Organic & Biomolecular Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Journal Name

RSCPublishing

ARTICLE

RNA-peptide conjugate synthesis by inverse-electron demand Diels-Alder reaction

Cite this: DOI: 10.1039/x0xx00000x

Sandeep Ameta,[‡a] Juliane Becker[‡a] and Andres Jäschke[a]*,

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Here we report an efficient method for the synthesis of RNA-peptide conjugates by inverse-electron demand Diels-Alder reaction. Various dienophiles were enzymatically incorporated into RNA and reacted with a chemically synthesized diene-modified peptide. The Diels-Alder reaction proceeds with near-quantitative yields in aqueous solution with stoichiometric amounts of reactants, even at low micromolar concentrations.

Introduction

RNA, being a simple molecule, plays a major role in catalysis,[1] metabolite sensing,[2] and gene regulation.[3] To better understand the structural and functional dynamics of RNA molecules, their site-specific labelling is often required. Examples for such labels are affinity handles,[4] fluorescent moieties[5] or peptide tags.[6] The labelling of RNA molecules with peptides has been exploited for cell-specific targeting and cellular delivery of small interfering RNAs.[7] Moreover, RNA-peptide conjugates have been used to understand the dynamics of the ribosome elongation cycle.[6, 8] Therefore, there is a strong interest in developing efficient chemical conjugation methods for the covalent labelling of RNA with peptides.

Our laboratory has a long-standing interest in evolving RNA catalysts with protease activity,[9] and RNA-peptide conjugates are essential tools for this purpose. To attach a substrate peptide to RNA, we earlier developed a conjugation strategy based on aldehyde-hydrazine chemistry.[10] RNA molecules were prepared by transcription, initiated with an aldehyde-containing guanosine and then conjugated to a hydrazine-bearing peptide. Although being simple and feasible, this approach required 50- to 100-fold excess of the peptide over RNA, long reaction times (4-6 h), and still yielded no more than 70% conjugate. Moreover, it also involved acid treatment for deprotection of the aldehyde which gave rise to partial degradation of RNA. Therefore, there is need for faster, milder, and more selective conjugation methods based on an efficient chemistry in aqueous solution.

`Click´ reactions are now considered as state-of-the-art for labelling biomolecules. While there are elegant demonstrations of using copper-catalysed azide alkyne cycloadditions (CuAAC) for oligonucleotide labelling,[11] copper-free cycloadditions have also been exploited for the labelling of biomolecules to circumvent undesired effects of copper ions.[11b, 12] Moreover, click reactions based on inverse-electron demand Diels-Alder (DAinv) cycloadditions were developed as a bioorthogonal labelling approach for biomolecules. Here, a strained dienophile reacts with an electron-deficient diene, where the diene can be a diazine[13], triazine[14] or a tetrazine[15, 16]. As tetrazines are most reactive, they are usually the substrate of choice for the bioconjugation reactions. Recently, DAinv cycloadditions were reported as a tool for labelling biomolecules *in vitro* as well as *in vivo,*[17] and we extended this approach to DNA labelling.[18] Lately we developed a RNA labelling method based on DAinv where RNA molecules were labelled, starting from either chemically or enzymatically synthesized norbornene-carrying oligonucleotides.[19] While norbornene is an efficient and selective dienophile in Diels-Alder-based labelling, its moderate reactivity is limiting its application at low micromolar concentrations. Recently, more reactive dienophiles have been reported, such as tricyclooctene **2**, bicyclononyne **3** and *trans*-cyclooctene **4** (Figure 1A).[17c, 20]

In the current study we developed a RNA-peptide conjugation strategy based on DAinv by synthesizing and enzymatically incorporating faster reacting dienophiles into RNA, followed by their reaction with a tetrazine-derivatized peptide. For the best system, we observed quantitative RNA-peptide conjugate formation even at 1:1 stoichiometry of RNA and peptide at low micromolar concentrations.

Results and discussion

First, we designed a guanosine monophosphate (GMP) initiator nucleotide based on tricyclooctene, i.e., a norbornene fused to a cyclopropane ring (modification **2**, Figure 1A), as such fusions had been reported to increase the reactivity in DAinv.[16, 21] The GMP initiator nucleotide bearing this modification was synthesized by using an alcohol of **2** and a 5'-phosphoramidite of guanosine using standard synthesis protocols (Figure 1B). Purification by reversed-phase chromatography and subsequent

ion-exchange resulted in 45% yield of the final compound **GMP2**. Next, the incorporation of this initiator nucleotide at the 5'-end of a 19mer RNA using *in vitro* transcription by T7

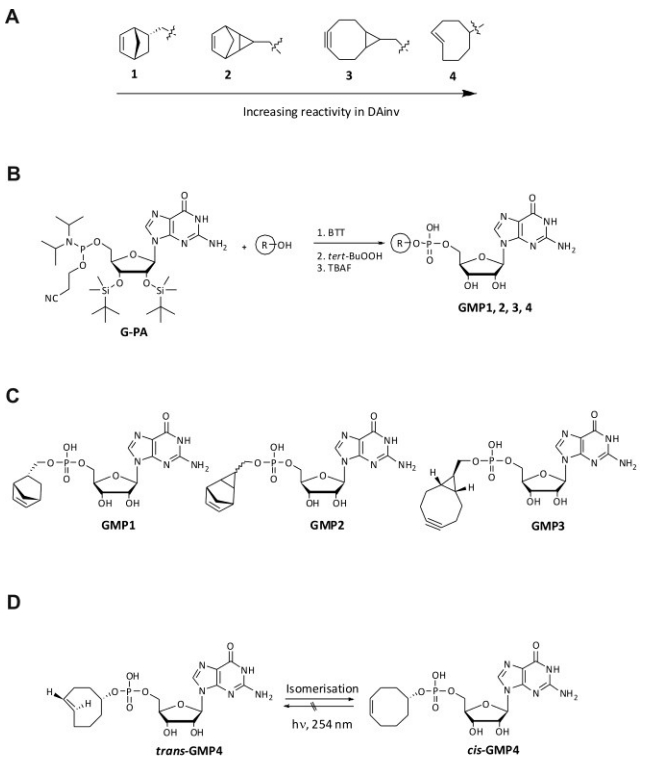

Figure 1. Dienophiles and initiator nucleotides used in this study. (A) Chemical structures of different dienophiles. **(B)** Synthesis of initiator nucleotides. **(C)** Structures of initiator nucleotides used. **(D)** *Cis-trans* isomerization of cyclooctene.

