

# Stabilization of RNA hairpins using non-nucleotide linkers and circularization

Agnieszka Kiliszek<sup>†</sup>, Leszek Błaszczuk<sup>†</sup>, Ryszard Kierzek and Wojciech Rypniewski\*

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Received November 29, 2016; Revised January 18, 2017; Editorial Decision February 08, 2017; Accepted February 09, 2017

## ABSTRACT

**An RNA hairpin is an essential structural element of RNA. Hairpins play crucial roles in gene expression and intermolecular recognition but are also involved in the pathogenesis of some congenital diseases. Structural studies of the hairpin motifs are impeded by their thermodynamic instability, as they tend to unfold to form duplexes, especially at high concentrations required for crystallography or nuclear magnetic resonance spectroscopy. We have elaborated techniques to stabilize the RNA hairpins by linking the free ends of the RNA strand at the base of the hairpin stem. One method involves stilbene diether or hexaethylene glycol linkers and circularization by T4 RNA ligase. Another method uses click chemistry to stitch the RNA ends with a triazole linker. Both techniques are efficient and easy to perform. They should be useful in making stable, biologically relevant RNA constructs for structural studies.**

## INTRODUCTION

An RNA hairpin structure is the most abundant motif in RNA molecules. In 16S rRNA nearly 70% of the nucleotides form hairpins (1). Although its architecture is simple (consisting of a stem and apical loop), hairpins exhibit substantial structural complexity. The main differentiating factor is the hairpin stem which can be fully complementary and double-stranded or contain bulges, internal loops or even a quadruplex platform (2). The apical loop exhibits various architecture depending on size and nucleotide content. These components determine the structural diversity and function of RNA hairpins and makes them important research objects and molecular tools in biomedicine.

As independent elements RNA hairpins are involved in attenuation of gene expression in bacteria by forming stable structures that cause a premature termination of transcription. When located in the 3' untranslated regions of mRNA they protect transcripts from degradation. Short hairpins are engaged in subcellular localization of some mRNA, in

promoting correct cleavage of mRNA or in viral replication (1,3). As parts of larger structures hairpins are important components of riboswitches, regulators of gene expression whose structure and function is modulated by interactions with specific metabolites or cofactors (4). In internal ribosome entry sites of specific cellular and viral mRNAs, hairpins form platforms for interaction with protein factors, participating in the cap-independent translation (5). Hairpin structures are also precursors of microRNAs which are likely the most common RNA silencing inducers in mammals (6). Finally, as parts of rRNA, hairpins associate with a number of proteins to form ribosomes (7).

The structural and functional richness of RNA hairpins results in many applications. They can be aptamers—*in vitro* selected oligomers exhibiting high affinity and selectivity for target molecules—used in molecular biology, medicine (therapy and diagnostics) and as biosensors to detect drugs or pollution in food and the environment (8–11). Engineered RNA hairpins have also been implemented in the assembly of nanomaterials by utilizing loop–loop interactions between two hairpin motifs that have been taken from naturally existing molecules such as pRNA of the bacteriophage phi29 DNA-packaging motor or the retroviral kissing loop hairpins (12).

Hairpin structures are also involved in the development of many neurodegenerative disorders such as myotonic dystrophy type 1 and 2, syndromes linked with fragile X chromosome, several spinocerebellar ataxias or Huntington's disease. The major subset of pathogenic sequences are the trinucleotide CNG repeats (N stand for one of the four natural nucleotides) but also tetra-, penta- or hexanucleotides belongs to this group (13). The number of repeats can undergo abnormal expansion from several to even thousands units. The RNA containing expanded runs forms stable hairpin structures (14). It has been shown that these structures disorganize cellular processes by sequestering important proteins into insoluble foci (15,16). This exemplifies negative consequences of RNA's natural ability to fold.

In view of their biological importance, RNA hairpins are interesting research objects. Although a number of biophysical methods allow characterization of hairpin properties, they are not easily available for structural studies including

\*To whom correspondence should be addressed. Tel: +48 61 852 8503; Fax: +48 61 852 0532; Email: wojtekr@ibch.poznan.pl

<sup>†</sup>These authors contributed equally to the paper as first authors.

crystallography. The main reason is that hairpin structures are thermodynamically unstable under high concentrations of RNA and/or salt required for crystallization. Under such conditions they tend to unfold and form duplexes. This is a thermodynamically driven process and altering the sequence does not help because any hairpin, having complementary ends, can equally well form a duplex with another such strand (17,18). This instability is reflected in the near-absence in literature of crystal structures of RNA hairpins. In the Protein Data Bank (PDB) or in the Nucleic Acid Data Base there are only a dozen or so crystal models of isolated RNA hairpins. In our recent work on the structure of CUG repeats we crystallized an oligomer consisting of eight CUG repeats and obtained a duplex instead of a hairpin (PDB code: 5MWI).

In order to overcome the difficulties in structural studies of RNA hairpins we have developed, tested and optimized techniques to stabilize RNA in the hairpin form. The stabilization has been achieved by using circularization and non-nucleotide linkers. The hairpin structure has been circularized by enzymatic or click chemistry approaches. The enzymatic method is based on ligation by T4 RNA ligase 1 which joins the 5' phosphate and 3' OH groups (19–21). While particular structural and sequential requirements must be fulfilled for efficient ligation, this method preserves the native backbone of the RNA hairpin. The click chemistry involves linking azide and alkynes groups in the presence of Cu(I) ions (22). It is well known for its high efficiency but requires an introduction of functional groups into the oligomer.

The non-nucleotide linkers are located on the opposite site of the apical loop forming the base of the hairpin. The non-nucleotide spacers has been shown to increase the thermal stability of the structure of duplexes, triplexes or tetraplexes (23–29). They were also implemented in nuclear magnetic resonance (NMR) (30–34) and crystallographic studies (29,35,36) but so far such linkers has not been used to stabilize the native structure of RNA hairpins.

## MATERIALS AND METHODS

### Monoalkyne solid support and bromophosphorodiamidites for click oligomers

Preparation of the monoalkyne solid support and bromophosphorodiamidites was carried out according to Lietard *et al.* (37). The only difference was that one more step was added to obtain the modified support in which a succinic group was incorporated in 1-Propargyl-2-[(4,4'-dimethoxytrityl)oxymethyl]-2-methylpropane-1,3-diol and then the LCAA-CPG support was coupled.

