DNA Recognition

DOI: 10.1002/ange.200600804

Four-Color, Enzyme-Free Interrogation of DNA Sequences with Chemically Activated, 3′-Fluorophore-Labeled Nucleotides**

*Niels Griesang, Kerstin Gießler, Tanja Lommel, and Clemens Richert**

The technology for interrogating nucleic acid sequences affects the progress of genetics, medical diagnostics, and biotechnology. A recent review on DNA sequencing technologies states: "Current sequencing technologies are too expensive, labor intensive, and time consuming for broad application in human sequence variation studies".[1] The vision of the "$1000 human genome" calls for significant improvements over existing sequencing and genotyping techniques.[2] The benchmark method for interrogating DNA is Sanger sequencing[3] with fluorophore-labeled dideoxynucleoside triphosphates[4] and capillary electrophoresis, monitored by a charge-coupled-device (CCD) camera. It involves enzymes and substrates that are costly. Furthermore, it produces an oligonucleotide for every position to be interrogated in reactions that must be run in individual vessels. An entirely chemical approach to sequencing by synthesis promises a less-expensive access to sequence information, particularly if one was to perform a nonenzymatic primer extension in a microarray format.

To the best of our knowledge, no enzyme-free protocol for sequencing by stepwise primer extension currently exists. Nonenzymatic primer extension has been previously demonstrated, however. Studies focused on prebiotic chemistry have shown that RNA strands complementary to a template can be formed in the absence of enzymes.[5] A number of chemically activated nucleic acids have been shown to undergo chemical primer extension,[6] and a number of related reactions have been shown to occur in a templated fashion.[7] A chemical primer-extension method for nonenzymatic genotyping of single-nucleotide polymorphisms was recently reported.[8] This method employs azaoxybenzotriazolides of 2′-deoxynucleotides (OAt-dNMPs) as activated monomers and short "helper oligonucleotides" to achieve primer conversion within hours,[9] but requires mass spectrometry as a read-out technique.

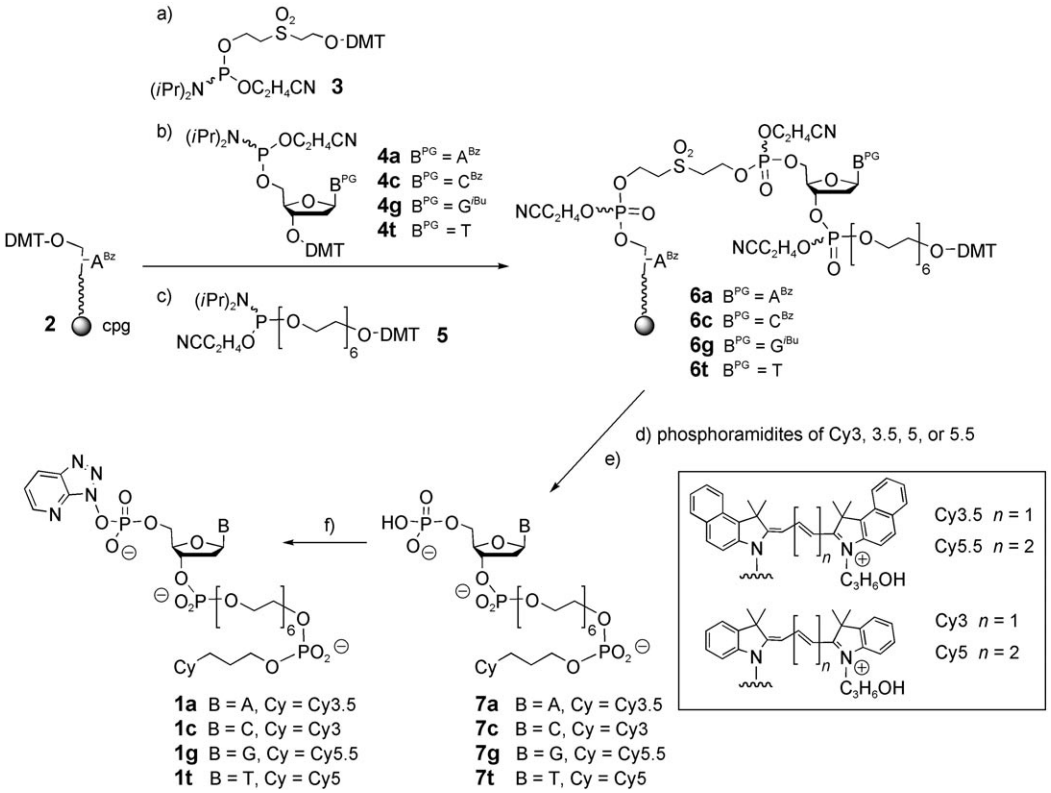
Herein we report that OAt-dNMP-driven primer extensions can be carried out with fluorophore-labeled monomers. The labeled nucleotides allow for optical read out.[10] The focus of this study was on genotyping, that is, the extension of primers by a single nucleotide. Syntheses for all four dNMPs, each labeled with a different dye, are reported together with a protocol for their activation. We also report the first activated monomers in which the fluorophore is attached through a photolabile linker.[10c] This allows it to be removed after the optical read out through irradiation with UV light.

