BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

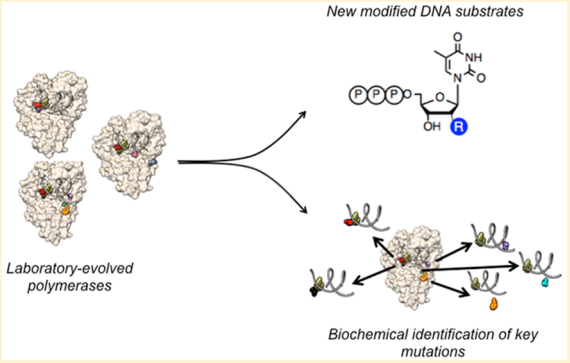
Taq DNA Polymerase Mutants and 2′-Modified Sugar Recognition

Hayley J. Schultz, Andrea M. Gochi, Hannah E. Chia, Alexie L. Ogonowsky, Sharon Chiang, Nedim Filipovic, Aurora G. Weiden, Emma E. Hadley, Sara E. Gabriel, and Aaron M. Leconte*

W. M. Keck Science Department, Claremont McKenna, Pitzer, and Scripps Colleges, Claremont, California 91711, United States

Supporting Information

ABSTRACT: Chemical modifications to DNA, such as 2′ modifications, are expected to increase the biotechnological utility of DNA; however, these modified forms of DNA are limited by their inability to be effectively synthesized by DNA polymerase enzymes. Previous efforts have identified mutant *Thermus aquaticus* DNA polymerase I (Taq) enzymes capable of recognizing 2′-modified DNA nucleotides. While these mutant enzymes recognize these modified nucleotides, they are not capable of synthesizing full length modified DNA; thus, further engineering is required for these enzymes. Here, we describe comparative biochemical studies that identify useful, but previously uncharacterized, properties of these enzymes; one enzyme, SFM19, is able to recognize a range of 2′-modified nucleotides much wider than that previously examined, including fluoro, azido, and amino modifications. To understand the molecular origins of these differences, we also identify specific amino acids and combinations of amino acids that contribute most to the previously evolved unnatural activity. Our data suggest that a negatively charged amino acid at 614 and mutation of the steric gate residue, E615, to glycine make up the optimal combination for modified oligonucleotide synthesis. These studies yield an improved understanding of the mutational origins of 2′-modified substrate recognition as well as identify SFM19 as the best candidate for further engineering, whether via rational design or directed evolution.



In addition to encoding all known life, DNA is a valuable biotechnological tool because of its ability to be easily amplified without a loss of information. DNA polymerases perform this enzymatic amplification with high efficiency and fidelity, allowing a number of foundational biological technologies, including the polymerase chain reaction (PCR)[1] and Sanger sequencing,[2] as well as a number of emerging biological technologies, such as high-throughput DNA sequencing,[3] DNA-encoded libraries,[4] and many DNA nanotechnologies.[5] While the remarkable biochemical properties of DNA polymerases permit these technologies, they also restrict them; DNA polymerases recognize very few modified substrates, limiting many emerging applications.[6] Thus, researchers have spent significant time and effort to broaden the substrate repertoire of these enzymes through directed evolution.[7]

Oligonucleotides possessing altered sugar structures, such as 2′ modifications, are often refractory to nuclease digestion, improving the utility of these nucleic acids in biological settings, both *in vivo* and *ex vivo*.[7−9] For instance, RNAi and other short, therapeutic oligonucleotides bearing 2′ modifications have been chemically synthesized and shown to possess improved properties relative to those of unmodified DNA or RNA. Unfortunately, there are no known native DNA polymerases that recognize these modified nucleotides, limiting the use of modified nucleotides in nucleic acids longer than those amenable to chemical synthesis in an array of applications, including, but not limited to, the evolution of modified DNA.[10] Considering the utility of these modified nucleotides, as well as the limited success of rational design of enzymes in this area, several research groups have used directed evolution to identify mutant DNA polymerases capable of using 2′-modified nucleotides.

To date, directed evolution efforts have largely focused on DNA polymerase I from *Thermus aquaticus* (Taq)[11−13] and, more recently, B family polymerases such as Therminator[14] and the replicative DNA polymerase from *Thermococcus gorgonarius* (Tgo).[15] Taq has been evolved to improve recognition of 2′-modified nucleotides on three different occasions;[11−13] in each case, a mutant enzyme with multiple mutations (Table 1) and significantly improved properties was identified. Each of these three enzymes (AA40, SFR3, and SFM19) is marked by a remarkable ability to use 2′-modified nucleotides; both AA40 and SFR3 were evolved to recognize ribonucleotides, and SFM19 was evolved to recognize 2′-methoxy-modified nucleotides. However, all three enzymes possess a common inability to synthesize 2′-modified DNA beyond approximately six to eight modified nucleotides. While these results are promising, to date, these Taq variants have not been further engineered or characterized beyond the initial reports.

Received: June 19, 2015
Revised: September 2, 2015
Published: September 3, 2015

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

Table 1. Mutations Observed in Parent Enzymes

amino acid	mutant[a]	additional observations[b]
A597	T (SFR3)	T (reverse transcription, 3′-modified substrates)
E602	V (AA40)	G (hydrophobic nucleobases)
W604	R (SFR3)	R (reverse transcription)
L605	Q (SFR3)	none
A608	V (AA40)	V (hydrophobic nucleobases, mismatch extension)
I614	T (SFR3), M (AA40), E (SFM19)	M (hydrophobic nucleobases, ancient DNA), F (unnatural base pair)
E615	G (SFR3, AA40, SFM19)	none

[a]Mutation identified in a parent enzyme; identity of the parent enzyme in parentheses. [b]Mutation identified in an evolution experiment for an activity other than 2′-modified nucleotide recognition; unnatural activity in parentheses.

More recently, several groups have evolved B family DNA polymerases to recognize 2′-modified nucleotides. Marx and co-workers identified a mutant Therminator capable of synthesizing RNA of up to 70 bp.[14] Most recently, Holliger and co-workers identified a Tgo mutant (TGK) that is capable of readily synthesizing RNA well beyond six modified nucleotides.[15] The Tgo mutant can also use a range of 2′-modified nucleotides, although the characterization of recognition of these alternative nucleotide structures was limited to a single, qualitative experiment, and the preliminary results indicate that only a small fraction of the modified primer was extended. Further, the fidelity of both Therminator and TGK is at least 1 order of magnitude worse than those of Taq and previously characterized Taq mutants. Although characterization and application of these B family enzymes have been limited to these single reports, these mutants are among the most promising enzymes to date for the synthesis of 2′-modified forms of DNA.

