was at 447 nm.

stopped-flow trace best fit a double exponential equation, with an increase in fluorescence at an exponential decay rate of 136 s$^{-1}$ (the same rate obtained for the experiment with the DNA$_{dd}$ substrate), followed by a slower decrease in fluorescence back to the original value at a decay rate of 100 s$^{-1}$.

Equilibrium measurements show that the change in fluorescence intensity is correlated with the formation of the closed enzyme state.

Stopped-flow kinetics

Because of the rapid rate of nucleotide incorporation by the T7 DNA polymerase, a stopped-flow instrument was used to measure the fluorescence signal from 7-HCou upon addition of nucleotide and all experiments were performed at 4 °C. To initially simplify the reaction scheme to only nucleotide binding steps (Scheme 1), the DNA$_{dd}$ substrate was used to measure the nucleotide binding rate at a fixed nucleotide concentration, with the observed rate a function of $k_{pol}$ and $K_{d,app}$ for each nucleotide (30). These parameters are important because the ratio $k_{pol}/K_{d,app}$ determines enzyme specificity and the ratio of $(k_{pol}/K_{d,app})_{Mismatch}$ versus $(k_{pol}/K_{d,app})_{Correct}$ defines the fidelity of the enzyme. To measure $k_{pol}$ and $K_{d,app}$ for the correct nucleotide, rapid-quench pre-steady-state burst experiments were performed (Fig. 9), which allowed reaction times down to 2.5 ms to be accurately measured. The temperature of the reaction was maintained at 4 °C with a circulating water bath. T7 DNA polymerase (60 nM WT and 75 nM E514Cou) with thioredoxin (2 μM), BSA (0.1 mg/ml), and 6-carboxyfluorescein (FAM)-labeled-27/45-18T (with T as the templating base) DNA (125 nM WT and 150 nM E514Cou) were mixed with various concentrations of dATP (2.5–110 μM) and Mg$^{2+}$ (12.5 mM) to start the reaction. EDTA from the quench syringe was used to stop the reaction, and products were analyzed by denaturing PAGE or capillary electrophoresis. The exponential “burst” phase arises from the fast rate of incorporation of nucleotide by the DNA-bound enzyme, relative to the slow rate of DNA dissociation, before binding a new substrate DNA molecule to allow further nucleotide incorporation. Data were fit using KinTek Explorer with the model in Scheme 2 to extract the parameters $k_{pol}$ and $K_{d,app}$. For comparison, the results of equation-based data fitting are shown in Fig. S5. Global data fitting returned values with standard errors based on confidence contour analysis to define $k_{pol} = 181 \pm 25$ s$^{-1}$ and $K_{d,app} = 28 \pm 8$ μM for WT, and $k_{pol} = 202 \pm 14$ s$^{-1}$ and $K_{d,app} = 15 \pm 3$ μM E514Cou T7 DNA polymerase. A $\chi^2$ threshold of 0.96 was used to define confidence limits based on the number of parameters and the number of data points in the fitting (14, 31). Both enzymes have almost identical values of $k_{pol}$, whereas the apparent $K_d$ for the nucleotide appears slightly lower for the E514Cou variant. Nonetheless, these data indicate that incorporation of 7-HCou has a minimal effect on the kinetics of correct nucleotide incorporation.
Unnatural fluorescent amino acid to monitor polymerase dynamics

Figure 8. Stopped-flow dATP (correct nucleotide) binding and incorporation kinetics at 4 °C. A, stopped-flow dATP binding-rate experiment. T7 DNA polymerase E514Cou (750 nM), thioredoxin (15 μM), and 27/45-18T DNA (1 μM) were mixed with dATP (10 μM) and Mg2+ (12.5 mM) in a KinTek SF-300 stopped-flow instrument. The single exponential fit to the data (black line) gives an observed decay rate of 136 ± 1 s⁻¹. B, stopped-flow dATP incorporation experiment. T7 DNA polymerase E514Cou (750 nM), thioredoxin (15 μM), and 27/45-18T DNA (1 μM) were mixed with dATP (10 μM) and Mg2+ (12.5 mM). The fit of the data to a double exponential function is shown as a black line. With a fast phase at 136 s⁻¹, the slower phase showing a decrease in fluorescence has an observed decay rate of 100 ± 2 s⁻¹. For comparison, the exponential phase of the quench-flow burst data at 10 μM dATP gave an observed rate of 92 ± 10 s⁻¹ (Fig. 9).

Figure 9. dATP (correct nucleotide) burst kinetics at 4 °C for WT and E514Cou T7 DNA polymerase by rapid quench. A, pre-steady-state burst kinetics for WT T7 DNA polymerase. WT T7 DNA polymerase (60 nM), thioredoxin (1.2 μM), BSA (0.1 mg/ml), and FAM-27/45-18T DNA (150 nM) were mixed with Mg2+ (12.5 mM) and varying concentrations of dATP (2.5–110 μM) in a KinTek RQF-3 quench-flow instrument, and the reaction was quenched with EDTA from a quench syringe. Data for each dATP concentration are shown in different colors. The solid lines through the data points are the best fit obtained by fitting the data globally in KinTek Explorer using the model in Scheme 2. The best-fit values derived from the fitting are kpol: 181 ± 25 s⁻¹ and Kd,app: 28 ± 8 μM. B, pre-steady-state burst kinetics for T7 DNA polymerase E514Cou. T7 DNA polymerase E514Cou (75 nM), thioredoxin (1.5 μM), BSA (0.1 mg/ml), and FAM-27/45-18T DNA (150 nM) were mixed with Mg2+ (12.5 mM) and varying concentrations of dATP (2.5–110 μM) in the quench-flow to start the reaction and then with EDTA to stop the reaction at various times. Data for each dATP concentration are shown in different colors. The solid lines are the best global fit of the data using KinTek Explorer with the model in Scheme 2 to derive the parameters kpol: 202 ± 14 s⁻¹ and Kd,app: 15 ± 3 μM.

