Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 16 (2006) 3902–3905
Bioorganic & Medicinal Chemistry Letters

A diversity oriented synthesis of 3′-*O*-modified nucleoside triphosphates for DNA 'sequencing by synthesis'

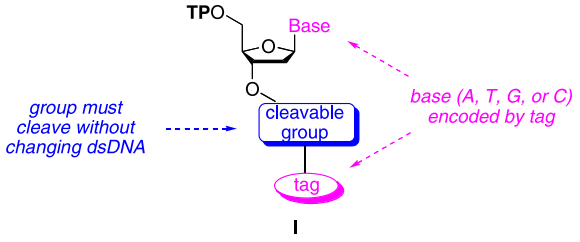
Genliang Lu and Kevin Burgess*

Department of Chemistry, Texas A&M University, PO Box 300012, College Station, TX 77842-3012, USA

Received 29 March 2006; revised 10 May 2006; accepted 12 May 2006
Available online 6 June 2006

Abstract—The nucleoside triphosphates **1**, containing a photochemically cleavable group, and **2**, having one that may be cleaved via palladium catalysis, were prepared as a prelude to investigating sequencing of DNA via sequencing by synthesis.

With notable exceptions,[1] throughput in practical DNA sequencing technologies tends to be limited by the requirement of electrophoresis.[2] In response, several groups,[3–11] including ours,[12,13] have been motivated to investigate 'sequencing by synthesis' (SBS) methods that circumvent electrophoresis. A typical embodiment of this idea would feature derivatives of the natural nucleoside triphosphates (dATP, dTTP, dCTP, and dGTP) modified such that the 3′-hydroxyl functionalities carry a cleavable group attached to a fluorescent label, that is compound types **I**.



I

TP = triphosphate, $P_3O_9H_4$

Four nucleobase analogs **I** would be required, each with resolvable fluorescent labels that reveal the nucleobase with which they are associated. Sequencing would then be achieved via the following, iterative, process: (i) polymerase-mediated incorporation of the appropriate, complementary base analog; (ii) detection of the fluorescent label (thereby revealing the complementary base on the template); (iii) rupture of the cleavable group giving a natural 3′-hydroxy terminus; and, (iv) repetition of the process (Fig. 1).

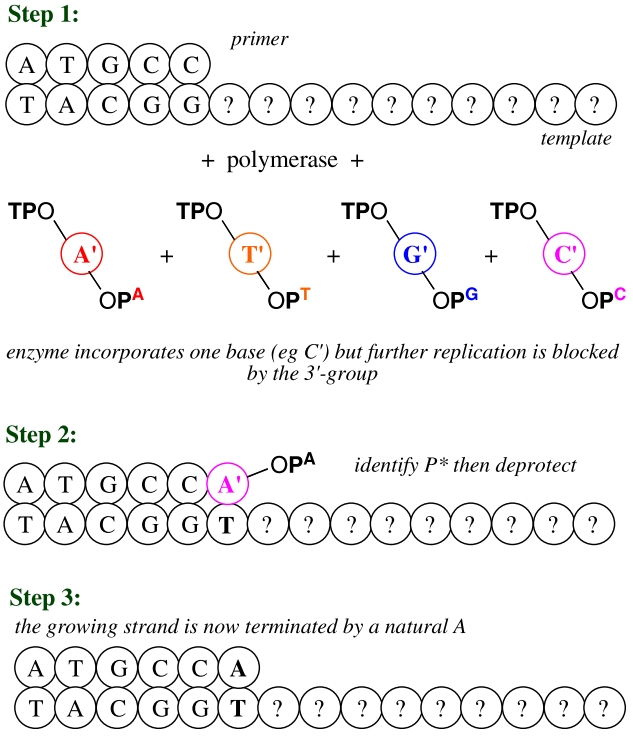


Figure 1. The concept of sequencing by synthesis.

Keywords: DNA; Sequencing; Nucleoside; Triphosphate; Photolabile.
* Corresponding author. Tel.: +1 979 845 4345; fax: +1 979 845 1881; e-mail: burgess@tamu.edu

0960-894X/$ - see front matter
doi:10.1016/j.bmcl.2006.05.035

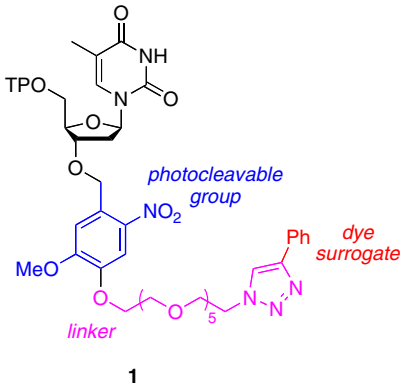
The main obstacle to reducing Figure 1 to practice is that polymerase enzymes are extremely intolerant of modifications at the 3′-position: they tend not to incorporate these bases efficiently. Until the problem of incorporation is solved, other critical questions like the fidelity of the incorporation cannot be addressed. Consequently, it is clear that successful implementation of sequencing by synthesis as shown in Figure 1 will require libraries of compounds **I** to be screened against libraries of polymerase mutants that are engineered to tolerate changes at the nucleotide 3′ site.

