Polymerase Reactions that Involve Modified Nucleotides

Masayasu Kuwahara, Kenta Hagiwara, and Hiroaki Ozaki

Contents

Abstract Polymerases are known to catalyze the synthesis of long DNA/RNA chains and play important roles in replication and translation in biological systems. Because their enzymatic activities are versatile, they have been widely employed for medical diagnoses and criminal investigations and also as research tools in biological studies. To date, various polymerases have been genetically engineered and are commercially available for selected applications that involve various chemically modified nucleoside triphosphate analogs. For example, dye-terminator sequencing, which made a great contribution to sequence determination in the human genome project, uses four kinds of $2',3'$-dideoxynucleoside-$5'$-triphosphate analogs that contain a fluorophore attached to the base moiety. Concomitantly, polymerase variants that can efficiently accept those analogs as substrates were developed. Various triphosphate analogs modified at their base/sugar/phosphate moieties have been designed and synthesized for the development of nucleic acid aptamers as therapeutic drugs, diagnostic agents, and molecular

M. Kuwahara (✉) • K. Hagiwara • H. Ozaki
Graduate School of Science and Technology, Gunma University, 1-5-1 Tenjin-cho, Kiryu, Gunma 376-8515, Japan
e-mail: mkuwa@gunma-u.ac.jp

indicators. Because nucleoside analogs include xenonucleic acids (XNAs), which have unique sugar backbones, drastic modifications in polymerase engineering are being boldly pursued. In this chapter, we are focusing on polymerase reactions that involve chemically modified substrates and their applications to innovative life sciences and technologies.

1 Introduction

While adenine (A), guanine (G), cytosine (C), and thymine/uracil (T/U) are well known as the standard bases of DNA/RNA, modified bases can also be found in messenger, transfer, and ribosomal RNAs (Fig. 1a). To date, more than 100 kinds of modified bases, including 7-methyl guanine, pseudouracil, dihydrouracil, hypoxanthine, and large bases with amino acid adducts, have been reported (Limbach et al. 1994). Sugar modifications, such as 2′-methoxy (OMe) and 1″,3-(5-″-phosphoryl)-ribosyl groups, have also been discovered. These modifications are considered to stabilize the three-dimensional structures of RNAs, allowing them to evade facile degradation by nucleases. Modified nucleotides are generally synthesized by modifying enzymes, such as methyltransferases and pseudouridine synthases, after transcription. For example, a methyltransferase initiates a nucleophilic attack by the thiolate (-S$^-$) of the cysteine residue at position 375 (Cys 375)
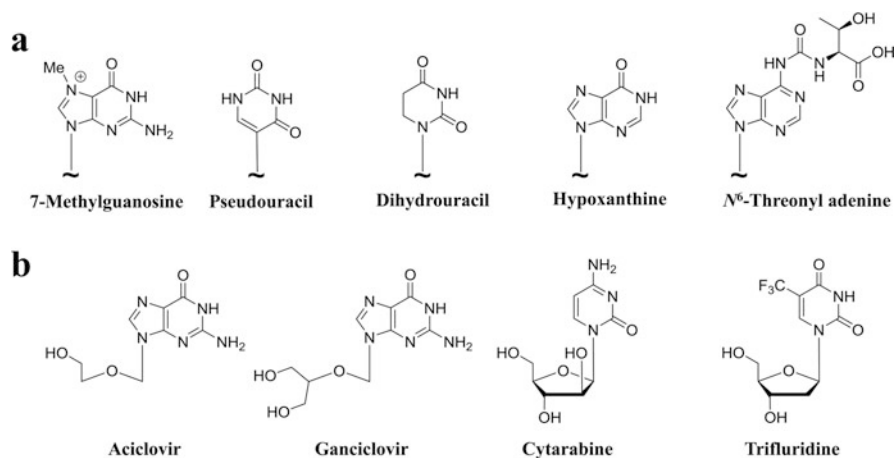
Fig. 1 (**a**) Modified bases found in messenger, transfer, and ribosomal RNAs; (**b**) polymerase inhibitors for antiviral drugs and antineoplastic agents

on the C6 position of cytosine, forming a Michael addition product that activates the C5 position of cytosine. Subsequent transfer of a methyl group from cofactor S-adenosyl-L-methionine yields 5-methylcytidine (Boschi-Muller and Motorin 2013).

Artificially synthesized nucleoside analogs, which were actively researched and developed from the 1960s to 1970s as polymerase inhibitors, are well known. These analogs, which include antiviral drugs (e.g., aciclovir and ganciclovir) and antineoplastic agents (e.g., cytarabine and trifluridine) (Fig. 1b), are phosphorylated through cellular metabolism and thereafter specifically inhibit polymerase activities to suppress virus proliferation and tumor growth (De Clercq 2001; Kufe et al. 1980). The other typical examples are fluorescence-labeled nucleoside triphosphate analogs, which are used in sequencing and microarray technologies for genetic research. Similar to fluorescent labeling, the introduction of foreign substituents to confer additional functionality, such as cell membrane permeability, nuclease resistance, and electroconductivity, has been studied for diverse applications.

2 Kinds of Polymerases

In general, DNA polymerases can extend a poly(oligo)nucleotide strand with a sequence complimentary to that of its template strand using nucleoside 5′-triphosphates as substrates. Some DNA polymerases have a proofreading 3′ to 5′ exonuclease activity. Since the time A. Kornburg discovered *E. coli* DNA polymerase I, various types of DNA polymerases have been isolated from prokaryotes, archaea, eukaryotes, and retroviruses. They are mainly classified into the following eight evolutionary families: A, B, C, D, E, X, Y, and reverse transcriptase (RT). For example, *E. coli* DNA polymerase I and mitochondrial DNA polymerase γ belong to family A; eukaryotic DNA polymerases α, δ, ε, and *E. coli* DNA polymerase II (pol II) are classified into family B; and *E. coli* DNA polymerase III (pol III) is a member of family C. Avian myeloblastosis virus (AMV) reverse transcriptase, which is an RNA-dependent DNA polymerase, can be mentioned as an example of an RT.

Highly thermostable DNA polymerases show optimal activity at temperatures around 75 °C. They are widely employed for polymerase chain reaction (PCR) to amplify/replicate DNA target sequences. To date, various engineered polymerases have been developed and supplied by manufacturers. One of the most popular thermostable DNA polymerases is *Taq* DNA polymerase, which belongs to family A and possesses 5′ to 3′ exonuclease activity, but not 3′ to 5′ exonuclease activity. It was isolated by A. Chien in 1976 (before the first report of the PCR technique in 1987) from an extremely thermophilic bacterium, *Thermus aquaticus* YT1, that was first discovered in the lower geyser basin of Yellowstone National Park, USA, in 1969 (Chien et al. 1976). The other examples of thermostable family A DNA polymerases are *AmpliTaq* and *Tth* DNA polymerases. The former, which was reported by F. C. Lawyer in 1989, is a modified form of *Taq* DNA polymerase obtained by expression of the gene in an *E. coli* host (Lawyer et al. 1989). The

latter, which was isolated from the extremely thermophilic bacterium *Thermus thermophilus* HB8, is used for reverse transcription (RT)-PCR owing to its very efficient RT activity in the presence of Mg^{2+} ion (Rüttimann et al. 1985). Meanwhile, *rTth* DNA polymerase, a recombinant DNA polymerase, is known to exhibit efficient RT activity in the presence of Mn^{2+} ion.

Hyperthermostable DNA polymerases that belong to family B generally exhibit excellent proofreading ability during DNA chain extension, owing to their 3′ to 5′ exonuclease activity. These polymerases, such as *Vent* polymerase from *Thermococcus litoralis*, *KOD* from *Thermococcus kodakaraensis* KOD1, $9°N_m$ from *Thermococcus* species 9°N-7, *Tgo* from *Thermococcus gorgonarius* (Miroshnichenko et al. 1998), *Deep vent* from *Pyrococcus* species GB-D (Jannasch et al. 1992), *Pfu* from *Pyrococcus furiosus* (Fiala and Stetter 1986), and *Pwo* from *Pyrococcus woesei* (Rüdiger et al. 1995), were obtained from hyperthermophilic archaea found in deep ocean vents, volcanic marine mud, and solfataras on the seashore. Owing to their higher fidelity (i.e., lower error rate) and heat stability, compared with those of *Taq*, they and their variants have been well studied, and some of them are commercially available. *Vent* DNA polymerase, which was reported in 1991, was the first thermostable DNA polymerase having a 3′ to 5′ proofreading exonuclease activity (Mattila et al. 1991). Its D141A and E143A variant, which was engineered to eliminate the exonuclease activity and is known as *Vent*(exo-) DNA polymerase, was developed to improve the yield of PCR products and to be applied to dideoxy sequencing reactions (Kong et al. 1993). *KOD* DNA polymerase not only possesses excellent proofreading ability, with about 50-fold higher fidelity than *Taq* DNA polymerase, but can also elongate an oligonucleotide strand with about a fivefold higher reaction rate than the other family B DNA polymerases, such as *Pfu* and *Deep vent* DNA polymerases (Takagi et al. 1997). The N210D variant of *KOD*, known as *KOD*(exo-) DNA polymerase, which possesses one thousandth of the 3′ to 5′ exonuclease activity of *KOD* DNA polymerase, has also been developed (Nishioka et al. 2001). The blend of *KOD* and *KOD*(exo-) DNA polymerases, which is commercially available under the product names *KOD Dash* or *KOD XL*, enables the production of long double-stranded DNAs (~15 kbp) by PCR. *KOD FX* and *AccuPrimePfx* DNA polymerases were developed from *KOD* DNA polymerase for use in hot start PCR. This is an improved PCR technique that can evade nonspecific amplification of DNA at lower temperatures by inactivating the polymerase with its antibody. The $9°N_m$ DNA polymerase is the E143D variant of wild-type 9°N-7 DNA polymerase (Southworth et al. 1996). This $9°N_m$ variant exhibits reduced 3′ to 5′ exonuclease activity (0.4–5 % of wild-type exonuclease activity). Therminator DNA polymerase, the D141A, E143A, and A485L variant of 9°N-7 DNA polymerase, has no 3′ to 5′ exonuclease activity. This DNA polymerase exhibits enhanced incorporation of modified nucleotides, that is, efficient single-base incorporation of dideoxy and acyclonucleotides. Thus, various B family DNA polymerases have been discovered and genetically engineered to be applied in a number of applications.

