DNA Replication

DOI: 10.1002/ange.200501794

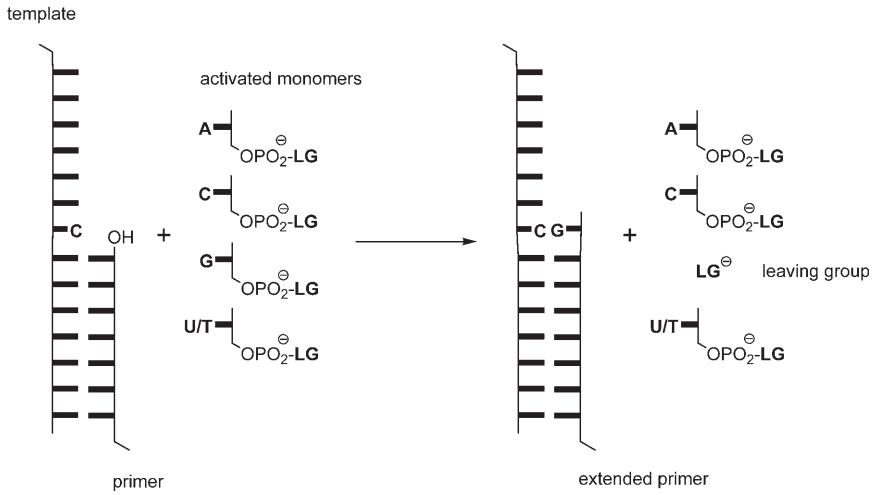
Chemical Primer Extension: Efficiently Determining Single Nucleotides in DNA**

*Patrizia Hagenbuch, Eric Kervio, Annette Hochgesand, Ulrich Plutowski, and Clemens Richert**

The replication and transcription of genetic information requires the stepwise extension of DNA or RNA primers in sequence-specific, template-directed reactions (Scheme 1). Template-directed primer extension also underlies PCR,[1] dideoxy sequencing,[2] and attempts to recreate life from inanimate materials.[3] Primer extension presents a formidable chemical challenge. Four different, richly functionalized monomers (the mononucleotides) have to be converted into good electrophiles that react in sequence-selective oligomerization reactions in aqueous solution without polymerizing or hydrolyzing rapidly. Nature employs pyrophosphates as leaving groups and enzymes with proofreading capabilities (polymerases) to meet this challenge. Entirely chemical, non-enzymatic systems employing imidazolides as leaving groups show a certain level of spontaneous primer extension, but successful replication has remained elusive.[3]

Known non-enzymatic replication reactions with monomers are slow. For ribonucleotides and 2-methylimidazolides as activated monomers, reaction times of days are required, even if the monomers are employed near their solubility limit ($\geq$ 50 mM)[4] under extreme salt conditions (up to 1M Mg^{2+}).[5] Weakly base pairing nucleotides (A or U/T as nucleobase) are incorporated more slowly than those that form stable base pairs (C or G as nucleobase).[6] Even the ligation of oligonucleotides requires the templating of G and C residues.[7] This makes non-enzymatic primer extension seemingly unattractive for practical applications, such as the genotyping of single-nucleotide polymorphisms (SNPs) by mass spectrometry.[8] Herein we demonstrate how these reactions can be accelerated, so that the determination of nucleotides at selected sites of DNA within hours becomes feasible, starting from subpicomole quantities of analytes.

We first used 2-methylimidazolides[9] (LG = MeIm) of 2′-deoxynucleosides **1a–t** as activated monomers **2a–t** (Scheme 2), which were treated with 3′-amino-terminal primers. Amines are known to react faster than alcohols with activated nucleotides,[10,11] so that singly extended primers rather than the product mixtures typical for chemical replication reactions are formed. Assays were monitored by



Scheme 1. Primer extension.