### DNA/RNA synthesis and purification

Oligonucleotides were chemically synthesized in an Applied Biosystems DNA/RNA synthesizer, using the cyanoethyl phosphoramidite chemistry or purchased from Future Synthesis. All nucleotides and hexaethylene glycol phosphoramidites with 2'-O-tertbutyldimethylsilyl were purchased from Glen Research, Azco, Proligo. The stilbene diether linker was synthesized using a custom synthesis service. Before the deprotection step, the click oligomers were treated with LiN<sub>3</sub> (300eq) in dimethylformamide (DMF) for 2 h at

65°C. RNA oligomers were cleaved from the solid support using ammonium hydroxide/ethanol (3:1 v/v) at 55°C for 16 h. Deprotection was carried out with triethylamine trihydrofluoride at 55°C for 3 h followed by the *n*-butanol precipitation for 1 h at 4°C. Oligomers were desalted on illustra NAP-25 columns (GE Healthcare) and purified by gel electrophoresis in denaturing conditions. In the case of the DNA oligomer it was cleaved and deprotected by ammonium hydroxide at 55°C for 16 h, desalted and purified by the denaturing gel electrophoresis.

### T4 RNA ligase 1 expression and purification

The plasmid containing T4 RNA ligase gene was a generous gift from Prof. Peter J. Unrau from Simon Fraser University in Canada (38). The T4 RNA ligase open reading frame was re-cloned into pMCSG9 vector (Midwest Center for Structural Genomics) containing an N-terminal His6 tag. The construct with pMCSG9 T4 RNA ligase gene was obtained by ligase-independent cloning (39). The identity of the clone was confirmed by sequencing. The enzyme was produced in *Escherichia coli* BL21 Magic cells in Luria-Bertani medium (Midwest Center for Structural Genomics). The expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and carried out for 16 h at 19°C. The cells were harvested by centrifugation (4000 g, 15 min, 4°C) and resuspended in the lysis buffer (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20 mM imidazole, 10 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub>, 5% glycerol, 1 mM tris(2-carboxyethyl)phosphine (TCEP) and a protease inhibitors cocktail). The cells were lysed by sonication: 2 s pulse followed by 10 s pause (25 min at 4°C). Cell debris was removed by centrifugation (15 000 g, 30 min, 4°C). The supernatant was mixed with 6 ml of Ni Sepharose High Performance (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 10 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> and 5% glycerol. Sepharose beads were washed with 50 ml of buffer I (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 40 mM imidazole, 10 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub>, 5% glycerol, 1 mM TCEP) and 200 ml of buffer II (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 40 mM imidazole, 5% glycerol, 1 mM TCEP). The protein was eluted with 2 × 10 ml of buffer III (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 300 mM imidazole, 5% glycerol, 1 mM TCEP) and dialyzed into 50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20% glycerol and 1 mM TCEP and next into the storage buffer (20 mM HEPES-KOH pH 7.5, 50 mM KCl, 1 mM dithiothreitol (DTT) and 50% glycerol). The purity of the recombinant protein was assessed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).

### Circularization of RNA oligomers using T4 RNA ligase 1

RNA oligomers (20–50 μM) were denatured in 95°C for 2 min in 50 mM KCl and slowly cooled to 25 or 37°C. The ligation reaction was carried using 0.25–0.40 mg/ml T4 RNA ligase 1 in buffer (50 mM Tris pH 7.5, 10 mM MgCl<sub>2</sub> and 10 mM DTT) with 0.5 mM ATP at 37°C for 4 h. The solution was desalted using Bio-Gel polyacrylamide P-6 columns (Biorad) and resolved on 6–10% denaturing gel electrophoresis. Gels were stained using toluidine blue or SYBR Green II. For the UV melting measurements

circularized oligomers were eluted by the crash-and-soak method, desalted and concentrated using centrifugal filters.

### RNA dephosphorylation and $^{32}\text{P}$ -labeling

The dephosphorylation reaction of linear or circular RNA was performed using alkaline phosphatase (Thermo) at  $37^\circ\text{C}$  for 10 min in a buffer containing 100 mM Tris-HCl pH 8.0, 50 mM  $\text{MgCl}_2$ , 1 M KCl, 0.2% Triton X-100 and 1 mg/ml Bovine serum albumin. The enzyme was deactivated by heating at  $75^\circ\text{C}$  for 10 min. The RNA was either ethanol precipitated or loaded directly on the gel.

RNA oligomers were 5'-end labeled with  $[\gamma\text{-}^{32}\text{P}]$  ATP (Hartmann-Analytic) using T4 Polynucleotide Kinase (Thermo). Prior to reaction the RNA was denatured for 5 min at  $90^\circ\text{C}$  and incubated on ice for 10 min. The reaction was carried out in a buffer containing 50 mM Tris-HCl pH 7.6, 10 mM  $\text{MgCl}_2$ , 5 mM DTT, 0.1 mM spermidine supplemented with  $1\ \mu\text{l}$  of  $[\gamma\text{-}^{32}\text{P}]$  ATP (0.02 mCi) and 10 U of enzyme. The reaction was incubated for 30 min at  $37^\circ\text{C}$ . Labeled RNAs were purified by denaturing gel electrophoresis and quantified using a scintillation counter.

### Alkaline RNA hydrolysis

One volume of 5'  $^{32}\text{P}$ -labeled RNA was mixed with five volumes of 10 mM  $\text{MgCl}_2$  in formamide and incubated at  $100^\circ\text{C}$  for 10 min. The reaction was stopped by adding one volume of loading buffer (formamide, 0.02% xylene cyanol) and frozen on dry ice. The reaction products were separated on denaturing gels and visualized using autoradiography with an intensifying screen on Fluor Imager FLA-5100 (FujiFilm).

### UV melting of oligonucleotides

UV thermal melting studies were performed according to Pasternak and Wengel (2011) (40) on DU-640 spectrometer with a thermoprogrammer (Beckman). RNA oligomers were dissolved in a buffer containing 10 mM sodium chloride, 20 mM sodium cacodylate and 0.5 mM disodium ethylenediaminetetraacetate (EDTA), pH 7.0 or 5.2. Each oligomer was prepared in nine different concentrations in the range  $10^{-5}$ – $10^{-6}$  M. Concentrations of single-

stranded oligomers were calculated from the high temperature ( $>80^\circ\text{C}$ ) absorbance and single strand extinction coefficients approximated by the nearest-neighbor model. The UV absorption versus temperature was measured at 260 nm at the heating rate of  $1^\circ\text{C}/\text{min}$  in the range of  $20$ – $90^\circ\text{C}$ . The melting curves were analyzed and the thermodynamic parameters calculated using MeltWin 3.5.