In addition to the highly thermostable and hyperthermostable DNA polymerases, *E. coli* DNA polymerase I, Klenow fragment (KF), *Bst* DNA polymerase, φ29 DNA

polymerase, and T7 DNA polymerase are well known. They display optimal activity at 30–37 °C, except for *Bst* DNA polymerase, which has an optimal temperature of 60–65 °C. KF is obtained as a large protein fragment by enzymatic cleavage of *E. coli* DNA polymerase I using the protease subtilisin (Klenow and Henningsen 1970). Although KF has lost its 5′ to 3′ exonuclease activity, it retains its 3′ to 5′ exonuclease activity; furthermore, its D355A and E357A variant, known as KF (3′–5′ exo-), lacks 3′ to 5′ exonuclease activity (Bebenek et al. 1990). *Bst* DNA polymerase is the large fragment of DNA polymerase I from the thermophilic bacterium *Bacillus stearothermophilus*, which is found in soil, hot springs, and ocean sediment and is generally unable to grow at temperatures below 35 °C (Kiefer et al. 1998). *Bst* DNA polymerase lacks 3′ to 5′ exonuclease activity, but possesses strong strand displacement activity, which allows isothermal DNA amplifications such as loop-mediated amplification (LAMP) and rolling circle amplification (RCA). Meanwhile, φ29 DNA polymerase, which is derived from the *Bacillus subtilis* phage phi29 (Φ29), is known to exhibit 3′ to 5′ proofreading exonuclease activity and extreme processivity, in addition to strong strand displacement activity (Blanco et al. 1989). Bacteriophage T7 DNA polymerase, which belongs to family A, also possesses 3′ to 5′ proofreading exonuclease activity (Grippo and Richardson 1971; Campbell et al. 1978; Adler and Modrich 1979). This DNA polymerase is known to be a complex comprising phage-encoded gene 5 protein and *E. coli* host thioredoxin, which enhances the processivity of the polymerase.

In eukaryotes, there are three main types of RNA polymerase: RNA polymerase I, which transcribes ribosomal RNA but not 5S rRNA; RNA polymerase II, which transcribes precursors of mRNA, snRNA, and microRNA; and RNA polymerase III, which transcribes ribosomal 5S rRNA, tRNA, and other small RNAs. Bacteriophage T7 (Davanloo et al. 1984), T3 (Majumder et al. 1979), and SP6 (Kotani et al. 1987) RNA polymerases can be mentioned as examples of commercially available DNA-dependent RNA polymerases. These RNA polymerases bind to their cognate promoters with very high sequence specificities and thereafter transcribe the DNA template downstream of the promoter to generate the complimentary single-stranded RNA. Some variants of T7 RNA polymerase have been developed to enhance the incorporation of modified nucleotides.

3   Enzymatic Synthesis of Modified Nucleotides

3.1   Base Modification

A variety of base-modified nucleotides, particularly C5-modified uridine analogs, have been reported to date (Fig. 2). Research results have demonstrated that nucleoside triphosphates bearing a C5-modified pyrimidine ring or a C7-modified 7-deazapurine ring can be more acceptable polymerase substrates than nucleoside
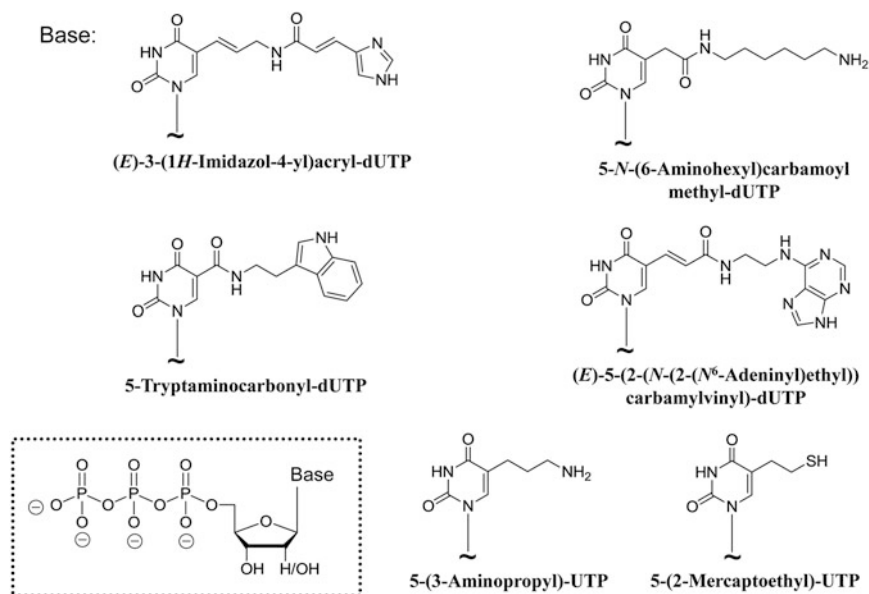
Fig. 2 Base-modified nucleoside triphosphates

triphosphates with substituents introduced at the other positions. Using these modified analogs in the presence of the four natural nucleoside triphosphates (dNTPs) allows foreign functionalities, such as fluorophores or biotin to be thinly incorporated into nucleic acid strands that can serve as probes or capture specific nucleotide targets. For example, in 1992, T. Ried et al. reported fluorescence in situ hybridization (FISH) probes, which were synthesized using *AmpliTaq* DNA polymerase-catalyzed PCR that incorporated dUTPs labeled with fluorescein, biotin, digoxigenin, and 2,4-dinitrophenol; these modified DNA probes were used to specifically visualize the centromeres of human chromosomes (Ried et al. 1992).

In 1998, K. Sakthivel et al. first reported the enzymatic synthesis of modified DNAs in which foreign functionalities were densely incorporated during PCR. They used ten different C5-modified 2′-deoxyuridine-5′-triphosphates (dUTPs) in the presence of the natural dNTPs, except for thymidine-5′-triphosphate (TTP) (Sakthivel and Barbas III 1998). Four different DNA polymerases (*Taq*, *Vent*, *Pfu*, and *rTth*) were examined. The C5-modified dUTP bearing an (*E*)-3-(1*H*-imidazol-4-yl)acryl group was accepted as a good PCR substrate for all of these DNA polymerases and provided the corresponding modified DNA very efficiently. The experimental results particularly emphasized that the rigid and extended α,-β-unsaturated arm adjacent to the imidazoyl group has a great influence on the incorporation efficiency of the modified dUs. Furthermore, modified UTPs bearing 5-(3-aminopropyl) or 5-(2-mercaptoethyl) groups were also synthesized and densely incorporated into RNAs with T7 RNA polymerase (Vaish et al. 2000).

Using these analogs instead of UTP, it should be possible to perform modified RNA aptamer/ribozyme selections.

In 2001, D. M. Williams and coworkers focused on the linker length and rigidity of substituents and systematically analyzed the substrate properties of ten different C5-modified dUTPs in the absence of TTP during PCR catalyzed by *Taq* DNA polymerase (Lee et al. 2001). The data showed that C5-modified dUTPs with linker arms containing rigid alkynyl and *trans*-alkenyl groups in the vicinity of uracil base were significantly superior to those with linker arms containing *cis*-alkenyl or alkyl groups. Furthermore, they synthesized C7-modified 7-deaza-dATP analogs with alkynyl, *cis*-alkenyl, and alkyl linker arms and found that the analog with the alkynyl linker arm acts as the best substrate for PCR catalyzed by *Taq* DNA polymerase (Gourlain et al. 2001). Intriguingly, using the preferred dUTP and 7-deaza-dATP analogs, modified dUs and 7-deaza-dAs could be simultaneously incorporated during PCR. In the same year, H. Sawai et al. first demonstrated, using PCR experiments incorporating dUTPs modified with a methylene linker at C5, that *KOD Dash* DNA polymerase is one of the most promising candidates as a catalyst for enzymatic syntheses of modified DNAs (Sawai et al. 2001). Indeed, *KOD* DNA polymerase-related products (e.g., *KOD XL*, *KOD FX*, *AccuPrimePfx*, and *KOD Dash*) are widely employed in applications involving enzymatic syntheses of modified DNAs.

In 2002, H. A. Held et al. demonstrated that *Pfu* and *Pwo* DNA polymerases, which are members of family B, as well as *KOD* DNA polymerase, can efficiently produce modified DNAs by primer extension (PEX) and PCR, using dUTPs modified at C5 with protected thiol groups (Held and Benner 2002). In successive incorporations of modified dUs, family B polymerases (*Pfu*, *Pwo*, *Vent*, and *Deep vent*) were found to be preferable to family A polymerases (*Taq*, *Tfl*, Hot *Tub*, and *Tth*). Thereafter, M. Kuwahara et al. systematically analyzed the PCR-based synthesis of modified DNA strands using C5-modified dUTPs and C5-modified dCTPs as substrates and family A (*Taq*, *Tth*) and B [*Vent*(*exo-*), *KOD Dash*, and *KOD*(*exo-*)] polymerases as enzymes and arrived at the same conclusion (Kuwahara et al. 2006). Furthermore, their kinetic studies using modified primers/templates/substrates revealed that modified group(s) adjacent to the extending terminus of the primer can greatly reduce catalytic efficiency, which resulted in low product yields with successive modified nucleotide incorporations.