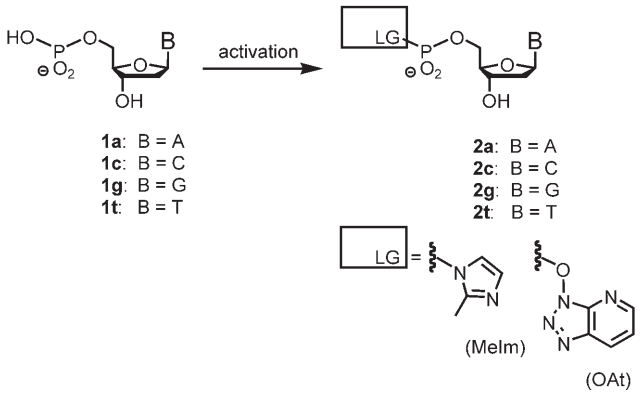
quantitative MALDI-TOF mass spectrometry,[12] thus allowing competitive reactions with all four activated nucleotides in the same solution and selective detection of products in mass spectra.[13] Second-order rate constants were derived from fits to the kinetic data.[13] We tested whether non-enzymatic primer extension was detectable with any of the four possible bases (A, C, G, T) at the templating position of 40-mer templates **4a–t** (Scheme 3).

Sequence-selective extension was observed for each of the four reactions (Table 1). Although this established that chemical primer extension was sufficiently selective to

[*] Dr. P. Hagenbuch, Dr. E. Kervio, A. Hochgesand, U. Plutowski, Prof. C. Richert
Institut für Organische Chemie
Universität Karlsruhe (TH)
76131 Karlsruhe (Germany)
Fax: (+49) 721-608-4825
E-mail: cr@rrg.uka.de

[**] The authors thank Dr. J. Rojas Stütz for helpful discussions. This work was supported by the DFG (grant no. RI 1063/1-3) and the Fonds der Chemischen Industrie (project 164431).

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



Scheme 2. Activated monomers **2a–t**.

Table 1: Results from primer extension reactions with a mixture of the four 2-methylimidazolides **2a–t**/MeIm.[a]

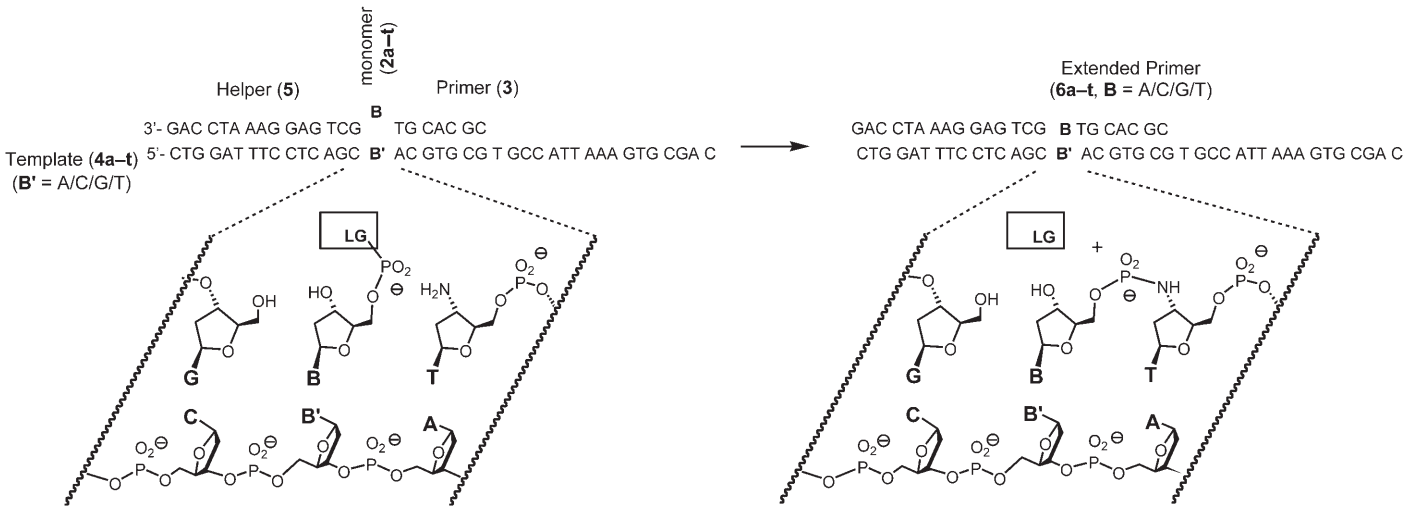
Primer/ template	k' [h^{-1} M^{-1}][b]	Product ratio (primer+C/T/A/G)
3/4a	46	3:**84**:5:8
3/4c	84	4:5:8:**83**
3/4g	38	**84**:5:10:1
3/4t	43	6:9:**74**:11