Our first efforts to prepare activated, fluorophore-labeled mononucleotides for sequence determination led to an imidazolide of thymidine 5′-monophosphate with a directly linked carboxyfluorescein moiety in a lengthy, solution-phase synthesis.[11] This imidazolide showed low reactivity in primer-extension assays, prompting us to use azaoxybenzotriazolides[8] for the current study. The current nucleotides (**1a**, **c**, **g**, **t**, Scheme 1) also feature a hexaethyleneglycol linker to minimize steric interference from the dyes. The fluorophores were appended to the 3′-position to prevent undesired multiple extension reactions. The ability to decorate the 3′-position with a substituent is one of the advantages of chemical primer extension. The active sites of polymerases, on the other hand, put an α helix on the base pair between the incoming dNTP and the templating nucleotide[12] leaving little space for substituents or protecting groups at the 3′-position,

[*] Dr. N. Griesang, Dipl.-Chem. K. Gießler, T. Lommel, Prof. C. Richert
Institut für Organische Chemie
Universität Karlsruhe (TH)
76131 Karlsruhe (Germany)
Fax: (+49) 721-608-4825
E-mail: cr@rrg.uka.de
Homepage: http://chip.chemie.uni-karlsruhe.de

[**] The authors thank U. Plutowski, C. Ahlborn, and S. Vogel for helpful discussions. This work was supported by the DFG (grant No. 1063/1-3) and the Fonds der Chemischen Industrie.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.



Scheme 1. Synthesis of labeled active esters of mononucleotides **1a**, **c**, **g**, **t**. Conditions: a) Coupling cycle with **3**, b) coupling cycle with **4a**, **c**, **g**, **t**, c) coupling cycle with **5**, d) coupling cycle with phosphoramidites of cyanine dyes, e) NH_4OH, f) HOAt, EDC·HCl; pH 5. Bz = benzoyl, PG = protecting group.

which is the most obvious position for labeling and (reversibly) blocking extensions.[13]

The synthesis of active esters **1a**, **c**, **g**, **t** (Scheme 1) started from controlled pore glass (cpg) loaded with a deoxynucleoside presenting a dimethoxytrityl(DMT)-protected hydroxy group.[14] Four phosphoramidite coupling cycles, including oxidation, capping, and detritylation, were run on a 2-μmol scale on a DNA synthesizer using phosphorylating reagent **3**,[15] 5′-phosphoramidites **4a**, **c**, **g**, **t**,[16,17] hexaethyleneglycol linker reagent **5**,[18] and the phosphoramidite of one of the four fluorophores chosen (Cy3, 3.5, 5, or 5.5). Nucleotides **7a**, **c**, **g**, **t** were obtained in 17–56% overall yields and were characterized by NMR spectroscopy, MS, and UV/Vis spectroscopy. The normalized fluorescence spectra are shown in Figure 1.

Each of the labeled nucleotides were converted into its active ester (**1a**, **c**, **g**, **t**). Best results were obtained with aqueous buffers containing *N*′-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide hydrochloride (EDC)[19] and hydroxyazabenzotriazole.[8] Compounds **1a**, **c**, **g**, **t** are (necessarily) labile and evidence for their formation was obtained by MALDI MS (see the Supporting Information). Solutions of **1a**, **c**, **g**, **t**, when diluted, were directly used for primer extension.

