Universal initiator nucleotides for the enzymatic synthesis of 5'-amino- and 5'-thiol-modified RNA

Jörg C. Schlatterer 1, Andres Jäschke *

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

Received 30 March 2006
Available online 19 April 2006

Abstract

We report the chemical synthesis of 5'-amino- and 5'-thiol-hexaethylene glycol guanosine nucleotides and their enzymatic incorporation into RNA, followed by chemical modifications at their nucleophilic ends. By using two similar routes, the conjugates of guanosine-5'-monophosphate and hexaethylene glycol with attached reactive groups (—SH or —NH2) were synthesized using phosphoramidite chemistry, and characterized by MALDI TOF mass spectrometry. These initiator molecules were efficiently incorporated into RNA at the 5'-end by run-off transcription using T7 RNA polymerase. The potential of these RNA conjugates for a broad reaction range with electrophiles is shown here, thereby enabling their use for diverse biochemical applications.

Keywords: Transcription; Conjugation; Ribozymes; In vitro selection

RNA molecules play an important role in cellular processes. They are transmitters of information, regulators, and catalysts, and they hold great potential for various diagnostic and biochemical applications. Site-specific substitution and derivatization provide powerful tools for understanding structure and function of RNA [1–5], and conjugation strategies allow the incorporation of various reporters or functional elements, like fluorophores, crosslinkers, chelators, signal peptides, and localization signals. While short oligoribonucleotides can be easily synthesized and conjugated by chemical means, the conjugation of longer and enzymatically synthesized RNAs is much more complicated. The most common strategies for the site-specific modification of enzymatically synthesized RNA are: (i) modification of the 5'-terminus [6], (ii) modification of the 3'-terminus [7], and (iii) more complex ligation strategies involving both transcripts and chemically synthesized modified oligonucleotides [8]. While in chemical oligonucleotide synthesis a number of standard building blocks are widely used for the incorporation of reactive amino and thiol groups that can be utilized for site-specific postsynthetic derivatization, such general tools are generally not available for enzymatic synthesis.

Our prime motivation for developing new conjugation strategies comes from RNA catalysis. While the naturally occurring ribozymes catalyze only a narrow spectrum of chemical reactions, artificial ribozymes isolated from synthetic combinatorial RNA libraries have been shown to accelerate a broad range of reactions, from amide bond formation to ester hydrolysis to redox chemistry to C—C bond formation [9,10]. A key step in the iterative in vitro selection procedures used to isolate such catalytic RNAs is the conjugation of a potential reactant to the RNA molecules, preferably via a flexible polymeric tether. These species are then reacted with a second reactant carrying an affinity tag, e.g., a biotin group. RNA molecules that accelerate the reaction of the attached first reactant with the tagged second reactant will carry the tag and can then be isolated by affinity chromatography. The isolated RNAs are then copied and amplified by reverse transcription and PCR, and used as input for the next selection cycle.
which starts with the synthesis of the RNA conjugates. This process requires the RNA to be synthesized enzymatically, thus limiting the choices for site-specific modification.

Our laboratory has previously described different strategies for the solution of this problem [7,11–14], as well as the successful application to the selection of Diels–Alder ribozymes [15–20]. The most convenient approach utilizes initiator nucleotides and takes advantage of the well-known fact that RNA polymerase accepts 5′-modified guanosine derivatives instead of GTP during the initiation, but not during the elongation steps [6,14,21]. So far, however, complicated multi-step syntheses were carried out for the preparation of the initiator nucleotide in each project.

To create new universal tools for 5-labeling and modification of RNAs, we developed the synthesis of a 5′-amino- and a 5′-thiol-hexaethylene glycol guanosine. We show the acceptance of the 5′-amino- and a 5′-thiol-hexaethylene glycol guanosine derivatives by T7 RNA polymerase, their incorporation into RNA, as well as their successful chemical modification.

Materials and methods

General procedures. Hexaethylene glycol (HEG) was purchased from Fluka, (2-cyanoethyl-N,N-diisopropyl)chlorophosphoramidite, sulfo-NHS-biotin, and biotin maleimide from Sigma-Aldrich; [2-13C]CTP from Amersham Biosciences, NTPs from Fermentas, and T7 RNA polymerase from Stratagene; all other chemicals, buffers, solvents, and reagents obtained were of the highest commercial grade available and used without further purification. Reactions were carried out at room temperature unless stated otherwise. 2′-O,3′-O,N2′-trisobutylxynogausine-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite 3 was synthesized as described [14]. All compounds containing hexaethylene glycol residues were detected on silica gel TLC plates by Dragendorff’s reagent. Silica gel (0.063–0.2 mm) for column chromatography was obtained from J.T. Baker.