### Click reaction

Prior to reaction 1 nmol of DNA or RNA oligomer was denatured in water for 2 min at  $95^\circ\text{C}$  and placed on ice for 10 min. The reaction was carried out in 200 mM NaCl,  $500\ \mu\text{M}$  Copper(II)-tris(benzyltriazolylmethyl)amine (Cu(II)-TBTA) complex and freshly prepared 5 mM ascorbic acid. The click reaction was incubated at  $37^\circ\text{C}$  for 4 h. The reaction was desalted using Bio-Gel polyacrylamide P-6 columns and analyzed on 10–15% denaturing polyacrylamide gels.

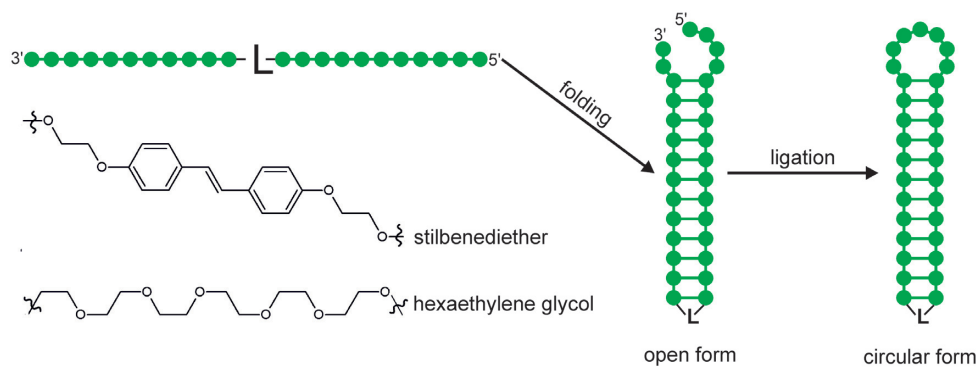
### HpaII digestion

Linear or circular DNA oligomer ( $10\ \mu\text{g}$ ) was digested in a buffer containing  $1\times$  Tango buffer and 40U of HpaII enzyme at  $37^\circ\text{C}$  overnight. The enzyme was deactivated by heating in  $65^\circ\text{C}$  for 20 min. The reaction mixture was desalted using Bio-Gel polyacrylamide P-6 columns or precipitated using lithium chloride and five volumes of EtOH/acetone (1:1 v/v) mix. The reaction products were separated on 20% denaturing polyacrylamide gel and visualized by UV shadowing.

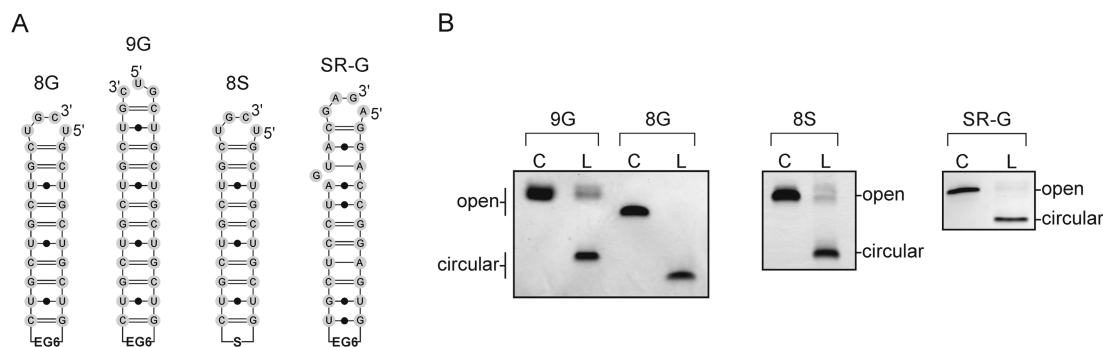
## RESULTS

### Circularization by T4 RNA ligase 1

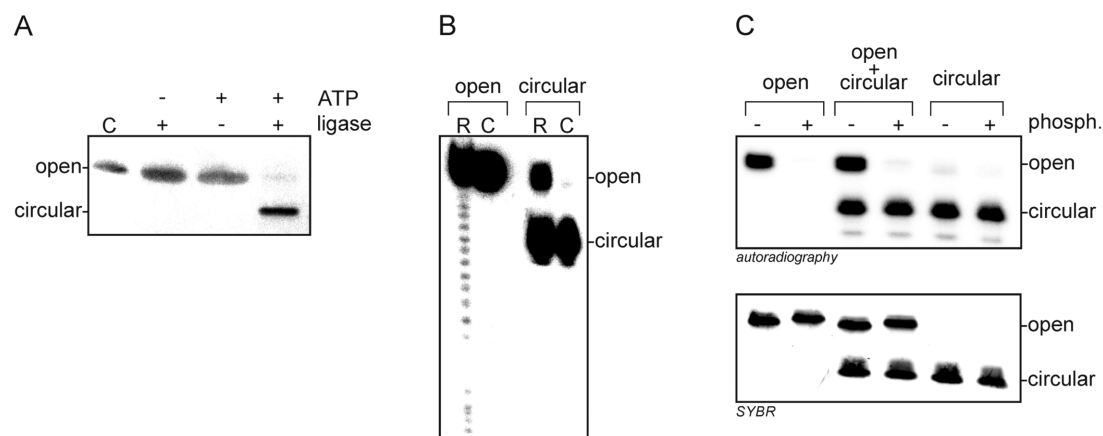
In the first method, we circularized RNA using T4 RNA ligase 1. The tested oligomers were composed of two stretches of RNA connected with a non-nucleotide linker: hexaethylene glycol (EG6) or stilbene diether (Figure 1). The EG6 linker is composed of six ethylene glycol units while stilbene diether contains two aromatic rings. The oligomers were designed to fold into a duplex having the non-nucleotide linker at one side and the 5' and 3' ends on the opposite side, creating a ligation region. To facilitate the self-assembly of the



**Figure 1.** Stabilization of the hairpin structure by non-nucleotide linkers (-L-) following enzymatic circularization. Structures of the stilbene diether and hexaethylene glycol (EG6) linkers are depicted.



