

# Towards mRNA with superior translational activity: synthesis and properties of ARCA tetraphosphates with single phosphorothioate modifications†‡

Malwina Strenkowska,<sup>§a</sup> Joanna Kowalska,<sup>§a</sup> Maciej Lukaszewicz,<sup>a</sup> Joanna Zuberek,<sup>a</sup> Wei Su,<sup>b</sup> Robert E. Rhoads,<sup>b</sup> Edward Darzynkiewicz<sup>a</sup> and Jacek Jemielity<sup>\*a</sup>

Received (in Montpellier, France) 6th November 2009, Accepted 15th January 2010

First published as an Advance Article on the web 10th March 2010

DOI: 10.1039/b9nj00644c

We describe the chemical synthesis and preliminary biophysical and biochemical characterization of a series of mRNA 5'-end (cap) analogs designed as reagents for obtaining mRNA molecules with augmented translation efficiency and stability *in vivo* and as useful tools to study mRNA metabolism. The analogs share three structural features: (i) the 5',5'-bridge is elongated to a tetraphosphate to increase their affinity to translation initiation factor eIF4E; (ii) a single phosphorothioate modification at either the  $\alpha$ ,  $\beta$ ,  $\gamma$  or  $\delta$ -position of the tetraphosphate has been made to decrease their susceptibility to enzymatic degradation and/or to modulate their interaction with specific proteins; and (iii) a 2'-O-methyl group has been added to the ribose of 7-methylguanosine (a characteristic of Anti-Reverse Cap Analogs (ARCAs), which are incorporated into mRNA during *in vitro* transcription exclusively in the correct orientation). The dinucleotides bearing the modified tetraphosphate bridge were synthesized by ZnCl<sub>2</sub>-mediated coupling between two mononucleotide subunits with isolated yields of 30–65%. The preliminary biochemical results show that mRNAs capped with new analogs are 2.5–4.5 times more efficiently translated in a cell-free system than m<sup>7</sup>GpppG-capped mRNAs, which makes them promising candidates for RNA-based therapeutic applications such as gene therapy and anticancer vaccines.

## Introduction

Dinucleoside 5',5'-polyphosphates fulfill numerous regulatory, signaling, or energetic functions in biological systems. A particular representative of this group of biologically important molecules is the m<sup>7</sup>G-containing cap present on the 5'-end of all eukaryotic messenger RNAs. Different types of cap structures exist in nature,<sup>1</sup> but they all contain guanosine methylated at the N7-position connected by a 5',5'-triphosphate bridge to the first nucleotide of the RNA transcript (Fig. 1A).

The cap plays a key role in a variety of cellular processes related to mRNA metabolism, including splicing, nucleo-cytoplasmic transport, intracellular targeting, translation, translational repression, and turnover.<sup>1,2</sup> Therefore, synthetic analogs of the cap are widely employed as tools for studying these physiological

processes as well as being useful in biotechnology. There are also promising potential clinical applications.

One application that has been intensively explored by our groups is the use of cap analogs as reagents for *in vitro* synthesis of capped mRNAs. Capped mRNAs may be employed for expression of proteins in either *in vitro* translation systems or in cultured cells. Recently, mRNAs have emerged as an attractive alternative to classical DNA-vector based gene delivery due to several advantages; the use of mRNA avoids the risk of insertional mutagenesis, is effective in non-dividing cells, and does not require introduction into the nucleus because mRNA is translated in the cytoplasm.<sup>3</sup>

Synthesis of capped mRNA can be achieved by transcribing a DNA template *in vitro* with either a bacterial<sup>4</sup> or bacteriophage<sup>5,6</sup> RNA polymerase in the presence of all four NTPs and a dinucleotide cap analog such as m<sup>7</sup>GpppG. These polymerases normally initiate transcription with nucleophilic attack by the 3'-OH of GTP on the  $\alpha$ -phosphate of the next nucleoside triphosphate specified by the DNA template. If a cap analog such as m<sup>7</sup>GpppG is present in the reaction mixture at a ~5:1 ratio to GTP, the transcription is initiated mainly by attack of 3'-OH of the cap dinucleotide rather than that of GTP, leading to formation of capped transcripts of the form m<sup>7</sup>GpppGpNp(Np)<sub>n</sub>. Unfortunately, attack can occur by the 3'-OH of either the Guo or m<sup>7</sup>Guo moieties of the cap dinucleotide, resulting in one-third to one-half of the transcripts being capped in a reversed orientation, *i.e.*, Gpppm<sup>7</sup>GpNp(Np)<sub>n</sub>.<sup>7</sup> Such reverse-capped transcripts decreased the overall translational

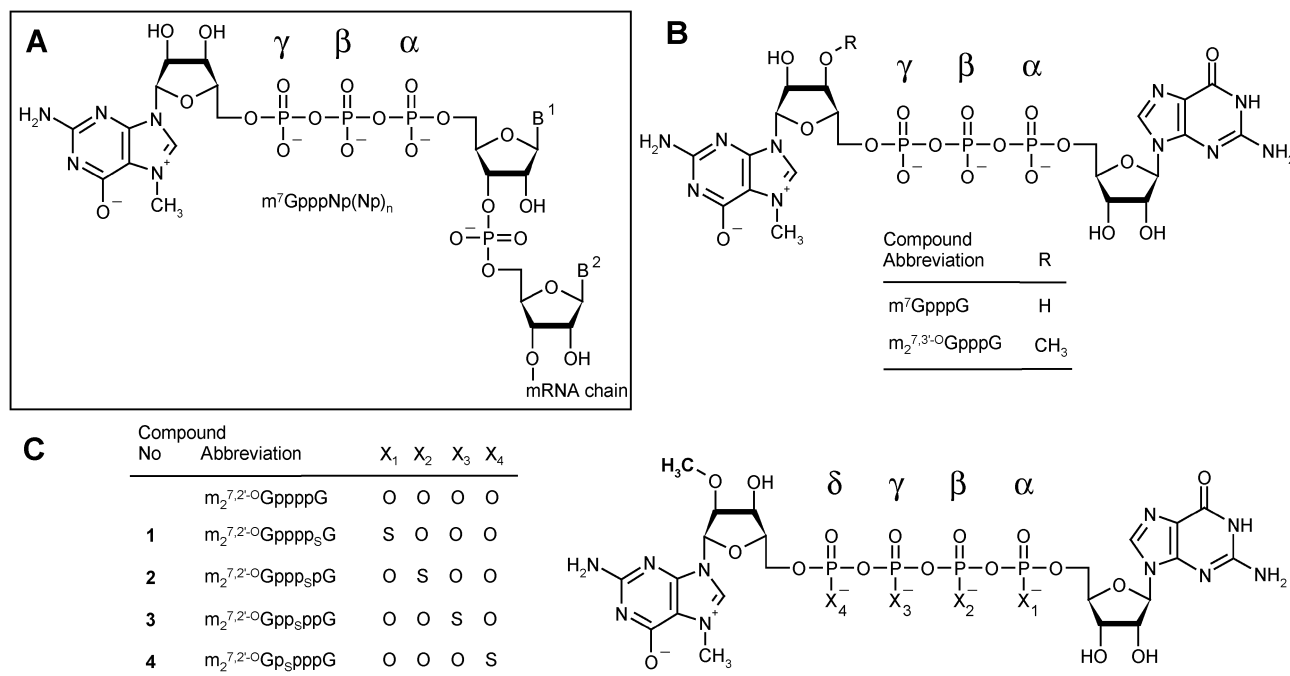
<sup>a</sup> Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland. E-mail: jacek@biogeo.uw.edu.pl; Fax: +48 22 5540771; Tel: +48 22 5540774

<sup>b</sup> Department of Biochemistry and Molecular Biology, Louisiana State University Health Sciences Center, Shreveport, Louisiana, USA 71130-3932, USA

† Electronic supplementary information (ESI) available: Table S1, HPLC profiles, <sup>1</sup>H and/or <sup>31</sup>P NMR spectra of compounds 1–6. See DOI: 10.1039/b9nj00644c

‡ This article is part of a themed issue on Biophosphates, and is dedicated to Professor Wojciech J. Stec on the occasion of his 70th birthday.

§ These authors contributed equally to this work.



**Fig. 1** Cap structures. (A) A general depiction of the eukaryotic mRNA 5'-end. (B) Structures of standard triphosphate analogs: m<sup>7</sup>GpppG and m<sub>2</sub><sup>7,3'-O</sup>GpppG (ARCA). (C) Structures of m<sub>2</sub><sup>7,2'-O</sup>GppppG and analogs synthesized in this study.

activity of mRNA. This problem was overcome several years ago by the introduction of anti-reverse cap analogs (ARCAs) bearing either 3'-O-methyl, 3'-H or 2'-O-methyl modifications in the m<sup>7</sup>Guo moiety, ensuring 100% correct orientation.<sup>8–10</sup> ARCA-capped mRNAs were shown to have higher translational efficiency than m<sup>7</sup>GpppG capped mRNAs both *in vitro*<sup>8,10</sup> and in cultured cells.<sup>11–13</sup>

One important advantage of the ARCA method over the post-transcriptional enzymatic capping is the option to introduce additional chemical modifications into the mRNA 5'-end that can modify the behavior and properties of the mRNAs. Among dozens of modified synthetic cap analogs examined, several have been identified that augment mRNA translational efficiency, both *in vitro* and in cultured cells.<sup>10,13</sup> Two features of the modified cap structure have emerged that influence the overall yield of protein produced from the mRNA, particularly in cultured cells.<sup>11,13</sup> The first one is the cap's intrinsic ability to compete with natural mRNAs for recruitment to the translational machinery, and the second is its susceptibility to enzymatic decapping.

During recruitment of mRNA to the translational machinery, the cap is specifically recognized by eukaryotic initiation factor 4E (eIF4E). mRNA recruitment to form the 48S pre-initiation complex is the rate-limiting step for initiation of translation under normal circumstances (absence of virus infection, cellular stress, *etc.*) and plays a central role in regulation of translation.<sup>14</sup> The presence of the cap greatly increases the rate of translation. Affinity of the cap for eIF4E, and hence for the translational machinery, can be increased by modifying the 5',5'-triphosphate bridge. One class of compounds with modifications of this type are cap analogs in which the 5',5'-bridge is extended from triphosphate to tetraphosphate, *e.g.* m<sup>7</sup>GppppG, m<sub>2</sub><sup>7,2'-O</sup>GppppG or m<sub>2</sub><sup>7,3'-O</sup>GppppG.<sup>10</sup>

Generally, these analogs have binding affinities to eIF4E that are ~9–10-fold higher than for the corresponding triphosphates and promote translation 20–30% more efficiently when introduced into mRNA. However, it should be mentioned that further increasing the affinity to eIF4E may exert opposite effect on mRNA translational efficiency. Surprisingly, it was found that extending the 5',5'-bridge to pentaphosphate increases the binding affinity of cap analogs to eIF4E by a factor of ~3.5–5 compared to tetraphosphates; nonetheless, mRNAs capped with these analogs are translated with efficiency comparable or even lower than for corresponding triphosphates.<sup>10</sup>

A second important role of the mRNA cap is protecting mRNA against degradation.<sup>15</sup> Exonucleolytic mRNA cleavage in the 5'→3' direction cannot take place until the cap is removed from mRNA 5'-end by a decapping pyrophosphatase termed Dcp1/Dcp2 (a heterodimer consisting of a regulatory and catalytic subunit). The enzyme cleaves the cap, provided it is attached to an RNA of at least 20 nt, between the α and β phosphate moieties to release m<sup>7</sup>GDP and 5-phosphorylated mRNA chain. This is then degraded by the exonuclease Xrn1.<sup>16,17</sup> Both the 5'→3' and 3'→5' pathways play major roles in mRNA degradation.<sup>15</sup> We have recently demonstrated that mRNAs capped with an ARCA analog containing a phosphorothioate moiety at the β-position of the triphosphate bridge, m<sub>2</sub><sup>7,2'-O</sup>GppspG (D2),<sup>18</sup> were resistant to decapping by Dcp2 *in vitro*, resulting longer half-lives in cultured mouse mammary cells (257 min compared to 155 min for m<sub>2</sub><sup>7,2'-O</sup>GpppG-capped mRNA).<sup>13</sup> An unexpected further benefit was that m<sub>2</sub><sup>7,2'-O</sup>GppspG (D2)-capped mRNA was 2.4 times more efficiently translated in these cells than m<sub>2</sub><sup>7,2'-O</sup>GpppG-capped mRNA.

In the present work, we aimed at obtaining analogs that would combine both features that can enhance the translational yield of mRNA: high affinity for eIF4E and resistance to cleavage by Dcp2, in the expectation that these two features would independently improve the translational properties of mRNA in living cells. We therefore synthesized a series of tetraphosphate ARCA analogs modified with the phosphorothioate group at various positions of the 5',5'-tetraphosphate bridge (Fig. 1C).

The new tetraphosphate analogs could potentially be useful in other spheres of mRNA metabolism, based on our recent findings with triphosphate phosphorothioate ARCA analogs modified at different positions. Those modified at the  $\gamma$ -position were resistant to another decapping enzyme, DcpS, which acts on the products of 3'  $\rightarrow$  5' mRNA decay. DcpS cleaves only short capped 5'-oligonucleotides or m<sup>7</sup>GpppN dinucleotides released after processive degradation of mRNA by the exosome initiated from the 3'-end. Cleavage is between the  $\beta$ - and  $\gamma$ -phosphates of the 5,5-triphosphate bridge, releasing m<sup>7</sup>GMP and a downstream (oligo)nucleotide.<sup>19</sup> Resistance to DcpS is an important factor in design of long-lasting cap analogs that can inhibit translation.