Mono-bromo-hexaethylene glycol 1. Hexaethylene glycol (2.96 ml, 11.6 mmol) was dissolved in dry toluene (30 ml). Triethylamine (0.97 ml) was added under inert conditions. Triionylbromide (0.45 ml, 5.8 mmol) in 1.7 ml dry toluene was added dropwise within 1 h. After 1 h of reflux, the reaction mixture was stirred for 45 h at room temperature. The pellet was removed by filtering through Celite 545, the solution was tempered at reaction mixture was stirred for 45 h at room temperature. The pellet was removed by filtering through Celite 545, the solution was tempered at 50°C with 5 g active carbon and filtered through Celite 545 again. The solvent was removed by rotary evaporation, and the resulting oil subjected to silica gel chromatography (MeOH/CH2Cl2: 1.9 v/v). Concentration in vacuum afforded mono-bromo-hexaethylene glycol 1 with a yield of 30%, referred to HEG.


Tritylmercapto-hexaethylene glycol 2. Mono-bromo-hexaethylene glycol (0.5 g, 1.45 mmol) 1 was dissolved in 12 ml ethanol. NaOH (0.146 ml, 10 M) and 0.4 g (1.45 mmol) triphenylmethylcarbene were added and stirred for 15 h. After cooling the reaction with an ice/water bath, the solution was filtered using Celite 545. Following evaporation to dryness, product 2 was chromatographically purified by silica gel (MeOH/CH2Cl2: 1:9 v/v). The yield of tritylmercapto-hexaethylene glycol 2 was 93% with respect to 1.


Protected initiator-(S)-nucleotide 4. All glassware was carefully dried before use. 0.64 g (1.18 mmol) tritylmercapto-hexaethylene glycol 2 and 0.69 equivalents (0.82 mmol) of the trisobutylxynogausine-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite 3 were co-distilled with pyridine twice and dried over P2O5 for 12 h. The substrates were dissolved in dry acetone (10 ml). In an inert atmosphere, 8.5 ml tetrazole in acetonitrile (0.47 M, 4 mmol) was added. After stirring for 15 min, 12 ml of an iodine solution (1.2 mmol, 0.1 M in tetrahydrofuran/collidine/water 2:2:1) was added and stirred for 5 min. The solution was quenched by a sodium-thiosulfate (1 M) until the brown color of iodine disappeared. The crude product was diluted by 70 ml dichloromethane and extracted by saturated NaHCO3 and NaCl solutions before the organic phase was dried by Na2SO4. The solvents were removed under reduced pressure, the crude product was purified by silica gel chromatography (MeOH/CH2Cl2: 95:5 v/v) and was deprotected ammonically (33% aqueous ammonia) at 55°C overnight. HPLC purification and vacuum concentration yielded to ~78% of the protected initiator-(S)-nucleotide 4, referred to 3.

MALDI+-MS: calculated for C14H25N2O6PS [M+H]+ = 585.9, found 588.9.

Initiator-(S)-nucleotide 5. Protected initiator-(S)-nucleotide (11.7 mmol) 4 in aqueous solution (117 mM) was mixed with 25 ml of 1 M AgNO3 solution. After incubating for 1 h and shaking, 50 µl of 0.5 M DTT was added, the solution was spun down, and the supernatant extracted with 500 µl ethylacetate. Purification was accomplished by HPLC and led to a yield of ~90% of initiator-(S)-nucleotide 5 referred to 4 after drying in vacuum.

MALDI+-MS: calculated for C14H25N2O6PS [M+H]+ = 643.2, found 644.

Phthalimido-hexaethylene glycol 6. Mono-bromo-hexaethylene glycol (1.22 g, 3.53 mmol) 1 was dissolved in 22 ml dry dimethylformamide. Potassium phthalimide (0.82 g, 4.4 mmol) was added. After 28 h of reflux, the reaction mixture cooled down slowly and was diluted by 30 ml CH2Cl2. Extraction with water, saturated NaHCO3, and NaCl solutions followed, and the organic phase was dried by Na2SO4. After evaporation to dryness, product 6 was chromatographically purified on silica gel (MeOH/CH2Cl2: 1.9 v/v). The yield of phthalimido-hexaethylene glycol 6 was 75%, referred to 1.

1H NMR (250 MHz, [D1]CHCl3, 25°C): δ = 3.45–3.75 (m, 22H, CH2-O), 3.85 (t, 2H, CH2-N), 7.65 (t, 2H, Ar-H), 7.75 (d, 2H, Ar-H). 13C NMR (63 MHz, [D1]CHCl3, 25°C): δ = 38.5, 61.0, 68.1, 70.9, 73.1, 124.0, 132.0. FAB+-MS: calculated for C14H25N2O6 [M+H]+ = 411.45, found 412.

