ORGANIC LETTERS

2007
Vol. 9, No. 15
2851−2854

Synthesis of Rigid Homo- and Heteroditopic Nucleobase-Terminated Molecules Incorporating Adenine and/or Thymine

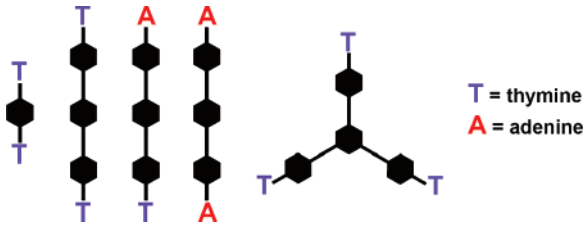
Mikkel F. Jacobsen, Casper S. Andersen, Martin M. Knudsen, and Kurt V. Gothelf*

Centre for DNA Nanotechnology, Department of Chemistry, University of Aarhus, Langelandsgade 140, 8000 Aarhus C, Denmark

kvg@chem.au.dk

Received April 30, 2007

ABSTRACT



A series of homo- and heteroditopic thymine- and/or adenine-terminated molecules incorporating rigid aryl or oligo(phenylene ethynylene) linkers has been efficiently synthesized. The key steps involved in the synthesis are the construction of the *N*-arylated nucleobases using the Chan–Lam–Evans-modified Ullman coupling and their further elaboration using the Sonogashira coupling. Furthermore, the synthesis of a rigid tripodal thymine derivative is reported.

The high fidelity observed in the base-pairing between complementary DNA strands allows for the predetermined formation of artificial nanostructures.[1] In addition, a wide range of artificial nucleobase derivatives has been employed in the creation of supramolecular self-assemblies using molecular recognition via base-pairing.[2] Even though the Watson−Crick mode of bonding is prevalent in natural systems, e.g., DNA double helices, other hydrogen-bonding motifs are available and expand the possibility for creation of different structural networks. In this respect, previous studies have revealed that simple mixtures of adenine and thymine can form well-ordered supramolecular nanopatterns on surfaces featuring a unique hydrogen-bonding setup based on reverse Hoogsteen A−T−A−T quartets adjacent to homochiral chains of adenine duplets.[3] Furthermore, several homo- and heteroditopic nucleobase derivatives (bolaamphiphiles) have been demonstrated to spontanously homo- and heteroassemble to form well-defined supramolecular fibers.[4] Therefore, the ability to create artificial polymeric systems incorporating several nucleobase moieties can serve

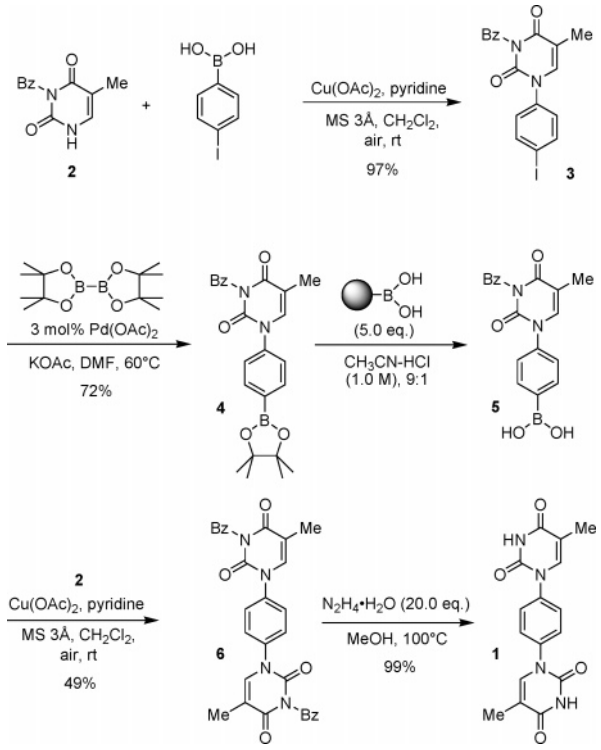
(1) (a) Rothemund, P. W. K. *Nature* **2006**, *440*, 297. (b) Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H. *Science* **2003**, *301*, 1882. (c) Mao, C. D.; LaBean, T. H.; Reif, J. H.; Seeman, N. C. *Nature* **2000**, *407*, 293. (d) Yan, H.; Feng, L. P.; LaBean, T. H.; Reif, J. H. *J. Am. Chem. Soc.* **2003**, *125*, 14246. (e) Winfree, E.; Liu, F.; Wenzler, L.; Seeman, N. C. *Nature* **1998**, *394*, 539. (f) LaBean, T.; Yan, H.; Kopatsch, J.; Liu, F.; Winfree, E.; Reif, J. H.; Seeman, N. C. *J. Am. Chem. Soc.* **2000**, *122*, 1848. (g) Chelyapov, N.; Brun, Y.; Gopalkrishnan, M.; Reishus, D.; Shaw, B.; Adleman, L. *J. Am. Chem. Soc.* **2004**, *126*, 13924. (h) Ding, B.; Sha, R.; Seeman, N. C. *J. Am. Chem. Soc.* **2004**, *126*, 10230. (i) Rothemund, P.; Papadakis, N.; Winfree, E. *Pub. Lib. Sci. Biol.* **2004**, *2*, 2041. (j) Liu, Y.; Ke, Y.; Yan, H. *J. Am. Chem. Soc.* **2005**, *127*, 17140. (k) Rothemund, P.; Ekani-Nkodo, A.; Papadakis, N.; Kumar, A.; Fygenson, D.; Winfree, E. *J. Am. Chem. Soc.* **2004**, *126*, 16344. (l) Brun, Y.; Gopalkrishnan, M.; Reishus, D.; Shaw, B.; Chelyapov, N.; Adleman, L. *FNANO* **2004**, 2.

