JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Site-Specific Labeling of DNA and RNA Using an Efficiently Replicated and Transcribed Class of Unnatural Base Pairs

Young Jun Seo,[†] Denis A. Malyshev,[†] Thomas Lavergne,[†] Phillip Ordoukhanian, and Floyd E. Romesberg*

Department of Chemistry and Center for Protein and Nucleic Acid Research, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037, United States

Supporting Information

ABSTRACT: Site-specific labeling of enzymatically synthesized DNA or RNA has many potential uses in basic and applied research, ranging from facilitating biophysical studies to the in vitro evolution of functional nucleic acids and the construction of various nanomaterials and biosensors. As part of our efforts to expand the genetic alphabet, we have developed a class of unnatural base pairs, exemplified by d**5SICS**-d**MMO2** and d**5SICS**-d**NaM**, which are efficiently replicated and transcribed, and which may be ideal for the site-specific labeling of DNA and RNA. Here, we report the synthesis and analysis of the ribo- and deoxyribo-variants, (d)**5SICS** and (d)**MMO2**, modified with free or protected propargylamine linkers that allow for the site-specific modification of DNA or RNA during or after enzymatic synthesis. We also synthesized and evaluated the α-phosphorothioate variant of d**5SICS**TP, which provides a route to backbone thiolation and an additional strategy for the postamplification site-specific labeling of DNA. The deoxynucleotides were characterized via steady-state kinetics and PCR, while the ribonucleosides were characterized by the transcription of both a short, model RNA as well as full length tRNA. The data reveal that while there are interesting nucleotide and polymerase-specific sensitivities to linker attachment, both (d)**MMO2** and (d)**5SICS** may be used to produce DNA or RNA site-specifically modified with multiple, different functional groups with sufficient efficiency and fidelity for practical applications.

1. INTRODUCTION

The sequence-specific amplification of DNA and RNA has revolutionized biotechnological and biomedical research, making techniques such as cloning and sequencing routine and also enabling a variety of emerging techniques with important applications in diverse fields ranging from synthetic biology[1] to nanomaterials[2,3] and medical diagnostics.[4,5] In addition, the use of polymerases to amplify DNA molecules has made it possible to evolve functional DNAs or RNAs with desired ligand recognition or catalytic properties.[6,7] However, the potential physical and functional properties of nucleic acids are restricted by the limited diversity of their constituent nucleotides. While the potential diversity of DNA or RNA may be expanded via chemically synthesizing the oligonuclotides,[8−10] this approach is incompatible with enzymatic amplification. Thus, efforts have been directed toward the modification of one or more of the natural triphosphates with linkers that enable the attachment of different functional groups without ablating polymerase recognition.[11−21] The best sites of linker attachment have proven to be the nucleobases: specifically, the C5 position of the pyrimidines and the C7 position of the 7-deaza purine scaffold. With the pyrimidine scaffold, the detailed structure of the linker is critical,[13,15] with rigid alkynes or *trans*-alkenes being optimal.[13,19,22,23] At least with the propargylamide-like linkers and A-family polymerases (such as *Taq*), this appears to result at least in part from favorable hydrogen-bonding interactions between the polymerase and the amide carbonyl group of the linker.[24]

While the linker modification of one or more natural dNTPs allows for the polymerase-mediated synthesis of highly functionalized DNA, the efficiency of modified triphosphate incorporation is generally reduced compared to their natural counterparts, resulting in lower yields of amplified product as well as strong sequence dependences.[14,25] For oligonucleotide evolution experiments, this sequence-dependent efficiency is especially problematic because it decouples amplification level from copy number, and thus from fitness, during the amplification step required after or during selection. Moreover, high functional group density is unavoidable in any nucleotide-specific labeling strategy (i.e., all or none of a given nucleotide must be modified), but the majority of functional groups are unlikely to contribute to any desired property while still contributing to biases in replication. Finally, it is noteworthy that proteins rarely (if ever) employ such a high density of functional groups but instead typically have only one or a few that are embedded in a controlled environment where their activity (i.e., pK_a, nucleophilicity, redox potential, etc.) can be manipulated. Thus, methods to site-specifically label nucleic acids with one or a few functional groups attached to the fifth and sixth nucleotides of an expanded genetic alphabet would enable less biased selections and potentially the construction of more protein-like environments.

The development of an efficiently replicated and transcribed unnatural base pair would make possible the site-specific modification of DNA and RNA. While the four-letter genetic alphabet based on complementary hydrogen-bonding (H-bonding) is conserved throughout nature, and orthogonal H-bonding

Received: August 21, 2011
Published: October 08, 2011

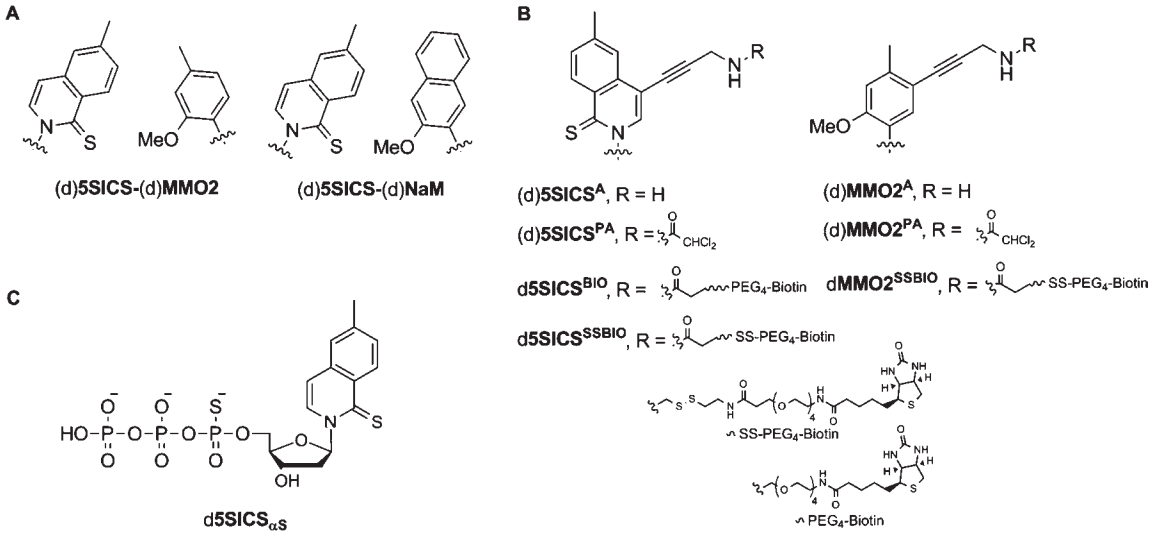

Figure 1. (A) Parental unnatural base pairs, (B) linker-modified analogues, and (C) α-phosphorothioate variant of d**5SICS**TP. Sugar and phosphate backbone are omitted for clarity in A and B.

patterns have been used to design unnatural base pairs,[26−29] we[30−41] and others[43−47] have demonstrated the potential of hydrophobic and packing forces to control the efficient and selective replication of unnatural base pairs. However, as with the natural nucleotides, the general use of an unnatural base pair for the site-specific labeling of DNA or RNA requires that the unnatural nucleotides bear linkers or other functionalities that allow different groups to be attached without ablating polymerase recognition.

Two of the most promising unnatural base pairs that we have developed are those formed between d**5SICS** and d**MMO2** (d**5SICS**-d**MMO2**) or d**NaM** (d**5SICS**-d**NaM**) (Figure 1A). Both unnatural base pairs can be amplified by PCR with good to excellent efficiency and fidelity using a variety of DNA polymerases, even when embedded within difficult to replicate sequences.[30,34,38,42] While d**5SICS**-d**NaM** is replicated and transcribed with greater efficiency, our efforts to develop an unnatural base pair for in vitro applications include both d**5SICS**-d**NaM** and d**5SICS**-d**MMO2**, because unlike the annular ring scaffold of d**NaM**, the d**MMO2** scaffold may be modified with linkers at a position that is analogous to the C5 position of the natural pyrimidines.

Here, we report the synthesis and evaluation of protected and free propargylamine linker-derivatized variants of both the deoxy- and ribonucleosides of **MMO2** and **5SICS** ((d)**MMO2**$^{\mathbf{PA}}$, (d)-**MMO2**$^{\mathbf{A}}$, (d)**5SICS**$^{\mathbf{PA}}$, and (d)**5SICS**$^{\mathbf{A}}$), as well as the biotinylated deoxynucleotides d**MMO2**$^{\mathbf{SSBIO}}$, d**5SICS**$^{\mathbf{BIO}}$, and d**5SICS**$^{\mathbf{SSBIO}}$ (Figure 1B). The deoxynucleotides were characterized via both steady-state kinetics and PCR, and the ribonucleosides were characterized by the transcription of short model RNAs as well as a full length amber suppressor tyrosyl tRNA ($tRNA^{Tyr}_{CUA}$) from *Methanocaldococcus jannaschii*.[48] Both the protected and free amino variants were evaluated to fully explore double labeling strategies, including those involving postsynthetic modification. The data reveals that the modifications of the unnatural nucleotides have varying effects on the different steps of replication and transcription and that the linkers are generally less perturbative within the (d)**MMO2** scaffold. We also report the synthesis and evaluation of the α-phosphorothioate variant of d**5SICS**TP (d**5SICS**$_{\alpha\mathbf{S}}$TP, Figure 1C) and show that it is an efficient route to site-specifically introduce a thiolate for functional group attachment. Overall, we demonstrate that several of the modified nucleotides are sufficiently well recognized by polymerases to enable practical applications based on the site-specific labeling of nucleic acids with multiple different functionalities.