\[
\text{ED}_{dd} + \text{dNTP} \overset{k_1}{\underset{k_{-1}}{\rightleftharpoons}} \text{ED}_{dd}N \overset{k_2}{\underset{k_{-2}}{\rightleftharpoons}} \text{FD}_{dd}N
\]

Scheme 2. Kinetic pathway for nucleotide binding.

reaction for consistency with the correct incorporation reactions. T7 DNA polymerase (225 nM), thioredoxin (4.5 μM), BSA (0.1 mg/ml), and FAM 27/45-18A DNA (75 nM, with an A as the templating base) were mixed with a solution of Mg2⁺ (12.5 mM) and Mg2⁺ (12.5 mM) to start the reaction. Reactions were quenched by adding EDTA, and products were analyzed by electrophoresis as described above for correct incorporation. The time course of misincorporation at varying nucleotide concentrations for both WT and E514Cou T7 DNA polymerase are shown in Fig. 10. The reaction progress curves are biphasic, even though the experiment was set up with a large excess of enzyme over DNA. This can be explained with the minimal model in Scheme 3, where chemistry (k₂) is reversible and pyrophosphate release is slow (k₃). Traditional methods of fitting data using equations for biphasic reaction progress curves are highly error prone and difficult to interpret. In contrast, this type of reaction scheme can be easily modeled in KinTek Explorer including statistical analysis to determine how well each parameter is constrained. A good fit to the data was obtained for the model in Scheme 3; however, the individual rate constants were not well-defined. Rather, \( k_{pol}/K_{d,app} \equiv k_jk_jk_j(k_jk_j+k_j(k_j+k_j)) \) was obtained from fitting the data. Values of \( k_{pol}/K_{d,app} \) of 3.90 ± 0.12 M⁻¹ s⁻¹ and 21.0 ± 0.7 M⁻¹ s⁻¹ from the confidence contour analysis were obtained for the WT and E514Cou variant, respectively, at a \( \chi^2 \) threshold of 0.96. This result is sufficient to define the fidelity of both the WT and E514Cou T7 DNA polymerase variant when combined with the values extracted from the analysis of the pre-steady-state burst experiments for the correct nucleotide.

Fidelity of the WT enzyme versus the E514Cou variant

The parameter \( k_{pol}/K_{d,app} \) is equivalent to \( k_{cat}/K_m \) the specificity constant. The ratio of this parameter for the correct
Unnatural fluorescent amino acid to monitor polymerase dynamics

Figure 10. Misincorporation kinetics for WT and E514Cou T7 DNA polymerase at 4 °C. A, time course of T7 DNA polymerase misincorporation. T7 DNA polymerase (225 nM), thioredoxin (4.5 μM), BSA (0.1 mg/ml), and FAM-27/45-18A DNA (75 nM) were mixed with Mg²⁺ (12.5 mM) and varying concentrations of dATP (0.25–3 mM), and the reaction was quenched at various time points with EDTA. Data for each dATP concentration are shown in different colors, and the solid lines through the data show the results of global fitting using KinTek Explorer with the model in Scheme 3. For this experiment, kpol and Kd,app were not individually determined because Kd,app is too high to measure; however, the specificity constant, kpol/Kd,app, was accurately determined (3.90 ± 0.12 M⁻¹ s⁻¹) from the fit by simulation. B, time course of T7 DNA polymerase E514Cou misincorporation. T7 DNA polymerase E514Cou (225 nM), thioredoxin (4.5 μM), BSA (0.1 mg/ml), and FAM-27/45-18A DNA (75 nM) were mixed with Mg²⁺ (12.5 mM) and varying concentrations of dATP (0.25–3 mM). Data for each dATP concentration are shown in different colors, and the solid lines through the data show the results of global data fitting by simulation using KinTek Explorer with the model in Scheme 3. The parameter kpol/Kd,app was determined to be 21.0 ± 0.7 M⁻¹ s⁻¹, as in (A).

Scheme 3. Kinetic pathway for nucleotide incorporation with slow PPI release.

Discussion

In the spirit of Arthur Kornberg’s Fourth Commandment (32), “Do not waste clean thinking on dirty enzymes,” we began by optimizing the expression and purification of the WT enzyme without any tags. Various strategies such as low growth temperature and growth media additives were tested (not shown) to increase the yield of soluble protein. However, we found that coexpression of thioredoxin resulted in a very high level of soluble overexpression of WT enzyme without additional expression strategies.

Alternative failed strategies

In preliminary studies not reported here, inspection of the T7 DNA polymerase crystal structure suggested that most of the cysteines on the protein were not surface-exposed, so we tested whether we could achieve site-specific labeling of the T7 DNA polymerase with MDCC with fewer mutations. After multiple rounds of site-directed mutagenesis, expression, purification, MDCC labeling, and mass spectrometry, we realized that we were chasing a moving target. One cysteine in the WT protein was labeled, but after mutagenesis of this residue, other cysteines became labeled, and the cycle continued until we reached the eight cysteine mutations previously reported (18).

To circumvent the problems associated with cysteine-maleimide conjugation for protein labeling, we progressed to the amber suppression system to incorporate unnatural amino acids. We first tested amber suppression efficiency with p-acetyl-phenylalanine. This unnatural amino acid has the potential for site-specific labeling of the acetyl group through an oxime ligation reaction (33). Incorporation of p-acetyl-phenylalanine was successful and had high efficiency; however, difficulties arose when attempting the labeling reaction, despite numerous optimization efforts (labeling time, catalysts (34, 35), temperature, etc.).

Labeling with 7-HCou

Based on prior failures with other methods, we opted to directly incorporate 7-HCou into the enzyme, eliminating the labeling step required with the other systems. Others have used this unnatural amino acid to measure folding of myoglobin (22), detect antibody-antigen interactions (36), improve CFP with a eCFP(Cou) variant with a large stokes shift (37), and measure NaK channel preference for K⁺ by fluorescence lifetime and anisotropy analysis (38), among other uses (39–41).