(2) (a) Sivakova, S.; Rowan, S. J. *Chem. Soc. Rev.* **2005**, *34*, 9. (b) Sessler, J. L.; Jayawickramarajah, J. *Chem. Commun.* **2005**, 1939. (c) Sessler, J. L.; Lawrence, Candace, M.; Jayawickramarajah, J. *Chem. Soc. Rev.* **2007**, *36*, 314. (d) Davis, J. T.; Spada, G. P. *Chem. Soc. Rev.* **2007**, *36*, 296.

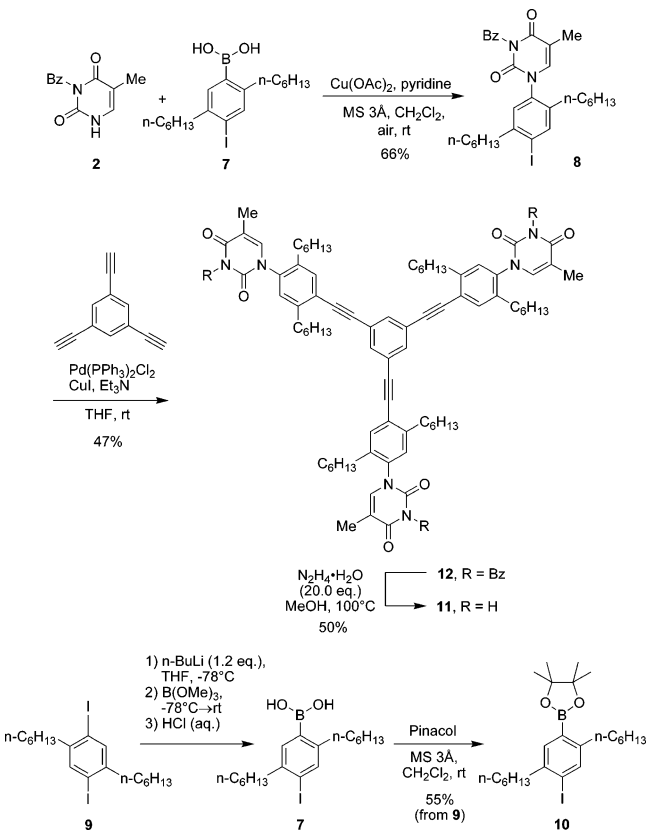
(3) Mamdouh, W.; Dong, M.; Xu, S.; Rauls, E.; Besenbacher, F. *J. Am. Chem. Soc.* **2006**, *128*, 13305.

10.1021/ol071006x CCC: $37.00
Published on Web 06/29/2007

Scheme 1. Synthesis of Short Thymine−Thymine Derivative **1**



Scheme 2. Synthesis of the Tripodal Thymine Derivative **11** and Boronic Acid **7**



as a means to study their inter- and intramolecular hydrogen-bonding properties and gain further insight into the mechanisms behind the assembly of complementary DNA strands.