RNA polymerase was analysed at different ratios of initiator *vs.* GTP (1:1, 5:1, 10:1). Electrophoretic analysis of the transcription mixtures revealed that the incorporation of **GMP2** increased with the increasing ratio, with the best incorporation efficiency of ~ 70% at a ratio of 10:1 (see Supplementary Information, Figure S1), corresponding well with previous reports on transcription initiation. [10, 19] LC-MS analysis of a transcription reaction using a 10:1 ratio of GMP2 *vs.* GTP showed the mass of the expected initiated RNA (Table 1, RNA_**GMP2**).

Table 1 LC-MS analysis of initiated RNA-oligonucleotides

RNA	R_t [min]	[M] calculated	[M] deconvoluted	Δ [ppm]
RNA_**GMP2**	20.6	6130.8509	6130.8275	3.8
RNA_**GMP3**	17.7	6144.8660	6144.8444	3.5
RNA_**AMP3**	19.6	6473.9190	6473.8784	6.3

In order to obtain a peptide harbouring a tetrazine moiety, an arbitrary peptide sequence was synthesized using standard solid-phase peptide synthesis (Fmoc chemistry) and coupled on-resin to a tetrazine NHS ester. The tetrazine-derivatized peptide was cleaved from the solid support by standard procedures and purified by reversed-phase HPLC (see Supplementary Information, Figure S2). For DAinv, the *in vitro* transcribed **GMP2**-initiated RNA was mixed with an excess of tetrazine peptide (>5 equivalents) in water and incubated at room temperature for 50 minutes. The resulting products were analysed by LC-MS which confirmed the formation of the expected DAinv product (Figure 2A).

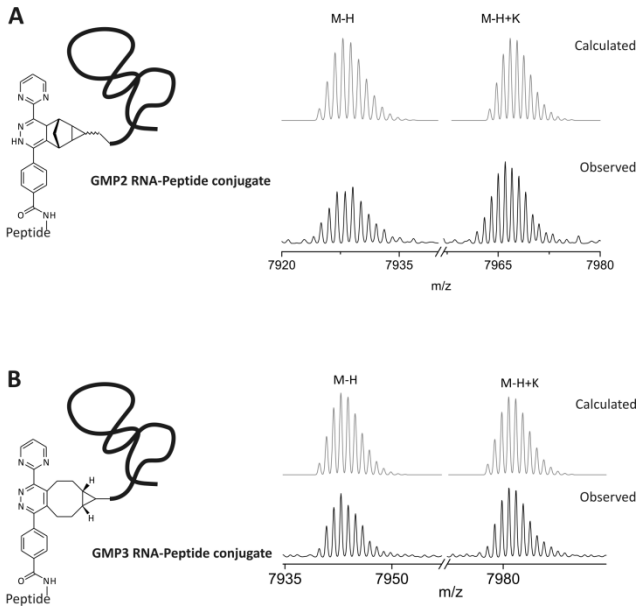

Figure 2. Mass analysis of RNA-peptide conjugation. Conjugate formation between a tetrazine peptide and **(A) GMP2**-initiated 19mer RNA or **(B) GMP3**-initiated 19mer RNA.

In order to speed up the conjugation, initiator nucleotides containing faster dienophiles, namely bicyclononyne **3** and *trans*-cyclooctene **4,** were synthesized (see Figure 1A). Compound **3** was found to be stable under the synthesis conditions and the guanosine initiator **GMP3** could be isolated in 24% yield over two steps. The highly reactive *trans*-cyclooctene did not withstand the harsh reaction conditions during initiator nucleotide synthesis. Instead of the desired *trans*-**GMP4,** only the corresponding *cis*-isomer, *cis*-**GMP4**, could be isolated (Figure 1D) which is known to be less reactive in DAinv.[22] Attempts of re-isomerising *cis*-**GMP4** to *trans*-**GMP4** by UV-light (254 nm) failed (data not shown). Therefore, no further study on **GMP4** was performed.

Similar to **GMP2**, the incorporation efficiency of **GMP3** was found to be ~70% at a 10:1 ratio of **GMP3**:GTP (see Supplementary Information, Figure S1). Mass analysis

confirmed successful RNA initiation (Table 1, RNA_**GMP3**) as well as RNA-peptide conjugate formation (Figure 2B). Since cyclooctyne **3** is a very reactive dienophile in DAinv, as a next step, we evaluated the amount of tetrazine-diene required for quantitative product formation by incubating 3 µM of **GMP3**-initiated 19mer RNA with varying concentrations of tetrazine-peptide for 10 min at room temperature. Gel-electrophoretic analysis revealed near-quantitative product formation already at 1:1 stoichiometry (Figure 3). In comparison to the previously reported Diels-Alder conjugation on RNA using the norbornene-carrying initiator nucleotide **GMP1** (Figure 1C)[19], where a 5-fold excess of tetrazine and higher absolute concentrations were required, the conjugation is much faster using bicyclononyne **GMP3** as dienophile.