Our initial attempts to incorporate 7-HCou resulted in very
low yield of labeled protein, a common issue with some of these unnatural amino acid systems. Various strategies to optimize low-yield systems have worked in a large number of these systems; however, the data were unclear as to how to optimize this particular evolved AARS/tRNA pair. Adding an inducible copy of CouRS to complement a constitutive copy on the plasmid has proven effective in multiple instances (28, 30), and we found that this strategy greatly increased the amber suppression efficiency for CouRS. The tRNA used also has shown to have a major impact on amber suppression efficiency and in multiple unnatural amino acid systems (23, 25). We found that a single copy of optimized tRNA outperformed six copies of the original tRNA (23) and eliminated issues of plasmid stability and toxicity that were caused by high expression of foreign tRNA. Finally, the condensed culture strategy (24) greatly enhanced the yield of wet cell pellet and greatly reduced the amount of 7-HCou used per expression, with minimal effects on protein expression level.

It should also be noted that we used heat to induce expression of proteins, but this strategy causes many proteins to form inclusion bodies. However, this can be circumvented with our expression plasmid by using naldixic acid to induce expression of RecA and allow expression from the λ promoter at lower temperatures. Future work could also optimize the sequence surrounding the amber codon, which in some cases has been shown to have a large effect on amber suppression efficiency (42). Also, future studies could try using a different amber-less E. coli strain and/or a strain with an altered RF1 (43). Our yield for T7 DNA polymerase E514Cou was only ∼10–20% of WT yield; however, this was still sufficient for our extensive kinetic studies because we had first optimized soluble overexpression of the WT enzyme. The reduced expression level of the 7-HCou variant may limit studies with this system on enzymes where the expression level of WT enzyme is low.

We then showed that our T7 DNA polymerase E514Cou variant was sensitive to the nucleotide-induced conformational change through equilibrium fluorescence emission scans and stopped-flow experiments. A roughly 20% change in fluorescence was observed upon nucleotide binding, which was sufficient to measure fast reactions in the stopped-flow instrument. Because the enzyme provides a signal to measure conformational changes in the stopped flow, the other important question is whether the fidelity of the enzyme was affected by 7-HCou incorporation. Although the MDCC-labeled 8-Cys-lite T7 DNA polymerase variant provided valuable insights into the issue of conformational dynamics of DNA replication, the extensive mutagenesis required to site-specifically label the enzyme made the enzyme more difficult to work with (reduced solubility and stability) and reduced its fidelity (4). Whereas $k_{pol}$ stayed the same, the fidelity was lowered by at least one order of magnitude ($3 \times 10^{5} k_{pol} = 230 \text{ s}^{-1}$) (4). The conformational dynamics of slower, lower-fidelity DNA polymerases have been characterized (12), so our goal was to study truly high-fidelity replication on a more robust system. We therefore performed the important pre-steady-state kinetic experiments for both the correct incorporation and misincorporation to determine whether the fidelity of the enzyme was significantly affected by insertion of 7-HCou and found that the effect of 7-HCou incorporation on fidelity for our T7 DNA polymerase variant was minimal. Overall, the fidelity of the variant is within 3-fold of the WT enzyme, which is the highest fidelity labeled DNA polymerase that we are aware of to date. This work lays the foundation for extensive kinetic analysis to establish the role of enzyme conformational dynamics in specificity and error correction by high-fidelity DNA polymerases and other enzymes. Moreover, our optimized labeling strategies should reduce the effort needed to pursue fluorescent labeling of other enzymes to examine the role of conformational dynamics in enzyme catalysis and specificity.

### Table 1

| Enzyme | Nucleotide | $k_{pol}$ (s$^{-1}$) | $k_{dapp}$ (µM) | $k_{pol}/k_{dapp}$ (M$^{-1}$ s$^{-1}$) | Fidelity |
|--------|------------|----------------------|-----------------|--------------------------------------|----------|
| WT     | Correct (dATP:T) | 181 (154, 221) | 28.1 (20.4, 41.0) | 6.4 × 10^{5} (3.9, 10.8) | 0.61 × 10^{-6} (1.03, 0.35) |
|        | Mismatch (dATP:A) | – | – | 3.90 (3.78, 4.02) | – |
| E514Cou | Correct (dATP:T) | 202 (188, 225) | 15.0 (12.8, 18.2) | 13.5 × 10^{5} (10.3, 17.6) | 1.56 × 10^{-6} (2.10, 1.56) |
|        | Mismatch (dATP:A) | – | – | 21.0 (20.3, 21.7) | – |

Parameter (best-fit [lower boundary, upper boundary]).

### Experimental procedures

#### Preparation of reagents

The fluorescent unnatural amino acid 7-HCou was purchased from Sigma-Aldrich and was prepared as a 10 mM stock in H$_2$O and titrated with NaOH to roughly pH 9.5. The stock solution was aliquoted and stored at −80 °C. Lysozyme, PEI, and PMSF were from Sigma-Aldrich. All chromatography resin/pre-packed columns and an AKTA Pure HPLC with multiwavelength detection were from GE Healthcare. Unless otherwise noted, all buffer components, E. coli growth media, and antibiotics were purchased from either Sigma-Aldrich or Thermo Fisher Scientific. Phusion DNA polymerase, T4 ligase, T4 polynucleotide kinase, BSA, dNTPs, restriction enzymes, and T4 DNA polymerases were purchased from New England Biolabs. SDS-PAGE was performed by the method of Laemmli (44). Plasmid maps were constructed with SnapGene (GSL Biotech, RRID:SCR_015052). All plasmid maps are included in the supporting data files accompanying this article. Protein structures were prepared with the PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC). Sanger sequencing was performed by the University of Texas at Austin’s sequencing facility in the Center for Biomedical Research Support at all cloning steps to ensure mutations were not introduced during PCR.