Our continued interest in DNA structures, including its use for controlling chemical reactivity and creating self-assembled nanostructures,[5] controlling photosensitized processes,[6] and detection of DNA,[7] has driven us to create an efficient methodology for the synthesis of *N*-arylated and *N*-alkenylated nucleobases employing the Chan−Lam−Evans-modified Ullmann coupling.[8] Thus, the *N*-arylated and *N*-alkenylated nucleobases are obtained via the direct coupling of protected or masked nucleobase derivatives with alkenyl or aryl boronic acids. We envisaged that such artificial molecules incorporating both multiple nucleobases and an extended rigid aromatic system may be suitable for the controlled formation of supramolecular structures. Such structures may be formed on surfaces by Watson−Crick and other types of hydrogen bonding between the bases or in solution by a combination of hydrogen bonding and π-stacking. In another application such structures may interact with DNA strands and form extended helices by Watson−Crick base pairing in analogy with the structures reported by Iwaura et al.[9]

As a starting point in this realm, we now report on the synthesis of a series of novel rigid aryl or oligo(phenylene ethynylene) structures connected via bisaryl-like linkages to multiple nucleobase moieties.

At the outset, we aimed for the synthesis of a short conjugated T−T derivative **1**, in which the two thymine moieties are separated only by one benzene ring, which was accomplished by two sequential Chan-Lam-Evans reactions (Scheme 1). The previously described N^3-benzoyl thymine **2**[9] was converted to *N*-aryl derivative **3** as previously described.[8] By subjecting **3** to the modified Miyuara protocol the aryl pinacolboronate **4** was obtained in good yield.[10] The efficient hydrolysis of pinacolboronates can be troublesome. However, the recently described method employing a solid-

(4) (a) Iwaura, R.; Yoshida, K.; Mitsutoshi, M.; Yase, K.; Shimizu, T. *Chem. Mater.* **2002**, *14*, 3047. (b) Shimizu, T.; Iwaura, R.; Masuda, M.; Hanada, T.; Yase, K. *J. Am. Chem. Soc.* **2001**, *123*, 5947. (c) Iwaura, R.; Minamikawa, H.; Shimizu, T. *J. Colloid Interface Sci.* **2004**, *277*, 299. (d) Shimizu, T.; Masuda, M.; Minamikawa, H. *Chem. Rev.* **2005**, *105*, 1401. (e) Fuhrhop, J.-H.; Wang, T. *Chem. Rev.* **2004**, *104*, 2901. (f) Shimizu, T. *Macromol. Rapid Commun.* **2002**, *23*, 311. (g) Shimizu, T.; Masuda, M. *J. Am. Chem. Soc.* **1997**, *119*, 2812. (h) Itojima, Y.; Ogawa, Y.; Tsuno, K.; Handa, N.; Yanagawa, H. *Biochemistry* **1992**, *31*, 4757.

(5) (a) Gothelf, K. V.; Thomsen, A. H.; Nielsen, M.; Cló, E.; Brown, R. S. *J. Am. Chem. Soc.* **2004**, *126*, 1044. (b) Nielsen, M.; Thomsen, A. H.; Cló, E.; Kirpekar, F.; Gothelf, K. V. *J. Org. Chem.* **2004**, *69*, 2240. (c) Brown, R. S.; Nielsen, M.; Gothelf, K. V. *Chem. Commun.* **2004**, 1464. (d) Gothelf, K. V.; LaBean, T. H. *Org. Biol. Chem.* **2005**, *3*, 4023. (e) Gothelf, K. V.; Brown, R. S. *Chem. Eur. J.* **2005**, *11*, 1062. (f) Nielsen, M.; Dauksaite, V.; Kjems, J.; Gothelf, K. V. *Bioconjugate Chem.* **2005**, *16*, 681.

(6) Cló, E.; Snyder, J. W.; Voigt, N. V.; Ogilby, P. R.; Gothelf, K. V. *J. Am. Chem. Soc.* **2006**, *128*, 4200.

