Ternary Conjugates of Guanosine Monophosphate as Initiator Nucleotides for the Enzymatic Synthesis of 5′-Modified RNAs

Burckhard Seelig and Andres Jäschke*

Institut für Biochemie, Freie Universität Berlin, Thielallee 63, D-14195 Berlin. Received July 23, 1998; Revised Manuscript Received October 26, 1998

We give a detailed account on the enzymatic synthesis of RNA conjugates by T7 RNA polymerase using modified initiator nucleotides during transcription. Following two different routes, ternary conjugates of guanosine-5′-monophosphate, poly(ethylene glycol), and anthracene were synthesized via phosphoramidite intermediates and characterized by a variety of spectroscopic techniques. Up to a degree of polymerization n_{PEG} of about 17, these conjugates were efficiently incorporated into RNA by T7 RNA polymerase at the 5′-termini, thereby giving access to RNA conjugates required for biochemical studies as well as for the exploration of the catalytic potential of ribonucleic acids. The resulting conjugates are intact and functional.

INTRODUCTION

Oligonucleotide conjugates have found many uses in biochemical research and medical diagnostics, and a large number of conjugate groups have been used (Agrawal, 1994; Beaucage and Iyer, 1993; Goodchild, 1990). The most common strategies for site-specific conjugation of oligonucleotides are (i) transformation of the conjugate group into a phosphoramidite and direct incorporation during automated solid-phase synthesis of oligonucleotides and (ii) incorporation of a reactive linker during oligonucleotide synthesis and postsynthetic attachment of the conjugate group to this linker.

While these conjugation strategies work well with short oligonucleotides accessible by chemical synthesis, the possibilities to conjugate enzymatically synthesized long nucleic acid sequences are very limited [e.g., Oh and Pace (1994) and Tong et al. (1993)]. This is especially problematic with ribonucleic acids, where chemical synthesis is normally limited to a chain lengths of about 50 nucleotides. Most RNAs of biological interest are much longer, and for biochemical studies they are either isolated from biological sources or prepared by enzymatic transcription from DNA templates using viral RNA polymerases (Milligan et al., 1987). To study RNA processing, translocation, or ribozyme reaction mechanisms, it is highly desirable to have simple and reliable methods available that allow the site-specific incorporation of conjugate groups (Bevilacqua et al., 1994; Igloi, 1996; Moore and Sharp, 1992).

Recently, ternary RNA conjugates have successfully been used for the in vitro selection of new catalytic RNA species (Tarasow et al., 1997; Wiegand et al., 1997; Seelig and Jäschke, 1999). A key feature of these strategies is the use of synthetic RNA conjugate libraries in which a small organic residue as a potential reactant is attached to each individual sequence of a combinatorial RNA pool via a flexible polymeric tether. These RNA conjugates are then incubated with a second reactant carrying a biotin tag. If an RNA molecule catalyzes the reaction of the attached reactant with the biotinylated free reactant, the reaction product would also carry the biotin tag and could be isolated by affinity chromatography. The RNA part of the isolated conjugate is then enzymatically amplified. The cycle of conjugation, reaction, selection, and amplification is repeated until the active species are sufficiently enriched from the ancestral RNA pool. This approach was successfully applied to the selection of RNA Diels–Alderases and RNA amide synthases (Tarasow et al., 1997; Wiegand et al., 1997; Seelig and Jäschke, 1999).

Toward the same end, we have described the synthesis of multifunctional RNA conjugates in which a small organic molecule is site-specifically attached to the 3′-terminal hydroxyl group of the RNA via a long flexible poly(ethylene glycol) tether harboring a photocleavage site (Hausch and Jäschke, 1997). In that case, we made use of an enzymatic two-step procedure involving preparation of the unmodified RNA molecule by enzymatic transcription and subsequent addition of a conjugated dinucleotide by T4 RNA ligase.

In a recent short communication, we have described an enzymatic one-step procedure which permits an easy access to RNA 5′-conjugates (Seelig and Jäschke, 1997). This procedure is based on the known tolerance of T7 RNA polymerase toward certain chemical modifications in the initiation step (Pitulle et al., 1992). The conjugate group is added at the 5′-end during transcription by T7 RNA polymerase, which utilizes ternary conjugates of guanosine monophosphate, poly(ethylene glycol), and anthracene (as a potential organic reactant) exclusively during the initiation, but not during the elongation steps. While a number of reports have been published on the use of different initiator nucleotides, a systematic investigation of this method has not been carried out so far. Here, we give a detailed account on the chemical synthesis of ternary initiator nucleotides with varying lengths of the poly(ethylene glycol) linker, an investiga-

* To whom correspondence should be addressed. Phone: (49) 30 838 6023. Fax: (49) 30 838 6413. E-mail jaschke@chemie.fu-berlin.de.

[1] Abbreviations: DMF, *N*,*N*-dimethylformamide; DNA, deoxyribonucleic acid; dsDNA, double-stranded DNA; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid); MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight (mass spectrometry); NTPs, (ribo)nucleoside 5′-triphosphates; PCR, polymerase chain reaction; PEG, poly(ethylene glycol); RNA, ribonucleic acid.

Published on Web 00/00/0000

tion of their enzymatic incorporation, and a detailed characterization of the enzymatically synthesized conjugates.

MATERIALS AND METHODS

General. Poly(ethylene glycol) was purchased from Fluka; all other chemicals, buffers, solvents and reagents were obtained as the highest commercial grade available and used without further purification. Reactions were carried out at room temperature unless stated otherwise. Silica gel (0.063−0.200 mm) for column chromatography was obtained from J. T. Baker. HPLC purification was performed on a Beckman System Gold chromatograph using an Ultrasphere ODS reversed-phase column (4.6 × 250 mm, flow 1 mL/min). NMR spectra were recorded using a Bruker A250 (^{13}C) and a Bruker AMX500 (^{1}H, ^{31}P). ESI mass spectra were recorded on a Finnigan MAT 7M spectrometer, and MALDI TOF mass spectra were obtained using a Bruker Reflex spectrometer. UV spectra were recorded with a Shimadzu UV-160A.