#### Cloning and mutagenesis

All experiments in this paper using T7 GP5 refer to an exonuclease-deficient variant (D5A, E7A) to allow measurements of the polymerization reaction without complications from the
Table 2

Cloning/mutagenesis DNA oligos

| Oligo No. | Description | Sequence (5′ → 3′) | Method |
|-----------|-------------|--------------------|--------|
| 1 | pCI(T7x-, bicis.) E514TAG F | GCTACCTACCGCGATAACGCG | IPCR |
| 2 | pCI(T7x-, bicis.) E514TAG R | TAAGCAGCATATCCTGGTTCTAGT | IPCR |
| 3 | MjTyrRS Y32E F | GATGAAAAATCTGCTGAGATAGGTTTTGAACCAAGTGG | QC |
| 4 | MjTyrRS Y32E R | CAATCTTTGATATATTATATCAAATCCAGCATTTTGTAAATC | QC |
| 5 | MjTyrRS D286R (29) mutant of CouRS was made with oligos 11 and mutations in MjYRS was designated pSUP-(CouRS-6TRN). | ACCAAATGGTATATTTATATCAATGCGCATTGGGTTAATC | IPCR |
| 6 | MjTyrRS D158G/L162G R | TAATGAATGGGATTAACCTGCATTATTGGATAGATACTTCAG | IPCR |
| 7 | MjTyrRS D158G/L162G F | GATTTAGGCGCCTATTTAAACCAGAAAGG | IPCR |
| 8 | MjTyrRS F108Y/Q109H R | ATATTCACTTCCATAAAAAATTTTTGCGGACTTATTTGCAATTTTGC | IPCR |
| 9 | MjTyrRS F108Y/Q109H F | GATGAAAAATCTGCTGAGATAGGTTTTGAACCAAGTGG | QC |
| 10 | MjTyrRS D158G/L162G | ATATTCACTTCCATAAAAAATTTTTGCGGACTTATTTTGAATATTTTGA | QC |
| 11 | pCI(T7x-, bicis.) E514TAG R | TTACGCATACTAGTGTGCAATGTAACATCAGAGATTTTG | IPCR |
| 12 | pCI(T7x-, bicis.) E514TAG F | TTCGCTGTGCATATGGACGAATTTGAAATGATAAAG | IPCR |
| 13 | pcI EcoRI F | GGATTTAGGCGCCTATTTAAACCAGAAAGG | IPCR |
| 14 | pcI EcoRI R | ATTTCACTTCCATAAAAAATTTTTGCGGACTTATTTTGAATATTTTGA | IPCR |
| 15 | CouRS NdeI F | TTCGCTGTGCATATGGACGAATTTGAAATGATAAAG | IPCR |
| 16 | CouRS NdeI R | ATTTCACTTCCATAAAAAATTTTTGCGGACTTATTTTGAATATTTTGA | IPCR |
| 17 | pSUFCouRS BamHI F | TTTACTAGCTGGATCCCAGGAAACAGCTATGACCATGATTAC | IPCR |
| 18 | pSUFCouRS BamHI R | TTCGCTGTGCATATGGACGAATTTGAAATGATAAAG | IPCR |
| 19 | pCI EcoRI F | TTTACTAGCTGGATCCCAGGAAACAGCTATGACCATGATTAC | IPCR |
| 20 | pCI EcoRI R | TTCGCTGTGCATATGGACGAATTTGAAATGATAAAG | IPCR |
| 21 | pcI NdeI F | TTACGCATACTAGTGTGCAATGTAACATCAGAGATTTTG | IPCR |
| 22 | pcI NdeI R | TTCGCTGTGCATATGGACGAATTTGAAATGATAAAG | IPCR |
| 23 | pcI SpeI F | TTACGCATACTAGTGTGCAATGTAACATCAGAGATTTTG | IPCR |
| 24 | pcI SpeI R | TTCGCTGTGCATATGGACGAATTTGAAATGATAAAG | IPCR |

Table 3

M. jannaschii amber suppressor tRNA variants

| tRNA | Sequence (5′ → 3′) | Method |
|------|--------------------|--------|
| tRNA_Mj | CCGCCGCUAGUUCACAGCCGCAAGACCGGGCCCCAUCAAAAACGCGCCCGCGGCA | | |
| tRNA_opt | CCGCCGCUAGUUCACAGCCGCAAGACCGGGCCCCAUCAAAAACGCGCCCGCGGCA | | |

Sequences in green are the bases of the anticodon loop. Bases in red are mutations in the optimized tRNA.

Expression tests to monitor amber suppression efficiency

Starter cultures of LB (1% (w/v) Bacto-Tryptone, 0.5% (w/v) yeast extract, and 1% (w/v) NaCl) plus 100 µg/ml ampicillin and 30 µg/ml chloramphenicol were inoculated with BL21 E. coli (flhuA2 [lon] ompT gal [dcm] ΔhsdS), harboring both pcI5-
Unnatural fluorescent amino acid to monitor polymerase dynamics

(T7 G5x- E514TAG, trxA) and various pSUP plasmids, and grown overnight at 30 °C. Starter cultures were inoculated into expression cultures of sterile-filtered supplemented M9 media (48 mM Na2HPO4, 22 mM KH2PO4, 8.6 mM NaCl, 19 mM NH4Cl, 0.126 mg/ml arginine, 0.02 mg/ml glutamine, 10 μM FeSO4, 5 mM MgSO4, 0.2 mM CaCl2, 0.1 mM thiamine–HCl, 1% (v/v) glycerol, and 0.5% (w/v) casamino acids, pH 7.4) with ampicillin and chloramphenicol and incubated at 30 °C with shaking at 250 rpm until the A600 reached 0.5. A “Pre-Induction” sample was taken by pelleting a volume of cells equal to 1 ml at an A600 of 3, aspirating the supernatant, and storing the pellet at −80 °C. Expression cultures were then moved to a water bath set to 42 °C for 15 min to induce expression. Cultures were incubated at 38 °C for 3 h, and then a 1-ml volume of cells at an A600 of 3 was harvested as the Post-Induction sample and stored at −80 °C. Samples were processed for separation by SDS-PAGE as previously described (21). Following electrophoresis, gels were stained with Coomassie Brilliant Blue.