**Figure 2.** Circularization of RNA using T4 RNA ligase. (A) Schematic representation of RNA substrates: three oligomers containing CUG repeats and a sarcin-ricin (SR-G) sequence from *Escherichia coli* 23S rRNA. (B) Control (C) and ligation reactions (L) resolved on an 8% polyacrylamide gel in denaturing conditions. Open and circular forms of the RNA oligomers are indicated.



**Figure 3.** Identification of the circular form of the RNA oligomer. (A) Circularization reaction was carried out in the absence of ATP or T4 RNA ligase and separated on an 8% polyacrylamide gel. C—RNA oligomer only. (B) Autoradiogram shows separation of the products of alkaline hydrolysis to the linear and circular forms of the RNA oligomer (R). C—control reaction. (C) Identification of the circular RNA form using dephosphorylation. The reaction products were separated on an 8% polyacrylamide gel and visualized by autoradiography followed by staining with SYBR Green II.

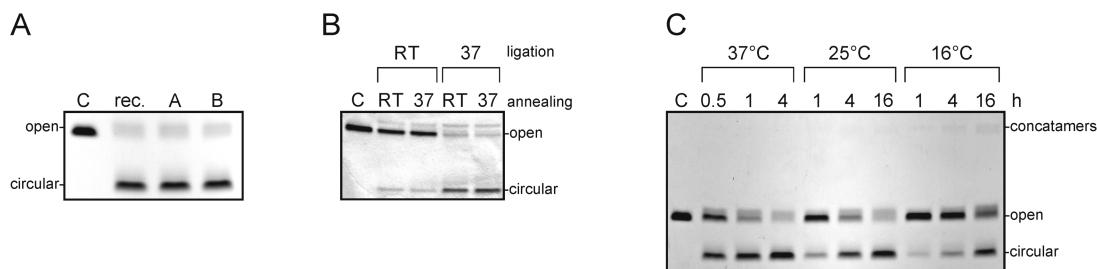
RNA oligomer into the hairpin form the RNA part contained CUG repeats which were earlier shown to fold into hairpin structures having a stem composed of blocks of G–C and C–G pairs interrupted by the non-canonical U–U pairs (41,42) (Figure 2A). Among the CNG repeats the CUG repeats are intermediate in terms of stability, less stable than CGG, similar to CAG and more stable than CCG (42,43). The sequence is also intermediate in terms of meeting the T4 RNA ligase requirements, the ligation site possessed a U residue at the 5' end (donor) and a C residue at the 3' end (acceptor) (19–21). The uracil contained a 5' mono-phosphate group which was incorporated during the chemical synthesis. After the enzymatic circularization the apical loop of the hairpin was reconstructed, consisting of 3 or 4 nt, CUG or UGCU, depending on whether the oligomer consisted of an odd or even number of CUG repeats (Figures 1 and 2A).

The ligation products were separated in the denaturing polyacrylamide gels. It turned out that circularization using T4 RNA ligase was very efficient since more than 95% of the open oligomer had been converted into the circular form (Figure 2B). Additionally, migration of the circularized oligomers was accelerated in relation to the open

substrate, in agreement with formation of a more compact structure.

In order to test the applicability of this circularization method to sequences other than tri-nucleotide repeats, the reaction was performed on a SR-G hairpin from the *E. coli* 23SrRNA (44) and a similar high yield was achieved (Figure 2).

Circularization of the RNA oligomers was confirmed in three types of tests. In the first test, the circularization reaction was carried out either without adenosinetriphosphate (ATP) or the enzyme (Figure 3A). In the absence of one of the essential components the RNA migrated similarly to the control reaction indicating the presence of only the linear substrate. In the second test, the  $^{32}\text{P}$ -labeled linear substrate or the circularized product were partially hydrolyzed in the presence of  $\text{Mg}^{2+}$  ions (Figure 3B). The reaction conditions were chosen to break only one phosphodiester bond in the RNA oligomer at a time. The hydrolysis of the linear substrate resulted in a 'ladder' of radioactive RNA fragments of different lengths. In the case of the circularized product we observed only some additional band whose migration was similar to the linear substrate. This indicated that the covalent bond in the circular oligomer was broken forming a mixture of linear RNA oligomers of the same length. The



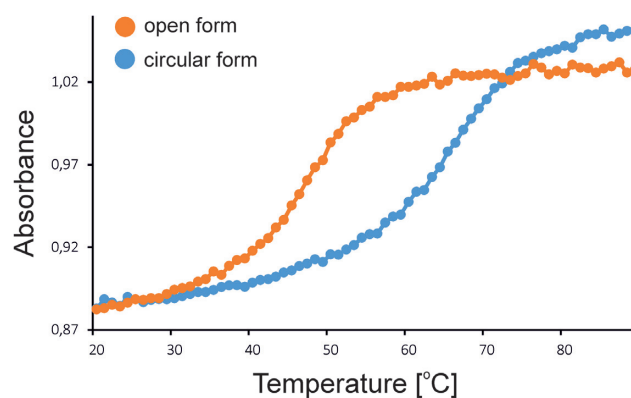
**Figure 4.** (A) Circularization efficiency of RNA using recombinant (rec.) or commercially available enzymes (lanes A and B). The ligation efficiency is not affected by the annealing temperature (B) but depends on the ligation temperature. (C) Open and circular forms of the RNA oligomer are indicated.

goal of the third test was to remove the 5'  $^{32}\text{P}$  phosphate group using alkaline phosphatase (Figure 3C). The enzyme catalyzes the release of the 5'- and 3'-phosphate groups from linear DNA, RNA and nucleotides. Thus, radioactive signal should disappear only in case of the linear substrate having an accessible 5' phosphate group. The autoradiography of the polyacrylamide gel revealed that the radioactive signal was no longer detected when the linear  $^{32}\text{P}$ -labeled was used which indicated that the lower migrating band was in fact a circular form. We confirmed this result by staining the gel using SYBR Gold which detected both linear and circular forms of the RNA oligomer.

Milligram quantities of RNA are usually needed for crystallization. Thus, a large quantity of T4 RNA ligase is necessary. In order to lower the cost we produced the enzyme in the laboratory. Typically 8–10 mg of T4 RNA ligase was obtained from 500 ml of cell culture. The activity of the enzyme was comparable with the commercially available enzymes (Figure 4A). The ligation reaction was performed according to well-established protocols (19). We found that the efficiency of circularization was insensitive to renaturation conditions but affected by the temperature of the ligation reaction (Figure 3B and C). The maximum concentration of RNA in the ligation reaction was 50  $\mu\text{M}$  in order to avoid formation of duplexes at higher concentrations. Typically over 90% RNA molecules were converted into the circular form.