Mono-[9-(methyl)anthracene]-PEG 600 Conjugate and Mono-[9-(methyl)anthracene]-PEG 1500 Conjugate 1. PEG 600 (3.6 g, 6 mmol) or PEG 1500 (9 g, 6 mmol) was dried in vacuo (P_4O_{10}) and dissolved in dry benzene (30 mL). *n*-Butyllithium (2.5 mL, 6.2 mmol, 2.5 M solution in hexane) was added dropwise under vigorous stirring. After further stirring for 2 h, benzene was removed by rotary evaporation, and the product was redissolved in dry *N*,*N*-dimethylformamide (45 mL). 9-(Chloromethyl)anthracene (1.36 g, 6 mmol) was added, and the mixture was stirred for 2 h at room temperature. The solvent was removed under reduced pressure, and the resulting oil subjected to silica gel chromatography (CH_2Cl_2/MeOH 96:4 v/v for PEG 600 conjugate and CH_2Cl_2/MeOH 9:1 v/v for PEG 1500 conjugate). Concentration in vacuo afforded mono-[9-(methyl)anthracene]-PEG 600 **1** with a yield of 39% [mono-[9-(methyl)anthracene]-PEG 1500 with a yield of 48%], referred to PEG. ^{1}H NMR (250 MHz, [D$_1$]CHCl$_3$, 25 °C): δ = 3.5−3.8 (m, 56 H; CH_2−O for n = 14), 5.6 (s, 2H; CH_2−Ar), 7.48 (m, 4H; H-2/H-3/H-6/H-7), 8.0 (d, 2H; H-1/H-8), 8.4 (d, 2H; H-4/H-5), 8.42-(s, 1H; H-10). ^{13}C NMR (63 MHz, [D$_1$]CHCl$_3$, 25 °C): δ = 131.2, 130.8, 128.8, 128.6, 128.1, 125.9, 124.7, 124.3. ESI$^+$-MS: calculated for $C_{43}H_{68}O_{15}$ (n = 14) [MH$^+$] = 825.0, found 825.

Mono-[9-(methyl)anthracene]-PEG 600-(2-cyanoethyl-*N*,*N*-diisopropyl)-phosphoramidite 2. Mono-[9-(methyl)anthracene]-PEG 600 **1** (260 mg, 325 μmol) was coevaporated twice from CH_2Cl_2/pyridine (9:1 v/v) and dried in vacuo overnight (P_4O_{10}). *N*,*N*-Diisopropylethylamine (585 μL, 3.25 mmol) and dichloromethane (1 mL, refluxed with CaH_2 and distilled in the dark immediately before use) were added under argon. (2-Cyanoethyl-*N*,*N*-diisopropyl)chlorophosphoramidite (73 μL, 325 μmol) was added dropwise, and the mixture was stirred for 30 min in the dark. A total of 50 μL of anhydrous methanol was added, and the reaction mixture was diluted with CH_2Cl_2 to 30 mL. After extraction with saturated $NaHCO_3$ solution and with water, the organic phase was dried over Na_2SO_4 and solvent was removed under reduced pressure to yield 190 mg (58%) of phosphoramidite **2**, which was used for coupling without further purification.

^{31}P NMR (202 MHz, [D$_1$]CHCl$_3$, 20 °C): δ = 146.05 ppm.

2′-O,3′-O,*N*^{2}-Triisobutyrylguanosine 3. N^2-Isobutyryl-5′-*O*-(4,4′-dimethoxytrityl)guanosine (3.28 g, 5 mmol) was added to a solution of isobutyric anhydride (8.33 mL, 50 mmol) in pyridine (50 mL) and stirred for 20 h. Pyridine was distilled off, and the residue was dissolved in water (50 mL) and extracted with dichloromethane (200 mL). After extraction with saturated $NaHCO_3$ solution, saturated NaCl solution and water, the organic phase was dried over Na_2SO_4 and solvent was removed under reduced pressure. The residue was subjected to silica gel chromatography (CH_2Cl_2/MeOH/triethylamine 96:4:0.25 v/v). The pooled fractions were evaporated to dryness and detritylated with trichloroacetic acid (3%) in dichloromethane for 2 min. The solution was extracted with saturated $NaHCO_3$ solution, saturated NaCl solution, and water. After evaporation, the final purification by silica gel chromatography (CH_2Cl_2/MeOH 98:2 v/v) gave 775 mg (31%) of 2′-O,3′-O,N^2-triisobutyrylguanosine **3**. ^{1}H NMR (500 MHz, [D$_1$]CHCl$_3$, 20 °C): δ 1.0−1.23 (m, 18 H; 6 × CH_3-Ibu), 2.4, 2.6 (2 × septet, 2 × 1H; CH-2′,3′-Ibu), 2.78 (septet, 1H; CH-N^2-Ibu), 3.74 (dd, 1H; H-5′), 3.92 (dd, 1H; H-5′), 4.2 (q, 1H; H-4′), 5.63 (dd, 1H; H-3′), 5.86 (d, 1H; H-1′), 5.91 (t, 1H; H-2′), 7.8 (s, 1H; H-8). ^{13}C NMR (63 MHz, [D$_1$]CHCl$_3$, 25 °C): δ = 18.5-18.8 (9 × CH_3), 33.5, 33.7 (CH-2′,3′-Ibu), 36.1 (CH-N^2-Ibu), 61.7 (C-5′),71.1 (C-4′), 72.3 (C-2′), 84.4 (C-1′), 87.4 (C-3′), 122.1 (C-5), 138.9 (C-8), 147.6 (C-4), 148 (C-6), 155.3 (C-2), 175.3 (C=O, 2′-Ibu), 175.8 (C=O, 3′-Ibu), 179.4 (C=O, N^2-Ibu). ESI$^+$-MS calculated for $C_{22}H_{31}O_8N_5$ [M$^+$]: 493. Found: 493.

2′-O,3′-O,*N*^{2}-Triisobutyrylguanosine-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite 4. 2′-O, 3′-O, N^2-Triisobutyrylguanosine **3** (493 mg, 1 mmol) was dried in vacuo overnight (P_4O_{10}) and dissolved in dichloromethane (4 mL, refluxed with CaH_2 and distilled in the dark immediately before use) under argon. *N*,*N*-Diisopropylethylamine (1.22 mL, 7 mmol) was added, and (2-cyanoethyl-*N*,*N*-diisopropyl)chlorophosphoramidite (340 μL, 1.5 mmol) was added dropwise and the mixture was stirred for 30 min in the dark. A total of 100 μL of anhydrous methanol was added, and the reaction mixture was diluted with CH_2Cl_2 to 30 mL. After extraction with saturated $NaHCO_3$ solution and with water, the organic phase was dried over Na_2SO_4 and solvent was removed under reduced pressure. The yielded phosphoramidite **4** was used for coupling without further purification. TLC in $CHCl_2$/MeOH/triethylamine (90:10:0.25 v/v) showed the product spot (R_f = 0.68) and the pale spot of the starting material (R_f = 0.48), indicating almost complete conversion.