Large-scale expression of T7 DNA polymerase WT and E514Cou

For WT T7 DNA polymerase expression, C2984H E. coli (F' proA' B' lacIq ΔlacZM15/fluA2 Δ(lac-proAB) gltV' galK16 galE15 R(2gb-210::Tn10)TetS endA1 thi-1 Δ(hsdS-mcrB8), harboring pclA5-(T7 G5x- , trxA), were inoculated into a starter culture of LB plus 100 μg/ml ampicillin and incubated overnight with shaking at 30 °C. The following overnight culture was inoculated into 1 liter of Terrific Broth (2.4% (w/v) yeast extract, 2% (w/v) trypptone, 0.4% (v/v) glycerol, 17 mM KH2PO4 and 72 mM K2HPO4) plus 100 μg/ml ampicillin and incubated with shaking at 30 °C, 250 rpm until the A600 reached 1.5. The culture was then incubated at 42 °C for 30 min with shaking to induce expression, followed by incubation at 38 °C for 3 h with shaking. The cells were harvested by centrifugation at 6,000 × g for 15 min at 4 °C and stored at −80 °C until purification. For expression of T7 DNA polymerase E514Cou, BL21 E. coli harboring pSUP-(2×CouRS–tRNAopt) and pclA5-(T7 G5x- E514TAG, trxA) were inoculated into an overnight starter culture of LB plus 100 μg/ml ampicillin and 25 μg/ml chloramphenicol and incubated with shaking at 30 °C, 250 rpm.

The following morning, 20 ml of overnight culture was inoculated into each of 6 × 1-liter of sterile-filtered supplemented M9 media plus 100 μg/ml ampicillin plus 25 μg/ml chloramphenicol in 2.8-liter baffled bottomed flasks. The cultures were incubated at 30 °C with shaking at 200 rpm. When the A600 of the cultures reached 0.5, the cultures were pelleted at 3,500 × g for 12 min at 4 °C. The cells were then resuspended in a total of 600 ml of sterile filtered supplemented M9 plus 1 mM 7-HCou, 100 μg/ml ampicillin, and 25 μg/ml chloramphenicol in 2.8-liter baffled bottomed flask. The condensed culture was incubated for 30 min with shaking at 30 °C. The incubator was then set to 42 °C with shaking for 30 min to induce expression. The culture was then incubated at 38 °C with shaking for 3 h, and the cells were harvested as described for the WT enzyme. The condensed 600-ml culture generally yielded a pellet of 6–9 g.

Purification of T7 DNA polymerase

WT T7 DNA polymerase and T7 DNA polymerase E514Coul were purified using the same protocol, and all steps were performed at 4 °C unless otherwise noted. The frozen pellet was thawed and resuspended in lysis buffer (50 mM Tris–HCl, pH 8, 150 mM NaCl, 2.5 mM EDTA, 0.1 mM DTT, and 1 mM β-mercaptoethanol) at 5 ml of buffer/gram of E. coli, and PMSF was added to a final concentration of 10 mM. Lysozyme was added to 0.3 mg/ml, and the mixture was stirred for 15 min at room temperature. The lysate was sonicated on ice for 20 min with constant stirring, followed by addition of sodium deoxycholate (Sigma-Aldrich) to 0.1% (w/v) and NaCl to 500 mM. The lysate was stirred for 30 min, and then cellular debris was pelleted by ultracentrifugation at 100,000 × g for 30 min. The clarified lysate was diluted with dilution buffer (50 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, and 5 mM DTT) to bring the concentration of NaCl to 100 mM. A solution of 10% (v/v) PEI, pH 8, was added to a final concentration of 0.1% (v/v), then stirred for another 15 min. The precipitated protein was collected by centrifugation at 4,000 × g for 15 min, and then protein was extracted and re-extracted in a small volume of PEI extraction buffer (dilution buffer plus 0.5 M NaCl) and the supernatants were pooled. Ammonium sulfate was slowly added to 35% saturation, the mixture was stirred for 30 min, and the sample was then centrifuged at 24,000 × g for 20 min. Ammonium sulfate was slowly added to the supernatant to 70% saturation and stirred for 30 min, and then the precipitated protein was collected as above. The pellet was dissolved in a small volume of resuspension buffer (dilution buffer plus 10% (v/v) glycerol) and dialyzed against two changes of the same buffer using 50,000-Da molecular weight cutoff dialysis tubing (Spectra-Por) for 1 h each. Resuspension buffer was added to bring the conductivity of the sample down to the conductivity of buffer A (resuspension buffer plus 100 mM NaCl). The sample was filtered through a 0.2-μm syringe filter and loaded on a 50-ml Q Sepharose Fast Flow column, monitoring absorbance at both 280 nm and 335 nm. The column was washed with 10% buffer B (buffer A with 400 mM NaCl), then eluted with a linear gradient of 10–70% buffer B over 20 column volumes. Fractions containing T7 DNA polymerase were pooled and loaded on a 5-ml HiTrap heparin column equilibrated in buffer A. The column was washed with 20% buffer C (buffer A with 1 M NaCl), and then the protein was eluted with a linear gradient of 20–70% buffer C over 20 column volumes. Fractions containing T7 DNA polymerase were pooled, diluted 6-fold with resuspension buffer, and loaded on a 5-ml HiTrap SP Sepharose Fast Flow column equilibrated in buffer A. The column was washed with buffer A, and then the protein was eluted with a linear gradient of 0–40% buffer C over 20 column volumes. Fractions containing T7 DNA polymerase were pooled and concentrated using an Ultracel 30,000 centrifugal concentrator (Amicon). The protein was dialyzed into storage buffer (40 mM Tris-HCl, pH 7.5, 50 mM NaCl, 0.1 mM EDTA, 50% (v/v) glycerol, and 1 mM DTT) using 1,000-Da molecular weight cutoff dialysis tubing (Spectra-Por). The concentration of T7 DNA polymerase (T7 GP5 plus thioredoxin) was determined by absorbance at 280 nm using the extinction coefficient ε280: 150,000 M−1 cm−1.
calculated based on the amino acid sequence of each polypeptide. The protein was aliquoted, flash-frozen in liquid nitrogen, and stored at \(-80^\circ\text{C}\). The final purity of the purified enzyme was \(>99%\) as estimated by SDS-PAGE. Purification from 1 liter of cells for the WT enzyme yielded 50–100 mg of purified protein, whereas purification from the 0.6-liter condensed culture for the unnatural amino acid yielded between 5 and 10 mg of purified protein.