[a] [**2a–t**] = 19.3 mM; primer/template: 360 μM; pH 7.9; buffer: HEPES (200 mM), NaCl (400 mM), $MgCl_2$ (80 mM). [b] k' = second-order rate constant for the formation of desired product; oligonucleotide complexes treated as one reactant, activated nucleotide as the other.

identify a templating nucleotide in DNA, the assay was not attractive for genotyping or analyzing epigenomic methylation patterns,[14,15] as it requires 1.8 nmol of template for a 5-μL assay, that is, more template than is available in typical clinical samples after PCR. We therefore decreased the concentration of template and primer tenfold and reduced the concentration of the activated monomers from 19.3 mM to 3.6 mM. Dilution favors hydrolysis of the activated nucleotides, an unavoidable competing reaction, and disfavors the association of monomers and template. Under these conditions, even a strongly base pairing nucleotide at the templating position, when offered the matched monomer alone, gave a second-order rate constant that was almost an order of magnitude lower than that measured at higher concentration (compare Table 1, entry 3 with Table 2, entry 1). In the absence of competing reactions and/or changes in binding equilibria, second-order rate constants should, of course, be independent of concentration.

We then focused on identifying alternative leaving groups for the activated deoxynucleotides. Ferris and co-workers had previously shown that oligomerization of ribonucleotides on mineral surfaces can be accelerated through judicious choice of leaving groups,[16] but template-directed reactions have proven to be a challenge to accelerate.[17] With EDC[18] alone, template **4a** gives non-sequence-selective extensions. Attempts to improve extension reactions through activation in situ with EDC in the presence of possible leaving groups or through addition of organic catalysts to methylimidazolides failed (see Supporting Information). Proflavine[19] increased the conversion with **2c**–MeIm minimally from 16 to 21% after 4 h, and this effect was observed only at the highest concentration tested (0.67 mM).

Therefore, we focused on the activation of **1c** in organic solvents and the addition of the activated monomers to **3**/**4g** in aqueous buffer. The tetrazolide of **2c** gave a slight acceleration over **2c**–MeIm, but the only substantial acceleration was observed after activation with HATU[20] and HOAt, reagents originally developed for peptide coupling. Optimized activation conditions gave an active ester in satisfactory yield after precipitation (see Supporting Information). Azabenzotriazolide **2c**–OAt reacts four times faster than **2c**–MeIm with **3**/**4g** (Table 2) and does not undergo significant side reactions with the primer. Interestingly, an even more dramatic acceleration of primer extension is observed in an all-RNA system when switching from methylimidazolides to azabenzotriazolides, even though the nucleophile is a hydroxy group in that case.[21]



Scheme 3. Primer-extension system employed.

Table 2: Effect of activation of monomers, helper, and pH shift on the rate of non-enzymatic single-nucleotide-extension reactions.[a]

Oligomers	Monomer(s)/ LG/reagent[b]	pH	k' [h^{-1} M^{-1}][c]	Product ratio (primer + C/T/A/G)[d]	$t_{1/2}$ primer [h]
noncompetitive reactions					
3/**4g**	**2c**/MeIm	7.9	5	–	38.5
3/**4g**	**2c**/OAt	7.9	19	–	10.0
3/**4g**/**5**	**2c**/MeIm	7.9	20	–	9.5
3/**4g**/**5**	**2c**/OAt	7.9	112	–	1.7
3/**4g**/**5**	**2c**/OAt[e]	7.9	205	–	0.9
3/**4g**/**5**	**2c**/OAt[e]	8.9	381	–	0.5
3/**4g**/**5**	**2c**/OAt[e]	9.5	393	–	0.5
competitive reactions					
3/**4a**/**5**	**2a**–**t**/OAt	8.9	94	4:**64**:14:18	2.0
3/**4c**/**5**	**2a**–**t**/OAt	8.9	259	1:1:5:**93**	0.7
3/**4g**/**5**	**2a**–**t**/OAt	8.9	212	**86**:1:8:5	0.9
3/**4t**/**5**	**2a**–**t**/OAt	8.9	217	1:1:**89**:9	0.9
competitive reactions, one-pot activation/extension					
3/**4a**/**5**	**2a**–**t**/OAt/EDC	8.9	81	4:**75**:18:3	2.4
3/**4c**/**5**	**2a**–**t**/OAt/EDC	8.9	177	1:3:1:**95**	1.1
3/**4g**/**5**	**2a**–**t**/OAt/EDC	8.9	230	**97**:1:1:1	0.8
3/**4t**/**5**	**2a**–**t**/OAt/EDC	8.9	134	1:1:**97**:1	1.4
competitive one-pot assays with dT-boosted monomer mixture					
3/**4a**/**5**	**2a**–**t**/OAt/EDC	8.9	47	2:**90**:5:3	0.8
3/**4c**/**5**	**2a**–**t**/OAt/EDC	8.9	146	1:1:1:**97**	1.3
3/**4g**/**5**	**2a**–**t**/OAt/EDC	8.9	217	**97**:1:1:1	0.9
3/**4t**/**5**	**2a**–**t**/OAt/EDC	8.9	117	1:1:**97**:1	1.6