The problem of producing even small numbers of nucleoside triphosphates **I** is non-trivial for several reasons. Formation of nucleoside triphosphates tends to be moderate-to-low-yielding reactions requiring HPLC purification.[14] Purification is necessary because tests for incorporation by polymerases can give deceptive results (both false positives and negatives) if small amounts of impurities are present in the triphosphate sample.

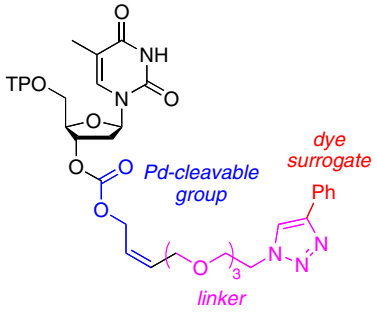
The considerations outlined above led us to consider methods for syntheses of the compounds **1** and **2** (Chart 1). These model compounds have the following attributes. First, both compounds have groups that could be cleaved efficiently under conditions that should not disrupt ds (double strand) DNA. Second, they both have oligoethylene glycol linkers that remove the steric bulk of the dye (or dye surrogate in this case) from the nucleoside component. Third, instead of including a fluorescent dye, molecules **1** and **2** each have a phenyl group. All the fluorescent dyes that are usable in this methodology are larger than a phenyl group, so if these surrogates are not incorporated by the polymerases then ones containing fluorescent dyes are unlikely to be incorporated too.

Compounds **1** and **2** have another important attribute, but this relates to the overall strategy rather than to the compound structures. The triazole functionality enables single advanced precursors to be prepared, then derivatized[15] with a range of fluorescent dyes that have pendant terminal alkynes. In other words, the synthesis is diversity oriented.

The synthesis of compound **1** is described in Scheme 1. The valeraldehyde-derived alcohol **3** was alkylated with