Primer **8**, featuring a 3′-amino-2′,3′-dideoxynucleotide at the 3′-terminus, was employed for assays (Scheme 2), as amines are known to be reactive toward chemically activated nucleotides and nucleic acids.[20] Assays used equimolar mixtures of OAt-esters **1a**, **c**, **g**, **t** (3.6 mM each), a buffer

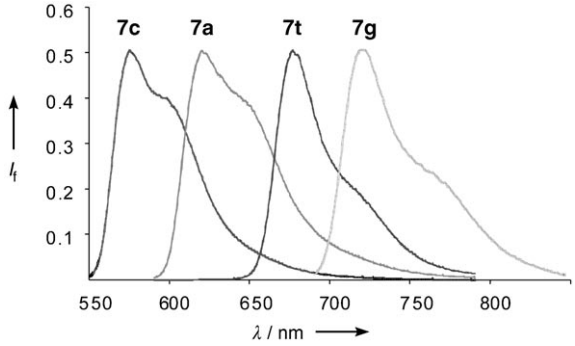
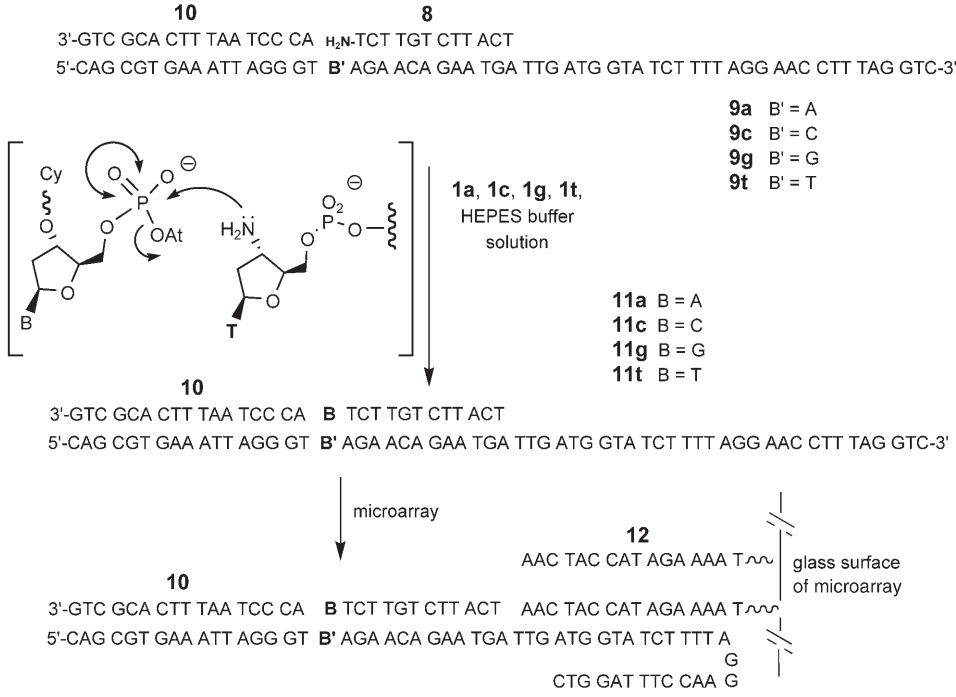


Figure 1. Fluorescence spectra of **7a**, **c**, **g**, **t**. Conditions: **7a** (0.39 μM in water), excitation wavelength: 585 nm; **7c** (0.31 μM in water) excitation wavelength: 545 nm; **7g** (0.23 μM in water) excitation wavelength: 680 nm; **7t** (0.18 μM in water), excitation wavelength: 643 nm. I_f = fluorescence intensity.

containing 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES; 0.2 M), NaCl (0.4 M), and $MgCl_2$ (0.08 M) at pH 8.9, and were performed at room temperature. Four templates (**9a**, **c**, **g**, **t**), displaying any of the four nucleobases at the site to be interrogated, were reacted with OAt-dNMPs in the presence of downstream-binding helper oligonucleotide **10**, which helped to retain nucleotides at the reaction site.[8,9] First, activated nucleotides without fluorophore labels were reacted to establish the kinetics of primer extension for the sequence motif chosen under conditions identical to those used earlier.[8] Half-life times for primer conversion between



Scheme 2. Determining nucleotides in DNA templates through non-enzymatic primer extension with fluorophore-labeled, activated nucleotides **1a**, **c**, **g**, **t**. Oligodeoxynucleotide **10** serves as a "helper oligonucleotide",[8] generating additional stacking surfaces for the activated nucleotide. The nucleobase in the template (B′) can be determined through fluorescence (Figure 3).

