ORGANIC LETTERS

2001
Vol. 3, No. 8
1105−1108

3′-Nitrophenylpropyloxycarbonyl (NPPOC) Protecting Groups for High-Fidelity Automated 5′ → 3′ Photochemical DNA Synthesis

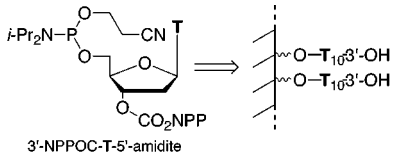
Michael C. Pirrung,* Laixin Wang, and Michael P. Montague-Smith

Department of Chemistry, Levine Science Research Center, Duke University, Box 90317, Durham, North Carolina 27708

pirrung@chem.duke.edu

Received November 22, 2000 *(Revised Manuscript Received March 13, 2001)*

ABSTRACT



3′-NPPOC-T-5′-amidite

The most powerful DNA microarrays would be prepared by photolithography with free 3′-ends that could be processed enzymatically. A photoremovable group that could be removed in quantitative yield would ensure high purity of the synthesized probes. We have developed new pyrimidine building blocks for 5′ → 3′ DNA synthesis with high cycle yields using the NPPOC (3′-nitrophenylpropyloxycarbonyl) protecting group. These phosphoramidites were proved in automated photochemical DNA synthesis on a modified synthesizer.

DNA microarray technology has had a profound effect on the way nucleic acid analysis is conducted[1] and, consequently, on the way biological research is performed in the post-genomic era. Many different problems are being addressed through single-nucleotide polymorphism (SNP) genotyping,[2] RNA/gene expression analysis,[3] mutation detection/diagnostics,[4] genomic DNA sequencing and resequencing,[5] and population genetics.[6]

Initial excitement about the field was generated[7] using arrays prepared by light-directed chemical synthesis,[8] where DNA is formed one base at a time in situ. Many more methods are now available to fabricate microarrays, involving either in situ DNA synthesis by spatially directed delivery methods such as ink-jet printing or spotting of prepared DNA sequences (both short, synthetic oligonucleotides (<50 bases) and longer, biologically derived DNAs, such as PCR products).[9] The size of the elements in physically directed DNA arrays is inherently limited by the volume of reagent that can be accurately delivered. Thus far, volumes of ∼100 pL have been dispensed, creating spot sizes of ∼100 μm diameter after spreading on the surface. For comparison, arrays prepared photolithographically currently have 20 μm diameter probe sites, giving them 25-fold more probes per unit area. Improvements in the lithographic technology may also be forthcoming, as semiconductor manufacturing currently uses 0.3 μm photolithography for memory chip production. Photolithographic DNA synthesis is becoming more convenient with replacement of masks by digital programmable spatial light director devices.[10] It also has the

(1) Gerhold, D.; Rushmore, T.; Caskey, C. T. *Trends Biochem. Sci*. **1999**, *24*, 168−73.
(2) Guo, Z.; Guilfoyle, R. A.; Thiel, A. J.; Wang, R.; Smith, L. M. *Nucleic Acids Res*. **1994**, *22*, 5456−5465.
(3) Lockhart, D. J.; Winzeler, E. A. *Nature* **2000**, *405*, 827−36. Sohail, M.; Akhtar, S.; Southern, E. M. *RNA* **1999**, *5*, 646−55.
(4) Whitcombe, D.; Newton, C. R.; Little, S. *Curr. Opin. Biotechnol*. **1998**, *9*, 602−608.
(5) Hacia J. G. *Nat. Genet*. **1999**, *21* (Suppl), 42−7. Kurg, A.; Tonisson, N.; Georgiou, I.; Shumaker, J.; Tollett, J.; Metspalu, A. *Genet. Test*. **2000**, *4*, 1−7.
(6) Cronin, M. T.; Fucini, R. V.; Kim, S. M.; Masino, R. S.; Wespi, R. M.; Miyada, C. G. *Hum. Mutat*. **1996**, *7*, 244−55. Kozal, M. J.; Shah, N.; Shen, N.; Yang, R.; Fucini, R.; Merigan, T. C.; Richman, D. D.; Morris, D.; Hubbell, E.; Chee, M.; Gingeras, T. R. *Nat. Med*. **1996**, *2*, 753−759.
(7) Pease, A. C.; Solas, D.; Sullivan, E. J.; Cronin, M. T.; Holmes, C. P.; Fodor, S. P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5022−5026.
(8) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Liu, A. T.; Solas, D. *Science* **1991**, *251*, 767.
(9) Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. *Science* **1995**, *270*, 467−470.