Recently, a cap analog modified at the  $\alpha$ -position, m<sub>2</sub><sup>7,2'-O</sup>GpppsG (D2), was shown to enhance microRNA-mediated translational repression by a factor of 2 with no change in translation efficiency, providing new insights into the mechanism of this recently discovered but poorly understood process.<sup>20</sup>

## Results

### Novel teraphosphate S-ARCA

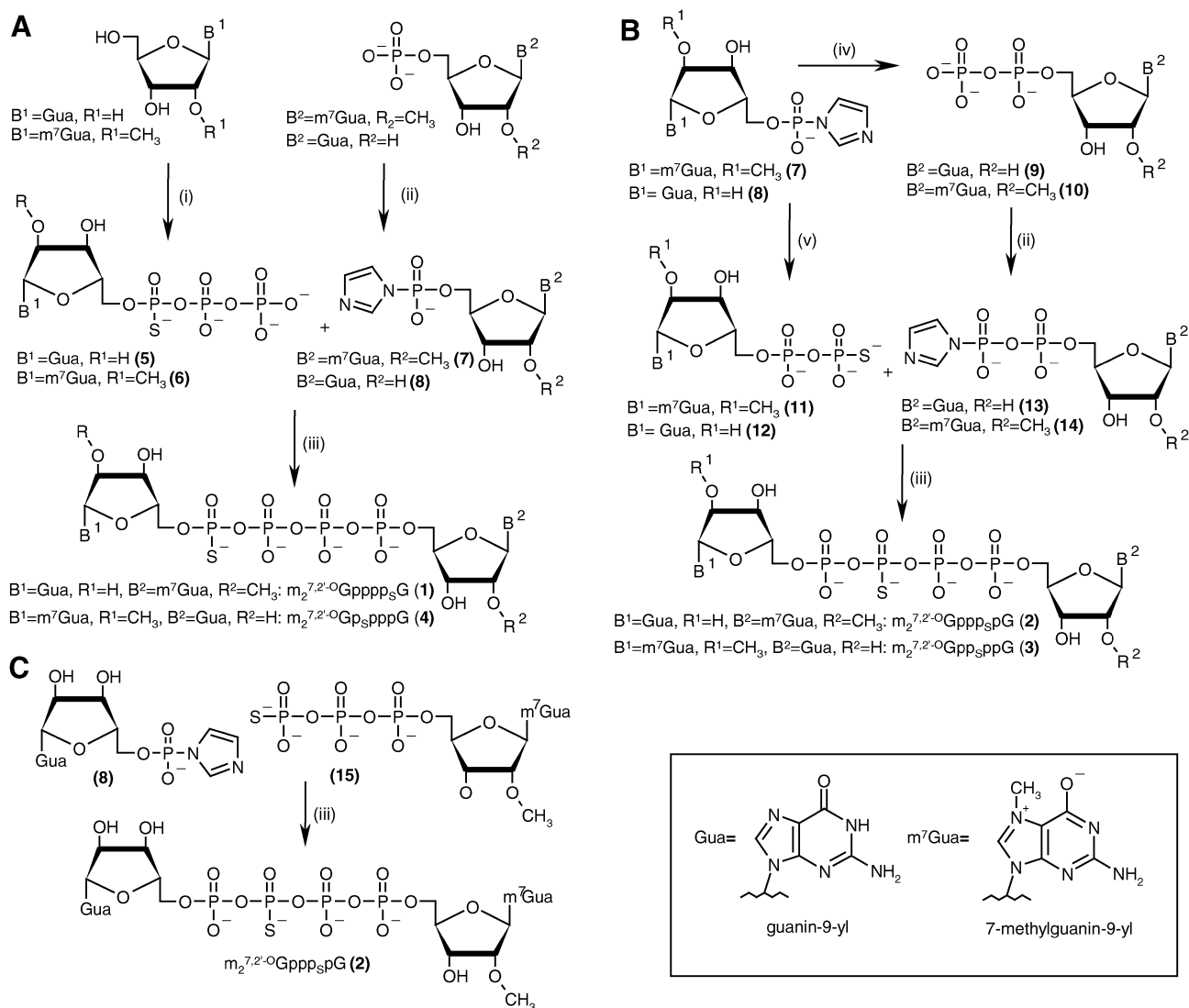
We designed and prepared a series of four cap analogs **1–4** bearing a single phosphorothioate moiety at either the  $\alpha$ -,  $\beta$ -,  $\gamma$ -, or  $\delta$ -positions of the tetraphosphate chain and 2-*O*-methyl group (ARCA modification) in the m<sup>7</sup>Gua moiety. We refer to these as 4P-S-ARCAs (Fig. 1C). The synthetic routes provide these analogs with good yields and can also be applied to the synthesis of a variety of other biologically relevant dinucleoside polyphosphates. Due to the presence of a P-stereogenic center, each 4P-S-ARCA exists as two P-diastereomers, designated D1 and D2 based on their elution order during reverse-phase (RP) HPLC. Two pairs of diastereomers were successfully resolved by RP HPLC, but we could not achieve this for the other two. Subsequently, the new analogs were subjected to biophysical and biochemical characterization. We determined their binding affinities for eIF4E, susceptibility to decapping by DcpS, and translational efficiencies in a rabbit reticulocyte lysate (RRL) system. Other biochemical properties of mRNAs capped with these analogs, including susceptibility to the human decapping enzyme Dcp2, translational efficiency in cultured mammalian cells, and half-life in mammalian cells, are currently under investigation. Preliminary results suggest that mRNA capped with a mixture of diastereomers (D1 plus D2) of m<sub>2</sub><sup>7,2'-O</sup>GpppsG (compound **2**) is less susceptible to *in vitro* decapping by hDcp2 than mRNA capped with m<sub>2</sub><sup>7,2'-O</sup>GppspG (compound **3**).<sup>21</sup>

### Chemical synthesis of new analogs

The key step in the preparation of each new analog (**1–4**) was formation of a thio-modified tetraphosphate bridge by ZnCl<sub>2</sub>-mediated coupling between a nucleoside 5'-*O*-(thiopolyphosphate) and a second nucleotide activated by conversion into the P-imidazolidine. Generally, this approach offers the possibility of synthesizing each compound three ways<sup>22</sup> (see Discussion). For each analog we chose the most straightforward synthetic pathway (Scheme 1). Analogs **1** and **4**, which are modified at the  $\alpha$ - and  $\delta$ -positions, were obtained by the “3 + 1\*” strategy, *i.e.*, by coupling a nucleoside 5'-*O*-(1-thiotriphosphate) with a nucleoside 5'-monophosphate P-imidazolidine. Analogs **2** and **3**, which are modified at the  $\beta$ - and  $\gamma$ -positions, were obtained by the “2 + 2\*” strategy, *i.e.*, by coupling a nucleoside 5'-(2-thiodiphosphate) with a nucleoside 5'-diphosphate imidazolidine. An alternative “3 + 1\*” route was also tested for analog **2**.

The guanosine 5'-*O*-(1-thiotriphosphate) **5** was obtained by 5'-thiophosphorylation of guanosine in trimethylphosphate in the presence of a non-nucleophilic base, 2,6-lutidine,<sup>23</sup> followed by addition of tributylammonium pyrophosphate. As monitored by RP HPLC (see Experimental section), the thiophosphorylation proceeded smoothly, and more than 90% conversion of nucleoside was observed within 5 h. After addition of pyrophosphate solution in DMF and weakly-alkaline hydrolysis (pH 7.5–8.0) a mixture of products was obtained consisting mainly of the desired guanosine 5'-*O*-(1-thiotriphosphate) (~60%), guanosine 5'-*O*-thiophosphate (~10%) and guanosine 5'-*O*-(1-thiopentaphosphate) (~15%). The products were separated by DEAE Sephadex A-25 ion-exchange chromatography, after which **5** was isolated as 7 : 10 (D1 : D2) diastereomeric mixture with 46% yield. The synthesis of **6** was similar to that of **5**, but the reaction was less efficient for two reasons: (i) thiophosphorylation proceeded significantly slower, the final conversion (after 24 h) not exceeding 80%, and (ii) the byproduct, 7,2'-*O*-dimethylguanosine 5'-*O*-(1-thiopentaphosphate), was formed more efficiently (~25%). Moreover, <sup>31</sup>P NMR analysis unfortunately revealed that the desired product co-eluted with unreacted pyrophosphate, probably due to the partial neutralization of the negative charge on the  $\alpha$ -phosphate by the positive charge in the m<sup>7</sup>Gua moiety. Consequently, additional HPLC purification was necessary before the synthesis scheme could be continued. Compound **6** was finally isolated as a 7 : 10 (D1 : D2) diastereomeric mixture with 30% yield. Both **5** and **6** were used in the subsequent coupling reactions as diastereomeric mixtures.

The final couplings between **5** and **7** or **6** and **8** leading to **1** or **4**, respectively, were mediated by ~16-fold ZnCl<sub>2</sub> excess. The P-imidazolidine derivatives, **7** and **8**, synthesized using a dithiodipyridine/triphenylphosphine activating system,<sup>24</sup> were used in 1.5–2-fold excess to maximize consumption of the second reactant. Both coupling reactions proceeded smoothly, and essentially complete conversion of **5** or **6** into tetraphosphate-bridged dinucleotides was observed by RP HPLC after several hours (Fig. 2A). The respective HPLC conversions and isolated yields achieved after ion-exchange purification are shown in Table 1.



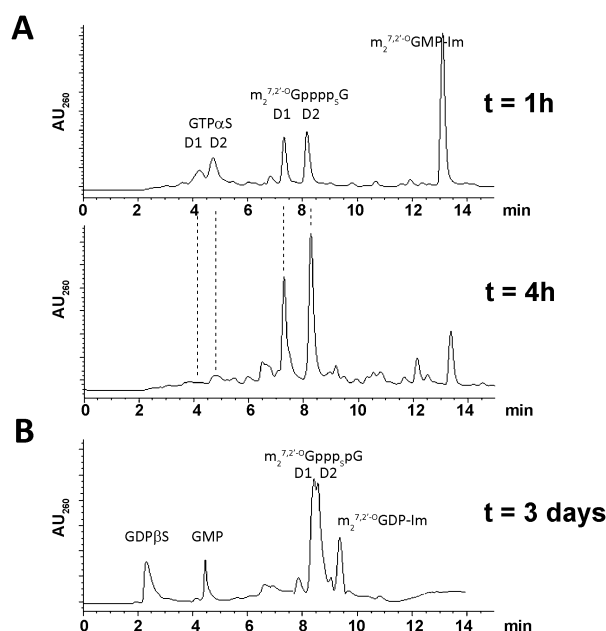
**Scheme 1** The synthesis of tetraphosphate S-ARCA 1–4. (A) Synthesis of **1** and **4** by the “3 + 1\*” strategy. (B) Synthesis of **2** and **3** by the “2 + 2\*” strategy. (C) Synthesis of **2** by the “3 + 1\*” strategy (the asterisk denotes the activated nucleotide subunit). Reagents: i. (1)  $\text{PSCl}_3$ , trimethylphosphate, 2,6-lutidine, 0 °C; (2) 0.5 M tributylammonium pyrophosphate in DMF; ii. imidazole, 2,2′-dithiodipyridine,  $\text{PPh}_3$ , DMF; iii.  $\text{ZnCl}_2$ , DMF; iv. triethylammonium phosphate,  $\text{ZnCl}_2$ , DMF; v. triethylammonium thiophosphate,  $\text{ZnCl}_2$ , DMF.

The synthesis of analogs modified at internal positions of the tetraphosphate bridges, *i.e.*, **2** and **3**, involved formation of mixed phosphate–thiophosphate anhydride bonds in the final step (Scheme 1B and C). This required the nucleoside 5′-(2-thiodiphosphates) **11** and **12** or 5′-(3-thiotriphosphate) **15**, which were synthesized with 80–85% yields by a procedure reported previously by our group.<sup>25</sup> In the case of mixed anhydrides, the formation of dinucleotides proceeded slower than for unmodified ones. The reaction progress usually stopped after 3 d despite some reactants remaining. Partial hydrolysis of **11**, **12** or **15** also occurred due to extended reaction times; consequently, HPLC conversions and isolated yields were lower than in the case of **1** and **4** but were still acceptable (Table 1, Fig. 2B). In the case of analog **2**, both synthetic pathways proved to be equally efficient.

Analog **1**–**4** were isolated after ion-exchange chromatography as diastereomeric mixtures. The diastereomers of **1** and **4**, in

which the stereogenic P-center is next to the nucleoside moiety, were well separated by either analytical or preparative RP HPLC (see Experimental section). The differences between diastereomers of **2** and **3** were less distinct, perhaps because the P-stereogenic center is located in the more flexible part of tetraphosphate bridge and is more distant from the nucleoside stereogenic centers. Hence, we were unable to find conditions in which these diastereomers could be separated from each other on a preparative basis, although different mobile phases and two types of columns were tested. However, in case of **2** we managed to separate small amounts of compound sufficient to determine some biochemical properties (~0.5 mg). Thus analogs **1** and **4** were subjected to biological studies as diastereomerically pure samples (**1a** and **4a** = D1, **1b** and **4b** = D2), but **3** was studied as a diastereomeric mixture. Compound **2** was studied both as a diastereomeric mixture (**2**) and as pure diastereomers (**2a** = D1, **2b** = D2).<sup>26</sup>





**Fig. 2** Typical HPLC profiles of reaction mixtures. (A) Progress in synthesis of **1** after 1 h and 4 h. (B) Synthesis of **2** after 3 days.