**Expression and purification of thioredoxin**

Thioredoxin was expressed and purified as described previously (47) with two modifications. A heat denaturation step for 10 min at 63 °C before the DEAE column was used and a Superdex 200 column (GE Healthcare) was used for gel filtration. Concentration of purified thioredoxin was determined by absorbance at 280 nm, using an extinction coefficient of 13,980 M\(^{-1}\) cm\(^{-1}\). Purity was determined to be greater than 99% as estimated by SDS-PAGE. Roughly 35 mg of purified thioredoxin was obtained per liter of cells.

**MS**

MS analyses including protein digest, LC–MS/MS, and database searching were conducted by the University of Texas Austin Center for Biomedical Research Support’s Biological MS facility. Analysis of MDCC-labeled Cys mutants was performed as described by Tsai et al. (18). For WT and E514Cou T7 DNA polymerase, purified protein was separated by SDS-PAGE and the band for T7 gene product 5 was excised, alkylated with iodoacetamide, and digested with trypsin. The extracted peptides were analyzed by LC–MS/MS on a Dionex Ultimate 3000 RSLChano Nanoflow UPLC connected to an Orbitrap Fusion instrument (Thermo Fisher Scientific) with a 60-min gradient. Proteome Discoverer v2.4 with Sequest HT (Thermo Fisher Scientific) was used for data processing by searching against a database consisting of 12 common keratin contaminants and T7 gene product 5. Two missed cleavages by trypsin were allowed. Cysteine carbamidomethylation was used as a fixed modification, and variable methionine oxidation and Glu → 7-HCou (+C\(_2\)H\(_4\)O) modifications were used. The mass tolerance for precursor ions was 10 ppm (monoisotopic), and the mass tolerance for fragment ions was 0.8 Da (monoisotopic).

Data analysis and visualization were performed with Scaffold v4.8 (RRID:SCR_019109), which was also used for preparing figures. The peptide threshold used was 99.9%, resulting in 1,723 total spectra for the WT enzyme and 1,762 spectra for the E514Cou variant. Sequence coverage for both samples was 90%. The peptide \(^{509}\)NQIAAXLPTDRNAK \(^{522}\) (where X is either glutamic acid of 7-HCou at position 514) ionized well and had a high abundance for both the peptide with glutamic acid in the WT sample and the peptide with 7-HCou for the labeled protein sample. To test for 7-HCou incorporation at other positions in T7 DNA polymerase, purified protein was digested in solution with TPCK-trypsin (Promega) and the peptides were separated by reverse phase HPLC on a C18 column with acetonitrile/TFA mobile phase (not shown). The two peaks with an absorbance at 335 nm were analyzed by MALDI-MS in reflector mode with \(\alpha\)-cyano-4-hydroxy-cinnamic acid as a matrix. Data analysis was performed with the program mmass (48) using the same search parameters as above, with a 0.1-Da error tolerance.

**Preparation of oligonucleotide substrates**

All synthetic oligonucleotides were synthesized by Integrated DNA Technologies with standard desalting, and oligos used for kinetics assays were further purified in-house by denaturing PAGE and had a final purity of \(>99%\) full-length oligo. Concentrations of purified oligos for kinetics assays were determined by absorbance at 260 nm, using the extinction coefficients given in Table 4. DsDNA substrates for kinetic assays were prepared by mixing the appropriate 27-nt primer with the 45-nt template oligos (Table 4) at a 1:1.05 ratio in annealing buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl, and 1 mM EDTA). Primer/template complexes were then heated to 95 °C, slowly cooled to room temperature over two hours, and stored at \(-20^\circ\text{C}\).

**Kinetic measurements**

All experiments were carried out in T7 reaction buffer (2) (40 mM Tris-HCl, pH 7.5, 50 mM NaCl, 12.5 mM MgCl\(_2\), 1 mM DTT, and 1 mM EDTA). Reactions with 2,3′ dideoxy-CMP–terminated 27-nt primer (27\(_{\text{dd}}\) in Table 4) annealed to a 45-nt template (45–18T in Table 4, where the T represents the templating base) designated DNA\(_{\text{dd}}\) allowed measurements of nucleotide binding without the chemistry step. In all experiments, a 20-fold molar excess of thioredoxin over T7 DNA polymerase was included. All reactions were performed with 12.5 mM Mg\(^{2+}\) during the reaction. Reactions were set up by preincubating enzyme with DNA in the absence of MgCl\(_2\) in one syringe and mixing with a solution of dNTP and MgCl\(_2\) from the other syringe to start the reaction. No difference in kinetic data was observed for the reactions where both syringes had Mg\(^{2+}\) versus one syringe with Mg\(^{2+}\) at 2× the final concentration mixed with a syringe containing no Mg\(^{2+}\). Concentrations of reaction components given in the text are concentrations after mixing. Rapid-quench experiments were performed on an RQF-3.
Unnatural fluorescent amino acid to monitor polymerase dynamics

