

Barcoded Nucleotides**

Anna Baccaro, Anna-Lena Steck, and Andreas Marx*

DNA as an information storage system is simple and at the same time complex owing to the various different arrangements of the four natural nucleotides.^[1] The DNA synthesis by DNA polymerases is intriguing, since these enzymes are able to catalyze the elongation of the primer strand by recognizing the DNA template and selecting the corresponding nucleotide.^[1b,2] This feature can be exploited to diversify the four-base-code by substitution of the natural substrates with modified analogues.^[3] Nucleotide analogues equipped with various marker groups (e.g. dyes, tags, or spin labels^[4]) can be employed in DNA polymerase catalyzed reactions to increase the application scope of DNA (e.g. sequencing, structural characterization, and immobilization^[4d,5]). The “information” embedded in the marker groups allow conclusions to be drawn from the evaluation of the resulting signals. A significant gain in information would result by embedding a marker that exhibits the properties of a barcode. Typically, the barcode label bears no descriptive data but it consists of a series of signs which code for the deposited information (the term was used in other contexts with DNA before).^[6] For universal adoption the barcode should be simple, affixed easily, and allow a reliable assignment of the deposited information. Oligodeoxynucleotides (ODNs) meet the requirements of a barcode label to a great extent, since they have a simple code and can be distinguished by characteristics such as self-assembly and hybridization specificity. For a simple introduction of these DNA barcode labels into DNA, an enzyme-mediated approach utilizing ODN-modified nucleotides would be desirable.^[7] However, the acceptance of these modified nucleotides by DNA polymerases should be hampered by the steric demand of the ODN-modified nucleotides. Herein, we show that despite the steric demand the enzymatic synthesis of barcoded DNA is feasible by using ODN-modified nucleoside triphosphates that are about 40-times larger than the natural nucleotides and longer than the diameter of a DNA polymerase (Figure 1 A).

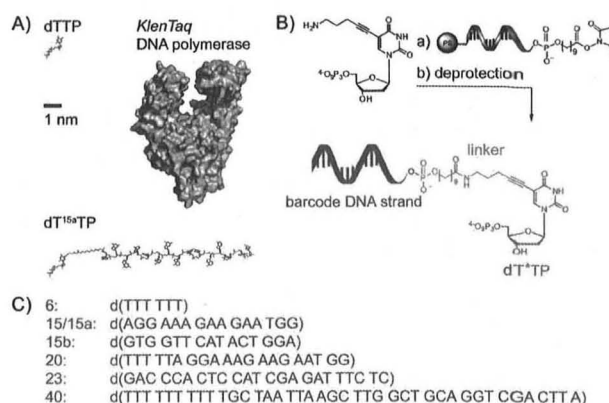


Figure 1. A) Schematic depiction for the comparison of sizes. dTTP versus dT^{15a}TP compared with *KlenTaq* DNA polymerase. B) Reaction pathway for the synthesis of ODN-modified dTTP. Yields and DNA sequences are listed in Table S1 of the Supporting Information. C) Sequences of barcode DNA strands. The numbers indicate the nucleotide lengths.

Herein, we introduce 2'-deoxyribonucleotide analogues, containing an ODN at the nucleobase (Figure 1B), as substrates for DNA polymerase mediated reactions. We chose the C5 position for pyrimidines and the C7 position for 7-deaza-purines to introduce the DNA strand at the nucleobase, since modifications at these positions have been accepted by DNA polymerases in several cases.^[3,8] To ODN-barcode-label nucleotides, an ODN strand was activated with a commercial available carboxy modifier at the 5'-end while still on solid support and then coupled to the amine-functionalized triphosphates (Figure 1B, see Supporting Information). After deprotection and cleavage from the solid support, these ODN-functionalized nucleotides were tested in DNA polymerase promoted primer-extension reactions (yields and DNA sequences are listed in Figure 1C and Supporting Information, Table S1). We examined the acceptance of the ODN-modified thymidine analogues by DNA polymerases in primer-extension reactions (Figure 2A for *Therminator* DNA polymerase, Supporting Information Figure S1 for *KlenTaq* DNA polymerase). We used a 23-nucleotide (nt) primer with a ³²P-label at the 5'-end and a 35-nt template, which contains a single A residue at position 27, coding for insertion of a thymidine analogue after extending the primer by three nucleotides (Figure 2A). Incubation with a DNA polymerase in absence of a thymidine analogue resulted in a primer elongation that is predominantly paused at position 27 without generating significant amounts of full-length product (Figure 2A, lane 1), while the reaction including all four natural deoxynucleoside triphosphates (dNTPs) showed full-length product (Figure 2A,