While evolved B family polymerase mutants such as Therminator and Tgo are highly promising, there is still good reason to pursue variants of Taq as a potential enzyme engineered to synthesize 2′-modified DNA. Taq, the proteolytic Klenow fragment of the highly homologous DNA polymerase I from *Escherichia coli* (KF), and the equivalent proteolytic fragment of Taq (Stoffel Fragment, SF) have been extensively biochemically[16−19] and structurally[20−22] characterized, which may allow future engineering efforts. Further, the wide-ranging use of Taq in biotechnological applications has resulted in significant development of the enzyme, producing known variants of Taq with a range of useful biochemical properties, including but not limited to cold sensitivity,[23] increased thermostability,[24] increased resistance to known inhibitors such as heparin,[24,25] human plasma,[26] and soil,[26] and the ability to use of an array of modified nucleobases.[27−30] Additionally, several mutations that increase the fidelity of KF[31] and SF[32] have been identified, which might be used to recover fidelity that might be lost during engineering. Thus, one could imagine further engineering of Taq might be able to draw on this rich biochemical and biotechnological knowledge to continue to engineer this enzyme for a number of downstream applications. Notably, this stands in contrast to work on B family DNA polymerases, which are significantly less well characterized and also have sometimes suffered from not being robust.[33]

To date, there has been no further characterization of the evolved Taq mutants beyond the initial studies, which largely focused on the process of their identification. It is clear that a practically useful mutant will require further engineering; comparative biochemical characterization, described here, can help guide that future engineering. Here, we describe the evaluation of these previously evolved mutants (here, termed "parents") for their ability to recognize a range of 2′-modified substrates; previously, only AA40 was examined for its ability to recognize modified nucleotides other than the selection targets. Our studies suggest that one enzyme in particular, SFM19, is the most promising for further engineering. Further, we examine the genetic origins of the novel function of these three enzymes, revealing the minimal mutations needed to allow modified substrate recognition. While each previously identified enzyme has multiple amino acid mutations (Table 1), it is not clear if all of them are necessary for unnatural function; the relationship between these recurrent mutations and the acquisition of unnatural activity has not been previously explored qualitatively or quantitatively despite their frequent appearance. Collectively, these studies provide foundational information for the engineering of Taq mutants capable of synthesizing 2′-modified DNA.

MATERIALS AND METHODS

Molecular Cloning. All constructs were built onto a SF-pET vector provided by the Romesberg group (The Scripps Research Institute, La Jolla, CA). All polymerase chain reactions (PCRs) used the following materials: Phusion DNA polymerase (Fisher), manufacturer-provided buffers (Fisher), dNTPs (New England Biolabs), and primer oligonucleotides (Integrated DNA Technologies) or synthetic genes (gBlocks, Integrated DNA Technologies). All PCR products and plasmids were purified using commercially available kits (Qiagen). All transformations were performed using chemically competent HB101 cells (Promega). All DNA sequences were verified using DNA sequencing (Eurofins MWG Operon). For further details, including oligonucleotides used and specific conditions used, see the Supporting Information.

Protein Expression and Purification. Expression and purification were performed using previously published methods.[26] See the Supporting Information for additional experimental details.

General Nucleotide Incorporation Assay. Enzymes were characterized using a six-nucleotide incorporation assay. For all figures, two co-authors independently prepared and assayed the enzymes, giving two completely independent replicates. A single representative gel and heat map is shown for each. In every assay, dATP was measured in parallel as a positive control. All reactions were run using the following conditions: 5′ IRDye700-labeled 40 nM primer (5′-dTAATACGACTCACTATAGGGAGA, TriLink Biotechnologies), 80 nM template (5′-dCGCTAGGACGGCATTGTTTTTTTCTCCCTATAGTGAGTCGTATTA, TriLink Biotechnologies), 50 mM Tris buffer (pH 8.5, Research Products International), 6.5 mM $MgCl_2$ (Sigma-Aldrich), 0.05 mg/mL Ac-BSA (Promega), 50 mM KCl (Sigma-Aldrich), variable quantities of nucleoside triphosphate, and variable quantities of enzyme. All nucleotides were obtained from commercial sources (dATP and rATP from Fisher Scientific and 2′FATP, 2′NH_3^+ATP, and 2′AzATP from TriLink Biotechnologies). For specific conditions used, see Table S2. All reaction mixtures were incubated at 50 °C for variable times and reactions quenched using 2 volume

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

equivalents of quenching buffer: 95% formamide (Acros), 12.5 mM EDTA (Sigma-Aldrich), and trace amounts of Orange G powder (<1 mg, Sigma-Aldrich).

Steady-State Kinetic Assay. The primer−template and buffer conditions were identical to those described in General Nucleotide Incorporation Assay. Steady-state rate constants were measured using previously described methods.[34] The data presented are the average of three replicates.

Modified Primer Incorporation Assay. Experiments with modified primers were identical to those described in General Nucleotide Incorporation Assay with the exception of the primer used. A primer bearing two terminal ribonucleotides [5′-dTAATACGACTCACTATAGGGAGA(rA) (rA)] was postsynthetically labeled with IRDye700DX through a 5′ amine modifier by TriLink Biotechnologies as this synthesis strategy was more compatible with RNA/DNA hybrid chemical synthesis.

Analysis and Reporting of Results. All assays were visualized on a 20% polyacrylamide gel consisting of 20% acrylamide (Research Products International), 1× TBE (Research Products International), and 7 M urea (Research Products International). Acrylamide was polymerized with 10% (w/v) ammonium persulfate (Fisher Scientific) and 1% TEMED (Fisher Scientific). Gels were imaged using Li-Cor Odyssey CLx. Results were quantified using ImageStudio (Li-Cor). To minimize the effect of user-defined manipulation of contrast, data are presented as heat maps generated using MatLab (Mathworks); see the note in the Supporting Information for further details. Gel images used to generate heat map are provided in the Supporting Information.

RESULTS

Comparative Characterization of AA40, SFR3, and SFM19. Initially, we cloned the three parent enzymes (SFR3, SFM19, and AA40) into a common construct (the catalytic Stoffel fragment, SF) and expressed and purified the enzymes using previously published methods.[27] Past studies examined these enzymes under relatively forcing conditions, including, typically, high enzyme concentrations and the use of manganese as a divalent metal, which is known to make DNA polymerases more promiscuous. These past studies revealed similar activity; typically, the evolved enzyme incorporated approximately six modified nucleotides, after which the parent enzyme is no longer able to continue modified DNA synthesis. We characterized the enzymes under more mild conditions (i.e., lower enzyme concentrations without manganese) to better observe biochemical differences between these enzymes. We initially assayed the ability of each of the three parent enzymes as well as the wild-type enzyme to incorporate multiple 2′-modified ATP nucleotides on a template encoding the addition of six A nucleotides. We incubated the enzyme with either rATP (Figure 1A) or OMeATP (Figure 1B) because each was a target substrate of the initial selections. In this experiment, as with all of our experiments, we also incubated the enzymes with dATP as a control (Figure S3).

Interestingly, we observe significant differences between the three enzymes under these conditions. Although it was originally evolved to recognize 2′OMeNTPs, SFM19 is the most active enzyme for rATP incorporation. Even at 5 min, the $n + 6$ product is the major product; both SFR3 and AA40 are not able to achieve this even after 60 min. SFR3 was the second most active with rATPs and synthesized a mixture of $n + 5$ and $n + 6$ products after 60 min. AA40 was the least active as it was unable to synthesize the $n + 6$ product, even after 60 min.