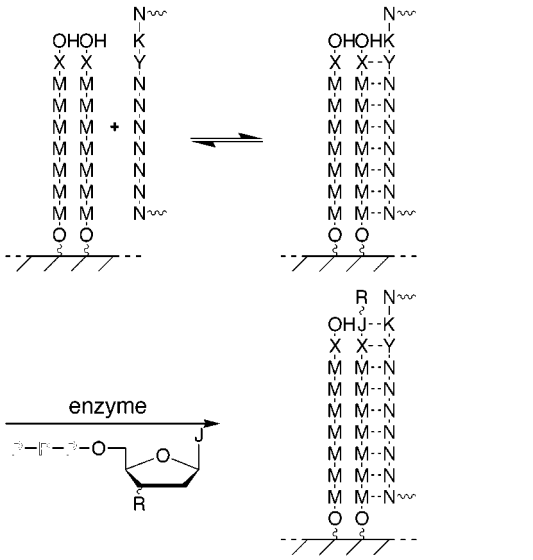
10.1021/ol0069150 CCC: $20.00
Published on Web 03/29/2001

advantage of being a parallel, combinatorial process, where the number of sequences prepared far exceeds the number of steps in the synthesis. Spatial delivery, on the other hand, is an inherently serial process. The most complex DNA microarrays known ($>10^5$ probe sites) are commercially prepared using photolithography.

DNA microarrays are most commonly used for hybridization to analyte nucleic acid (DNA or RNA). With a plethora of oligonucleotide probes on an array and a high-complexity target significantly longer than the probes, a number of different hybrids (perfect match and single-base mismatch) might be formed at each probe site. The difference in their stability may be quite small, depending on the location of the mismatch, with mismatches at the end of the probe being hardest to discriminate against. Analysis of the hybridization of target to the chip is primarily based on perfect match hybrids, so signal from mismatches could confound the analysis. Interestingly, chips prepared in situ by photolithography with a high density of probe sequences within each probe site have an unusual, nonclassical binding behavior (as compared to hybridization events in solution) that makes the discrimination against mismatches greater than might be expected.[11] This is indeed fortunate, as low cycle yields (~90%) in commercial DNA chip production lead to impure probe sequences that could further confound the analysis.[12] Thus, better methods for fabrication and use of DNA arrays are needed to give high discrimination against false signals.

Approaches to increase the fidelity of microarray analysis are based on enzymatic processing (Scheme 1). They exploit

Scheme 1. Enzymatic Processing on DNA Microarrays



the fact that enzymes that replicate and join DNA have evolved to maintain a very low error rate. While hybridization can occur to give an imperfect match at the reaction site (X & Y ≠ complements), polymerase and ligase enzymes operate on such mismatches with greatly reduced efficiency and rate. When there is a perfect match between probe and target site (X & Y ≡ complements), reaction with a triphosphate occurs (J & K ≡ complements, in ligation R = DNA chain (deoxyoligonucleotide 5′-triphosphate), in primer extension R = H or OH (nucleotide 5′-triphosphate)). The covalent attachment of J to the probe permits stringent washing steps that minimize background in imaging of the probe site. While equilibrium binding may permit the formation of end-mismatched duplexes, they are not attached and do not contribute to background.

Primer extension methods for DNA arrays have been reported by other groups[13] as well as our own APEX effort.[14] Ligation methods are also known.[15] Both require a free 3′-hydroxyl. Conventional DNA synthesis, as used in current photolithographic array preparation, attaches probes to the array at their 3′-end, preventing ligation or primer extension from being performed. Ligation/primer extension methods have therefore been limited to arrays made by spotting. Methods to synthesize DNA starting from the 5′-end are known.[16] We earlier reported efforts to develop new methods for reverse photochemical DNA synthesis.[17]

Approaches to increase the fidelity of photolithographic DNA microarray preparation have focused on improved photochemically removable protecting groups. Pfleiderer developed the NPPOC (nitrophenylpropyloxycarbonyl) group that is deprotected by a β-elimination reaction[18] (Scheme

(10) Singh-Gasson, S.; Green, R. D.; Yue, Y.; Nelson, C.; Blattner, F.; Sussman, M. R.; Cerrina, F. *Nat. Biotechnol*. **1999**, *17*, 974−978. LeProust, E.; Pellois, J. P.; Yu, P.; Zhang, H.; Gao, X.; Srivannavit, O.; Gulari, E.; Zhou, X. *J. Comb. Chem*. **2000**, *2*, 349.

