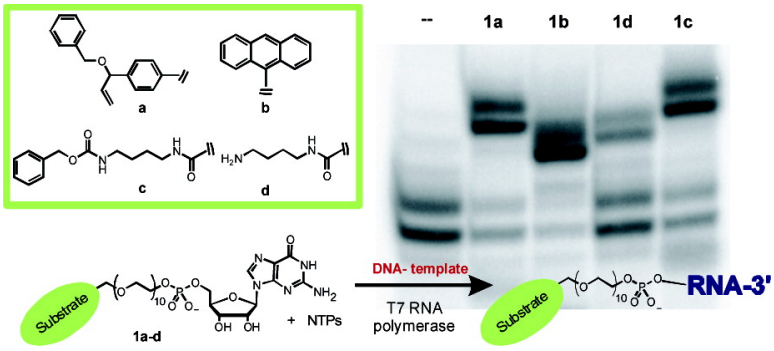
Article

Efficient Preparation of Organic Substrate–RNA Conjugates via in Vitro Transcription

Roberto Fiammengo, Kamil Muslek, and Andres Jschke

J. Am. Chem. Soc., **2005**, 127 (25), 9271-9276• DOI: 10.1021/ja051179m • Publication Date (Web): 03 June 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 4 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 06/03/2005

Efficient Preparation of Organic Substrate−RNA Conjugates via in Vitro Transcription

Roberto Fiammengo, Kamil Musílek, and Andres Jäschke*

Contribution from the Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany

Received February 24, 2005; E-mail: jaeschke@uni-hd.de

Abstract: A concise synthetic way has been developed for the preparation of guanosine monophosphate derivatives carrying a decaethylene glycol spacer at their 5′-oxygen to which are attached a range of organic substrates. The four different compounds, prepared via a convergent synthetic strategy, carry a tethered benzylallyl ether residue (**1a**), an anthracene (**1b**), a benzyl carbamate residue (**1c**), or a primary amino group (**1d**), respectively. All four compounds have been successfully incorporated at the 5′-end of a 25-mer long RNA transcript via T7 RNA polymerase, and no inhibition of chain elongation could be observed. Under proper conditions, **1a** and **1b** can be incorporated up to 90−95% and **1c** up to 68%. The amino-terminated initiator **1d** is incorporated less efficiently although still up to 49%. These results show that the more hydrophobic the guanosine monophosphate derivative is, the higher is its enzymatic incorporation.

Introduction

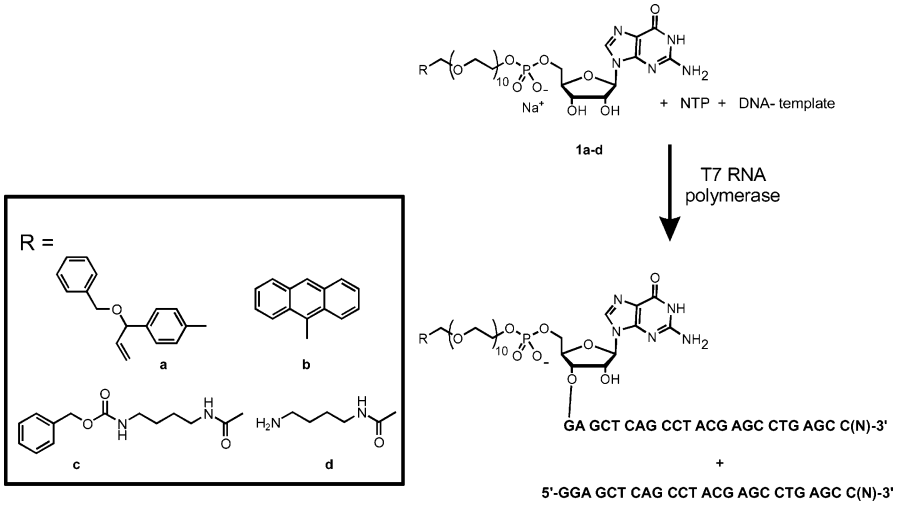
In vitro selection and in vitro evolution techniques[1,2] have been used to isolate a large number of RNA catalysts[3−5] and RNA aptamers from combinatorial libraries (pools).[6−8] If new RNA catalysts are to be discovered via in vitro selection, a substrate must be covalently attached to every sequence of the RNA pool. This is especially true when attempting the direct selection of ribozymes that catalyze a reaction between small organic substrates.[9,10] In these cases, the biopolymeric part is catalyzing the reaction between a covalently linked reactant and a free reactant without being itself modified. The occurrence of the desired reaction should nevertheless bring about a change which allows preparative isolation of the active RNA sequences, i.e. those that catalyze the reaction. For instance, the reaction of the RNA-attached substrate with a second substrate could allow tagging of the active sequences with biotin.[10−12] The biotinylated sequences can then be isolated from the rest of the pool, enzymatically copied and amplified, and used as input for the next round of the iterative in vitro selection process. While chemical methods are known for the covalent attachment of small organic moieties to RNA, most of them cannot be used during in vitro selection, as they are either incompatible with the enzymatic steps or not selective or reliable enough for derivatizing complex mixtures such as combinatorial pools of nucleic acids.