D. M. Williams and coworkers applied the dual modification technique to modified DNAzyme selection, using modified dU and 7-deaza-dA analogs (Sidorov et al. 2004). Meanwhile, in 1999, D. M. Perrin et al. reported dual modification using C5-modified dUTP and C8-modified dATP in a PEX technique catalyzed using Sequenase Version 2.0 DNA polymerase, which is a genetically engineered form of T7 DNA polymerase with virtually no $3'$ to $5'$ exonuclease activity (Perrin et al. 1999; 2001). Using this technique, followed by in vitro selection, they were the first to produce a modified DNAzyme that acted as a metal-independent RNAase A mimic with two different functional groups ($k_{cat} = 0.044$ min^{-1}) (Perrin et al. 2001). They also produced a modified DNAzyme with three different functional groups (i.e., amine, imidazole, and guanidine) which had an improved k_{cat}

value (0.134 min^{-1}) (Hollenstein et al. 2009). In 2003, T. Tasara et al. synthesized C5-modified dUTPs, C5-modified dCTPs, C7-modified 7-deaza-dATPs, and C7-modified 7-deaza-dGTPs with foreign functional groups, including biotin, Rhodamine Green, Cyanine 5, Evoblue 30, and Gnothis Blue 3 (Tasara et al. 2003). They examined successive incorporations of modified nucleotides in the absence of all four natural dNTPs using PEX. A reaction using *Vent(exo-)* polymerase and biotinyl dNTPs with four different bases provided a modified DNA with a 40-mer elongated strand, in which biotinyl dNs were successively incorporated. Similarly, in 2005, S. Jäger et al. synthesized modified dNTPs with four different bases possessing basic, acidic, and lipophilic substituents and examined their successive incorporations (Jäger et al. 2005). Full-length modified DNAs comprising 40-mer elongated strands were produced by successive modified dN incorporations using PEX catalyzed by *Pwo* DNA polymerase. Furthermore, these oligonucleotides, which were modified at high density, were found to be reverse-transcribed to natural cDNAs in a PEX procedure using *Pwo* DNA polymerase, four natural dNTPs, and a GC-rich reaction buffer containing 1.5 mM dimethyl sulfoxide. Eventually, the technical challenges in successive incorporations have encouraged researchers to explore various applications of enzymatically synthesized modified DNAs.

Postsynthetic derivatization is a convenient and alternative method for the preparation of long DNA strands with high density and/or bulky modifications (Fig. 3). In 2003, H. Sawai and coworkers reported the enzymatic synthesis of a 108-mer modified DNA containing 5-methoxycarbonylmethyl-dUs in place of natural T using PCR and its subsequent derivatizations via amide bond formation by ester–amide exchange reactions using amino functionalities such as tris (2-aminoethylamine), histamine, and hexamethylenediamine (Mehedi Masud et al. 2003). Conversion rates of the derivatizations were unfortunately not very high, i.e., 56 %, 72 %, and 76 %, respectively, owing to the hydrolysis of the methyl ester that occurred during the addition reactions. Thereafter, various postsynthetic derivatizations using different chemistries have been developed. For example, pyranosyl sugar-modified DNA derivatized from a 300-mer precursory modified DNA containing 5-ethynyl-dUs by Cu-catalyzed alkyne–azide cycloaddition, i.e., a "click" reaction (Gierlich et al. 2007); biotin-modified DNA derivatized from a 35-mer precursory modified DNA containing a 7-[5-{(4-azidobutyl)amino}-5-oxopent-1-yn-1-yl]-7-deaza-dA by Staudinger ligation using biotinylated phosphine (Weisbrod and Marx 2007); hydrazone-modified DNA derivatized from a 98-mer precursory modified DNA containing 5-(5-formylthiophene-2-yl)-dCs by condensation reactions of aldehydes with arylhydrazines (Raindlová et al. 2010); and biotin-modified DNA derivatized from a 414-mer precursory modified DNA containing 7-vinyl-7-deaza-dAs by inverse electron demand Diels–Alder reaction (Bußkamp et al. 2014). The precursory modified DNAs were prepared in good yields by PCR/PEX preferably using *KOD XL*, *Pwo*, and *KlenTaq* DNA polymerases; some postsynthetic derivatizations exhibited significantly improved conversion rates, i.e., >90 %.
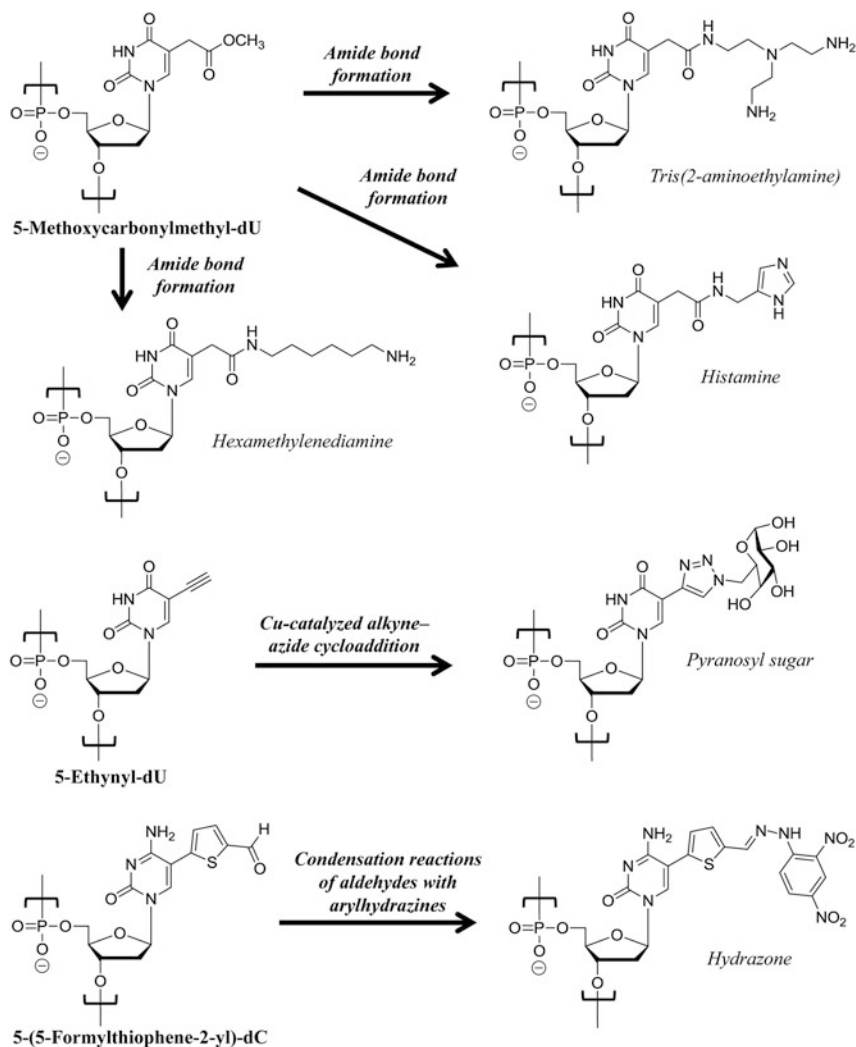
Fig. 3 Base modifications employed for postsynthetic derivatization

3.2 Special Base Modification (Artificial Base Pairs)

Artificial base pairs have great potential for expanding the genetic information system of life on earth. Such ambitious studies were considered to be able to pursue the basic principles of life phenomena and origin of life as well as technological applications.

In the late 1980s, S. A. Benner et al. first proposed expanded DNA alphabets (Switzer et al. 1989). In 1990, they demonstrated that artificial base pair formation,
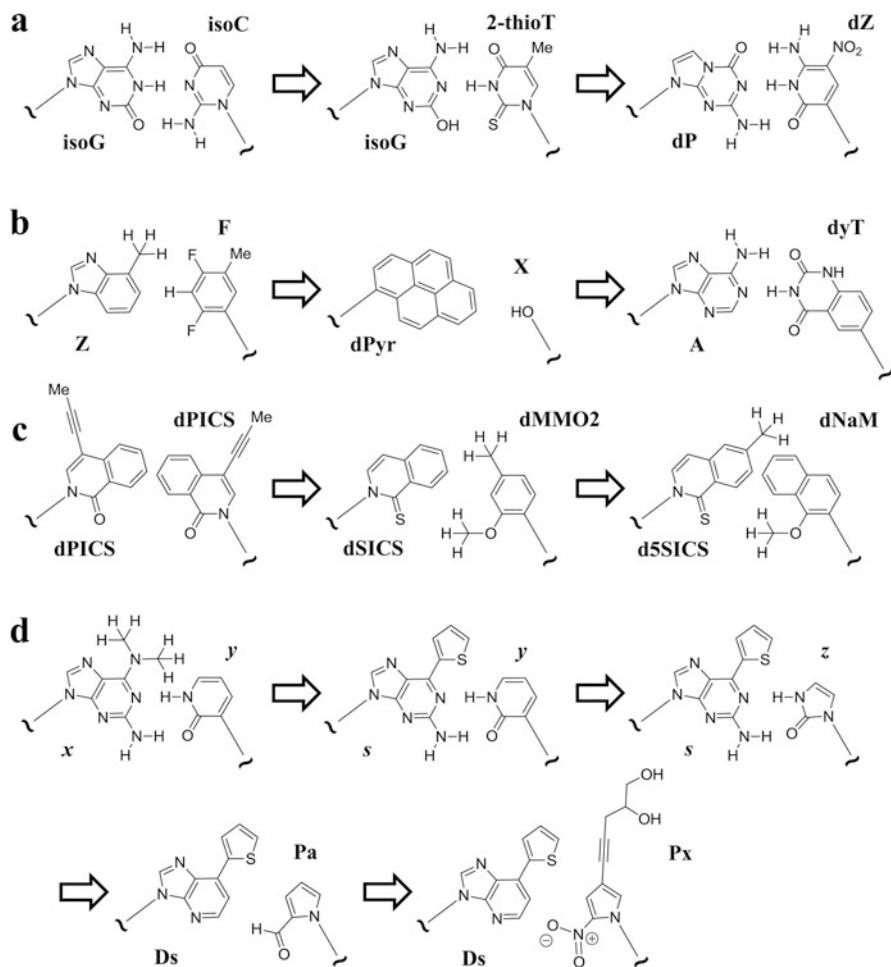
Fig. 4 Artificial base pairs reported by (**a**) Benner's, (**b**) Kool's, (**c**) Romesberg's, and (**d**) Hirao's groups

i.e., isoG–isoC (Fig. 4a), which differs from standard Watson–Crick base pairs, i.e., A–T(U) and G–C in the pattern of hydrogen bond formation, is possible during enzymatic DNA and RNA syntheses catalyzed by KF and T7 RNA polymerases, respectively (Piccirilli et al. 1990). Unfortunately, the incorporation of T opposite the enolic tautomer of isoG was also observed. To exclude the formation of the undesirable isoG–T pairing, they used 2-thioTTP instead of TTP, in addition to the other five triphosphate substrates, i.e., dATP, dGTP, dCTP, isoGTP, and isoCTP. As expected, the correct incorporation of isoC at a single isoG on a 51-mer template during PCR using *KlenTaq* DNA polymerase increased the value of fidelity-per-round in PCR from approximately 93 % to 98 % (Sismour and Benner 2005).