[a] Template and primer: 36 μM; activated **2**: 3.6 mM; aqueous solution; buffer: HEPES (200 mM), NaCl (400 mM), $MgCl_2$ (80 mM). [b] LG = leaving group; MeIm = 2-methylimidazole; OAt = oxyazabenzotriazolide, EDC = excess carbodiimide from one-pot activation/extension assay. [c] k' = second-order rate constant; oligonucleotide complexes treated as one reactant, activated nucleotide as the other. [d] Perfectly matched primer-extension products are highlighted in boldface. [e] Activated monomer purified by HPLC.

Next, we studied whether an additional or helper oligonucleotide **5** that binds immediately downstream from the templating base would create a tighter binding site for the incoming monomer (Scheme 3). The rate was accelerated several fold for both **2c**–MeIm and **2c**–OAt (Table 2). For azabenzotriazole-activated **2c**–OAt, the rate enhancement over the helper-free reaction with **2c**–MeIm is 22-fold. A further twofold acceleration resulted from purification of the azabenzotriazolide **2c**–OAt by HPLC prior to use in our assay, an effort invested only for the kinetic study. The azabenzotriazolide-driven primer-extension reaction benefits further from elevated pH values. At pH 9.5, the acceleration of the helper-assisted reaction over that of the unassisted reaction with **2c**–MeIm is 79-fold (Table 2). A pH value of 8.9 gives a similar effect, but lowers the risk of losing sequence selectivity through deprotonation of dT and dG.

We then established HATU/HOAt activation for monomers **1a**, **1g**, and **1t** (CH_3CN was required as solvent for **1t** as cleaner activations were observed when starting from slurries). Hydrolysis of OAt esters, despite their reactivity towards amino and ribo primers,[21] is sufficiently slow. An exploratory study of the hydrolysis of **2t**–OAt in the assay buffer in D_2O by NMR spectroscopy shows a half-life of 16.2 h (see Supporting Information). Mixtures of all four OAt esters employed in competitive reactions give rapid primer extension and acceptable sequence selectivities for all four templates **4a**–**t** (Table 2).

To make the chemical primer extension reactions more attractive for biomedical applications, we investigated one-pot activation/primer-extension protocols. When solutions of **1a**–**t** (0.4 M) were treated with EDC (5 equiv)[16] and HOAt (3 equiv) at pH 5, the solution contained > 90% of OAt esters after 2 h, as determined by ^{31}P NMR. Dilution and adjustment of the pH value to 8.9 followed by addition of DNA strands led to reactions similar to those with HATU-activated OAt esters (Table 2), at least in the presence of **5**. Because the excess of activation reagents allows reactivation of nucleotides, the kinetics of these reactions are more complex, but the one-pot procedure makes assays easier to perform by personnel not trained as synthetic chemists.