[*] Dr. A. Baccaro,^[†] Dipl.-Chem. A.-L. Steck,^[†] Prof. Dr. A. Marx
Department of Chemistry and Konstanz Research School Chemical
Biology, University of Konstanz
Universitätsstrasse 10, 78457 Konstanz (Germany)
E-mail: andreas.marx@uni-konstanz.de
Homepage: <http://www.chemie.uni-konstanz.de/~agmarx/>

[†] These authors contributed equally to this work.

[**] We gratefully acknowledge support by the Konstanz Research School Chemical Biology; the group of C. Hauck, University of Konstanz, for providing equipment, and the Ministerium für Wissenschaft, Forschung und Kunst, Baden-Württemberg for funding within the programme Bionik.

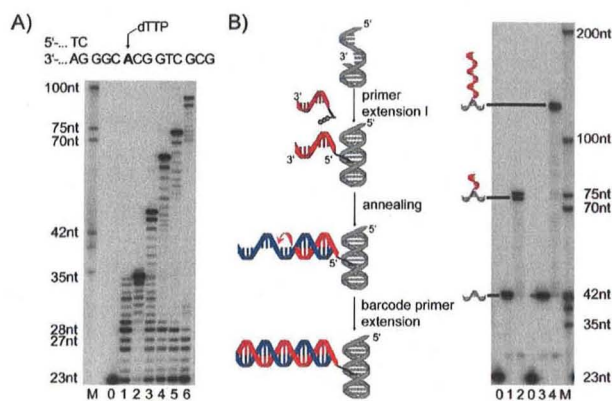


Figure 2. A) Partial DNA sequences of primer and template (see Supporting Information for more information) and PAGE analysis of the primer-extension studies using *Therminator* DNA polymerase, a 23-nt primer, a 35-nt template, and $10\ \mu\text{M}$ dNTPs. M: DNA marker; lane 0: $5'$ - ^{32}P -labeled primer only; lane 1: primer extension performed in the presence of dATP, dCTP, and dGTP; lane 2: same as lane 1, but in the presence of dTTP; lane 3: as lane 1, but in the presence of dT⁶TP; lane 4: as lane 1, but in the presence of dT^{15a}TP; lane 5: as lane 1, but in the presence of dT²³TP; lane 6: as lane 1, but in the presence of dT⁴⁰TP. B) Elongation of one incorporated dT²³MP. Left side: Reaction sequence used in this experiment (see Supporting Information). Right side: PAGE analysis of the primer-extension studies using *KlenTaq* DNA polymerase. M: DNA marker; lane 0: $5'$ - ^{32}P -labeled primer only, lane 1: primer-extension reaction I performed in the presence of dATP, dCTP, dGTP, and dTTP; lane 2: in the presence of dATP, dCTP, dGTP, and dT²³TP; lane 3: barcode primer-extension reaction performed with natural dNTPs and unmodified DNA, lane 4: barcode primer-extension reaction performed with natural dNTPs and dT²³MP modified DNA.

lane 2). By substitution of natural thymidine with one of the modified triphosphates (dT⁶TP, dT^{15a}TP, dT²³TP or dT⁴⁰TP; note: the superscript numbers represent the ODN-label length; DNA sequences are listed in Figure 1c and Supporting Information, Table S1) full-length product was obtained (Figure 2A, lanes 3–6). Double bands were observed arising from non-templated nucleotide addition to the 3'-termini of the blunt-ended DNA strand, which has been reported before.^[9] As expected, these reaction products migrated significantly more slowly in denaturing polyacrylamide gel electrophoresis (PAGE) than the unmodified full-length reaction product, indicating that the provided bulky nucleotide is incorporated. The lower mobility that increased with the size of the label, is explained by the additional bulk of the incorporated barcode DNA strand. Similar findings of lower mobility for modified reaction products have been reported before.^[10]