rapid-quench Instrument (KinTek Corporation) with a circulating water bath set to 4 °C. All experiments were performed with 0.6 mM EDTA in the quench syringe (giving a final concentration of 0.2 mM) and reaction buffer without magnesium in the drive syringes. All stopped-flow experiments were performed on an SF-300 instrument with a circulating water bath set to 4 °C, a 150-watt xenon lamp as the light source, and a dead time of 1.3 ms. Stopped-flow data shown in the main text are averages of at least eight individual traces. Stopped-flow experiments were performed on at least three separate occasions to ensure reproducibility, and each replicate yielded similar rates. Fluorescence experiments containing T7 DNA polymerase E514Cou were performed with excitation at 295 nm and emission at 445 nm, observed with a 45-nm bandpass filter (Semrock). Misincorporation reactions were performed by hand-mixing a solution containing the enzyme-DNA complex with a solution containing dATP plus Mg2+ to start the reaction. Reactions were quenched by the addition of EDTA to a final concentration of 0.2 mM. Samples collected by hand-mixing and by rapid-quench were analyzed either by denaturing PAGE or by capillary electrophoresis. Both methods rely upon detecting the fluorescence of the FAM label to quantify the concentrations of products and reactants at each time point. A control pre-steady-state burst experiment with [32P]-labeled DNA gave the same rates and amplitudes as the 5′-[6-FAM]-labeled DNA (data not shown), so all subsequent experiments were performed with the FAM-DNA for convenience. Samples separated by PAGE were imaged on a Typhoon FLA 9500 Scanner (GE Healthcare) and analyzed with the program ImageQuant (GE Healthcare). Samples separated by capillary electrophoresis were diluted into HiDi formamide (Thermo Fisher Scientific) with a 28-nt Cy3 internal standard oligo added for sizing. Samples were then injected onto a 36-cm capillary array filled with POP-6™ polymer (Thermo Fisher Scientific) on a 3130xl Genetic Analyzer (Applied Biosystems). Samples analyzed by either method gave identical results, and the capillary electrophoresis method will be elaborated in a future paper.

Steady-state fluorescence excitation/emission scans

Fluorescence excitation and emission spectra of various T7 DNA polymerase E514Cou complexes were collected at room temperature (≈25 °C) on a Horiba Fluorolog3 fluorimeter. Final concentrations of reaction components are given in the main text. Excitation scans were performed by monitoring emission at 460 nm (10-nm bandpass) while varying the excitation wavelength from 250–425 nm (bandpass: 5 nm, integration time: 0.5 s). Emission spectra were collected by scanning emission wavelengths (integration time: 0.5 s, bandpass: 5 nm) with a fixed excitation wavelength of either 298 nm or 335 nm (bandpass: 8 nm). Excitation and emission spectra were analyzed with the program a|e v1.2 (FluorTools, RRID:SCR_019108). All spectra shown in the main text have been corrected by subtracting the spectrum for the buffer alone in the cuvette, collected under the same instrument settings and experimental conditions. Emission scans were replicated three times to ensure reproducibility.

Data analysis

Data fitting/analysis was performed with the simulation software KinTek Explorer (49, 50) v9. This software was also used in preparing figures for kinetic data. Stopped-flow traces were fit in the software using the afit function to either a single or double exponential equation. The equation for a single exponential used is \( y = A_0 + A_1 (1 - e^{-b_1 t}) \), where \( A_0 \) is the starting fluorescence, \( A_1 \) is the amplitude of the fluorescence change, \( b_1 \) is the decay rate, and \( t \) is time. The equation for a double exponential used is \( y = A_0 + A_1 (1 - e^{-b_1 t}) + A_2 (1 - e^{-b_2 t}) \), where \( A_0 \) is the starting fluorescence and \( A_1, A_2, b_1, \) and \( b_2 \) are the amplitudes and decay rates for the first and second phases, respectively. After global data fitting (Fig. 9 and Fig. 10), confidence contour analysis was performed for each dataset using the Fitspace (31) feature of KinTek Explorer. Parameter boundaries are reported in Table 1, using a \( \chi^2 \) cutoff of 0.96 in the Fitspace calculation, which was recommended by the software as a reasonable limit based on the number of parameters and number of data points used in the fitting.

Data availability

The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (51) partner repository with the dataset identifier PXD021676. All remaining data are contained within this article and in the supporting information.

Author contributions—T. L. D. and K. A. J. conceptualization; T. L. D. and K. A. J. formal analysis; T. L. D. investigation; T. L. D. methodology; T. L. D. writing—original draft; K. A. J. resources; K. A. J. software; K. A. J. supervision; K. A. J. funding acquisition; K. A. J. project administration; K. A. J. writing—review and editing.

Funding and additional information—This work was supported by NIGMS, National Institutes of Health, Grant 5R01GM114223 (to K. A. J.) and the Welch Foundation Grant F-1604 (to K. A. J.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—K. A. J. is president of KinTek Corporation, which provided the SF300x stopped-flow and RQF-3 rapid quench-flow instruments and KinTek Explorer software used in this study.

Abbreviations—The abbreviations used are: 7-HCou, (7-hydroxy-4-coumarin-yl)-ethylglycine; MDCC, 7-diethylamino-3-(((2-methylimidyl)ethyl)amino)carbonyl)coumarin; GP5, gene product 5; FAM, carboxyfluorescein; CmR, gene for chloramphenicol acetyltransferase enzyme, conferring chloramphenicol resistance; trxA, gene for E. coli thioredoxin proK, E. coli prolyl tRNA synthetase gene; glnS, mutant promoter for the E. coli glucose-tRNA ligase gene; tRNA<sup>M<sub>th</sub></sup>, M. jannaschii tyrosyl-tRNA with CUA anticodon loop; tRNA<sup>opt</sup>, tRNA<sup>M<sub>th</sub></sup> with three mutations in one of the stem loops that enhances amber suppression efficiency; AARS, aminoacyl tRNA synthetase; MiYRS, M. jannaschii tyrosyl-tRNA synthetase; CouRS, MiYRS mutant with specificity for 7-HCou.
References

1. Roettger, M. P., Bakhtina, M., and Tsai, M.-D. (2008) Mismatched and matched dNTP incorporation by DNA polymerase β proceed via analogous kinetic pathways. *Biochemistry* **47**, 9718–9727 CrossRef Medline

2. Patel, S. S., Wong, I., and Johnson, K. A. (1991) Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry* **30**, 511–525 CrossRef Medline

3. Wang, I., Patel, S. S., and Johnson, K. A. (1991) An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. *Biochemistry* **30**, 526–537 CrossRef Medline