2. RESULTS

2.1. Synthesis of Modified Unnatural Nucleotides. With natural triphosphates, propargylamide linkers are relatively well tolerated by different polymerases.[13,23] To explore the use of similar linkers with d**5SICS**-d**NaM** and d**5SICS**-d**MMO2**, we synthesized (d)**5SICS**$^{\mathbf{PA}}$TP, (d)**5SICS**$^{\mathbf{A}}$TP, d**5SICS**$^{\mathbf{BIO}}$TP, d**5SICS**$^{\mathbf{SSBIO}}$TP, (d)**MMO2**$^{\mathbf{PA}}$TP, (d)**MMO2**$^{\mathbf{A}}$TP, and (d)-**MMO2**$^{\mathbf{BIO}}$TP (Figure 1B), as described in detail in the Supporting Information. Briefly, d**5SICS**$^{\mathbf{PA}}$ was generated from the isosteryl compound, which was iodinated and then sulfonylated. The modified base was coupled to (2*R*,5*R*)-5-chloro-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl 4-methylbenzoate, resulting in anomeric mixtures of nucleosides, with the pure β-anomer obtained by column chromatography. After removal of the toluoyl group, the iodo functionality was used to attach the dichloro acetyl protected propargylamine via Sonogashira coupling. Phosphorylation under Ludwig conditions[49] provided d**5SICS**$^{\mathbf{PA}}$TP, which was purified by anion exchange chromatography followed by HPLC, and then deprotected to provide d**5SICS**$^{\mathbf{A}}$TP. The biotinylated analogues, d**5SICS**$^{\mathbf{BIO}}$TP and d**5SICS**$^{\mathbf{SSBIO}}$TP, were synthesized by coupling d**5SICS**$^{\mathbf{A}}$TP to NHS-PEG$_4$-biotin or NHS-SS-PEG$_4$-biotin, respectively. To explore the use of a site-specifically incorporated backbone thiolate, d**5SICS**$_{\alpha\mathbf{S}}$TP (Figure 1C) was synthesized via the phosphorylation of the corresponding nucleoside in the presence of thiophosphoryl chloride.[50] Note that this usually results in a mixture of Sp and Rp stereoisomers and that only the Sp is recognized by DNA polymerases.[51] For d**MMO2** derivatives, we iodinated the free nucleoside, which was synthesized as described previously,[40] and then attached the dichloro acetyl protected propargylamine to provide the desired protected nucleoside.[52] The nucleoside was then phosphorylated to provide d**MMO2**$^{\mathbf{PA}}$TP, which was subsequently deprotected to provide d**MMO2**$^{\mathbf{A}}$TP. The

Table 1. Steady-State Kinetic Data for Kf-Mediated Insertion of Modified Triphosphates (dYTP) Opposite Their Cognate Nucleotide in the Template (X)

5′-d(TAATACGACTCACTATAGGGAGA)
3′-d(ATTATGCTGAGTGATATCCCTCT**X**GCTAGGTTACGGCAGGATCGC)

X	Y	k_{cat} (min^{-1})	K_M (μM)	$k_{cat}/K_M \times 10^5$ (M^{-1} min^{-1})
T	A[a]	4.1 ± 0.3	0.0053 ± 0.0004	7700
5SICS	**NaM**[a]	14.6 ± 0.8	0.25 ± 0.04	580
	MMO2[a]	14 ± 2	35.8 ± 0.6	4.0
	MMO2^{PA}	11 ± 2	15 ± 2	7.2
	MMO2^{A}	6.7 ± 0.3	69 ± 8	0.97
	MMO2^{SSBIO}	1.5 ± 0.3	45 ± 10	0.34
NaM	**5SICS**[a]	8.3 ± 1.1	0.039 ± 0.004	2100
	5SICS^{PA}	8.1 ± 0.9	1.0 ± 0.1	83
	5SICS^{A}	0.085 ± 0.05	19 ± 4	0.045
	5SICS$_{\alpha S}$[b]	6.5 ± 1.5	0.42 ± 0.16	160
	5SICS^{BIO}	0.55 ± 0.13	31 ± 12	0.18
	5SICS^{SSBIO}	1.3 ± 0.3	29 ± 9	0.45

[a] Ref 42. [b] Mixture of Sp and Rp diastereomers.

cleavable biotinylated analogue, d**MMO2**^{SSBIO}TP, was synthesized via coupling of d**MMO2**^{A}TP with NHS-SS-PEG$_4$-biotin.

For the **5SICS** ribonucleoside derivatives, 5-methyl isocarbostyril was first coupled with (2*R*)-2-((benzoyloxy)methyl)-5-oxotetrahydrofuran-3,4-diyl dibenzoate. The pure β anomer was purified from an anomeric mixture by column chromatography to provide the desired benzoyl protected nucleoside. Iodination, sulfonylation, and benzoyl deprotection, followed by coupling with the dichloro acetyl propargylamine, then provided **5SICS**^{PA}; phosphorylation provided **5SICS**^{PA}TP, and deprotection provided **5SICS**^{A}TP. For the **MMO2** derivatives, the free nucleoside was synthesized as reported previously,[39] iodinated, and then finally coupled to the dichloro acetyl protected propargylamine to produce **MMO2**^{PA}. Phosphorylation and deprotection proceeded as with d**5SICS**^{A} to provide **MMO2**^{PA}TP and **MMO2**^{A}TP.

2.2. Steady-State Kinetic Analysis of Linker-Modified Unnatural Base Pair Synthesis. To characterize polymerase recognition of the modified unnatural base pairs, we first explored the efficiency with which the exonuclease deficient Klenow fragment of *E. coli* DNA polymerase I (Kf) inserts the linker-modified derivatives of d**5SICS**TP opposite d**NaM** (Table 1). For comparison, Kf inserts dATP opposite dT and d**5SICS**TP opposite d**NaM** with an efficiency of 7.7×10^8 M^{-1} min^{-1} and 2.1×10^8 M^{-1} min^{-1}, respectively.[42] We found that Kf inserts d**5SICS**^{PA}TP with an efficiency of 8.3×10^6 M^{-1} min^{-1}. The 25-fold decreased insertion efficiency relative to d**5SICS**TP is due entirely to an elevated apparent K_M. d**5SICS**^{BIO}TP and d**5SICS**^{SSBIO}TP were inserted less efficiently, with second-order rate constants of 1.8×10^4 M^{-1} min^{-1} and 4.5×10^4 M^{-1} min^{-1}, respectively, again, largely because of elevated apparent K_M values. The insertion of d**5SICS**^{A}TP opposite d**NaM** by Kf is even less efficient, proceeding with a second-order rate constant of only 4.5×10^3 M^{-1} min^{-1}. In this case, the reduced insertion efficiency relative to d**5SICS**TP is due to both a decreased apparent k_{cat} and an increased apparent K_M.

We then characterized the efficiency with which Kf inserts the linker-modified derivatives of d**MMO2**TP opposite d**5SICS** (Table 1). For comparison, Kf inserts d**MMO2**TP opposite d**5SICS** with an efficiency of 4×10^5 M^{-1} min^{-1}.[31] Interestingly, we found that d**MMO2**^{PA}TP is inserted more efficiently ($k_{cat}/K_M = 7.2 \times 10^5$ M^{-1} min^{-1}) than d**MMO2**TP, because of a 2-fold decrease in the apparent K_M, while d**MMO2**^{A}TP is inserted only slightly less efficiently ($k_{cat}/K_M = 9.7 \times 10^4$ M^{-1} min^{-1}), because of small changes in both the apparent k_{cat} and K_M. Because d**5SICS**^{SSBIO}TP was more efficiently recognized than d**5SICS**^{BIO}TP, we also characterized d**MMO2**^{SSBIO}TP and found that it is inserted by Kf opposite d**5SICS** with an efficiency of 3.4×10^4 M^{-1} min^{-1}, because of a 9-fold decrease in the apparent k_{cat}.

2.3. PCR Amplification of Modified DNA. To characterize the effect of the linkers on the PCR amplification of DNA containing the unnatural base pair, we incorporated d**MMO2**-d**5SICS** into the middle of a 149-mer DNA duplex. The duplex DNA was then PCR amplified using DeepVent (exo$^+$) DNA polymerase and a wide variety of different combinations of the modified triphosphates. The amplification level was determined for each combination of triphosphates, and then the amplicons were sequenced to determine replication fidelity (Table 2, and Figure S1, Supporting Information). In most cases, the amplification efficiency approaches that of the fully natural control (996-fold amplification), and the fidelity as measured by sequencing is in excess of 99% per doubling. The single exceptions are the amplifications with d**5SICS**^{A}TP, where the amplification level was only ~200-fold and the fidelity only ~90% per round. This agrees well with the steady-state kinetic data, which demonstrated that d**5SICS**^{A}TP insertion is inefficient. However, despite the slight reduction in steady-state insertion efficiency of d**5SICS**^{PA}TP and d**5SICS**$_{\alpha S}$TP, relative to d**5SICS**TP, no significant or systematic differences were apparent in amplification efficiency or fidelity with these three triphosphates. Fidelities of PCR amplifications involving biotin-modified triphosphates were also determined by streptavidin gel shift (Figure 2). These fidelities paralleled those determined by sequencing but were generally somewhat reduced. We attribute the differences to incomplete binding of streptavidin to the biotin tag due to its location in the middle of the large duplex where it is more obscured than when present at the terminus, as has been more commonly examined.