### Thermal stability of circularized RNA

We developed the method in order to stabilize thermodynamically RNA in a hairpin form. Thus, it was important to characterize the effect of circularization on the thermodynamic stability of RNA oligomers. UV melting studies were conducted on two RNA oligomers (8G and 9G) containing eight and nine CUG repeats, respectively (Figure 2A). Both oligomers possessed the EG6 linker. The thermodynamic parameters: enthalpy ( $\Delta H$ ), entropy ( $\Delta S$ ) and free energy ( $\Delta G_{37}$ ) were calculated from fits of individual melting curves (Table 1). It turned out that thermodynamic stability of circularized oligomers increased which is reflected in a change of the free energy  $\Delta G_{37}$  by 1.94 kcal/mol for 8G and 0.8 kcal/mol for 9G. Consequently, the melting point of the circular oligomers is 20°C higher than for the open form (Figure 5). Moreover, the melting temperature of the open and circular forms was concentration independent confirm-



**Figure 5.** The normalized UV melting curves of the open (orange) and circular (blue) form of the 8G RNA oligomer.

ing the presence of the hairpin structure in experimental conditions.

### Purification of circularized RNA

A suitable method of purification of the circular RNA from the linear substrate should be selected as a compromise between quantity and quality of the pure product. This is important especially in terms of resolution since the only difference between the circular and linear forms of the RNA is a single oxygen and a hydrogen atom which are released during the ligation. The most common large-scale purification technique in molecular biology is PAGE in denaturing conditions. The advantage of this method is low cost, good resolution and simplicity. However, significant drawback is the yield of recovery of the material from the gel matrix which can be low for larger RNA molecules (>40 nt). Alternatively, chromatographic techniques are used which are more efficient but selection of an appropriate column, packing material and separation conditions can be time consuming and difficult to establish. We employed both approaches and only the PAGE allowed unequivocal separation of the circular product from the linear substrate. We were able to purify the circular form even on low percentage gels (6%). The accelerated migration of the circular form of the RNA oligomer suggested that despite denaturing conditions the hairpin form was preserved. Purification using RP-HPLC, gel filtration with Sephadex 75, 200 or hydroxyapatite in different separation conditions gave no positive results as both

**Table 1.** Thermodynamic parameters of open and circular forms of 8G and 9G oligomers

Oligomer	$-\Delta H$ (kcal/mol)	$-\Delta S$ (eu)	$\Delta G_{37}$ (kcal/mol)	$T_m$ ( $^{\circ}C$ )
8G				
open	$51.8 \pm 2.1$	$161.5 \pm 6.5$	$1.78 \pm 0.08$	48.0
closed	$41.4 \pm 2.0$	$121.4 \pm 6.0$	$3.72 \pm 0.17$	67.6
9G				
open	$58.3 \pm 2.6$	$182.1 \pm 8.1$	$1.84 \pm 0.10$	47.1
closed	$28.9 \pm 1.5$	$84.7 \pm 4.4$	$2.64 \pm 0.16$	68.1

circular and linear form migrated as a single peak. This is probably due to the same length, molecular weight and nucleotide composition of the circular and linear forms.

### Circularization by click chemistry

The second method used for oligomer cyclization uses click chemistry. The reaction involves the 1,3-dipolar cycloaddition between an alkyne and an azide function in the presence of Cu(I) ions resulting in a triazole linkage (Figure 6). We used the method described by Morvan *et. al.* (37) which allows incorporation of alkyne and an azide functions into an oligomer during the chemical synthesis in a sequence-independent manner. In the first step, solid support is coupled with tris(hydroxymethyl)ethane containing an alkyne group. Subsequently, standard phosphoramidite elongation cycles are performed and likewise, the bromohexyl phosphoramidite is introduced at the 5' end of the oligomer. In the next step the azide group substitutes the bromine atom and the oligomer is ready for deprotection and purifications.

We tested the efficiency of the click method on a DNA oligomer. The DNA substrate was designed to fold into a hairpin having a stem composed of G–C and C–G pairs and a poly-T apical loop (Figure 7B). In order to provide conformational flexibility during the triazole linkage formation two mismatched thymidine residues were present at the 5' and 3' end of the hairpin. The click reaction was carried out according to Paredes and Das (45) in the 200 mM NaCl solutions with the of Cu(II)–TBTA complex reduced to Cu(I) state by ascorbic acid. The reaction mixture was desalted and separated in a polyacrylamide gel. The circularization of the oligomer resulted in a non-nucleotide triazole linkage at the hairpin base. More than 90% of the linear oligomer was converted to the circular form (Figure 7A). Similarly to the ligation method, the closed form was characterized by an accelerated migration in the gel.

In order to confirm the presence of the circular form of the DNA hairpin the double-stranded stem was designed into a recognition site for HpaII restriction enzyme (Figure 7B). The pattern of restriction products resolved on the polyacrylamide gel allowed distinguishing the circular and linear specimens. The HpaII enzyme should cut the open form into three fragments of 5, 7 and 16 nt (Figure 7B). However, we detected only two fragments (7 and 16 nt) which suggested that the shortest fragment was lost during the precipitation step or co-migrated with the 7-nt fragment (Figure 7C). Performing precipitation of the reaction products with acetone, helping recovery of short oligomers, did not allow identification of the 5-nt fragment. In the case of circular DNA, two digestion products are observed (12 and

16 nt long) confirming the presence of the circular form of DNA hairpin (Figure 7C).