All newly synthesized compounds were characterized by MS ESI (–),  $^1\text{H}$  and  $^{31}\text{P}$  NMR. Each pair of diastereomers showed some small but significant differences in their  $^1\text{H}$  chemical shifts, particularly for H8 of the nucleobase and H1' of the ribose moiety of either guanosine or 7,2'-*O*-dimethylguanosine. For **3** distinctive differences were also observed for the  $m_2^{7,2'-O}\text{Guo}$  H3' and both methyl groups. In case of all compounds modified with sulfur at a phosphate neighboring a nucleoside (both nucleoside triphosphates **5** and **6** and dinucleoside tetraphosphates **1** and **4**) the H8 proton of the diastereomer with shorter RP HPLC retention time (D1) was more up-field shifted than that of D2, with  $\Delta\delta_{\text{H}}$  chemical shift differences being in the range of 0.2–0.7 ppm (see Experimental section and Table S1). For ATP $\alpha$ S and its analogs, Major *et al.*<sup>27</sup> showed that the faster diastereomer in RP HPLC, with a less shielded H8 proton, has the  $S_{\text{P}}$  configuration; their model is that the H8 shift depends on its distance from the S atom, which is different for the two diastereomers. The same correlation between the RP HPLC retention time, H8 proton shift and the absolute configuration was subsequently reported for other structurally related compounds, including guanosine nucleotides.<sup>28</sup> Some minor differences between diastereomers were also noted in the  $^{31}\text{P}$  NMR spectra of **1**–**6**. One

observation was that the signal of the phosphorothioate moiety in compounds **1**, **4**, **5** (GTP $\alpha$ S) and **6** was 0.2–0.3 ppm up-field shifted for D1 compared to D2 (Table S1). The absolute configuration of GTP $\alpha$ S was first reported by Conolly *et al.*;<sup>29</sup> the RP HPLC-faster diastereomer was assigned as  $S_{\text{P}}$ . Later Ludwig and Eckstein reported that the  $\alpha$ -phosphorus signals in GTP $\alpha$ S (and other NTP $\alpha$ Ss) were  $\sim 0.2$  ppm more up-field for D1 ( $S_{\text{P}}$ ) than for D2 ( $R_{\text{P}}$ ).<sup>30</sup> The above-mentioned similarities in the  $^1\text{H}$  and  $^{31}\text{P}$  spectra of compounds **1** and GTP $\alpha$ S (**5**) suggest that their D1 diastereomers have the same, *i.e.*  $S_{\text{P}}$ , absolute configuration. This is further supported by the observation that **1a** and **1b** are formed in the same diastereomeric ratio as the **5a** and **5b** ratio before reaction. On the basis of analogous observations made for compounds **4** and **6**, one can presume that in their cases the isomer D1 is also of  $S_{\text{P}}$  configuration. However, this assumption needs to be further confirmed, *e.g.* by X-ray crystallography.

### Binding affinities to eIF4E

The binding affinities of new analogs for murine eIF4E were determined by fluorescence quenching (Fig. 3).<sup>31</sup> Experimental curves obtained upon titration of eIF4E with increasing concentrations of cap analogs are depicted in Fig. 3A. The descending fragment of each curve is due to the quenching of intrinsic Trp fluorescence in eIF4E upon cap binding, whereas the increasing fluorescence signal after reaching the plateau originates from fluorescence emission of the unbound cap analog.<sup>31</sup> The  $K_{\text{AS}}$  values and free energies of binding ( $\Delta G^\circ$ ) for the new analogs and reference compounds are presented in Table 2 and Fig. 3B. As expected, all new analogs have binding affinities at least 10-fold higher than the unmodified cap structure,  $m^7\text{GpppG}$ . The comparison to  $m_2^{7,2'-O}\text{GppppG}$  reveals that the phosphorothioate moiety generally stabilizes complexes with eIF4E. Modifications in the  $\delta$ -,  $\gamma$ -, and  $\beta$ -positions (analogs **4a**, **4b**, **3**, **2a**, **2b**) give additional energetic gains which are similar to those reported for the corresponding modifications in the triphosphate series.<sup>32</sup> By contrast, the phosphorothioate substitution in the  $\alpha$ -position, which represent the 'extra' phosphate moiety (analogs **1a** and **1b**), has virtually no influence on the binding affinity.

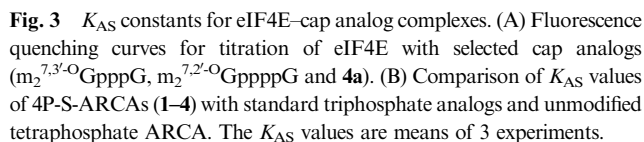
### Susceptibility to enzymatic degradation by specific and nonspecific pyrophosphatases

The susceptibility of new analogs to enzymatic hydrolysis by the cap-specific human DcpS pyrophosphatase was tested by

**Table 1** Characterization of coupling reactions of two nucleotide subunits leading to diastereomeric 4P-S-ARCA (1–4)

Compound				Yield (%)		
No.	Abbreviation	Imidazolidine	Nucleophile	HPLC	Isolated <sup>a</sup>	D1:D2 ratio <sup>b</sup>
1	m <sub>2</sub> <sup>7,2'-O</sup> GppppsG	m <sub>2</sub> <sup>7,2'-O</sup> GMP-Im (7)	GTPαS (5)	> 90	60	7 : 10
2	m <sub>2</sub> <sup>7,2'-O</sup> GppppsP	GMP-Im (8)	m <sub>2</sub> <sup>7,2'-O</sup> GTPγS (15)	56	39	9 : 10
3	m <sub>2</sub> <sup>7,2'-O</sup> GpppspG	m <sub>2</sub> <sup>7,2'-O</sup> GDP-Im (14)	GDPβS (12)	62	35	8 : 10
3	m <sub>2</sub> <sup>7,2'-O</sup> GpppsppG	GDP-Im (13)	m <sub>2</sub> <sup>7,2'-O</sup> GDPβS (11)	68	32	7 : 10
4	m <sub>2</sub> <sup>7,2'-O</sup> GpspppG	GMP-Im (8)	m <sub>2</sub> <sup>7,2'-O</sup> GTPαS (6)	> 90	65	7 : 10

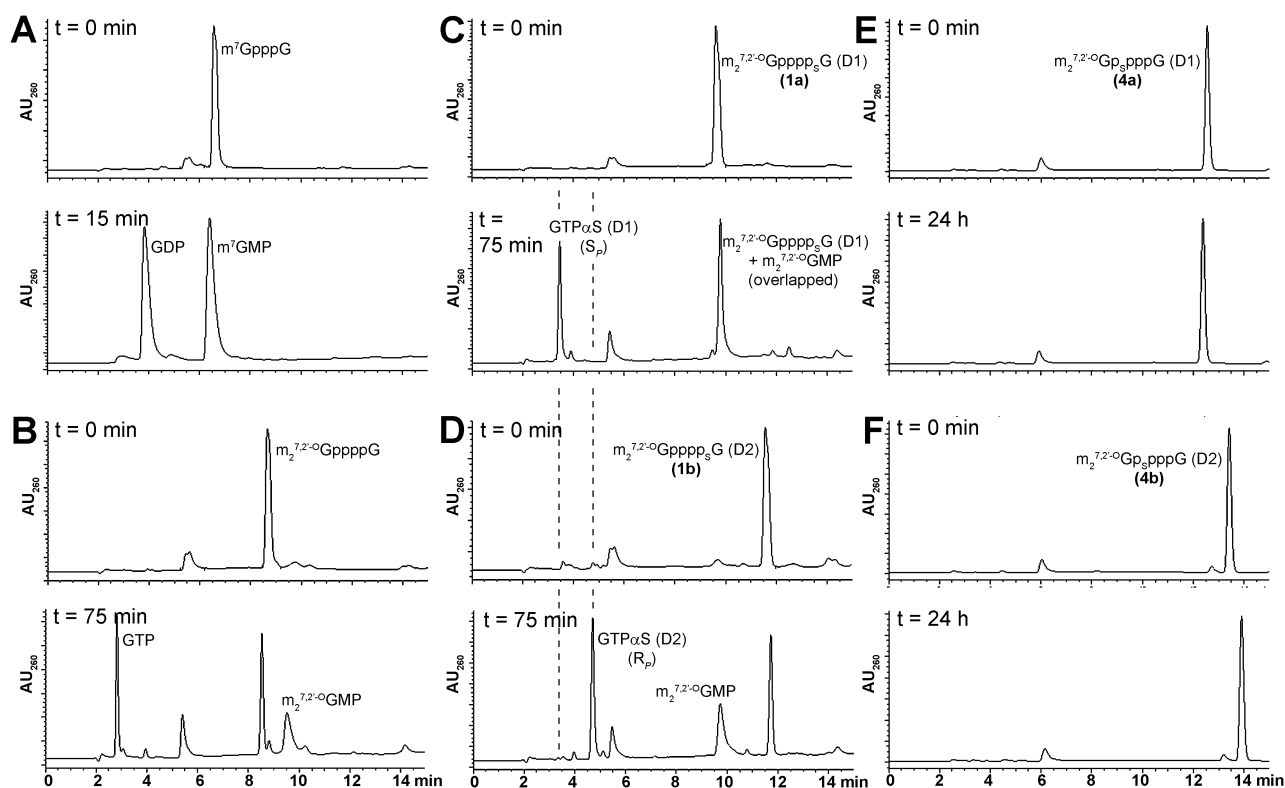
<sup>a</sup> Based on the optical density units at 260 nm of diastereomeric mixtures of products before their HPLC separation. <sup>b</sup> Based on HPLC and/or  $^1\text{H}$  NMR.



DcpS was adjusted to assure complete cleavage of m<sup>7</sup>GpppG to m<sup>7</sup>GMP and GDP over 15 min. Under these conditions, analogs bearing modification at the δ-position remained undigested after 24 h of incubation with DcpS, regardless of the P-center absolute configuration, and hence, were considered to be resistant to hydrolysis. Analogs **1–3** were hydrolyzed more than 80% by hDcpS to m<sub>2</sub><sup>7,2′-O</sup>GMP and the corresponding nucleoside triphosphate in 2 h, which is comparable to hydrolysis of m<sub>2</sub><sup>7,2′-O</sup>GppppG. The identification of hydrolysis product was confirmed by co-injections with authentic samples of m<sub>2</sub><sup>7,2′-O</sup>GMP, GTP, and GTPαS S<sub>P</sub> or R<sub>P</sub>. The analysis of DcpS degradation products of analogs **1a** and **1b** adds support to our contention that their absolute configurations around P-stereogenic centers are the same as those of GTPαS (D1) and (D2), respectively, *i.e.*, S<sub>P</sub> and R<sub>P</sub> (Fig. 4).

**Table 2** Biochemical properties of tetraphosphate S-ARCs

<sup>a</sup> D1 or D2 designate one of two P-diastereomers; for each pair D1 refers to the one eluted faster from RP HPLC column. <sup>b</sup> Values are means of 3 experiments. <sup>c</sup> The mean values were calculated from 2–3 assay repetitions for each of two independent mRNA syntheses. <sup>d</sup> Calculated by correcting the translation of each mRNA in a particular experiment by subtracting the translation of mRNA capped with ApppG (a nonfunctional cap structure). <sup>e</sup> Data from ref. 10. <sup>f</sup> The composition of D1/D2 mixtures was as shown in Table 1, except m<sub>2</sub><sup>7,2'-O</sup>Gp<sub>8</sub>pppG (mix), which was in a 1:1 ratio.



**Fig. 4** RP HPLC profiles of mRNA cap dinucleotides digestion by hDcpS. A)  $m^7\text{GpppG}$  B)  $m_{2,2'}\text{-O-GppppG}$  C) **1a** D) **1b** E) **4a** F) **4b**. Assay conditions are given in the Experimental section.

dephosphorylation of GMP and cleavage of  $m_{2,2'}\text{-O-GTP}$  to  $m_{2,2'}\text{-O-GMP}$ . Surprisingly, when analogs modified at the  $\alpha$ -position (**1a** and **1b**) were subjected to SVPDE digestion, the regioselectivity (regardless of the configuration of the P-stereogenic center) was opposite to that observed for  $m^7\text{GpppG}$  and  $m_{2,2'}\text{-O-GppppG}$ . The complete cleavage of both **1a** and **1b** occurred within 1.5 h and the only products observed were  $m_{2,2'}\text{-O-GMP}$  and  $\text{GTP}\alpha\text{S}$  (either D1 or D2). The reaction mixture remained practically unchanged after further incubation with enzyme, so unfortunately, under these conditions, we were not able to observe stereoselectivity of the cleavage of the phosphorothioate moieties. The analogs modified at the  $\delta$ -position, **4a** and **4b**, were cleaved to release GMP and  $m_{2,2'}\text{-O-GTP}\alpha\text{S}$  (D1) or (D2), respectively. As for **1a** and **1b**, no further cleavage of  $\text{NTP}\alpha\text{S}$  occurred. Nonetheless, this experiment adds support also to the assumption of the absolute configurations of **4a** and **4b**, and identifies SVPDE as a new tool for studying modified dinucleotide cap analogs, since its specificity is complementary to that of DcpS.

#### Translational efficiencies of 4P-S-ARCs

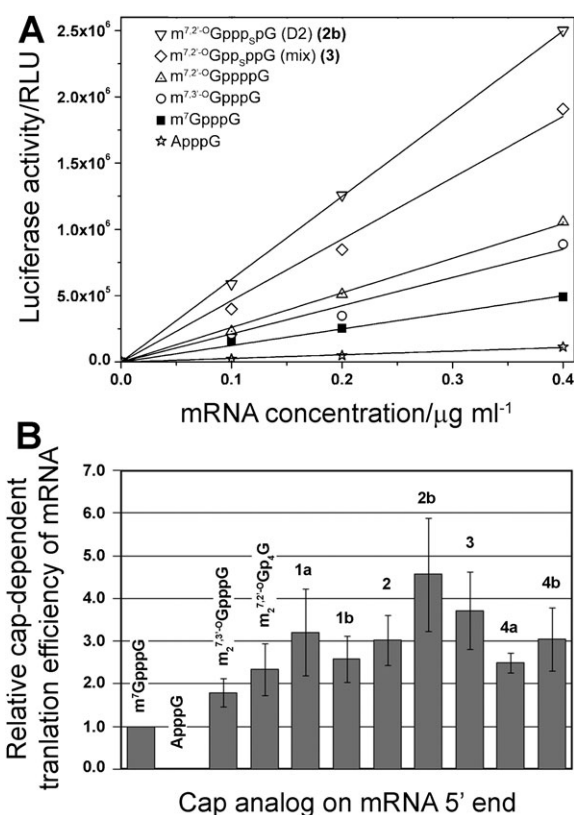
Finally, the new analogs were investigated for their utility as reagents for increasing the translational activity of mRNA. Transcripts encoding firefly luciferase were synthesized by SP6 RNA polymerase in the presence of a DNA template, four NTPs, and a cap dinucleotide (**1a**, **1b**, **2**, **2b**, **3**, **4a**, **4b**). The assays included three reference compounds:  $m^7\text{GpppG}$ ,  $m_{2,3'}\text{-O-GpppG}$ , and  $m_{2,2'}\text{-O-GppppG}$ . *In vitro*-synthesized mRNAs were translated in a micrococcal nuclease-treated RRL system and the activity of luciferase per  $\mu\text{g}$  of RNA

was determined by luminometry after 60 min (Fig. 5A). Cap-dependent translational efficiencies were calculated by correcting the translation of each mRNA in particular experiment by subtracting the translation of mRNA capped with ApppG (a nonfunctional cap structure), which is considered to be cap-independent. The overall translational efficiencies and cap-dependent translational efficiencies (both normalized to the translation of  $m^7\text{GpppG}$ -capped mRNA) are shown in Table 2.