In vitro transcription from DNA templates by DNA-dependent T7 RNA polymerase, on the other hand, is a simple and reliable enzymatic method allowing the synthesis of virtually any RNA sequence.[13,14] Using this method guanosine monophosphate derivatives can be incorporated selectively at the 5′-end of a transcribed oligoribonucleotide.[15−17] Transcription reactions in the presence of guanosine derivatives carrying covalently appended an organic moiety (the reaction substrate during the selection process) result therefore in the site-specific covalent attachment of this moiety to the RNA. Most importantly, the conjugation occurs under very mild conditions, which are in principle compatible with a wide variety of substrates. Although very powerful, this method has so far found only limited application,[10,18−20] probably because of the laborious multistep synthesis required for the preparation of these guanosine monophosphate derivatives, often named initiator nucleotides.

We report here a more versatile and concise synthetic strategy for the preparation of initiator nucleotides bearing the potential substrate attached to the end of a decaethylene glycol spacer (**1a**−**d**). The adopted synthetic scheme is highly convergent and

(1) Bittker, J. A.; Phillips, K. J.; Liu, D. R. *Curr. Opin. Chem. Biol.* **2002**, *6*, 367−374.
(2) Breaker, R. R. *Chem. Rev.* **1997**, *97*, 371−390.
(3) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611−647.
(4) Pan, T. *Curr. Opin. Chem. Biol.* **1997**, *1*, 17−25.
(5) Williams, K. P.; Bartel, D. P. In *Catalytic RNA*; Eckstein, F., Lilley, D. M. J., Eds.; Springer: Berlin, 1996; Vol. 10, pp 367−381.
(6) Murphy, M. B.; Fuller, S. T.; Richardson, P. M.; Doyle, S. A. *Nucleic Acids Res.* **2003**, *31*, e110.
(7) Famulok, M.; Mayer, G.; Blind, M. *Acc. Chem. Res.* **2000**, *33*, 591−599.
(8) Famulok, M. *Curr. Opin. Struct. Biol.* **1999**, *9*, 324−329.
(9) Keiper, S.; Bebenroth, D.; Seelig, B.; Westhof, E.; Jäschke, A. *Chem. Biol.* **2004**, *11*, 1217−1227.
(10) Sengle, G.; Eisenfuhr, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001**, *8*, 459−473.
(11) Jadhav, V. R.; Yarus, M. *Biochemistry* **2002**, *41*, 723−729.
(12) Seelig, B.; Jäschke, A. *Chem. Biol.* **1999**, *6*, 167−176.
(13) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783−8798.
(14) Milligan, J. F.; Uhlenbeck, O. C. *Methods Enzymol.* **1989**, *180*, 51−62.
(15) Martin, C. T.; Muller, D. K.; Coleman, J. E. *Biochemistry* **1988**, *27*, 3966−3974.
(16) Contreras, R.; Cheroutre, H.; Degrave, W.; Fiers, W. *Nucleic Acids Res.* **1982**, *10*, 6353−6362.
(17) Konarska, M. M.; Padgett, R. A.; Sharp, P. A. *Cell* **1984**, *38*, 731−736.
(18) Eisenfuhr, A.; Arora, P. S.; Sengle, G.; Takaoka, L. R.; Nowick, J. S.; Famulok, M. *Bioorg. Med. Chem.* **2003**, *11*, 235−249.
(19) Zhang, L.; Sun, L.; Cui, Z.; Gottlieb, R. L.; Zhang, B. *Bioconjugate Chem.* **2001**, *12*, 939−948.
(20) Seelig, B.; Jäschke, A. *Bioconjugate Chem.* **1999**, *10*, 371−378.

Scheme 1. Guanosine Monophosphate Derivatives Used as Initiator Nucleotides in T7 RNAP Transcriptions (N, Nontemplated Nucleotide Added at the 3′-end of the Transcripts)



relies on two common intermediates, viz., monoprotected decaethylene glycol **2** and base-unprotected guanosine 5′-phosphoramidite **3**. Furthermore screening of the transcription reaction conditions has allowed an unprecedented efficient incorporation of all synthesized compounds at the 5′-end of a 25-mer RNA transcript. Thus up to 90−95% of the transcribed RNA molecules carry a benzylallyl ether moiety or an anthracene residue when the transcription mixtures contain **1a** and **1b**, respectively. The incorporation of initiator nucleotides **1c** and **1d** are up to 68 and 49%, suggesting that more hydrophobic substrates are incorporated more efficiently, all the other parameters remaining unchanged (spacer nature and length, transcription conditions).