0.13 and 2.5 h and sequence-selective reactions were measured (see the Supporting Information). Next, 90 pmol of each template was reacted with an equimolar mixture of fluorescent nucleotides **1a**, **c**, **g**, **t** (3.6 mM), and the formation of extended, fluorophore-labeled primers **11a**, **c**, **g**, **t** was measured.

MALDI mass spectra confirmed that in each case the correct nucleotide was incorporated opposite the templating base with high fidelity (Figure 2). The same result was obtained when the oligonucleotides were bound on microarrays[21] featuring capture oligonucleotides (**12**) on spots of their surface. After incubation and washing, scans of the microarray surface with four different filter settings and appropriate corrections for the relative fluorescence intensities of the cyanine dyes allowed for unambiguous base calls for each of the templates (Figure 3). (The filter settings and correction factors had been established separately in calibration experiments with Cy-labeled oligonucleotides; see the Supporting Information.) Furthermore, primer extension with the unrelated template sequence 5′-GGAATCACACGTGCG-3′ and labeled monomers **1a**, **c**, **g**, **t** again showed high selectivity (see the Supporting Information).

Finally, we performed exploratory assays to determine the potential for reading more than one nucleotide in the target strands. For this, **13** was prepared, whose cyanine fluorophore is linked to the 3′-amino-2′,3′-dideoxynucleotide through a photolabile linker[10c] (Scheme 3, see also the Supporting Information). Thymidines, such as **13**, are the most difficult cases for chemical primer extension, as their nucleobase

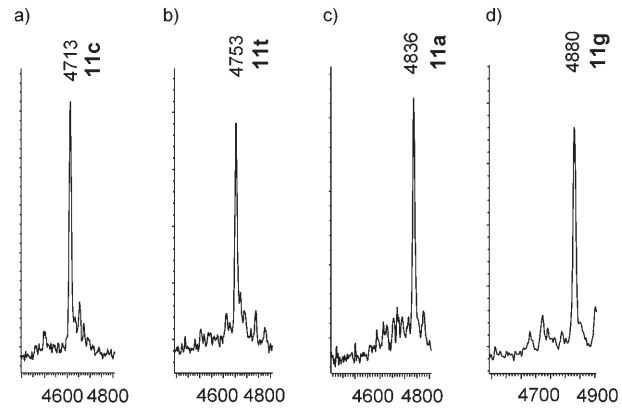


Figure 2. MALDI-TOF spectra of products of primer extension with labeled nucleotides **1a**, **c**, **g**, **t** at a 36 μM concentration of oligonucleotides, after 14 h at room temperature; a) template **9g**, b) template **9a**, c) template **9t**, d) template **9c**.