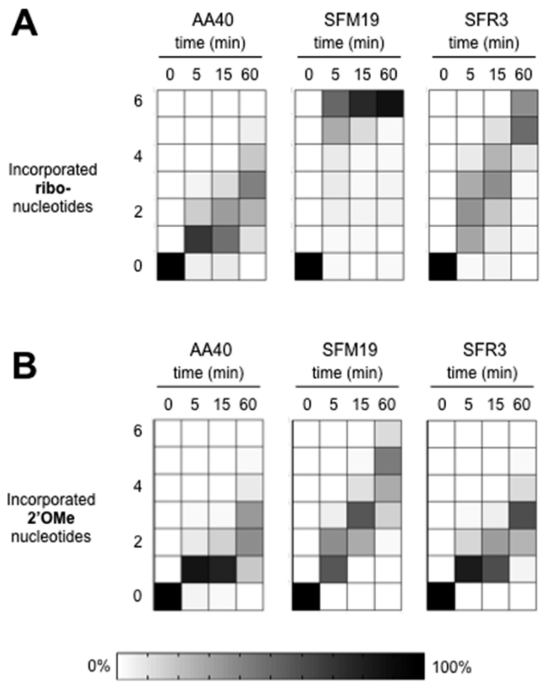


Figure 1. Recognition of the 2′-modified nucleotide by parent enzymes (AA40, SFM19, and SFR3). Enzymes were incubated with DNA encoding six A additions and 50 μM 2′-modified nucleotide substrate for varying amounts of time: (A) 2 nM enzyme with rATP substrate and (B) 50 nM enzyme with 2′OMeATP substrate. Under these conditions, wild-type SF did not show detectable activity. For corresponding gels, see Figures S1 and S2.

We also examined the ability of these enzymes to incorporate 2′OMeATP on the same template. The activity of all of the enzymes was substantially less than for rATP incorporation, so we used a higher enzyme concentration to better observe differences between the enzymes. Once again, SFM19 was the most active among the three enzymes; however, it is much less active on 2′OMeATP than it was on rATP, in spite of its initial selection for 2′OMeNTPs. Also mirroring the results with ribonucleotides, SFR3 is more active than AA40 but less active than SFM19. Thus, the general trends of activity are similar among the two substrates. Interestingly, for all three of the enzymes, the first time point shows a majority of the primer converted to at least $n + 1$ product, which implies that the limiting factor for recognition might be the use of the 2′OMe-modified primer as a substrate rather than the modified nucleoside triphosphate.

Considering that each enzyme recognizes both of the selection substrates (2′OMeNTPs and rNTPs) significantly better than the native enzyme, we wondered if the parents could also recognize other 2′-modified nucleotides. We tested the ability of the three parent enzymes to incorporate nucleotides containing fluoro, amino, and azido 2′ modifications (Figure 2). Interestingly, for all three modified nucleotides, SFM19 appears to be the most active enzyme and, thus, the best candidate enzyme for further engineering. While AA40 and SFR3 are able to incorporate only one or two amino-ATP nucleotides, SFM19 has a major product of $n + 3$ with some $n + 4$ product. SFM19 is also able to incorporate six azido-ATP nucleotides, while the second best, SFR3, is able to incorporate only four. Both SFM19 and SFR3 are able to add

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

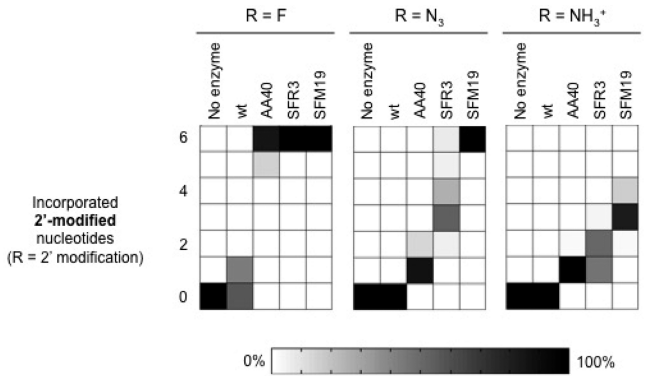


Figure 2. Recognition of novel 2′-modified nucleotides by parent enzymes (AA40, SFR3, and SFM19) and wild-type SF (wt). Each enzyme (5 nM) was incubated with DNA encoding six A additions and 50 μM 2′-modified nucleotides for 60 min. For corresponding gels, see Figures S4–S6.

six F-ATP nucleotides in the shortest time period observed, 15 min. Previously, SFM19 had only been characterized for its ability to recognize 2′OMeNTPs; here, we have shown that SFM19 can recognize a variety of 2′ modifications and is the most active of the parents for each 2′ modification.

Interestingly, each enzyme recognizes the different modified nucleotides with the same relative specificity; each parent recognizes the 2′-fluoro modification most effectively, followed by, in descending order, azido, hydroxyl, amino, and methoxy substitutions. Thus, it appears that the three parents may have a common mechanism for recognizing modified triphosphates and differences between substrates might be more dependent on the properties of the nucleotide modification than the enzyme. These data suggest that the current enzymes might be already able to synthesize 2′-fluoro-modified DNA without further mutation and highlight the value of examining mutant enzymes for their ability to recognize substrates other than their selection substrates. Future engineering efforts would benefit from concurrent synthetic optimization of the nucleotides and protein engineering of the polymerases.

Biochemical Origins of Differences in Activity. To better understand the biochemical origins of the activity of SFM19, we quantified the rates of both deoxyribonucleotide and ribonucleotide incorporation by wild-type SF and SFM19 (Table 2). The rate of addition of ribonucleotides and deoxyribonucleotides catalyzed by SFR3 and AA40 had both been measured previously, so we compared our data with

Table 2. Relative Steady-State Rate Constants of Incorporation of the Nucleotide onto an Unmodified Primer by DNA Polymerases[a]

enzyme	$rCTP_{mut}/dCTP_{wt}$[b]	$rCTP_{mut}/dCTP_{mut}$[c]	ref
wild type	–	<0.001	12
SFM19	0.02	0.2	this study
AA40	0.4	1.3	13
SFR3	0.4	1.9	12

[a]Ratios relative to wild type were used to account for differences in conditions and constructs used for measurements in the various studies. Rate constants are listed in Table S3. [b]Rate of rCTP incorporation by the mutant enzyme relative to the rate of dCTP incorporation by the wild-type enzyme measured under identical conditions. [c]Rate of rCTP incorporation relative to dCTP incorporation for the mutant enzyme under identical conditions.