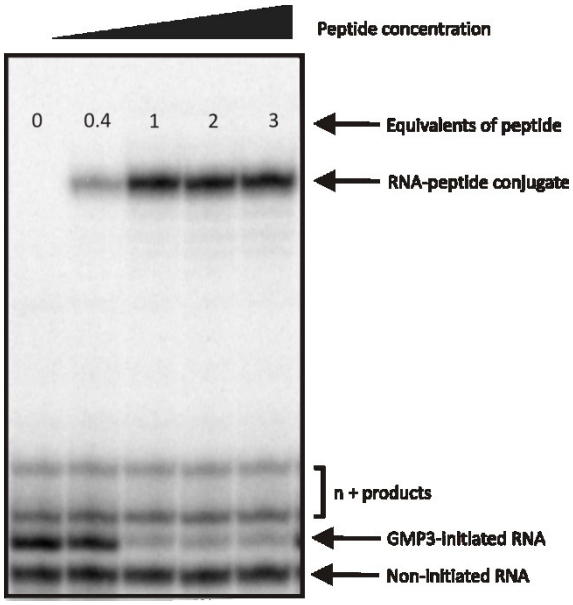

Figure 3. Diels-Alder conjugation at different RNA-peptide ratios. Sequencing PAGE analysis of RNA-peptide conjugate formation conducted at different RNA/peptide ratios. 3 µM of **GMP3**-initiated 19mer RNA was reacted for 10 min. RNA was radioactively labelled using α-^{32}P CTP during *in vitro* transcription and analysed on a 18% denaturing sequencing PAGE gel. 'n+1' products are formed by the non-templated addition of NTPs by T7 RNA polymerase during transcription.[23]

While guanosine-based initiators have found broad applications for terminally modifying natural as well as non-natural RNAs, some promoters require a different nucleotide, e.g. adenosine, for the initiation. To expand the scope of DAinv-based RNA-peptide conjugation, we synthesized an adenosine-based initiator nucleotide bearing bicyclononyne dienophile **3,** followed by transcription initiation. Although incorporation of the adenosine-based initiator **AMP3** was found to be less efficient than the guanosine-based **GMP3,** it was well in the range reported in a previous study[24], and successful incorporation could be verified via LC-MS analysis (Table 1). Diels-Alder conjugation on the initiated RNA yielded the expected RNA-peptide conjugate (Figure 4).

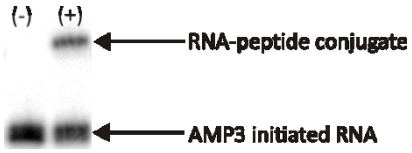

Figure 4. Gel analysis of RNA-peptide conjugation by AMP3 initiated RNA. 15% denaturing PAGE analysis of the conjugate formation between tetrazine peptide and **AMP3** initiated 19mer RNA. **(-)** and **(+)** represents presence or absence of tetrazine peptide in the reaction, respectively.

As our goal is the application of this conjugation strategy in order to evolve protease ribozymes, the aforementioned approach has to function efficiently on longer RNAs. Therefore, a 232 nucleotides long RNA, RNA19 ribozyme,[9b] was *in vitro* transcribed in the presence of **GMP3** and subsequently conjugated with the tetrazine peptide (in 1:1 stoichiometry). As the attachment of a small peptide to a rather long RNA creates only a small change in the electrophoretic mobility of the RNA, a 10-23 DNAzyme[25] was designed to cleave off a 29mer from the 5'-end of the transcribed RNA (Figure 5A, Supplementary Information, Figure S3). When the DAinv product of 233mer RNA was subjected to DNAzyme cleavage and analysed by denaturing PAGE, the formation of a DAinv product was confirmed as the 29mer fragment

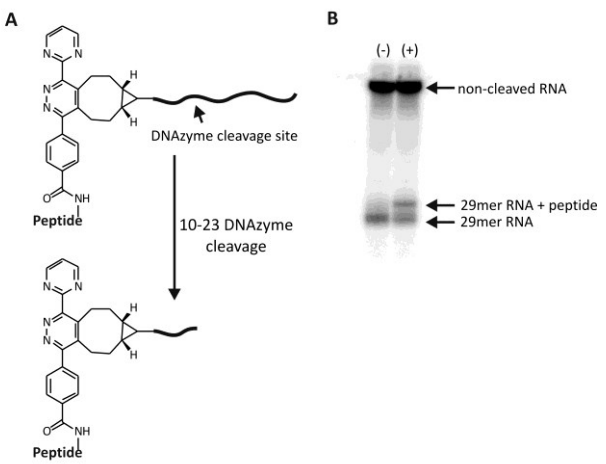

Figure 5. Peptide conjugation of long RNAs. (A) Scheme of the cleavage of a 232mer RNA conjugated to the tetrazine peptide by a 10-23 DNAzyme. **(B)** PAGE analysis of DNAzyme cleavage of 232mer RNA showing the peptide conjugated 29mer RNA fragment. RNA has been radioactively labelled using α-^{32}P CTP during *in vitro* transcription and analysed on 18% denaturing PAGE. **(-)** and **(+)** represents presence or absence of tetrazine peptide in the reaction, respectively.

conjugated to the tetrazine peptide was shifted, compared to the control (Figure 5B).

Conclusions

In conclusion, RNA-peptide formation using DAinv and cyclooctyne as dienophile is a fast and efficient conjugation method and results in near-quantitative conversion at 1:1 stoichiometry and low micromolar concentrations. The mild reaction conditions render this versatile approach potentially useful for diverse applications. The current study highlights the potential of this method for attaching synthetic peptides to RNAs of different length and sequence. It is, however, to be noted that in the current study we used a peptide substrate rich in positively charged amino acids (Arg, Lys), as such peptides are attractive vehicles for the cellular delivery of RNA molecules (e.g., TAT peptide).[7] The positive charge might – *via* electrostatic interaction with the RNA – also increase the effective molarity and, thus, the reactivity in DAinv reactions. Other peptide substrates may therefore be slightly less reactive.
As the DAinv of strained cyclooctynes is several orders of magnitude faster than their strain-promoted azide-alkyne cycloaddition (SPAAC), the recently reported undesired side reactions in SPAAC of cysteines by thiol-yne addition[26] are likely much less prominent. Therefore, future studies will explore the applicability of this RNA-peptide conjugation method to full-size proteins and cell lysates, as well as its orthogonality with other modification strategies.