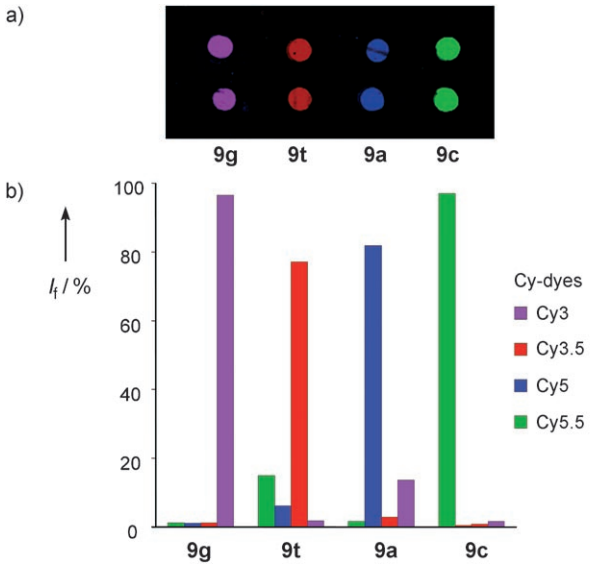
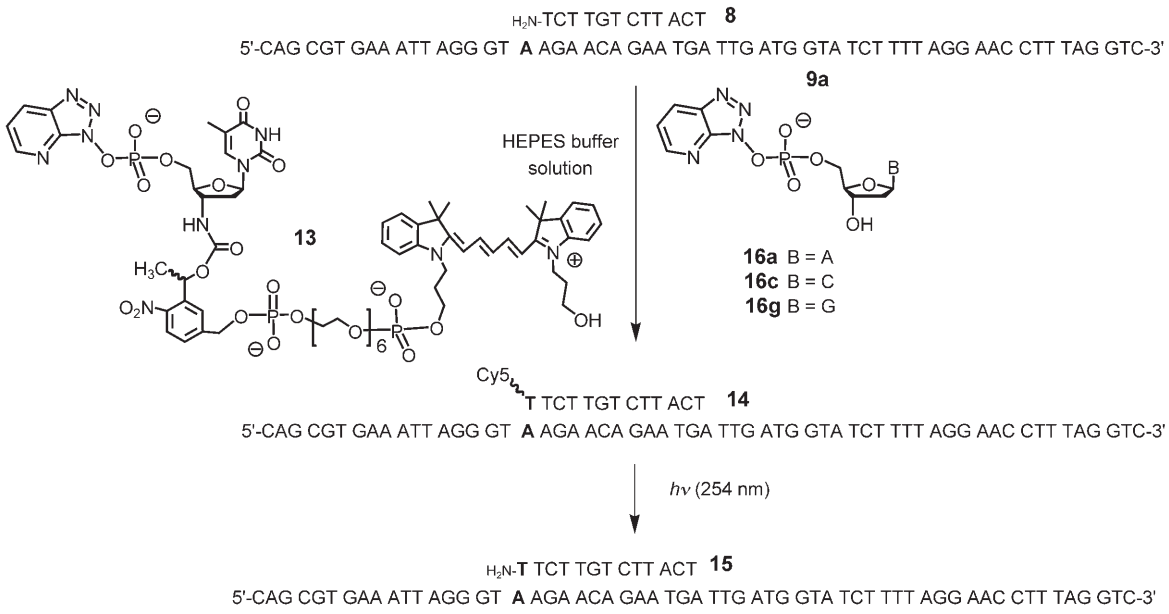


Figure 3. Scans of a microarray with the capture oligonucleotide **12** after hybridization with products from primer extensions with templates **9a**, **9c**, **9g**, or **9t**. a) Overlay of false-color fluorescence images, generated with the settings for each of the fluorophores (Cy3/3.5/5/5.5) using filters, correction factors, and exposure times established in calibration experiments (see the Supporting Information). Products were analyzed in duplicate (two spots vertically above each other). b) Integration of the fluorescence signal obtained in every channel for each pair of spots; the small bars accompanying those for the correct signals are believed to result from mismatch incorporation.



Scheme 3. Enzyme-free primer extension with **13**, whose fluorophore is linked through a photolabile linker. In a competitive reaction with equimolar amounts of **16a**, **c**, **g** and **13** (3.6 mM/18 mM), the residue of **13** is appended to primer **8** to give extended primer **14**. The adenosine base in template **9a** was successfully identified through a fluorescence scan on a microarray (not shown). Subsequent irradiation of **14**:**9a** with UV light cleaves the photolabile linker, removing the Cy5 label and producing the amino-terminal primer **15**, which is suitable for a subsequent primer-extension reaction. Compare Figure 4 for mass spectra.

forms only two hydrogen bonds and the binding site at the template offers little stacking surface for the pyrimidine. Even for the unlabeled OAt-dNMPs, primer extension is, therefore, an order of magnitude slower than for the strongly base pairing nucleotides (see the Supporting Information). Despite the odds, which were worsened by the lack of a helper oligonucleotide in the "multiple-extension case," **13** reacted selectively with primer **8** to give **14** in the presence of competitor nucleotides **16a**, **c**, **g** (Scheme 3), as seen in MALDI-TOF mass spectra (Figure 4) and fluorescence images of a microarray read-out experiment. When the oligonucleotide complex containing **14** was irradiated with UV light (254 nm) from a simple, hand-held lamp for 5 min, the linker with the fluorophore was removed, generating the extended primer **15**, which was suitable for a subsequent interrogation of the next position in the template (Figure 4). An exploratory assay with the deoxycytidine analogue of **13** showed that a dual primer extension is feasible (see the Supporting Information). We are in the process of optimizing conditions to ensure full primer conversion, even for weakly base-pairing nucleotides, and developing a capping scheme that ensures blocking of failure strands.