SFM19 to the previously published data. Each study measured the rates of incorporation using somewhat different conditions; thus, to mitigate differences between assays, here, we compare the rate of incorporation of the modified nucleotide by the parent to the rate of incorporation of the unmodified nucleotide by the wild type measured under identical conditions. Interestingly, all three parents recognize the modified nucleoside triphosphate with a rate constant that is similar to the rate of addition of the unmodified nucleotide by the wild type. Further, all three parents catalyze the incorporation of a modified nucleotide approximately as well as addition of an unmodified nucleotide. Relative to SFR3 and AA40, SFM19 appears to be slightly less active. While there are subtle differences between the parent enzymes, it is possible that differing assay conditions may cause these slight differences. These data do not explain the large observed differences in activity among the parents.

To examine the role of a modified primer in the activity of the parent enzymes, we characterized the addition of unmodified nucleotides onto modified primers bearing two ribonucleotides on the primer terminus (Figure 3). Because the

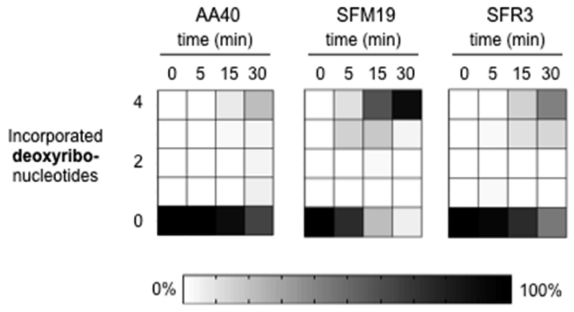


Figure 3. Differential ability to extend ribonucleotide-modified primers. Parent enzymes (1 nM) were incubated for variable times with unmodified dATP (50 μM) and DNA containing a modified primer bearing two ribonucleotides at the primer terminus. For corresponding gels, see Figure S8.

nucleoside triphosphates are unmodified, this assay independently probes the role of the modified primer in synthesis. This assay shows a trend that is highly similar to that in the initial qualitative assays that contain only ribonucleotides; SFM19 is most active, while AA40 is least active. These data suggest that modified primer recognition is most likely the cause of the different activities of these three parent enzymes.

E615G and Unnatural Function. To understand the mutational basis of these molecular differences, we further characterized these enzymes. We chose to focus our initial studies on ribonucleotides both because it was the selection substrate in two of the three original reports and because the higher activity on these substrates suggested that we might be able to better observe mutations that impart only minor activity. E615G is the only common mutation found in all three parents, so we began by characterizing the role of E615G in these evolved enzymes. E615 is well-known to be the classical steric gate residue, competing for the physical space that the 2′-modified nucleotide inhabits;[19] thus, it is unsurprising that it is mutated to glycine in the most active mutant enzyme in all three studies. We cloned the E615G mutant enzyme and compared it to the parent enzymes to understand how much of the unnatural activity of each parent can be purely attributed to this single, common mutation. Characterization of the E615G

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999–6008

mutant (Figure 4A) in comparison to the parents (Figure 1A) shows that E615G is responsible for a large portion of the unnatural activity observed in the parents, but it is not responsible for the entirety of the unnatural activity. The E615G mutant is slightly more active than AA40, as has been previously observed,[13] is slightly less active than SFR3, and is much less active than SFM19.

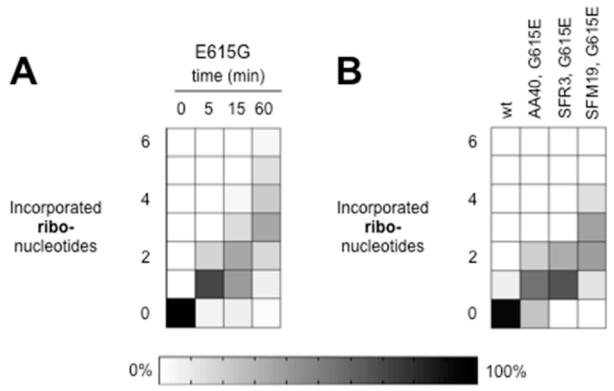


Figure 4. Role of E615G in 2′-modified substrate recognition. (A) E615G (2 nM) was incubated with DNA encoding six A additions and 50 μM 2′ rATP for varying times. (B) G615E reversion mutants (10 nM) were incubated with DNA encoding six A additions and 50 μM 2′ rATP for 60 min. For corresponding gels, see Figures S9 and S11.

To examine the combined role of the other mutations in the absence of the critical E615G mutation, we reverted the G615 amino acid back to the wild-type glutamate. Characterization of these mutant enzymes relative to the wild-type enzyme showed that, under identical conditions, SFR3/G615E, AA40/G615E, and SFM19/G615E all show activity significantly higher than that of the wild-type enzyme, although the conditions required enzyme concentrations higher than that of E615G (Figure 4B). Interestingly, the relative activities of the revertants are similar to the relative activities of the parent enzymes; SFM19/G615E is the most active revertant, and SFM19 is the most active parent. Collectively, both experiments with E615G and with the G615E revertants suggest that E615G is not the only determinant of unnatural function and is not absolutely required for recognition.

Other Mutations Important for Unnatural Function. To understand which individual mutations are responsible for the unnatural activity in the absence of E615G, we dissected each parent enzyme into its component mutations, cloning each additional individual mutation found in the prior studies onto the wild-type enzyme. We then measured the ability of each mutant to incorporate rATP relative to the wild type, the parent mutant, and the parent mutant containing the G615E reversion. We began by creating and characterizing the single-amino acid mutants identified in AA40 [E602V, A608V, I614M, and E615G (Figure 5A)]. The E602V and A608V mutations did not show detectable activity greater than that of the wild-type enzyme. The I614M mutant possesses detectable activity above that of the wild type and activity comparable to that of the AA40/G615E revertant. These data suggest that the I614M mutation contributes most to unnatural activity in the absence of E615G while the E602V and A608V mutations either do not contribute to catalysis or contribute only in the presence of a second mutation.

We also made all of the possible single mutants from SFR3 [A597T, W604R, L605Q, I614T, and E615G (Figure 5B)]. Similar to AA40, only the I614T mutant showed substantial activity above that of the wild-type enzyme. A597T, W604R, and L605Q did not show significant activity. I614T showed slightly less activity than the SFR3/G615E reversion mutant. From these data, we conclude that the I614T mutation contributes to unnatural activity in the absence of other mutations and that the other mutations either do not contribute to catalysis or require a second mutation to have an effect.

SFM19 is composed of only two mutations, I614E and E615G. The SFM19/G615E mutant is the I614E mutant characterized in Figure 3B. Notably, because the I614E mutant is more active than the AA40/G615E revertant or the SFR3/G615E revertant, each of which is comparable in activity to their respective mutations at I614, we can conclude that the I614E mutation increases the level of unnatural substrate recognition more than the I614M or I614T mutation. Collectively, these data show that mutations at I614 produce the most active mutants aside from E615G and further imply that other mutations either do not contribute or require the key mutations at position 614 or 615 to function.