Experimental Section

List of abbreviations. TBAF = tetrabutylammoniumfluoride, TBDMS = *tert*-butyldimethylsilane, ACN = acetonitrile, min = minutes, h = hours, DA = Diels-Alder, DAinv = inverse electron-demand Diels-Alder reaction, TEA = triethylamine, PAGE = polyacrylamide gel electrophoresis.

General procedures and methods

All reagents were purchased from Sigma-Aldrich and used without further purification. Reversed-phase HPLC purification was performed on an Agilent 1100 Series HPLC system equipped with a diode-array detector using a semi-preparative Phenomenex Luna C18 column (5 µm, 15.0 × 250 mm) at a flow rate of 5 ml/min and eluted with a gradient of 100 mM triethylammonium acetate *p*H 7.0 (buffer A) and 100 mM triethylammonium acetate in 80% acetonitrile (buffer B). LC-MS experiments were performed on a Bruker microTOFQ-II ESI mass spectrometer connected to an Agilent 1200 Series HPLC system equipped with a multi-wavelength detector. A Phenomenex Kinetex C18 column (2.6 µm, 2.1 × 100 mm) was used and eluted at a flow rate of 0.2 ml/min using a gradient of 100 mM hexafluoroisopropanol/ 8.6 mM triethylamine *p*H 8.3 and methanol (LC-MS grade). Analysis of the LC-MS measurements was carried out using Hyphenation Star PP (Version 3.2.44.0) and DataAnalysis (Version 4.0, SP 4) software (Bruker Daltonics). MS-spectra were deconvoluted using Maximum Entropy deconvolution. For high-resolution mass spectra, internal calibration was performed (enhanced quadratic mode) using ESI Tunemix (Fluka) as calibrant. Calculated molecular weights refer to the m/z values given by the DataAnalysis software.
NMR spectra were recorded on a Varian Mercury Plus 500 MHz spectrometer. The assignment of proton and carbon resonances is based on two-dimensional correlation experiments (COSY, GHSQC, GHMBC). Flash purification was done on a Varian IntelliFlash 310 discovery scale flash purification system. Radioactive gels were exposed to store-phosphor screens (Amersham Biosciences) were scanned on a Typhoon 9400 Imager (GE Healthcare). Analysis and calculation of conversions were carried out with Image Quant software (Molecular Dynamics, Version 5.2).

Synthesis of G-initiator nucleotides

5′-*O*-(2-cyanoethyl-*N,N*′-diisopropylaminophosphino)-2',3'-di-*O*-(*tert*butyldimethylsilyl)-guanosine (232 mg, 324 µmol, 1.2 eq) [28] was dissolved in abs. THF (3.63 ml) and added to the respective alcohol (270 µmol of modification 2, 3 and 4). Benzylthiotetrazole (52.6 mg, 270 µmol) was added and the solution stirred for 1h at RT. 0.36 ml aqueous *tert*BuOOH was added and stirring continued for 30 min. The solvent was evaporated and then TBAF (1M in THF, 1.64 ml) was added. After stirring overnight at room temperature, the solvent was evaporated and the crude product was purified by reversed-phase flash chromatography using water/ACN as eluent system. Gradient: 5 minutes 5% ACN; increase to 30% ACN over 15 minutes; increase to 80% ACN over 5 min.
Fractions containing the initiator nucleotide were lyophilized and one round of ion exchange was performed (Na+ cycle).
5'-O-(Tricyclo[3.2.1.02,4]oct-6-en-3-ylmethyl)-guanosine monophosphate (GMP2). Yield: 58 mg, 121 µmol, 45%, mixture of *exo*-/*endo*-**GMP2** (ratio 2:1). ^{1}H-NMR (500 MHz, D_2O, 25 °C, TMS): δ = 0.54 (d, *J* = 7.65 Hz, 1H, H_{exo}), 0.64 (d, *J* = 7.65 Hz, 1H, H_{exo}), 0.70 (d, *J* = 9.25 Hz, 1H, H_{endo}), 0.78 (tt, *J* = 6.95, 2.27 Hz, 1H, H_{endo}), 0.92 (d, *J* = 9.25 Hz, 1H, H_{endo}), 1.06-1.16 (m, 2H, H_{exo}), 1.52 (d, *J* = 6.90 Hz, 1H, H_{endo}), 1.60 (td, *J* = 6.90, 1.61 Hz, 1H, H_{endo}), 1.94-1.98 (m, 1H, H_{exo}), 2.59 (bs, 1H, H_{exo}), 2.64 (bs, 2H, H_{endo}, Hexo), 2.70 (bs, 1H, H_{endo}), 3.13-3.31 (m, 1H, H_{exo}), 3.40 (t, *J* = 6.95 Hz, 2H, H_{endo}), 3.53 (td, J = 10.71, 6.13 Hz, 1H, H_{exo}), 4.02-4.07 (m, 2H, H5'), 4.29-4.34 (m, 1H, H4'), 4.55 (dd, *J* = 5.35, 3.75 Hz, 1H, H3'), 4.89 (td, *J* = 12.96, 5.35 Hz, 1H, H2'), 5.62 (dd, *J* = 5.40, 3.09 Hz, 1H, H_{endo}), 5.70 (dd, *J* = 5.40, 3.09 Hz, 1H, H_{endo}), 5.88-5.92 (m, 1H, H1'), 6.26 (dd, *J* = 5.40, 2.92 Hz, 1H, H_{exo}), 6.30 (dd, *J* = 5.40, 2.92 Hz, 1H, H_{exo}), 8.07 (s, 1H, H_{endo}), 8.08 (s, 1H, H_{exo}).
^{13}C {^{1}H}-NMR (125 MHz, D_2O, 25 °C, TMS): d = 26.7, 27.4, 29.4, 30.8, 38.4, 40.9, 41.0, 41.6, 42.0, 63.1, 64.7, 64.8, 67.1 (d, J_{P-C} = 5.58 Hz), 68.1 (d, J_{P-C} = 5.54 Hz), 70.3, 70.5, 72.8, 73.0, 83.6, 83.7, 83.8, 86.5, 86.8, 86.9, 116.3, 131.5, 131.7, 137.1, 137.3, 141.0, 141.7, 151.8, 151.9, 153.8, 158.7. ^{31}P{^{1}H}-NMR (121 MHz, D_2O, 25 °C, H_3PO_4): d = -0.25, -0.13. UV (H_2O): λ_{max} = 254 nm (ε = (10660 ± 620) $Lmol^{-1}cm^{-1}$). LC-MS (HR-