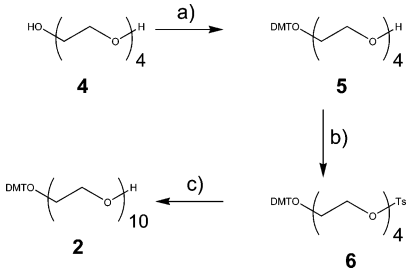
Results and Discussion

Recent developments in ribozyme catalysis require chemoenzymatic access to complex RNA conjugates. The incorporation of guanosine monophosphate derivatives at the 5′-end of RNA transcripts has been used as a valuable tool for attaching potential substrates to each individual sequence of a RNA pool during in vitro selection of ribozymes.[12,18] It is therefore interesting to understand how versatile this method could be and under which conditions best results concerning conjugation could be achieved. To this end, four guanosine monophosphate derivatives **1a**−**d** have been prepared displaying different functional groups at the end of a decaethylene glycol spacer (Scheme 1). These functional groups are benzylallyl ether, anthracene, benzyl carbamate, and a primary amino group for **1a**, **1b**, **1c**, and **1d**, respectively.

The novel synthetic strategy here reported allows the preparation of these four compounds using two common intermediates, namely, mono-DMT-decaethylene glycol **2** and 2′,3′-bisTBDMS protected guanosine 5′-*O*-phosphoramidite **3**. The potential substrates, corresponding to the R groups of compounds **1a**−**1c**, are linked to the terminal hydroxyl group of the ethylene glycol spacer via alkylation reactions, while **1d** was obtained by catalytic hydrogenolysis of the Cbz protected precursor **1c**.

Synthesis of Common Intermediates. The synthesis of **2** started from inexpensive and commercially available tetraethylene glycol which was mono-DMT protected in 79% yield. Slow addition over 16 h of dimethoxytrityl chloride (0.23 equiv)

Scheme 2. Synthesis of Monoprotected Decaethylenethylene Glycol **2**[a]



[a] Reagents and conditions: (a) DMTCl (0.23 equiv), Et_3N (0.37 equiv), DMAP (0.01 equiv), CH_2Cl_2, rt, 16 h, 79%; (b) TsCl (1.1 equiv), Et_3N (3.0 equiv), DMAP (0.05 equiv), CH_2Cl_2, rt, 12 h, 99%; (c) hexaethylene glycol (3.0 equiv), NaH (2.0 equiv), DMF, rt, 13 h, 73%. DMT = 4,4′-dimethoxytriphenylmethyl, DMAP = 4-(dimethylamino)pyridine.

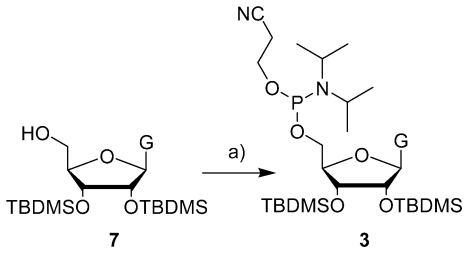
to the glycol (**4**) gave substantially higher yields of mono- DMT-tetraethylene glycol **5** than the 63% so far reported.[21] Tosylation (giving **6**) followed by monoalkylation of commercially available hexaethylene glycol afforded **2** in 73% yield under optimized conditions. Multigram amounts of this internediate are therefore readily available involving only two chromatographic steps for the purification of **5** and **2,** respectively (Scheme 2). This recursive synthesis of oligoethylene glycols should allow the simple preparation of monodisperse tethers of virtually any length just by iterating the tosylation and alkylation step.

2′,3′-bisTBDMS-protected guanosine **7**[22,23] is readily available from unprotected guanosine and was easily converted to the corresponding 5′-*O*-phosphoramidite **3** in 98% yield using 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite as phosphitylating agent. Notably no protecting groups have been used on the nucleobase in this reaction (Scheme 3).

Preparation of Substrate-Carrying Decaethylene Glycol Derivatives. One of the key synthetic steps allowing the preparation of initiator nucleotides with different potential substrates appended at the end of a long tether is the alkylation

(21) Rumney, S. I. V.; Kool, E. T. *J. Am. Chem. Soc.* **1995**, *117*, 5635−5646.
(22) Kotch, F. W.; Sidorov, V.; Lam, Y.-F.; Kayser, K. J.; Li, H.; Kaucher, M. S.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 15140−15150.
(23) Zhu, X.-F.; Williams, H. J.; Scott, A. I. *J. Chem. Soc., Perkin Trans. 1* **2000**, 2305−2306.

Scheme 3. Synthesis of Guanosine Phosphoramidite **3**[a]



[a] Reagents and conditions: (a) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.05 equiv), DIPEA (3.2 equiv), CH_2Cl_2, 0 °C→rt, 1 h, 98%. DIPEA = ethyldiisopropylamine. Compound **3** is a mixture of two diastereoisomers (ca. 1:1).

of **2** with the desired substrate in the form of an alkylating agent. Compounds **11** and **14** were therefore prepared in a few-step synthesis starting from 4-chloromethyl benzyl alcohol and 1,4-diaminobutane (Scheme 4).