In our experiment we could reproduce results for cap analogs that were previously tested.<sup>10</sup> For instance, mRNA capped with  $m_{2,3'}\text{-O-GpppG}$  and  $m_{2,2'}\text{-O-GppppG}$  were translated  $1.79 \pm 0.33$  and  $2.33 \pm 0.61$  more efficiently than  $m^7\text{GpppG}$ -capped mRNA, respectively, which is similar to the previously reported values of  $1.88 \pm 0.40$  and  $2.56 \pm 0.18$ . Importantly, mRNAs capped with all of the new 4P-S-ARCs were not translated less efficiently than mRNA capped with the parent compound  $m_{2,2'}\text{-O-GppppG}$ . Their relative translation efficiencies ranged from 2.5 to 4.5 compared to  $m^7\text{GpppG}$  (Table 2, Fig. 5B), which makes them attractive candidates for further studies in living cells. mRNAs capped with two analogs, **3** and **2b**, were translated more than 3-fold more efficiently than  $m^7\text{GpppG}$ -capped mRNA.

#### Discussion

Nucleoside and dinucleoside polyphosphates are biologically important molecules. Numerous methods for their synthesis have been developed, several of them providing synthetic routes to dinucleoside tri- or tetraphosphates modified at



**Fig. 5** Translational efficiency of mRNA transcripts *in vitro*. (A) Example experiment of translational efficiency of mRNA transcript encoding firefly luciferase capped with various dinucleotides. (B) Comparison of relative cap-dependent translational efficiency of mRNA transcripts. The mean values were calculated from 2–3 assay repetitions for each of two independent mRNA syntheses.

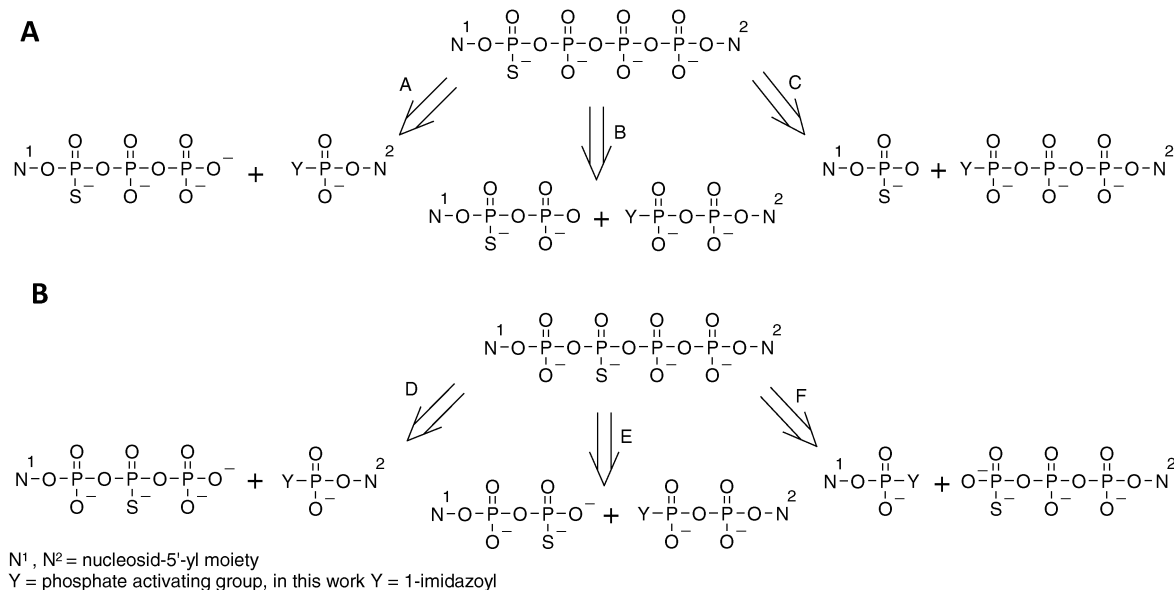
non-bridging positions in the phosphate group(s).<sup>29,35–40</sup> However, most of these methods have drawbacks with regard

to synthesis of mRNA cap analogs. These include limited structural variety of possible products available by a particular method (*e.g.*, limitation to symmetrical compounds or those modified at the  $\alpha$ -position with respect to nucleoside) or requirement for protected nucleosides as starting materials. Due to the presence of positive charge in 7-methylguanine,  $m^7\text{Guo}$  is unusually polar and thus poorly soluble in organic solvents. It is also susceptible to hydrolysis both in acidic and basic aqueous solutions,<sup>41</sup> which are required for removal of the majority of protecting groups. These facts make synthesis of  $m^7\text{Guo}$ -containing nucleotides even more challenging.

As first noted by Kadokura *et al.*<sup>42</sup> and then explored by our laboratory,<sup>8,10,32,43</sup> and others,<sup>44</sup> the formation of pyrophosphate bonds *via* coupling of nucleotides with nucleotide P-imidazolid derivatives in DMF is greatly accelerated by the presence of excess of anhydrous zinc chloride. The rationale for  $\text{ZnCl}_2$  catalysis is at least two-fold. First, it increases solubility of nucleotides in DMF, probably through complexing negatively charged phosphates. Due to that, it is possible to obtain completely homogenous DMF solutions of nucleotide triethylammonium salts or even sodium salts in the case of P-imidazolid derivatives. Second,  $\text{ZnCl}_2$  is thought to act as Lewis acid to activate imidazole as a leaving group.

The synthesis of 4P-S-ARCA has also been carried out taking advantage of this phenomenon. The formation of dinucleoside tetraphosphates modified at either the  $\alpha$ - or  $\beta$ -position with respect to the closest nucleoside can be generally achieved by any of the pathways depicted in Scheme 2. The choice of particular pathway should depend on the availability of the appropriate substrates as well as on the purification method to be applied.

For the analogs **1** and **4** pathways A, B and C were taken into consideration. Pathway B was excluded, as nucleoside 5'-*O*-(1-thiodiphosphates) can be obtained by procedures similar to those for 5'-*O*-(1-thiotriphosphates), albeit with



**Scheme 2** Possible disconnections of tetraphosphate-bridged thio-modified dinucleotides. (A) Tetraphosphates with phosphorothioate modification in the external position. (B) Tetraphosphates with phosphorothioate modification in the internal position.



lower yields.<sup>28,45</sup> Moreover, nucleoside diphosphates are available from the corresponding activated monophosphates used in pathway A. This would ultimately lengthen the synthesis, giving no obvious advantage over pathway A. Pathway C would involve coupling of nucleoside 5'-monothiophosphates (which are easily available) and imidazolidine derivatives of nucleoside triphosphates (which are possible to obtain,<sup>10</sup> although – due to decreased solubility in DMF – with lower yields than those for mono- and diphosphates). Moreover, our experience gained during the synthesis of thio-modified dinucleoside triphosphates shows that formation of a mixed phosphate-thiophosphate anhydride bond is much slower than for an unmodified pyrophosphate bond, which would presumably enhance by-product formation, hampering product purification. Hence, the pathway A appeared to us as favorable, since it involves the most available substrates and, additionally, the formation of the tetraphosphate bridge involves coupling between two unmodified phosphate moieties. This route proved to be efficient, as the coupling reactions leading to analogs **1** and **4** were relatively rapid (several hours), with HPLC conversions exceeding 90%.

In regard to the analogs **2** and **3**, pathways D, E and F should be considered. Pathway D, although it shares some of the advantages of pathway A, is less favorable since nucleoside 5'-(2-thiotriphosphates) can be obtained only by complicated procedures. Pathways E and F seemed to be both acceptable and both proved to be similarly efficient, with HPLC conversions from 50 to 70%.

It is worth emphasizing that the described synthetic method can be also applied to other biologically important dinucleoside polyphosphates.<sup>46</sup> It would be particularly advantageous for the synthesis of unsymmetrical compounds, for instance analogs of uridine adenosine tetraphosphate,<sup>47</sup> or for compounds bearing highly polar nucleosides.

The determination of binding affinity to eIF4E by means of fluorescence quenching titration is a relatively simple yet very useful biophysical test that enables one to predict the functionality of any newly synthesized cap analog. A sufficiently high binding affinity for eIF4E is essential for the recognition of mRNA by translational machinery and efficient competing of mRNAs possessing modified caps with natural mRNAs.

In one of our previous studies,<sup>11</sup> we identified a triphosphate ARCA analog modified with a methylene group between the  $\alpha$  and  $\beta$  phosphates,  $m_2^{7,3'-O}GppCH_2pG$ , which protected mRNA against hydrolysis by Dcp2-decapping pyrophosphatase. Due to this, the mRNA had a prolonged half-life when transfected into cultured cells. However, the methylene modification coincidentally decreased the binding affinity to eIF4E from  $10.2 \pm 0.3$  (for  $m_2^{7,3'-O}GpppG$ ) to  $4.41 \pm 0.2 \mu M^{-1}$ . This resulted in the mRNA capped with  $m_2^{7,3'-O}GppCH_2pG$  (despite being more stable *in vivo*) being translated even less efficiently than that capped with regular ARCA.<sup>11</sup> Another example of the influence of the binding affinity to eIF4E on mRNA translation is the tetraphosphate ARCA analog,  $m_2^{7,2'-O}GppppG$ , which due to the additional phosphate moiety has  $K_{AS}$  much higher than the corresponding triphosphate, because of which it was shown to enhance mRNA translation efficiency *in vitro*.<sup>10</sup> It is worth mentioning

that the combination of methylene modification with the presence of the additional, fourth, phosphate moiety has been shown recently to compensate for the loss of binding affinity.<sup>43</sup>

The high binding affinity for eIF4E of new tetraphosphate S-ARCA constituted thus an important premise that all those analogs should be at least as efficient in promoting translation as  $m_2^{7,2'-O}GppppG$ . The particular  $K_{AS}$  values were dependent both on the position of the phosphorothioate modification and on the absolute configuration of the P-stereogenic center. Their comparison to  $m_2^{7,2'-O}GppppG$  and to the S-ARCA triphosphate series reveals that the phosphorothioate moiety in the  $\delta$ -,  $\gamma$ - and  $\beta$ -positions of tetraphosphate correspond to  $\gamma$ -,  $\beta$ - and  $\alpha$ -modifications in the triphosphate bridge.<sup>32,48</sup> This implies that eIF4E 'counts' the phosphates starting from  $m^7Guo$ , which is in agreement with the well-established fact that the  $m^7Guo$  moiety is crucial for mRNA 5'-end recognition by eIF4E. For instance, the determined  $\Delta G^\circ$  binding energy for **4a** was 0.40 and 0.54 kcal mol<sup>-1</sup> lower than that of **4b** and  $m_2^{7,2'-O}GppppG$ , respectively. In the triphosphate series the  $\Delta \Delta G^\circ$  differences between  $m_2^{7,2'-O}GpsppG$  (D1) and either its D2 counterpart or  $m_2^{7,2'-O}GpppG$  were 0.59 and 0.69 kcal mol<sup>-1</sup>, respectively. This observation indicates that the energetic gains arising from the presence of an additional phosphate and an oxygen-to-sulfur substitution are additive. For analog **3** we were only able to determine the binding affinity for the D1/D2 mixture. In the triphosphate S-ARCA series, modification at the corresponding position ( $\beta$ ) resulted in notably different  $K_{AS}$  values for D1 and D2 ( $43.1 \pm 1.4$  versus  $19.3 \pm 2.2 \mu M^{-1}$ , respectively). From the titration performed on the 1 : 1 mixture of **4a** and **4b**, which also differ in their  $K_{AS}$ , we found that the  $K_{AS}$  value determined for a diastereomeric mixture is roughly a mean of the  $K_{AS}$  values for pure diastereomers. Interestingly, the  $K_{AS}$  value for analog **3** ( $282 \pm 12 \mu M^{-1}$ ) is similar to that of pentaphosphate ARCA analog,  $m_2^{7,3'-O}Gp_5G$ . As mentioned in the Introduction, this analog, despite its high  $K_{AS}$  value ( $299 \pm 20 \mu M^{-1}$ ), was not effective in promoting translation when incorporated into the mRNA 5'-end.<sup>10</sup> The proposed explanation was that the extraordinarily high binding affinity interferes with release of the cap at the end of translation initiation, and thereby inhibits further translation events. However, the high translation efficiency of analog **3** suggests that some other factors may contribute to this decreased efficiency of mRNA capped with  $m_2^{7,3'-O}Gp_5G$ , e.g. low capping percentage during *in vitro* transcription, or non-specific interactions of highly charged pentaphosphate chain with other proteins.

The results of *in vitro* translation of mRNAs capped with new 4P-S-ARCA further support the initial assumption that combining the phosphorothioate modification with elongation of the 5',5'-bridge to tetraphosphate may produce analogs with beneficial properties in terms of preparation of translationally efficient mRNA. The results suggest that the largest increase in mRNA translation efficiency may be caused by modifications in the internal positions of the tetraphosphate chain (*i.e.*  $\beta$  and  $\gamma$ ). The somehow unexpected difference between translation efficiency of **2** (diastereomeric mixture) and **2b** (diastereomer D2) underlines the need for finding a method of separating it into pure diastereomers in order to explore this phenomenon.