ESI$^-$): Rt = 11.13 min, 13.85 min; m/z (Rt = 11.13 min): 480.1320, m/z (Rt = 13.85 min): 480.1316 (calculated for $[C_{19}H_{24}N_5O_8P_1\text{-}H]^-$ 480.1290).

5'-*O*-(*exo*-bicyclo[6.1.0]non-4-yn-9-ylmethyl)-guanosine monophosphate (GMP3). For synthesis of respective alcohol see ref. [29]. Yield: 32.0 mg, 65.0 µmol, 24%. After ion exchange, a final purification was performed using preparative HPLC. Gradient: Increase from 15% buffer B to 27.5% buffer B over 25min. ^{1}H-NMR (500 MHz, D_2O, 25°C): δ = 0.13-0.25 (m, 2H), 0.52 (td, *J* = 5.69, 10.88 Hz, 1H), 1.06-1.16 (m, 2H), 1.99-2.24 (m, 6H), 3.22-3.27 (m, 1H), 3.84-3.88 (m, 1H), 4.00-4.07 (m, 2H), 4.31-4.33 (m, 1H), 4.53 (dd, *J* = 5.21, 3.32 Hz, 1H), 4.94 (t, *J* = 5.81 Hz, 1H), 5.90 (d, *J* = 6.22 Hz, 1H), 8.12 (s, 1H). ^{13}C {^{1}H} NMR (75 MHz, D_2O, 25°C, TMS): δ = 23.0, 24.1, 25.2, 35.0. 67.5, 73.2, 73.5, 75.1, 86.5, 88.8, 102.5, 119.0, 139.5, 154.7, 156.6, 161.4. ^{31}P-NMR (121 MHz, D_2O, 25°C, H_3PO_4): δ = -0.22. UV (H_2O): λ_{max} = 254 nm (ε = (12410 ± 530) Lmol^{-1}cm^{-1}). MS (HR-ESI$^-$): m/z 494.1419 (calculated for $[C_{20}H_{26}N_5O_8P_1\text{-}H]^-$ 494.1446).

Synthesis of A-initiator nucleotide

5'-*O*-(2-Cyanoethyl-*N*,*N*-diisopropylaminophosphino)-2',3'-O-(*ter*-butyldimethylsilyl) adenosine Under argon atmosphere, 2',3'-*O*-(*tert*-butyldimethylsilyl)-adenosine[30] (500 mg, 1.01 mmol, 1.09 eq) was diluted in 3.5 mL dry DCM and mixed with diisopropylethylamine (550 µL, 3.5 eq). The reaction mixture was cooled to 0 °C and 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (210 µL, 0.93 mmol) was added. After stirring the solution for one hour at room temperature the crude product was directly purified over a silica column (SiO_2, cyclohexane/ acetone = 2:1, 2% TEA, R_f = 0.37) and dried under high vacuum.

Yield: 280 mg, 0.40 mmol, 43%. ^{1}H-NMR (500 MHz, $CDCl_3$, 25 °C, TMS): δ = -0.19 - -0.15 (m, 3H), -0.06 - 0.01 (m, 3H), 0.68-0.12 (m, 6H), 0.79-0.83 (m, 9H), 0.93-0.95 (m, 9H), 1.20-1.23 (m, 12H), 2.65 (2t, *J* = 6.51, 1.98 Hz, 2H), 3.62-4.01 (m, 6H), 4.20-4.40 (m, 2H, $H_{2'}$, $H_{4'}$), 4.77 (t, *J* = 4.32 Hz, 1H, $H_{3'}$), 5.51 (bs, 2H, NH_2), 5.97 (d, *J* = 4.54 Hz, 1H, $H_{1'}$), 8.18 (s, 1H), 8.33 (s, 1H). ^{13}C{^{1}H}-NMR (125 MHz, $CDCl_3$, 25 °C, TMS): δ = -4.9, -4.7, -4.4, 17.9, 18.1, 20.4, 24.6, 24.7, 25.7, 25.8, 43.1 (d, $J_{P\text{-}C}$ = 12.40 Hz), 58.7 (d, $J_{P\text{-}C}$ = 20.90 Hz), 62.0 (d, $J_{P\text{-}C}$ = 14.50 Hz, $C_{5'}$), 71.9 ($C_{3'}$), 75.1 ($C_{2'}$), 84.0 ($C_{4'}$), 89.0 ($C_{1'}$), 117.4 (CN), 120.4, 140.0, 149.8, 152.9, 155.3. ^{31}P{^{1}H}-NMR (121 MHz, $CDCl_3$, 25 °C, H_3PO_4): δ = 148.4, 148.5. MS (HR-ESI$^+$): m/z 696.3851 (calculated for $[C_{31}H_{58}N_7O_5PSi_2\text{+H}]^+$ 696.3848).

