Synthesis of a bifunctional cytidine derivative and its conjugation to RNA for in vitro selection of a cytidine deaminase ribozyme

Nico Rublack and Sabine Müller[*]

Abstract

Over the past 20 years, the generation of functional RNAs by in vitro selection has become a standard technique. Apart from aptamers for simple binding of defined ligands, also RNAs for catalysis of chemical reactions have been selected. In the latter case, a key step often is the conjugation of one of the two reactants to the library, requiring suitable strategies for terminal or internal RNA functionalization. With the aim of selecting a ribozyme for deamination of cytidine, we have set up a selection scheme involving the attachment of the cytidine acting as deamination substrate to the 3'-terminus of the RNAs in the library, and library immobilization. Here, we report the synthesis of a bifunctional cytidine derivative suitable for conjugation to RNA and linkage of the conjugated library to a streptavidine-coated surface. Successful conjugation of the cytidine derivative to the 3'-terminus of a model RNA is demonstrated.

Introduction

Since the discovery of the first catalytic RNA in the ciliate *Tetrahymena thermophilia* in 1982 [1], a number of naturally occurring ribozymes have been described [2]. Whereas all of these natural ribozymes accelerate transesterifications or, as in the case of the ribosome, the peptide bond formation, artificial ribonucleic acids developed by in vitro selection have been shown to catalyze a wide variety of organic chemical reactions [3-5]. Moreover, several of these developed ribozymes promote reactions of metabolic relevance in modern life organisms.

Impressive examples are RNA catalysts that support an aldol reaction between an aldehyde and a ketone, relevant to the synthesis of sugars [6], or the linkage of ribose to nucleobases to generate nucleotides [7]. Such ribozymes are seen as important functional entities underscoring the RNA world theory, where ribonucleic acids are suggested acting as the carrier of genetic information as well as functional players enabling RNA amplification and processing [8]. Even though a number of ribozymes supporting the RNA world theory are known, there is a wide

field of protein-catalyzed reactions in present-day organisms, for which no ribozyme analogue has yet been found. This applies for example to the transformation of cytidine to uridine, which is a well-known RNA editing event in modern cellular chemistry [9]. This process is catalyzed by cytidine deaminases (CDA, EC 3.5.4) belonging to a family of enzymes found in pro- and eukaryotes [10]. In the active center of all CDAs, a zinc ion is responsible for the activation of a water molecule that acts as nucleophile attacking the C4 carbon center of the cytosine residue and thus facilitating deamination [11]. The development of a ribozyme supporting the same kind of reaction would be a valuable addition to the repertoire of RNA activities with relevance to the RNA world theory.

A very useful and often applied technique for the generation of catalytic RNA structures is a variation of the classical SELEX approach (Systematic Evolution of Ligands by EXponential Enrichment) [12,13]. A typical procedure for the selection of a ribozyme, which enhances the reaction between two substrates, involves the conjunction of either of them with the members of an RNA-library. This is mostly achieved by transcription priming [14], which however, allows attachment of a specific reactant merely to the 5'-terminus of library RNAs [15-17]. Alternatively, post-transcriptional protocols can address both

the 5'- and the 3'-terminus. In general, post-transcriptional 5'-modification may be achieved by thiophosphorylation with T4-polynucleotide kinase followed by derivatization of the introduced terminal thiophosphate [18,19], or by chemical conversion of the 5'-terminal primary OH group into an amine or azide to be used for further conjugation with NHS-esters [20] or with alkynes [21]. Alternatively, natural and modified nucleosides can be attached to the 3'-terminus by the use of enzymes like poly(U)-polymerase (PUP), poly(A)-polymerase (PAP) or the terminal deoxynucleotidyl transferase (TdT) [22-25]. However, since most enzymatic techniques achieve only moderate yields, chemical strategies may be advantageous. Chemical 3'-end modification uses the unique properties of the RNA's 3'-terminal *cis*-diol, which can be specifically oxidized with metaperiodate [26-28], followed by reaction of the produced dialdehyde with amines or hydrazines linked to a desired functional entity [29].