This more efficient mRNA translation *in vitro* may benefit biotechnological *in vitro* protein production. The determination of translation efficiency in RRL is also a useful preliminary test, enabling selection of most attractive analogs for expensive studies in cultured cells. However, conventionally used rabbit reticulocyte lysates (RRLs), are micrococcal nuclease pre-treated to deplete any endogenous mRNAs present in the lysate. Therefore, this assay does not fully reflect the cellular conditions since: (i) there is no competition of investigated mRNA with natural mRNAs, and (ii) the influence of its stability (half-life) on overall translation yield is rather marginal. Both of these factors have a large impact on the overall protein synthesis *in vivo*.<sup>13</sup> The results from our previous studies suggest that in cellular conditions the differences in translation yield between mRNAs capped with modified and unmodified analogs should be more pronounced.

The ability to produce more stable and more efficiently translated *in vivo* mRNAs can be advantageous for applications in mRNA-based gene therapy. Several approaches have been exploited, to improve the stability and/or translation efficiency of mRNA vectors *in vivo*. This includes extending the polyA length, alterations of the mRNA 3'-end, and introducing into mRNA modified nucleosides, which retain coding properties of natural ones.<sup>3,12,49,50</sup> Modification of mRNA cap structure is another possibility to facilitate mRNA-based gene delivery. A particularly attractive application of mRNA in therapy is transfection of the patient's own dendritic cells with mRNAs encoding tumor-associated antigens to evoke auto-immunization against cancer cells.<sup>51</sup>

We believe that the 4P-S-ARCA's will also enable some new insights into mRNA-related physiological processes. The synthetic cap analogs have been commonly used to investigate mechanism of initiation of protein biosynthesis, and the discovery of the involvement of decapping enzymes in the mRNA degradation pathways expanded further the scope of processes that have been studied by means of cap analogs. One of the still unsolved problems is the detailed mechanism of the Dcp2 decapping enzyme, which belongs to the NuDiX family of pyrophosphatases.<sup>16</sup> It is known that the enzyme utilizes cap structures on intact mRNAs and cleaves between the  $\alpha$ - and  $\beta$ -phosphates of the triphosphate bridge. The finding that the phosphorothioate group in the  $\beta$ -position protected mRNA from decapping suggests that the  $\beta$ -phosphate is the one directly involved in the catalytic mechanism.<sup>13</sup> However, since Dcp2 requires both cap structure and mRNA chain for its activity, it is unclear whether the enzyme recognizes the phosphate groups 'counting' them from the m<sup>7</sup>Gua site or from the mRNA body. The susceptibility to Dcp2 of tetraphosphate S-ARCA modified in the  $\beta$  and  $\gamma$ -positions, which is currently under investigation, may address this issue.

On the other hand, the analogs modified in the  $\delta$ -position proved to be resistant to the DcpS decapping enzyme, which cleaves cap structures lacking an mRNA chain. DcpS-resistant cap analogs with high affinity for eIF4E are important in the context of developing stable inhibitors of cap-dependent translation which could counteract elevated levels of eIF4E in cancer cells.<sup>52</sup> It is assumed that DcpS functions in the cytoplasm to prevent the accumulation of complexes between eIF4E and cap structures that would otherwise accumulate

following 3'→5' mRNA decay,<sup>53</sup> which implies that only DcpS-resistant analogs could act as potent eIF4E inhibitors in cellular conditions. Moreover, it was also found that DcpS fulfils also some other biological functions, and constitutes an interesting object for investigation by itself. Its presence has also been detected in the nucleus, where it is implicated in mRNA splicing, and recently, it has been identified as therapeutic target for spinal muscular atrophy.<sup>54,55</sup> Non-hydrolysable cap analogs such as m<sub>2</sub><sup>7,2'-O</sup>GppppG may serve as invaluable tools to investigate these processes, for instance by serving as inhibitors of DcpS activity.

Finally, it has been recently discovered that by modifying the 5',5'-phosphate bridge in the mRNA cap structure one can observe enhancement in translation repression mediated by microRNAs.<sup>20</sup> In the quoted work, two cap analogs have been identified, which in *Drosophila melanogaster* embryos extract evoke ~2-fold stronger miRNA-mediated inhibition of translation initiation but are "neutral" towards general cap-dependent translation: the triphosphate ARCA modified at the  $\alpha$ -position, m<sub>2</sub><sup>7,2'-O</sup>GpppsG (D2), and the analog with the 5',5'-bridge extended to hexaphosphate, m<sup>7</sup>Gp<sub>6</sub>G. Thus a 4P-S-ARCA modified at the  $\alpha$ -position, which somehow combines the structural features of these two analogs, may be an interesting tool for studying mechanism of microRNA action.

## Conclusions

A general methodology for the synthesis of unsymmetrical dinucleotides bearing a single phosphorothioate modification in either position of the tetraphosphate chain has been developed. Four pairs of tetraphosphate S-ARCA's have been synthesized with good yields. Biophysical and biochemical characterization of new analogs indicates that 4P-S-ARCA's are good candidates for reagents to produce mRNA transcripts with high translational activity. Stabilizing effects of the additional phosphate as well as the phosphorothioate substitutions on the cap-eIF4E complex are additive, which is reflected in  $K_{AS}$  values much higher than for triphosphate cap analogs and notably higher than for unmodified tetraphosphate ARCA. This property of 4P-S-ARCA's should be reflected in the kinetics of translation initiation complex formation and make synthetic transcripts more competitive than endogenous mRNA *in vivo*.

It is also shown that phosphorothioate modification in the  $\delta$ -position of the tetraphosphate bridge protects the dinucleotide cap analog against hydrolysis by DcpS. Similarly, the modification in the appropriate position of the tetraphosphate bridge should decrease the susceptibility of 4P-S-ARCA-containing mRNA transcripts to hydrolysis by Dcp2/Dcp1, which would affect their half-life and give an additional contribution to translational activity *in vivo*.

The results of studies on translation *in vitro* show that mRNAs capped with new analogs are 2.5–4.5 times more efficiently translated in a cell-free system than m<sup>7</sup>GpppG-capped mRNAs. On the basis of our previous studies, one can predict that the effect of these novel caps on translational efficiency should be more pronounced when measured in cultured cells. These observations indicate that 4P-S-ARCA's are promising

candidates for studies on therapeutic applications of mRNA, which include mRNA-based gene delivery and antitumor vaccination. In addition to the biotechnological and potential therapeutic applications, 4P-S-ARCA appear to be useful tools for elucidating mechanisms of various mRNA-related physiological processes.

## Experimental section

### Chemical syntheses

#### General procedures

Solvents and other reagents were purchased from Sigma-Aldrich and used without further treatment, unless otherwise stated. Reaction mixtures were analyzed by analytical reverse-phase (RP) HPLC, which was performed on a Agilent Technologies Series 1200 apparatus using a Supelcosil LC-18-T RP column (4.6 × 250 mm, 5 μm, flow rate 1.3 mL/min) developed with a 0–50% linear gradient of methanol in 0.05 M ammonium acetate buffer (pH 5.9) over 30 min, UV-detection at 260 nm and fluorescence detection (excitation at 280 nm and detection at 337 nm). Samples for HPLC analysis were prepared by dissolving 5 μL of a reaction mixture in 100 μL of either water or, in the case of ZnCl<sub>2</sub>-containing reactions, an aqueous solution of EDTA (10 mg/mL) and NaHCO<sub>3</sub> (5 mg/mL).

Nucleotides were purified and isolated from reaction mixtures by ion-exchange chromatography on DEAE-Sephadex A-25 (HCO<sub>3</sub><sup>−</sup> form) column using a linear gradient of triethylammonium bicarbonate (TEAB) in deionized water (1.6 L of each) and, after evaporation under reduced pressure with repeated additions of ethanol and drying in a vacuum desiccator over P<sub>2</sub>O<sub>5</sub>, isolated as triethylammonium (TEA) salts.

The final products were additionally purified (and separated) by semi-preparative RP HPLC, which was performed on either a Waters 600E Multisolvant Delivery System apparatus using Discovery RP Amide C-16 HPLC column (25 cm × 21.2 mm, 5 μm, flow rate 5.0 mL/min) (or Ascentis C18 HPLC column (25 cm × 21.2 mm, 5 μm, flow rate 5.0 mL/min) developed with linear gradients of methanol in 0.05 M ammonium acetate (AAC) buffer (pH 5.9). UV-detection at 260 nm. After repeated freeze-drying, the products were isolated as ammonium salts. For compound **6**, triethylammonium acetate (TEAAC) buffer (pH 7.0) was used in order to isolate compounds as TEA salts.

Yields were preferably calculated based on optical density milliunits at 260 nm (mODU<sub>260</sub>) of substrates and isolated products, measured in 0.1 M phosphate buffer of pH either 6 (for m<sup>7</sup>Guo mononucleotides) or pH = 7 (for cap dinucleotides and Guo nucleotides). Optical density milliunits mODU<sub>260</sub> are defined as absorption of compound solution in 0.1 M phosphate buffer of appropriate pH at 260 nm multiplied by the volume of the solution in mL. Extinction coefficients taken for calculations were ε<sub>pH6</sub> = 11 400 M<sup>−1</sup> cm<sup>−1</sup> for m<sup>7</sup>Guo mononucleotides, ε<sub>pH7</sub> = 22 600 M<sup>−1</sup> cm<sup>−1</sup> for cap dinucleotides and ε<sub>pH7</sub> = 12 000 M<sup>−1</sup> cm<sup>−1</sup> for Guo mononucleotides.

The structures and homogeneities of synthesized nucleotides were confirmed by chromatography on analytical RP HPLC, mass spectrometry using negative electrospray ionization

(MS ESI<sup>−</sup>) and both <sup>1</sup>H and <sup>31</sup>P NMR spectroscopy. <sup>1</sup>H NMR and <sup>31</sup>P NMR spectra were recorded in D<sub>2</sub>O at 25 °C for samples at 1.5–3.0 mg mL<sup>−1</sup> concentrations on a Varian UNITY-plus spectrometer at 399.94 MHz and 161.90 MHz, respectively. <sup>1</sup>H NMR chemical shifts in ppm were reported to sodium 3-trimethylsilyl[2,2,3,3-D<sub>4</sub>]propionate (TSP) in D<sub>2</sub>O as an internal standard. <sup>31</sup>P NMR chemical shifts in ppm were reported to 80% phosphorus acid in D<sub>2</sub>O as an external standard. *J* values are given in Hz. Mass spectra were recorded on a Micromass QToF 1 MS spectrometer.

2′-*O*-Methylguanosine was synthesized as described by Kusmirek and Shugar.<sup>56</sup> 7,2′-*O*-Dimethylguanosine was obtained from 2′-*O*-methylguanosine by methylation with CH<sub>3</sub>I analogously to that described earlier for 7-methylguanosine.<sup>57</sup> Tributylammonium pyrophosphate was prepared according to Ludwig and Eckstein.<sup>30</sup> Thiophosphate TEA salt was obtained from its trisodium salt as described earlier.<sup>25</sup> Guanosine 5′-monophosphate disodium salt was purchased from Fluka and before use converted into its triethylammonium (TEA) salt by passing through Dowex 50W X8 (100–200 mesh), evaporating collected fractions to dryness and drying in vacuum over P<sub>2</sub>O<sub>5</sub>. Nucleotides: guanosine 5′-diphosphate (**9**), 7,2′-*O*-dimethylguanosine 5′-monophosphate and 7,2′-*O*-dimethylguanosine 5′-diphosphate (**10**) (all as TEA salts) as well as imidazolidine derivatives **7**, **8**, **13** and **14** (all as sodium salts) were synthesized as described previously.<sup>10,43</sup>

**Guanosine 5′-(1-thiotriphosphate), GTPαS (**5**).** To a suspension of guanosine (7710 mODU<sub>260</sub>, 250 mg, 0.64 mmol; dried overnight *in vacuo* over P<sub>4</sub>O<sub>10</sub>) in trimethyl phosphate (3.75 mL) cooled to 0 °C on ice/water bath were added 2,6-dimethylpyridine (460 μL, 4 mmol) and PSCl<sub>3</sub> (270 μL, 2.65 mmol). The reaction progress was under HPLC control. After 6 h tributylammonium pyrophosphate solution (15.9 mL of 0.5 M solution) in DMF and tributylamine (0.61 mL) were added. The reaction was maintained at RT for 15 min and then quenched by addition of 145 mL of water and a portion of 0.7 M TEAB to pH 8, and extracted twice with ether. The product was separated by DEAE Sephadex chromatography using a linear gradient of 0–1.2 M TEAB to yield **5** (3550 mODU<sub>260</sub>, 310 mg, 0.30 mmol, 46%) as a 72:100 S<sub>P</sub>:R<sub>P</sub> (D1:D2, **5a:5b**) diastereomeric mixture. ESI MS (−) *m/z* 537.98 (Calc. for C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>S: 537.96). **5a**: Analytical RP HPLC *R*<sub>t</sub> = 3.5 min δ<sub>H</sub> 8.29 (1H, s, H(8) G); 5.95 (1H, d, *J*<sub>1′,2′</sub> 6.3, H(1′) G); 4.85 (1H, ~t, *J*<sub>1′,2′</sub> 6.3 *J*<sub>2′,3′</sub> 4.8, H(2′)); 4.60 (1H, dd, *J*<sub>2′,3′</sub> 4.8, *J*<sub>3′,4′</sub> 3.2, H(3′)); 4.39 (1H, m, H(4′)); 4.30 (2H, m, H(5′) H(5′′)). δ<sub>P</sub> 43.12 (1P, d, *J*<sub>P1,P2</sub> 27.0, P1), −10.81 (1P, d, *J*<sub>P2,P3</sub> 19.6, P3), −23.98 (1P, dd, *J*<sub>P1,P2</sub> 27.0, *J*<sub>P2,P3</sub> 19.6, P2). **5b**: Analytical RP HPLC *R*<sub>t</sub> = 4.1 min δ<sub>H</sub> 8.22 (1H, s, H(8)); 5.95 (1H, d, *J*<sub>1′,2′</sub> 6.3, H(1′)); 4.85 (1H, ~t, *J*<sub>1′,2′</sub> 6.3 *J*<sub>2′,3′</sub> 4.8, H(2′)); 4.60 (1H, dd, *J*<sub>2′,3′</sub> 4.8, *J*<sub>3′,4′</sub> 3.2, H(3′)); 4.39 (1H, m, H(4′)); 4.30 (2H, m, H(5′) H(5′′)), δ<sub>P</sub> 43.64 (1P, d, *J*<sub>P1,P2</sub> 27.0, P1), −10.81 (1P, d, *J*<sub>P2,P3</sub> 19.6; P3), −24.02 (1P, dd, *J*<sub>P1,P2</sub> 27.0, *J*<sub>P2,P3</sub> 19.6; P2).