4. Tsai, Y.-C., and Johnson, K. A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry* **45**, 9675–9687 CrossRef Medline

5. Doublé, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**, 251–258 CrossRef Medline

6. Donlin, M. J., Patel, S. S., and Johnson, K. A. (1991) Kinetic partitioning between the exonuclease and polymerase sites in DNA error correction. *Biochemistry* **30**, 538–546 CrossRef Medline

7. Showalter, A. K., and Tsai, M.-D. (2002) A reexamination of the nucleotide incorporation fidelity of DNA polymerases. *Biochemistry* **41**, 10571–10576 CrossRef Medline

8. Wang, H., Huang, N., Dangerfield, L., Hao, X., Zhang, Q., Hu, B., Brittain, S. M., Brinker, A., Caldwell, J., Bursulaya, B., Spraggon, G., Brock, A., Ryu, Y., Uno, T., Schultz, P. G., et al. (2008) In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* **130**, 9268–9281 CrossRef Medline

9. Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K., and Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.* **10**, 425–432 CrossRef Medline

10. Kati, W. M., Johnson, K. A., Jerva, L. F., and Anderson, K. S. (1992) Mechanism and fidelity of HIV reverse transcriptase. *J. Biol. Chem.* **267**, 25988–25997 CrossRef Medline

11. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) FitSpace: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal. Biochem.* **387**, 30–41 CrossRef Medline

12. Kornberg, A. (2000) Ten commandments: lessons from the enzymology of DNA replication. *J. Bacteriol.* **182**, 3613–3618 CrossRef Medline

13. Wang, L., Zhang, Z., Brock, A., and Schultz, P. G. (2003) Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 56–61 CrossRef Medline

14. Hahn, A., Reschke, S., Leimkühler, S., and Risse, T. (2014) Ketoamine coupling of p-acetylphenylalanine at neutral pH for site-directed spin labeling of human sulfite oxidase. *J. Phys. Chem. B* **118**, 7077–7084 CrossRef Medline

15. Wendeler, M., Grinberg, L., Wang, X., Dawson, P., and Baca, M. (2014) Enhanced catalysis of oxime-based bioconjugations by substituted anilines. *Bioconjug. Chem.* **25**, 93–101 CrossRef Medline

16. Mills, J. H., Lee, H. S., Liu, C. C., Wang, J., and Schultz, P. G. (2009) A genetically encoded direct sensor of antibody-antigen interactions. *ChemBioChem* **10**, 2162–2164 CrossRef Medline

17. Kuhn, S. M., Rubini, M., Müller, M. A., and Skerra, A. (2011) Biosynthesis of a fluorescent protein with extreme pseudo-stokes shift by introducing a genetically encoded non-natural amino acid outside the fluorophore. *J. Am. Chem. Soc.* **133**, 3708–3711 CrossRef Medline

18. Liu, S., Lv, P., Li, D., Guo, X., Zhang, B., Yu, M., Li, D., Xiong, Y., Zhang, L., and Tian, C. (2015) K+ preference at the NaK channel entrance revealed by fluorescence lifetime and anisotropy analysis of site-specifically incorporated (7-hydroxycoumarin-4-yl)ethylglycine. *Chem. Commun.* **51**, 15971–15974 CrossRef Medline

19. Chen, S., Fahmi, N. E., Wang, L., Bhattacharya, C., Benkovic, S. J., and Hecht, S. M. (2013) Detection of dihydrofolate reductase conformational change by FRET using two fluorescent amino acids. *J. Am. Chem. Soc.* **135**, 12924–12927 CrossRef Medline

20. Dean, S. F., Whalen, K. L., and Spies, M. A. (2015) Biosynthesis of a novel glutamate racemase containing a site-specific 7-hydroxycoumarin amino acid. Enzyme–ligand promiscuity revealed at the atomic level. *ACS Cent. Sci.* **1**, 364–373 CrossRef Medline
Unnatural fluorescent amino acid to monitor polymerase dynamics

41. Ugwumba, I. N., Ozawa, K., Xu, Z.-Q., Ely, F., Foo, J.-L., Herlt, A. J., Coppin, C., Brown, S., Taylor, M. C., Ollis, D. L., Mander, L. N., Schenk, G., Dixon, N. E., Otting, G., Oakeshott, J. G., et al. (2011) Improving a natural enzyme activity through incorporation of unnatural amino acids. J. Am. Chem. Soc. 133, 326–333 CrossRef Medline

42. Schwark, D. G., Schmitt, M. A., and Fisk, J. D. (2018) Dissecting the contribution of release factor interactions to amber stop codon reassignment efficiencies of the Methanocaldococcus jannaschii orthogonal pair. Genes 9, 546 CrossRef

43. Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., Mercer, J. A., Wang, H. H., Carr, P. A., Mosberg, J. A., Rohland, N., Schultz, P. G., Jacobson, J. M., Rinehart, J., Church, G. M., et al. (2013) Genomically recoded organisms expand biological functions. Science 342, 357–360 CrossRef Medline

44. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685 CrossRef Medline

45. Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. Genetics 120, 621–623 Medline

46. Zheng, L., Baumann, U., and Raymond, J.-L. (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Res. 32, e115 CrossRef Medline

47. Schwark, D. G., Schmitt, M. A., and Fisk, J. D. (2018) Dissecting the contribution of release factor interactions to amber stop codon reassignment efficiencies of the Methanocaldococcus jannaschii orthogonal pair. Genes 9, 546 CrossRef

48. Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., Mercer, J. A., Wang, H. H., Carr, P. A., Mosberg, J. A., Rohland, N., Schultz, P. G., Jacobson, J. M., Rinehart, J., Church, G. M., et al. (2013) Genomically recoded organisms expand biological functions. Science 342, 357–360 CrossRef Medline

49. Johnson, K. A. (2009) Fitting enzyme kinetic data with KinTek global kinetic explorer. Methods Enzymol. 467, 601–626 CrossRef Medline

50. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. Anal. Biochem. 387, 20–29 CrossRef Medline

51. Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q. W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 44, D447–D456 CrossRef Medline