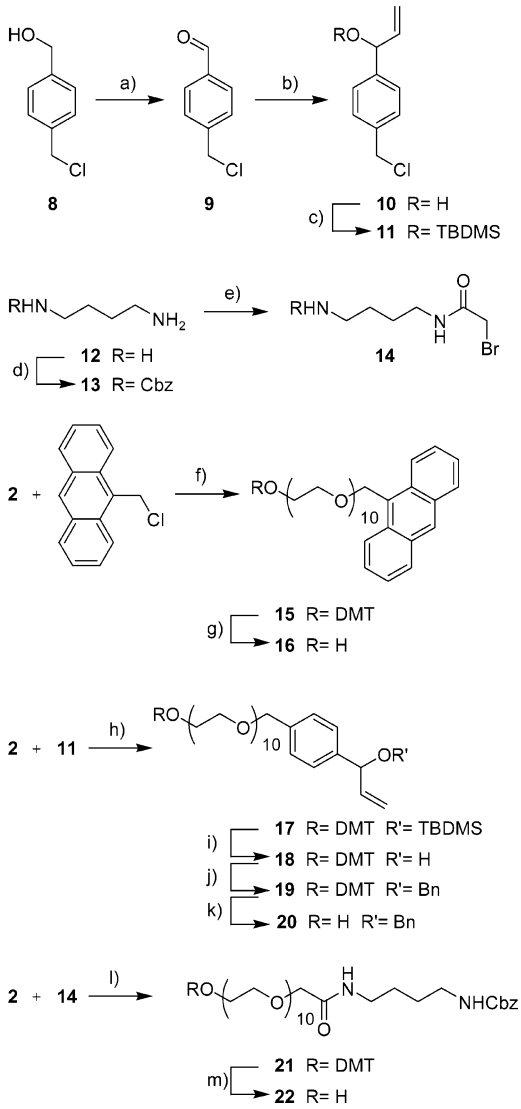
More in detail, oxidation of 4-chloromethylbenzyl alcohol **8** to the corresponding aldehyde **9**[24] followed by reaction with vinylmagnesium bromide gave benzyl allyl alcohol **10**. The hydroxyl group was then protected with TBDMSCl affording **11**. 1,4-Diaminobutane **12** was monoprotected with *N*-(benzyloxycarbonyloxy)succinimide in 44% yield and successively reacted with bromoacetyl bromide under Schotten–Baumann conditions affording **14** in 75% yield.

Alkylation of **2** with 9-chloromethylanthracene, **11**, and **14** in acetonitrile in the presence of NaH gave the desired adduct **15**, **17**, and **21**, in 73, 90, and 62% yield, respectively (Scheme 4). NaI was used as catalyst only when the leaving group was a chloride. Further elaboration on the protecting groups was then required in order to obtain the hydroxyl-terminated derivatives to be used in the phoshoramidite coupling reaction. Therefore the TBDMS group of **17** was removed and replaced by a benzyl to give **19** before removal of the DMT group affording the desired alcohol **20**. Removal of the DMT group from **15** and **21** afforded directly alcohols **16** and **22**.

Synthesis of Guanosine Monophosphate Derivatives. According to several literature reports even relatively long oligonucleotides (e.g. 60-mers) were synthesized via the phosphoramidite method without nucleoside base protection.[25,26] Moreover, it has been observed that guanosine does not undergo *N*-phosphitylation during phosphoramidite coupling in the presence of activators such as imidazolium triflate, benzimidazolium triflate, 1-*H*-tetrazole, or 5-(*p*-nitrophenyl)-1*H*-tetrazole.[25] To keep the synthetic effort to a minimum, we decided therefore to perform our coupling reactions also in the absence of additional protecting groups on the nucleoside base (guanine). Guanosine monophosphate derivatives were obtained by coupling **3** with hydroxyl-terminated decaethylene glycol building blocks (**16**, **20**, and **22**) carrying the substrate at the other end of the tether in the presence of activators such as 4,5-dicyanoimidazole, 5-benzylthiotetrazole, or imidazolium triflate (Schemes 5 and 6).

The coupling reactions were quenched by addition of *tert*-butylhydroperoxide, giving a phosphotriester intermediate as a mixture of diastereoisomers due to the stereogenic phosphor atom. However, the phosphodiester group in the final products **1** (after removal of the 2-cyanoethyl protecting group) is not stereogenic. It proved therefore more convenient to directly treat the phosphotriester intermediates with TBAF to simultaneously remove all protecting groups (TBDMS and 2-cyanoethyl). The crude products were then purified by RP-chromatography giving **1a**, **1d**, and **1c** in 42, 31, and 35% yield, respectively.