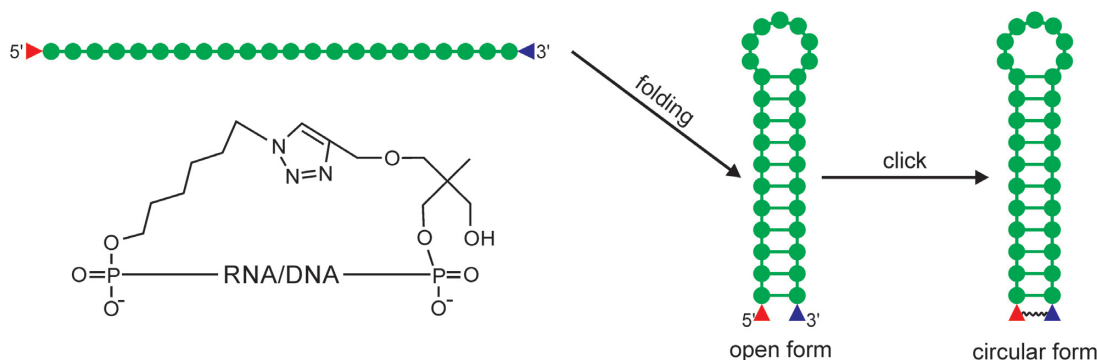
Having established the click reaction for DNA, we applied it for circularization of RNA oligomers. The RNA oligomer possessed the same sequence as its DNA counterpart. It turned out that the RNA was almost entirely converted into the circular form which indicated that click circularization could be applied to both DNA and RNA (Figure 7A).

### DISCUSSION

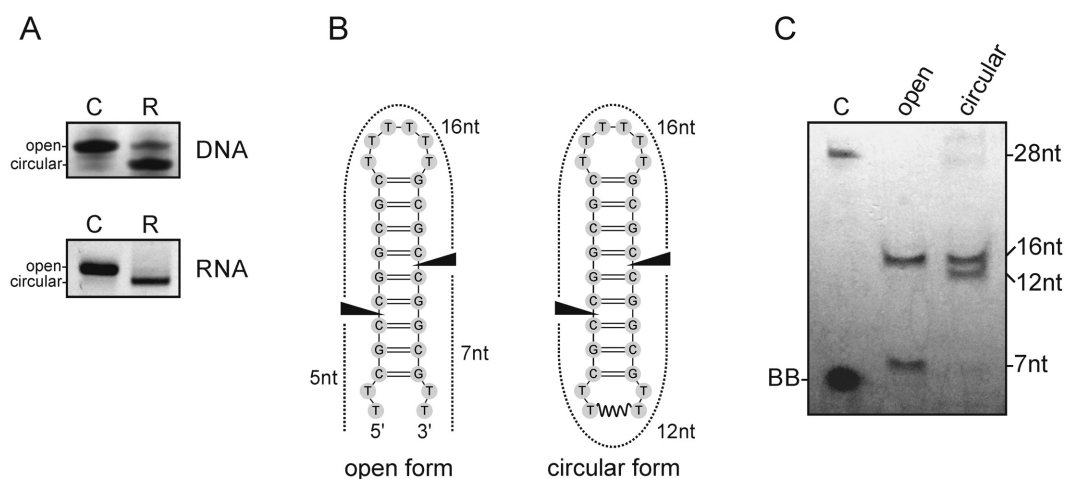
In this study we presented two alternative methods for stabilizing of RNA oligomers in the hairpin form which can be used in biochemical and structural studies. At high concentrations required for crystallization, hairpins tend to unfold and form duplexes which is a serious impediment in crystallographic studies of this important class of biological structures and is reflected in a near-absence in literature of crystal structures of RNA hairpins. In our methods, the hairpin structure is stabilized by cross-linking the ends of the stem by a non-nucleotide linker. Next, the oligomer is circularized and thus unable 'to open' and switch to the duplex form. The RNA is closed using either an enzymatic or chemical reaction.

Engineering RNA molecules for crystallographic studies is a common procedure. Many different constructs are usually tested before useful crystals are obtained. The RNA constructs can be a truncated form of the wild type RNA, can have removed or mutated loops or a sequence for binding a helper protein. Our approach adds another possibility in engineering RNA which allows preserving the native sequence and structure of the hairpin stem and apical loop.

Introducing the EG6 spacer into the RNA chain was dictated by several reasons. Biochemical studies showed that if EG6 was linking the ends of the duplex, it did not influence its native structure and biological activity (25,30). Moreover, the thermodynamic stability of the duplex increases compared to that without EG6 (27,28). The other advantage is that the linker is chemically stable and does not increase the price or lower the yield of chemical synthesis. Also the length of the polyethylene glycol spacer can be modulated according to needs. For double-stranded nucleic acids the best length is 5–7 ethylene units. The above characteristics for the EG linkers can be applied to the RNA hairpins which share the common motif with the duplexes—the double-helical stem. In the literature there are few examples where EG6 was used in structural studies. To date, the spacer was used only in NMR studies showing a stabilization effect on the DNA duplexes and a triplex structure (30–34). The NMR models revealed that EG6 did not interfere



**Figure 6.** Stabilization of the hairpin structure by the click reaction. Red and blue triangles represent the 5' and 3' end modifications required for circularization using the click reaction. The triazole linkage is depicted.



**Figure 7.** Circularization of DNA and RNA oligomers using the click chemistry. (A) Separation of the click reaction products (R) along with the linear form of the oligomer (C). (B) Schematic representation of the open and closed forms of DNA oligomer. The HpaII restriction sites and length of the digestion products are marked. (C) 20% polyacrylamide gel electrophoresis (PAGE) of the restriction products of open and circular forms of the DNA oligomer. C—control lane.

with the native DNA structure and showed conformational flexibility.

The stilbene diether linker has also been used in biochemical and structural studies (24,35). Thermodynamic studies revealed unusual stabilization properties of this linker. The melting temperatures of short DNA duplexes composed of A–T pairs and containing a stilbene diether linker reached 30–66°C while for unmodified duplexes did not exceed 0°C. Stilbene diether linker was successfully used in crystallographic studies of a DNA duplex (35). It turned out that the aromatic groups of the linker formed stacking interactions with neighboring nucleobases or symmetry related molecules, increasing crystallization potential.

Our biochemical and thermodynamic experiments have confirmed that a circularized hairpin structure is formed using both methods. In the ligation method the presence of the non-nucleotide linker probably enhanced the proper folding of the ‘open’ form, by bringing the free 5' and 3' ends close to each other. It is likely to have increased the ligation efficiency and prevented the formation of undesired products, such as concatamers. UV melting experiments confirmed that the open duplex and circularized hairpin, both hav-

ing a non-nucleotide linker, maintained their structure since the melting point was concentration-independent. However, the circularized RNA hairpin possessed substantially higher thermodynamic stability than the open form which was reflected in lower free energy values accompanied by a concomitant increase of the melting temperature by 20°C. The higher thermodynamic stability of the circular form is also supported by its migration in gel matrix. Accelerated migration suggests that the circular hairpin form is maintained despite the presence of 8 M urea.