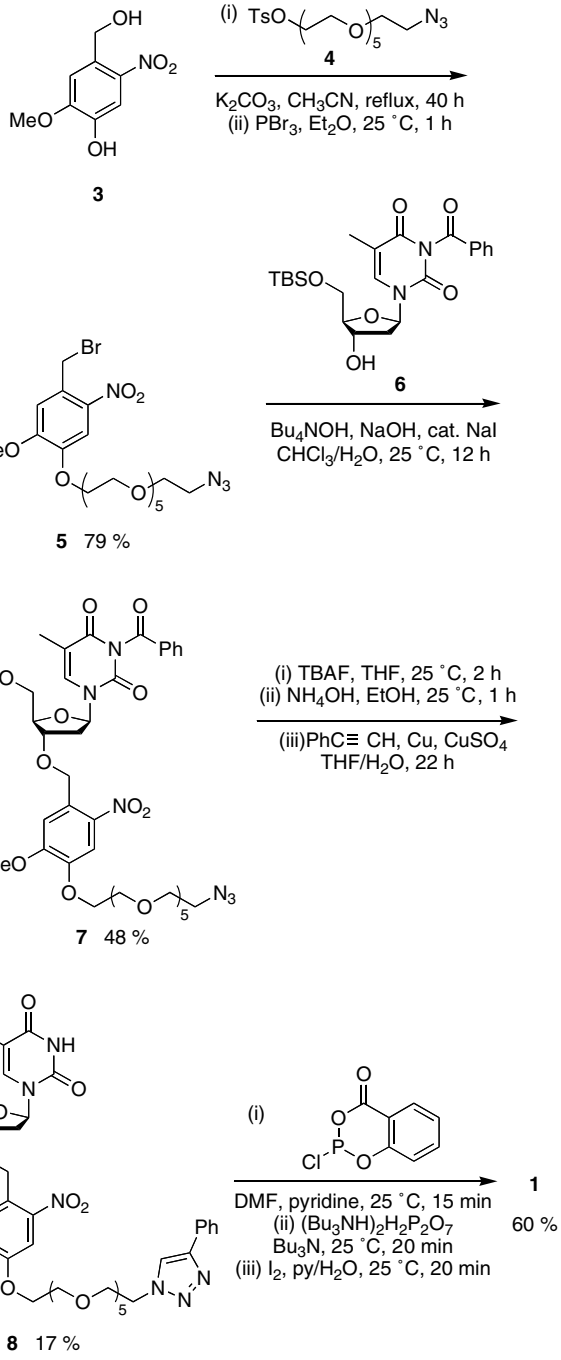


1



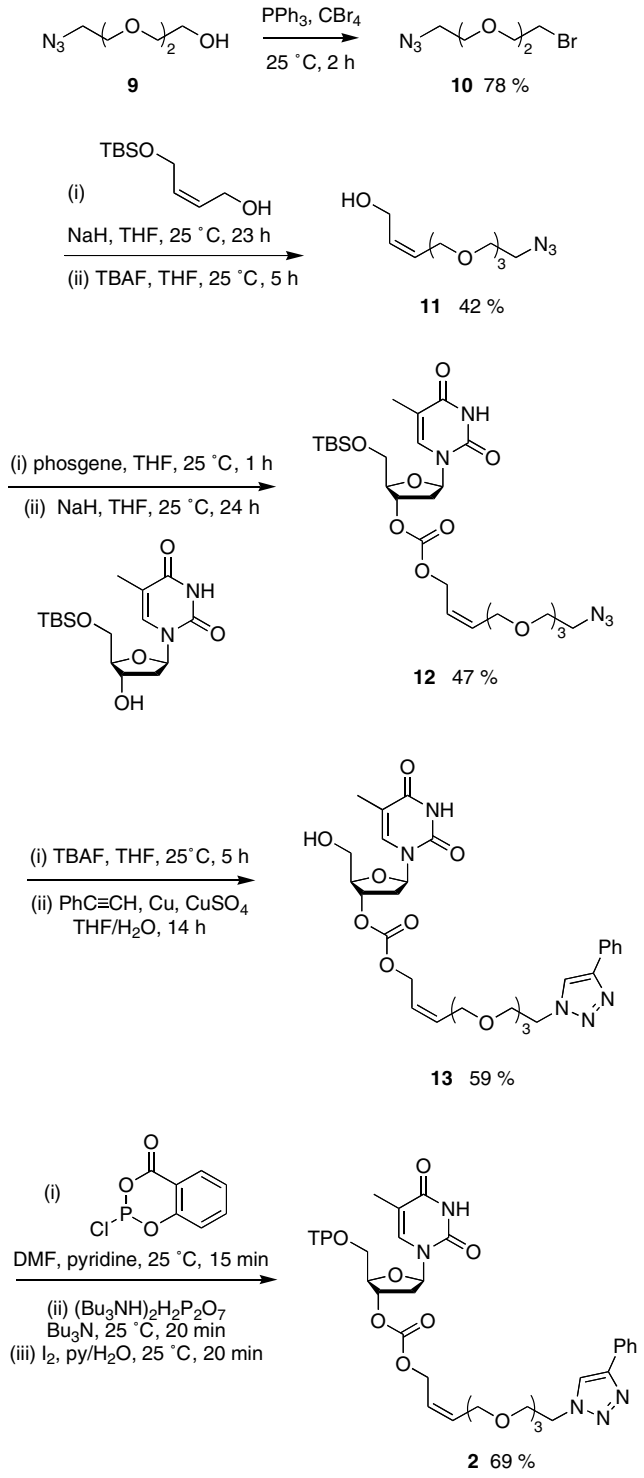
2

Chart 1. Target nucleoside triphosphates **1** and **2**.



Scheme 1. A synthesis of nucleoside triphosphate **1**.

an azido ethylene glycol linker **4**[16] giving an oligoether that was then converted into the benzylic bromide **5**. *N*-Protection of the thymidine base is essential for derivatization of the 3′-oxygen, otherwise the base is alkylated preferentially.[17,18] Consequently, the known *N*-benzyl thymidine derivative **6** was coupled with the benzylic bromide **5** under phase transfer conditions to give the 3′-ether **7**. Desilylation of the 5′-*O*-TBS and hydrolysis of the 3-*N*-benzoyl group afforded a nucleoside ready for copper-mediated cycloaddition. This proceeded efficiently to give a nucleoside that was then triphosphorylated using the Ludwig–Eckstein method, that is phosphorylation with 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one followed by pyrophosphate addition and oxidation.[19]



Scheme 2. A synthesis of nucleoside triphosphate **2**.