The lower reactivity and selectivity of the reaction templated by adenine was then addressed. We increased the concentration of **2t**–OAt fivefold to compensate for the

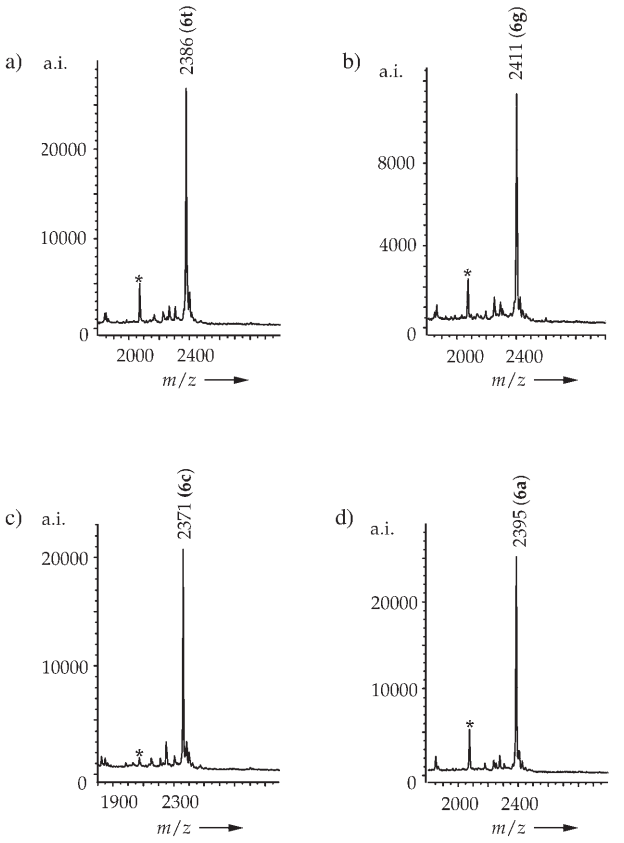


Figure 1. MALDI-TOF mass spectra from competitive primer extension reactions with **3**/**4a**–**t**/**5** and a dT-boosted mixture of **2a**–**t**/OAt at pH 8.9, 8 h. a) Template **4a**, b) template **4c**, c) template **4g**, d) template **4t**. The asterisk denotes the peak for unconverted primer.

lower target affinity. The dT-boosted mixture gives sequence-selective reactions for all four templates **4a–t** in which the correctly extended primer constitutes $\geq 90\%$ of the elongation products. Furthermore, the half-life of the reactions is ≤ 1.6 h and within a factor of 2 for all four reactions (Table 2). More-refined reactivity-adjusted mixtures may make the rate differences even smaller. Representative MALDI spectra and kinetics are shown in Figures 1 and 2.

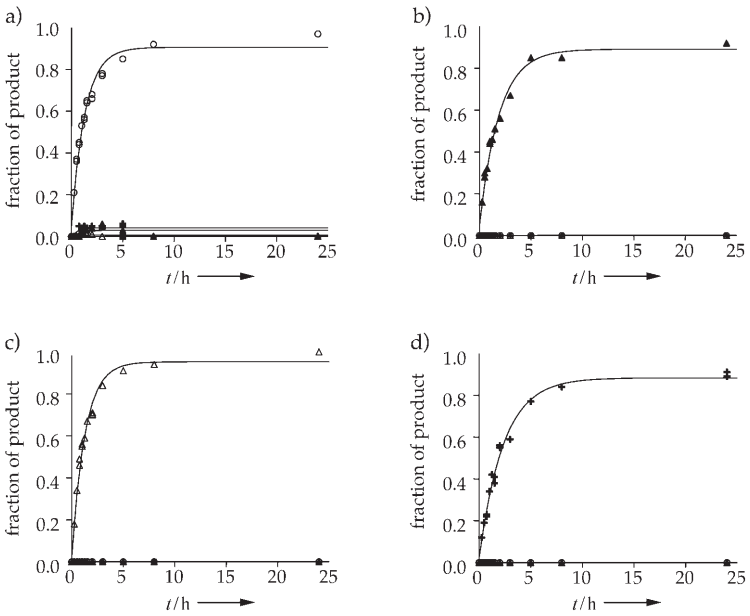


Figure 2. Kinetics of competitive primer-extension reactions with **3**/**4a–t**/**5** and dT-boosted mixture of **2a–t**/OAt at pH 8.9. a) Template **4a**, b) template **4c**, c) template **4g**, d) template **4t**. △ **6c**, ○ **6t**, + **6a**, ▲ **6g**.