Initiator Nucleotide 6. Route 1: 2′-O,3′-O,N^2-Triisobutyrylguanosine **3** (30 mg, 60 μmol) and 1.5 equiv of mono-[9-(methyl)anthracene]-PEG 600-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite **2** (90 mg, 90 μmol) was coevaporated twice from pyridine and dried in vacuo overnight (P_4O_{10}). The two components were dissolved in anhydrous acetonitrile (5 mL), and a solution of tetrazole in acetonitrile (510 μL/0.47 M, 240 μmol) was added under argon. The mixture was stirred for 15 min, and iodine (450 μL/0.1 M in tetrahydrofurane/collidine/water 2:2:1 v/v, 45 μmol) was added. After 5 min, the mixture was diluted with an aqueous solution of sodium thiosulfate (5 mL/1 M, 5 mmol), and the product was extracted with dichloromethane (50 mL). The organic phase was extracted with saturated $NaHCO_3$ solution and water and dried over Na_2SO_4. The solvent was removed under reduced pressure, and the residue was subjected to silica gel chromatography (CH_2Cl_2/MeOH 85/15). The fractions **5** with the typical additive anthracene/guanosine UV spectrum were pooled and deprotected with aqueous ammonia (4 mL, 33%) at 55 °C overnight. After evaporation to dryness, product **6** was finally purified and fractionated according to the degree of polymerization of PEG by isocratic HPLC using 38% acetonitrile in 100 mM

triethylammonium acetate (TEAA), pH 7.0. The fractions were lyophilized.

Route 2: Mono-[9-(methyl)anthracene]-PEG 600 **1** (1.2 g, 1.5 mmol) or mono-[9-(methyl)anthracene]-PEG 1500 (2.25 g, 1.5 mmol) and 0.67 equiv of 2′-O,3′-O,N^2-triisobutyrylguanosine-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite **4** (1 mmol) and were coevaporated twice from pyridine and dried in vacuo overnight (P_4O_{10}). The two components were dissolved in anhydrous acetonitrile (10 mL) and a solution of tetrazole in acetonitrile (8.5 mL/0.47 M, 4 mmol) was added under argon. The mixture was stirred for 15 min and iodine (12 mL/0.1 M in tetrahydrofurane/collidine/water 2:2:1 v/v, 1.2 mmol) was added. After 5 min, the mixture was diluted with an aqueous solution of sodium thiosulfate (1 M) until the brown color of iodine disappeared. The product was extracted with dichloromethane (70 mL). The organic phase was extracted with saturated $NaHCO_3$ solution and with water and dried over Na_2SO_4. The solvent was removed under reduced pressure and the product **5** was chromatographically purified, deprotected, and finally HPLC-purified exactly as described for route 1.

MALDI$^+$-MS calcd for $C_{45}H_{65}O_{18}N_5P$ ($n = 10$) [MH^+]: 994.4. Found: 994. Calcd for $C_{93}H_{159}O_{42}N_5P$ ($n = 34$) [MH^+]: 2050.27. Found: 2049.6. UV (H_2O): λ_{max}= 254.5, 348, 365, 385, 277 nm (shoulder). Fluorescence spectrophotometry (H_2O): emission at $\lambda = 419$ nm upon exitation at $\lambda = 254$ nm (Figure 1).

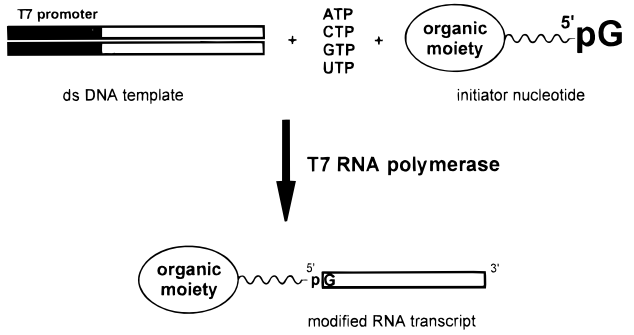
T7 Transcriptions. T7 transcriptions were performed essentially as described (Milligan et al., 1987). DNA templates for T7 transcription were synthesized by standard phosphoramidite chemistry and purified by reversed-phase HPLC (C18) using a gradient from 8 to 32% acetonitrile in 100 mM triethylammonium acetate, pH 7.0, in 30 min. The DNA template for the 25 nt transcript was a duplex of a 44-mer (3′-AG ATT ATG CTG AGT GAT ATC CTC GAG TCG GAA GTG ACG AGG TGG-5′) and a 38-mer (5′-TC TAA TAC GAC TCA CTA TAG GAG CTC AGC CTT CAC TGC-3′), prepared by heating a mixture of the two DNA strands to 95 °C for 3 min and then cooling to room temperature in 1 h. The DNA templates for the randomized 113 and 159 nt transcripts were prepared essentially as described (Hausch and Jäschke, 1997) with completely randomized regions of 74 and 120 nt, respectively, flanked by constant sequences at the termini.

The transcription reactions were run in 80 mM HEPES (pH 7.5), 1 mM spermidine, 22 mM $MgCl_2$, 10 mM dithiothreitol, 120 μg/mL bovine serum albumin, 4 mM NTPs (except GTP), 3 μM [α-^{32}P]CTP (0.2 μCi/μL from Amersham Buchler), 1 μM dsDNA template, 0.5 units/μL ribonuclease inhibitor (MBI Fermentas), and 1 unit/μL T7 RNA polymerase (Stratagene). The concentrations of GTP and initiator nucleotide **6** were varied as indicated in each experiment. The reaction mixtures were incubated at 37 °C for 3 h. After transcription, the RNA was heated in loading buffer containing 50% formamide at 95 °C for 2 min immediately prior to loading on an analytical 15% (25 nt transcript) or 10% (113/159 nt transcript) denaturing polyacrylamide gel. Products were separated on a 24 cm vertical gel apparatus, and the band was located by autoradiography, cut out, and Cerenkov counted in a Beckman LS 6000 SC liquid scintillation counter to determine yields.