Initiator-(N)-nucleotide 7. Phthalimido-hexaethylene glycol (0.3 g, 0.73 mmol) 6 and 0.69 equivalents (0.44 mmol) of the trisobutylxynogausine-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite 3 were co-distilled with pyridine twice and dried over P2O5 for 12 h. The substrates were dissolved in 10 ml dry acetonitrile.

In an inert atmosphere, 9 ml tetrazole in acetonitrile (0.47 M, 4 mmol) was added. After stirring for 15 min, 8 ml of an iodine solution (0.8 mmol, 0.1 M in tetrahydrofuran/collidine/water 2:2:1) was added and stirred for 5 min. The solution was quenched by a sodium-thiosulfate (1 M) until the brown color of iodine disappeared. The crude product was diluted by 50 ml dichloromethane and extracted by saturated NaHCO3 and NaCl solutions before the organic phase was dried by Na2SO4. The solvents were removed under reduced pressure, the crude product purified by silica gel chromatography (MeOH/CH2Cl2: 95:5 v/v) and deprotected ammonically (33% aqueous ammonia) at 55°C overnight. HPLC purification and vacuum concentration yielded to ~85% of the protected initiator-(N)-nucleotide 7 referred to 3.

MALDI+-MS: calculated for C14H25N2O6P [M+H]+ = 626.5, found 627.

T7 transcriptions. T7 transcriptions were performed essentially as described [22]. The DNA templates for T7 transcriptions were purchased...
Universally modifiable RNAs are useful tools for the investigation of the catalytic potential of RNA, and transcription initiation offers various advantages over other synthetic strategies. Most importantly, it is a one-step process that occurs during enzymatic polymerization and does not require additional handling or purification steps. For in vitro selection, the use of tethered (or linker-coupled) reactants has been most successful. In this work, a hexaethylene glycol tether was used. It is a flexible molecule, so the reactant can move and position itself with only little restrictions. The extended length of a hexaethylene glycol tether is about 21 Å, allowing the attached reactant to move within a rather large volume. Hexaethylene glycol does not form higher-order structures and is unlikely to influence the tertiary structure of the RNA. The molecule is stable and features a high solubility in water.

For the preparation of universal building blocks, we chose amino- and thiol groups as the reactive groups to be attached. Both are reactive nucleophiles that can be converted in numerous standard reactions (e.g., Michael addition, NHS-ester chemistry). The thiol group could furthermore be derivatized by redox chemistry.

Synthesis and characterization of initiator-(S) and initiator-(N)-nucleotides

The initiator-(S)- and initiator-(N)-nucleotides 5 and 7 were synthesized by coupling the protected guanosine phosphoramidite 3 with the respective mono-functionalized and protected amino- and thio-hexaethylene glycols 2 and 6, respectively (Scheme 1).

First, hexaethylene glycol was mono-brominated using thionyl bromide, yielding 1. For the initiator-(S)-nucleotide synthesis, nucleophilic substitution of 1 with triphenylmercaptan led to tritylmercapto-hexaethylene glycol 2. Triisobutyl guanosine phosphoramidite 3 [14] was coupled with 2 using tetrazole as activator, and the phosphate triester oxidized by aqueous iodine solution. Ammonia deprotection of the isobutyl-protected hydroxyl and amino functions of the guanosine and the subsequent HPLC purification led to the tritylated initiator-(S)-nucleotide 4. Finally, compound 4 was treated with silver nitrate solution and afterwards with DTT [23]. Reaction product 5 was purified by reversed-phase chromatography, and the identity confirmed by MALDI TOF mass spectrometry. The appearance of dimers was observed, but the monomers could be re-obtained by treatment with DTT.

Initiator-(N)-nucleotide 7 was synthesized by Gabriel synthesis. Compound 1 was reacted with potassium phthalimide, and product 6 reacted with guanosine derivative 3 in the presence of tetrazole. Iodine oxidation and deprotection led to initiator-(N)-nucleotide 7 which was purified by HPLC. MALDI TOF mass spectrometry confirmed the structure of 7.

Thiol-reactive biotin conjugation and attachment to neutravidin

Initiator-(S)-nucleotide 5 was incorporated into RNA by T7 transcription using dsDNA templates, according to standard protocols. Desalted 32P-labelled RNA-conjugate/RNA transcription mixtures (10 nucleotide transcript length) were analyzed on an 18% denaturing polyacrylamide gel. The autoradiogram (Fig. 1A) indicates a concentration-dependent appearance of a band with lower mobility (b), consistent with the incorporation of the initiator-(S)-nucleotide. The maximal incorporation yield of ~60% was obtained by using 4 mM initiator-(S)-nucleotide and 1 mM GTP (lane 3).
To obtain a more direct proof for the presence of the thiol group in conjugates synthesized by the above procedure, various transcription mixtures were reacted with the thiol-reactive reagent biotin maleimide. Fig. 1B shows a 25 nt transcript/conjugate mixture. The presence of the biotin moiety could be confirmed by addition of the biotin-binding protein neutravidin (an engineered version of avidin), inducing the disappearance of the RNA-biotin conjugate band d and the appearance of a new band e with almost no electrophoretic mobility, consistent with neutravidin binding. Other bands corresponding to unbiotinylated RNA remained unaffected by the neutravidin treatment.