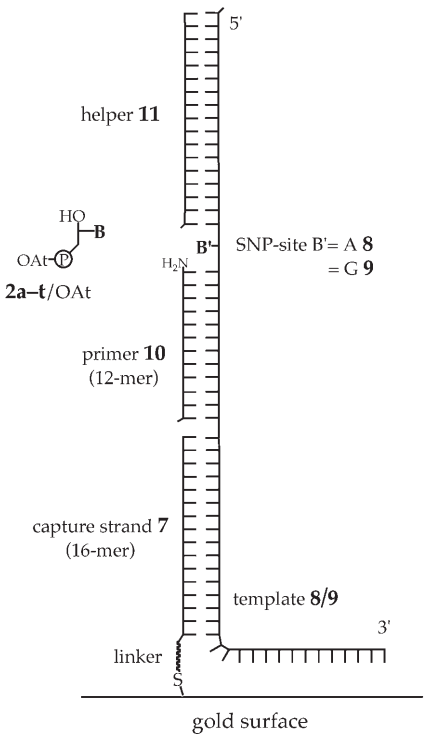


Figure 3. System for determining single-nucleotide identity on DNA microarrays for on-chip MALDI-TOF mass spectrometry.[22] DNA strands used: **7** (capture strand): 5′-taaaagataccatcaa-3′; **8** (dA template): 5′-cagcgtgaaattagggtAagaacagaatgattgatggtatcttttaggaaccttaggtc-3′; **9** (dG template): 5′-cagcgtgaaattagggtGagaacagaatgattgatggtatcttttaggaaccttaggtc-3′; **10** (3′-amino-primer): 5′-tcattctgttct-3′; **11** (helper): 5′-accctaatttcacgctg-3′. Lower-case letters denote deoxynucleotides, SNP sites are highlighted as upper-case letters.

Finally, we performed exploratory primer extension reactions with the microarray system shown schematically in Figure 3. It consists of capture oligonucleotide **7**, immobilized on spots of flame-smoothed gold surfaces background-passivated by oligo(ethylene glycol ether) self-assembled monolayers.[22] The spots were exposed to solutions containing one of two different DNA 60-mers **8** or **9** (which contain either nucleotide A (**8**) or G (**9**) at the site of primer extension), primer **10**, and helper **11**. After hybridization, template-directed reactions were induced that employed EDC-activated mononucleotides (3.6 mM each) in the extension buffer at pH 8.9 and 8 °C. After 160 min, the surface was washed, treated with MALDI matrix solution, and subjected to mass-spectrometric analysis. Figure 4 shows spectra obtained in assays starting from solutions containing templates **8** or **9** (500 fmol). Even at 50 fmoles, an SNP call could be made for the A-template (see Supporting Information).

Our optimized activation of nucleotides is a one-step process employing commercial reagents. No significant non-nucleoside extensions of primers have been detected. Given the rate accelerations observed, practical applications of chemical-primer extension, such as SNP genotyping[23] can be envisioned. Chemical-primer extension on a chip does not require purification after the elongation step and avoids sample transfer, as all steps are performed on the same surface. This should lower the cost of high-throughput assays. Heavily modified nucleotides, including those that form non-natural base pairs, may be incorporated into primers without the constraints imposed by active sites of polymerases. An extension of this work to assays with fluorophore-labeled monomers that allow optical detection on DNA microarrays

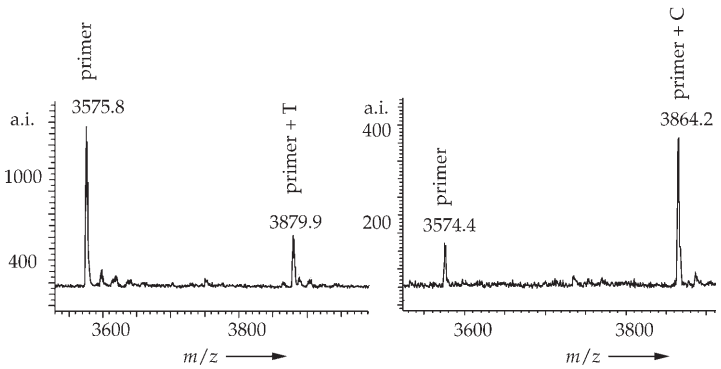


Figure 4. MALDI-TOF mass spectra from primer extensions performed with the system shown in Figure 3 and 60 mer templates **8**/**9** featuring the nucleotides A (left) or G (right) at the SNP site after 160 min. Spectra were obtained directly from the surface on which the primer-extension reaction had been performed. Surfaces had been incubated with 500 fmol of template **8** or **9**. See Supporting Information for further details.