Table 2. PCR Fidelities and Amplification Efficiencies[a]

dXTPs incorporated	amplification	fidelity (sequencing)[b]	fidelity (gel shift)[c]
d**NaM** and d**5SICS**	735	>99.7	–
d**MMO2** and d**5SICS**	609	99.6	–
d**NaM** and d**5SICS**$_{\alpha S}$[d]	662	99.6	–
d**MMO2**^{PA} and d**5SICS**	489	>99.7	–
d**MMO2**^{A}, d**5SICS**	528	99.5	–
d**NaM** and d**5SICS**^{PA}	888	>99.7	–
d**NaM** and d**5SICS**^{A}	160	91.1	–
d**MMO2**^{PA} and d**5SICS**^{PA}	960	>99.7	–
d**MMO2**^{A} and d**5SICS**^{A}	279	90.9	–
d**MMO2**^{A} and d**5SICS**^{PA}	624	98.7	–
d**MMO2**^{PA} and d**5SICS**^{A}	167	91.4	–
d**MMO2**^{A} and d**5SICS**$_{\alpha S}$[d]	378	99.4	–
d**MMO2**^{SSBIO} and d**5SICS**	351	99.2	95.0
d**NaM** and d**5SICS**^{SSBIO}	690	>99.7	95.4
d**NaM** and d**5SICS**^{BIO}	624	>99.7	96.5
d**NaM** only	164[e]	–[e]	–
d**5SICS** only	217[e]	–[e]	–

[a] See Materials and Methods for experimental details. [b] Calculated as average fidelities in both directions per doubling (see Materials and Methods for details). [c] Calculated from gel mobility assay (see text). [d] d**5SICS**$_{\alpha S}$ was used as a mixture of Sp and Rp diastereomers. [e] Unnatural base pair lost during amplification.

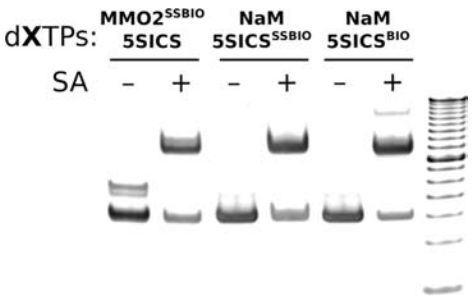

Figure 2. PCR fidelity determination via gel mobility. Based on the intensity of the DNA band that shifts in the presence of added streptavidin (SA), the biotin incorporation level is 65% for d**5SICS**-d**MMO2**^{SSBIO}, 64% for d**5SICS**^{SSBIO}-d**NaM**, and 72% for d**5SICS**^{BIO}-d**NaM**. (Note that the slightly slower migrating band in lane 1 corresponds to unbiotinylated single-stranded DNA resulting from incomplete annealing after PCR.) The overall incorporation levels are converted to fidelities (Table 2) by normalizing by the number of doublings. A 50 bp DNA ladder is loaded in the rightmost lane.

This conclusion is consistent with the immobilization studies described below. However, we cannot exclude the possibility that some of the biotin may have been lost by disulfide cleavage or exchange during PCR.

2.4. Postamplification Modification of DNA. To examine the postamplification labeling of DNA with a single functional group, we first explored strategies based on the PCR amplification of DNA with d**5SICS**^{A}TP and d**NaM**TP or d**5SICS**TP and d**MMO2**^{A}TP (Figure 3A,B). The resulting duplexes were biotinylated using sulfo-NHS-SS-biotin and the labeling efficiency was determined by streptavidin gel shift (Figure 3D). We found that the reactions were complete after 1 h, with 60−70% of the duplexes modified. We next attempted to immobilize the biotin-labeled DNA to streptavidin solid support; however, unlike conventional end-labeling via biotinylated primers,[53] we found that the nature and length of the spacer arm used to attach the biotin is critical. Virtually all of the biotinylated DNA was bound to the solid support when the linker was conjugated to biotin via a hydrophilic PEG_4 spacer (29 Å spacer arm) or an SS-PEG_4 spacer (38 Å spacer arm including a cleavable disulfide bridge, Figure 4A, B) coupled to d**MMO2**^{A}, but only small amounts of the DNA were bound when shorter and more hydrophobic spacers were used (for example, after conjugation to EZ-Link sulfo-NHS-SS-biotin which has a 24 Å spacer arm, data not shown). In the case of SS-PEG_4-immobilized DNA, after washing to remove any contaminating natural DNA and cleaving the disulfide with DTT, the double-stranded DNA was efficiently released and then reconjugated to iodoacetyl-PEG_2-biotin. Gel-shift assays before and after the final conjugation revealed that greater than 50% of the duplexes were labeled (Figure 4C).

To explore the postamplification labeling of DNA with two different groups, we pursued two strategies. First, we examined the amplification of DNA containing either d**5SICS**^{PA}-d**MMO2**^{A} or d**5SICS**^{A}-d**MMO2**^{PA} (Figure 5A,B). After amplification, the free amino groups were labeled with NHS-SS-PEG_4-biotin and characterized by gel mobility. After propargylamine deprotection, the newly liberated amine was labeled with sulfo-NHS-biotin, with the labeling efficiency again characterized via gel mobility. Labeling efficiencies were 60−85% for each step of labeling via d**5SICS**^{PA}-d**MMO2**^{A} and d**5SICS**^{A}-d**MMO2**^{PA} (Figure 5D). A second strategy for double labeling was based on the amplification of DNA using d**5SICS**^{A}TP and d**MMO2**^{SSBIO}TP (Figure 5C). The level of biotinylation via direct d**MMO2**^{SSBIO} incorporation was characterized via streptavidin gel shift, and a second biotin was then attached via coupling (DTT resistant) NHS-PEG_4-biotin to d**5SICS**^{A}. After treatment with DTT to selectively remove the biotin from d**MMO2** (quantitative removal based on control reactions, Figure 5E), the level of biotinylation at d**5SICS** was characterized. Labeling efficiencies were 85% and 68%, respectively, for the first and second steps via d**5SICS**^{A}-d**MMO2**^{SSBIO} (Figure 5E).

2.5. Thiolation of DNA as a Route to Site-Specific Labeling. We also explored the use of d**5SICS**$_{\alpha S}$TP to site-specifically incorporate a reactive center into the backbone of DNA (Figure 3C). As expected, steady-state kinetics revealed that the α-phosphorothioate is only slightly perturbative, with Kf inserting d**5SICS**$_{\alpha S}$TP opposite d**NaM** with an efficiency of 1.6×10^7 M^{-1} min^{-1} (Table 1). To explore the use of labeling with d**5SICS**$_{\alpha S}$ for postamplification modification, we amplified DNA with d**NaM**TP and d**5SICS**$_{\alpha S}$TP. As expected, d**5SICS**$_{\alpha S}$-d**NaM** amplified with excellent efficiency and fidelity (Table 2). To explore postsynthetic labeling, the resulting duplex DNA was conjugated to EZ-Link iodoacetyl-PEG_2-biotin. Gel-shift assays with streptavidin demonstrated that 70% of the duplexes were labeled (Figure 3D), in good agreement with values reported in the literature for chemically synthesized α-phosphorothioate DNA.[54,55]

2.6. Site-Specifically Modified RNA. To explore the site-specific modification of RNA, we first characterized the ability of the RNA polymerase from T7 bacteriophage (T7 RNAP) to incorporate the linker-modified unnatural ribonucleotide triphosphates into 17 nt transcripts (Figure 6A). We first examined the transcription of a template containing d**5SICS** with the natural ribotriphosphates and either **MMO2**^{PA}TP or **MMO2**^{A}TP.

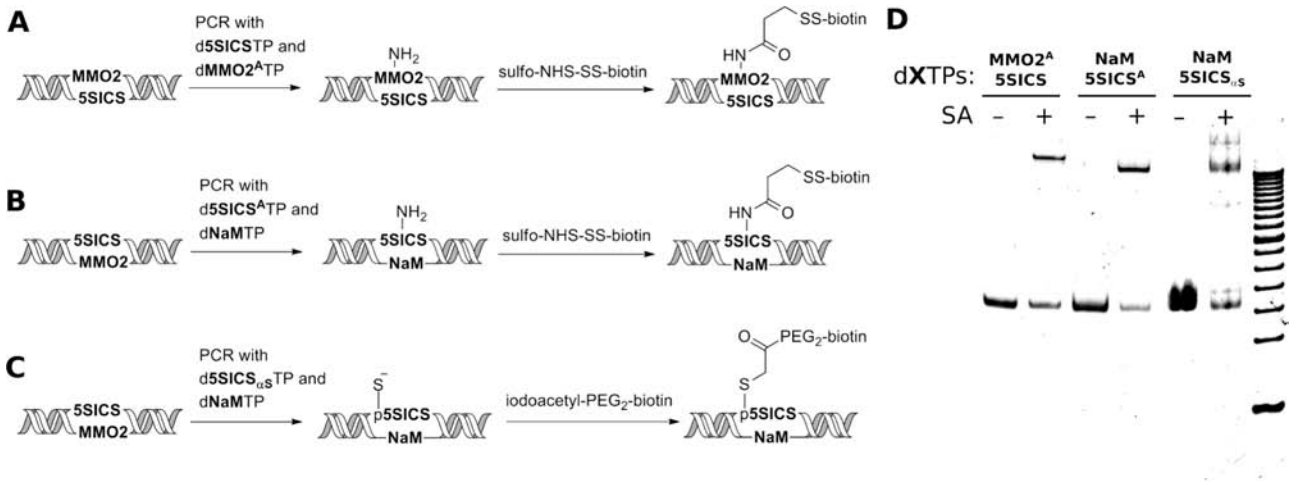

Figure 3. Postamplification DNA labeling with single functional groups, using (A) d**5SICS**$^{\mathbf{A}}$, (B) d**MMO2**$^{\mathbf{A}}$, and (C) d**5SICS**$_{\alpha S}$. (D) Determination of labeling efficiency via streptavidin gel shift. The biotin incorporation level is 55% for d**5SICS**-d**MMO2**$^{\mathbf{A}}$, 70% for d**5SICS**$^{\mathbf{A}}$-d**NaM**, and 70% for d**5SICS**$_{\alpha S}$-d**NaM**. A 50 bp DNA ladder is included in the rightmost lane of the gel. The faster migrating, strong band corresponds to dsDNA, while the slower migrating band corresponds to the 1:1 complex between dsDNA and streptavidin. The faint and most slowly migrating band in lane 6 corresponds to the 2:1 complex of dsDNA and streptavidin.