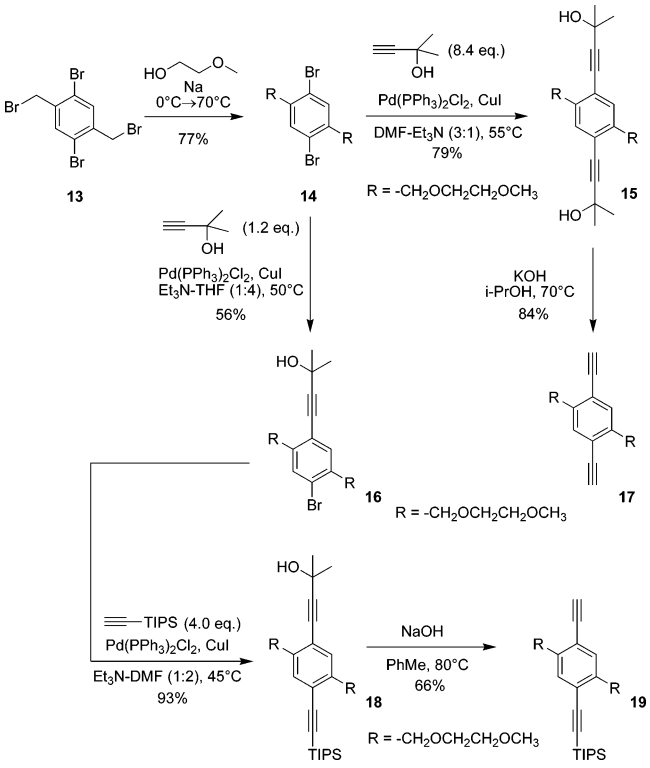
(7) Hansen, J. A.; Mukhopadhyay, R.; Hansen, J. Ø.; Gothelf, K. V. *J. Am. Chem. Soc.* **2006**, *128*, 3860.

(8) Jacobsen, M. F.; Knudsen, M. M.; Gothelf, K. V. *J. Org. Chem.* **2006**, *71*, 9183.

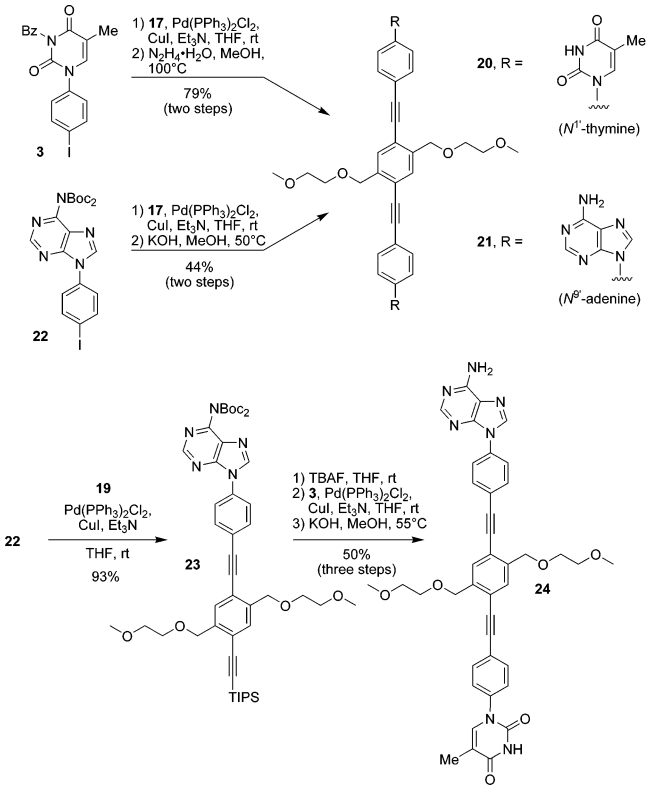
(9) Frieden, M.; Giraud, M.; Reese, C. B.; Song, Q. *J. Chem. Soc., Perkin Trans. 1* **1998**, 2827.

(10) Zhu, L.; Duquette, J.; Zhang, M. *J. Org. Chem.* **1999**, *68*, 3729.

Scheme 3. Synthesis of Oligo(phenylene ethynylene) Linkers **17** and **19**



Scheme 4. Completion of T−T (**20**), A−A (**21**), and A−T (**24**) Derivatives



supported boronic acid in a transesterification reaction gratifyingly afforded the required boronic acid **5**,[11] which was used without further purification in excess in the Chan−Lam−Evans reaction with **2** to yield the protected T−T derivative **6**. Facile removal of the Bz-protection groups was accomplished with methanolic hydrazine to yield the desired T−T derivative **1**.

The notorious low solubility of the parent nucleobases in organic media necessitates the careful design of molecules incorporating multiple nucleobase moieties. Although **1** has some solubility in organic solvents (alcohols, DMF, DMSO), we worried that a longer oligo(phenylene ethynylene) linker may lower the solubility even further. Therefore, we chose a design in which either PEG or alkyl side chains are incorporated. The introduction of alkyl side chains has been extensively used in the synthesis of large oligo(phenylene ethynylene) networks in order to improve the solubility in organic media.[12]

Thus, **2** was reacted with excess of crude boronic acid **7** having two C6-alkyl chains to afford **8** in good yield (Scheme 2). The boronic acid was easily synthesized from the known symmetrical diiodide **9** by borate trapping of the mono-organolithium species.[13] The boronic acid **7** could be further converted to the pinacol boronic ester **10** for characterization purposes. With **8** in hand, we aimed for the preperation of a tripodal thymine derivative **11**. To our delight, the triple Sonogashira reaction of **8** with 1,3,5-triethynylbenzene provided the star-shaped trimeric derivative **12** in reasonable yield.[14] The deprotection of **12** was accomplished with methanolic hydrazine to yield **11**, which has a considerably improved solubility profile compared to **1** in organic media such as CH_2Cl_2.