is being actively pursued in these laboratories. A fluorophore-labeled form of activated **1t** suitable for non-enzymatic primer extension reactions was recently described.[24] Exploratory experiments with 3′-Cy3-labeled **2c**–OAt, and **3**/**4g**/**5** gave 42% primer conversion after 1 h and 72% conversion after 3.5 h,[25] demonstrating the feasibility of fluorophore-based genotyping.

Received: May 24, 2005
Revised: July 1, 2005
Published online: September 21, 2005

Keywords: DNA recognition · non-enzymatic replication · oligonucleotides · primer extension · single-nucleotide polymorphisms

[1] K. B. Mullis, *Angew. Chem.* **1994**, *106*, 1271 – 1276; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1209 – 1213.
[2] F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5463 – 5467.
[3] L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99 – 123.
[4] M. Hey, C. Hartel, M. W. Göbel, *Helv. Chim. Acta* **2003**, *86*, 844 – 854.
[5] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 317 – 322.
[6] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 7963 – 7969.
[7] A. Luther, R. Brandsch, G. von Kiedrowski, *Nature* **1998**, *396*, 245 – 248.
[8] J. Tost, I. G. Gut, *Mass Spectrom. Rev.* **2002**, *21*, 388 – 418.
[9] T. Inoue, L. E. Orgel, *J. Am. Chem. Soc.* **1981**, *103*, 7666 – 7667.
[10] L. E. Orgel, R. Lohrmann, *Acc. Chem. Res.* **1974**, *7*, 368 – 377.
[11] A. Kanavarioti, M. W. Stronach, R. J. Ketner, T. B. Hurley, *J. Org. Chem.* **1995**, *60*, 632 – 637.
[12] D. Sarracino, C. Richert, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2543 – 2548.
[13] J. A. Rojas Stütz, C. Richert, *J. Am. Chem. Soc.* **2001**, *123*, 12718 – 12719.
[14] C. Grunau, W. Hindermann, A. Rosenthal, *Hum. Mol. Genet.* **2000**, *9*, 2651 – 2663.
[15] G. Egger, G. N. Liang, A. Aparicio, P. A. Jones, *Nature* **2004**, *429*, 457 – 463.
[16] a) K. J. Prabahar, T. D. Cole, J. P. Ferris, *J. Am. Chem. Soc.* **1994**, *116*, 10914 – 10920; b) K. J. Prabahar, J. P. Ferris, *J. Am. Chem. Soc.* **1997**, *119*, 4330 – 4337; c) W. Huang, J. P. Ferris, *Chem. Commun.* **2003**, 1458 – 1459.
[17] L. Rodriguez, L. E. Orgel, *J. Mol. Evol.* **1991**, *32*, 278 – 281.
[18] J. Sulston, R. Lohrmann, L. E. Orgel, H. T. Miles, *Proc. Natl. Acad. Sci. USA* **1968**, *59*, 726 – 733.
[19] S. S. Jain, F. A. Anet, C. J. Stahle, N. V. Hud, *Angew. Chem.* **2004**, *116*, 2038 – 2042; *Angew. Chem. Int. Ed.* **2004**, *43*, 2004 – 2008.
[20] L. A. Carpino, *J. Am. Chem. Soc.* **1993**, *115*, 4397 – 4398.
[21] S. R. Vogel, C. Deck, C. Richert, *Chem. Commun.* **2005**, DOI: 10.1039/B510775.
[22] U. Plutowski, C. Richert, *Angew. Chem.* **2005**, *117*, 627 – 631; *Angew. Chem. Int. Ed.* **2005**, *44*, 621-625.
[23] K. Tang, D.-J. Fu, D. Julien, A. Braun, C. R. Cantor, H. Köster, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10016 – 10020.
[24] N. Griesang, E. Kervio, C. Richert, *Synthesis* **2005**, 2327 – 2334.
[25] N. Griesang, C. Richert, unpublished results.