Nucleotide linkers could be used as an alternative to the non-nucleotide linkers proposed in this work. However, due to the natural propensity of nucleotides to interact with other nucleotides, they would have to be designed carefully to minimize the possibility of undesirable interference with the native sequence. One example is a linker consisting of four thymidine residues (T4 loop), designed not to form alternative conformers with the rest of the structure. This linker was found to stabilize the structure less than non-nucleotide linkers (28).

The click reaction is characterized by a high rate and yield. It is catalyzed by the copper ion which has to be

reduced to the first oxidation state. Typically, freshly prepared ascorbic acid is used but in some protocols TCEP is also employed. Unfortunately, despite the presence of reducing agents in the reaction the copper cations change the oxidation state back to (II). This has a negative effect on the reaction and mostly on the purification of the product. Copper cations interact strongly with nucleic acids which makes them hard to remove. This can make gel electrophoresis difficult. The migration of nucleic acids is distorted and the band of the product is smeared. In our experience it was necessary to employ the TBTA reagent which forms a complex with copper and stabilizes it in the (I) oxidation state. The use of acetone or N,N,N',N'',N''-pentamethyldiethylenetriamine (PMDETA) or degassing the solution did not improve the reaction nor removed the copper cations from the solvent.

By bringing together different methods we hope to make RNA hairpins amenable to biochemical and especially to crystallographic studies. This can also be useful in NMR spectroscopy in which the oligomers are also brought to high concentrations during the measurement and again, the coexistence of hairpins with duplexes causes difficulties in interpreting the resonance spectra in terms of a three-dimensional structure. The presented methodologies are generally applicable, easy to perform and do not require a special equipment. They should be useful in preparing stable, biologically relevant RNA constructs for structural studies. In particular they should help reduce the deficit of detailed structures of the RNA hairpin motifs.

## FUNDING

National Science Centre (Poland) [UMO-2011/01/B/NZ1/04429, UMO-2013/08/A/ST5/00295]; Ministry of Science and Higher Education (Poland) [01/KNOW2/2014]. Funding for open access charge: Ministry of Science and Higher Education (Poland) [01/KNOW2/2014].  
*Conflict of interest statement.* None declared.

## REFERENCES

1. Bevilacqua, P.C. and Blose, J.M. (2008) Structures, kinetics, thermodynamics, and biological functions of RNA hairpins. *Annu. Rev. Phys. Chem.*, **59**, 79–103.
2. Warner, K.D., Chen, M.C., Song, W.J., Strack, R.L., Thorn, A., Jaffrey, S.R. and Ferre-D'Amare, A.R. (2014) Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat. Struct. Mol. Biol.*, **21**, 658–663.
3. Svoboda, P. and Di Cara, A. (2006) Hairpin RNA: a secondary structure of primary importance. *Cell Mol. Life Sci.*, **63**, 901–908.
4. Serganov, A. and Nudler, E. (2013) A decade of riboswitches. *Cell*, **152**, 17–24.
5. Filbin, M.E. and Kieft, J.S. (2009) Toward a structural understanding of IRES RNA function. *Curr. Opin. Struct. Biol.*, **19**, 267–276.
6. Ha, M. and Kim, V.N. (2014) Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.*, **15**, 509–524.
7. Steitz, T.A. (2008) A structural understanding of the dynamic ribosome machine. *Nat. Rev. Mol. Cell Biol.*, **9**, 242–253.
8. Germer, K., Leonard, M. and Zhang, X. (2013) RNA aptamers and their therapeutic and diagnostic applications. *Int. J. Biochem. Mol. Biol.*, **4**, 27–40.
9. Ouellet, J. (2016) RNA fluorescence with light-up aptamers. *Front. Chem.*, **4**, 29.
10. Pfeiffer, F. and Mayer, G. (2016) Selection and biosensor application of aptamers for small molecules. *Front. Chem.*, **4**, 25.
11. Zhou, J., Bobbin, M.L., Burnett, J.C. and Rossi, J.J. (2012) Current progress of RNA aptamer-based therapeutics. *Front. Genet.*, **3**, 234.
12. Guo, P. (2010) The emerging field of RNA nanotechnology. *Nat. Nanotechnol.*, **5**, 833–842.
13. Mirkin, S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932–940.
14. Galka-Marciniak, P., Urbanek, M.O. and Krzyzosiak, W.J. (2012) Triplet repeats in transcripts: structural insights into RNA toxicity. *Biol. Chem.*, **393**, 1299–1315.
15. Todd, P.K. and Paulson, H.L. (2010) RNA-mediated neurodegeneration in repeat expansion disorders. *Ann. Neurol.*, **67**, 291–300.
16. Wojciechowska, M. and Krzyzosiak, W.J. (2011) Cellular toxicity of expanded RNA repeats: focus on RNA foci. *Hum. Mol. Genet.*, **20**, 3811–3821.
17. Holbrook, S.R., Cheong, C.J., Tinoco, I. and Kim, S.H. (1991) Crystal-structure of an RNA double helix incorporating a track of non-Watson-Crick base-pairs. *Nature*, **353**, 579–581.
18. Rypniewski, W., Adamiak, D.A., Milecki, J. and Adamiak, R.W. (2008) Noncanonical G(syn)-G(anti) base pairs stabilized by sulphate anions in two X-ray structures of the (GUGGUCUGAUGAGGCC) RNA duplex. *RNA*, **14**, 1845–1851.
19. Frilander, M.J. and Turunen, J.J. (2008) RNA ligation using T4 DNA ligase. In: Hartmann, R.K., Bindereif, A., Schön, A. and Westhof, E. (eds). *Handbook of RNA Biochemistry*. Wiley-VCH Verlag GmbH, Weinheim, pp. 36–52.
20. Persson, T., Willkomm, D.K. and Hartmann, R.K. (2008) T4 RNA Ligase. In: Hartmann, R.K., Bindereif, A., Schön, A. and Westhof, E. (eds). *Handbook of RNA Biochemistry*. Wiley-VCH Verlag GmbH, Weinheim, pp. 53–74.
21. Petkovic, S. and Muller, S. (2015) RNA circularization strategies *in vivo* and *in vitro*. *Nucleic Acids Res.*, **43**, 2454–2465.
22. Kolb, H.C., Finn, M.G. and Sharpless, K.B. (2001) Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. Engl.*, **40**, 2004–2021.
23. Doktycz, M.J., Paner, T.M. and Benight, A.S. (1993) Studies of DNA Dumbbells. IV. Preparation and melting of a DNA dumbbell with the 16-base-pair sequence 5'-G-T-A-T-C-C-C-T-C-T-G-G-A-T-A-C-3' linked on the ends by dodecyl chains. *Biopolymers*, **33**, 1765–1777.
24. Lewis, F.D., Wu, Y.S. and Liu, X.Y. (2002) Synthesis, structure, and photochemistry of exceptionally stable synthetic DNA hairpins with stilbene diether linkers. *J. Am. Chem. Soc.*, **124**, 12165–12173.
25. Ma, M.Y.X., Mccallum, K., Climie, S.C., Kuperman, R., Lin, W.C., Summersmith, M. and Barnett, R.W. (1993) Design and synthesis of RNA miniduplexes via a synthetic linker approach. 2. generation of covalently closed, double-stranded cyclic HIV-1 TAR RNA analogs with high TAT-binding affinity. *Nucleic Acids Res.*, **21**, 2585–2589.
26. Ng, P.S., Laing, B.M., Balasundaram, G., Pingle, M., Friedman, A. and Bergstrom, D.E. (2010) Synthesis and evaluation of new spacers for use as dsDNA end-caps. *Bioconjug. Chem.*, **21**, 1545–1553.
27. Pils, W. and Micura, R. (2000) Flexible non-nucleotide linkers as loop replacements in short double helical RNAs. *Nucleic Acids Res.*, **28**, 1859–1863.
28. Rumney, S. and Kool, E.T. (1995) Structural optimization of non-nucleotide loop replacements for duplex and triplex DNAs. *J. Am. Chem. Soc.*, **117**, 5635–5646.
29. Thevarpadam, J., Bessi, I., Binas, O., Goncalves, D.P.N., Slavov, C., Jonker, H.R.A., Richter, C., Wachtveitl, J., Schwalbe, H. and Heckel, A. (2016) Photoresponsive formation of an intermolecular minimal g-quadruplex motif. *Angew. Chem. Int. Ed. Engl.*, **55**, 2738–2742.
30. Altmann, S., Labhardt, A.M., Bur, D., Lehmann, C., Bannwarth, W., Billeter, M., Wuthrich, K. and Leupin, W. (1995) NMR studies of DNA duplexes singly cross-linked by different synthetic linkers. *Nucleic Acids Res.*, **23**, 4827–4835.
31. Asensio, J.L., Brown, T. and Lane, A.N. (1998) Comparison of the solution structures of intramolecular DNA triple helices containing adjacent and non-adjacent CG center dot C+ triplets. *Nucleic Acids Res.*, **26**, 3677–3686.
32. Kozerski, L., Mazurek, A.P., Kawecki, R., Bocian, W., Krajewski, P., Bednarek, E., Sitkowski, J., Williamson, M.P., Moir, A.J.G. and Hansen, P.E. (2001) A nicked duplex decamer DNA with a PEG(6) tether. *Nucleic Acids Res.*, **29**, 1132–1143.