A synthesis of nucleotide triphosphate **2** was completed as shown in Scheme 2. A triethylene glycol chlorocarbonate was constructed as shown, then coupled to a 5′-protected thymidine derivative. In this case protection of the thymidine part is not necessary because the chlorocarbonate adds to the 3′-hydroxyl selectively. Desilylation and copper-mediated cycloaddition chemistry with phenylethyne (our dye surrogate) gave the nucleoside **13** that was then triphosphorylated as described above.

It is important that 3′-protection of the nucleoside triphosphates **1** and **2** can be cleaved under relatively mild conditions that would not perturb dsDNA. Photolysis of compound **8** was performed using a relatively simple strip light designed to emit at 360 nm (Southern New England Ultraviolet Company). Acetonitrile was used as solvent and ethanolamine as an additive. HPLC anal-

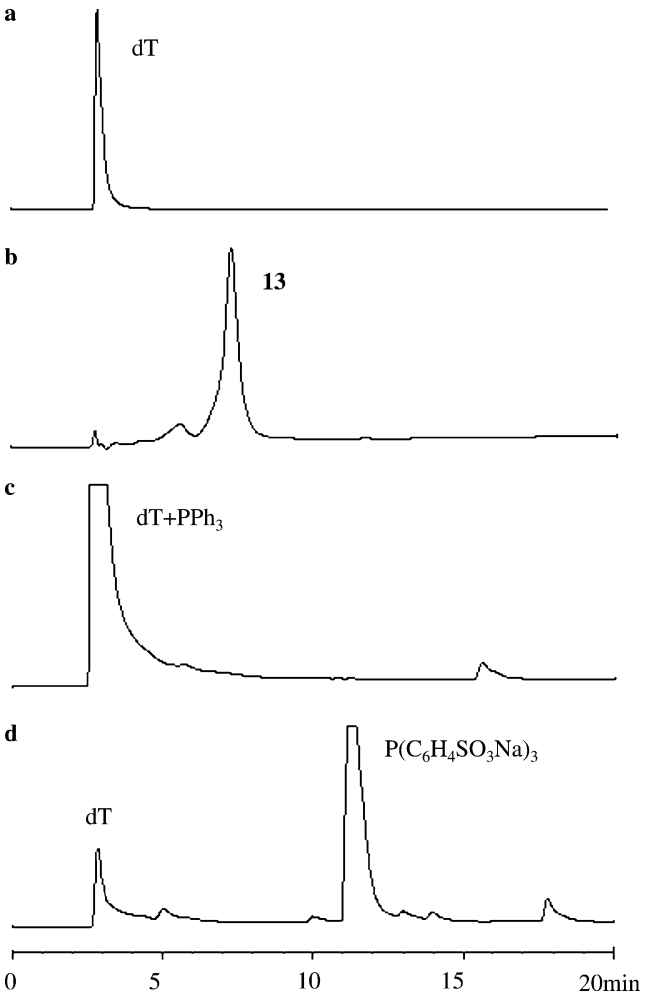


Figure 2. HPLC data from the deprotection of nucleoside **13**: (a) deoxythymidine; (b) nucleoside (13); (c) deprotection of **13** under $Pd(OAc)_2$ and PPh_3; (d) deprotection of **13** under $Pd[P(C_6H_4SO_3Na)_3]_4$. The unlabeled peaks were unidentified.

ysis showed that the 3′-*O*-photolabile group of deoxythymidine **8** was completely deprotected in 15 min giving dT. Ethanolamine was necessary, however, since without it the deprotection was only partial under otherwise identical conditions.[20] Deprotection of **13** was examined under two sets of conditions involving palladium catalysis (1, $HCOOH/BuNH_2$ (1:1), $Pd(OAc)_2$, PPh_3, H_2O/CH_3CN (1:1); and, 2, $Pd[P(C_6H_4SO_3Na)_3]_4$, H_2O). Both these methods resulted in complete removal of the ALLOC-based protecting group of **13** in fewer than 10 min and neutral pH giving dT also. Figure 2 shows illustrative HPLC data from the palladium catalysis deprotection.

Much work needs to be done to reach a viable sequencing scheme. Production and testing of polymerase mutants is at least as demanding as the synthetic challenges surrounding the triphosphates to be tested. However, the work summarized here represents a significant milestone. It illustrates that nucleosides like the photolabile system **8** and the Pd-cleavable system **13** can potentially be combined with a series of fluorescent dyes containing alkyne handles, then converted to triphosphates. This approach will facilitate the arduous task of making a library of photocleavable, fluorescent, 3′-protected nucleoside triphosphates for testing in SBS.