(11) Forman, J. E.; Walton, I. D.; Stern, D.; Rava, R. P.; Trulson, M. O. *ACS Symp. Ser*. **1998**, *682*, 206−228.

(12) Pirrung, M. C.; Bradley, J.-C. *J. Org. Chem*. **1995**, *60*, 6270. Pirrung, M. C.; Fallon, L.; McGall, G. *J. Org. Chem*. **1998**, *63*, 241.

(13) Broude, N. E., Sano, T.; Smith, C. L.; Cantor, C. R. *Proc. Natl. Acad. Sci. U.S.A*. **1994**, *91*, 3072−3076. Livshits, M. A.; Forentive, V. L.; Mirzabekov, A. D. *J. Biomol. Struct. Dynam*. **1994**, *11*, 783−795. Dubiley, S.; Kirillov, E.; Mirzabekov, A. *Nucleic Acids Res*. **1999**, *27*, e19. Tonisson, N.; Kurg, A.; Kaasik, K.; Lohmussaar, E.; Metspalu, A. *Clin. Chem. Lab. Med*. **2000**, *38*, 165−70.

(14) Shumaker, J. M.; Metspalu, A.; Caskey, C. T. *Laboratory Protocol for Mutation Detection* Landegren, U., Ed.; Oxford University Press: Oxford, U.K., 1996; pp 93−95. Tollett, J. J.; Kurg, A.; Shah, A.; Roa, B. B.; Richards, C. S.; Nye, S. H.; Pirrung, M.; Metspalu, A.; Shumaker, J. M. *Am. J. Hum. Gen*. **1997**, *61* (Suppl. S), 1322. Kurg, A.; Tollett, J. J.; Shah, A.; Roa, B. B.; Richards, C. S.; Nye, S. H.; Pirrung, M.; Metspalu, A.; Shumaker, J. M. *Am. J. Hum. Gen*. **1997**, *61* (Suppl. S), 68. Pirrung, M. C.; Connors, R. V.; Montague-Smith, M. P.; Odenbaugh, A. L.; Walcott, N. G.; Tollett, J. J. *J. Am. Chem. Soc*. **2000**, *122*, 1873.

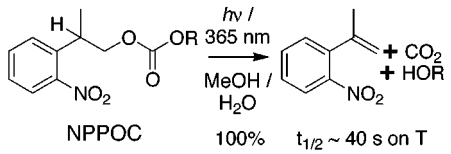
(15) Drmanac, R.; Drmanac, S.; Strezoska, Z.; Pauneska, T.; Labat, I.; Zeremski, M.; Snoddy, J.; Funkhouser, W. K.; Koop, B.; Hood, L.; Crkvenjakov, R. *Science* **1993**, *260*, 1649−1652. Gunderson, K. L.; Huang, X. C.; Morris, M. S.; Lipshutz, R. J.; Lockhart, D. J.; Chee, M. S. *Genome Res*. **1998**, *8*, 1142−1153. Drmanac, S.; Kita, D.; Labat, I.; Hauser, B.; Schmidt, C.; Burczak, J. D.; Drmanac, R. *Nat. Biotechnol*. **1998**, *16*, 54−8.

(16) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc*. **1990**, *112*, 2435−2437. Roble, J.; Pedroso, E.; Grandas, A. *Nucleic Acids Res*. **1995**, *23*, 4151. Shchepinov, M.; Case-Green, S.; Southern, E. *Nucleic Acids Res*. **1997**, *25*, 1155. Hudson, R.; Robidoux, S.; Damha, M. *Tetrahedron Lett*. **1998**, *39*, 1299.

(17) Pirrung, M. C.; Fallon, L.; Lever, D. C.; Shuey, S. W. *J. Org. Chem*. **1996**, *61*, 2129.