Furthermore, in 2006, they designed a novel third base pair comprising 2-amino-8-(2′-deoxy-β-D-erythro-pentofuranosyl)-imidazo[1,2-a]-1,3,5-triazin-4(8H)-one (dP) and 6-amino-5-nitro-3-(1′-β-D-2′-deoxyribofuranosyl)-2(1H)-pyridone (dZ) (Yang et al. 2006). PCR experiments using a 51-mer template containing a single dP residue showed high incorporation accuracies without using 2-thioTTP when dPTP and dZTP were used with the four natural dNTPs. The estimated values for fidelity-per-round were 94.4 %, 97.5 %, and 97.5 %, respectively, when *Taq*, *Vent* (*exo-*), and *Deep vent*(*exo-*) DNA polymerases were used (Yang et al. 2007). In the meantime, E. T. Kool, F. E. Romesberg, and I. Hirao were independently designing and developing their own candidates for the third base pair.

In 1998, E. T. Kool and coworkers reported a F(2,4-difluorotoluene)–Z (4-methylbenzimidazole) base pair (Fig. 4b), which does not form hydrogen bonds between the bases (Morales and Kool 1998). They demonstrated that efficiencies for a single incorporation of the Z residue opposite the F residue on a 28-mer DNA template were 130–1900-fold greater than those of the four natural dNs when KF (3′–5′ exo-) was used. Their results indicate that hydrogen bonding is not necessary for base pair formation, and, if anything, the size and shape of paired bases are more important to the adoption of artificial base pairs by an oligonucleotide duplex. Indeed, thereafter, it was demonstrated that pyrene deoxynucleoside (dPyr) as a nonpolar base analog can selectively be incorporated opposite an abasic nucleoside (X) or a tetrahydrofuran abasic analog (φ) to form a dPyr–X/φ pair during PEX (Matray and Kool 1999). Moreover, they developed "yDNA" (an abbreviation of "wide DNA"), which involves benzopyrimidine deoxynucleosides (dyT and dyC) bearing size-expanded pyrimidines, i.e., yT and yC, designed to form yT–A and yC–G pairs (Lee and Kool 2005). They then examined whether or not dyT and dyC can store and transfer genetic information in vitro and in bacterial cells (Chelliserrykattil et al. 2008). The results showed that the correct nucleotides could be inserted opposite yDNA residues in PEX using KF (3′–5′ exo-) and *Vent*(*exo-*) DNA polymerases. Furthermore, the first example of an encoding protein (GFP; green fluorescent protein) in a living organism, i.e., *E. coli*, by unnatural DNA base pair architecture was exhibited in 2008.

In 1999, F. E. Romesberg and coworkers reported that a stable 7-propynyl isocarbostyril nucleoside (dPICS) self-pair can be formed in duplex DNA (Fig. 4c), and dPICS triphosphate can be incorporated opposite dPICS on the template by KF (3′–5′ exo-) with reasonable efficiency (McMinn et al. 1999). However, after the dPICS incorporation, synthesis proceeded inefficiently. Thereafter, they determined the best pair from the 3600 (60×60) combinations of unnatural DNA base analogs, i.e., dSICS–dMMO2 (Leconte et al. 2008), and subsequently achieved d5SICS–dMMO2 and d5SICS–dNaM pairs, which exhibited the high values of 85.7–99.8 % for fidelity-per-round in PCR using *Taq*, *Deep vent*, and *Phusion high-fidelity* DNA polymerases (Seo et al. 2009; Malyshev et al. 2009). Recently, they finally managed to create a semisynthetic organism with an expanded genetic alphabet involving d5SICS–dNaM as the third base pair; the genetically engineered organism was *E. coli* that expresses an algal nucleotide triphosphate transporter, which has the efficient uptake of the triphosphates of

d5SICS and dNaM, and thereby accurately replicates a plasmid containing d5SICS–dNaM (Malyshev et al. 2014).

In 2000, I. Hirao and coworkers designed and synthesized 2-amino-6-(N, N-dimethylamino)purine (denoted by x) and pyridin-2-one (denoted by y) deoxynucleoside analogs (Fig. 4d) (Ishikawa et al. 2000). They anticipated that the steric hindrance between the dimethyl at the N6 position of x and the 4-keto group of T would interfere with the formation of an x–T mismatch pair and, furthermore, that the unique pattern of hydrogen bonding between N1 and N2 on x and N1 and O2 on y would form a stable and specific x–y base pair. In PEX, using KF and KF ($3'$–$5'$ exo-), y was selectively incorporated opposite x on the template, which unfortunately was also erroneously incorporated opposites A and G. A ribonucleoside-$5'$-triphosphate analog of y was also synthesized, and the single incorporation of y opposite x in transcription was assessed using T7 RNA polymerase (Ohtsuki et al. 2001). As a result, y was incorporated opposite x with 95 % accuracy, while the erroneous incorporation of U opposite x was only occasionally observed (<5 % of instances). Thereafter, to improve incorporation efficiency and selectivity, they developed 2-amino-6-(2-thienyl)purine (denoted by s) and imidazolin-2-one (denoted by z) analogs to form s–y (Hirao et al. 2002)and s–z base pairs, respectively, which involve the formation of two hydrogen bonds (Hirao et al. 2004). Furthermore, the improvement of s provided the 7-(2-thienyl)-imidazo [4,5-b]pyridine (Ds) analog and of z yielded pyrrole-2-carboxaldehyde (Pa) (Hirao et al. 2006) and 2-nitro-4-propynylpyrrole (Px) analogs, which can form Ds–Pa and Ds–Px base pairs, respectively (Kimoto et al. 2009). The Ds–Px base pair in particular exhibited a high value of 99.9 % for fidelity-per-round in PCR using *Deep vent* DNA polymerase. Intriguingly, as models such as Kool's and Romesberg's possessing the F–Z and d5SICS–dNaM pairs, these base pairs do not involve hydrogen bonds. Using an analog bearing, i.e., a foreign functionality via the 4-propynyl group of Px, they have recently performed selections of modified DNA aptamers containing nucleoside analogs with the 5th base (discussed later).

3.3 Sugar Modification

In general, sugar-modified nucleic acids are called xenonucleic acids (XNAs). XNAs containing modified sugars with $2'$-substituents such as methoxy (–OMe), fluoro (–F), amino (–NH$_2$), and azido (–N$_3$) groups are typical examples. Moreover, XNAs based on C2$'$-stereoisomers such as arabinonucleic acids (ANAs) and $2'$-fluoroarabinonucleic acids (FANAs) have been extensively studied. Furthermore, XNAs containing unconventional sugars such as hexitol nucleic acids (HNAs), α-L-threofuranosyl nucleic acids (TNAs), cyclohexenyl nucleic acids (CeNAs), and $2'$-O,$4'$-C-methylene-bridged/locked nucleic acids ($2'$,$4'$-BNAs/LNAs; hereinafter referred to as "LNA") have attracted the attention of researchers as informational biopolymer alternatives to DNA and RNA (Fig. 5).
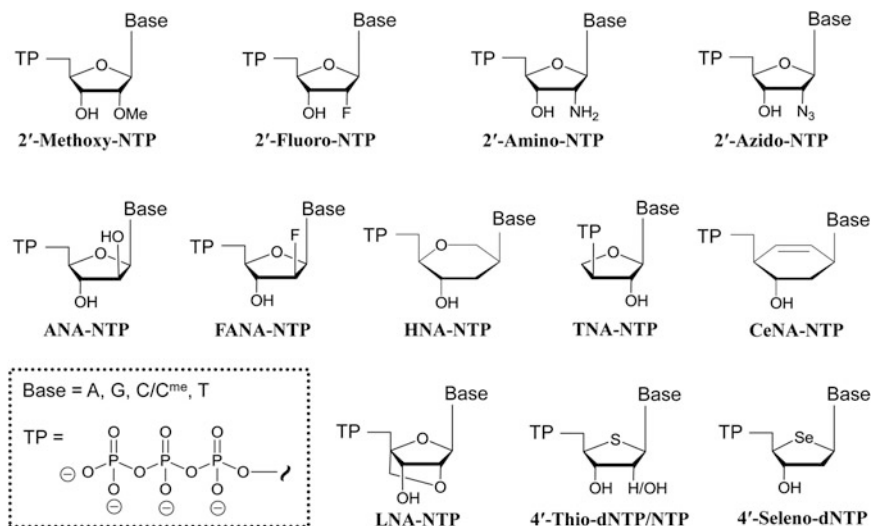
Fig. 5 Sugar-modified nucleoside triphosphates

In the late 1990s, to incorporate 2′-modified nucleotide analogs, screening of DNA polymerases that can incorporate 2′–F analogs was attempted, and *Vent*(*exo*-) and *Deep Vent*(*exo*-) DNA polymerases were found to be efficient catalysts (Ono et al. 1997). Moreover, the engineering of RNA polymerases was attempted. For example, T7 RNA polymerase variants, Y639F and Y639F/H784A, which enabled an efficient enzymatic incorporation of 2′–F or 2′–NH$_2$ analogs and those of the 2′–OMe or 2′–N$_3$ analogs, respectively, were developed (Padilla and Sousa 1999, 2002).