We set out to select from a random library catalytically active RNAs that support deamination of cytidine to uridine. For this purpose, a bifunctionalized cytidine derivative (Figure 1) was synthesized. Via its 5'-OH group, the cytidine derivative is linked to a hexaethylene glycol tether bearing a primary amino group. At the C4-position of the base, a short linker connected
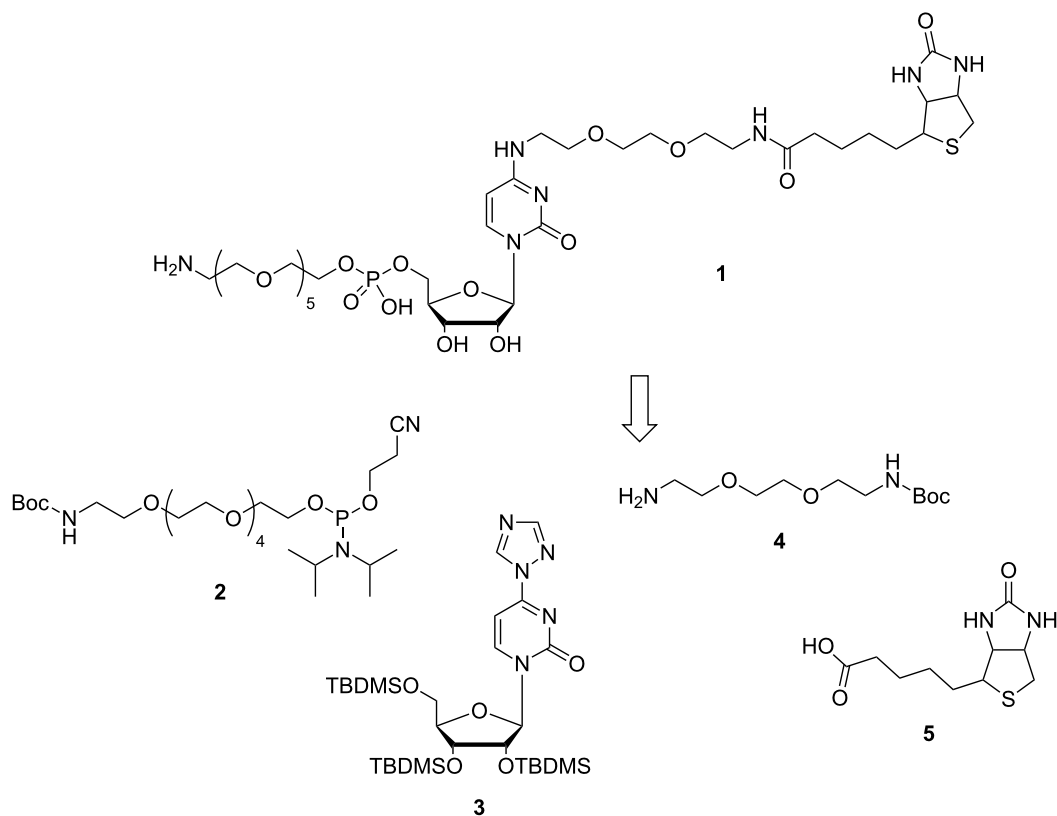

Figure 1: Retrosynthetic analysis of the bifunctional cytidine derivative **1** for functionalization of a periodate-oxidized RNA library.
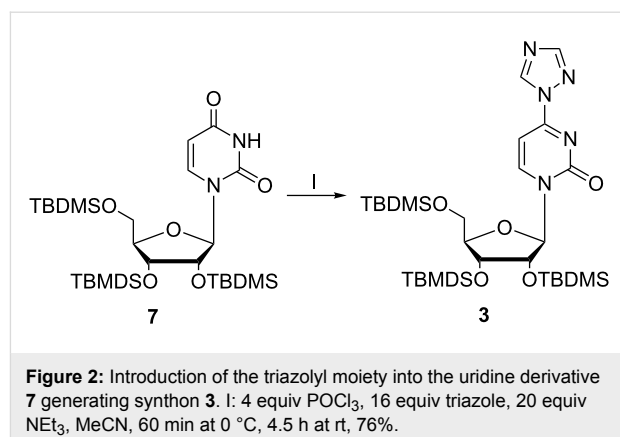
to a biotin was introduced. Upon periodate oxidation of the 3'-*cis* diol of the RNA library molecules, the produced dialdehyde is thought to react with the primary amino group of the cytidine derivative allowing immobilization of the resulting modified RNAs onto a solid surface through biotin–streptavidine interaction. All RNA sequences that can catalyze the desired deamination reaction of cytidine to uridine will be cleaved off and thus released into solution. Upon recovery, beneficial variants can be subjected to the next round of selection. Here we describe the synthetic route to the bifunctionalized cytidine derivative and its successful conjugation to RNA.

Results and Discussion

Synthesis of the bifunctional cytidine derivative **1** started from uridine making use of four synthons: a protected hexaethylene glycol linker phosphoramidite **2** bearing a primary amine to be used later for RNA functionalization, protected and suitably activated uridine **3**, and a short mono protected diamino linker **4** for the attachment of a biotin moiety **5** in order to immobilize the conjugated RNA molecules onto a streptavidine-coated surface (Figure 1).

First, the sugar hydroxy functions of uridine (**6**) were fully protected with *tert*-butyldimethylsilyl (TBDMS) groups and the tris-silylated uridine **7** [30] was further reacted with $POCl_3$, triazole and triethylamine (Figure 2).
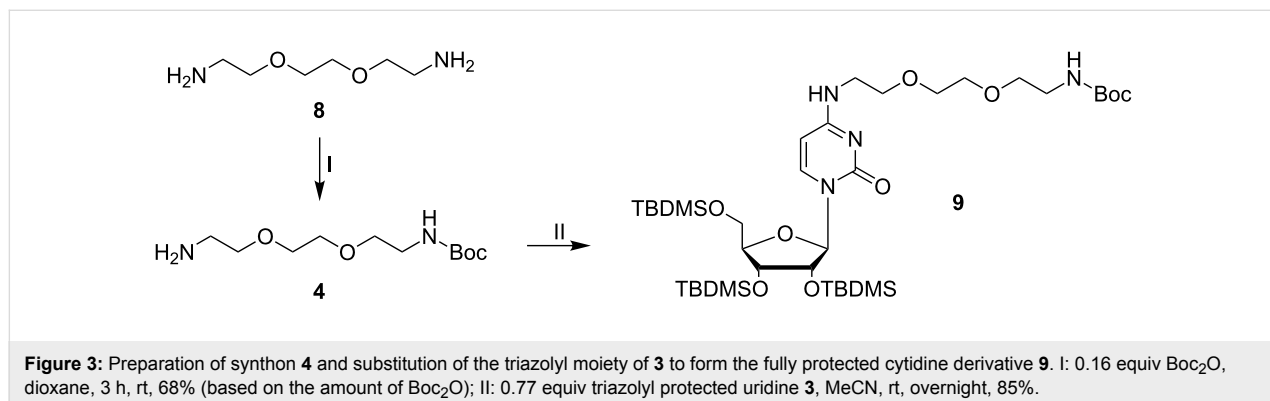
Among the procedures described for conversion of uridine into cytidine derivatives [31-34] we have chosen the path via a triazolyl activated uridine [31], because the triazolyl derivatives are known being stable enough to allow convenient handling and isolation, and they have shown very good reactivity with aliphatic amines. Synthon **3** could be isolated by recrystallization in 76% yield. Next, we introduced a short linker to the N^4-position of **3** in order to attach a biotin unit to the modified nucleoside. Linker **4** was synthesized from 2,2'-(ethylenedioxy)diethylamine (**8**) with di-*tert*-butyl dicarbonate (Boc_2O) in dioxane [35] (Figure 3).