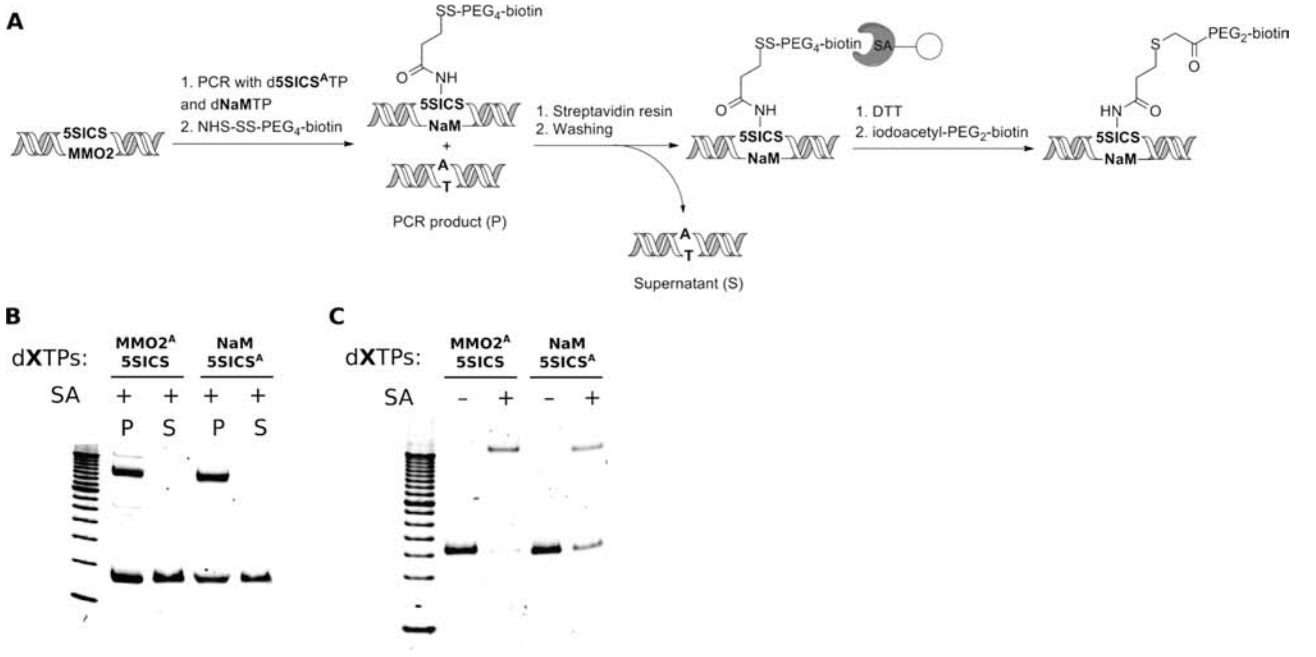

Figure 4. (A) Immobilization of biotinylated dsDNA on streptavidin affinity resin. (B) Gel mobility assay of PCR amplicons labeled with NHS-SS-PEG$_4$-biotin (P) compared to the unbound fraction remaining in the supernatant (S) after binding to the streptavidin solid support. Biotinylation levels are 53% for d**5SICS**-d**MMO2**$^{\mathbf{A}}$ and 70% for d**5SICS**$^{\mathbf{A}}$-d**NaM**. A 100 bp DNA ladder is loaded in the leftmost lane. (C) Conjugation of dsDNA to iodoacetyl-PEG$_2$-biotin after release from the streptavidin affinity resin via DTT treatment. Biotin incorporation levels are 89% for d**5SICS**-d**MMO2**$^{\mathbf{A}}$ and 53% for d**5SICS**$^{\mathbf{A}}$-d**NaM**. A 50 bp DNA ladder is loaded in the leftmost lane. In B and C, the faster migrating, strong band corresponds to dsDNA, while the slower migrating band corresponds to the 1:1 complex between dsDNA and streptavidin.

Under the conditions employed, no full length product was observed in the absence of an unnatural triphosphate or in the presence of only a **5SICS**TP derivative, and most of the truncated product corresponded to the termination of transcription immediately before d**5SICS** in the template (Figure 6B). In contrast, when either **MMO2**$^{\mathbf{PA}}$TP or **MMO2**$^{\mathbf{A}}$TP was present, we observed efficient conversion to full-length product. We then characterized transcription of the unnatural base pair in the opposite strand context by examining the ability of d**NaM** to template the transcription of RNA containing **5SICS**$^{\mathbf{PA}}$ or **5SICS**$^{\mathbf{A}}$. Again, in the absence of an unnatural triphosphate, virtually no full-length product was observed, but addition of either cognate unnatural triphosphate, **5SICS**$^{\mathbf{PA}}$TP or **5SICS**$^{\mathbf{A}}$TP, resulted in the efficient production of the full-length transcription product (Figure 6B). To characterize the fidelity of transcription, the experiments were again run with [α-^{32}P]ATP and the

Table 3. Nucleotide Composition Analysis of T7 RNAP Transcription Products[a]

		normalized composition of nucleotides incorporated 5′ to A during transcription of 17 nt RNA				
template	YTP	Ap	Gp	Cp	Up	Yp
d**5SICS**	d**MMO2**^{A}	1.02 ± 0.01 [1]	1.94 ± 0.03 [2]	n.d. [0][b]	n.d. [0][b]	1.03 ± 0.02 [1]
d**5SICS**	d**MMO2**^{PA}	0.99 ± 0.03 [1]	1.99 ± 0.03 [2]	n.d. [0][b]	n.d. [0][b]	1.01 ± 0.04 [1]
d**NaM**	d**5SICS**^{A}	1.03 ± 0.02 [1]	2.00 ± 0.01 [2]	n.d. [0][b]	n.d. [0][b]	0.96 ± 0.02 [1]
d**NaM**	d**5SICS**^{PA}	1.07 ± 0.01 [1]	1.95 ± 0.02 [2]	n.d. [0][b]	n.d. [0][b]	0.95 ± 0.01 [1]
		normalized composition of nucleotides incorporated 5′ to A during transcription of $tRNA^{Tyr}_{CUA}$				
template	YTP	Ap	Gp	Cp	Up	Yp
d**5SICS**	**MMO2**^{A}	4.06 ± 0.13 [4]	3.03 ± 0.07 [3]	5.91 ± 0.17 [6]	0.98 ± 0.03 [1]	1.02 ± 0.05 [1]
d**NaM**	**5SICS**^{A}	4.16 ± 0.16 [4]	3.10 ± 0.04 [3]	5.86 ± 0.06 [6]	1.01 ± 0.02 [1]	0.92 ± 0.04 [1]

[a] Values were determined by dividing the radioactivity observed for a particular monophosphate by the product of the total radioactivity and the expected number of nucleotides. Predicted values assuming 100% fidelity are shown in brackets. The error reported is the standard deviation of at least three independent determinations. [b] Below detection limit.