RESULTS AND DISCUSSION

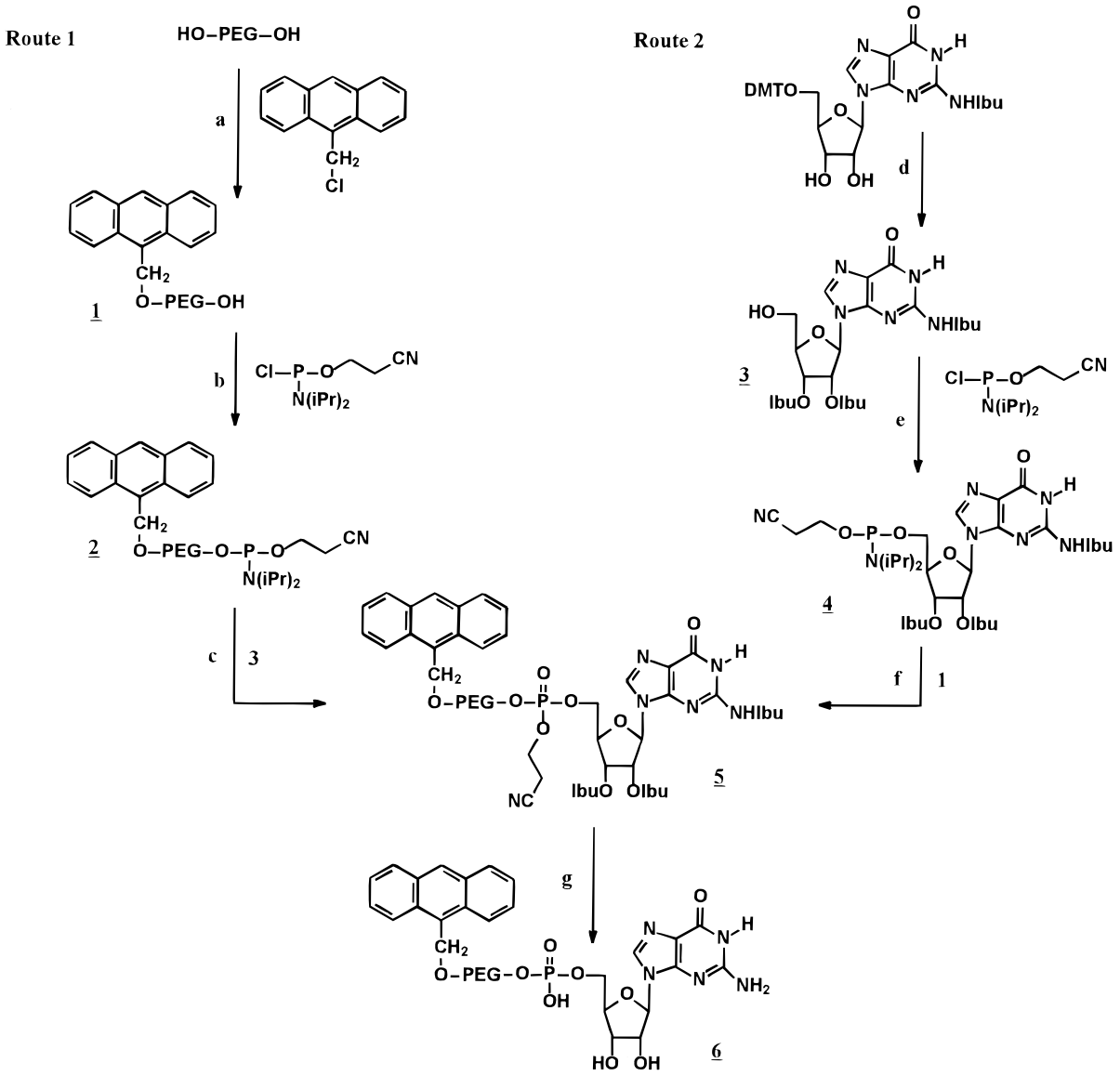
Design, Synthesis, and Characterization of the Initiator Nucleotides. The most common method to synthesize RNA in the laboratory is in vitro transcription. The DNA-dependent RNA polymerase of bacteriophage T7 recognizes and binds to specific double-stranded DNA promotor sequences and starts transcription right behind the promotor by using ribonucleotide-5′-triphosphates according to the sense strand of the DNA duplex. Synthesis stops at the end of the DNA template (runoff transcription). The first incorporated nucleotide is always GTP. The polymerase has been shown to be rather tolerant toward certain modifications at the 5′-position of this first guanosine in the initiation step (Martin and Coleman, 1989). Therefore, the use of 5′-modified guanosines as "initiator nucleotides" has been described for the incorporation of cap-analogues, nucleoside monophosphates, or thiophosphates as well as short oligonucleotides and fluorescence-labeled nucleotides at 5′-terminal positions of RNA transcripts (Burgin and Pace, 1990; Darzynkiewicz et al., 1988; Logsdon et al., 1992; Pitulle et al., 1992). These initiator nucleotides cannot be incorporated during elongation, since they are lacking the essential 5′-triphosphate.

Scheme 1. General Scheme for 5′-Conjugation by Runoff Transcription Using Ternary Initiator Nucleotides



To develop a generally applicable method to attach labels, fluorophores, cross-linkers or reactive groups to long RNA molecules without interfering with the overall RNA structure, we chose to investigate ternary conjugates in which the organic moiety is attached to the 5′-position of guanosine via a polymeric tether (Scheme 1). As the organic moiety, we chose anthracene primarily for use of the resulting conjugates in selection experiments similar to those described by Tarasow et al. (1997), i.e., for the selection of RNA molecules catalyzing Diels−Alder reactions from combinatorial libraries (Seelig and Jäschke, 1999). Additionally, anthracene turned out to be a convenient molecule to study because (i) it is a highly sensitive fluorophore and can thus be easily detected, (ii) attachment of anthracene to an RNA molecule causes significant shifts in reversed-phase HPLC, and (iii) the specific reactivity of anthracene provides an additional means to study the properties of the conjugates. We used poly(ethylene glycol) as a tether because it is highly flexible, water soluble, and its chemical properties are known to be compatible with RNA chemistry (Jäschke, 1997; Jäschke et al., 1996). To select catalysts from libraries without prior structural knowledge, the use of polydisperse PEG was thought to be beneficial because a range of different tether lengths would allow the system to find the optimum tether length itself. We designed initiator nucleotide **6**, consisting of guanosine monophosphate, esterified with the terminal PEG hydroxyl group, while the other end of the PEG chain was connected to the anthracene ring system via an ether linkage (Scheme 2).