Figure 2: Introduction of the triazolyl moiety into the uridine derivative **7** generating synthon **3**. I: 4 equiv $POCl_3$, 16 equiv triazole, 20 equiv NEt_3, MeCN, 60 min at 0 °C, 4.5 h at rt, 76%.
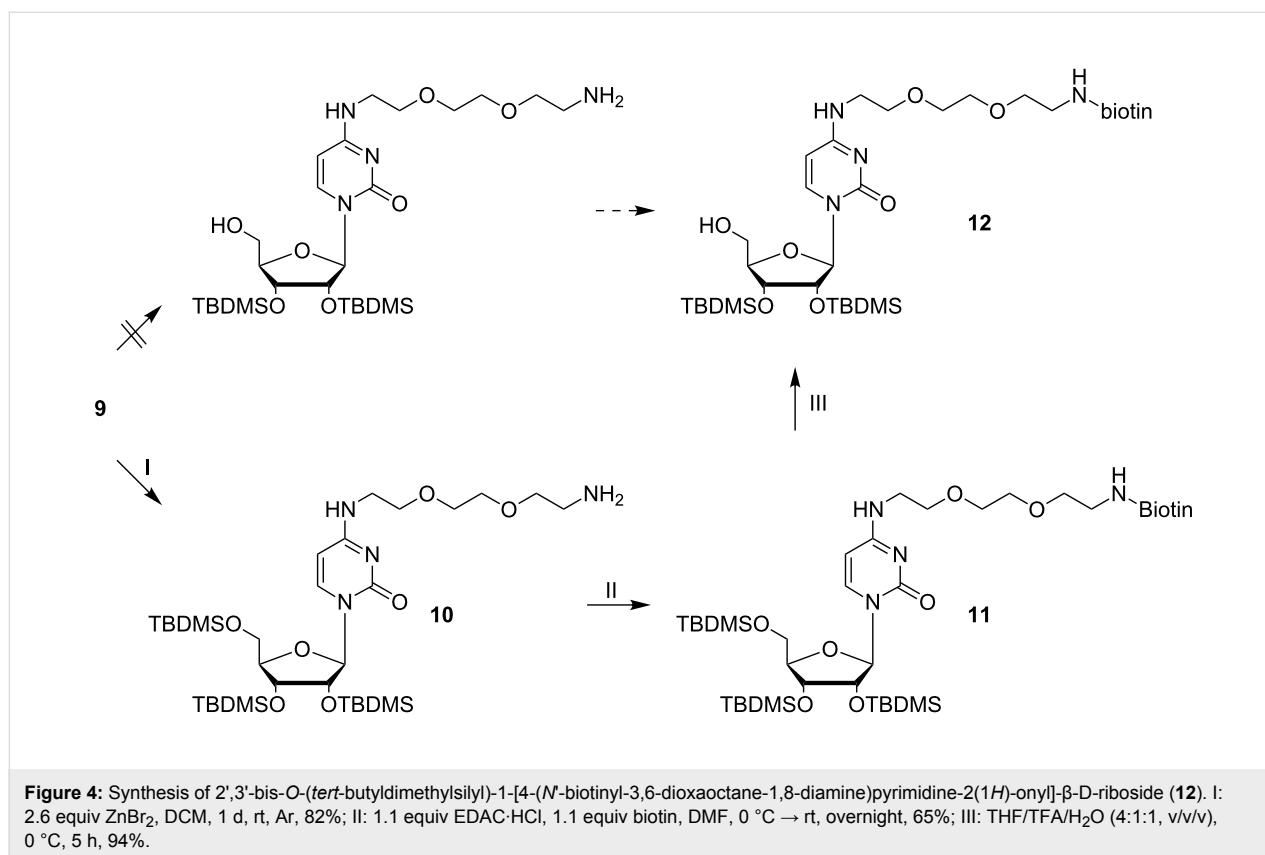
The mono protected diamino linker **4** was purified by chromatography on silica gel and subsequently used for reaction with the triazolyl protected uridine derivative **3** to produce compound **9** in 85% yield. Upon removal of the Boc group and sugar deprotection, the remaining aliphatic amine should be used for conjugation to biotin, followed by coupling of the 5'-hydroxy group to a linker phosphoramidite.

In terms of synthesis efficiency, we first tried to simultaneously deprotect the amino functionality at the nucleobase and the sugar 5'-hydroxy group to generate derivative **12**. Thereafter, the greater nucleophilicity of the primary amine over the alcohol should be used for selective amide bond formation with biotin, provided for reaction as *N*-hydroxysuccinimidyl (NHS-) ester, and subsequently the 5'-*O*-phosphoramidite should be prepared. Unfortunately, preparation of **12** in the suggested one step procedure (Figure 4, upper path) was not possible.