5'-*O*-(*exo*-Bicyclo[6.1.0]non-4-in-9-ylmethyl)-adenosine monophosphate (AMP3)

AMP3 was synthesized similar to **GMP3** using 5'-O-(2-cyanoethyl-*N,N*-diisopropylaminophosphino)-2',3'-*O*-(*tert*butyl dimethylsilyl) adenosine instead of 5'-*O*-(2-cyanoethyl-*N,N*'-diisopropylaminophosphino)-2',3'-di-*O*-(*tert*butyldimethylsilyl) guanosine. Yield: 6.00 mg, 109 µmol, 5%. After ion exchange, a final purification was performed using preparative HPLC. Gradient: Increase from 15% buffer B to 27.5% buffer B over 25min; R_t (AMP3) = 20.7 min. ^{1}H-NMR (500 MHz, DMSO, 25 °C, TMS): δ = 0.50-0.58 (m, 3H), 1.10-1.25 (m, 2H), 1.98-2.05 (m, 2H), 2.12-2.27 (m, 4H), 3.07 (q, *J* = 7.25 Hz, 1H), 3.53 (dd, *J* = 14.42, 7.25 Hz, 1H), 3.83-3.93 (m, 2H, $H_{5'}$), 4.03 (dd, *J* = 7.38, 3.66 Hz, 1H, $H_{4'}$), 4.18-4.20 (m, 1H, $H_{3'}$), 4.60 (t, *J* = 5.48 Hz, 1H, $H_{2'}$), 5.91 (d, *J* = 5.48 Hz, 1H, $H_{1'}$), 8.14 (s, 1H), 8.40 (s, 1H). ^{13}C{^{1}H}-NMR (125 MHz, DMSO, 25 °C, TMS): δ = 20.8, 21.9, 22.0, 32.8, 57.6, 64.8, 70.6, 73.5, 83.5, 86.9, 98.9, 118.8, 139.4, 149.5, 152.6, 155.8.

^{31}P{^{1}H}-NMR (121 MHz, DMSO, 25 °C, H_3PO_4): δ = -0.16. MS (HR-ESI$^-$): m/z 478.1491 (calculated for $[C_{20}H_{26}N_5O_7P_1\text{-}H]^-$ 478.1497).

Incorporation of initiator nucleotides into 19mer RNA

***In vitro* transcription:** For *in vitro* transcription, a DNA template was formed by annealing two DNA fragments (5'-TCTAATACGACTCACTATA-3', 5'-GGAGTGAAGGCTG ATGCCTATAGTGAGTCGTATTAGA-3' where underlined Gs are 2'-OMe modified[31] in Tris-buffer (10 mM Tris-HCl *p*H 7.5, 80 mM NaCl) as reported earlier.[19] *In vitro* transcription was performed by mixing 0.8 µM of dsDNA template in transcription buffer (40 mM Tris *p*H 8.1, 1 mM spermidine, 22 mM $MgCl_2$, 0.01% Triton-X-100) with 10 mM DTT, 0.01 mg/mL bovine serum albumin, 4 mM of each ATP, CTP & UTP and 10 U/µL of T7 RNA polymerase (lab prepared stock). In the case of GMP initiation, 0.4 mM GTP was used along with 0.4, 2, 4 mM of the respective initiator for 1:1, 1:5; 1:10 ratio, respectively. In the case of **AMP3**, 0.4 mM ATP was used at 4 mM of other three NTPs along with 0.4, 2, 4 mM of **AMP3** for 1:1, 1:5; 1:10 ratio, respectively. To visualize the transcripts, the transcription reaction was doped with 30 µCi of radioactive CTP (^{32}P α-CTP, 10 µCi/µL, Hartmann Analytics, Germany). The transcription reaction was incubated at 37°C for 4 h. After quenching the reaction with gel loading buffer (10% TBE in formamide containing xylene cyanol and bromophenol blue), the transcripts were purified over 10% denaturing polyacrylamide gel using standard electrophoresis conditions (1X TBE buffer, run at 25 W for 1h 20 min). The gels were scanned using storage-phosphor screens and a Typhoon Imager (GE Healthcare). Transcript bands were excised, eluted in 0.3 M Na-Acetate *p*H 5.5 overnight at 25°C, ethanol precipitated and then dissolved in neutral water (Millipore water, MilliQ) for DAinv.

LC-MS analysis of the *in vitro* transcribed RNA: In order to subject transcripts to LC-MS analysis, transcriptions were performed without addition of radioactive CTP on a 400 µL scale. After reaction, DNA was digested by addition of 40 µL DNAse I buffer (10X, Fermentas) and 5 µL DNAse I (Fermentas, 50 U/µL). After incubation for 15 min at 37 °C the reaction was quenched with gel loading buffer (10% TBE in formamide containing xylene cyanol and bromophenol blue) and the transcripts were purified over a 10% denaturing polyacrylamide gel using standard electrophoresis conditions (1X TBE buffer, run at 17 W for 1h 40 min). Transcript bands were detected via UV shadowing, excised and eluted in 0.3 M Na-Acetate *p*H 5.5 overnight at 25°C. After precipitation with isopropanol the pellet was dissolved in water and 45-65 pmol of

the transcribed RNA was diluted in 50 µl of water, and 40 µl were injected into the LC-MS system (HR-ESI). Gradient used: increase of methanol from 5% to 35% over 30 min.

Incorporation ratio: To calculate the incorporation ratio of initiators at the 5'-end of 19mer RNA, at each ratio of GTP or ATP:Initiator, transcriptions were performed in triplicates. In order to resolve the incorporation of single nucleotides, the purified transcripts were separated on 15% denaturing polyacrylamide sequencing gels under standard electrophoresis conditions (1X TBE buffer, run at 40 W for 4 h 50 min). The gels were exposed to storage-phosphor screens and scanned on a Typhoon Imager (GE Healthcare). The % incorporation was calculated from these gels using ImageQuant software (Molecular Dynamics) and plotted as graph (Figure S1).