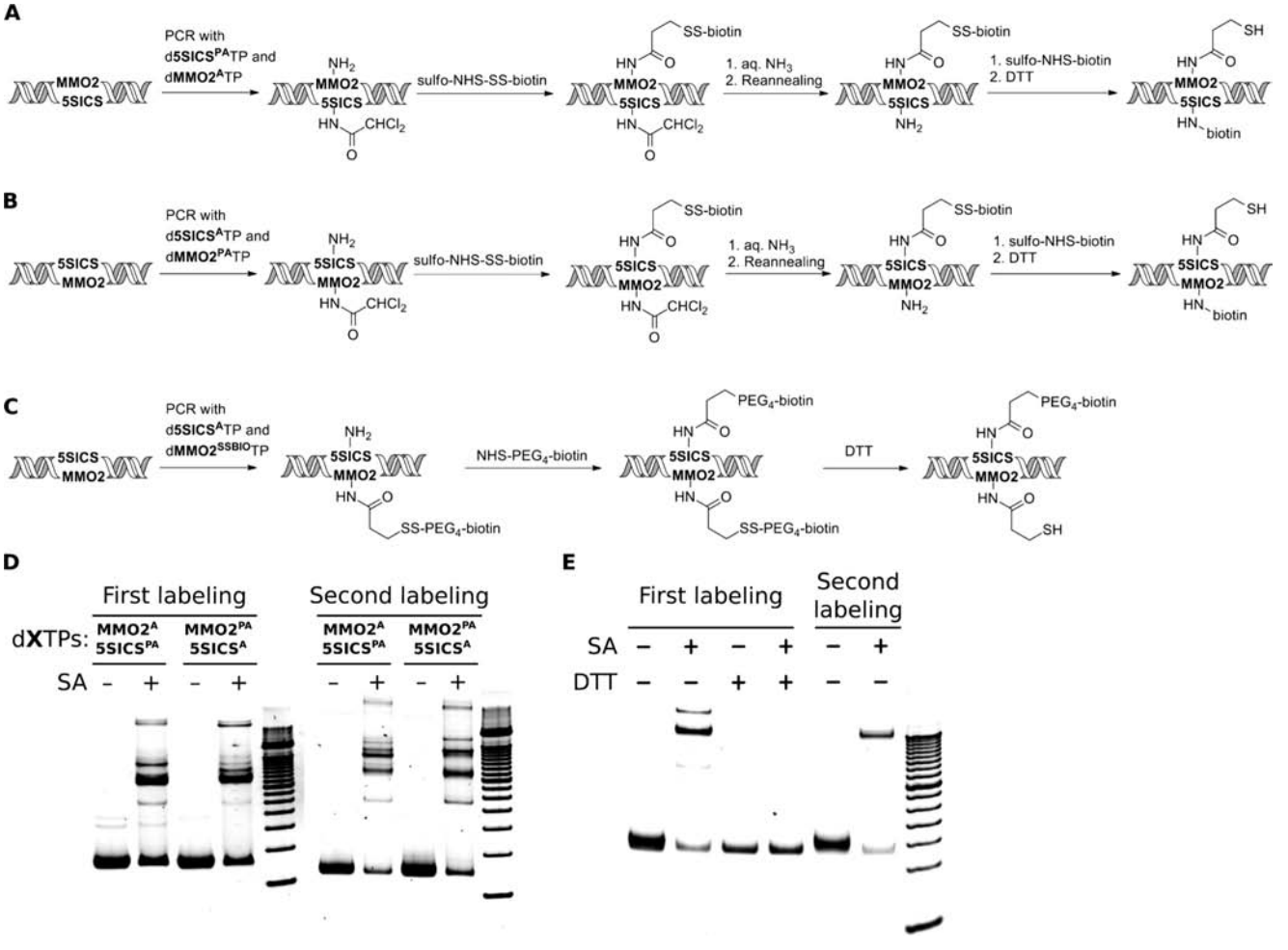

Figure 5. Postamplification DNA labeling with two different functional groups, using (A) d**5SICS**^{PA}-d**MMO2**^{A}, (B) d**5SICS**^{A}-d**MMO2**^{PA}, and (C) d**5SICS**^{A}-d**MMO2**^{SSBIO}. (D) Efficiencies of first and second labeling of DNA from containing d**5SICS**^{PA}-d**MMO2**^{A} and d**5SICS**^{A}-d**MMO2**^{PA}. First labeling efficiencies are 61% for d**5SICS**^{PA}-d**MMO2**^{A} and 66% for d**5SICS**^{A}-d**MMO2**^{PA}, and second labeling efficiencies are 85% for d**5SICS**^{PA}-d**MMO2**^{A} and 78% for d**5SICS**^{A}-d**MMO2**^{PA}. A 100 bp ladder is loaded in the rightmost lane of each gel. (E) First and second labeling efficiencies of DNA containing d**5SICS**^{A}-d**MMO2**^{SSBIO}. The first labeling efficiency is 85% in the absence of DTT and 0% in the presence of DTT (due to linker cleavage). The second labeling efficiency is 68%. A 50 bp DNA ladder is loaded in the rightmost lane. In D and E, the faster migrating, strong band corresponds to dsDNA, while the slower migrating band corresponds to the 1:1 complex between dsDNA and streptavidin. The faint and more slowly migrating bands correspond to higher order complexes of dsDNA and streptavidin.

resulting transcripts were digested to nucleoside 3′-phosphates, which were separated by 2D-TLC (Figure S2, Supporting Information), and the relative amount of each was quantified via densitometry (Table 3). Nucleotide-composition analysis confirmed that both unnatural nucleotides direct the transcription of RNA containing a cognate unnatural nucleotide with good selectivity. Interestingly, unlike with the replication of the deoxy-variants discussed above, we observed little difference in the efficiency or fidelity of **5SICS**^{A}TP or **5SICS**^{PA}TP incorporation.

To further explore the T7 RNAP-mediated transcription of RNA for more practical applications, we synthesized a 96-nt DNA fragment encoding the 77-nt *M. jannaschii* $tRNA^{Tyr}_{CUA}$ with either d**NaM** or d**5SICS** at the third position of the anticodon (nucleotide 37). Full length product was efficiently produced in the presence of each natural triphosphate and the cognate unnatural triphosphate. However, unlike with the shorter model template, we were unable to identify conditions where product formation depended only upon the addition of the correct unnatural ribotriphosphate, likely due to the higher concentrations of natural NTPs required for efficient transcription of the longer RNA. Thus, we characterized the fidelity of transcription of the DNA template containing d**NaM** with **5SICS**^{A} and of the template containing d**5SICS** with **MMO2**^{A} by 2D-TLC and densitometry (Table 3, and Figure S3, Supporting Information). The incorporation of **MMO2**^{A} opposite d**5SICS** proceeded with excellent fidelity, while the incorporation of **5SICS**^{A} opposite d**NaM** proceeded with a more modest fidelity of 92%, suggesting that as with the deoxy variants and DNA polymerases, the free amine is tolerated better in the **MMO2** scaffold than in the **5SICS** scaffold.

Finally, we explored the post-transcription site-specific labeling of RNA (Figure 7). Purified $tRNA^{Tyr}_{CUY}$ product with **Y** = **MMO2**^{A} at the third position of the anticodon loop was first unfolded by heating to 85 °C in the presence of 20% DMSO and then reacted with NHS-PEG_4-biotin at 37 °C for 1 h. Based on streptavidin gel shift, the efficiency of site-specific biotin attachment was greater than 75%.

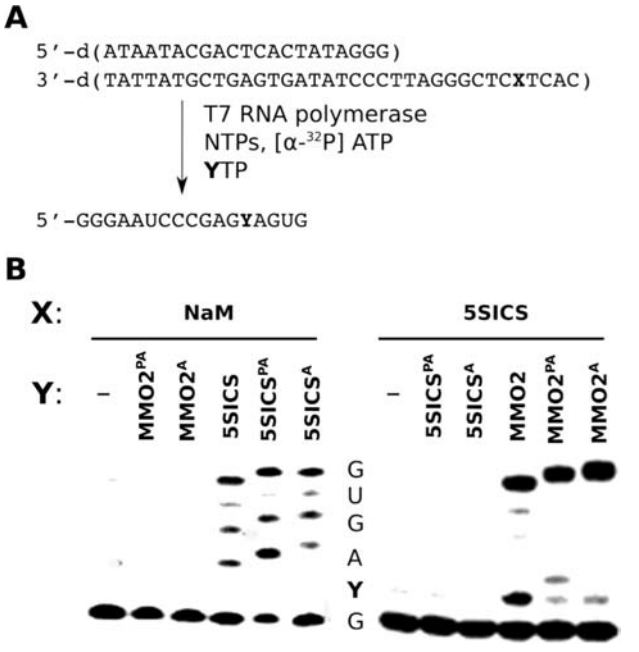

Figure 6. T7 RNAP transcription of site-selectively modified RNA. (A) Sequences of DNA template and transcription product. (B) Denaturing PAGE analysis with sequence indicated. Note that the bands for the linker-modified full length transcripts are shifted slightly more than those of the unlabeled full length product.

3. DISCUSSION

The development of an unnatural base pair is the first step toward creating semisynthetic organisms with increased potential for information storage and retrieval but would likely find more immediate use in different applications based on the production of site-specifically modified DNA or RNA. While functional unnatural base pairs that make site-specific labeling possible have recently been identified,[29,34,39,44,46] little is known about their ability to accommodate the linkers required to attach the various functional groups of interest. Presumably, just as the development of the unnatural base pairs required extensive optimization, so too will development of the linkers. Previously,

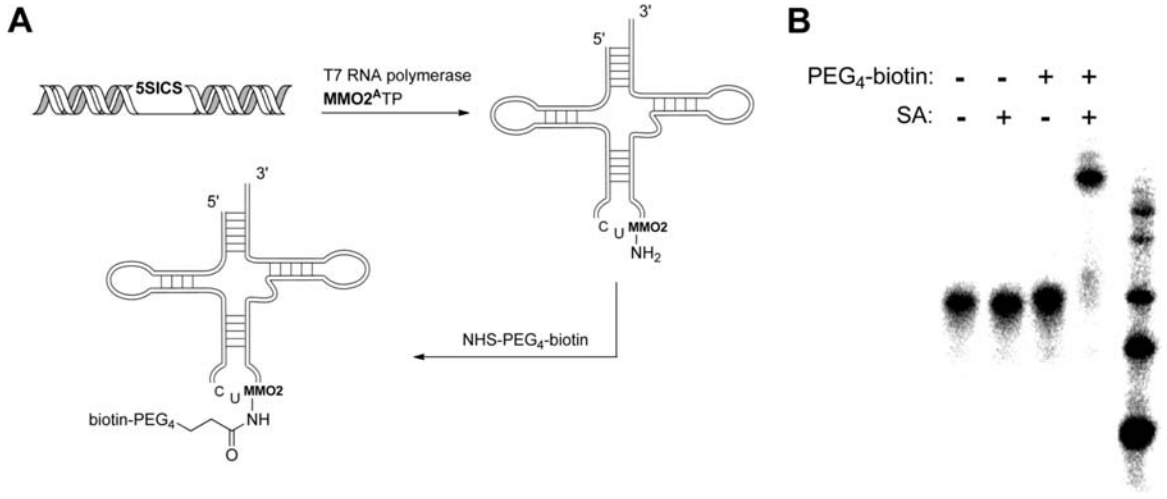

Figure 7. (A) Transcription of *M. jannaschii* $tRNA^{Tyr}_{CUA}$ site-selectively modified with **MMO2**^{A}. (B) PAGE analysis of the modified $tRNA^{Tyr}_{CUA}$ containing **MMO2**^{A} before (left two lanes) and after (right two lanes) coupling to NHS-PEG_4-biotin. Coupling efficiency calculated as greater than 75%. A 25 bp DNA ladder is loaded in the rightmost lane.

the Hirao group demonstrated the use of the same propargylamine linker used here, along with an acetamidohexanamide spacer, to attach a single fluorophore to a promising unnatural base pair within DNA, as well as to directly attach a biotin tag to a constituent nucleotide in RNA.[44,46] To further explore the potential of linker-modified unnatural base pairs to site-specifically label DNA or RNA with one or more different moieties, either during or after enzymatic synthesis, we have explored the replication and transcription of d**5SICS**-d**MMO2**/d**NaM** with unnatural triphosphates bearing propargylamine-based linkers or an α-phosphorothioate.