Interactions between I614 Mutations and E615G. Although all three parent enzymes possess a mutation at amino acid 614 to methionine (M), threonine (T), or glutamate (E), a single mutation to any of these amino acids only imparted relatively modest unnatural activity; to observe the unnatural activity required a relatively high enzyme concentration and long time periods. These data imply that mutation at E615G is

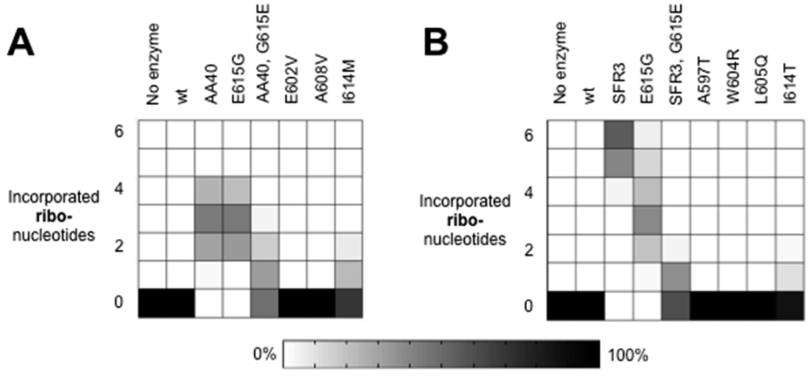


Figure 5. Ribonucleotide recognition by single-point mutants from (A) AA40 and (B) SFR3. Five nanomolar enzyme was incubated with DNA encoding six A additions and 50 μM 2′ rATP for 60 min. For corresponding gels, see Figures S13 and S15.

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

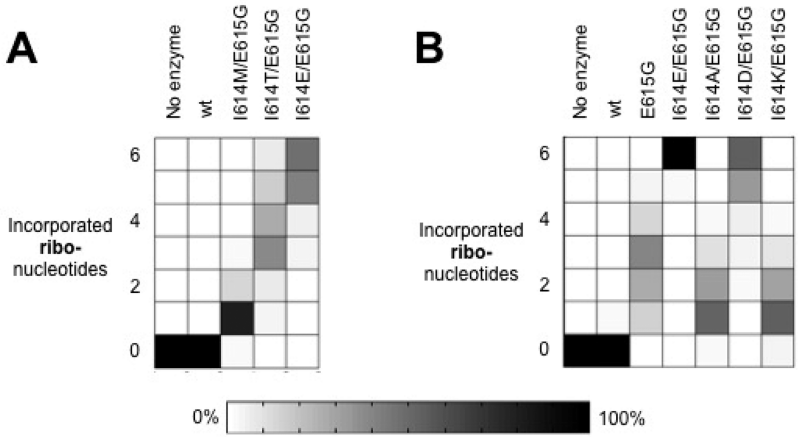


Figure 6. Synergistic interactions between I614 mutations and E615G. (A) I614X mutants identified from prior studies (I614E, -M, and -T) and (B) I614X mutants created for this study (I614A, -D, and -K). One nanomolar enzyme was incubated with DNA encoding six A additions and 50 μM 2′ rATP for 60 min. For corresponding gels, see Figures S17 and S19.

the primary driver of unnatural function; thus, we wondered if I614 mutations might have a more substantial effect in the presence of E615G. Notably, interactions between amino acids, particularly between amino acids close in physical space, can often lead to nonadditive effects. To test whether the I614M, -E, or -T mutation contributes more to RNA synthesis in the presence of E615G, we created and characterized all possible double mutants (i.e., I614M/E615G, I614E/E615G, and I614T/E615G). Interestingly, mutation at I614, when coupled to E615G, appears to have significantly variable and synergistic effects (Figure 6A). In particular, addition of the I614E mutation to the E615G background significantly increased 2′ substrate recognition. Additionally, although I614M and I614T had similar effects in isolation, they possessed differing effects in the context of E615G. The I614M/E615G double mutant was actually less active than the E615G mutant in isolation, which might explain the decreased activity of the AA40 mutant relative to that of E615G. Meanwhile, the I614T/E615G double mutant had activity comparable to that of E615G.

Considering the variable and nonadditive effects of mutations at I614 when combined with the E615G mutation, we wondered how other mutations at I614 would combine with mutation at E615G. Thus, we designed, cloned, and characterized three additional I614X/E615G double mutant enzymes with varying mutations at I614 (A, D, and K) to better understand this relationship. We chose these amino acids because of their chemical variance as well as the known substrate broadening properties of I614K.[34] As a control, we also constructed all of the I614X single mutants to observe the effects of these mutations in the absence of the E615G mutation. Each set of mutants was assayed for the ability to incorporate either dATP or rATP on a template encoding addition of six A's. As previously observed, mutation of I614 to K enhances the ability to incorporate modified ribonucleotides (Figure S21); however, this effect is weaker than that of other I614 mutants such as I614D and I614E. Interestingly, I614D was more active on ribonucleotide substrates than any other characterized mutation other than I614E, suggesting that negatively charged amino acids are particularly helpful at I614 for imparting unnatural function.

Interestingly, the effects upon addition of these mutations to E615G were highly variable (Figure 6B). I614K, which when added to the wild-type enzyme has been previously shown to increase the promiscuity toward 2′-modified substrates, actually decreases the activity of the enzyme when it is added to E615G. Notably, mutation of I614 to aspartate shows a substantial increase in activity when added to E615G; I614D/E615G was the only new mutant that generated measurable $n + 6$ products. These data are similar to those for addition of I614E to E615G and, collectively, imply that a negatively charged amino acid at position 614 generally facilitates 2′-modified nucleotide recognition, especially in the presence of E615G. Mutations to either T, A, M, or K had weak or slightly negative effects upon addition to E615G. These data suggest a complex, but sensitive, interaction between these two commonly mutated amino acids and suggest design principles for the minimal mutations needed to access unnatural activity for future generations of enzymes.

DISCUSSION

SFM19. Here, we have characterized three enzymes that have been previously evolved to recognize 2′-modified substrates; SFM19 was previously selected for its ability to recognize 2′OMeNTPs, while AA40 and SFR3 were previously selected for their ability to recognize ribonucleotides. Interestingly, we find SFM19 is the most active of the three enzymes on all possible substrates, including 2′-hydroxyl modifications. Although it was only previously characterized for 2′OMeNTP addition, we find that SFM19 is the most efficient enzyme at recognizing a wide range of modified nucleotides, including fluoro, amino, and azido modifications and ribonucleotides. The original reports for SFM19 did not consider the substrate scope of this enzyme; our studies here show that it is an active enzyme on a number of substrates and suggest that it is much more broadly applicable to the synthesis of 2′-modified DNA than previously observed.