According to a previous report by Zhu et al. [36], the 5'-OH group of tris-silylated nucleosides can be selectively removed by treatment with a THF/TFA/H_2O mix (4:1:1, v/v/v) at 0 °C. Since cleavage of Boc-groups requires strong Brønsted acids [37-39], we varied the protocol of Zhu et al. with respect to the acid concentration and to reaction time. However, we could not

Figure 3: Preparation of synthon **4** and substitution of the triazolyl moiety of **3** to form the fully protected cytidine derivative **9**. I: 0.16 equiv Boc_2O, dioxane, 3 h, rt, 68% (based on the amount of Boc_2O); II: 0.77 equiv triazolyl protected uridine **3**, MeCN, rt, overnight, 85%.

Figure 4: Synthesis of 2',3'-bis-*O*-(*tert*-butyldimethylsilyl)-1-[4-(*N'*-biotinyl-3,6-dioxaoctane-1,8-diamine)pyrimidine-2(1*H*)-onyl]-β-D-riboside (**12**). I: 2.6 equiv ZnBr$_2$, DCM, 1 d, rt, Ar, 82%; II: 1.1 equiv EDAC·HCl, 1.1 equiv biotin, DMF, 0 °C → rt, overnight, 65%; III: THF/TFA/H$_2$O (4:1:1, v/v/v), 0 °C, 5 h, 94%.

identify conditions that led to removal of the Boc group, and leaving the TBDMS groups intact; multiple desilylated products were detected in all cases. These findings forced us to synthesize **12** in an alternative way. Apart from protic acids, the *tert*-butoxycarbonyl moiety of Boc-protected amines can be selectively cleaved with Lewis acids [40]. Thus, stirring of **9** in anhydrous DCM with ZnBr$_2$ delivered the free amine **10** with a yield of 82% (Figure 4). To this, biotin was coupled by in situ activation of the carboxylic acid as an NHS-ester using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDAC·HCl). Next, the 5'-*O*-TBDMS group was removed by stirring **11** in a mixture of THF/TFA/H$_2$O (4:1:1, v/v/v) at 0 °C, and solely the 5'-*O*-deprotected cytidine derivative **12** was obtained in 94% yield (Figure 4). The long amino linker **13** was synthesized from hexaethylene glycol with an overall yield of 52% over three steps following a standard literature protocol [41]. After introduction of a Boc group for protection of the linker amino function, the resulting compound **14** was reacted with (2-cyanoethyl-*N,N*-diisopropyl)chlorophosphoramidite to yield the phosphoramidite **2**, which was coupled with synthon **12** in anhydrous THF in the presence of 5-benzylmercapto-1*H*-tetrazole (BMT) as activator (Figure 5).
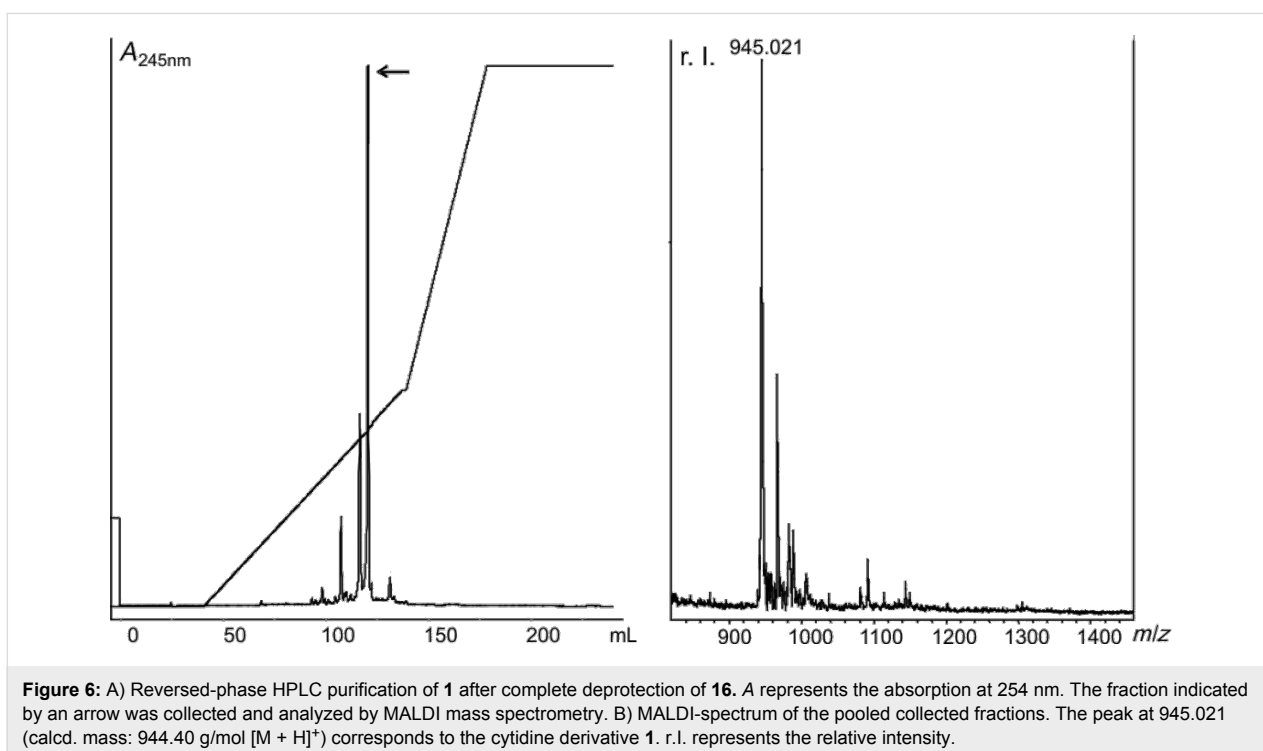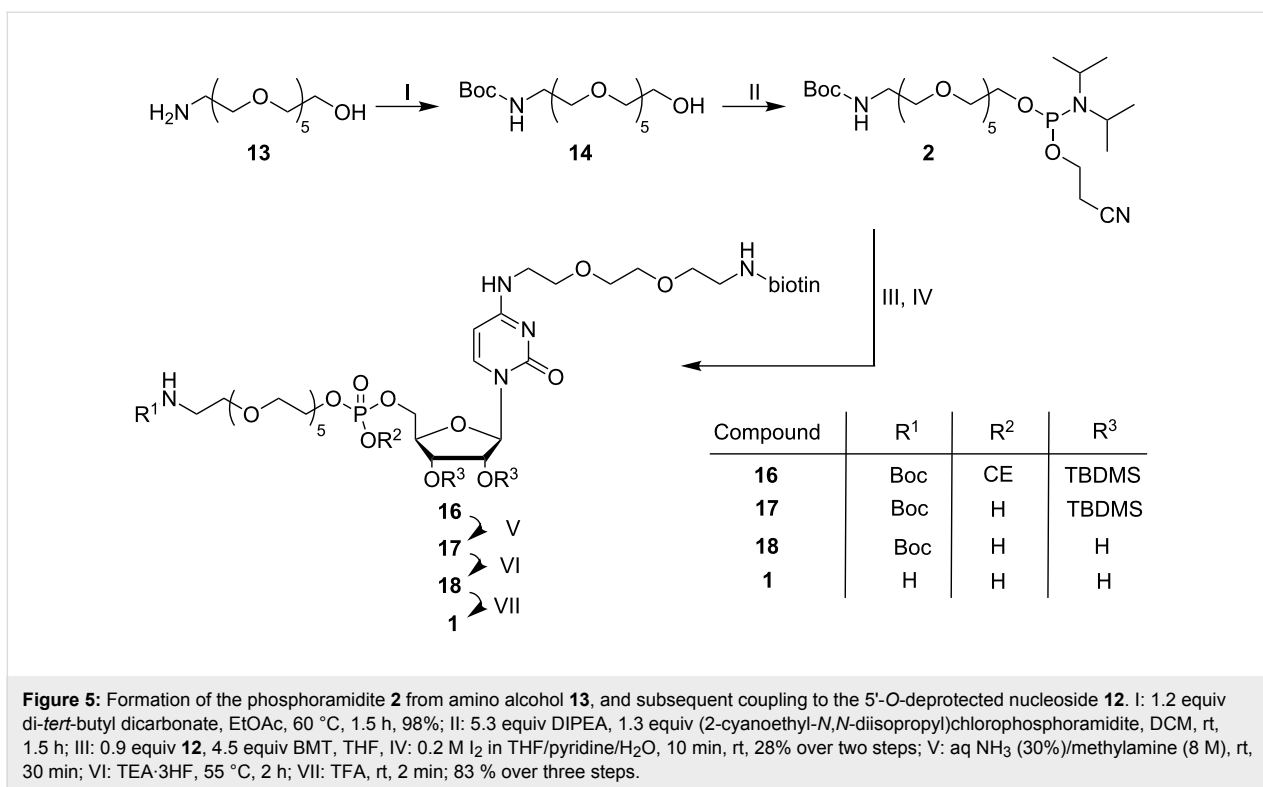
The coupling product **16** was isolated with a yield of 28%. Finally, all remaining protecting groups were cleaved off. The base-labile β-cyanoethyl (CE) group was removed by stirring **16** in a 1:1 mixture of concentrated aqueous ammonia and methylamine. Subsequent treatment of **17** with triethylamine trihydrofluoride (TEA·3HF) at 55 °C for 2 h delivered compound **18**, which was reacted with neat TFA. Each deprotection step was monitored by MALDI mass spectrometry. The bifunctionalized cytidine derivative **1** as the final product was isolated by reversed-phase chromatography with a yield of 83% over the last three steps (Figure 6).