Scheme 2. Synthesis of Initiator Nucleotides[a]



[a] (a) benzene, *n*-butyllithium; DMF; (b) CH_2Cl_2, *N*,*N*-diisopropylethylamine; (c) acetonitrile, tetrazole; iodine; (d) pyridine, isobutyric acid; CH_2Cl_2, trichloroacetic acid (e) CH_2Cl_2, *N*,*N*-diisopropylethylamine; (f) acetonitrile, tetrazole; iodine; (g) 33% aqueous ammonia.

Initiator nucleotides **6** were synthesized by two different ways (Scheme 2). In the first route, PEG 600 was converted into the alcoholate and reacted with 9-chloromethylanthracene. The resulting mixture was chromatographically fractionated, and monoanthracene-PEG **1** was reacted with (2-cyanoethyl-*N*,*N*-diisopropyl)chlorophosphoramidite to give phosphoramidite **2**. 2′-O,3′-O,N^2-Triisobutyrylguanosine **3** was prepared from 5′-*O*-dimethoxytrityl-N^2-isobutyrylguanosine and reacted with **2** in the presence of tetrazole. After coupling, the reaction mixture was oxidized and purified by column chromatography on silica gel. **5** was deprotected using aqueous ammonia giving initiator nucleotide **6**, which was fractionated by reversed-phase HPLC according to the number of ethylene glycol units into 12 product peaks.

Alternatively, 2′-O,3′-O,N^2-triisobutyrylguanosine **3** was reacted with (2-cyanoethyl-*N*,*N*-diisopropyl)chlorophosphoramidite to give a protected guanosine-5′-phosphoramidite **4** which allowed an easier chromatographic purification, compared to polydisperse **2**. After reaction with **1** and tetrazole, oxidation, deprotection, and purification, **6** was obtained. Total yields were generally higher with route 2, which was also used for the synthesis of derivatives with higher n_{PEG}, starting from PEG 1500.

As observed with oligonucleotide−PEG conjugates (Jäschke et al., 1993, 1994, 1996), the degree of polymerization (number of CH_2CH_2O units) of the coupled PEG influences the molecular properties of the conjugates. For initiator nucleotides **6**, single degrees of polymerization were eluted sequentially in reversed-phase HPLC, giving a characteristic elution pattern. Figure 1a shows the chromatogram of initiator nucleotides derived from PEG 1500. For PEG 600, similar chromatograms were observed at lower acetonitrile concentrations (Seelig and Jäschke, 1997). The UV spectra of the conjugates show the typical features of both anthracene (absorbance maximum at 254.5 nm, "fingers" at 348, 365, and 385 nm) and guanosine (shoulder at 277 nm, see Figure 1b), and no influence of n_{PEG} on the spectroscopic properties was observed. Upon UV excitation at 254 nm, the initiator nucleotides show a typical blue anthracene fluorescence with an emission maximum at 419 nm (Figure 1c). The correct composition was corroborated by MALDI mass spectrometry (Figure 1d), and the mea-

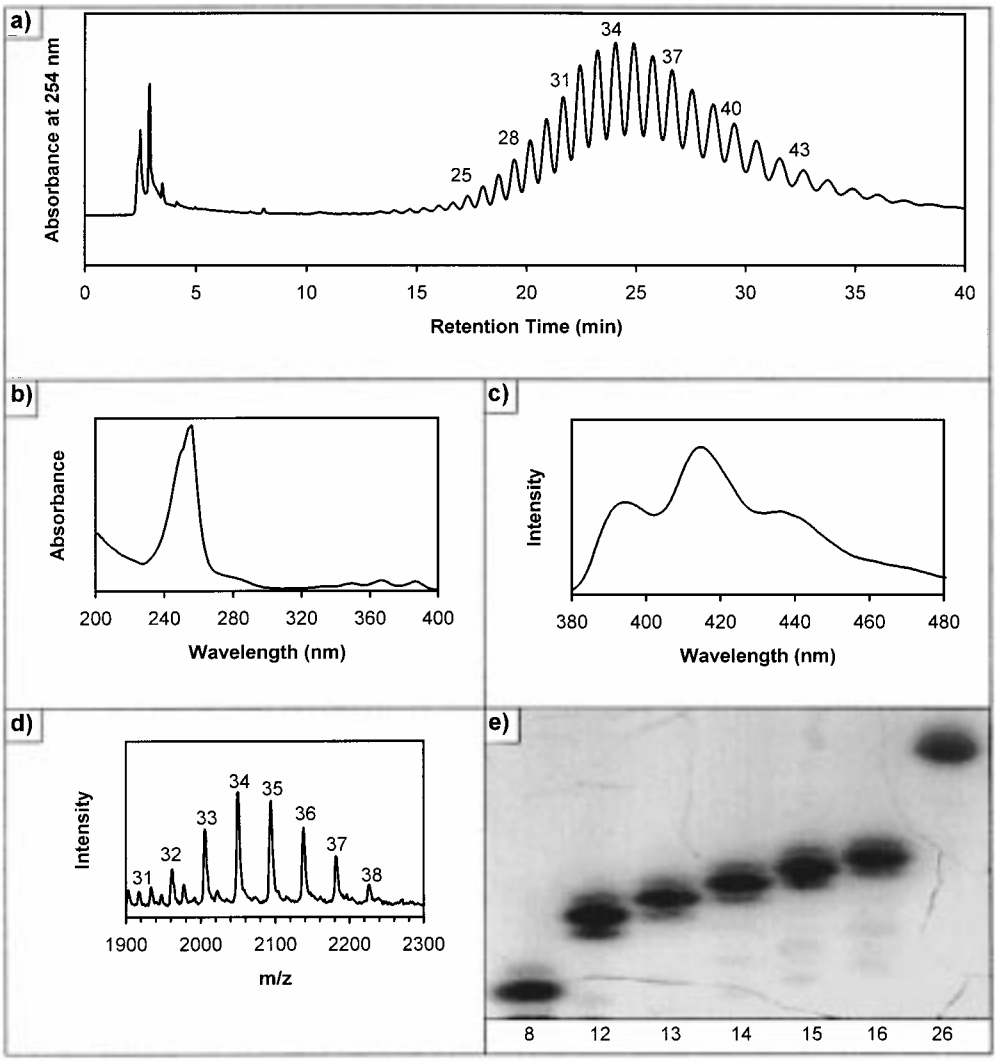


Figure 1. (a) Reversed-phase HPLC of initiator nucleotides **6** (PEG 1500). Eluent: 38% acetonitrile in 100 mM triethylammonium acetate, pH 7.0. The figures at the peaks indicate the respective degree of polymerization. (b) UV spectrum of the peak eluted at 13.2 min. (c) Fluorescence spectrum measured on excitation at 254 nm at a concentration of 1 μM. (d) MALDI mass spectrum of polydisperse **6** (PEG 1500). (e) Electrophoretic analysis of **6** with varying n_{PEG}. Two A_{260} units each of HPLC-purified **6** (n_{PEG} = 8−35) were loaded in 50% formamide and ran on a 15% denaturing polyacrylamide gel. Bands were visualized by epiillumination at 254 nm and photographed.

sured *m/z* values were in good agreement with the calculations (typically ±0.5 Da). The molecular masses of neighboring peaks differ by 44 Da, i.e., by the size of one CH_2CH_2O unit. Electrophoretic mobility decreases with increasing size of the PEG chain (Figure 1e). Generally, the electrophoretic mobility is rather low, due to the low charge-to-mass ratio. On a 15% denaturing polyacrylamide gel, **6** with n_{PEG} = 8 has about the mobility of an 32-mer oligonucleotide.