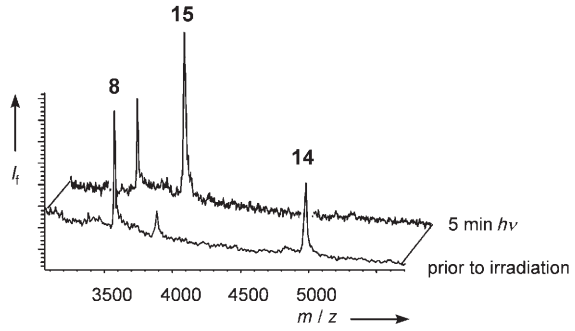


Figure 4. MALDI-TOF mass spectra of samples drawn from the reaction shown in Scheme 3. Assay conditions: template (**9a**; 36 μM) and primer (**8**); HEPES buffer solution (0.2 M), NaCl (0.4 M), and $MgCl_2$ (0.08 M), pH 8.9, 26 h, RT. Please note that longer sequences give smaller signals as a consequence of lower desorption/ionization efficiencies.

Together, these results suggest that chemical primer extension with inexpensive, fluorescent substrates is suitable for determining nucleotides in DNA sequences, including read-out in a parallel fashion through microarrays. In favorable cases, the colors of the fluorophores may be detected with the naked eye, providing a simplified form of the current assays (Scheme 2) that might serve as a "lacmus test" for detecting single-nucleotide polymorphisms. The more-sophisticated form involving microarrays and a chip scanner might prove an inexpensive alternative to current genotyping systems. Modifications in the helper[22] and primer,[23] as well as adjustments in the nucleotide concentration and activation chemistry might eventually lead to assays that go to completion in minutes, rather than in hours. Though far from competitive at the moment, assay formats involving removable fluorophores might thus become suitable for reading short stretches of DNA. Photolabile linkers have the advantage of enabling the use of light to control when the next nucleotide is admitted. Unlike most chemical deprotection agents, light does not induce the dissociation of primer–template duplexes. We are currently working on syntheses for nucleotides containing all four nucleobases (A/C/G/T) with photolably linked fluorophores. Independent of

further technical advances, the current results again demonstrate that nucleic acids in themselves have a strong propensity to undergo spontaneous replication reaction sequence selectively.

Received: March 1, 2006
Published online: August 22, 2006

Keywords: DNA recognition · DNA replication · oligonucleotides · primer extension · single-nucleotide polymorphism