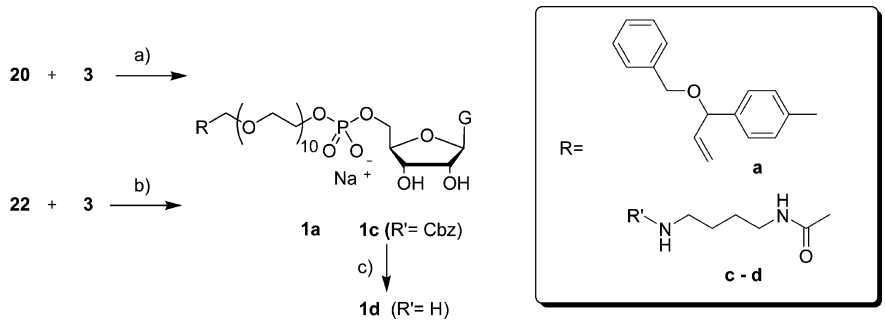
Scheme 4. Synthesis of Substrate–Decaethylene Glycol Derivatives[a]



[a] Reagents and conditions: (a) PCC (1.2 equiv), CH_2Cl_2, rt, 2 h, 94%; (b) vinylmagnesium bromide (1.05 equiv), THF, −15 °C, 30 min, 83%; (c) TBDMSCl (1.3 equiv), imidazole (3.3 equiv), DMF, rt, 1 h, 99%; (d) Z-OSu (0.25 equiv), DMAP (0.005 equiv), THF, rt, 23 h, 44%; (e) bromoacetylbromide (1.5 equiv), aq Na_2CO_3/CH_2Cl_2 1:1, 0 °C, 1 h, 75%; (f) **2** (1.0 equiv), 9-chloromethylanthracene (1.1 equiv), NaH (1.6 equiv), NaI (1.2 equiv), ACN, 40 °C, 0.5 h, 73%; (g) 3% TCA in DCE (excess), rt, 5 min, 88%; (h) **2** (1.0 equiv), **11** (1.5 equiv), NaH (1.6 equiv), NaI (1.1 equiv), ACN, 40 °C, 1.5 h, 90%; (i) TBAF (2.0 equiv), THF, rt, 1 h, 90%; (j) BnBr (3.0 equiv), NaH (1.5 equiv), ACN, rt, 1h, 91%; (k) 3% TCA in DCE (excess), rt, 5 min, 89%; (l) **2** (1.0 equiv), **14** (1.5 equiv), NaH (1.6 equiv), ACN, 40 °C, 1.5 h, 62%; (m) 3% TCA in DCE (excess), rt, 5 min, 93%. PCC = pyridinium chlorochromate, Z-OSu = *N*-(benzyloxycarbonyloxy)succinimide, ACN = acetonitrile, TBAF = tetra-*n*-butylammonium fluoride, Bn = benzyl, TCA = trichloroacetic acid, DCE = 1,2-dichloroethane. Compound **20** and its precursors are racemic.

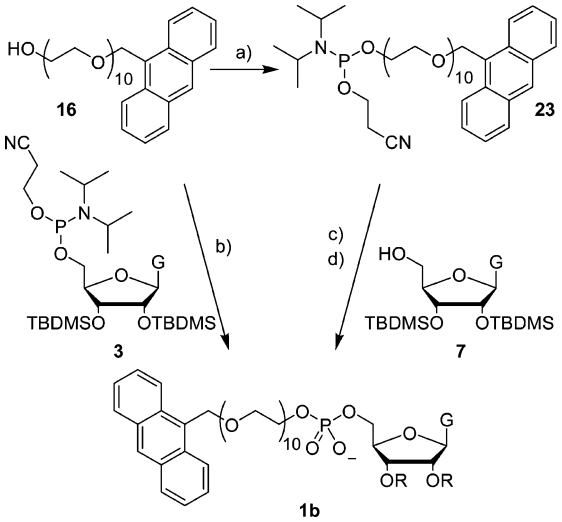
(24) Pierce, M. E.; Harris, G. D.; Islam, Q.; Radesca, L. A.; Storace, L.; Waltermire, R. E.; Wat, E.; Jadhav, P. K.; Emmett, G. C. *J. Org. Chem.* **1996**, *61*, 444−450.
(25) Hayakawa, Y.; Kataoka, M. *J. Am. Chem. Soc.* **1998**, *120*, 12395−12401.
(26) Ohkubo, A.; Ezawa, Y.; Seio, K.; Sekine, M. *J. Am. Chem. Soc.* **2004**, *126*, 10884−10896.

Scheme 5. Synthesis of Guanosine Phosphate Derivatives **1a**, **1c**, and **1d**[a]



[a] Reagents and conditions: (a) **20** (1.0 equiv), **3** (1.4 equiv), IMT (0.56 equiv), ACN/THF 1:1, rt, 1 h, then TBHP (10 equiv), rt, 10 min; volatiles removal, and then TBAF (10 equiv), THF, rt, 1h, 41% over two steps; (b) **22** (1.0 equiv), **3** (1.2 equiv), DCI (3.0 equiv), THF, rt, 1 h, then TBHP (10 equiv), rt, 10 min; volatiles removal, and then TBAF (10 equiv), THF, rt, 3h, 36% over two steps; (c) H_2, Pd/C, MeOH, rt, 2 h, 94%. IMT = imidazolium triflate, TBHP = *tert*-butylhydroperoxide, DCI = 4,5-dicyanoimidazole.