**7,2′-*O*-Dimethylguanosine 5′-(1-thiotriphosphate), m<sub>2</sub><sup>7,2′-*O*</sup>-GTPαS (**6**).** To a suspension of 7,2′-*O*-dimethylguanosine (4500 mODU<sub>260</sub>, 150 mg, 0.39 mmol; dried overnight *in vacuo*



over  $P_4O_{10}$  in trimethyl phosphate (2.25 ml) cooled to 0 °C in an ice/water bath were added 2,6-dimethylpyridine (340  $\mu$ l, 2.9 mmol) and  $PSCl_3$  (150  $\mu$ l, 1.4 mmol). The reaction progress was under HPLC control. After 24 h tributylammonium pyrophosphate solution (14 ml of 0.5 M solution in DMF) and tributylamine (0.67 ml) were added to the reaction mixture. The reaction was maintained at RT for 15 min and then quenched by addition of 120 ml of water and a portion of 0.7 M TEAB to pH 8 and extracted twice with ether. The product was separated by DEAE Sephadex chromatography using a linear gradient of 0–1.1 M TEAB to yield **6** (1580 mODU<sub>260</sub>, 270 mg, 0.14 mmol, 25%) as a 7:10  $S_P:R_P$  (D1:D2, **6a:6b**) diastereomeric mixture. The  $^{31}P$  NMR analysis revealed that the isolated product was contaminated with  $\sim 1.5$  eq. of pyrophosphate. Consequently, the product was further purified by semi-preparative RP HPLC using a linear gradient of 0–50% MeOH in TEAAC (pH 7.0) over 45 min and the diastereomers were collected together yielding, after repeated freeze-drying, **6** (1350 mODU<sub>260</sub>, 120 mg, 0.12 mmol, 30%) as a 7:10 diastereomeric mixture. Additionally, a small portion of **6** was purified using a linear gradient of 0–25% MeOH in AAC (pH 5.9) over 120 min, the diastereomers were collected separately (**6a**  $R_t$  = 41 min ( $S_P$ , D1), **6b**  $R_t$  = 45 min ( $R_P$ , D2)) and after repeated freeze-drying isolated as  $NH_4^+$  salts. ESI MS (–)  $m/z$  566.02 (calc. for  $C_{12}H_{19}N_5O_{13}P_3S$ : 565.99). **6a**: Analytical RP HPLC  $R_t$  = 8.5 min  $\delta_H$  9.33 (1H, s, H(8)); 6.15 (1H, d,  $J_{1',2'}$  2.6, H(1')); 4.70 (1H,  $\sim t$ ,  $J_{2',3'}$  =  $J_{3',4'}$   $\sim 5.3$ , H(3')); 4.42–4.33 (4H, overlapped m, H(2'), H(4'), H(5'), H(5'')); 4.17 (3H, s,  $NCH_3$ );  $\delta_P$  43.54 (1P, d,  $J_{P1,P2}$  28.0, P1), –10.32 (1P, d,  $J_{P2,P3}$  20.0, P3), –23.73 (1P, dd,  $J_{P1,P2}$  28.0,  $J_{P2,P3}$  20.0, P2). **6b**: Analytical RP HPLC  $R_t$  = 9.0 min,  $\delta_H$  9.30 (1H, s, H(8)); 6.14 (1H, d,  $J_{1',2'}$  2.1, H(1')); 4.66 (1H,  $\sim t$ ,  $J_{2',3'}$  =  $J_{3',4'}$   $\sim 5.3$ , H(3')); 4.46–4.27 (4H, overlapped m, H(2'), H(4'), H(5'), H(5'')); 4.16 (3H, s,  $NCH_3$ );  $\delta_P$  43.33 (1P, d,  $J_{P1,P2}$  27.6, P1), –10.55 (1P, d,  $J_{P2,P3}$  19.8, P3), –23.73 (1P, dd,  $J_{P1,P2}$  27.6,  $J_{P2,P3}$  19.8, P2).

#### General procedure for nucleoside 5'-(2-thiodiphosphates) and 5'-(3-thiotriphosphates) – according to ref. 25

An appropriate nucleotide imidazolide derivative (1 eq.) was mixed with thiophosphate TEA salt (3–4 eq.) in DMF followed by immediate addition anhydrous  $ZnCl_2$  (at least 8 eq.) and vigorous shaking until a homogenous solution was obtained. The reaction was quenched after 15–25 min by being diluted with  $\sim 20$  volumes of aqueous solution of disodium EDTA dihydrate (in 1:1 ratio to  $ZnCl_2$ ) and  $NaHCO_3$  (half of the amount of EDTA in mg). If necessary the pH was adjusted to 7–7.5 with an additional portion of  $NaHCO_3$ . The product was isolated by ion exchange chromatography. Product-containing fractions were poured together and evaporated to dryness with repeated additions of 96% ethanol and, finally, absolute ethanol until a white solid resulted.

**7,2'-O-Dimethylguanosine 5'-(2-thiodiphosphate),  $m_2^{7,2'-O}$ -GDP $\beta$ S (**11**).** Obtained starting from **7** (2110 mODU<sub>260</sub>, 100 mg, 0.18 mmol), thiophosphate TEA salt (obtained from 200 mg of trisodium salt hydrate,  $\sim 0.5$  mmol) and  $ZnCl_2$  (235 mg, 1.73 mmol) in 3 ml of DMF. Ion exchange

purification was accomplished with 0–0.9 M linear gradient of TEAB to yield 1790 mODU<sub>260</sub> (130 mg, 0.16 mmol, 84%) of **11** (TEA salt). ESI MS (–)  $m/z$  486.03 (Calc. for  $C_{12}H_{18}N_5O_{10}P_2S$ : 486.02)  $\delta_P$  35.80 (1P, d,  $J_{P1,P2}$  30.7, P2), –11.66 (1P, d,  $J_{P1,P2}$  30.7, P1).

**Guanosine 5'-(2-thiodiphosphate), GDP $\beta$ S (**12**).** Obtained starting from **8** (3570 mODU<sub>260</sub>, 150 mg, 0.30 mmol), thiophosphate TEA salt (prepared from 400 mg of trisodium salt hydrate,  $\sim 0.75$  mmol) and  $ZnCl_2$  (350 mg, 2.6 mmol) in 3 ml of DMF. Ion exchange purification was accomplished with 0–1.0 M linear gradient of TEAB to yield 2820 mODU<sub>260</sub> (180 mg, 0.23 mmol, 79%) of **12** (TEA salt). ESI MS (–)  $m/z$  457.89 (Calc. for  $C_{10}H_{14}N_5O_{10}P_2S$ : 457.99)  $\delta_P$  34.69 (1P, d,  $J_{P1,P2}$  31.4, P2), –11.57 (1P, d,  $J_{P1,P2}$  31.4 Hz, P1).

**7,2'-O-Dimethylguanosine 5'-(3-thiotriphosphate),  $m_2^{7,2'-O}$ -GTP $\gamma$ S (**15**).** Obtained starting from **14** (1600 mODU<sub>260</sub>, 100 mg, 0.14 mmol), thiophosphate TEA salt (obtained from 200 mg of trisodium salt hydrate,  $\sim 0.5$  mmol) and  $ZnCl_2$  (240 mg, 1.76 mmol) in 3.5 ml of DMF. Ion exchange purification was accomplished with a 0–1.1 M linear gradient of TEAB to yield 1290 mODU<sub>260</sub> (110 mg, 0.11 mmol, 80%) of **15** (TEA salt). ESI MS (–)  $m/z$  565.98 (Calc. for  $C_{12}H_{19}N_5O_{13}P_3S$ : 565.99)  $\delta_P$  39.18 (1P, d,  $J_{P2,P3}$  28.7, P3), –11.24 (1P, d,  $J_{P1,P2}$  20.4, P1), –23.83 (1P, dd,  $J_{P2,P3}$  28.7,  $J_{P1,P2}$  20.4, P2).

**P1-(7,2'-O-Dimethylguanosin-5'-yl) P4-guanosin-5'-yl (4-thio-tetraphosphate),  $m^{7,2'-O}$ Gpppp $\beta$ G (**1**).** To a suspension of **7** (1800 ODU<sub>260</sub>, 93 mg, 0.16 mmol) and **5** (1590 ODU<sub>260</sub>, 0.13 mmol) in DMF (5 ml) anhydrous  $ZnCl_2$  (244 mg, 2.54 mmol) was added and the mixture was shaken until reagents dissolved (1–2 min). The reaction was maintained at RT for 6 h and then quenched by being diluted with a solution of 940 mg (2.54 mmol) of disodium EDTA in 100 ml of water and adjusted to pH 6 with solid  $NaHCO_3$ . The product was isolated on DEAE Sephadex using 0–1.3 M gradient of TEAB to yield 1800 ODU<sub>260</sub> (0.080 mmol, 60%) of **1** (TEA salt) as a 7:10  $S_P:R_P$  diastereomeric mixture. The diastereomers were separated using 0–25% gradient of MeOH in AAC (pH = 5.9) over 120 min (**1a**:  $R_t$  = 43 min ( $S_P$ , D1), **1b**:  $R_t$  = 49 min ( $R_P$ , D2)) to yield 484 ODU<sub>260</sub> (19.4 mg, 0.02 mmol, 16%) of D1 (**1a**) and 761 ODU<sub>260</sub> (31 mg, 0.033 mmol, 25%) of D2 (**1b**) as  $NH_4^+$  salts. ESI MS (–)  $m/z$  911.05 (Calc. for  $C_{22}H_{31}N_{10}O_{20}P_4S_1$ : 911.04). **1a**: Analytical RP HPLC  $R_t$  = 6.8 min;  $\delta_H$  9.19 (1H, s, H(8)m<sup>7</sup>G); 8.06 (1H, s, H(8)G); 6.03 (1H, d,  $J_{1',2'}$  2.4, H(1')m<sup>7</sup>G); 5.83 (1H, d,  $J_{1',2'}$  6.0, H(1')G); 4.72 (1H,  $\sim t$ , H(2')G); 4.61 (1H,  $\sim t$ , H(3')m<sup>7</sup>G); 4.52 (1H,  $\sim t$ , H(3')G); 4.45–4.42 (7H, overlapped m, H(4')m<sup>7</sup>G, H(4')G, H(2')m<sup>7</sup>G, H(5')G, H(5'')G, H(5'), H(5'')m<sup>7</sup>G); 4.10 (3H, s,  $N-CH_3$ ); 3.58 (3H, s,  $O-CH_3$ ),  $\delta_P$  43.70 (1P, d,  $J_{P3,P4}$  26.0, P4), –11.11 (1P, d,  $J_{P1,P2}$  18.0, P1), –23.80 (1P,  $\sim t$ ,  $J_{P1,P2}$  18.0,  $J_{P2,P3}$  17.2, P2), –24.90 (1P, dd,  $J_{P3,P4}$  26.0,  $J_{P2,P3}$  17.2, P3). **1b**: Analytical RP HPLC  $R_t$  = 7.2 min;  $\delta_H$  9.16 (1H, s, H(8)m<sup>7</sup>G); 8.05 (1H, s, H(8)G); 5.99 (1H, d,  $J_{1',2'}$  2.4, H(1')m<sup>7</sup>G); 5.80 (1H, d,  $J_{1',2'}$  6.0, H(1')G); 4.75 (1H,  $\sim t$  overlapped with HDO, H(2')G); 4.60–4.45 (2H, overlapped m, H(3')m<sup>7</sup>G; H(3')G); 4.40–4.30 (3H, overlapped m, H(4')G, H(5')G, H(5'')G); 4.30–4.22 (3H, overlapped m, H(4')m<sup>7</sup>G,



H(5'), H(5'')m<sup>7</sup>G); 4.10 (3H, s, N-CH<sub>3</sub>); 3.59 (3H, s, O-CH<sub>3</sub>),  $\delta_P$  43.47 (1P, d,  $J_{P3,P4}$  26.3, P4), -11.27 (1P, d,  $J_{P1,P2}$  18.0; P1), -23.00 (1P, ~t,  $J_{P1,P2}$  18.0,  $J_{P2,P3}$  17.0; P2), -24.07 (1P, dd,  $J_{P3,P4}$  26.3,  $J_{P2,P3}$  17.0, P3).

**P1-(7,2'-O-Dimethylguanosin-5'-yl) P4-guanosin-5'-yl (3-thio-tetraphosphate), m<sub>2</sub><sup>7,2'-O</sup>Gppp<sub>s</sub>pG (2)**

**Method I.** To a suspension of **14** (Na salt, 102 mg; 1165 mODU<sub>260</sub>, 0.10 mmol) and **12** (TEA salt, 1730 mODU<sub>260</sub>, 0.14 mmol) in DMF (3 ml) anhydrous ZnCl<sub>2</sub> (250 mg, 1.84 mmol) was added and the mixture was shaken until reagents dissolved (1–2 min). The reaction was maintained at RT for 3 d and then quenched by addition of 685 mg (1.84 mmol) of EDTA in 100 ml of water and adjusted to pH 6 with solid NaHCO<sub>3</sub>. The product was isolated on DEAE Sephadex using 0–1.3 M gradient of TEAB to yield 795 mODU<sub>260</sub> (0.035 mmol, 35%) of **2** (TEA salt). The product was additionally purified by semi-preparative RP HPLC using isocratic 10% MeOH and 90% AAC (pH 5.9) yielding 640 mODU<sub>260</sub> (31 mg, 0.031 mmol, 28%) of **2**, NH<sub>4</sub><sup>+</sup> salt) as 8 : 10 (D1:D2, **2a:2b**) diastereomeric mixture.