To evaluate the efficiency of incorporation of the modified nucleotides in comparison to the natural nucleotides we conducted single-nucleotide incorporation experiments in which the modified nucleotides (dT⁶TP, dT²⁰TP) directly compete for incorporation with their natural counterparts (Supporting Information, Figure S2). The ratio of unmodified versus modified nucleotide incorporation is easily accessible by PAGE through the significantly different retention times caused by the incorporation of the bulky modification. This

setup was previously used for the same purpose^[8a] as well as to study DNA polymerase selectivity.^[11] We found that *Therminator* DNA polymerase incorporates the investigated nucleotides with approximately 6- and 16-fold lower efficiency than the natural nucleotide while for *KlenTaq* DNA polymerase 33- and 66-fold lower efficiencies were observed. The observed efficiencies compare well to recently studied C5-modified dTTP analogues.^[8a]

We investigated the feasibility of multiple incorporations (Supporting Information, Figure S3). Using dT²⁰TP and a template coding for the insertion of 46 TMPs in the primer extension reaction, a highly branched reaction product is generated with at least 7 modified nucleotides in a row. Encouraged by these results we synthesized dATP, dCTP, and dGTP analogues (see Supporting Information) and tested them as well in the primer extension reaction (Figure S4). All the analogues were accepted by *Therminator* DNA polymerase and the primer was extended to full-length.

We tested the ability of DNA polymerases to utilize the incorporated barcode DNA strand as a primer in primer extension reactions. For this purpose, we performed primer extension reactions with natural dNTPs as a control reaction, and another reaction with dT²³TP instead of dTTP using a 24-nt primer and a 42-nt template coding for the insertion of one dTMP. These reaction products were hybridized with a second template (69-nt) complementary to the incorporated barcode DNA strand and incubated with a DNA polymerase and dNTPs for 1 h at 60°C performing the barcode primer extension reaction (Figure 2B). We observed complete disappearance of the initial band (Figure 2B, lanes 2 and 4) and the appearance of a new band shifted to lower mobility, indicating that the incorporated barcode DNA strand was used as the primer and elongated to a full-length product (Figure 2B, lane 4). As expected, in the control reaction with natural dNTPs, the mobility of the reaction product after the first primer extension was not altered on incubation under the same conditions. In addition, we tested the elongation of the incorporated DNA strand by rolling circle amplification^[12] (RCA) in solution and found extension as well (Supporting Information, Figure S5).

We investigated whether the ODN-modified nucleotides can be used as diagnostic tools for enzymatic reactions on solid supports. Therefore, we evaluated the feasibility for the detection of single nucleotide variations in the sequence context of the B type Raf kinase (BRAF) gene. The BRAF somatic T1796A mutation is encountered to a high extent in malignant melanomas and human cancers.^[13] Genome dissimilarities, such as single nucleotide polymorphisms (SNPs), are often responsible for a predisposition to the diseases^[13,14] and different drug efficiencies in certain individuals.^[15] For the SNP detection system, primer probes were covalently bound to an aminopropyl PDITC (1,4-phenylene diisothiocyanate) activated glass substrate.^[16] First single incorporation of ODN-modified nucleotides was performed using a template coding for the insertion of a dTMP (Figure 3A). Therefore, two reaction blocks of nine primer loci were incubated in the presence of a DNA polymerase, template, and with dA¹⁵TP or dT^{15a}TP. After incubation, the slides were washed and subsequently incubated with Cy3-labeled oligonucleotides