Acknowledgments

Support for this project was provided by The NIH: HG003573 and GM72041, and The Robert Welch Foundation.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2006.05.035.

References and notes

1. Margulies, M.; Egholm, M.; Altman, W. E.; Attiya, S.; Bader, J. S.; Bemben, L. A.; Berka, J.; Braverman, M. S.; Chen, Y.-J.; Chen, Z.; Dewell, S. B.; Du, L.; Fierr, J. M.; Gomes, X. V.; Godwin, B. C.; He, W.; Helgesen, S.; Ho, C. H.; Irzyk, G. P.; Jando, S. R. C.; Alenquer, M. L. I.; Jarvie, T. P.; Jirage, K. B.; Kim, J.-B.; Knight, J. R.; Lanza, J. R.; Leamon, J. H.; Lefkowitz, S. M.; Lei, M.; Li, J.; Lohman, K. L.; Lu, H.; Makhijani, V. B.; McDade, K. E.; McKenna, M. P.; Myers, E. W.; Nickerson, E.; Nobile, J. R.; Plant, R.; Puc, B. P.; Ronan, M. T.; Roth, G. T.; Sarkis, G. J.; Simons, J. F.; Simpson, J. W.; Srinivasan, M.; Tartaro, K. R.; Tomasz, A.; Vogt, K. A.; Volkmer, G. A.; Wang, S. H.; Wang, Y.; Weiner, M. P.; Yu, P.; Begley, R. F.; Rothberg, J. M. *Nature* **2005**, *437*, 376.
2. Hunkapiller, T.; Kaiser, R. J.; Koop, B. F.; Hood, L. *Science* **1991**, *254*, 59.
3. Sarfati, R. S.; Namane, A.; Munier, H.; Barzu, O. *Tetrahedron Lett.* **1991**, *36*, 4699.
4. Sarfati, S. R.; Pochet, S.; Guerreiro, C.; Namane, A.; Huynh-Dinh, T.; Igolen, J. *Tetrahedron* **1987**, *43*, 3491.
5. Canard, B.; Sarfati, R. S. *Gene* **1994**, *148*, 1.
6. Bai, X.; Li, Z.; Jockusch, S.; Turro, N. J.; Ju, J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 409.
7. Li, Z.; Bai, X.; Ruparel, H.; Kim, S.; Turro, N. J.; Ju, J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 414.
8. Ruparel, H.; Bi, L.; Li, Z.; Bai, X.; Kim, D. H.; Turro, N. J.; Ju, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5932.
9. Seo, T. S.; Bai, X.; Kim, D. H.; Meng, Q.; Shi, S.; Ruparel, H.; Li, Z.; Turro, N. J.; Ju, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5926.
10. Aksyonov, S. A.; Bittner, M.; Bloom, L. B.; Reha-Krantz, L. J.; Gould, I. R.; Hayes, M. A.; Kiernan, U. A.; Niederkofler, E.; Pizziconi, V.; Rivera, R. S.; Williams, D. J. B.; Williams, P. *Anal. Biochem.* **2006**, *348*, 127.
11. Bi, L.; Kim, D. H.; Ju, J. *J. Am. Chem. Soc.* **2006**, *128*, 2542.
12. Welch, M. B.; Martinez, C. I.; Zhang, A. J.; Jin, S.; Gibbs, R.; Burgess, K. *Chem. Eur. J.* **1999**, *5*, 951.
13. Metzker, M. L.; Raghavachari, R.; Richards, S.; Jacutin, S. E.; Civitello, A.; Burgess, K.; Gibbs, R. A. *Nucleic Acids Res.* **1994**, *22*, 4259.
14. Burgess, K.; Cook, D. *Chem. Rev.* **2000**, *100*, 2047.
15. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004.
16. Lu, G.; Lam, S.; Burgess, K. *Chem. Commun.* **2006**, 1652.
17. Welch, M.; Martinez, C.; Zhang, A.; Jin, S.; Gibbs, R.; Burgess, K. *Chem. Eur. J.* **2005**, *11*, 7136.
18. Welch, M. B. *Nucleosides Nucleotides Nucleic Acids* **2006**, *25*, 119.
19. Ludwig, J.; Eckstein, F. *J. Org. Chem.* **1989**, *54*, 631.
20. Burgess, K.; Jacutin, S. E.; Lim, D.; Shitangkoon, A. *J. Org. Chem.* **1997**, *62*, 5165.