In 2000, K. Vastmans et al. demonstrated that *Vent*(*exo*-) DNA polymerase can elongate a 6-mer HNA strand on a DNA template using an HNA triphosphate (hATP) bearing an adenine base (Vastmans et al. 2000). In 2003, J. C. Chaput et al. examined DNA synthesis on a TNA template and TNA synthesis on a DNA template using various DNA polymerases (Chaput and Szostak 2003). Thereafter, they found that the Therminator DNA polymerase can polymerize TNA oligomers that are at least 80 nt in length using the following four TNA triphosphates, i.e., tTTP, tGTP, tCTP, and tDTP, which bear thymine, guanine, cytosine, and 2,6-diaminopurine, respectively, as a nucleobase (Ichida et al. 2005). In 2005, V. Kempeneers et al. reported that seven efficient successive CeNA incorporations were possible in the DNA-dependent CeNA polymerization using a CeNA triphosphate (CeATP) bearing an adenine base and *Vent*(*exo*-) DNA polymerase under conditions supplemented with 1 mM Mn^{2+} (Kempeneers et al. 2005). In 2007, R. N. Veedu et al. first examined the enzymatic incorporation of LNA nucleotides and observed that *Phusion High-Fidelity* DNA polymerase could accept LNA-triphosphates bearing thymine and adenine bases and catalyze primer extension reactions to yield DNA-based LNA strands (Veedu et al. 2007). In 2008,

M. Kuwahara et al. first demonstrated that *KOD Dash* DNA polymerase was superior to *Phusion High-Fidelity* DNA polymerase because of its reduced 3′,5′-exonuclease activity and could be applied for the synthesis of not only LNA but also other types of LNA, i.e., 2′,4′-BNA^{COC} and 2′,4′-BNA^{NC} (Kuwahara et al. 2008).

In 2012, P. Holliger et al. successfully created some XNA polymerases variants of *Tgo* DNA polymerase using compartmentalized self-tagging (CST) selection, which was performed on libraries of *TgoT* DNA polymerase that contained four amino acid mutations (V93Q, D141A, E143A, and A485L) (Pinheiro et al. 2012). For example, HNA polymerase was additionally mutated at V589A, E609K, I610M, K659Q, E664Q, Q665P, R668K, D669Q, K671H, K674R, T676R, A681S, L704P, and E730G. CeNA/LNA polymerase was additionally mutated at E654Q, E658Q, K659Q, V661A, E664Q, Q665P, D669A, K671Q, T676K, and R709K. ANA/FANA polymerase was additionally mutated atL403P, P657T, E658Q, K659H, Y663H, E664K, D669A, K671N, and T676I. These *TgoT* variants can produce the corresponding 72-mer XNA strands on a DNA template using four XNA triphosphates with different bases (A, G, C, and T).

As examples of other types of XNAs, 4′-modified analogs such as 4′-thioribonucleoside-5′-triphosphates, 2′-deoxy-4′-thionucleoside-5′-triphosphates, and 2′-deoxy-4′-selenonucleoside-5′-triphosphates, which have been developed by N. Minakawa and coworkers, are mentioned (Kato et al. 2005; Kojima et al. 2013; Tarashima et al. 2015). They have found that the first ones can be accepted by T7 RNA polymerase, whereas the second and third ones can act as good substrates for *KOD Dash* DNA polymerase.

Thus, continuous efforts in the screening and engineering of polymerases have enabled an efficient enzymatic production of various XNAs.

3.4 Phosphate Modification

Phosphorothioate nucleotides are well-studied phosphate-modified analogs in which one of the non-bridging α-phosphate oxygen is replaced by sulfur because introduction of phosphorothioate linkages to oligonucleotides can enhance nuclease resistance and antisense activity (Fig. 6). In 1968, F. Eckstein et al. first demonstrated enzymatic incorporation of a phosphorothioate using UTPαS (mixture of Sp- and Rp-isomers) and *Escherichia coli* DNA-dependent RNA polymerase (Matzura and Eckstein 1968). Thereafter, using optically pure Sp- or Rp-ATPαS, they discovered that the Sp-isomer could be more efficiently incorporated than the Rp-isomer during RNA strand elongation catalyzed by *E. coli* DNA-dependent RNA polymerase (Eckstein et al. 1976). In 1988, H. P. Vosberg et al. showed that the *Taq* DNA polymerase-catalyzed polymerase chain reaction (PCR) with Sp-dATPαS and dNTPs, except dATP, could efficiently amplify 310-mer double-stranded DNA fragments that contained multiple phosphorothioates (Nakamaye et al. 1988). The same results were also obtained when Sp-dNTPαS was used
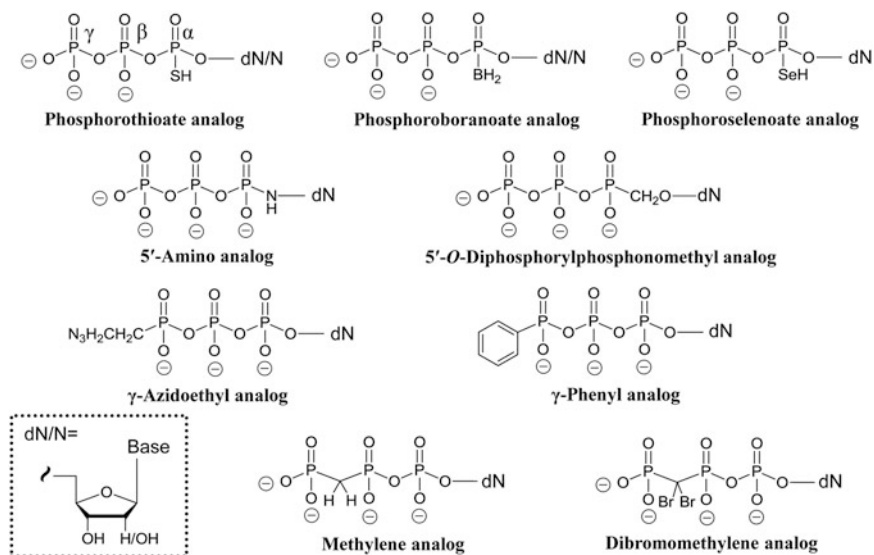
Fig. 6 Phosphate-modified nucleoside triphosphates

with a different base (G, C, or T). Other examples of non-bridging α-phosphate oxygen substituents, phosphoroboranoate and phosphoroselenoate analogs of 2′-deoxynucleoside-5′-triphosphates, have been reported. Unlike phosphorothioate analogs, the Rp-isomer of phosphoroboranoates was preferably incorporated using Sequenase DNA polymerase. One should note, however, that the Sp-isomer of phosphorothioates corresponds to the Rp-isomer of phosphoroboranoates in the configuration of the four substituents, i.e., $-P_2O_5^{3-}$, $=O$, nucleoside, and $-SH/-BH_{3-}$, that bond to the asymmetric phosphorus atom (Li et al. 1995; He et al. 1999). Meanwhile, both isomers of phosphoroselenoates were equally accepted as substrates by KF to produce corresponding DNA strands (Carrasco and Huang 2004).

Replacement of the bridging oxygen between α-phosphorus and the sugar 5′ carbon in a dNTP by other substituents is one of the means to enzymatically introduce chemical modifications into oligonucleotide backbones. J. L. Wolfe et al. reported that in the presence of natural dNTPs, KF (3′–5′ exo-) could accept 5′-amino-2′,5′-dideoxynucleoside-5′-N-triphosphates to produce the corresponding oligodeoxynucleotide strand in which multiple oxygen atoms in the 5′-position were replaced with imino (–NH–) groups (Wolfe et al. 2002). Similarly, P. Herdewijn et al. demonstrated the insertion of the methyleneoxy (–CH₂O–) group at this position using 5′-O-diphosphorylphosphonomethyl-2′-dA with Therminator DNA polymerase (Renders et al. 2007).

To analyze the effects of chemical modifications of the pyrophosphate moiety as a leaving group on polymerase activities, γ-phosphate oxygen substituents, i.e., γ-substituted dNTPs, were examined using AMT RT and DNA polymerase α (Alexandrova et al. 1998). Consequently, γ-substituted dNTPs with azidoethyl,

aminoethyl, phenyl, or 2,4-dinitrophenyl groups could moderately be accepted only by AMT RT, although the latter two were poorer substrates than the former two owing to the bulkiness of the substituents. Furthermore, phosphate-modified analogs in which the bridging oxygen between β- and γ-phosphorus is replaced by the methylene (–CH_2–) or dibromomethylene (–CBr_2–) group were also examined. The results indicated that these modifications are more sensitive to the action of AMT RT than γ-substitutions.