(18) Giegrich, H.; Eisele-Buhler, S.; Hermann, C.; Kvasyuk, E.; Charubala, R.; Pfleiderer, W. *Nucleosides Nucleotides* **1998**, *17*, 1987. Hasan, A.; Stengele, K.-P.; Gegrich, H.; Cornwell, P.; Isham, K.; Sachleben, R.; Pfleiderer, W. Foote, R. *Tetrahedron* **1997**, *53*, 4247.

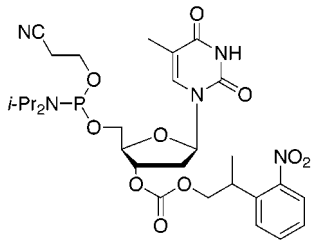
Scheme 2. Photochemical Removal of NPPOC Groups As Reported by Pfleiderer[18] Occurs by β-Elimination



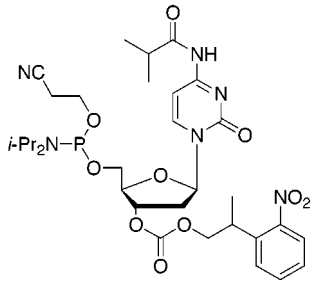
2). Beier and Hoheisel reported nucleoside phosphoramidites with NPPOC groups at the 5′ position that can be deprotected photochemically in the presence of limited amounts of amine base (piperidine or DBU).[19] Arrays with free 5′-ends result from DNA synthesis with these NPPOC amidites. They undergo coupling and photochemical deprotection in essentially quantitative cycle yields, far superior to amidites with other photoremovable groups such as MeNPOC. Given this high-fidelity method of DNA array synthesis and the interest in high-fidelity array analysis methods requiring free 3′-ends, we have adapted the NPPOC method to reverse DNA synthesis.

Two pyrimidine 5′-phosphoramidites **1** and **2** (Chart 1) were prepared from the 5′-*tert*-butyldiphenylsilylated nucleosides by treatment with NPPOC-Cl (generated from phosgene and the alcohol, 0 °C → rt, 12 h, pyridine, 88−92%), deprotection with triethylamine trihydrofluoride[20] (10 equiv, THF, rt, 12 h, 91%), and treatment with β-cyanoethyl *N*,*N*-bis(diisopropylamino)chlorophosphine (*i*-Pr_2EtN, 3 h, CH_2Cl_2, 0 °C → RT, 70−77%). The resulting amidites were purified by precipitation from methylene chloride/hexane.

Chart 1. Pyrimidine Nucleoside NPPOC Phosphoramidites



1



2

Testing of these phosphoramidites as a prelude to array preparation requires a method to perform photochemical deprotection during automated synthesis. Commercial controlled-pore glass-5′-**T**-3′-DMTr support (Glen Research) was loaded into a synthesis cartridge fabricated from polymer filters and a clear fluoropolymer PFA tube and then deprotected by manual delivery of 3% TCA/acetonitrile. An ABI 392 automated DNA synthesizer was modified such that an external events control relay usually used to trigger a fraction collector (to assess dimethoxytrityl cation release) instead opens a shutter in an optical train (Figure 1). This optical train delivers ~4 mW/cm^2 of 365 nm light from an arc lamp through a liquid light guide to the cartridge, of which ~25% is transmitted to the support. The synthesis program is modified so the usual acid deprotection step is replaced with the photochemical removal of an NPPOC group in the presence of base. The base solution (50 mM piperidine in CH_3CN) replaces the acid solution on the synthesizer. The synthesizer triggers the external event relay during the deprotection step instead of afterward.

Earlier work on 5′ → 3′ synthesis with DMTr protection showed that increased coupling time is required compared to 5′-DMTr-amidites, presumably as a result of the lower nucleophilicity of the 3′-hydroxyl group. We verified this observation with commercial 3′-DMTr-5′-amidites, showed that 120 s coupling provides >99% cycle yield, and used this coupling time in our photochemical reverse synthesis.