Scheme 1. Synthesis of initiator nucleotides. (a) Toluene, triethylamine, and thionylbromide; (b) toluene, triphenylmercaptane; (c) acetonitrile, triisobutylguanosine-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite 3, tetrazole, iodine, and 33% aqueous ammonia; (d) water, AgNO₃; (e) DMF, potassium phthalimide; (f) acetonitrile, triisobutylguanosine-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite 3, tetrazole, iodine, and 33% aqueous ammonia.
Amine-reactive biotin conjugation and attachment to neutravidin

The incorporation of initiator-(N)-nucleotide 7 could also be demonstrated by concentration-dependent in vitro transcription studies. Transcriptions were carried out using three different initiator 7/GTP (1 mM) ratios. The autoradiogram of an 8% denaturing polyacrylamide gel demonstrates the effective incorporation of the initiator (Fig. 2A) into a 10-mer transcript. Lane 1 shows the RNA synthesis without 7, lane 2 the transcription with 2 mM initiator, and lane 3 that with 4 mM initiator-(N)-nucleotide. The maximal incorporation of ~65% was accomplished by using 4 mM initiator-(N)-nucleotide and 1 mM GTP.

The ability of the N-conjugate to form RNA derivatives by NHS-ester chemistry is shown in Fig. 2B. Lane 4 represents the transcription with a 4:1 initiator:GTP ratio, leading to the appearance of two bands, assigned to the transcript (c) and the conjugate (d). Reaction with an excess of biotin-sulfo-NHS ester (lane 5) yielded only a small mobility shift in this case. Addition of neutravidin, however, caused complete disappearance of the conjugate band d and the formation of a high-molecular weight complex (band e), while the band corresponding to the unmodified transcript remained unaffected.

Discussion

In this work, we have described the synthesis, enzymatic incorporation, and derivatization of initiator nucleotides containing tethered primary amino and thiol groups, respectively. This allows for a straightforward access to enzymatically synthesized RNA conjugates in which a variety of different residues can be attached to the 5’-end of RNA via a tether. Tethering may be advantageous for applications in which a direct coupling causes problems, e.g., quenching of a fluorophore by adjacent RNA nucleotides, or folding problems with RNA immobilized directly to a matrix. The homogeneity of the monodisperse hexamethylene glycol tether will prevent band broadening in chromatography and electrophoresis, simplify workup and purification, and also prevent mechanistic ambiguities caused by the presence of different tether lengths in a sample [20].

Zhang et al. developed 5’-sulphhydryl initiators which contain two to four ethylene glycol units as a tether. The successful selections of Diels–Alderase and amide synthase ribozymes [16,24,25], however, have used much longer tethers to ensure that tethered reactants can actually reach the catalytic center. Hexameric ethylene glycol with an
extended tether length of 21 Å appears to be a good compromise between length and synthetic simplicity. A decacethylene glycol-tethered amino-initiator was recently synthesized in our laboratory using a different synthetic strategy, and may provide an alternative if longer spacers are needed [11].

The described initiators 5 and 7 may act as multifunctional tools in RNA chemistry. On the one hand, the transcription products of these initiators can be used directly for the in vitro selection of ribozymes that catalyze the reaction between the tethered functional group and another substrate. For example, the 5'-thiol-modified ribonucleic acids could be used for selecting ribozymes that catalyze Michael or redox reactions, while 5'-amino-modified RNA may be used to isolate amide synthases and enzymes for general nucleophilic substitutions. These reaction types are all quite attractive in the context of the RNA world hypothesis.

On the other hand, a broad range of compounds are commercially available in the form of active esters, maleimides, acyl halogenides, and disulfides that can be easily attached to primary amines and thiols, respectively, in aqueous solution, and without the need for chemistry equipment and expertise. This opens the door to attach virtually anything to the RNA in a site-specific manner, and to search for catalysts for an extensive range of chemical reactions.

Finally, such conjugates may be of interest for investigations of biologically active RNA molecules. For example, the prokaryotic and eukaryotic splicing machinery as well as different RNA-protein interactions and their dynamics [26–28] can be studied by coupling N- and S-reactive chromophores, fluorophores, and cross-linkers to amino- and thiol-RNA conjugates.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (Grant No. Ja 794/3) and the Bundesministe- rium für Bildung und Forschung (Biofuture, Grant No. BEO 0311861).

References