Collectively, our data do not suggest that any of the enzymes have evolved specific, positive interactions with 2′-modified substrates. For each individual enzyme, the relative ability to recognize different 2′ modifications is identical; 2′-fluoro is most effectively recognized followed by, in descending order, hydroxyl, azido, amino, and methoxy substitutions. These data imply that the enzymes have lost, to varying degrees, the negative discrimination against 2′ modifications rather than acquired an ability to recognize a specific 2′ modification. These experiments also show the relative ability of mutant SF

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

enzymes to recognize these substrates; it is possible that the differences in recognition of modified substrates reflect intrinsic properties of the modified nucleotides rather than the enzymes. More specifically, it is possible that the differences in recognition reflect the ability of 2′ modifications to be tolerated upon being incorporated into a primer terminus. Importantly, 2′-fluoro modifications to DNA have been previously shown to impart nuclease resistance;[8] our data suggest that fluoro substitutions might be an easier target for DNA polymerases than, for example, methoxy substitutions. Future efforts would benefit from further synthetic optimization of the sugar scaffold to examine the full breadth of recognition of modified substrates by these evolved mutants.

It is interesting to speculate about why SFM19, the only enzyme not evolved to recognize ribonucleotides, is superior at recognizing ribonucleotides. The relatively poor recognition of 2′-methoxy-modified nucleotides suggests that this might be one of the most biochemically demanding modifications available. Thus, the more challenging selective pressure of adding a 2′-methoxy-modified nucleotide and, perhaps, the even more challenging selective pressure of adding a nucleotide to a 2′-methoxy-modified primer were sufficient to select for a more active enzyme, for a number of different substrates. This suggests, in turn, that SFR3 and AA40 were selected on the basis of secondary properties rather than being the most effective polymerase for their substrate.

Biochemical and Mutational Origins of Unnatural Function. We have also extensively analyzed the origins of the mutational differences between these enzymes. We observe that mutations at I614 and E615 can explain the majority of unnatural function; differences in how mutations at I614 and E615 interact predominantly explain differences in function between these enzymes. Notably, while many of the mutations observed here are observed in multiple enzymes that recognize 2′ modifications, they are also found in mutant enzymes for other functions, implying that lessons learned here might have broader implications for polymerase engineering.

While mutations at positions 614 and 615 appear to be the primary determinants of unnatural function, it is noteworthy that nearly all of the other identified mutations have been identified in other studies, as well (Table 1). For example, A597T, a mutation in SFR3, has also been identified in an enzyme capable of reverse transcription[35] as well as a different polymerase capable of incorporating 3′-modified substrates.[36] A608V, a mutation in AA40, has also been identified in a polymerase with increased tolerance for hydrophobic nucleobases,[28] and a different polymerase capable of extending aberrant primer termini.[37] Similarly, mutations of W604[35] and E602[28] have been identified in other studies, as well. Interestingly, we did not observe measurable activity for any of these mutations, and we conclude that these mutations are unlikely to play a prominent functional role in catalysis of 2′-modified nucleic acids. While it is possible that these recurrent mutations occur only because they are well-tolerated and thus do not contribute to fitness, it is also plausible that they may have an important function such as improving protein stability or the kinetics of folding. These properties would not be identified in the assays performed here; to assess the importance of these mutations to noncatalytic functions requires further experiments. Such studies would have important implications for polymerase engineering.

Through systematic characterization of point mutants, we have shown that E615G is the primary driver of unnatural function in all three mutants. The high activity of E615G is perhaps unsurprising; prior studies have shown that mutation of E710, the steric gate amino acid in the proteolytic fragment of the highly homologous DNA polymerase I from *E. coli* (Klenow fragment), similarly increases the level of recognition of 2′-modified nucleotides.[19] Similar steric gate mutations are commonly needed for acquisition of the ability to recognize 2′-modified nucleotides[15] and are often an initial important step in the engineering of further function. Notably, there is strong selective pressure for DNA polymerases to possess such a steric gate to prevent incorporation of ribonucleotides into genomic DNA, which would compromise the stability of the genome.

In the absence of other mutations, the role of I614 in substrate broadening, especially in isolation, is more complex. Mutations of I614 have also been previously shown to broaden substrate specificity in Klenow fragment[38] and have been identified in a number of evolved variants of Taq for broadly ranging properties such as improved recognition of unnatural DNA base pairs,[29] hydrophobic nucleobase analogues,[28] and ancient DNA.[39] Although these previous directed evolution studies have not directly demonstrated that I614 specifically broadens acceptance of these substrates, the recurrent evolution of enzymes containing mutations at I614 implies that there might be a role for I614 in substrate broadening, generally. Our data show that a range of mutations to I614, including several mutations that have not been identified in prior studies, impart broadened substrate recognition toward 2′ substrates. These data suggest that I614 might be particularly effective at broadening substrate recognition at 2′ modifications and suggest that mutations at this position should be included in future evolution efforts, especially considering their synergistic interactions with E615G. Notably, from our data, it is not clear if these mutations specifically alter recognition of 2′-modified substrates or more nonspecifically broaden recognition, which might be the cause of its frequent recurrence in directed evolution experiments. Previously, scientists have posited that I614 mutations create space in the active site; however, our data do not suggest a relationship between the relative size of the mutation and the promiscuity of the enzyme. It is possible that mutations at I614 play a broader role, such as impacting enzyme dynamics. Further experiments are required to more closely examine the biochemical and biophysical role of I614.

Our study is the first to systematically combine mutations at I614 and E615 to observe the potential synergistic effects of their combination. Our data suggest that these two mutations comprise the bulk of the activity for the mutant enzymes previously identified; the relative activities of the double mutants closely parallel the relative activity of the parents. Thus, while the genetic identities of the three previously identified mutants are substantially different, the core mutations required for unnatural function of these enzymes are relatively similar, relying predominantly on mutations at I614 and E615. Interestingly, we observed highly synergistic interactions between mutations at these amino acids. Most importantly, negatively charged amino acids, either aspartate or glutamate, significantly increased the activity of the enzyme upon being added to E615G, implying that a negative charge is most beneficial at position 614.

Our biochemical data characterizing the parent enzymes suggest a possible rationale for the synergistic interactions between residues 614 and 615. We did not observe substantial differences in the respective ability of each parent enzyme to add a ribonucleotide to an unmodified primer; however, the

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

parents possess highly varying abilities to incorporate unmodified nucleoside triphosphates onto a modified primer. This suggests that modified primer recognition differentiates these three enzymes more than modified nucleoside triphosphate recognition. Thus, we propose a biochemical model for the differences: E615G both broadly enables 2′-modified nucleoside triphosphate recognition and weakens the ability to extend modified primer termini. Mutation of residue 614 to a negatively charged amino acid can compensate for the weakened ability to extend modified primer termini while not impacting the ability to recognize modified triphosphates.