4 Applications

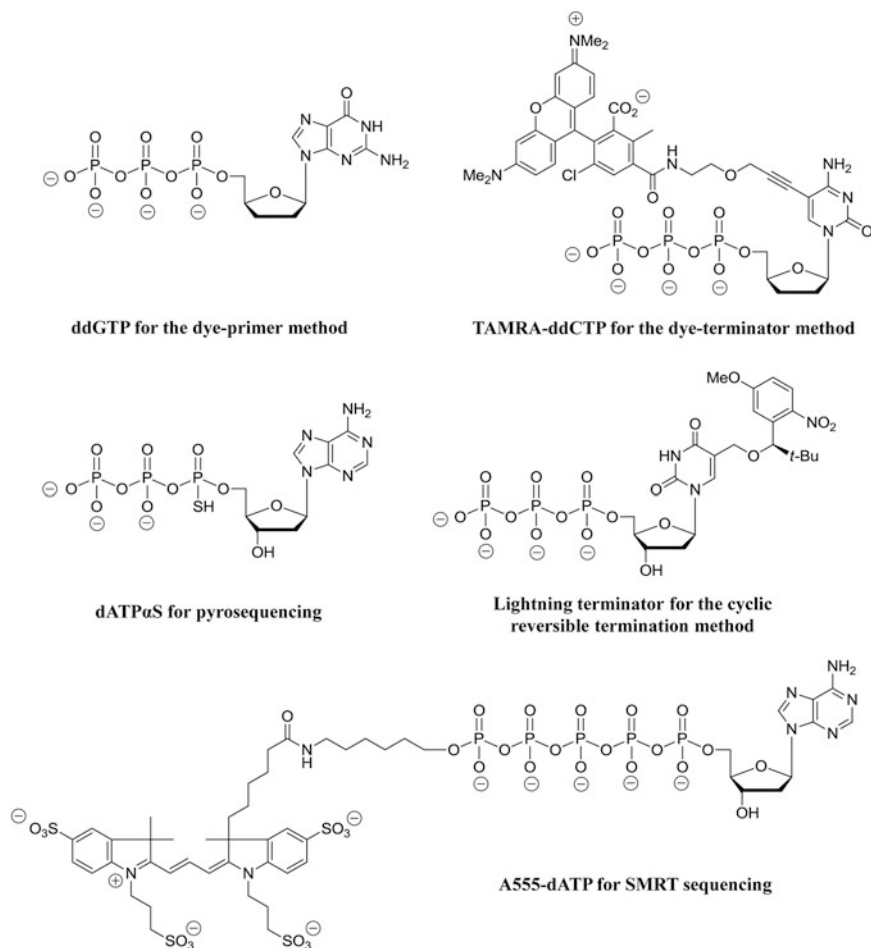## 4.1 DNA Sequencing

DNA sequencing methods can roughly be classified into three categories, i.e., first-, second-, and third-generation technologies. First-generation sequencing refers to fluorescent DNA sequencing using capillary electrophoresis, which greatly contributed to the Human Genome Project, 1990–2003. Fluorescent DNA sequencing methods, namely, the dye-primer method, which uses fluorophore-labeled primers, and the dye-terminator method, which uses fluorophore-labeled terminators, i.e., four $2',3'$-dideoxynucleoside-$5'$-triphosphates (ddNTPs) tagged with different fluorescent dyes, were devised in the 1980s (Fig. 7). Initially, the dye-primer method, which could provide more satisfactory long-read sequencing data, was mainly used, and KF ($3'$–$5'$ exo-) and Sequenase DNA polymerase were employed as enzymes for sequencing reactions. However, after *AmpliTaq* DNA polymerase FS, which exhibits a very weak $5'$ to $3'$ exonuclease activity and readily incorporates ddNTPs, had been developed, the dye-primer method was replaced with the dye-terminator method.

The second-generation sequencing technologies enable massively parallel sequencing. For example, Roche 454 sequencing technologies based on the pyrosequencing method employ dATPαS instead of dATP because luciferase accepts natural dATP as a substrate, giving rise to false-positive signals. As with dATPαS, some dATP analogs with base modifications can reduce false luciferase positives (Kajiyama et al. 2011). In pyrosequencing, KF ($3'$–$5'$ exo-) (Ronaghi et al. 1996), Sequenase Version 2.0 (Gharizadeh et al. 2004) DNA polymerase, and *Bst* DNA polymerase (large fragment) are used (Margulies et al. 2005). Lasergen developed other sequencing technologies based on the cyclic reversible termination method using $3'$-OH unblocked nucleotides called Lightning Terminators. These triphosphates have a terminating 2-nitrobenzyl moiety attached to hydroxymethylated nucleobases and are efficiently incorporated by Therminator DNA polymerase (Gardner et al. 2012) (Fig. 7).

The third-generation sequencing technologies involve single-molecule real-time (SMRT) sequencing, which enables direct analysis of DNA/RNA extracted from

Fig. 7 Modified nucleoside triphosphates employed for DNA sequencing technologies

biological samples such as cells, tissues, and organs. For example, Pacific Biosciences developed an SMRT sequencing system using four modified dNTPs consisting of a γ-phosphate with different fluorophores and a complex of φ29 DNA polymerase and an analyte DNA template. The complex is solely immobilized at the bottom of a zero-mode waveguide, and attenuated light from the excitation beam penetrates only the lower 20–30 nm of each waveguide to provide a light microscope with a detection volume of only 20^{-21} L (Eid et al. 2009) (Fig. 7).

Thus, development of modified triphosphates and engineered polymerases continues to greatly contribute to the advancement of sequencing technologies.

4.2 Aptamer Development

To enhance nuclease resistance and improve target-binding affinities and specificities, various nucleoside triphosphates have been employed for the aptamer selection process, so-called systematic evolution of ligands by exponential enrichment (SELEX) (Tuerk and Gold 1990; Ellington and Szostak 1990). Macugen (pegaptanib sodium injection) for the treatment of wet age-related macular degeneration is a typical example of RNA-based aptamers. This modified RNA aptamer targeting vascular endothelial growth factor (VEGF) was selected from a modified RNA library. The library was enzymatically synthesized using T7 RNA polymerase and 2′-F pyrimidine nucleoside triphosphates (U and C) as well as natural ATP and GTP, followed by derivatizations, including post-SELEX modification, by which all natural purine nucleotides, except for two adenosine residues, could be replaced with 2′-OMe analogs (Ruckman et al. 1998). Furthermore, the use of a T7 RNA polymerase variant (Y639F, H784A, and K378R) enabled direct selection of modified RNA aptamers totally replaced with 2′-OMe nucleotides, which was demonstrated with VEGF as a target (Burmeister et al. 2005). Other examples of modified RNA aptamers created by SELEX using the corresponding triphosphate analogs and T7 RNA polymerase are those containing 2′-NH_2 pyrimidine nucleosides (U and C) for human neutrophil elastase (HNE) (Lin et al. 1994), those containing 4′-thioribonucleosides (A, G, C, and U) for human thrombin (Minakawa et al. 2008), those containing 5-iodouridines (photoaptamer) for the HIV-1 Rev protein (Jensen et al. 1995), those containing 5-(3-aminopropyl)uridines for ATP (Vaish et al. 2003), and those containing phosphorothioates (A, G, C, and U) for basic fibroblast growth factor (bFGF) (Jhaveri et al. 1998).

For DNA-based aptamer selection, *Vent* and *KOD*-related DNA polymerases have mainly been used. For example, modified DNA aptamers containing 5-(1-pentynyl)-2′-deoxyuridines for human thrombin (Latham et al. 1994) and those containing 5-(3-aminopropyl)-2′-deoxyuridines for ATP were obtained by SELEX using *Vent* DNA polymerase (Battersby et al. 1999). Meanwhile, modified DNA aptamers containing 5-*N*-(6-aminohexyl)carbamoylmethyl-2′-deoxyuridines for (R)-thalidomide(Fig. 2) (Shoji et al. 2007); those containing (*E*)-5-(2-(*N*-(2-(N^6-adeninyl)ethyl))carbamylvinyl)-2′-deoxyuridines for camptothecin derivatives (Fig. 2) (Imaizumi et al. 2013); those containing 5-tryptaminocarbonyl-2′-deoxyuridines (SOMAmers) for various protein targets such as fractalkine, osteoprotegerin, and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) (Fig. 2) (Gold et al. 2010); those containing the 2′-deoxynucleosides bearing Ds, i.e., one of the artificial bases from the Ds–Px base pair for VEGF and interferon-γ (Fig. 4d) (Kimoto et al. 2013); and those containing locked nucleic acid (LNA) nucleosides bearing thymine for human thrombin (Kasahara et al. 2013) were obtained by SELEX using *KOD*-related DNA polymerases such as *KOD Dash*, *KOD XL*, and *AccuPrimePfx* DNA polymerases (Fig. 5). In addition, SELEX using TaKaRa *Taq* Hot Start (HS) DNA polymerase was reported to have provided DNA-based aptamers comprised of six letters, i.e., A, G, C, T, Z (2(1H)-pyridone),

and P (imidazo[1,2-a]-1,3,5-triazin-4(8H)-one), for HepG2 liver cancer cells (Zhang et al. 2015).

Furthermore, selection of the following XNA aptamers has been achieved using DNA polymerases: threose nucleic acid (TNA) aptamers for human thrombin (Therminator DNA polymerase) (Yu et al. 2012), hexitol nucleic acid (HNA) aptamers for hen egg lysozyme and HIV trans-activating response RNA (*TgoT* variant) (Pinheiro et al. 2012), and 2′-fluoroarabinonucleic acid (FANA) aptamers for HIV-1 reverse transcriptase (*TgoT* variant) (Alves Ferreira-Bravo et al. 2015).

5 Conclusions and Future Outlook

Since the 1980s, various nucleoside triphosphate analogs and polymerase variants have been developed. In the absence of the corresponding NTP(s)/dNTP(s), it became possible to produce oligonucleotides with single, double, triple, and quadruple substitutions with modified nucleotide(s) as research on screening and engineering of RNA/DNA polymerases progressed. Furthermore, oligodeoxynucleotides with expanded genetic alphabets and artificial biopolymers, i.e., XNAs with sugar structures quite different from β-D-ribofuranose in RNA or 2-deoxy-β-D-ribofuranose in DNA, can be enzymatically synthesized. Except for the abovementioned examples, various enzymatically synthesized polynucleotides that exhibit unique properties owing to introduced modifications, e.g., solvatochromic (Riedl et al. 2012), viscosity-sensitive (Dziuba et al. 2015), electroconductive (Patolsky et al. 2002), and amphiphilic functionalities (Fujita et al. 2015), have been reported. Such artificial biopolymers will serve as sensor materials for environmental and biological analyses and as programmable nanocapsules for drug carriers and gene deliveries in the near future.