In initial experiments to optimize photochemical 5′ → 3′ synthesis, $\mathbf{T}_4$ was prepared on a 0.2 μmol scale (enabling product analysis by HPLC and MS) starting with cpg-5′-**T**-3′-OH. Coupling with **1** and photochemical deprotection were performed twice, and the terminal **T** was added by coupling with **1** (i.e., 2.5 cycles). The $\mathbf{T}_4$ was cleaved and fully deprotected with ammonia and directly analyzed by reverse-phase HPLC (Supelcosil LC-18S (5 μm, 1.6 mm × 250 mm), 20% → 40% linear gradient of 25 mM Et_3NHOAc, pH 6.5 in 2:3 CH_3CN/H_2O), which gives single-nucleotide resolution up to decanucleotides. The deprotection step was examined at increasing irradiation times, with >98% cycle yields observed in 20 min at ~1 mW/cm^2, 365 nm. The piperidine solution was refreshed 10× to remove the styrene byproduct and agitate the support, a measure that would not be necessary during DNA array fabrication. Substitution of DBU for piperidine, as reported by Beier and Hoheisel, led to reduced product yield and purity. This problem was traced to dark deprotection. A $\mathbf{T}_3$ synthesis omitting irradiation in the deprotection step using DBU as base still gives $\mathbf{T}_3$ (5% yield), while with piperidine as base only $\mathbf{T}_2$ is observed.

Several oligodeoxynucleotides were prepared by 5′ → 3′ synthesis using phosphoramidites **1** and **2** and analyzed by HPLC, with both the detection of truncation sequences and absolute peak integration against authentic calibration standards providing cycle and overall yields (Table 1). MALDI-TOF MS confirmed the fidelity of these products.

In conclusion, these novel phosphoramidite building blocks provide oligonucleotides in high cycle yields via reverse (5′ → 3′) light-based solid-phase synthesis. Under the photochemical deprotection conditions, isobutyroyl deoxycytidine

(19) Beier, M.; Hoheisel, J. D. *Nucleic Acids Res.* **2000**, *28*, e11.

(20) Pirrung, M. C.; Shuey, S. W.; Lever, D. C.; Fallon, L. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1345.

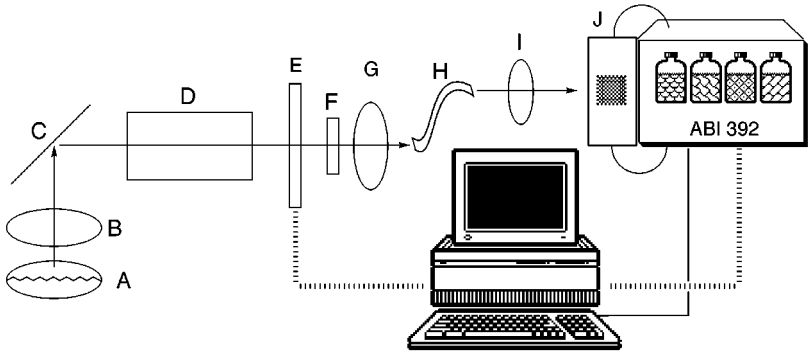


Figure 1. Automated photochemical DNA synthesizer. (A) Oriel Hg/Xe 1000 W Arc lamp (source). (B) Condenser lens (2 in.). (C) Surface-coated dichroic mirror (280−350 nm). (D) Water filter. (E) Electronically controlled shutter. (F) Narrow bandwidth (310 nm) interference filter. (G) Fiber optic focusing lens. (H) UV/vis liquid light guide (1 m × 5 mm). (I) Fiber optic output collimating lens. (J) DNA synthesis cartridge (target).

is unreactive. The purine building blocks remain to be added, but this methodology promises to enable the most powerful DNA microarray fabrication technology, photolithography, to be combined with enzymatic processing of probes to afford the highest fidelity in nucleic acid analysis.

Table 1. Oligodeoxynucleotides Synthesized Using **1** and **2**

sequence	av cycle yield[a]	total yield
5′ **TTT**	98.8	97.6
5′ **TTTT**	98.7	96.2
5′ **TTTTTTTTTT**	98.2	83.0
5′ **TCC**	98.7	97.5

[a] Photodeprotection was not performed after coupling of the 3′-base. The $\mathbf{T}_3$ synthesis requires two couplings but only one deprotection, etc.

Acknowledgment. This work was supported by NIH GM-46720 and an NIH postdoctoral fellowship (M.M.-S.).

Supporting Information Available: Experimental descriptions. This material is available free of charge via the Internet at http://pubs.acs.org.

OL0069150