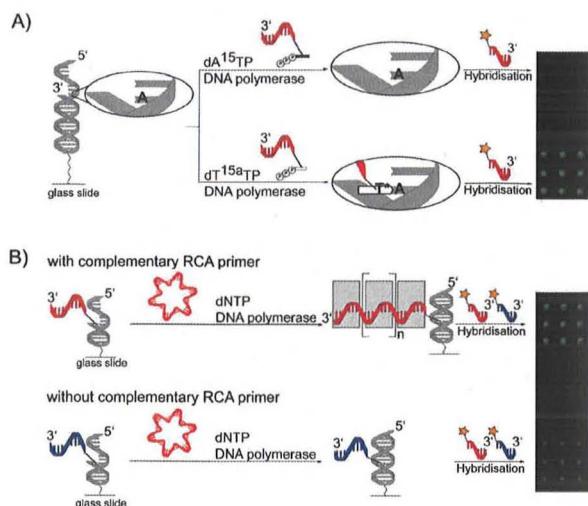


Figure 3. Microarray-based single-nucleotide-variation detection system. A) Reaction sequence performed on DNA microarray. Right side: Readout at 532 nm after hybridization with Cy3-labeled oligonucleotide. Reactions were conducted under the same conditions and on the same slide. B) Signal amplification by rolling circle amplification. Top: employing a complementary circular DNA template. Bottom: employing a non-complementary circular DNA template. Right side: Readout at 532 nm after hybridization with Cy3-labeled oligonucleotides. Reactions were conducted under the same conditions and on the same slide.

that bind to the oligonucleotide barcode of an incorporated dT^{15a}MP. Clearly, an intense fluorescence signal was only detected in cases where the canonical dT^{15a}MP was incorporated. To investigate signal amplification we incubated barcode-modified DNA complexes with a DNA polymerase in the presence of a circular template that binds to its complementary barcode DNA strand (Figure 3B). The circular template will enable the extension of the complementary primer strand by multiple copies of the sequence encoded in the template by RCA. Subsequently, for signal generation the slide was incubated with Cy3-modified oligonucleotides. As expected we could observe significant signal increase only at positions where barcodes complementary to the circular template were present.

Taken together, we introduce barcode-labeled dNTPs as substrates for DNA polymerases. We showed that commercially available DNA polymerases are able to process modified nucleotides that are up to 40-times larger than the natural substrate. The sequence-specific incorporation of barcode-modified nucleotides and the addressability of DNA by the simple hybridization of canonical DNA strands has potential for numerous applications. This method is very adaptable, so different techniques for further DNA manipulation and readout can be exploited, such as biotin-streptavidin chemistry,^[16a] nanoparticles,^[17] or branched DNA amplifiers (e.g. TSA detection kit, bDNA amplifier^[18]). The system has the potential to be expanded to a four-color detection system, using nucleotide analogues carrying unique sequences and the appropriate dye-labeled complementary DNA strands. The beneficial combination of microarray techniques

and sequence-specific introduction of barcode ODN-labels by enzymatic incorporation offers opportunities for future applications.

Keywords: DNA polymerase · enzymatic synthesis · microarray · nucleotides · oligonucleotides