[1] M. L. Metzker, *Genome Res.* **2005**, *15*, 1767 – 1776.
[2] a) J. Shendure, R. D. Mitra, C. Varma, G. M. Church, *Nat. Rev. Genet.* **2004**, *5*, 335 – 344; b) C. H. Arnaud, *Chem. Eng. News* **2005**, *83*(51), 60 – 61.
[3] F. Sanger, G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. A. Hutchison III, P. M. Slocombe, M. Smith, *Nature* **1977**, *265*, 687 – 695.
[4] B. B. Rosenblum, L. G. Lee, S. L. Spurgeon, S. H. Khan, S. M. Menchen, C. R. Heiner, S. M. Chen, *Nucleic Acids Res.* **1997**, *25*, 4500 – 4504.
[5] I. A. Kozlov, L. E. Orgel, *Mol. Biol.* **2000**, *34*, 781 – 789.
[6] Selected references: a) I. A. Kozlov, S. Pitsch, L. E. Orgel, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13448 – 13452; b) M. Koppitz, P. E. Nielsen, L. E. Orgel, *J. Am. Chem. Soc.* **1998**, *120*, 4563 – 4569; c) C. Hartel, M. W. Göbel, *Helv. Chim. Acta* **2000**, *83*, 2541 – 2549; d) J. C. Chaput, C. Switzer, *J. Am. Chem. Soc.* **2000**, *122*, 12866 – 12867; e) I. A. Kozlov, M. Zielinski, B. Allart, L. Kerremans, A. Van Aerschot, R. Busson, P. Herdewijn, L. E. Orgel, *Chem. Eur. J.* **2000**, *6*, 151 – 155; f) X. Wu, G. Delgado, R. Krishnamurthy, A. Eschenmoser, *Org. Lett.* **2002**, *4*, 1283 – 1286; g) S. Ficht, A. Mattes, O. Seitz, *J. Am. Chem. Soc.* **2004**, *126*, 9970 – 9981.
[7] X. Li, D. R. Liu, *Angew. Chem.* **2004**, *116*, 4956 – 4979; *Angew. Chem. Int. Ed.* **2004**, *43*, 4848 – 4870.
[8] P. Hagenbuch, E. Kervio, A. Hochgesand, U. Plutowski, C. Richert, *Angew. Chem.* **2005**, *117*, 6746 – 6750; *Angew. Chem. Int. Ed.* **2005**, *44*, 6588 – 6592.
[9] Compare also: S. R. Vogel, C. Deck, C. Richert, *Chem. Commun.* **2005**, 4922 – 4924.
[10] For selected recently developed sequencing methods that employ parallel optical read-outs, see, for example: a) J. Shendure, G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra, G. M. Church, *Science* **2005**, *309*, 1728 – 1732; b) M. Margulies, M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y.-J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. I. Alenquer, T. P. Jarvie, K. B. Jirage, J.-B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, J. M. Rothberg, *Nature* **2005**, *437*, 376 – 377; c) 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 – 5931.
[11] N. Griesang, E. Kervio, C. Richert, *Synthesis* **2005**, 2327 – 2334.
[12] S. Doublie, S. Tabor, A. M. Long, C. C. Richardson, T. Ellenberger, *Nature* **1998**, *391*, 251 – 258.
[13] H. Ruparel, L. Bi, Z. Li, X. Bai, D. H. Kim, N. J. Turro, J. Ju, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5932 – 5937.
[14] The nucleoside-loaded cpg was chosen as an inexpensive source of a DMT-protected, alcohol-displaying support for automated DNA synthesis, even though the deoxynucleoside loaded onto the support was, in itself, not required. The first nucleoside does serve as a useful internal standard for HPLC analyses of crudes, however.
[15] T. Horn, M. S. Urdea, *Tetrahedron Lett.* **1986**, *27*, 4705 – 4708.
[16] E. Uhlmann, J. Engels, *Tetrahedron Lett.* **1986**, *27*, 1023 – 1026.
[17] W. H. Connors, S. Narayanan, O. P. Kryatova, C. Richert, *Org. Lett.* **2003**, *5*, 247 – 250.
[18] S. Altmann, A. M. Labhardt, D. Bur, C. Lehmann, W. Bannwarth, M. Billeter, K. Wüthrich, W. Leupin, *Nucleic Acids Res.* **1995**, *23*, 4827 – 4835.
[19] K. J. Prabahar, T. D. Cole, J. P. Ferris, *J. Am. Chem. Soc.* **1994**, *116*, 10914 – 10920.
[20] a) R. Lohrmann, L. E. Orgel, *Nature* **1976**, *261*, 342 – 344; b) W. S. Zielinski, L. E. Orgel, *Nucleic Acids Res.* **1987**, *15*, 1699 – 1715; c) W. S. Zielinski, L. E. Orgel, *Nature* **1987**, *327*, 346 – 347; d) A. Kanavarioti, M. W. Stronach, R. J. Ketner, T. B. Hurley, *J. Org. Chem.* **1995**, *60*, 632 – 637; e) A. Luther, R. Brandsch, G. von Kiedrowski, *Nature* **1998**, *396*, 245 – 248.
[21] The microarrays were prepared as described in: K. L. Dombi, N. Griesang, C. Richert, *Synthesis* **2002**, 816 – 824, with the modifications mentioned in: Z. Dogan, R. Paulini, J. A. Rojas Stütz, S. Narayanan, C. Richert, *J. Am. Chem. Soc.* **2004**, *126*, 4762 – 4763.
[22] J. A. Rojas Stütz, C. Richert, *Chem. Eur. J.*, **2006**, *12*, 2472 – 2481.
[23] P. Baumhof, N. Griesang, M. Bächle, C. Richert, *J. Org. Chem.* **2006**, *71*, 1060 – 1067.