To set-up conditions for conjugation of the bifunctional cytidine derivative **1** to RNA, we synthesized a short model oligonucleotide (GUC AGC CGU CAG GAU CCG UG) corresponding to the 3'-terminal sequence of the envisaged RNA-library. The RNA was oxidized by treatment with sodium periodate [42], and the resulting dialdehyde was reacted with cytidine derivative **1**. The obtained intermediate was reduced with sodium cyanoborohydride to form the more stable amine as final product (Figure 7).

The reaction analysis was carried out by reversed-phase HPLC (Figure 8).

The RP-chromatogram of the crude product (A) shows two peaks (denoted as I) corresponding to the oxidized and

Figure 5: Formation of the phosphoramidite **2** from amino alcohol **13**, and subsequent coupling to the 5'-*O*-deprotected nucleoside **12**. I: 1.2 equiv di-*tert*-butyl dicarbonate, EtOAc, 60 °C, 1.5 h, 98%; II: 5.3 equiv DIPEA, 1.3 equiv (2-cyanoethyl-*N,N*-diisopropyl)chlorophosphoramidite, DCM, rt, 1.5 h; III: 0.9 equiv **12**, 4.5 equiv BMT, THF, IV: 0.2 M I$_2$ in THF/pyridine/H$_2$O, 10 min, rt, 28% over two steps; V: aq NH$_3$ (30%)/methylamine (8 M), rt, 30 min; VI: TEA·3HF, 55 °C, 2 h; VII: TFA, rt, 2 min; 83 % over three steps.