Scheme 6. Two Synthetic Strategies for Anthracene−Decaethylene Glycol Guanosine Phosphate **1b**[a]



[a] Reagents and conditions: (a) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.2 equiv), DIPEA (3.0 equiv), CH_2Cl_2, 0 °C→rt, 1 h, 98%; (b) **3** (1.2 equiv), DCI (1.0 equiv), THF, rt, 20 min, then TBHP (10 equiv), rt, 10 min; volatiles removal, and then TBAF (6.0 equiv), THF, rt, 1h, 31% over two steps; (c) **7** (1.2 equiv), BTT (10 equiv), ACN, rt, 9 h and then TBHP (10 equiv), rt, 10 min; (d) TBAF (5.0 equiv), THF, rt, 1h, 20% over two steps. BTT = 5-benzylthiotetrazole.

An alternative strategy was also considered for the preparation of guanosine monophosphate derivatives, namely, the coupling between **7** and phosphoramidites prepared from the decaethylene glycol derivatives (e.g. **23**). One disadvantage of this approach is that one different phosphoramidite should be synthesized for each of the different chains to be coupled. Nonetheless, the synthesis of **1b** (Scheme 6) was carried out following this second strategy (**23** + **7**, 20% yield). These experiments showed two major problems, namely, the very poor solubility of **7** in all the solvents typically used for coupling (ACN, THF, CH_2Cl_2) and a considerable more difficult purification of the desired product. In fact, several anthracene-carrying byproducts with polarities similar to that of the desired product were observed when phosphoramidite **23** was used, probably originated from side reactions at the phosphorus. In contrast the purification of **1b**, prepared according to the first strategy (**16** + **3**, 31% yield), was considerably simpler. Only two species containing anthracene were then present in the reaction mixture, i.e. the desired **1b** and unreacted **16,** and they could be easily separated by RP chromatography.

Finally phosphate **1d**, bearing a primary amino group at the end of the decaethylene glycol spacer, was obtained removing the Cbz group from **1c** via catalytic hydrogenation.

Incorporation of Compounds 1a−d at the 5′-end of RNA Transcripts. Bacteriophage RNA polymerases (RNAP) have been shown to be highly active for in vitro transcriptions from synthetic DNA templates.[13,27] In fact, T7 RNAP transcriptions can be used for the synthesis of sizable amounts of RNA (up to several milligram) starting from synthetic DNA templates.[14] It is also known that a number of modifications can be introduced at the 5′-end of RNA transcripts using bacteriophage RNAP. Thus, "capped" RNA transcripts are obtained simply by adding to the mixture of nucleotide triphosphates (required for polymerization) an appropriate guanosine monophosphate derivative.[16,17] Even dinucleotides and nucleotides with modified bases can be used as initiators. The incorporation occurs selectively at the 5′-end since during the elongation step only triphosphates can be used. On the basis of previous data from this laboratory,[20,28] and from others[19] it was known that guanosine monophosphate derivatives carrying long, flexible poly(ethylene glycol) spacers were accepted by T7 RNAP during in vitro transcriptions with moderate incorporation ratios. However, only very limited information exists regarding the influence of various parameters (like concentration of initiator nucleotide, of GTP, and of the other NTPs) on the incorporation efficiency and the total transcription yields.[18−20] Moreover, the influence of the chemical nature of the residue attached at the end of the spacer has not been explored at all.

We have systematically investigated the incorporation conditions for compounds **1a**,**b**,**d** using a double-stranded DNA template encoding for a 25-nucleotide long transcript. Then, under optimized conditions, all four initiator nucleotides **1a**−**d** have been incorporated in high yield.

The first set of experiments was performed keeping constant the concentrations of ATP, CTP, and UTP (each 1 mM) and the amounts of **1a**,**b**,**d** (4 mM), and varying the amount of GTP (which competes with the initiator nucleotide in the initiation step but is essential for elongation) between 0.2 and 2 mM (ratio **1**:GTP from 20:1 to 2:1). Gel electrophoresis analysis of the transcription reactions showed for **1a**,**b**,**d** a similar trend in total yield of RNA transcripts, i.e. the sum of both the desired

(27) Melton, D. A.; Krieg, P. A.; Rebagliati, M. R.; Maniatis, T.; Zinn, K.; Green, M. R. *Nucleic Acids Res.* **1984**, *12*, 7035−7056.
(28) Seelig, B.; Jäschke, A. *Tetrahedron Lett.* **1997**, *38*, 7729−7732.