While mutations of E615 to nonchelating amino acids have been shown to change the magnesium dependence of the enzyme,[19] structural studies of SF have not shown direct functional interactions between a magnesium ion and E615. Instead, structural studies have shown E615 interacting with R573 (Figure 7).[40,41] R573 is known to form an arginine fork

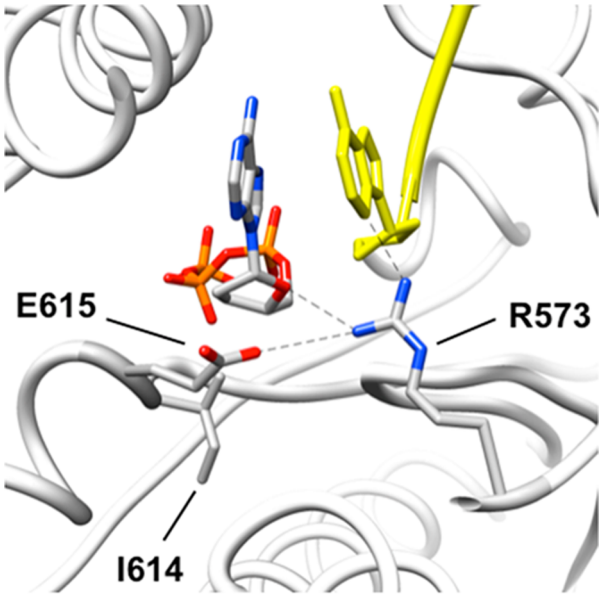


Figure 7. E615 and R573. Active site of Taq DNA polymerase (PDB entry 1QSY) with the primer strand colored yellow. R573, previously shown to form a functional arginine fork between the incoming dNTP and the minor groove of the terminal nucleobase of the primer strand, also forms a hydrogen bond to E615. The I614E mutation may help to restore this lost interaction after the E615G mutation.

that orients the incoming dNTP and the primer strand through H-bonds to the ribosyl oxygen of the incoming dNTP and the minor groove of the primer nucleobase;[42] E615 might help orient R573 to form an optimally catalytically competent state for the enzyme. Mutation of E615 might diminish this capacity but may be restored through mutation of I614 to aspartate or glutamate. Thus, we speculate that we have observed a previously unknown functional role for E615G in primer extension, which might be engendered through structurally observed interactions with R573. This function may be, in SFM19 and I614D/E615G, restored through mutation at I614, providing a possible biochemical explanation for the observed synergistic interactions. These observations not only inform polymerase design and biochemistry but also add to the growing literature about synergistic interactions between amino acid mutations.[43,44]

Implications for Engineering. In summary, our data show that the SFM19 (I614E/E615G) shows activity higher than that of SFR3 or AA40 on a broad range of 2′ modifications. Notably, previous research has shown that the fidelity of SFM19 is not significantly altered by these mutations, as steady-state kinetic assays did not detect mispair formation at a detectable level.[11] Considering both the broad ability to accept 2′-modified nucleotides, as well as the encouraging results regarding fidelity, these data suggest that SFM19 is likely the best starting point for future engineering efforts. Notably, in all three previous evolution experiments, the mutant DNA polymerase library focused diversity on motif A of DNA polymerase I and the selection required the addition of fewer than six nucleotides.[11−13] Considering both our biochemical data and the fact that three independent evolution experiments were conducted, the combination of I614E and E615G mutations is likely one of, if not the best, combination of mutations that might be found from such a library and selection. While SFM19 is the most active enzyme because of its ability to recognize modified DNA, extension of modified DNA still limits this enzyme; thus, SFM19 should be the basis for a library that adds additional diversity outside of motif A in regions previously shown to alter interactions between the DNA polymerase and mutant DNA. It is likely that, given appropriate selection conditions that directly select for synthesis onto 2′-modified primers, this might yield mutant Taq polymerases that are capable of synthesizing full length DNA bearing a range of 2′-modified substitutions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00689.

Supporting methods, Figures S1−S21, and Tables S1−S3 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: 909-607-9353. E-mail: aleconte@kecksci.claremont.edu.

Funding

This work was supported by the Research Corporation for Scientific Advancement (Cottrell College Scholar Award 22548) and by TriLink Biotechnologies (ResearchReward Grant G139).

Notes

The authors declare no competing financial interests.

■ ACKNOWLEDGMENTS

We thank the Romesberg group for providing the wild-type Stoffel fragment plasmid. We thank Scripps College Academy Math and Science Scholars Program students for assistance in creating mutant enzymes.

■ ABBREVIATIONS

Taq, *T. aquaticus* DNA polymerase I; SF, Stoffel fragment; wt, wild type; dNTP, deoxyribonucleoside triphosphate; rNTP, ribonucleoside triphosphate; OMeNTP, 2′-*O*-methyl ribonucleoside triphosphate.

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

REFERENCES

(1) Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science 239*, 487−491.

(2) Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A. 74*, 5463−5467.

(3) Shendure, J., and Ji, H. (2008) Next-generation DNA sequencing. *Nat. Biotechnol. 26*, 1135−1145.

(4) Clark, M. A. (2010) Selecting chemicals: the emerging utility of DNA-encoded libraries. *Curr. Opin. Chem. Biol. 14*, 396−403.

(5) Zhao, W., Ali, M. M., Brook, M. A., and Li, Y. (2008) Rolling circle amplification: applications in nanotechnology and biodetection with functional nucleic acids. *Angew. Chem., Int. Ed. 47*, 6330−6337.

(6) Kranaster, R., and Marx, A. (2010) Engineered DNA polymerases in biotechnology. *ChemBioChem 11*, 2077−2084.

(7) Chen, T., and Romesberg, F. E. (2014) Directed polymerase evolution. *FEBS Lett. 588*, 219−229.

(8) Layzer, J. M., McCaffrey, A. P., Tanner, A. K., Huang, Z., Kay, M. A., and Sullenger, B. A. (2004) In vivo activity of nuclease-resistant siRNAs. *RNA 10*, 766−771.

(9) Wilson, C., and Keefe, A. D. (2006) Building oligonucleotide therapeutics using non-natural chemistries. *Curr. Opin. Chem. Biol. 10*, 607−614.

(10) Keefe, A. D., and Cload, S. T. (2008) SELEX with modified nucleotides. *Curr. Opin. Chem. Biol. 12*, 448−456.

(11) Fa, M., Radeghieri, A., Henry, A. A., and Romesberg, F. E. (2004) Expanding the substrate repertoire of a DNA polymerase by directed evolution. *J. Am. Chem. Soc. 126*, 1748−1754.

(12) Xia, G., Chen, L., Sera, T., Fa, M., Schultz, P. G., and Romesberg, F. E. (2002) Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A. 99*, 6597−6602.

(13) Ong, J. L., Loakes, D., Jaroslawski, S., Too, K., and Holliger, P. (2006) Directed evolution of DNA polymerase, RNA polymerase and reverse transcriptase activity in a single polypeptide. *J. Mol. Biol. 361*, 537−550.

(14) Staiger, N., and Marx, A. (2010) A DNA polymerase with increased reactivity for ribonucleotides and C5-modified deoxyribonucleotides. *ChemBioChem 11*, 1963−1966.

(15) Cozens, C., Pinheiro, V. B., Vaisman, A., Woodgate, R., and Holliger, P. (2012) A short adaptive path from DNA to RNA polymerases. *Proc. Natl. Acad. Sci. U. S. A. 109*, 8067−8072.

(16) Goodman, M. F., Creighton, S., Bloom, L. B., Petruska, J., and Kunkel, T. A. (1993) Biochemical basis of DNA replication fidelity. *Crit. Rev. Biochem. Mol. Biol. 28*, 83−126.