The steady-state kinetics with Kf polymerase revealed that replication of the unnatural base pair is sensitive to the structure of the linker, with the protected amino linker being the best accommodated. However, the effects of linker attachment were not the same for both nucleotides. While modification of d**MMO2**TP with the free or the dithiol-linked biotin-modified linker reduced the efficiency of insertion opposite d**5SICS** 4- to 10-fold, modification with the protected amino linker actually increased insertion efficiency by 2-fold. In contrast, modification of d**5SICS**TP with any of the linkers examined resulted in reduced insertion efficiency opposite d**NaM**. Relative to d**5SICS**TP, the insertion efficiency of d**5SICS**$^{\mathbf{BIO}}$TP and d**5SICS**$^{\mathbf{SSBIO}}$TP is 3- to 4-orders of magnitude reduced, predominantly because of a reduction in the apparent binding of the triphosphate. The decrease was even larger for d**5SICS**$^{\mathbf{A}}$TP because of effects on both the apparent binding and turnover. In contrast, the insertion efficiency of d**5SICS**$^{\mathbf{PA}}$TP opposite d**NaM** was reduced only 25-fold (because of reduced apparent binding).

Relative to the d**MMO2**, the greater sensitivity of d**5SICS**TP insertion to linker modification may result from its being more efficiently recognized, making it more susceptible to perturbation. However, the increase in insertion efficiency observed with d**MMO2**$^{\mathbf{PA}}$, and the fact that d**5SICS**$^{\mathbf{A}}$TP is inserted less efficiently opposite d**NaM** than d**MMO2**$^{\mathbf{A}}$TP is inserted opposite d**5SICS**, suggests that the d**MMO2** scaffold is more tolerant of the modifications examined. Perhaps, hydrophobic and packing interactions between the π-rich triple bond of the linker and the flanking nucleobases within the major groove are more favorable when accessed from the d**MMO2** scaffold than from the d**5SICS** scaffold. This is consistent with the recently reported structure of the ternary complex of the large fragment of *Taq* DNA polymerase, which is homologous to Kf, bound to C5-propargylamide-modified dTTP, which shows that the amide moiety of the linker forms a stabilizing hydrogen-bonding interaction with the side chain of polymerase residue R660.[24] This Arg residue is conserved in Kf, and because the ring structure of d**MMO2** is similar in size and shape to a natural pyrimidine, it likely engages in a similar stabilizing interaction with the propargylamide linker when presented within the d**MMO2** scaffold. Perhaps the extra aromatic ring of the d**5SICS** scaffold disrupts this stabilizing interaction. The same model likely accounts for the decreased tolerance of the free amine linker, as it is expected to be protonated and thus to introduce electrostatic repulsion with the positively charged side chain of R660.

Regardless of the detailed rates with which the different triphosphates are inserted opposite their cognate unnatural nucleotide in the template, PCR experiments with DeepVent clearly demonstrate that different combinations of the propargylamine-modified nucleotides may be efficiently and site-selectively incorporated into DNA with reasonable fidelity. Thus, it should be straightforward to incorporate many different functional groups into duplex DNA via amide linkage to either unnatural triphosphate. Moreover, different combinations of unnatural triphosphates bearing protected or free amines may also be incorporated into DNA via PCR and thereby provide a versatile route to the postsynthetic labeling of DNA with two different functional groups, although with d**5SICS**, the most efficient replication requires that the amine is incorporated in a protected form. Thus, the most simple route to the labeling of DNA with two different functionalities would be either the use of triphosphates already bearing the functional groups of interest or the amplification of DNA containing d**5SICS**$^{\mathbf{PA}}$-d**MMO2**$^{\mathbf{A}}$, followed by modification of d**MMO2**$^{\mathbf{A}}$, deprotection, and finally modification of d**5SICS**$^{\mathbf{A}}$. The use of d**5SICS**$^{\mathbf{SSBIO}}$-d**MMO2**$^{\mathbf{A}}$ also provides a convenient route to purify DNA containing the unnatural base pair, and the use of spacers that are longer and more hydrophilic to connect the linkers to the functional groups of interest facilitates post-amplification modification.

While the efficiency of d**5SICS**$^{\mathbf{PA}}$ insertion is likely sufficient for practical labeling applications, we nonetheless demonstrated that the incorporation of d**5SICS**$_{\alpha S}$TP opposite d**NaM**, which proceeds with a 2-fold greater efficiency, also provides a convenient route to site-specific labeling via the modification of the α-phosphorothioate. While amplification with d**5SICS**$^{\mathbf{PA}}$TP, d**5SICS**$_{\alpha S}$TP, or d**5SICS**TP all proceeded with similar efficiencies and fidelities within the sequence context examined, it is possible that other sequences may be more sensitive to the specific triphosphate used. Also note that the α-phosphorothioate- and linker-based strategies are not mutually exclusive and when combined should allow for a given site to be simultaneously modified with up to three different functional groups, one attached to the nucleobase of d**5SICS**, a second attached to the nucleobase of d**MMO2**, and a third attached to the backbone immediately 5′ to an unnatural nucleotide (Figure 8).

As observed with the deoxy variants and Kf DNA polymerase, the linker-modified unnatural ribotriphosphates were found to be substrates for T7 RNAP. While we only characterized the transcription of the protected- and free-amine modified ribotriphosphates, the data suggest that just as with their deoxy counterparts, the direct attachment of different functional groups to the ribotriphosphates should be possible, providing a direct route to the transcription of RNA labeled with one or two functional groups. In addition, we demonstrated that the incorporation of ribotriphosphates bearing free amine groups should provide a general route to the post-transcription labeling of RNA with different functionalities of interest (Figure 8). In contrast to the

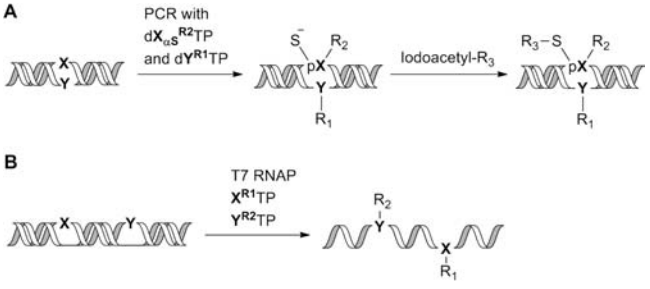

Figure 8. General scheme for labeling DNA with three different functional groups and RNA with two different functional groups. R_1 and R_2 may correspond to either amide-coupled functional groups of interest or a combination of protected and free amine linkers for further elaboration as described in the text.

differences in replication observed with d**5SICS**$^{\mathbf{A}}$TP and the other d**5SICS**TP derivatives, the transcription of short RNAs with **5SICS** and **5SICS**$^{\mathbf{A}}$ appears similar, although fidelity was somewhat reduced for the incorporation of **5SICS**$^{\mathbf{A}}$ into the tRNA, which is longer and more challenging to transcribe. Because alkylated α-phosphorothioates undergo strand cleavage in RNA, this route for backbone labeling is not possible, and the current methodology is limited to the site-specific labeling of RNA with only two different functional groups. However, the ability to induce site-specific RNA cleavage could also find interesting applications.[56,57]

Our goal is to expand the potential physical and functional properties of DNA and RNA without losing the inherent advantages possessed by these biopolymers relative to other materials. One property of DNA that has received much recent attention is its ability to act as a scaffold for the display of different functionalities with nanometer length-scale control for nanomaterial development.[2,3] The site-specific labeling approach described here seems well suited to contribute to these efforts, as many of the possible applications involve a level of distance control and functional group isolation that would be challenging or impossible to obtain with the higher functionalization density inherent to nucleotide-specific labeling approaches. From the perspective of creating (or evolving) controlled environments for molecular recognition or catalysis,[6,7] relative to the nucleotide-specific labeling approach, the site-specific approach is more analogous to that employed by proteins, where the position of one or only a few reactive centers are controlled within an environment provided by the remainder of the oligonucleotide. Moreover, including only one or a few modified nucleotides in different members of an oligonucleotide library is less likely to incur unnecessary replication biases than more heavily or fully functionalized DNA. Efforts to deploy the linker-modified unnatural base pairs in materials applications and for the selection of DNAs or RNAs with expanded functional potential are currently underway.

4. MATERIALS AND METHODS

4.1. General. All reactions were carried out in oven-dried glassware under inert atmosphere, and all solvents were dried over 4 Å molecular sieves with the exceptions of dichloromethane, which was distilled from CaH_2, and tetrahydrofuran, which was distilled from sodium metal. All other reagents were purchased from Aldrich and Acros. ^{1}H, ^{13}C, and ^{31}P NMR spectra were recorded on Varian Mercury 300, Varian Inova-400, or Bruker AMX-400 spectrometers. High-resolution mass spectroscopic data were obtained on an ESI-TOF mass spectrometer (Agilent 6200 Series) at the TSRI Open Access Mass Spectrometry Lab; MALDI-TOF mass spectrometry (Applied Biosystems Voyager DE-PRO System 6008) was performed at the TSRI Center for Protein and Nucleic Acid Research. Polynucleotide kinase and Kf were purchased from New England Biolabs, T7 RNAP from Takara USA, and [α-^{32}P]ATP and [γ-^{32}P]ATP from MP Biomedicals. DNA ladders were obtained from Invitrogen.