Figure 6: A) Reversed-phase HPLC purification of **1** after complete deprotection of **16**. *A* represents the absorption at 254 nm. The fraction indicated by an arrow was collected and analyzed by MALDI mass spectrometry. B) MALDI-spectrum of the pooled collected fractions. The peak at 945.021 (calcd. mass: 944.40 g/mol [M + H]$^+$) corresponds to the cytidine derivative **1**. r.I. represents the relative intensity.

cyanoborohydride reduced RNA (C), peak II corresponding to the cytidine derivative **1** (B), and a new peak (III) indicating the desired RNA–cytidine conjugate, as confirmed by MALDI mass spectrometry (see Supporting Information File 1).

Conclusion

We successfully synthesized a cytidine derivative for the functionalization of an RNA library following a synthetic route over 16 steps. Two linker moieties of different lengths and function-
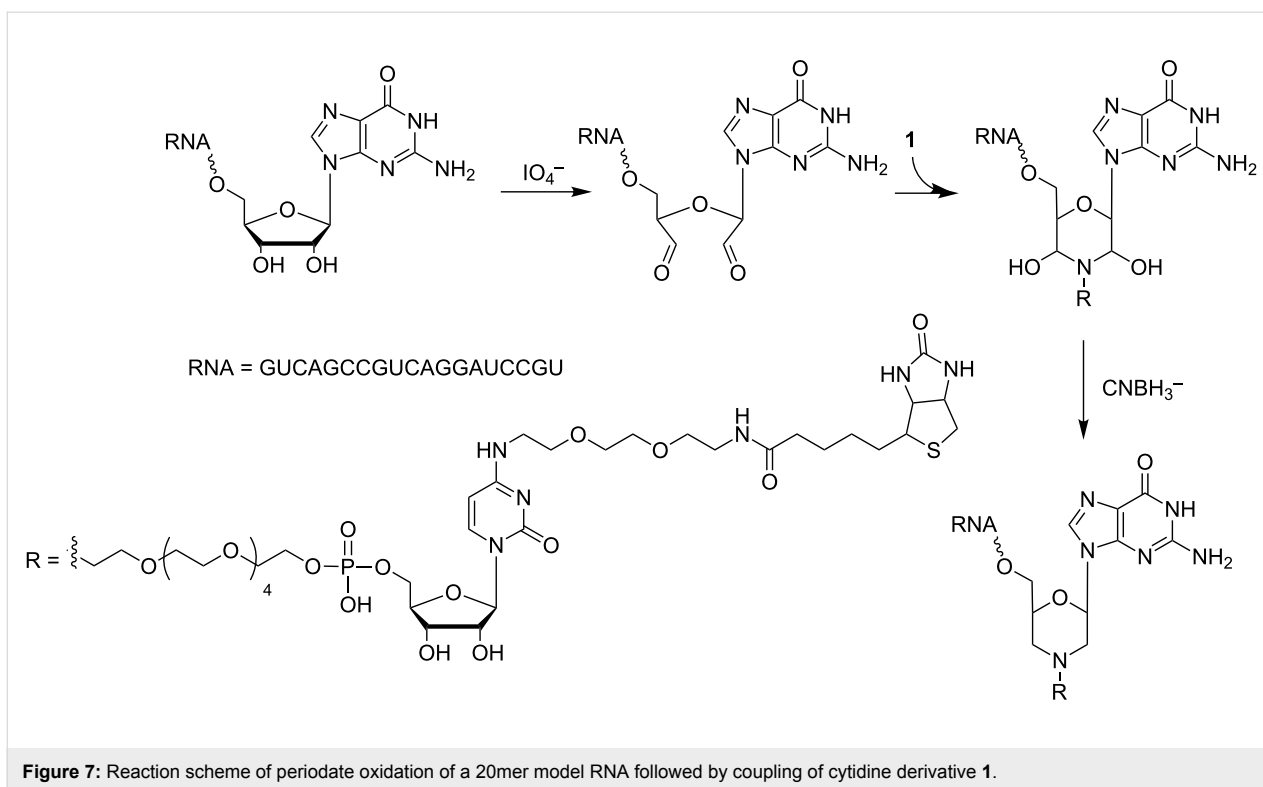
Figure 7: Reaction scheme of periodate oxidation of a 20mer model RNA followed by coupling of cytidine derivative **1**.
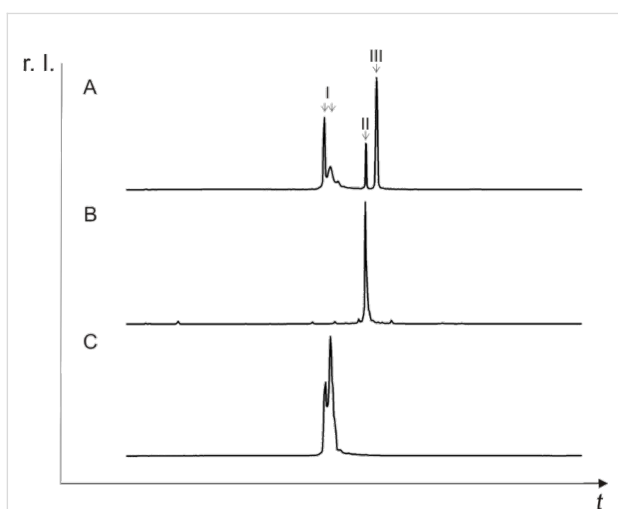

Figure 8: Reversed-phase HPLC analysis. A: Crude product of the coupling reaction between the 20mer model RNA and **1**. The peaks I correspond to the periodate oxidized and subsequently reduced RNA. Peak II matches the cytidine derivative **1**. Peak III indicates the RNA–cytidine **1** conjugate. B: Cytidine derivative **1**. C: 20mer model RNA after treatment with sodium periodate and reduction with sodium cyanoborohydride in the absence of cytidine derivative **1**. r. I.: relative intensity.