References

Adler S, Modrich P (1979) T7-induced DNA polymerase. Characterization of associated exonuclease activities and resolution into biologically active subunits. J Biol Chem 254:11605–11614

Alexandrova LA, Skoblov AY, Jasko MV et al (1998) 2′-Deoxynucleoside 5′-triphosphates modified at alpha-, beta- and gamma-phosphates as substrates for DNA polymerases. Nucleic Acids Res 26:778–786

Alves Ferreira-Bravo I, Cozens C, Holliger P et al (2015) Selection of 2′-deoxy-2′-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. Nucleic Acids Res 43:9587–9599

Battersby TR, Ang DN, Burgstaller P et al (1999) Quantitative analysis of receptors for adenosine nucleotides obtained via in vitro selection from a library incorporating a cationic nucleotide analog. J Am Chem Soc 121:9781–9789

Bebenek K, Joyce CM, Fitzgerald MP et al (1990) The fidelity of DNA synthesis catalyzed by derivatives of Escherichia coli DNA polymerase I. J Biol Chem 265:13878–13887

Blanco L, Bernad A, Lázaro JM et al (1989) Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. J Biol Chem 264:8935–8940

Boschi-Muller S, Motorin Y (2013) Chemistry enters nucleic acids biology: enzymatic mechanisms of RNA modification. Biochemistry (Mosc) 78:1392–1404

Burmeister PE, Lewis SD, Silva RF et al (2005) Direct in vitro selection of a 2′-O-methyl aptamer to VEGF. Chem Biol 12:25–33

Bußkamp H, Batroff E, Niederwieser A et al (2014) Efficient labelling of enzymatically synthesized vinyl-modified DNA by an inverse-electron-demand Diels-Alder reaction. Chem Commun (Camb) 50:10827–10829

Campbell JL, Richardson CC, Studier FW (1978) Genetic recombination and complementation between bacteriophage T7 and cloned fragments of T7 DNA. Proc Natl Acad Sci USA 75:2276–2280

Carrasco N, Huang Z (2004) Enzymatic synthesis of phosphoroselenoate DNA using thymidine 5′-(alpha-P-seleno)triphosphate and DNA polymerase for X-ray crystallography via MAD. J Am Chem Soc 126:448–449

Chaput JC, Szostak JW (2003) TNA synthesis by DNA polymerases. J Am Chem Soc 125:9274–9275

Chelliserrykattil J, Lu H, Lee AH et al (2008) Polymerase amplification, cloning, and gene expression of benzo-homologous "yDNA" base pairs. Chembiochem 9:2976–2980

Chien A, Edgar DB, Trela JM (1976) Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. J Bacteriol 127:1550–1557

Davanloo P, Rosenberg AH, Dunn JJ et al (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase. Proc Natl Acad Sci USA 81:2035–2039

De Clercq E (2001) Antiviral drugs: current state of the art. J Clin Virol 22:73–89

Dziuba D, Pohl R, Hocek M (2015) Polymerase synthesis of DNA labelled with benzylidene cyanoacetamide-based fluorescent molecular rotors: fluorescent light-up probes for DNA-binding proteins. Chem Commun 51:4880–4882

Eckstein F, Armstrong VW, Sternbach H (1976) Stereochemistry of polymerization by DNA-dependent RNA-polymerase from Escherichia coli: an investigation with a diastereomeric ATP-analogue. Proc Natl Acad Sci USA 73:2987–2990

Eid J, Fehr A, Gray J et al (2009) Real-time DNA sequencing from single polymerase molecules. Science 323:133–138

Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. Nature 346:818–822

Fiala G, Stetter KO (1986) Pyrococcus furiosus sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch Microbiol 145:56–61

Fujita H, Nakajima K, Kasahara Y et al (2015) Polymerase-mediated high-density incorporation of amphiphilic functionalities into DNA: enhancement of nuclease resistance and stability in human serum. Bioorg Med Chem Lett 25:333–336

Gardner AF, Wang J, Wu W et al (2012) Rapid incorporation kinetics and improved fidelity of a novel class of 3′-OH unblocked reversible terminators. Nucleic Acids Res 40:7404–7415

Gharizadeh B, Eriksson J, Nourizad N et al (2004) Improvements in pyrosequencing technology by employing sequenase polymerase. Anal Biochem 330:272–280

Gierlich J, Gutsmiedl K, Gramlich PM et al (2007) Synthesis of highly modified DNA by a combination of PCR with alkyne-bearing triphosphates and click chemistry. Chemistry 13:9486–9494

Gold L, Ayers D, Bertino J et al (2010) Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One 5:e15004

Gourlain T, Sidorov A, Mignet N et al (2001) Enhancing the catalytic repertoire of nucleic acids. II. Simultaneous incorporation of amino and imidazolyl functionalities by two modified triphosphates during PCR. Nucleic Acids Res 29:1898–1905

Grippo P, Richardson CC (1971) Deoxyribonucleic acid polymerase of bacteriophage T7. J Biol Chem 246:6867–6873

He K, Porter KW, Hasan A et al (1999) Synthesis of 5-substituted 2′-deoxycytidine 5′-(alpha-P-borano)triphosphates, their incorporation into DNA and effects on exonuclease. Nucleic Acids Res 27:1788–1794

Held HA, Benner SA (2002) Challenging artificial genetic systems: thymidine analogs with 5-position sulfur functionality. Nucleic Acids Res 30:3857–3869

Hirao I, Ohtsuki T, Fujiwara T et al (2002) An unnatural base pair for incorporating amino acid analogs into proteins. Nat Biotechnol 20:177–182

Hirao I, Harada Y, Kimoto M et al (2004) A two-unnatural-base-pair system toward the expansion of the genetic code. J Am Chem Soc 126:13298–13305

Hirao I, Kimoto M, Mitsui T et al (2006) An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. Nat Methods 3:729–735

Hollenstein M, Hipolito CJ, Lam CH et al (2009) A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M2+). Nucleic Acids Res 37:1638–1649

Ichida JK, Horhota A, Zou K et al (2005) High fidelity TNA synthesis by Therminator polymerase. Nucleic Acids Res 33:5219–5225

Imaizumi Y, Kasahara Y, Fujita H et al (2013) Efficacy of base-modification on target binding of small molecule DNA aptamers. J Am Chem Soc 135:9412–9419

Ishikawa M, Hirao I, Yokoyama S (2000) Synthesis of 3-(2-deoxy-β-d-ribofuranosyl)pyridin-2-one and 2-amino-6-(N, N-dimethylamino)-9-(2-deoxy-β-d-ribofuranosyl)purine derivatives for an unnatural base pair. Tetrahedron Lett 41:3931–3934

Jäger S, Rasched G, Kornreich-Leshem H et al (2005) A versatile toolbox for variable DNA functionalization at high density. J Am Chem Soc 127:15071–15082

Jannasch HW, Wirsen CO, Molyneaux SJ et al (1992) Comparative physiological studies on hyperthermophilic archaea isolated from deep-sea hot vents with emphasis on *Pyrococcus* strain GB-D. Appl Environ Microbiol 58:3472–3481

Jensen KB, Atkinson BL, Willis MC et al (1995) Using in vitro selection to direct the covalent attachment of human immunodeficiency virus type 1 Rev protein to high-affinity RNA ligands. Proc Natl Acad Sci USA 92:12220–12224

Jhaveri S, Olwin B, Ellington AD (1998) In vitro selection of phosphorothiolated aptamers. Bioorg Med Chem Lett 8:2285–2290

Kajiyama T, Kuwahara M, Goto M et al (2011) Optimization of pyrosequencing reads by superior successive incorporation efficiency of improved 2′-deoxyadenosine-5′-triphosphate analogs. Anal Biochem 416:8–17

Kasahara Y, Irisawa Y, Ozaki H et al (2013) 2′,4′-BNA/LNA aptamers: CE-SELEX using a DNA-based library of full-length 2′-O,4′-C-methylene-bridged/linked bicyclic ribonucleotides. Bioorg Med Chem Lett 23:1288–1292

Kato Y, Minakawa N, Komatsu Y et al (2005) New NTP analogs: the synthesis of 4′-thioUTP and 4′-thioCTP and their utility for SELEX. Nucleic Acids Res 33:2942–2951

Kempeneers V, Renders M, Froeyen M et al (2005) Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization. Nucleic Acids Res 33:3828–3836

Kiefer JR, Mao C, Braman JC et al (1998) Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. Nature 391:304–307

Kimoto M, Kawai R, Mitsui T et al (2009) An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules. Nucleic Acids Res 37:e14

Kimoto M, Yamashige R, Matsunaga K et al (2013) Generation of high-affinity DNA aptamers using an expanded genetic alphabet. Nat Biotechnol 31:453–457

Klenow H, Henningsen I (1970) Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from *Escherichia coli* B by limited proteolysis. Proc Natl Acad Sci USA 65:168–175

Kojima T, Furukawa K, Maruyama H et al (2013) PCR amplification of 4′-thioDNA using 2′-deoxy-4′-thionucleoside 5′-triphosphates. ACS Synth Biol 2:529–536

Kong H, Kucera RB, Jack WE (1993) Characterization of a DNA polymerase from the hyperthermophile archaea *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. J Biol Chem 268:1965–1975

Kotani H, Ishizaki Y, Hiraoka N et al (1987) Nucleotide sequence and expression of the cloned gene of bacteriophage SP6 RNA polymerase. Nucleic Acids Res 15:2653–2664

Kufe DW, Major PP, Egan EM et al (1980) Correlation of cytotoxicity with incorporation of ara-C into DNA. J Biol Chem 255:8997–9000

Kuwahara M, Nagashima J, Hasegawa M et al (2006) Systematic characterization of 2′-deoxynucleoside- 5′-triphosphate analogs as substrates for DNA polymerases by polymerase chain reaction and kinetic studies on enzymatic production of modified DNA. Nucleic Acids Res 34:5383–5394

Kuwahara M, Obika S, Nagashima J et al (2008) Systematic analysis of enzymatic DNA polymerization using oligo-DNA templates and triphosphate analogs involving 2′,4′-bridged nucleosides. Nucleic Acids Res 36:4257–4265

Latham JA, Johnson R, Toole JJ (1994) The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing 5-(1-pentynyl)-2′-deoxyuridine. Nucleic Acids Res 22:2817–2822

Lawyer FC, Stoffel S, Saiki RK et al (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. J Biol Chem 264:6427–6437

Leconte AM, Hwang GT, Matsuda S et al (2008) Discovery, characterization, and optimization of an unnatural base pair for expansion of the genetic alphabet. J Am Chem Soc 130:2336–2343

Lee AH, Kool ET (2005) A new four-base genetic helix, yDNA, composed of widened benzopyrimidine-purine pairs. J Am Chem Soc 127:3332–3338

Lee SE, Sidorov A, Gourlain T et al (2001) Enhancing the catalytic repertoire of nucleic acids: a systematic study of linker length and rigidity. Nucleic Acids Res 29:1565–1573

Li H, Porter K, Huang F et al (1995) Boron-containing oligodeoxyribonucleotide 14mer duplexes: enzymatic synthesis and melting studies. Nucleic Acids Res 23:4495–4501

Limbach PA, Crain PF, McCloskey JA (1994) Summary: the modified nucleosides of RNA. Nucleic Acids Res 22:2183–2196

Lin Y, Qiu Q, Gill SC et al (1994) Modified RNA sequence pools for in vitro selection. Nucleic Acids Res 22:5229–5234