4.2. Oligonucleotide Synthesis. Fully natural oligonucleotides were purchased from Integrated DNA Technologies. Oligonucleotides containing an unnatural nucleotide were prepared by the β-cyanoethyl phosphoramidite method on controlled pore glass supports (1 μmol) using an Applied Biosystems Inc. 392 DNA/RNA synthesizer. After automated synthesis, the oligonucleotides were cleaved from the support and deprotected by heating in aqueous ammonia at 55 °C for 12 h. After purification via 8 M urea 15% polyacrylamide gel electrophoresis (PAGE), the oligonucleotides were electroeluted and ethanol precipitated. Oligonucleotide concentration was determined by UV absorption.

4.3. Kinetic Assay. Primer oligonucleotides were 5′-radiolabeled with [γ-^{32}P]ATP and T4 polynucleotide kinase and annealed to template oligonucleotides by heating to 95 °C followed by slow cooling to room temperature. Reactions were initiated by adding a solution of 2× dNTP solution (5 μL) to a solution containing Kf polymerase (0.10−1.23 nM) and primer template (40 nM) in 5 μL of Kf reaction buffer (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, and 50 μg/mL acetylated BSA). After incubation at 25 °C for 3 − 10 min, the reactions were quenched with 20 μL of loading dye (95% formamide, 20 mM EDTA, and sufficient amounts of bromophenol blue and xylene cyanol). Reaction products were resolved by 8 M urea 15% polyacrylamide gel electrophoresis, and gel band intensities corresponding to primer and extended primer were quantified by phosphorimaging (Storm Imager, Molecular Dynamics) and Quantity One (BioRad) software. Plots of k_{obs} versus triphosphate concentration were fit to the Michaelis−Menten equation using the program Origin (Microcal Software) to determine V_{max} and K_M. k_{cat} was determined from V_{max} by normalizing by the total enzyme concentration. Each reaction was run in triplicate, and standard deviations for both kinetic parameters were determined.

4.4. PCR Amplification. dsDNA template D1 was prepared as described previously[34] (see Supporting Information for details) and amplified by PCR under the following conditions: 1 ng of the template, 1× ThermoPol reaction buffer (New England Biolabs), $MgSO_4$ adjusted to 6.0 mM, 0.6 mM of dNTP, 0.2 mM of each unnatural triphosphates, 1 μM of each primers, and 0.03 U/μL of DeepVent (exo+, New England Biolabs) in an iCycler Thermal Cycler (Bio-Rad) with a total volume of 25 μL under the following thermal cycling conditions: 94 °C, 30 s; 48 °C, 30 s; 65 °C, 8 min, 14 cycles. Sybr Green I (Invitrogen) was added to the final concentration of 0.5× to monitor amplification by qPCR. Upon completion, a 5 μL aliquot was analyzed by 2% agarose gel, and the remaining material was purified utilizing the PureLink PCR Purification Kit (Invitrogen). Amplification efficiency was quantified by fluorescent dye binding (Quant-iT dsDNA HS Assay kit, Invitrogen), and fidelity was quantified by either sequencing (3730 DNA Analyzer, Applied Biosystems) (see Supporting Information and Malyshev et al.[34]) or gel mobility (see below).

4.5. Gel Mobility Assays. DNA samples (10−50 ng) were mixed with 1 μg of streptavidin (Promega) in phosphate labeling buffer (50 mM sodium phosphate, pH 7.5, 150 mM NaCl, 1 mM EDTA), incubated for 30 min at 37 °C, mixed with 5× nondenaturing loading buffer (Qiagen), and loaded on 10% nondenaturing PAGE. The gel was run at 180 V for 25−40 min, then soaked in 1× Sybr Gold Nucleic Acid Stain (Invitrogen) for 30 min and visualized using a Molecular Imager Gel Doc XR+ equipped with 520DF30 filter (Bio-Rad). Strong bands corresponding to dsDNA (at ∼150 bp) and the 1:1 complex between dsDNA and streptavidin (at ∼400 bp) were apparent. Faint bands corresponding to higher order (slower migrating) complexes of DNA and streptavidin or from unbiotinylated, single-stranded DNA resulting from incomplete annealing after PCR in some cases were also apparent.

4.6. General Procedures for Postamplification DNA Labeling. For postenzymatic synthesis labeling, dsDNA with a free amino group was incubated with 10 mM EZ-Link sulfo-NHS-SS-biotin or EZ-Link NHS-PEG_4-biotin (Thermo Scientific) for 1 h at rt in phosphate labeling buffer and then purified using the Qiagen PCR purification kit. With either d**MMO2**$^{\mathbf{PA}}$ or d**5SICS**$^{\mathbf{PA}}$, the amine first required deprotection, which was accomplished by overnight incubation in a concentrated aqueous ammonia solution at rt. Ammonia was removed via a SpeedVac concentrator (water aspirator followed by oil vacuum pump). To cleave the disulfide-containing linkers (i.e., SS-biotin or SS-PEG_4-biotin), dsDNA was treated with DTT (final concentration of 30 mM) for 1 h at 37 °C. For backbone labeling, dsDNA with a backbone phosphorothioate was incubated with 25 mM EZ-Link iodoacetyl-PEG_2-biotin (Thermo Scientific) in phosphate labeling buffer overnight at 50 °C, and products

were purified with Qiagen PCR Purification Kit. All reactions manipulating attached biotin moieties were quantified by streptavidin gel-shift assays.

4.7. DNA Immobilization on Streptavidin Solid Support. Streptavidin Sepharose High Performance affinity resin (GE Healthcare) was washed twice with phosphate labeling buffer to remove ethanol. Site-specifically biotinylated dsDNA was added to the prewashed resin and incubated with occasional gentle mixing. After 1 h, the supernatant was removed (and analyzed by gel shift mobility to confirm the absence of biotinylated DNA), and the resin was washed three times with phosphate labeling buffer to remove any unbound DNA. dsDNA was recovered via DTT treatment (final concentration of 30 mM) for 1 h at 37 °C, followed by filtration and purification with the Qiagen PCR Purification Kit.

4.8. Transcription Experiments. For the transcription of short, 17 nt RNAs, the 35 nt DNA template 5′-CACT**X**CTCGGGATTCCC-TATAGTGAGTCGTATTAT (**X** = d**5SICS** or d**NaM**) and 21 nt DNA primer 5′-ATAATACGACTCACTATAGGG were annealed in a 1:1 ratio in 10 mM Tris-HCl buffer (pH 7.6) containing 10 mM NaCl, by heating at 95 °C for 3 min and slow cooling to 4 °C. Transcription was carried out at 37 °C in 10 μL reactions containing 100 nM DNA primer-template, 20 μM NTP, 0.25 mCi [α-^{32}P]ATP, 50 U T7 RNAP (Takara), and 10 μM **MMO2**$^{\mathbf{PA}}$TP, **MMO2**$^{\mathbf{A}}$TP, **5SICS**$^{\mathbf{PA}}$TP, or **5SICS**$^{\mathbf{A}}$TP in 1× Takara buffer (40 mM Tris-HCl, pH 8.0, 8 mM $MgCl_2$, 2 mM spermidine). All solutions were prepared with DEPC-treated and nuclease-free sterilized water (Fisher Bioreagents). After incubation for 2 h at 37 °C, the reactions were quenched by addition of gel loading dye solution (10 μL of 10 M urea, 0.05% bromophenol blue). This mixture was heated at 75 °C for 3 min and then separated on a 7 M urea 20% polyacrylamide gel using 1× TBE buffer. The gel was removed from the apparatus, and radioactivity was quantified by phosphorimaging (overnight exposure) using ImageQuant (Quantity One).

For transcription of *M. jannaschii* $tRNA^{Tyr}_{CUY}$, the template strand 5′-TGG TCC GGC GGG CCG GAT TTG AAC CAG CGC CAT GCG GAT T**X**A GAG TCC GCC GTT CTG CCC TGC TGA ACT ACC GCC GGT ATA GTG AGT CGT ATT ATC (**X** = d**5SICS** or d**NaM**) and the nontemplate strand 5′- GAT AAT ACG ACT CAC TAT ACC GGC GGT AGT TCA GCA GGG CAG AAC GGC GGA CTC TAA ATC CGC ATG GCG CTG GTT CAA ATC CGG CCC GCC GGA CCA were annealed in Tris-HCl buffer (pH 7.6) containing 10 mM NaCl, by heating at 95 °C for 3 min and cooling to 4 °C. Transcription was carried out at 37 °C in 20 μL reactions containing 2 μM DNA primer-template, 1 mM NTP, 0.25 mCi [α-^{32}P]ATP (MP Biomedicals, LLC, Solon, OH), 50 U T7 RNAP (Takara Bio Inc.), and 1 mM **MMO2**$^{\mathbf{A}}$TP or **5SICS**$^{\mathbf{A}}$TP in 1× Takara buffer (40 mM Tris-HCl, pH 8.0, 8 mM $MgCl_2$, 2 mM spermidine). All solutions were prepared with DEPC treated and nuclease-free sterilized water (Fisher Bioreagents). After incubation for 3 h at 37 °C, the reaction was quenched by the addition of gel loading dye solution (20 μL of 10 M urea, 0.05% bromophenol blue). This mixture was heated at 75 °C for 3 min and then separated on a 7 M urea 12% polyacrylamide gel using 1× TBE buffer. The gel was removed from the apparatus, and radioactivity was quantified by phosphorimaging (overnight exposure) using ImageQuant (Molecular Dynamics).