Peptide synthesis

All reagents for peptide synthesis were purchased from Novabiochem Merck. Peptide synthesis was performed on an automated peptide synthesizer (Applied Biosystems, 431 A) using standard reagents and standard protocols (Fmoc strategy). A Rink amide resin (100-200 mesh, binding capacity: 0.6-0.8 mmol/g) was used as solid support. After synthesis the solid support was dried under vacuum and then coupled to the tetrazine using NHS-tetrazine[32] (DMF, 15 h, RT). After washing with DMF (3x) and DCM (3x), the peptide was cleaved from the solid support by incubation with 95% TFA, 2.5% H_2O and 2.5% triisopropylsilane for 2 hours. The peptide was then precipitated using diethyl ether, lyophilized and then the crude product purified *via* HPLC (Gradient: 5% ACN/0.1% TFA for 5 min, increase to 52% ACN/0.1% TFA over 30 min), and the identity was confirmed by HR-ESI$^+$ (Figure S2).

Diels-Alder reaction with inverse-electron demand

DAinv between 19mer RNA and tetrazine-peptide: In order to perform DAinv on 19mer RNA, initiated RNA (**GMP2** or **GMP3** or **AMP3**) was mixed with tetrazine in water and incubated at 25°C. After incubation the reaction was analyzed by following:

LC-MS analysis (HR-ESI): To an aqueous solution of initiated RNA (c = 3.5 µM) tetrazine peptide (excess) was added and the solution incubated for 1 h. The reaction mixture was then diluted to a final volume of 50 µl directly injected into the LC-MS system (volume = 40 µl).

Gel analysis for fold-excess of tetrazine peptide required for the conjugation: For analyzing the fold excess of tetrazine peptide required for the efficient RNA-peptide conjugation, 3 µM of **GMP3** initiated RNA was mixed with different excess of tetrazine peptide (0.4-fold; 1.2 µM, 1-fold; 3 µM, 2-fold; 6 µM, 3-fold; 9 µM) in water at 5 µL reaction scale and incubated at 25°C for 10 min. After 10 min, 5 µL of gel-loading buffer was added and analyzed on 18% denaturing sequencing polyacrylamide gel under standard electrophoresis conditions (1X TBE buffer, run at 25W for 4 h 50 min). Gel was scanned over Typhoon Imager (GE Healthcare) after exposing to store-phosphor screen and analyzed by ImageQuant software (Molecular Dynamics).

Peptide conjugation for AMP3-initiated 19mer RNA: **AMP3** incorporation at 5'-end of the 20mer RNA was performed following the above mentioned *in vitro* transcription protocol using a template which has an A-initiation specific promoter 5'-GGAGTGAAGGCTGATGACCTAATAGTGAG TCGTATTAGA-3'. For DAinv reaction, 3 µM of **AMP3**-initiated 19mer RNA was mixed with 3 µM of tetrazine peptide and incubated at 25°C for 50 min in 5 µL reaction scale. After 50 min, 5 µL of gel-loading buffer was added and analyzed on 15% denaturing polyacrylamide gel under standard electrophoresis conditions (1X TBE buffer, run at 25W for 2 h 50 min). The gel was scanned on a Typhoon Imager (GE Healthcare) after exposing to storage-phosphor screen and analyzed by ImageQuant software (Molecular Dynamics). The gel analysis showed successful formation of the RNA-peptide conjugate between **AMP3**-initiated RNA and tetrazine peptide.

Peptide conjugation for a 232mer RNA: *In vitro* transcription: RNA19, a 232mer ribozyme identified by *in vitro* selection,[9b] was *in vitro* transcribed using a dsDNA template[9b] following the same protocol as above using 4 mM GMP3, 0.4 mM GTP and 30 µCi of ^{32}P α-CTP (Hartmann Analytics, Germany) on a 100 µL scale. The transcription reaction was incubated at 37°C for 4 h and purified on 8% denaturing polyacrylamide gels. Transcript bands were excised, eluted in 0.3 M Na-Acetate *p*H 5.5 overnight at 25°C, ethanol precipitated and dissolved in neutral water (Millipore water, MilliQ) for DAinv.

DAinv reaction: For DAinv, 3 µM of **GMP3** initiated 232mer RNA was mixed with 3 µM of tetrazine peptide and incubated at 25°C for 50 min. After incubation the RNA was ethanol precipitated from the reaction mixture and redissolved in neutral water (Millipore water, MilliQ) for DNAzyme cleavage.

DNAzyme cleavage of peptide conjugated RNA: To analyze the peptide conjugation on the 232mer RNA, a 10-23 DNAzyme[25] was designed which cleaves a 29mer RNA fragment from the 5'-end of RNA19 (232mer RNA, see Supplementary Information, Figure S3). For DNAzyme cleavage, 4 µM of both tetrazine peptide conjugated and non-conjugated RNA19 was mixed with 1X DNAzyme buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCL), 10-fold excess of DNAzyme (40 µM) in 10 µL reaction volume. The reaction was heated to 80°C for 30 sec, subsequently cooled down to 37°C at a rate of 0.5°C/sec and then incubated at 37°C for 3 min. To increase the amount of cleaved product, this process was cycled 5 times. After completion of reaction, 10 µL of gel loading buffer was added and cleaved 29mer peptide-attached RNA fragment was analyzed on a 18% denaturing polyacrylamide gel under standard electrophoresis conditions (1X TBE, run at 25 W for 2 h 20 min). The gel was scanned on a Typhoon Imager (GE Healthcare) after exposing to storage-phosphor screen and analyzed by ImageQuant software (Molecular Dynamics).

Acknowledgements

The authors thank Prof. M. Wießler (DKFZ Heidelberg) for providing the respective alcohol for synthesis of **GMP2**.