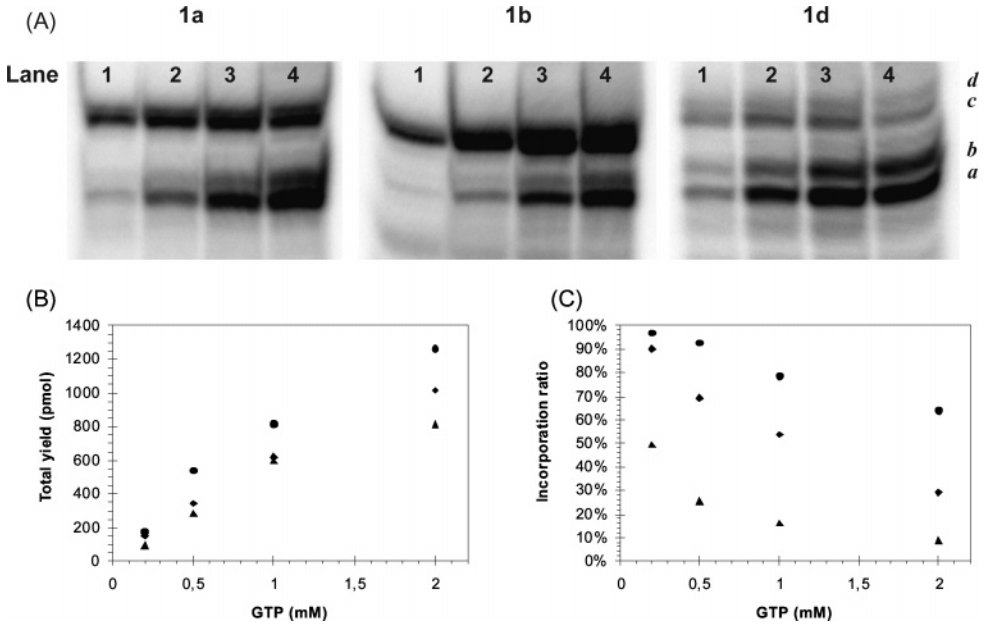


Figure 1. Enzymatic incorporation of initiator nucleotides **1a** (◆), **1b** (●), and **1d** (▲) (4 mM) via T7 RNAP transcription (2 h reaction time) of a 44-mer dsDNA template in the presence of ATP, CTP, and UTP (each 1 mM) and varying concentration of GTP. (A) Phosphorimages of the PAGE analyses. Lanes 1−4: GTP 0.2, 0.5, 1.0, and 2.0 mM. Transcription products *a* and *b*: 25- and 26-mer nonmodified transcripts. *c* and *d*: 25- and 26-mer modified transcripts. 26-mers result from nontemplated addition of one nucleotide to the 3′-end of transcribed RNA. (B) The total yields (a + b + c + d) increase with increasing GTP concentration, while the incorporation ratios (c + d divided by the total yield) decrease (C).

initiated transcript and the normal one with a G at its 5′-end, under identical conditions (Figure 1).

In fact, the total yield of RNA is increased by increasing the amount of GTP showing almost no saturation behavior at the used concentrations, as expected from previous reports.[13] However, it is interesting to note that the total yields are influenced by the nature of the different initiator nucleotides: systematically higher yields are obtained when compound **1b** was present in the reaction mixture. The incorporation yield, i.e. the amount of initiated transcripts divided by the total amount of transcripts, increased by lowering the amount of GTP (increasing the **1**:GTP ratio), but the effect was notably different for the various initiator nucleotides. At 2 mM GTP, compounds **1a** and **1d** are incorporated only modestly (29 and 9%), while **1b** is incorporated much more efficiently (64%). Decreasing the GTP concentration from 2 to 0.2 mM resulted in ∼3 and ∼5 times better incorporation yields (90 and 49%) for **1a** and **1d**, respectively. For **1b** the observed increment was ∼1.5 times, affording 97% of the initiated transcript.

It is also known that an increase in the concentration of nucleotide triphosphates (NTPs) up to 4 mM gives a substantial increment in the total transcription yields.[13] We therefore carried out a set of experiments keeping constant the ratio between **1a**,**b**,**d** and GTP (4 and 1 mM) and varying the concentrations of the other NTPs. The results showed indeed that the total yields increased increasing NTP concentrations. Unexpectedly also the incorporation ratio slightly increased (5−10%) (Figure 2). This result indicates that the conditions favoring higher transcription yields plays in the same direction as those favoring higher incorporation. Thus, it should be possible to increase the total transcription yields while still keeping high incorporation yields by working at higher NTP concentration (>1 mM) and increasing the GTP concentration (>0.2 mM). The next set of experiments confirmed our expectations (Figure 3). By raising the GTP concentration to 0.35 mM and the ATP, CTP, and UTP

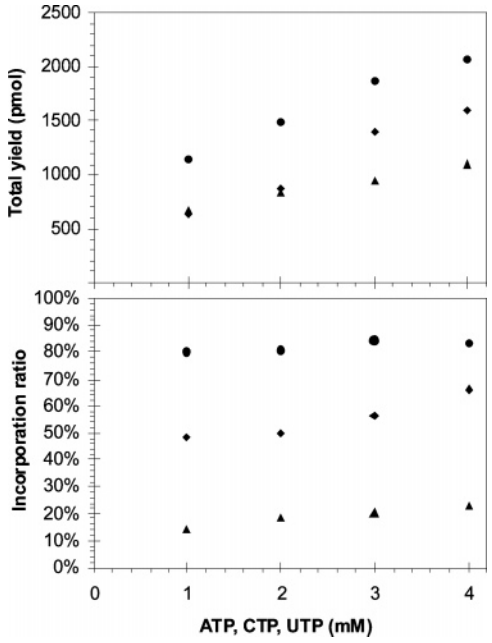


Figure 2. Enzymatic incorporation of initiator nucleotides **1a** (◆), **1b** (●), and **1d** (▲) (4mM) via T7 RNAP transcription (2 h reaction time) of a 44-mer dsDNA template: GTP, 1 mM; concentrations of ATP, CTP, and UTP (in equimolar amount), varying between 1 and 4 mM. Transcription yields (top) and the incorporation ratios (bottom) increase with increasing ATP, CTP, and UTP concentrations.