Majumder HK, Maitra U, Rosenberg M (1979) Termination of transcription by bacteriophage T3 RNA polymerase: homogeneous 3′-terminal oligonucleotide sequence of in vitro T3 RNA polymerase transcripts. Proc Natl Acad Sci USA 76:5110–5113

Malyshev DA, Seo YJ, Ordoukhanian P et al (2009) PCR with an expanded genetic alphabet. J Am Chem Soc 131:14620–14621

Malyshev DA, Dhami K, Lavergne T et al (2014) A semi-synthetic organism with an expanded genetic alphabet. Nature 509:385–388

Margulies M, Egholm M, Altman WE et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380

Matray TJ, Kool ET (1999) A specific partner for abasic damage in DNA. Nature 399:704–708

Mattila P, Korpela J, Tenkanen T et al (1991) Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase–an extremely heat stable enzyme with proofreading activity. Nucleic Acids Res 19:4967–4973

Matzura H, Eckstein F (1968) A polyribonucleotide containing alternation P=O and P=S linkages. Eur J Biochem 3:448–452

McMinn DL, Ogawa AK, Wu Y et al (1999) Efforts toward expansion of the genetic alphabet: DNA polymerase recognition of a highly stable, self-pairing hydrophobic base. J Am Chem Soc 121:11585–11586

Mehedi Masud M, Ozaki-Nakamura A, Kuwahara M et al (2003) Modified DNA bearing 5 (methoxycarbonylmethyl)-2′-deoxyuridine: preparation by PCR with thermophilic DNA polymerase and postsynthetic derivatization. Chembiochem 4:584–588

Minakawa N, Sanji M, Kato Y et al (2008) Investigations toward the selection of fully-modified 4′-thioRNA aptamers: optimization of in vitro transcription steps in the presence of 4-′-thioNTPs. Bioorg Med Chem 16:9450–9456

Miroshnichenko ML, Gongadze GM, Rainey FA et al (1998) Thermococcus gorgonarius sp. nov. and Thermococcus pacificus sp. nov.: heterotrophic extremely thermophilic archaea from New Zealand submarine hot vents. Int J Syst Bacteriol 48(Pt 1):23–29

Morales JC, Kool ET (1998) Efficient replication between non-hydrogen-bonded nucleoside shape analogs. Nat Struct Biol 5:950–954

Nakamaye KL, Gish G, Eckstein F et al (1988) Direct sequencing of polymerase chain reaction amplified DNA fragments through the incorporation of deoxynucleoside alpha-thiotriphosphates. Nucleic Acids Res 16:9947–9959

Nishioka M, Mizuguchi H, Fujiwara S et al (2001) Long and accurate PCR with a mixture of KOD DNA polymerase and its exonuclease deficient mutant enzyme. J Biotechnol 88:141–149

Ohtsuki T, Kimoto M, Ishikawa M et al (2001) Unnatural base pairs for specific transcription. Proc Natl Acad Sci USA 98:4922–4925

Ono T, Scalf M, Smith LM (1997) 2′-Fluoro modified nucleic acids: polymerase-directed synthesis, properties and stability to analysis by matrix-assisted laser desorption/ionization mass spectrometry. Nucleic Acids Res 25:4581–4588

Padilla R, Sousa R (1999) Efficient synthesis of nucleic acids heavily modified with non-canonical ribose 2′-groups using a mutant T7 RNA polymerase (RNAP). Nucleic Acids Res 27:1561–1563

Padilla R, Sousa R (2002) A Y639F/H784A T7 RNA polymerase double mutant displays superior properties for synthesizing RNAs with non-canonical NTPs. Nucleic Acids Res 30:e138

Patolsky F, Weizmann Y, Willner I (2002) Redox-active nucleic-acid replica for the amplified bioelectrocatalytic detection of viral DNA. J Am Chem Soc 124:770–772

Perrin DM, Garestier T, Hélène C (1999) Expanding the catalytic repertoire of nucleic acid catalysts: simultaneous incorporation of two modified deoxyribonucleoside triphosphates bearing ammonium and imidazolyl functionalities. Nucleosides Nucleotides 18:377–391

Perrin DM, Garestier T, Hélène C (2001) Bridging the gap between proteins and nucleic acids: a metal-independent RNAseA mimic with two protein-like functionalities. J Am Chem Soc 123:1556–1563

Piccirilli JA, Krauch T, Moroney SE et al (1990) Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. Nature 343:33–37

Pinheiro VB, Taylor AI, Cozens C et al (2012) Synthetic genetic polymers capable of heredity and evolution. Science 336:341–344

Raindlová V, Pohl R, Sanda M et al (2010) Direct polymerase synthesis of reactive aldehyde-functionalized DNA and its conjugation and staining with hydrazines. Angew Chem Int Ed Engl 49:1064–1066

Renders M, Emmerechts G, Rozenski J et al (2007) Enzymatic synthesis of phosphonomethyl oligonucleotides by therminator polymerase. Angew Chem Int Ed Engl 46:2501–2504

Ried T, Baldini A, Rand TC et al (1992) Simultaneous visualization of seven different DNA probes by in situ hybridization using combinatorial fluorescence and digital imaging microscopy. Proc Natl Acad Sci USA 89:1388–1392

Riedl J, Ménová P, Pohl R et al (2012) GFP-like fluorophores as DNA labels for studying DNA–protein interactions. J Org Chem 77:8287–8293

Ronaghi M, Karamohamed S, Pettersson B et al (1996) Real-time DNA sequencing using detection of pyrophosphate release. Anal Biochem 242:84–89

Ruckman J, Green LS, Beeson J et al (1998) 2′-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. J Biol Chem 273:20556–20567

Rüdiger A, Jorgensen PL, Antranikian G (1995) Isolation and characterization of a heat-stable pullulanase from the hyperthermophilic archaeon *Pyrococcus woesei* after cloning and expression of its gene in *Escherichia coli*. Appl Environ Microbiol 61:567–575

Rüttimann C, Cotoras M, Zaldivar J et al (1985) DNA polymerases from the extremely thermophilic bacterium *Thermus thermophilus* HB-8. Eur J Biochem 149:41–46

Sakthivel K, Barbas CF III (1998) Expanding the potential of DNA for binding and catalysis: highly functionalized dUTP derivatives that are substrates for thermostable DNA polymerases. Angew Chem Int Ed 37:2872–2875

Sawai H, Ozaki AN, Satoh F et al (2001) Expansion of structural and functional diversities of DNA using new 5-substituted deoxyuridine derivatives by PCR with superthermophilic KOD Dash DNA polymerase. Chem Commun 2604–2605

Seo YJ, Hwang GT, Ordoukhanian P et al (2009) Optimization of an unnatural base pair toward natural-like replication. J Am Chem Soc 131:3246–3252

Shoji A, Kuwahara M, Ozaki H et al (2007) Modified DNA aptamer that binds the (R)-isomer of a thalidomide derivative with high enantioselectivity. J Am Chem Soc 129:1456–1464

Sidorov AV, Grasby JA, Williams DM (2004) Sequence-specific cleavage of RNA in the absence of divalent metal ions by a DNAzyme incorporating imidazolyl and amino functionalities. Nucleic Acids Res 32:1591–1601

Sismour AM, Benner SA (2005) The use of thymidine analogs to improve the replication of an extra DNA base pair: a synthetic biological system. Nucleic Acids Res 33:5640–5646

Southworth MW, Kong H, Kucera RB et al (1996) Cloning of thermostable DNA polymerases from hyperthermophilic marine Archaea with emphasis on *Thermococcus* sp. 9 degrees N-7 and mutations affecting 3′-5′ exonuclease activity. Proc Natl Acad Sci USA 93:5281–5285

Switzer C, Moroney SE, Benner SA (1989) Enzymatic incorporation of a new base pair into DNA and RNA. J Am Chem Soc 111:8322–8323

Takagi M, Nishioka M, Kakihara H et al (1997) Characterization of DNA polymerase from *Pyrococcus* sp. strain KOD1 and its application to PCR. Appl Environ Microbiol 63:4504–4510

Tarashima N, Sumitomo T, Ando H et al (2015) Synthesis of DNA fragments containing 2′-deoxy-4′-selenonucleoside units using DNA polymerases: comparison of dNTPs with O, S and Se at the 4′-position in replication. Org Biomol Chem 13:6949–6952

Tasara T, Angerer B, Damond M et al (2003) Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. II. High-density labeling of natural DNA. Nucleic Acids Res 31:2636–2646

Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249:505–510

Vaish NK, Fraley AW, Szostak JW et al (2000) Expanding the structural and functional diversity of RNA: analog uridine triphosphates as candidates for in vitro selection of nucleic acids. Nucleic Acids Res 28:3316–3322

Vaish NK, Larralde R, Fraley AW et al (2003) A novel, modification-dependent ATP-binding aptamer selected from an RNA library incorporating a cationic functionality. Biochemistry 42:8842–8851

Vastmans K, Pochet S, Peys A et al (2000) Enzymatic incorporation in DNA of 1,5-anhydrohexitol nucleotides. Biochemistry 39:12757–12765

Veedu RN, Vester B, Wengel J (2007) Enzymatic incorporation of LNA nucleotides into DNA strands. Chembiochem 8:490–492

Weisbrod SH, Marx A (2007) A nucleoside triphosphate for site-specific labelling of DNA by the Staudinger ligation. Chem Commun (Camb) 1828–1830

Wolfe JL, Kawate T, Belenky A et al (2002) Synthesis and polymerase incorporation of 5′-amino-2′,5′-dideoxy-5′-N-triphosphate nucleotides. Nucleic Acids Res 30:3739–3747

Yang Z, Hutter D, Sheng P et al (2006) Artificially expanded genetic information system: a new base pair with an alternative hydrogen bonding pattern. Nucleic Acids Res 34:6095–6101

Yang Z, Sismour AM, Sheng P et al (2007) Enzymatic incorporation of a third nucleobase pair. Nucleic Acids Res 35:4238–4249

Yu H, Zhang S, Chaput JC (2012) Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. Nat Chem 4:183–187

Zhang L, Yang Z, Sefah K et al (2015) Evolution of functional six-nucleotide DNA. J Am Chem Soc 137:6734–6737