(17) Suzuki, M., Yoshida, S., Adman, E. T., Blank, A., and Loeb, L. A. (2000) Thermus aquaticus DNA polymerase I mutants with altered fidelity. Interacting mutations in the O-helix. *J. Biol. Chem. 275*, 32728−32735.

(18) Loh, E., and Loeb, L. A. (2005) Mutability of DNA polymerase I: implications for the creation of mutant DNA polymerases. *DNA Repair 4*, 1390−1398.

(19) Astatke, M., Ng, K., Grindley, N. D., and Joyce, C. M. (1998) A single side chain prevents Escherichia coli DNA polymerase I (Klenow fragment) from incorporating ribonucleotides. *Proc. Natl. Acad. Sci. U. S. A. 95*, 3402−3407.

(20) Li, Y., and Waksman, G. (2001) Crystal structures of a ddATP-, ddTTP-, ddCTP, and ddGTP- trapped ternary complex of Klentaq1: insights into nucleotide incorporation and selectivity. *Protein Sci. 10*, 1225−1233.

(21) Bergen, K., Steck, A. L., Strutt, S., Baccaro, A., Welte, W., Diederichs, K., and Marx, A. (2012) Structures of KlenTaq DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates. *J. Am. Chem. Soc. 134*, 11840−11843.

(22) Betz, K., Malyshev, D. A., Lavergne, T., Welte, W., Diederichs, K., Dwyer, T. J., Ordoukhanian, P., Romesberg, F. E., and Marx, A. (2012) KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry. *Nat. Chem. Biol. 8*, 612−614.

(23) Kermekchiev, M. B., Tzekov, A., and Barnes, W. M. (2003) Cold-sensitive mutants of Taq DNA polymerase provide a hot start for PCR. *Nucleic Acids Res. 31*, 6139−6147.

(24) Ghadessy, F. J., Ong, J. L., and Holliger, P. (2001) Directed evolution of polymerase function by compartmentalized self-replication. *Proc. Natl. Acad. Sci. U. S. A. 98*, 4552−4557.

(25) Zhang, Z., Kermekchiev, M. B., and Barnes, W. M. (2010) Direct DNA amplification from crude clinical samples using a PCR enhancer cocktail and novel mutants of Taq. *J. Mol. Diagn. 12*, 152−161.

(26) Baar, C., d'Abbadie, M., Vaisman, A., Arana, M. E., Hofreiter, M., Woodgate, R., Kunkel, T. A., and Holliger, P. (2011) Molecular breeding of polymerases for resistance to environmental inhibitors. *Nucleic Acids Res. 39*, e51.

(27) Leconte, A. M., Patel, M. P., Sass, L. E., McInerney, P., Jarosz, M., Kung, L., Bowers, J. L., Buzby, P. R., Efcavitch, J. W., and Romesberg, F. E. (2010) Directed evolution of DNA polymerases for next-generation sequencing. *Angew. Chem., Int. Ed. 49*, 5921−5924.

(28) Loakes, D., Gallego, J., Pinheiro, V. B., Kool, E. T., and Holliger, P. (2009) Evolving a polymerase for hydrophobic base analogues. *J. Am. Chem. Soc. 131*, 14827−14837.

(29) Leconte, A. M., Chen, L., and Romesberg, F. E. (2005) Polymerase evolution: efforts toward expansion of the genetic code. *J. Am. Chem. Soc. 127*, 12470−12471.

(30) Laos, R., Shaw, R., Leal, N. A., Gaucher, E., and Benner, S. (2013) Directed evolution of polymerases to accept nucleotides with nonstandard hydrogen bond patterns. *Biochemistry 52*, 5288−5294.

(31) Loh, E., Choe, J., and Loeb, L. A. (2007) Highly tolerated amino acid substitutions increase the fidelity of Escherichia coli DNA polymerase I. *J. Biol. Chem. 282*, 12201−12209.

(32) Summerer, D., Rudinger, N. Z., Detmer, I., and Marx, A. (2005) Enhanced fidelity in mismatch extension by DNA polymerase through directed combinatorial enzyme design. *Angew. Chem., Int. Ed. 44*, 4712−4715.

(33) Yu, H., Zhang, S., Dunn, M. R., and Chaput, J. C. (2013) An efficient and faithful in vitro replication system for threose nucleic acid. *J. Am. Chem. Soc. 135*, 3583−3591.

(34) Creighton, S., Bloom, L. B., and Goodman, M. F. (1995) Gel fidelity assay measuring nucleotide misinsertion, exonucleolytic proofreading, and lesion bypass efficiencies. *Methods Enzymol. 262*, 232−256.

(35) Vichier-Guerre, S., Ferris, S., Auberger, N., Mahiddine, K., and Jestin, J. L. (2006) A population of thermostable reverse transcriptases evolved from Thermus aquaticus DNA polymerase I by phage display. *Angew. Chem., Int. Ed. 45*, 6133−6137.

(36) Chen, F., Gaucher, E. A., Leal, N. A., Hutter, D., Havemann, S. A., Govindarajan, S., Ortlund, E. A., and Benner, S. A. (2010) Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection. *Proc. Natl. Acad. Sci. U. S. A. 107*, 1948−1953.

(37) Ghadessy, F. J., Ramsay, N., Boudsocq, F., Loakes, D., Brown, A., Iwai, S., Vaisman, A., Woodgate, R., and Holliger, P. (2004) Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat. Biotechnol. 22*, 755−759.

(38) Patel, P. H., and Loeb, L. A. (2000) Multiple amino acid substitutions allow DNA polymerases to synthesize RNA. *J. Biol. Chem. 275*, 40266−40272.

(39) d'Abbadie, M., Hofreiter, M., Vaisman, A., Loakes, D., Gasparutto, D., Cadet, J., Woodgate, R., Paabo, S., and Holliger, P. (2007) Molecular breeding of polymerases for amplification of ancient DNA. *Nat. Biotechnol. 25*, 939−943.

(40) Li, Y., Mitaxov, V., and Waksman, G. (1999) Structure-based design of Taq DNA polymerases with improved properties of dideoxynucleotide incorporation. *Proc. Natl. Acad. Sci. U. S. A. 96*, 9491−9496.

(41) Li, Y., Korolev, S., and Waksman, G. (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008

fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J. 17*, 7514−7525.
(42) Meyer, A. S., Blandino, M., and Spratt, T. E. (2004) Escherichia coli DNA polymerase I (Klenow fragment) uses a hydrogen-bonding fork from Arg668 to the primer terminus and incoming deoxynucleotide triphosphate to catalyze DNA replication. *J. Biol. Chem. 279*, 33043−33046.
(43) Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., and Cooper, T. F. (2011) Negative epistasis between beneficial mutations in an evolving bacterial population. *Science 332*, 1193−1196.
(44) Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M., and Tans, S. J. (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature 445*, 383−386.

DOI: 10.1021/acs.biochem.5b00689
Biochemistry 2015, 54, 5999−6008