33. Siegmund, K., Hariharan, M. and Lewis, F.D. (2011) Conformation of a dodecane DNA hairpin linker. multiple gauche bonds cover the bases. *J. Phys. Chem. B.*, **115**, 3740–3746.
34. Tarkoy, M., Phipps, A.K., Schultze, P. and Feigon, J. (1998) Solution structure of an intramolecular DNA triplex linked by hexakis(ethylene glycol) units: d(AGAGAGAA-(EG)(6)-TTCTCTCT-(EG)(6)-TCTCTCTT). *Biochemistry*, **37**, 5810–5819.
35. Egli, M., Tereshko, V., Mushudov, G.N., Sanishvili, R., Liu, X. and Lewis, F.D. (2003) Face-to-face and edge-to-face pi-pi interactions in a synthetic DNA hairpin with a stilbenediether linker. *J. Am. Chem. Soc.*, **125**, 10842–10849.
36. MacElrevey, C., Spitale, R.C., Krucinska, J. and Wedekind, J.E. (2007) A posteriori design of crystal contacts to improve the X-ray diffraction properties of a small RNA enzyme. *Acta. Crystallogr. D. Biol. Crystallogr.*, **63**, 812–825.
37. Lietard, J., Meyer, A., Vasseur, J.J. and Morvan, F. (2008) New strategies for cyclization and bicyclization of oligonucleotides by click chemistry assisted by microwaves. *J. Org. Chem.*, **73**, 191–200.
38. Wang, Q.S. and Unrau, P.J. (2002) Purification of histidine-tagged T4 RNA ligase from E-coli. *Biotechniques*, **33**, 1256–1260.
39. Kim, Y., Babnigg, G., Jedrzejczak, R., Eschenfeldt, W.H., Li, H., Maltseva, N., Hatzos-Skintges, C., Gu, M.Y., Makowska-Grzyska, M., Wu, R.Y. *et al.* (2011) High-throughput protein purification and quality assessment for crystallization. *Methods*, **55**, 12–28.
40. Pasternak, A. and Wengel, J. (2011) Unlocked nucleic acid - an RNA modification with broad potential. *Org. Biomol. Chem.*, **9**, 3591–3597.
41. Kiliszek, A., Kierzek, R., Krzyzosiak, W.J. and Rypniewski, W. (2009) Structural insights into CUG repeats containing the 'stretched U-U wobble': implications for myotonic dystrophy. *Nucleic Acids Res.*, **37**, 4149–4156.
42. Sobczak, K., Michlewski, G., de Mezer, M., Kierzek, E., Krol, J., Olejniczak, M., Kierzek, R. and Krzyzosiak, W.J. (2010) Structural diversity of triplet repeat RNAs. *J. Biol. Chem.*, **285**, 12755–12764.
43. Broda, M., Kierzek, E., Gdaniec, Z., Kulinski, T. and Kierzek, R. (2005) Thermodynamic stability of RNA structures formed by CNG trinucleotide repeats. Implication for prediction of RNA structure. *Biochemistry*, **44**, 10873–10882.
44. Moazed, D., Robertson, J.M. and Noller, H.F. (1988) Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. *Nature*, **334**, 362–364.
45. Paredes, E. and Das, S.R. (2011) Click chemistry for rapid labeling and ligation of RNA. *Chembiochem.*, **12**, 125–131.