**Method II.** To a suspension of **8** (Na salt, 107 mg; 2570 mODU<sub>260</sub>, 0.21 mmol) and **15** (TEA salt, 1290 mODU<sub>260</sub>, 110 mg, 0.11 mmol) in DMF (2 ml) anhydrous ZnCl<sub>2</sub> (775 mg, 5.69 mmol) was added and the mixture was shaken until reagents dissolved (1–2 min). The reaction was maintained at RT for 3 days and then quenched by addition of 2.11 g of EDTA (5.69 mmol) in 300 ml of water and neutralized with solid NaHCO<sub>3</sub>. The purification as in Method I yielded 1010 mODU<sub>260</sub> (0.045 mmol, 39%) of **2** (TEA salt) as 9 : 10 (**2a:2b**, D1:D2) diastereomeric mixture, and after additional HPLC purification 820 ODU<sub>260</sub> (37 mg, 0.036 mmol, 31%) of **2**. ESI MS (–) *m/z* 911.05 (calc. for C<sub>22</sub>H<sub>31</sub>N<sub>10</sub>O<sub>20</sub>P<sub>4</sub>S<sub>1</sub>: 911.04). **2a**: Analytical RP HPLC *R*<sub>t</sub> = 7.5 min,  $\delta_H$  9.04 (1H, s, H(8)m<sup>7</sup>G); 8.10 (1H, s, H(8)G); 5.97 (1H, d,  $J_{1',2'}$  2.9, H(1')m<sup>7</sup>G); 5.80 (1H, d,  $J_{1',2'}$  5.9, H(1')G); 4.70 (1H, ~t, H(2')G); 4.56 (1H, ~t, H(3')m<sup>7</sup>G); 4.50 (1H, ~t, H(3')G); 4.41 (1H, m, H(5')G); 4.34 (2H, m, overlapped H(4')m<sup>7</sup>G, H(4')G); 4.27 (2H, m overlapped, H(2')m<sup>7</sup>G, H(5'')G); 4.24 (2H, m, H(5'), H(5'')m<sup>7</sup>G); 4.08 (3H, s, N-CH<sub>3</sub>); 3.59 (3H, s, O-CH<sub>3</sub>).  $\delta_P$  30.49 (1P, t,  $J_{P2,P3}$  =  $J_{P3,P4}$  24.7, P3), -11.14 (1P, d,  $J_{P1,P2}$  17.6, P1), -11.88 (1P, d,  $J_{P3,P4}$  24.7, P4), -23.84 (1P, dd,  $J_{P2,P3}$  24.7,  $J_{P1,P2}$  17.6, P2) **2b**: Analytical RP HPLC *R*<sub>t</sub> = 7.6 min,  $\delta_H$  9.00 (1H, s, H(8)m<sup>7</sup>G); 8.08 (1H, s, H(8)G); 5.96 (1H, d,  $J_{1',2'}$  2.9, H(1')m<sup>7</sup>G); 5.81 (1H, d,  $J_{1',2'}$  5.9, H(1')G); 4.70 (1H, ~t, H(2')G); 4.55 (1H, ~t, H(3')m<sup>7</sup>G); 4.48 (1H, ~t, H(3')G); 4.40 (2H, overlapped m, H(4')m<sup>7</sup>G, H(5')G); 4.30 (3H, m (overlapped), H(2')m<sup>7</sup>G, H(4')G, H(5'')G); 4.25 (2H, m, H(5'), H(5'')m<sup>7</sup>G); 4.07 (3H, s, N-CH<sub>3</sub>), 3.62 (3H, s, O-CH<sub>3</sub>).  $\delta_P$  30.35 (1P, t,  $J_{P2,P3}$  =  $J_{P3,P4}$  24.7, P3), -11.14 (1P, d,  $J_{P1,P2}$  17.6, P1), -11.88 (1P, d,  $J_{P3,P4}$  24.7, P4), -23.86 (1P, dd,  $J_{P2,P3}$  24.7,  $J_{P1,P2}$  17.6, P2).

**P1-(7,2'-O-dimethylguanosin-5'-yl) P4-guanosin-5'-yl (2-thio-tetraphosphate), m<sub>2</sub><sup>7,2'-O</sup>Gpp<sub>s</sub>ppG (3).** To a suspension of **13** (80 mg, 1450 mODU<sub>260</sub>, 0.12 mmol) and **12** (895 mODU<sub>260</sub>, 0.078 mmol) in DMF (5 ml) anhydrous ZnCl<sub>2</sub> (172 mg, 1.27 mmol) was added and the mixture was shaken until

reagents dissolved (1–2 min). The reaction was maintained at RT for 3 days and then quenched by addition of 480 mg (1.27 mmol) of disodium EDTA in 45 ml of water and neutralized with solid NaHCO<sub>3</sub>. The product was isolated on DEAE Sephadex using 0–1.2 M gradient of TEAB to yield 560 mODU<sub>260</sub> (0.025 mmol, 32%) of **3**. The product was additionally purified by semi-preparative RP HPLC using isocratic 10% MeOH and 90% AAC (pH 5.9) yielding 435 ODU<sub>260</sub> (20 mg, 0.019 mmol, 24%) of **3**, NH<sub>4</sub><sup>+</sup> salt) as 7 : 10 (**3a:3b**, D1:D2) diastereomeric mixture. ESI MS (–) *m/z* 911.04 (calc. for C<sub>22</sub>H<sub>31</sub>N<sub>10</sub>O<sub>20</sub>P<sub>4</sub>S<sub>1</sub>: 911.04) **3a**: Analytical RP HPLC *R*<sub>t</sub> = 7.5 min,  $\delta_H$  9.04 (1H, s, H(8)m<sup>7</sup>G); 8.10 (1H, s, H(8)G); 5.97 (1H, d,  $J$  2.9, H(1')m<sup>7</sup>G); 5.80 (1H, d,  $J$  5.9, H(1')G); 4.70 (1H, ~t, H(2')G); 4.56 (1H, ~t, H(3')m<sup>7</sup>G); 4.50 (1H, ~t, H(3')G); 4.41 (1H, m, H(5')G); 4.34 (2H, m, overlapped H(4')m<sup>7</sup>G, H(4')G); 4.27 (2H, m overlapped, H(2')m<sup>7</sup>G, H(5'')G); 4.24 (2H, m, H(5'), H(5'')m<sup>7</sup>G); 4.08 (3H, s, N-CH<sub>3</sub>); 3.59 (3H, s, O-CH<sub>3</sub>),  $\delta_P$  30.21 (1P, t,  $J_{P1,P2}$  =  $J_{P2,P3}$  24.0; P2), -11.12 (1P, d,  $J_{P3,P4}$  18.5, P4), -11.96 (1P, d,  $J_{P1,P2}$  24.0, P1), -23.91 (1P, dd,  $J_{P2,P3}$  24.0,  $J_{P3,P4}$  18.5, P3). **3b**: Analytical RP HPLC *R*<sub>t</sub> = 7.6 min,  $\delta_H$  9.00 (1H, s, H(8)m<sup>7</sup>G), 8.08 (1H, s, H(8)G), 5.96 (1H, d,  $J$  2.9, H(1')m<sup>7</sup>G), 5.81 (1H, d,  $J$  5.9, H(1')G), 4.70 (1H, ~t, H(2')G); 4.55 (1H, ~t, H(3')m<sup>7</sup>G), 4.48 (1H, ~t, H(3')G), 4.40 (2H, overlapped m, H(4')m<sup>7</sup>G, H(5')G), 4.30 (3H, overlapped m, H(2')m<sup>7</sup>G, H(4')G, H(5'')G), 4.25 (2H, m, H(5'), H(5'')m<sup>7</sup>G), 4.07 (3H, s, N-CH<sub>3</sub>), 3.62 (3H, s, O-CH<sub>3</sub>),  $\delta_P$  30.21 (1P, t,  $J_{P1,P2}$  =  $J_{P2,P3}$  24.0, P2), -11.12 (1P, d,  $J_{P3,P4}$  18.5, P4), -11.96 (1P, d,  $J_{P1,P2}$  24.0, P1), -23.91 (1P, dd,  $J_{P2,P3}$  24.0,  $J_{P3,P4}$  18.5, P3).

**P1-(7,2'-O-dimethylguanosin-5'-yl) P4-guanosin-5'-yl (1-thio-tetraphosphate), m<sub>2</sub><sup>7,2'-O</sup>Gp<sub>s</sub>pppG (4).** To a suspension of **7** (Na salt, 65 mg; 1430 mODU<sub>260</sub>, 0.119 mmol) and **6** (TEA salt, 495 mODU<sub>260</sub>, 0.043 mmol) in DMF (6 ml) anhydrous ZnCl<sub>2</sub> (246 mg, 1.81 mmol) was added and the mixture was shaken until reagents dissolved (1–2 min). The reaction was maintained at RT for 6.5 h and then quenched by addition of 673 mg (1.81 mmol) of disodium EDTA in 100 ml of water and neutralized with solid NaHCO<sub>3</sub>. The product was isolated on DEAE Sephadex using 0–1.3 M gradient of TEAB to yield 635 mODU<sub>260</sub> (0.028 mmol, 65%) of **4** (TEA salt) as 7 : 10 (Sp:Rp, D1:D2, **4a:4b**) diastereomeric mixture. The diastereomers were separated by RP HPLC using 0–25% gradient of MeOH in AAC (pH = 5.9) over 120 min (**4a**: *R*<sub>t</sub> = 51 min (Sp, D1), **4b**: *R*<sub>t</sub> = 52 min (Rp, D2)) to yield 196 ODU<sub>260</sub> (8.5 mg, 0.0087 mmol, 20%) of **4a** and 308 ODU<sub>260</sub> (13.4 mg, 0.013 mmol, 31%) of **4b** (both as NH<sub>4</sub><sup>+</sup> salts). ESI MS (–) *m/z* 911.06 (calc. for C<sub>22</sub>H<sub>31</sub>N<sub>10</sub>O<sub>20</sub>P<sub>4</sub>S<sub>1</sub>: 911.04) **4a**: Analytical RP HPLC *R*<sub>t</sub> = 7.5 min,  $\delta_H$  9.04 (1H, s, H(8)m<sup>7</sup>G), 8.10 (1H, s, H(8)G), 5.97 (1H, d,  $J$  2.9, H(1')m<sup>7</sup>G), 5.80 (1H, d,  $J$  5.9, H(1')G), 4.70 (1H, ~t, H(2')G), 4.56 (1H, ~t, H(3')m<sup>7</sup>G), 4.50 (1H, ~t, H(3')G), 4.41 (1H, m, H(5')G), 4.34 (2H, overlapped m, H(4')m<sup>7</sup>G, H(4')G); 4.27 (2H, overlapped m, H(2')m<sup>7</sup>G, H(5'')G), 4.24 (2H, m, H(5'), H(5'')m<sup>7</sup>G), 4.08 (3H, s, N-CH<sub>3</sub>), 3.59 (3H, s, O-CH<sub>3</sub>).  $\delta_P$  43.92 (1P, d,  $J_{P1,P2}$  23.0, P1), -11.00 (1P, d,  $J_{P3,P4}$  16.6, P4), -22.71 (1P, ~t,  $J_{P3,P4}$  16.6,  $J_{P2,P3}$  16.5, P3), -23.63 (1P, dd,  $J_{P3,P4}$  23.0,  $J_{P2,P3}$  16.5, P2). **4b**: Analytical RP HPLC

$R_t = 7.8$  min:  $\delta_H$  9.00 (1H, s, H(8) m<sup>7</sup>G); 8.08 (1H, s, H(8)G); 5.96 (1H, d,  $J$  2.9, H(1')m<sup>7</sup>G); 5.81 (1H, d,  $J$  5.9, H(1')G); 4.70 (1H, ~t, H(2')G); 4.55 (1H, ~t, H(3')m<sup>7</sup>G); 4.48 (1H, ~t, H(3')G); 4.40 (2H, overlapped m, H(4')m<sup>7</sup>G, H(5')G); 4.30 (3H, overlapped m, H(2')m<sup>7</sup>G, H(4')G, H(5')G); 4.25 (2H, m, H(5'), H(5'')m<sup>7</sup>G); 4.07 (3H, s, N-CH<sub>3</sub>), 3.62 (3H, s, O-CH<sub>3</sub>).  $\delta_P$  43.52 (1P, d,  $J_{P1,P2}$  22.5, P4), -11.06 (1P, d,  $J_{P3,P4}$  15.6, P4), -22.89 (1P, ~t,  $J_{P3,P4}$  15.6,  $J_{P2,P3}$  16.5, P3), -23.74 (1P, dd,  $J_{P1,P2}$  22.5,  $J_{P2,P3}$  16.5, P2).

### Enzymatic reactions

Human DcpS was expressed in *Escherichia coli* according to the procedures described previously.<sup>58</sup> The protein of 15  $\mu$ M concentration was stored at -80 °C in 20 mM Tris buffer, pH 7.5, containing 50 mM KCl, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, and 20% glycerol. Snake Venom Phosphodiesterase I from *Crotalus adamanteus* (EC. 3.1.4.1) was purchased from Sigma (0.14 U/mg solid) and before use dissolved at 1 mg/mL concentration in 100 mM Tris buffer, pH 8.0, containing 100 mM NaCl, 14 mM MgCl<sub>2</sub> and 50% glycerol, and stored in this form at -20 °C.

hDcpS assay: A cap analog at 40  $\mu$ M concentration was treated with 3.0  $\mu$ L of DcpS in 500  $\mu$ L of 50 mM TRIS buffer, pH = 7.9, containing 20 mM MgCl<sub>2</sub> and 60 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> at 30 °C. After 15, 30, 75, 120 min and 24 h a 100  $\mu$ L sample was collected from the reaction mixture and deactivated by incubation in 90 °C for 3 min.

SVPDE assay: A cap analog at 40  $\mu$ M concentration was treated with 5.0  $\mu$ L of SVPDE in 500  $\mu$ L of 100 mM TRIS buffer, pH = 8.0, containing 14 mM MgCl<sub>2</sub> and 100 mM NaCl at 30 °C. After 15, 30, 90, 210, 330 min and 24 h a 50  $\mu$ L sample was collected from the reaction mixture and deactivated by incubation in 90 °C for 3 min.