Enzymatic Synthesis and Characterization of Ternary RNA Conjugates. To investigate whether T7 RNA polymerase incorporates the synthesized initiator nucleotides, HPLC-purified monodisperse **6** with n_{PEG} = 10 was added to standard transcription reactions from a short synthetic DNA template, while the concentration of GTP was reduced accordingly (Figure 2). Since T7 RNA polymerase always generates a mixture of correct transcripts and molecules with an additional nucleotide appended to the 3′-end, RNA transcripts of 25 (*n*-mer, band 5) and 26 nucleotides (*n*+1-mer, band 4), respectively, were generated. With increasing concentration of **6,** two new bands with lower electrophoretic mobility (1 and 2) appeared in the autoradiogram (Figure 2a). Upon excitation with 254 nm UV light, these bands showed the typical blue anthracene fluorescence. (Figure 2b). The third fluorescent band found on the photograph (band 3) does not contain RNA and is a decomposition product arising from **6** during incubation. Transcription reactions using polydisperse initiator nucleotides give complex mixtures of varying electrophoretic mobility with the longer oligomers moving slower than the shorter ones. The percentage of RNA molecules being coupled to anthracene is a function of the ratio of **6** to GTP, while the overall transcription yield drops with decreasing GTP concentration (Figure 2c). Bands were excised from the gel, eluted, and analyzed by MALDI mass spectrometry. The *n*-mer conjugate band 2 was found to have a molecular weight of 8623.3 (calculated 8623.7 for $[MH^+]$), indicating that no other modifications had taken place. A typical mass spectrum for a mixture of *n*+1-mers is shown in Figure 2d (n_{PEG} = 15), and three signals were found at *m/z* =9104.6, 9126.9, and 9145.3, while the calculated values with the arbitrarily appended "*n*+1"-nucleotide were 9106 (for N=C), 9107 (U), 9130 (A), and 9146 (G).

To investigate whether T7 RNA polymerase could also be used to fuse even larger molecules to RNA, we then systematically studied the dependence of the enzymatic

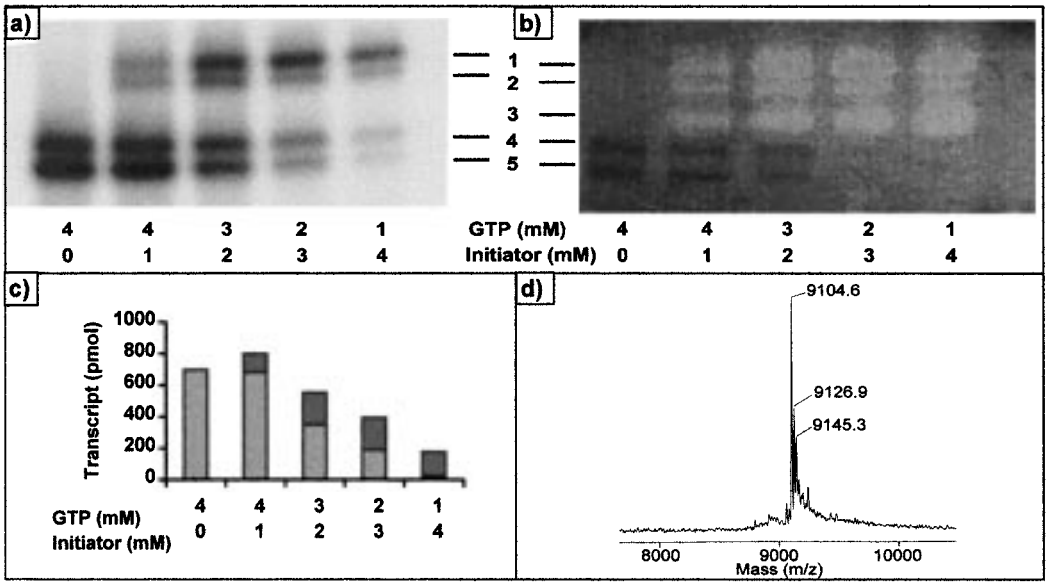


Figure 2. T7 transcriptions using initiator nucleotides **6** (n_{PEG} = 10). (a) Transcription by T7 RNA polymerase of a short DNA oligonucleotide duplex (see Materials and Methods). Autoradiogram of a 15% denaturing polyacrylamide gel. Nucleotide sequence of transcripts (bands 1,2,4,5), 5′(anthracene-PEG-)G GAG CUC AGC CUU CAC UGC UCC ACC (N)3′ (N = A, C, G or U); band 5, unmodified *n*-mer transcript; band 4, *n*+1-mer; band 2, anthracene-PEG-conjugate, *n*-mer; band 1, anthracene-PEG-conjugate; *n*+1-mer. (b) Fluorescence photograph of the same gel upon epiillumination at 254 nm. (c) Dependence of transcription yields on the composition of the transcription mixture. Light gray, unmodified RNA; dark filling, RNA conjugates. (d) MALDI mass spectrum of *n*+1-mer conjugates (equivalent to band 1 in Figure 2a) with n_{PEG} = 15. Bands were excised from the gel and eluted.