to 4 mM, initiator nucleotides **1a**−**c** were incorporated in 81, 91 and 68%, respectively, and the total yields for **1a** and **1b** increased by 70 and 80%, respectively. These increments compare well with the ones observed in previous experiments (Figures 1 and 2). Initiator **1d** was incorporated in a more modest 41% (the highest incorporation observed in the previous set of experiments was 49%), but the total yield increased almost 150%. It is also interesting to note that under these conditions

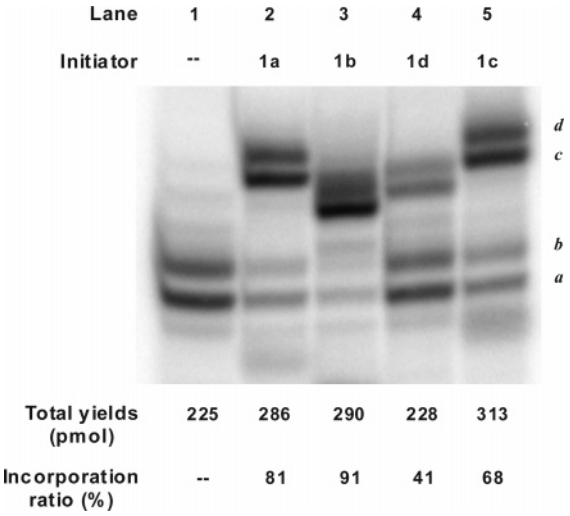


Figure 3. Optimized transcriptions initiation reactions (2 h reaction time) using a 44-mer dsDNA template. Initiator nucleotides **1a** (lane 2), **1b** (lane 3), **1c** (lane 5), and **1d** (lane 4), 4 mM; GTP, 0.35 mM; ATP, CTP, and UTP, each 4 mM. a and b: 25- and 26-mer nonmodified transcripts. c and d: 25- and 26-mer modified transcripts. In lane 1 the transcription reaction is run under identical conditions except for the absence of any initiator nucleotide.

the total transcription yields remained substantially constant for all initiator nucleotides investigated (Figure 3) and were comparable with the transcription yield obtained in the absence of any initiator nucleotide.

The reasons for the lower incorporation of **1d** are at present not fully understood. Most probably the primary amino group at the end of the decaethylene glycol chain is protonated at the pH used in the transcription experiments, significantly affecting the recognition of this substrate by T7-RNAP.

Conclusion

Incorporation of guanosine monophosphate derivatives at the 5′-end of RNA transcripts is an efficient, reliable method to produce RNA conjugates bearing covalently attached substrates for in vitro selection of novel ribozymes. The synthetic effort to prepare such molecules has been minimized by developing a versatile path which involves the two common intermediates **2** and **3**. It is likely that a wide variety of small molecule functionalities attached via a PEG-type spacer to the phosphate moiety of GMP can be prepared by making use of these intermediates. Intermediate **21** is expected to be of particular synthetic value, since it could be deprotected and reacted selectively at the amino terminus with active ester derivatives. Therefore even delicate substrates could then be attached to the decaethylene glycol spacer, without involving them directly during the harsh alkylation step. Efforts in this direction are currently underway in our laboratories. The data here presented also show that initiator nucleotides **1a**−**d** are well-accepted as substrates by the enzyme T7 RNAP in the runoff transcription reactions of synthetic dsDNA templates, without giving rise to any detectable inhibition. Under appropriate conditions, like GTP $\leq$ 0.35 mM, the majority of the full length transcript is initiated by the initiator nucleotide (up to 95%), the only exception being the initiator nucleotide carrying an unprotected amino group. Current work in our laboratory is now focusing on the development of a novel ribozyme which catalyses allylic substitution reactions in combination with transition metal complexes making use of initiator nucleotide **1a**.

Acknowledgment. This work was supported by the Deutsche Forschungsgemeinschaft (Grant SFB 623) and the Fonds der Chemischen Industrie. The authors thank Mihaela Caprioara for performing the MALDI TOF measurements and Prof. R. Krämer and Dr. A. Mokhir for helpful discussion and technical support. The authors also thank Nancy Sauer for technical assistance.

Supporting Information Available: Full Experimental Section including procedures and spectroscopic data for all new compounds and in vitro transcription procedures and MALDI TOF analysis of the transcripts. This material is available free of charge via the Internet at http://pubs.acs.org.

JA051179M