- [1] a) C. Bancroft, T. Bowler, B. Bloom, C. T. Clelland, *Science* **2001**, 293, 1763; b) E. T. Kool, J. C. Morales, K. M. Guckian, *Angew. Chem.* **2000**, 112, 1046; *Angew. Chem. Int. Ed.* **2000**, 39, 990.
- [2] a) J. D. Watson, F. H. Crick, *Nature* **1953**, 171, 737; b) Y. Li, S. Korolev, G. Waksman, *EMBO J.* **1998**, 17, 7514.
- [3] a) R. N. Veedu, B. Vester, J. Wengel, *ChemBioChem* **2007**, 8, 490; b) D. Summerer, A. Marx, *Angew. Chem.* **2001**, 113, 3806; *Angew. Chem. Int. Ed.* **2001**, 40, 3693; c) F. Seela, M. Zulauf, *Chem. Eur. J.* **1998**, 4, 1781; d) P. Kielkowski, H. Macickova-Cahova, R. Pohl, M. Hocek, *Angew. Chem.* **2011**, 123, 8886; *Angew. Chem. Int. Ed.* **2011**, 50, 8727.
- [4] a) Z. R. Zhu, J. Chao, H. Yu, A. S. Waggoner, *Nucleic Acids Res.* **1994**, 22, 3418; b) T. Ohbayashi, M. Kuwahara, M. Hasegawa, T. Kasamatsu, T. Tamura, H. Sawai, *Org. Biomol. Chem.* **2005**, 3, 2463; c) A. R. Kore, *Tetrahedron Lett.* **2009**, 50, 793; d) S. Obeid, M. Yulikov, G. Jeschke, A. Marx, *Angew. Chem.* **2008**, 120, 6886; *Angew. Chem. Int. Ed.* **2008**, 47, 6782; e) T. S. Seo, X. Bai, D. H. Kim, Q. Meng, S. Shi, H. Ruparel, Z. Li, N. J. Turro, J. Ju, *Proc. Natl. Acad. Sci. USA* **2005**, 102, 5926; f) U. Asseline, *Curr. Org. Chem.* **2006**, 10, 491.
- [5] a) D. R. Bentley et al., *Nature* **2008**, 456, 53; b) T. D. Harris et al., *Science* **2008**, 320, 106; c) R. Drmanac et al., *Science* **2010**, 327, 78; d) P. R. Langer, A. A. Waldrop, D. C. Ward, *Proc. Natl. Acad. Sci. USA* **1981**, 78, 6633; e) S. P. Liu, S. H. Weisbrod, Z. Tang, A. Marx, E. Scheer, A. Erbe, *Angew. Chem.* **2010**, 122, 3385; *Angew. Chem. Int. Ed.* **2010**, 49, 3313.
- [6] a) S. H. Um, J. B. Lee, S. Y. Kwon, Y. Li, D. Luo, *Nat. Protoc.* **2006**, 1, 995; b) N. J. Woodland, B. Silver, US 2612994, **1952**; c) J. Waugh, *Bioessays* **2007**, 29, 188; d) J. M. Nam, S. I. Stoeva, C. A. Mirkin, *J. Am. Chem. Soc.* **2004**, 126, 5932; e) J. M. Nam, C. S. Thaxton, C. A. Mirkin, *Science* **2003**, 301, 1884.
- [7] S. H. Weisbrod, A. Marx, *Chem. Commun.* **2008**, 5675.
- [8] a) S. Obeid, A. Baccaro, W. Welte, K. Diederichs, A. Marx, *Proc. Natl. Acad. Sci. USA* **2010**, 107, 21327; b) O. Thum, S. Jager, M. Famulok, *Angew. Chem.* **2001**, 113, 4112; *Angew. Chem. Int. Ed.* **2001**, 40, 3990; c) G. F. Kaufmann, et al., *Angew. Chem.* **2005**, 117, 2182; *Angew. Chem. Int. Ed.* **2005**, 44, 2144; d) M. Hocek, M. Fojta, *Org. Biomol. Chem.* **2008**, 6, 2233; e) S. Ikonen, H. Macickova-Cahova, R. Pohl, M. Sanda, M. Hocek, *Org. Biomol. Chem.* **2010**, 8, 1194; f) H. Weizman, Y. Tor, *J. Am. Chem. Soc.* **2002**, 124, 1568; g) G. Giller, T. Tasara, B. Angerer, K. Muhlegger, M. Amacker, H. Winter, *Nucleic Acids Res.* **2003**, 31, 2630; h) J. P. Anderson, B. Angerer, L. A. Loeb, *Biotechniques* **2005**, 38, 257; i) P. M. Gramlich, C. T. Wirges, A. Manetto, T. Carell, *Angew. Chem.* **2008**, 120, 8478; *Angew. Chem. Int. Ed.* **2008**, 47, 8350; j) S. Kumar, A. Sood, S. Rao, J. Nelson, US 7109316, **2006**; k) S. Jäger, M. Famulok, *Angew. Chem.* **2004**, 116, 3399; *Angew. Chem. Int. Ed.* **2004**, 43, 3337; l) P. Brázdilová, M. Vrabel, R. Pohl, H. Pivonkova, L. Havran, M. Hocek, M. Fojta, *Chem. Eur. J.* **2007**, 13, 9527; m) F. Seela, Y. M. Chen, *Chem. Commun.* **1996**, 2263; n) G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen, T. Carell, *J. Am. Chem. Soc.* **2006**, 128, 1398; o) P. M. Gramlich, C. T. Wirges, J. Gierlich, T. Carell, *Org.*