A series of nucleobase derivatives incorporating oligo-(phenylene ethynylene) linkers were now synthesized. The required linkers for the creation of T−T, A−A, and A−T derivatives were prepared as outlined in Scheme 3. The tetrabromo compound **13**[15] was subjected to reaction with the in situ prepared sodium alkoxide from 2-methoxyethanol to afford **14**. By the right choice of conditions, **14** could be converted to either the symmetrical or unsymmetrical oligo-(phenylene ethynylene) structures **15** or **16**, respectively, via Sonogashira couplings with 2-methylbut-3-yn-2-ol. The symmetrical oligo(phenylene ethynylene) **15** was deprotected at both ends by treatment with KOH to yield **17**. Bromide **16** was further subjected to a Sonogashira reaction to afford **18** with two orthogonal protection groups, and the 1,1-dimethylpropargylic alcohol protection was removed using NaOH providing **19**.[16]

(11) Pennington, T. E.; Kardiman, Hutton, C. C. A. *Tetrahedron Lett.* **2005**, *45*, 6657.

(12) Tour, J. M. *Molecular Electronics: Commercial Insights, Chemistry, Devices, Architecture and Programming*; World Scientific Publishing: Singapore, 2003.

(13) Hall, D. G. *Boronic Acids*; Wiley-VCH: Weinheim, 2005.

(14) (a) Negishi, E.-I. *Handbook of Organopalladium Chemistry for Organic Synthesis*; Wiley-Interscience: New York, 2002. (b) Thorand, S.; Krause, N. *J. Org. Chem.* **1998**, *63*, 8551.

(15) Bonifacio, M. C.; Robertson, C. R.; Jung, J.-Y.; King, B. T. *J. Org. Chem.* **2005**, *70*, 8522.

The symmetrical oligo(phenylene ethynylene) **17** was subsequently used in the formation of the symmetrical T−T and A−A nucleobase derivatives **20** and **21**, respectively, by subjecting it to a double Sonogashira reaction with either **3** or adenine derivative **22**[8] (Scheme 4). The T−T derivative **20** was obtained in good overall yield following a smooth deprotection step in methanolic hydrazine. On the other hand, the last deprotection step in the formation of A−A derivative **21** required prolonged heating with methanolic KOH to remove the four Boc-groups.[17] Conventional deprotection by treatment with TFA was not successful.[18]

In the same vein, the unsymmetrical oligo(phenylene ethynylene) **19** was subjected to a sequence of Sonogashira reactions and deprotection steps (Scheme 4). The first Sonogashira reaction of **19** with *N*-arylated adenine **22** to yield **23** was followed by removal of the TIPS-group with TBAF, a second Sonogashira reaction with thymine derivative **3**, and deprotection with KOH providing the heteroditopic A-T derivative **24**.

(16) Swindell, C. S.; Fan, W.; Klimko, P. G. *Tetrahedron Lett.* **1994**, *35*, 4959.

(17) Dey, S.; Garner, P. *J. Org. Chem.* **2000**, *65*, 7697.

(18) In general, treatment with strong acids may lead to degradation of any alkyne moieties present.

In summary, we have prepared a series of novel rigid thymine and adenine homo- and heteroditopic nucleobase derivatives and a tripodal thymine derivative using a combination of the Chan−Lam−Evans-modified Ullmann coupling and the Sonogashira reaction. Further efforts in our laboratories are now directed toward broadening the scope of the synthetic methodology and studies of the abilities of such nucleobase derivatives to form homo- and heteroassemblies in solution and on surfaces and to interact with DNA.

Acknowledgment. Financial support for this work by Danish Technical Research Council, the Danish National Research Foundation Centers: Centre for DNA Nanotechnology and Center for Catalysis, and the Carlsberg Foundation is gratefully acknowledged.

Supporting Information Available: Experimental procedures and spectroscopic and analytical data for all new compounds. UV spectra of products **1**, **11**, **20**, **21**, and **24**. This material is available free of charge via the Internet at http://pubs.acs.org.

OL071006X