ality were attached to the nucleobase and to the sugar residue of cytidine, thus enabling for one coupling of the nucleoside to periodate-oxidized RNAs, and for second linking the functionalized library to a surface. The 5'-*O*-coupled linker carrying an aliphatic amino group allows 3'-terminal conjugation of cytidine to the molecules of an RNA library, whereas the biotinylated linker attached to C4 of the nucleobase allows linking the modified library to a streptavidine coated surface. The suitability of the synthesized cytidine derivative was confirmed by successful conjugation to the 3'-terminus of a model RNA, as analyzed by HPLC and MALDI–MS. This set-up is going to be used for selection of a cytidine deaminase ribozyme supporting the conversion of uridine to cytidine. Active molecules will be cleaved from the solid phase and released into solution, such that those can be collected, reverse-transcribed and amplified to enter the next round of selection.

Experimental

The bifunctional cytidine derivative **1** was synthesized starting from uridine. All reaction steps and the characterization of the obtained products are described in detail in Supporting Information File 1. The RNA GUC AGC CGU CAG GAU CCG UG used as model for studying 3'-terminal conjugation of **1** was synthesized on an Applied Biosystems 394 DNA/RNA Synthesizer following the standard protocol for oligoribonucleotide chain assembly. The synthesized RNA was deprotected using aqueous ammonia (32%)/aq methylamine (40%) (1:1, v/v) at 65 °C for 30 min for removal of base and phosphate protecting groups and cleavage from the support, and TEA·3HF for removal of 2'-*O*-protecting groups, essentially as described previously [43].

Periodate oxidation

For the oxidation of the RNA's 3'-terminal *cis*-diol, 5 nmol of the 20mer RNA substrate were incubated in a total volume of 1 mL 30 mM sodium metaperiodate, 100 mM NaOAc (pH 5.4) for 1 hour at room temperature in the dark. The oxidized RNA was recovered by ethanol precipitation.

Conjugation of the cytidine derivative **1** to RNA

The oxidized RNA was dissolved in 250 μL water containing 20 mM imidazole (pH 8), 5 mM NaCNBH$_3$, 1 mM EDTA and 1 mM cytidine derivative **1**. The reaction was carried out at 37 °C. After 2 hours, 25 μL of 50 mM NaBH$_4$ were added, and the reaction mixture was incubated for additional 15 min. Ethanol precipitation yielded the crude coupling product, which was analyzed by reversed-phase HPLC on an Äkta Purifier (Amersham Bioscience). Column: Macherey Nagel EC 250/4 Nucleodur 100-5 C18 ec; Buffers: (A) 0.1 M triethyl-ammonium acetate (pH 7.0), 5% acetonitrile and (B) 0.1 M tri-ethylammonium acetate (pH 7.0), 30% acetonitrile; flow rate 0.5 mL min^{-1}; gradient: 0% → 85% (B) in 14 CV, 85% 4 CV, 85% → 100% (B), 100% (B) 4 CV, 100% → 0% (B) 2 CV. The product containing fraction was collected, RNA was lyophilized and analyzed by MALDI–MS: (calcd mass: 7339.9 g/mol, found: 7339.04 [M + 1]$^+$; 3668.515 [M + 2]$^+$/2, Supporting Information File 1).

Supporting Information

Supporting Information File 1

Experimental part.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-198-S1.pdf]