- Lett.* **2008**, *10*, 249; p) J. Gierlich, K. Gutsmedl, P. M. Gramlich, A. Schmidt, G. A. Burley, T. Carell, *Chem. Eur. J.* **2007**, *13*, 9486.
- [9] a) J. M. Clark, *Nucleic Acids Res.* **1988**, *16*, 9677; b) J. M. Clark, C. M. Joyce, G. P. Beardsley, *J. Mol. Biol.* **1987**, *198*, 123; c) K. A. Fiala, J. A. Brown, H. Ling, A. K. Kshetry, J. Zhang, J. S. Taylor, W. Yang, Z. Suo, *J. Mol. Biol.* **2007**, *365*, 590; d) H. Hwang, J. S. Taylor, *Biochemistry* **2004**, *43*, 14612; e) J. A. Peliska, S. J. Benkovic, *Science* **1992**, *258*, 1112; f) S. Obeid, N. Blatter, R. Kranaster, A. Schnur, K. Diederichs, W. Welte, A. Marx, *EMBO J.* **2010**, *29*, 1738.
- [10] S. Jäger, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum, M. Famulok, *J. Am. Chem. Soc.* **2005**, *127*, 15071.
- [11] J. G. Bertram, K. Oertell, J. Petruska, M. F. Goodman, *Biochemistry* **2010**, *49*, 20.
- [12] a) Z. Cheglakov, Y. Weizmann, A. B. Braunschweig, O. I. Wilner, I. Willner, *Angew. Chem.* **2008**, *120*, 132; *Angew. Chem. Int. Ed.* **2008**, *47*, 126; b) E. T. Kool, *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 1; c) Z. Deng, Y. Tian, S. H. Lee, A. E. Ribbe, C. Mao, *Angew. Chem.* **2005**, *117*, 3648; *Angew. Chem. Int. Ed.* **2005**, *44*, 3582; d) J. S. Hartig, S. Fernandez-Lopez, E. T. Kool, *ChemBioChem* **2005**, *6*, 1458; e) C. Lin, M. Xie, J. J. Chen, Y. Liu, H. Yan, *Angew. Chem.* **2006**, *118*, 7699; *Angew. Chem. Int. Ed.* **2006**, *45*, 7537; f) C. Lin, X. Wang, Y. Liu, N. C. Seeman, H. Yan, *J. Am. Chem. Soc.* **2007**, *129*, 14475.
- [13] H. Davies et al., *Nature* **2002**, *417*, 949.
- [14] a) S. A. DelRio-LaFreniere, R. C. McGlennen, *Mol. Diagn.* **2001**, *6*, 201; b) H. Engel, L. Zwang, H. H. van Vliet, J. J. Michiels, J. Stibbe, J. Lindemans, *Thromb. Haemostasis* **1996**, *75*, 267.
- [15] X. Wei, H. L. McLeod, J. McMurrrough, F. J. Gonzalez, P. Fernandez-Salguero, *J. Clin. Invest.* **1996**, *98*, 610.
- [16] a) R. Kranaster, A. Marx, *Angew. Chem.* **2009**, *121*, 4696; *Angew. Chem. Int. Ed.* **2009**, *48*, 4625; b) R. Kranaster, P. Ketzer, A. Marx, *ChemBioChem* **2008**, *9*, 694; c) J. Gaster, G. Rangam, A. Marx, *Chem. Commun.* **2007**, 1692.
- [17] J. J. Storhoff, A. D. Lucas, V. Garimella, Y. P. Bao, U. R. Muller, *Nat. Biotechnol.* **2004**, *22*, 883.
- [18] a) D. Kern et al., *J. Clin. Microbiol.* **1996**, *34*, 3196; b) M. L. Collins et al., *Nucleic Acids Res.* **1997**, *25*, 2979.