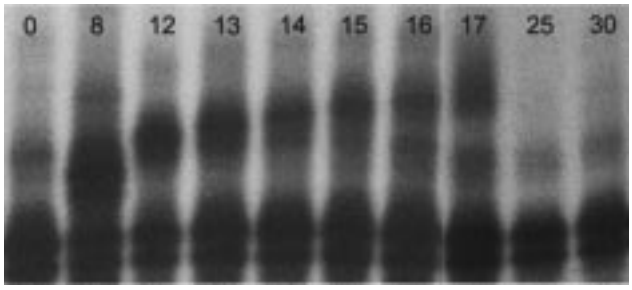


Figure 3. Electrophoretic analysis of transcripts prepared using **6** with varying n_{PEG}. Templates and conditions were identical to Figure 2. c_{GTP} was 4 mM, and $c_{Initiator}$ was 1 mM. HPLC-purified initiator nucleotide fractions were used. Bands were visualized by autoradiography.

incorporation on the length of the PEG tether. Transcription mixtures were prepared from individual HPLC fractions of initiator nucleotides derived from PEG 600 and PEG 1500, respectively. These fractions contained each at least 80% of one single degree of polymerization. To ensure sufficient yield of total transcript, GTP was kept at 4 mM, while initiators **6** were used at 1 mM. Figure 3 shows the autoradiograms of 15% denaturing polyacrylamide gels with n_{PEG} varying from 8 to 30. Up to n = 13, the initiators **6** are efficiently incorporated giving transcripts with the electrophoretic mobility decreasing with increasing n_{PEG}. From n = 14, a decline in the yield of conjugate occurs, while no significant effect on the yield of unmodified transcript is observed. From n = 25, no incorporation at all can be detected. These results demonstrate that initiator nucleotides **6** of this type can be used to introduce a poly(ethylene glycol) spacer of up to 50 atoms, while longer spacers are not incorporated.

Longer RNA transcripts up to 159 base pairs containing regions of randomized sequence for in vitro selection experiments were efficiently transcribed and conjugated

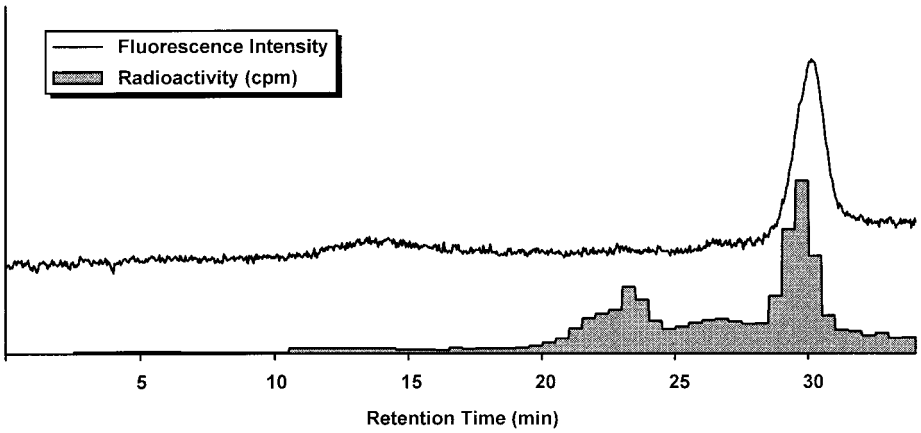


Figure 4. Reversed-phase HPLC analysis of a transcription reaction from a randomized ds DNA pool in the presence of polydisperse **6**. Nominal length of transcripts: 159 nucleotides. Separation was carried out on a C_4 column using a linear gradient of 1 to 32% acetonitrile in 40 min at a flow rate of 1 mL/min. Fluorescence detection was carried out with a Shimadzu RF 535 HPLC fluorescence detector using an excitation wavelength of 254 nm and an emission wavelength of 419 nm. Fractions of 500 μL were collected and analyzed by Cerenkov counting.

in the presence of initiator nucleotides **6** and using appropriate double-stranded DNA template molecules. A ratio of 3 mM GTP/2 mM **6** gave an average of 57% of the transcripts conjugated at their 5′-ends with only marginally lower total yield, compared with transcription in the absence of **6**. Quantitative separation, however, was not achieved under standard electrophoresis conditions due to the fact that the mass difference caused by coupling of anthracene−PEG is smaller than the mass dispersion caused by randomization of large parts of the molecules. Reversed-phase HPLC of transcript mixtures using a C4 column with fluorescence detection and Cerenkov counting of collected fractions turned out to be suitable for purification and quantification of incorporation (Figure 4). As expected, anthracene caused a significantly increased retention, and two peaks were detected containing radioactivity, with only the second showing anthracene fluorescence.

RNA-linker−anthracene conjugate libraries prepared by this method have successfully been used in selection experiments toward RNA Diels−Alderases. In short, these experiments involve incubation of the conjugate library with biotin maleimide, and biotinylated RNA species are isolated by affinity chromatography on immobilized streptavidin. These RNAs are then efficiently reverse transcribed and PCR amplified, demonstrating that conjugation does not interfere with recognition of RNA by polymerases. Repetitive rounds of conjugation, selection, and amplification were carried out, leading to the enrichment of highly active Diels−Alder catalysts (Seelig and Jäschke, 1999).

CONCLUSION

Methods for the site-specific incorporation of nonnucleotide moieties into RNA are valuable tools in RNA biochemistry and for the expansion of RNA catalysis beyond self-modification. The approach described here for coupling of anthracene−PEG derivatives allows the incorporation of a single conjugate group at the 5′-position of a transcript with high efficiency in a single step. The guanosine part of the initiator nucleotides represents the minimum substrate for T7 RNA polymerase (Martin and Coleman, 1989), and up to about $n_{PEG} = 13$, the poly(ethylene glycol) linker does not interfere with the transcription reaction. Combined with data from other publications (Martin and Coleman, 1989; Pitulle et al., 1992), the results indicate that a poly(ethylene glycol) linker attached to the 5′-phosphate of GMP could be used to incorporate various other molecules into RNA, too. The chemical synthesis of the described initiator nucleotides **6** is straightforward and can be easily carried out on a multigramm scale, thereby allowing the production of quantities sufficient for numerous preparative transcription reactions.

Using polydisperse poly(ethylene glycols), libraries of initiator nucleotides were synthesized differing only in the number of ethylene glycol units, thereby spanning a distance from about 20 to 60 Å from the 5′-end of the RNA to the attached moiety. This could be useful for finding the optimum length of the linker for applications were distances are critical (Cload et al., 1993) or as a mobility modifier for coding purposes (Baron et al., 1996; Grossman et al., 1994).

Finally, terminal coupling of PEG has been reported to increase the resistance of oligonucleotides to degradation in biological media and also to alter the pharmacokinetic properties, so enzymatic PEGylation strategies might provide an elegant alternative to the chemical methods described (Jäschke et al., 1994; Kawaguchi et al., 1995; Manoharan et al., 1995).

ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft (Grants Ja 794/1 and SFB 344, TP B10 to A.J.). We acknowledge additional support by Prof. V. A. Erdmann, and spectrometric measurements performed by Drs. P. Franke, A. Schäfer and E. Nordhoff.

LITERATURE CITED

Agrawal, S., Ed. (1994) *Protocols for oligonucleotide conjugates*, Humana Press, Totowa, NJ.

Baron, H., Fung, S., Aydin, A., Bähring, S., Luft, F. C., and Schuster, H. (1996) Oligonucleotide ligation assay (OLA) for the diagnosis of familial hypercholesterolemia. *Nat. Biotechnol. 14*, 1279−82.

Beaucage, S. L., and Iyer, R. P. (1993) The functionalization of oligonucleotides via phosphoramidite derivatives. *Tetrahedron 49*, 1925−1963.

Bevilacqua, P. C., Li, Y., and Turner, D. H. (1994) Fluorescence-detected stopped flow with a pyrene labeled substrate reveals that guanosine facilitates docking of the 5′ cleavage site into a high free energy binding mode in the Tetrahymena ribozyme. *Biochemistry 33*, 11340−11348.

Burgin, A. B., and Pace, N. R. (1990) Mapping the active site of ribonuclease P RNA using a substrate containing a photoaffinity agent. *EMBO J. 9*, 4111−4118.

Cload, S. T., Richardson, P. L., Huang, Y. H., and Schepartz, A. (1993) Kinetic and thermodynamic analysis of RNA binding by tethered oligonucleotide probes: alternative structures and conformational changes. *J. Am. Chem. Soc. 115*, 5005−5014.

Darzynkiewicz, E., Stepinsky, J., Ekiel, I., Jin, Y., Haber, D., Sijuwade, T., and Tahara, S. M. (1988) b-globin mRNAs capped with m7G, m2,2,7G or m2,3,2,7G differ in intrinsic translation efficiency. *Nucleic Acids Res. 16*, 8953−8962.

Goodchild, J. (1990) Conjugates of oligonucleotides and modified oligonucleotides: a review of their synthesis and properties. *Bioconjugate Chem. 1*, 165−187.

Grossman, P. D., Bloch, W., Brinson, E., Chang, C. C., Eggerding, F. A., Fung, S., Iovannisci, D. A., Woo, S., and Winn-Deen, E. S. (1994) High-density multiplex detection of nucleic acid sequences: oligonucleotide ligation assay and sequence-coded separation. *Nucleic Acids Res. 22*, 4527−4534.

Hausch, F., and Jäschke, A. (1997) Libraries of multifunctional RNA conjugates for the selection of new RNA catalysts. *Bioconjugate Chem. 8*, 885−890.

Igloi, G. (1996) Nonradioactive Labeling of RNA. *Anal. Biochem. 233*, 124−129.

Jäschke, A. (1997) Oligonucleotide-poly(ethylene glycol) conjugates−synthesis, properties and applications. In *Chemistry and Biological Applications of PEG* (J. M. Harris, and S. Zalipsky, Eds.) Vol. 680, pp 265−283, ACS Symp. Ser. Washington DC.

Jäschke, A., Bald, R., Nordhoff, E., Hillenkamp, F., Cech, D., Erdmann, V. A., and Fürste, J. P. (1996) Synthesis and analytical characterization of RNA-poly(ethylene glycol) conjugates. *Nucleosides Nucleotides 15*, 1519−1529.

Jäschke, A., Fürste, J. P., Cech, D., and Erdmann, V. A. (1993) Automated incorporation of poly(ethylene glycol) into synthetic oligonucleotides. *Tetrahedron Lett. 34*, 301−304.

Jäschke, A., Fürste, J. P., Nordhoff, E., Hillenkamp, F., Cech, D., and Erdmann, V. A. (1994) Synthesis and properties of oligodeoxyribonucleotide- poly(ethylene glycol) conjugates. *Nucleic Acids Res. 22*, 4810−4817.

Kawaguchi, T., Asakawa, H., Tashiro, Y., Juni, K., and Sueishi, T. (1995) Stability, specific binding activity, and plasma concentration in mice of an oligodeoxynucleotide modified at 5′-terminal with poly(ethylene glycol). *Biol. Pharm. Bull. 18*, 474−476.

Logsdon, N., Lee, C. G., and Harper, J. W. (1992) Selective 5′ modification of T7 RNA polymerase transcripts. *Anal. Biochem. 205*, 36−41.

Manoharan, M., Tivel, K. L., Andrade, L. K., Mohan, V., Condon, T. P., Bennet, C. F., and Cook, P. D. (1995) Oligonucleotide conjugates—alteration of the pharmacokinetic properties of antisense agents. *Nucleosides Nucleotides 14*, 969−973.

Martin, C. T., and Coleman, J. E. (1989) T7 RNA polymerase does not interact with the 5′-phosphate of the initiating nucleotide. *Biochemistry 28*, 2760−2762.

Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res. 15*, 8783−8798.

Moore, M. J., and Sharp, P. A. (1992) Site-specific modification of pre-mRNA: the 2′-hydroxyl groups at the splice sites. *Science 256*, 992−997.

Oh, B. K., and Pace, N. R. (1994) Interaction of the 3′-end of tRNA with ribonuclease P RNA. *Nucleic Acids Res. 22*, 4087−4094.

Pitulle, C., Kleineidam, R. G., Sproat, B., and Krupp, G. (1992) Initiator oligonucleotides for the combination of chemical and enzymatic RNA synthesis. *Gene 112*, 101−105.

Seelig, B., and Jäschke, A. (1997) Site-specific modification of enzymatically synthesized RNA: transcription initiation and Diels−Alder reaction. *Tetrahedron Lett. 38*, 7729−7732.

Seelig, B., and Jäschke, A. (1999) A small catalytic RNA motif with Diels−Alderase activity. *Chem. Biol. 6*, 167−176.

Tarasow, T. M., Tarasow, S. L., and Eaton, B. E. (1997) RNA-catalysed carbon−carbon bond formation. *Nature 389*, 54−57.

Tong, G., Lawlor, J. M., Tregear, G. W., and Haralambidis, J. (1993) The synthesis of oligonucleotide−polyamide conjugate molecules suitable as PCR primers. *J. Org. Chem. 58*, 2223−2231.

Wiegand, T. W., Janssen, R. C., and Eaton, B. E. (1997) Selection of RNA amide synthases. *Chem. Biol. 4*, 675−683.

BC980085H