Transcription fidelity was characterized by 2D TLC analysis. Briefly, gels containing transcription products were imaged (overnight exposure with Kodak X-OmatAR5 film), and the image was used to identify the radioactive spots corresponding to the full-length [^{32}P]-RNA transcripts. These portions of the gel were then excised and transferred to 2 mL microcentrifuge tubes for passive elution (6−12 h) at room temperature with 400 μL of sterilized water. Eluted full-length products were precipitated with cold ethanol (1.2 mL), CH_3COONa (0.3 M, 20 μL), and *E. coli* tRNA (1 mM, 1 μL, Sigma Aldrich) for 2 h. After centrifugation, the ethanol was removed and the product was evaporated to dryness with a SpeedVac concentrator. The transcripts were digested with RNaseI (10 U, 1 μL, Epicenter, WI) in 2 μL of 10× RNaseI dilution buffer in a final volume of 20 μL at 37 °C for 120 min. The digestion products were analyzed by 2D TLC using high performance thin layer chromatography (HPTLC) plates (100 × 100 mm) (Merck, Darmstadt, Germany) with the following developing solvents: isobutyric acid−ammonia−water (66:1:33 v/v/v) for the first dimension, and sodium phosphate (0.1 M, pH 6.8), ammonium sulfate, *n*-propanol (100:60:2 v/w/v) for the second dimension. The products on the gels and the TLC plates were analyzed by phosphorimaging using ImageQuant (Quantity One).

For post-transcriptional labeling of $tRNA^{Tyr}_{CUA}$, the purified full length RNA containing **MMO2**$^{\mathbf{A}}$ (60 μL, 2.2 ng/μL) was mixed with 15 μL of DMSO, heated at 85 °C for 5 min, and cooled immediately on ice. The resulting solution was incubated with EZ-Link NHS-PEG_4-biotin (Thermo Scientific) (1.5 mg) in 30 μL of 4× phosphate labeling buffer for 1 h at 37 °C. After ethanol precipitation, 10% of the resulting material (3 μL) was mixed with 3 μL of 5× denaturing loading buffer, heated to 85 °C, cooled on ice, mixed with 3 μL of streptavidin (1 mg/mL), and incubated for 30 min at 37 °C. The efficiency of labeling was determined by streptavidin gel shift on a 7 M urea 15% polyacrylamide gel using 1× TBE buffer (180 V for 25−40 min).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Nucleoside synthesis and characterization as well as details of kinetic experiments, sequencing, and transcription analysis.This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
floyd@scripps.edu

Author Contributions
[†]These authors contributed equally.

■ ACKNOWLEDGMENT

Funding for this work was provided by the National Institute for Health (GM060005).

■ REFERENCES

(1) Benner, S. A.; Sismour, A. M. *Nat. Rev. Genet.* **2005**, *6*, 533–543.
(2) Wang, H.; Yang, R.; Yang, L.; Tan, W. *ACS Nano* **2009**, *3*, 2451–2460.
(3) Chen, T.; Shukoor, M. I.; Chen, Y.; Yuan, Q.; Zhu, Z.; Zhao, Z.; Gulbakan, B.; Tan, W. *Nanoscale* **2011**, *3*, 546–556.
(4) Ruigrok, V. J.; Levisson, M.; Eppink, M. H.; Smidt, H.; van der Oost, J. *Biochem. J.* **2011**, *436*, 1–13.
(5) Heyduk, T. *Biophys. Chem.* **2010**, *151*, 91–95.
(6) Syed, M. A.; Pervaiz, S. *Oligonucleotides* **2010**, *20*, 215–224.
(7) Mayer, G. *Angew. Chem., Int. Ed.* **2009**, *48*, 2672–2689.
(8) Lonnberg, H. *Bioconjugate Chem.* **2009**, *20*, 1065–1094.
(9) Paredes, E.; Evans, M.; Das, S. R. *Methods* **2011**, *54*, 251–259.
(10) Lavergne, T.; Bertrand, J. R.; Vasseur, J. J.; Debart, F. *Chem.—Eur. J.* **2008**, *14*, 9135–9138.
(11) Battersby, T. R.; Ang, D. N.; Burgstaller, P.; Jurczyk, S. C.; Bowser, M. T.; Buchanan, D. D.; Kennedy, R. T.; Benner, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 9781–9789.
(12) Brakmann, S.; Lobermann, S. *Angew. Chem., Int. Ed.* **2001**, *40*, 1427–1429.
(13) Gourlain, T.; Sidorov, A.; Mignet, N.; Thorpe, S. J.; Lee, S. E.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1898–1905.

(14) Jager, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005**, *127*, 15071–15082.
(15) Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. *Nucleic Acids Res.* **2006**, *34*, 5383–5394.
(16) Mehedi Masud, M.; Ozaki-Nakamura, A.; Kuwahara, M.; Ozaki, H.; Sawai, H. *ChemBioChem* **2003**, *4*, 584–588.
(17) Perrin, D. M.; Garestier, T.; Helene, C. *Nucleosides Nucleotides* **1999**, *18*, 377–391.
(18) Roychowdhury, A.; Illangkoon, H.; Hendrickson, C. L.; Benner, S. A. *Org. Lett.* **2004**, *6*, 489–492.
(19) Sakthivel, K.; Barbas, C. F., III. *Angew. Chem., Int. Ed.* **1998**, *37*, 2872–2875.
(20) Tasara, T.; Angerer, B.; Damond, M.; Winter, H.; Dorhofer, S.; Hubscher, U.; Amacker, M. *Nucleic Acids Res.* **2003**, *31*, 2636–2646.
(21) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *Nucleic Acids Res.* **2009**, *37*, 1638–1649.
(22) Held, H. A.; Benner, S. A. *Nucleic Acids Res.* **2002**, *30*, 3857–3869.
(23) Lee, S. E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1565–1573.
(24) Obeid, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21327–21331.
(25) Zhu, Z.; Waggoner, A. S. *Cytometry* **1997**, *28*, 206–211.
(26) Horlacher, J.; Hottiger, M.; Podust, V. N.; Hübscher, U.; Benner, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6329–6333.
(27) Lutz, M. J.; Held, H. A.; Hottiger, M.; Hübscher, U.; Benner, S. A. *Nucleic Acids Res.* **1996**, *24*, 1308–1313.
(28) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33–37.
(29) Yang, Z.; Chen, F.; Chamberlin, S. G.; Benner, S. A. *Angew. Chem., Int. Ed.* **2010**, *49*, 177–180.
(30) Hwang, G. T.; Romesberg, F. E. *J. Am. Chem. Soc.* **2008**, *130*, 14872–14882.
(31) Leconte, A. M.; Hwang, G. T.; Matsuda, S.; Capek, P.; Hari, Y.; Romesberg, F. E. *J. Am. Chem. Soc.* **2008**, *130*, 2336–2343.
(32) Leconte, A. M.; Romesberg, F. E. In *Protein Engineering*; RajBhandary, C. K. a. U. L., Ed.; Springer-Verlag: Berlin, 2009; pp 291−314.
(33) Malyshev, D. A.; Pfaff, D. A.; Ippoliti, S. I.; Hwang, G. T.; Dwyer, T. J.; Romesberg, F. E. *Chem.—Eur. J.* **2010**, *16*, 12650–12659.
(34) Malyshev, D. A.; Seo, Y. J.; Ordoukhanian, P.; Romesberg, F. E. *J. Am. Chem. Soc.* **2009**, *131*, 14620–14621.
(35) Matsuda, S.; Fillo, J. D.; Henry, A. A.; Rai, P.; Wilkens, S. J.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Schultz, P. G.; Spraggon, G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 10466–10473.
(36) Matsuda, S.; Leconte, A. M.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 5551–5557.
(37) McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585–11586.
(38) Seo, Y. J.; Hwang, G. T.; Ordoukhanian, P.; Romesberg, F. E. *J. Am. Chem. Soc.* **2009**, *131*, 3246–3252.
(39) Seo, Y. J.; Matsuda, S.; Romesberg, F. E. *J. Am. Chem. Soc.* **2009**, *131*, 5046–5047.
(40) Seo, Y. J.; Romesberg, F. E. *ChemBioChem* **2009**, *10*, 2394–2400.
(41) Yu, C.; Henry, A. A.; Romesberg, F. E.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2002**, *41*, 3841–3844.
(42) Lavergne, T.; Malyshev, D. A.; Romesberg, F. E. *Chem. Eur. J.* **2011**, in press, DOI: 10.1002/chem.201102066.
(43) Hirao, I. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622–627.
(44) Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. *Nat. Methods* **2006**, *3*, 729–735.
(45) Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. *J. Am. Chem. Soc.* **2007**, *129*, 15549–15555.
(46) Kimoto, M.; Kawai, R.; Mitsui, T.; Yokoyama, S.; Hirao, I. *Nucleic Acids Res.* **2009**, *37*, e14.
(47) Mitsui, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S. *J. Am. Chem. Soc.* **2003**, *125*, 5298–5307.
(48) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79*, 413–444.
(49) Ludwig, J.; Eckstein, F. *J. Org. Chem.* **1989**, *54*, 631–635.
(50) Burgess, K.; Cook, D. *Chem. Rev.* **2000**, *100*, 2047–2060.
(51) Eckstein, F. *Annu. Rev. Biochem.* **1985**, *54*, 367–402.
(52) Kawai, R.; Kimoto, M.; Ikeda, S.; Mitsui, T.; Endo, M.; Yokoyama, S.; Hirao, I. *J. Am. Chem. Soc.* **2005**, *127*, 17286–17295.
(53) Rapley, R. *The Nucleic Acid Protocols Handbook*; Humana Press: New York, NY, 2000.
(54) Fidanza, J. A.; Ozaki, H.; McLaughlin, L. W. *Methods Mol. Biol.* **1994**, *26*, 121–143.
(55) Padmanabhan, S.; Coughlin, J. E.; Zhang, G.; Kirk, C. J.; Iyer, R. P. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1491–1494.
(56) Gish, G.; Eckstein, F. *Science* **1988**, *240*, 1520–1522.
(57) Strobel, S. A.; Shetty, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2903–2908.