Acknowledgements

References

1. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982,** *31,* 147–157. doi:10.1016/0092-8674(82)90414-7
2. Doudna, J. A.; Cech, T. R. *Nature* **2002,** *418,* 222–228. doi:10.1038/418222a
3. Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Höbartner, C.; Polonskaia, A.; Phan, A. T.; Wombacher, R.; Micura, R.; Dauter, Z.; Jäschke, A.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2005,** *12,* 218–224. doi:10.1038/nsmb906
4. Sengle, G.; Eisenführ, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001,** *8,* 459–473. doi:10.1016/S1074-5521(01)00026-6
5. Tsukiji, S.; Ramaswamy, K.; Suga, H. *Pure Appl. Chem.* **2004,** *76,* 1525–1536. doi:10.1351/pac200476071525
6. Fusz, S.; Eisenführ, A.; Srivatsan, S. G.; Heckel, A.; Famulok, M. *Chem. Biol.* **2005,** *12,* 941–950. doi:10.1016/j.chembiol.2005.06.008
7. Unrau, P. J.; Bartel, D. P. *Nature* **1998,** *395,* 260–263. doi:10.1038/26193
8. Gilbert, W. *Nature* **1986,** *319,* 618. doi:10.1038/319618a0
9. Powell, L. M.; Wallis, S. C.; Pease, R. J.; Edwards, Y. H.; Knott, T. J.; Scott, J. *Cell* **1987,** *50,* 831–840. doi:10.1016/0092-8674(87)90510-1
10. Navaratnam, N.; Sarwar, R. *Int. J. Hematol.* **2006,** *83,* 195–200. doi:10.1532/IJH97.06032
11. Carter, C. W., Jr. *Biochimie* **1995,** *77,* 92–98. doi:10.1016/0300-9084(96)88110-7
12. Ellington, A. D.; Szostak, J. W. *Nature* **1990,** *346,* 818–822. doi:10.1038/346818a0
13. Tuerk, C.; Gold, L. *Science* **1990,** *249,* 505–510. doi:10.1126/science.2200121
14. Sampson, J. R.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. U. S. A.* **1988,** *85,* 1033–1037. doi:10.1073/pnas.85.4.1033
15. Wolf, J.; Dombos, V.; Appel, B.; Müller, S. *Org. Biomol. Chem.* **2008,** *6,* 899–907. doi:10.1039/b716151d
16. Skipsey, M.; Hack, G.; Hooper, T. A.; Shankey, M. C.; Conway, L. P.; Schröder, M.; Hodgson, D. R. W. *Nucleosides, Nucleotides Nucleic Acids* **2013,** *32,* 670–681. doi:10.1080/15257770.2013.851393
17. Samanta, A.; Krause, A.; Jäschke, A. *Chem. Commun.* **2014,** *50,* 1313–1316. doi:10.1039/c3cc46132g
18. Harris, M. E.; Christian, E. L. *Methods* **1999,** *18,* 51–59. doi:10.1006/meth.1999.0756
19. Burgin, A. B.; Pace, N. R. *EMBO J.* **1990,** *9,* 4111–4118.
20. Davies, M. J.; Shah, A.; Bruce, I. J. *Chem. Soc. Rev.* **2000,** *29,* 97–107. doi:10.1039/a905230e
21. Paredes, E.; Das, S. R. *ChemBioChem* **2011,** *12,* 125–131. doi:10.1002/cbic.201000466
22. Rosemeyer, V.; Laubrock, A.; Seibl, R. *Anal. Biochem.* **1995,** *224,* 446–449. doi:10.1006/abio.1995.1068
23. Igloi, G. L. *Anal. Biochem.* **1996,** *233,* 124–129. doi:10.1006/abio.1996.0016
24. Martin, G.; Keller, W. *RNA* **1998,** *4,* 226–230.
25. Winz, M.-L.; Samanta, A.; Benzinger, D.; Jäschke, A. *Nucleic Acids Res.* **2012,** *40,* e78. doi:10.1093/nar/gks062
26. Malaprade, L. *Bull. Soc. Chim. Fr.* **1928,** *43,* 683–696.
27. House, H. O. *Modern synthetic reactions,* 2nd ed.; W. A. Benjamin Inc.: Menlo Park, California, 1972; pp 353–359.
28. Marinetti, G. V.; Rouser, G. *J. Am. Chem. Soc.* **1955,** *77,* 5345–5349. doi:10.1021/ja01625a042
29. Winston, S. E.; Fuller, S. A.; Evelegh, M. J.; Hurrell, J. G. R. Conjugation of Enzymes to Antibodies. *Current protocols in molecular biology;* John Wiley & Sons, 2001; Vol. 11, 11.1. doi:10.1002/0471142727.mb1101s50
30. Hisamatsu, Y.; Hasada, K.; Amano, F.; Tsubota, Y.; Wasada-Tsutsui, Y.; Shirai, N.; Ikeda, S.-i.; Odashima, K. *Chem. – Eur. J.* **2006,** *12,* 7733–7741. doi:10.1002/chem.200600099
31. Amantea, A.; Henz, M.; Strazewski, P. *Helv. Chim. Acta* **1996,** *79,* 244–254. doi:10.1002/hlca.19960790125
32. Sung, W. L. *Nucleic Acids Res.* **1981,** *9,* 6139–6152. doi:10.1093/nar/9.22.6139
33. Reese, C. B.; Ubazawa, A. *Tetrahedron Lett.* **1980,** *21,* 2265–2268. doi:10.1016/0040-4039(80)80020-7

34. Höbartner, C.; Kreutz, C.; Flecker, E.; Ottenschläger, E.; Pils, W.; Grubmayr, K.; Micura, R. *Monatsh. Chem.* **2003,** *134,* 851–873. doi:10.1007/s00706-003-0592-1

35. Boeijen, A.; Liskamp, R. M. J. *Eur. J. Org. Chem.* **1999,** 2127–2135. doi:10.1002/(SICI)1099-0690(199909)1999:9<2127::AID-EJOC2127>3.0.CO;2-T

36. Zhu, X.-F.; Williams, H. J.; Scott, A. I. *J. Chem. Soc., Perkin Trans. 1* **2000,** 2305–2306. doi:10.1039/b003562i

37. Stahl, G. L.; Walter, R.; Smith, C. W. *J. Org. Chem.* **1978,** *43,* 2285–2286. doi:10.1021/jo00405a045

38. Sakai, N.; Ohfune, Y. *J. Am. Chem. Soc.* **1992,** *114,* 998–1010. doi:10.1021/ja00029a031

39. Shendage, D. M.; Fröhlich, R.; Haufe, G. *Org. Lett.* **2004,** *6,* 3675–3678. doi:10.1021/ol048771l

40. Kaul, R.; Brouillette, Y.; Sajjadi, Z.; Hansford, K. A.; Lubell, W. D. *J. Org. Chem.* **2004,** *69,* 6131–6133. doi:10.1021/jo0491206

41. Miller, J. S.; Béthencourt, M. I.; Hahn, M.; Lee, T. R.; West, J. L. *Biotechnol. Bioeng.* **2006,** *93,* 1060–1068. doi:10.1002/bit.20809

42. Zahler, N. H.; Harris, M. E. In *Handbook of RNA biochemistry;* Hartmann, R. K.; Bindereif, A.; Schön, A.; Westhof, E., Eds.; Wiley-VCH: Weinheim, Germany, 2005; Vol. 1, pp 75–85.

43. Drude, I.; Strahl, A.; Galla, D.; Müller, O.; Müller, S. *FEBS J.* **2011,** *278,* 622–633. doi:10.1111/j.1742-4658.2010.07983.x