All samples were analyzed without further treatment by analytical RP HPLC using a 0–25% linear gradient of methanol in 0.1 M KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> buffer (pH = 6.0) over 30 min.

### Binding affinities for eIF4E

Murine eukaryotic Initiation Factor eIF4E (residues 28–217) was expressed in BL21(DE3) *Escherichia coli* strain. The protein was purified from inclusion bodies then guanidinium-solubilized protein was refolded by a one-step dialysis and purified by ion-exchange chromatography on a HiTrapSP column without contact with cap analogs. The concentration of eIF4E was determined spectrophotometrically ( $\epsilon_{280} = 53,400$  cm<sup>-1</sup>M<sup>-1</sup>). Fluorescence titration measurements were carried out on an LS-55 spectrofluorometer (Perkin Elmer Co.) in 50 mM HEPES/ KOH (pH 7.2), 100 mM KCl, 0.5 mM EDTA, 1 mM DTT at 20.0  $\pm$  0.2 °C. 1  $\mu$ L aliquots of cap analog solutions with increasing concentration were added to 1.4 mL of 0.1  $\mu$ M protein solutions. Fluorescence intensities (excitation at 280 nm with 2.5 nm bandwidth and detection at 337 nm with 4-nm bandwidth and 290 nm cut off filter) were corrected taking sample dilution and the inner filter effect into account. Equilibrium association constants ( $K_{AS}$ ) were determined by fitting the theoretical dependence of fluorescence intensity on the total concentration of cap analog to the experimental

data points, according to equation described previously.<sup>31</sup> The concentration of protein was fitted as a free parameter of equilibrium equation showing amount of “active” protein. The final  $K_{AS}$  was calculated as a weighted average of 3–4 independent titrations, with the weights taken as the reciprocals of the numerical standard deviations squared. Numerical nonlinear least-squares regression analysis was performed using ORIGIN 6.0 (Microcal Software Inc.). The Gibbs free energy of binding was calculated from the  $K_{AS}$  value according to the standard equation  $\Delta G^\circ = -RT \ln K_{AS}$ .

### Synthesis of mRNA transcripts and measuring of translation efficiency *in vitro*

Capped, polyadenylated luciferase mRNAs were synthesized *in vitro* on a dsDNA template (amplified by PCR reaction) that contains: SP6 promoter sequence of DNA-dependent RNA polymerase, 5'UTR sequence of rabbit  $\beta$ -globin, the entire firefly luciferase ORF and a string of 31 adenosines. A typical *in vitro* transcription reaction mixture (40  $\mu$ L) contained: SP6 transcription buffer (Fermentas), 0.7  $\mu$ g of DNA template, 1 U/ $\mu$ L RiboLock Ribonuclease Inhibitor (Fermentas), 0.5 mM ATP/CTP/UTP and 0.1 mM GTP and 0.5 mM dinucleotide cap analog (molar ratio cap analog:GTP 5 : 1). The reaction mixture was preincubated at 37 °C for 5 min before addition of SP6 RNA polymerase (Fermentas) to final concentration of 1 U/ $\mu$ L and reaction was continued for 45 min at 37 °C. After incubation, reaction mixtures were treated with DNase RQ1 (Promega) in transcription buffer, for 20 min at 37 °C at concentration 1 U per  $\mu$ g of template DNA. RNA transcripts were purified using NucAway Spin Columns (Ambion), the integrity of transcripts was checked on a non-denaturing 1% agarose gel and concentrations were determined spectrophotometrically.

A translation reaction in RRL was performed in 10  $\mu$ L volume for 60 min at 30 °C, under conditions determined as optimal for cap-dependent translation. A typical reaction mixture contained: 40% RRL lysate (Flexi Rabbit Reticulocyte Lysate, Promega), mixture of amino acids (0.01 mM), MgCl<sub>2</sub> (0.6 mM), potassium acetate (170 mM) and 5'-capped mRNA. Four different concentrations of each analyzed transcript were tested in an *in vitro* translation reaction. The activity of synthesized luciferase was measured with a luminometer.

### Acknowledgements

We are indebted to Mike Kiledjian (Rutgers University) for providing the hDcpS-encoding plasmid and Zbigniew M. Darzynkiewicz for expressing and purifying the protein, to the Laboratory of Biological NMR (Institute of Biochemistry and Biophysics of the Polish Academy of Sciences, IBB PAS) for access to the NMR apparatus, and to the Laboratory of Mass Spectrometry (IBB PAS) for recording MS spectra. Financial support from the Polish Ministry of Science and Higher Education (No. NN301 243 436), the National Science Support Project 2008–2010 (No. PBZMNiSW-07/I/2007), the Howard Hughes Medical Institute (to E.D.; No. 55005 604) and the National Institute for General Medical Sciences (to R.E.R.; No. GM20818) is gratefully acknowledged.

## References

- Y. Furuchi and A. J. Shatkin, *Adv. Virus Res.*, 2000, **55**, 135.
- R. E. Rhoads, *J. Biol. Chem.*, 2009, **284**, 16711.
- A. Yamamoto, M. Kormann, J. Rosenecker and C. Rudolph, *Eur. J. Pharm. Biopharm.*, 2009, **71**, 484.
- R. Contreras, H. Cheroutre, W. Degraeve and W. Fiers, *Nucleic Acids Res.*, 1982, **10**, 6353.
- M. M. Konarska, R. A. Padgett and P. A. Sharp, *Cell*, 1984, **38**, 731.
- J. K. Yisraeli and D. A. Melton, *Methods Enzymol.*, 1989, **180**, 42.
- A. E. Pasquinelli, J. E. Dahlberg and E. Lund, *RNA*, 1995, **1**, 957.
- J. Stepinski, C. Waddell, R. Stolarski, E. Darzynkiewicz and R. E. Rhoads, *RNA*, 2001, **7**, 1486.
- Z. H. Peng, V. Sharma, S. F. Singleton and P. D. Gershon, *Org. Lett.*, 2002, **4**, 161.
- J. Jemielity, T. Fowler, J. Zuberek, J. Stepinski, M. Lewdorowicz, A. Niedzwiecka, R. Stolarski, E. Darzynkiewicz and R. E. Rhoads, *RNA*, 2003, **9**, 1108.
- E. Grudzien, M. Kalek, J. Jemielity, E. Darzynkiewicz and R. E. Rhoads, *J. Biol. Chem.*, 2006, **281**, 1857.
- M. Mockey, C. Goncalves, F. P. Dupuy, F. M. Lemoine, C. Pichon and P. Midoux, *Biochem. Biophys. Res. Commun.*, 2005, **340**, 1062.
- E. Grudzien-Nogalska, J. Jemielity, J. Kowalska, E. Darzynkiewicz and R. E. Rhoads, *RNA*, 2007, **13**, 1745.
- A. C. Gingras, B. Raught and N. Sonenberg, *Annu. Rev. Biochem.*, 1999, **68**, 913.
- J. Collier and R. Parker, *Annu. Rev. Biochem.*, 2004, **73**, 861.
- E. van Dijk, N. Cougot, S. Meyer, S. Babajko, E. Wahle and B. Seraphin, *EMBO J.*, 2002, **21**, 6915.
- Z. Wang, X. Jiao, A. Carr-Schmid and M. Kiledjian, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12663.
- D2 is a designation for one of the two synthesized P-diastereoisomers of this compound.
- H. Liu, N. D. Rodgers, X. Jiao and M. Kiledjian, *EMBO J.*, 2002, **21**, 4699.
- A. Zdanowicz, R. Thermann, J. Kowalska, J. Jemielity, K. Duncan, T. Preiss, E. Darzynkiewicz and M. W. Hentze, *Mol. Cell*, 2009, **35**, 881.
- W. Su and R. E. Rhoads, unpublished work.
- Only the ways in which the unmodified nucleotide is activated are taken into account, since there is practically no efficient method for activation of phosphorothioate nucleotides with an imidazole group.
- J. R. Moran and G. M. Whitesides, *J. Org. Chem.*, 1984, **49**, 704.
- T. Mukaiyama and M. Hashimoto, *Bull. Chem. Soc. Jpn.*, 1971, **44**, 2284.
- J. Kowalska, M. Lewdorowicz, E. Darzynkiewicz and J. Jemielity, *Tetrahedron Lett.*, 2007, **48**, 5475.
- In the case of **2a**, we failed to obtain high quality transcripts in the first attempt, and the amount of compound was insufficient to repeat the experiment. Therefore, unfortunately only **2b** was tested for translation efficiency *in vitro*.
- D. T. Major, V. Nahum, Y. Wang, G. Reiser and B. Fischer, *J. Med. Chem.*, 2004, **47**, 4405.
- P. Li, Z. H. Xu, H. Y. Liu, C. K. Wennefors, M. I. Dobrikov, J. Ludwig and B. R. Shaw, *J. Am. Chem. Soc.*, 2005, **127**, 16782.
- B. A. Connolly, P. J. Romaniuk and F. Eckstein, *Biochemistry*, 1982, **21**, 1983.
- J. Ludwig and F. Eckstein, *J. Org. Chem.*, 1989, **54**, 631.
- A. Niedzwiecka, J. Marcotrigiano, J. Stepinski, M. Jankowska-Anyszka, A. Wyslouch-Cieszyńska, M. Dadlez, A. C. Gingras, P. Mak, E. Darzynkiewicz, N. Sonenberg, S. K. Burley and R. Stolarski, *J. Mol. Biol.*, 2002, **319**, 615.
- J. Kowalska, M. Lewdorowicz, J. Zuberek, E. Grudzien-Nogalska, E. Bojarska, J. Stepinski, R. E. Rhoads, E. Darzynkiewicz, R. E. Davis and J. Jemielity, *RNA*, 2008, **14**, 1119.
- P. M. Burgers, F. Eckstein and D. H. Hunneman, *J. Biol. Chem.*, 1979, **254**, 7476.
- F. R. Bryant and S. J. Benkovic, *Biochemistry*, 1979, **18**, 2825.
- H. T. Ho and P. A. Frey, *Biochemistry*, 1984, **23**, 1978.
- N. Puri, S. Hunsch, C. Sund, I. Ugi and J. Chattopadhyaya, *Tetrahedron*, 1995, **51**, 2991.
- G. M. Blackburn, X. Liu, A. Rosler and C. Brenner, *Nucleosides, Nucleotides Nucleic Acids*, 1998, **17**, 301.
- Q. Han, S. G. Sarafianos, E. Arnold, M. A. Parniak, B. L. Gaffney and R. A. Jones, *Org. Lett.*, 2007, **9**, 5243.
- Q. W. Han, S. G. Sarafianos, E. Arnold, M. A. Parniak, B. L. Gaffney and R. A. Jones, *Tetrahedron*, 2009, **65**, 7915.
- V. Nahum, M. Tulapurkar, S. A. Levesque, J. Seigny, G. Reiser and B. Fischer, *J. Med. Chem.*, 2006, **49**, 1980.
- S. Mikkola, S. Salomaki, Z. Zhang, E. Maki and H. Lonnberg, *Curr. Org. Chem.*, 2005, **9**, 999.
- M. Kadokura, T. Wada, C. Urashima and M. Sekine, *Tetrahedron Lett.*, 1997, **38**, 8359.
- A. M. Rydzik, M. Lukaszewicz, J. Zuberek, J. Kowalska, Z. M. Darzynkiewicz, E. Darzynkiewicz and J. Jemielity, *Org. Biomol. Chem.*, 2009, **7**, 4763.
- A. R. Kore, M. Shanmugasundaram, I. Charles, A. M. Cheng and T. J. Barta, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 5295.
- K. Misiura, D. Szymanowicz and W. J. Stec, *Org. Lett.*, 2005, **7**, 2217.
- V. Jankowski, M. van der Giet, H. Mischak, M. Morgan, W. Zidek and J. Jankowski, *Br. J. Pharmacol.*, 2009, **157**, 1142.
- V. Jankowski, M. Tolle, R. Vanholder, G. Schonfelder, M. van der Giet, L. Henning, H. Schluter, M. Paul, W. Zidek and J. Jankowski, *Nat. Med.*, 2005, **11**, 223.
- The differences in the phosphate designations between tri- and tetraphosphate cap analogs arise from the fact that due to biochemical nomenclature the  $\alpha$ -phosphate is always the one attached to 5'-hydroxyl of mRNA transcript after incorporation of the cap analog into mRNA.
- S. Holtkamp, S. Kreiter, A. Selmi, P. Simon, M. Koslowski, C. Huber, O. Tureci and U. Sahin, *Blood*, 2006, **108**, 4009.
- K. Karikó, H. Muramatsu, F. A. Welsh, J. Ludwig, H. Kato, S. Akira and D. Weissman, *Mol. Ther.*, 2008, **16**, 1833.
- D. A. Mitchell and S. K. Nair, *J. Clin. Invest.*, 2000, **106**, 1065.
- A. de Benedetti and J. R. Graff, *Oncogene*, 2004, **23**, 3189.
- S. W. Liu, X. Jiao, H. Liu, M. Gu, C. D. Lima and M. Kiledjian, *RNA*, 2004, **10**, 1412.
- V. Shen, H. Liu, S. W. Liu, X. Jiao and M. Kiledjian, *RNA*, 2008, **14**, 1132.
- J. Singh, M. Salcius, S. W. Liu, B. L. Staker, R. Mishra, J. Thurmond, G. Michaud, D. R. Mattoon, J. Printen, J. Christensen, J. M. Bjornsson, B. A. Pollok, M. Kiledjian, L. Stewart, J. Jarecki and M. E. Gurney, *ACS Chem. Biol.*, 2008, **3**, 711.
- J. Kusmierek and D. Shugar, *Nucleic Acids Res.*, Special Publ. No. 4, 73.
- J. W. Jones and R. K. Robins, *J. Am. Chem. Soc.*, 1963, **85**, 193.
- L. S. Cohen, C. Mikhli, C. Friedman, M. Jankowska-Anyszka, J. Stepinski, E. Darzynkiewicz and R. E. Davis, *RNA*, 2004, **10**, 1609.