Notes and references

[a] [a]*Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Im Neuenheimer Feld 364, Heidelberg 69120, Germany. Fax: +49 (0)6221 - 54 6430; Tel: +49 (0)6221 - 54 4851*

‡Both authors contributed equally. *E-mail: jaeschke@uni-hd.de

.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

1. M. J. Fedor and J. R. Williamson, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 399-412.
2. A. Roth and R. R. Breaker, *Annu. Rev. Biochem.*, 2009, **78**, 305-334.
3. J. S. Mattick, *FEBS Lett.*, 2011, **585**, 1600-1616.
4. A. E. Vazquez, L. Nie and E. N. Yamoah, *Methods Mol. Biol.*, 2009, **493**, 21-29.
5. S. Tyagi, *Nat. Methods*, 2009, **6**, 331-338.
6. H. Moroder, J. Steger, D. Graber, K. Fauster, K. Trappl, V. Marquez, N. Polacek, D. N. Wilson and R. Micura, *Angew. Chem. Int. Ed.*, 2009, **48**, 4056-4060.
7. J. H. Jeong, H. Mok, Y. K. Oh and T. G. Park, *Bioconjugate Chem.*, 2009, **20**, 5-14.
8. D. N. Wilson and K. H. Nierhaus, *Angew. Chem. Int. Ed.*, 2003, **42**, 3464-3486.
9. (a) S. Pfander, Ph.D. Thesis, Heidelberg University, 2008; (b) S. Ameta and A. Jäschke, *Chem. Sci.*, 2013, **4**, 957-964; (c) S. Ameta, Ph.D. Thesis, Heidelberg University, 2013.
10. S. Pfander, R. Fiammengo, S. I. Kirin, N. Metzler-Nolte and A. Jäschke, *Nucleic Acids Res.*, 2007, **35**.
11. (a) E. Paredes and S. R. Das, *ChemBioChem*, 2011, **12**, 125-131; (b) M. L. Winz, A. Samanta, D. Benzinger and A. Jäschke, *Nucleic Acids Res.*, 2012, **40**, e78; (c) A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, **39**, 1388-1405; (d) P. M. Gramlich, C. T. Wirges, A. Manetto and T. Carell, *Angew. Chem. Int. Ed.*, 2008, **47**, 8350-8358; (e) F. Seela and V. R. Sirivolu, *Nucleos. Nucleot. Nucl.*, 2007, **26**, 597-601.
12. (a) E. M. Sletten and C. R. Bertozzi, *Acc. Chem. Res.*, 2011, **44**, 666-676; (b) M. Shelbourne, X. Chen, T. Brown and A. H. El-Sagheer, *Chem. Commun*, 2011, **47**, 6257-6259.
13. (a) D. L. Boger, *Chem. Rev.*, 1986, 781-793; (b) S. N. Kessler and H. A. Wegner, *Org. Lett.*, 2010, **12**, 4062-4065.
14. E. D. Anderson and D. L. Boger, *Org. Lett.*, 2011, **13**, 2492-2494.
15. J. Sauer, D. K. Heldmann, J. Hetzenegger, J. Krauthan, H. Sichert and J. Schuster, *Eur. J. Org. Chem.*, 1998, 2885-2896.
16. M. Wiessler, E. Müller, P. Lorenz, C. Kliem and H. Fleischhacker, *Patent, WO/2007/144200*.
17. (a) N. K. Devaraj and R. Weissleder, *Acc. Chem. Res.*, 2011, **44**, 816-827; (b) D. S. Liu, A. Tangpeerachaikul, R. Selvaraj, M. T. Taylor, J. M. Fox and A. Y. Ting, *J. Am. Chem. Soc.*, 2012, **134**, 792-795; (c) A. C. Knall and C. Slugovc, *Chem. Soc. Rev.*, 2013, **42**, 5131-5142.
18. J. Schoch, M. Wiessler and A. Jäschke, *J. Am. Chem. Soc.*, 2010, **132**, 8846-8847.
19. J. Schoch, S. Ameta and A. Jäschke, *Chem. Commun.*, 2011, **47**, 12536-12537.
20. M. L. Blackman, M. Royzen and J. M. Fox, *J. Am. Chem. Soc.*, 2008, **130**, 13518-13519.
21. M. T. Taylor, M. L. Blackman, O. Dmitrenko and J. M. Fox, *J. Am. Chem. Soc.*, 2011, **133**, 9646-9649.
22. F. Thalhammer, U. Wallfahrer and J. Sauer, *Tetrahedron Lett.*, 1990, **31**, 6851-6854.
23. (a) J. F. Milligan, D. R. Groebe, G. W. Witherell and O. C. Uhlenbeck, *Nucleic Acids Res.*, 1987, **15**, 8783-8798; (b) J. A. Pleiss, M. L. Derrick and O. C. Uhlenbeck, *RNA*, 1998, **4**, 1313-1317.
24. F. Huang, G. Wang, T. Coleman and N. Li, *RNA*, 2003, **9**, 1562-1570.
25. (a) M. J. Cairns, A. King and L. Q. Sun, *Nucleic Acids Res.*, 2003, **31**, 2883-2889; (b) R. R. Breaker and G. F. Joyce, *Chem. Biol.*, 1994, **1**, 223-229.
26. R. van Geel, G. J. Pruijn, F. L. van Delft and W. C. Boelens, *Bioconjugate Chem.*, 2012, **23**, 392-398.
27. I. Nikic, T. Plass, O. Schraidt, J. Szymanski, J. A. Briggs, C. Schultz and E. A. Lemke, *Angew. Chem. Int. Ed.*, 2014.
28. R. Fiammengo, K. Musilek and A. Jäschke, *J. Am. Chem. Soc.*, 2005, **127**, 9271-9276.
29. J. Dommerholt, S. Schmidt, R. Temming, L. J. Hendriks, F. P. Rutjes, J. C. van Hest, D. J. Lefeber, P. Friedl and F. L. van Delft, *Angew. Chem. Int. Ed.*, 2010, **49**, 9422-9425.
30. X.-F. Zhu, H. J. Williams and A. I. Scott, *J. Chem. Soc., Perkin Trans.*, 2000, 2305-2306.
31. C. Kao, M. Zheng and S. Rudisser, *RNA*, 1999, **5**, 1268-1272.
32. J. Schoch, M. Staudt, A. Samanta, M. Wiessler and A. Jäschke, *Bioconjugate Chem.*, 2012, **23**, 1382-1386.