Organic & Biomolecular Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Journal Name

RSCPublishing

ARTICLE

Towards novel efficient and stable nuclear import signals: synthesis and properties of trimethylguanosine cap analogs modified within the 5′,5′-triphosphate bridge

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Malgorzata Zytek[a], Joanna Kowalska[a], Maciej Lukaszewicz[a], Blazej A. Wojtczak[b], Joanna Zuberek[a], Aleksandra Ferenc-Mrozek[b], Edward Darzynkiewicz[a,b], Anna Niedzwiecka[a,c], and Jacek Jemielity[a,b,*]

A trimethylguanosine (TMG) cap is present at the 5′ end of several small nuclear and nucleolar RNAs. Recently, it has been reported that the TMG cap is a potential nuclear import signal for nucleus-targeting therapeutic nucleic acids and proteins. The import is mediated by recognition of the TMG cap by the snRNA transporting protein, snurportin1. This work describes the synthesis and properties of a series of dinucleotide TMG cap ($m_3^{2,2,7}$GpppG) analogs modified in the 5′,5′-triphosphate bridge as tools to study TMG cap-dependent biological processes. The bridge was altered at different positions by introducing either bridging (imidodiphosphate, O to NH and methylenebisphosphonate, O to CH_2) or non-bridging (phosphorothioate, O to S and boranophosphate, O to BH_3) modifications, or by elongation to tetraphosphate. The stability of novel analogs in blood serum was studied to reveal that the α,β-bridging O to NH substitution ($m_3^{2,2,7}$GppNHpG) confers the highest resistance. Short RNAs capped with analogs containing α,β-bridging ($m_3^{2,2,7}$GppNHpG) or β-non-bridging ($m_3^{2,2,7}$Gpp$_S$pG D2) modifications were resistant to decapping pyrophosphatase, hNudt16. Preliminary studies on binding by human snurportin1, revealed that both O to NH and O to S substitutions support this binding. Due to favorable properties in all three assays, $m_3^{2,2,7}$GppNHpG was selected as a promising candidate for further studies on the efficiency of the TMG cap as a nuclear import signal.

Introduction

A distinctive structural feature of all RNAs synthesized by RNA polymerase II is the presence of a methylated 5’ cap.[1] The core feature of the cap is the presence of a distal N^7-methylguanosine (m^7G) connected to the 5’-terminal nucleoside of the RNA chain by a 5’,5’-triphosphate bridge (Fig. 1). Some additional methylations might occur either within the core cap structure or at the first transcribed nucleotides. The specific methylation pattern varies between different RNA types and different organisms.[2] For example, mRNAs usually carry a monomethylguanosine cap (MMG cap or m^7G cap), which may be further divided into cap 0, cap 1 and cap 2 subclasses, depending on the presence of additional methylations at the first two transcribed nucleotides (Fig. 1). The MMG cap participates in most steps of mRNA metabolism, including maturation, initiation of translation, intracellular transport and turnover.[3] Some small non-coding RNAs residing in the nucleus, such as snRNAs involved in pre-mRNA splicing or snoRNAs involved in rRNA synthesis, possess a trimethylguanosine cap (TMG or m_3G cap), which is double methylated at N^2 position in addition to the methylation at N^7 position of guanosine (Fig. 1). [4] TMG cap is also found at the 5´ end of nematode mRNAs (e.g., *Caenorhabditis elegans* and *Ascaris suum*)[5]. In higher organisms (including mammals), some nascent, immature snRNAs possessing an MMG cap are exported to the cytoplasm, where they undergo double N^2 methylation by the RNA-(guanosine-N2)-methyltransferase TGS1[6] and assemble with proteins to form matured snRNPs. The TMG cap subsequently serves as a nuclear localization signal (NLS) for snRNPs.[7] This nuclear import is mediated by snurportin1, an adaptor protein that selectively recognizes the TMG cap and binds to importin β, a receptor protein, which enables translocation through the nuclear pore complex.[8] Based on the natural mechanism of the TMG-mediated nuclear import of snRNPs, Moreno *et al.* proposed a new strategy for increasing the efficiency of nucleus-targeting therapeutics, such as splice-correcting oligonucleotides.[9]

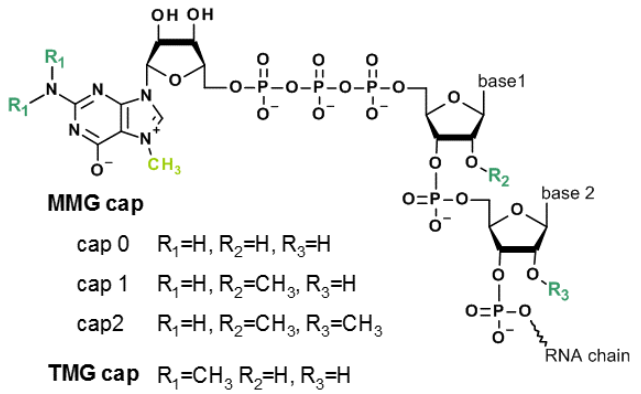
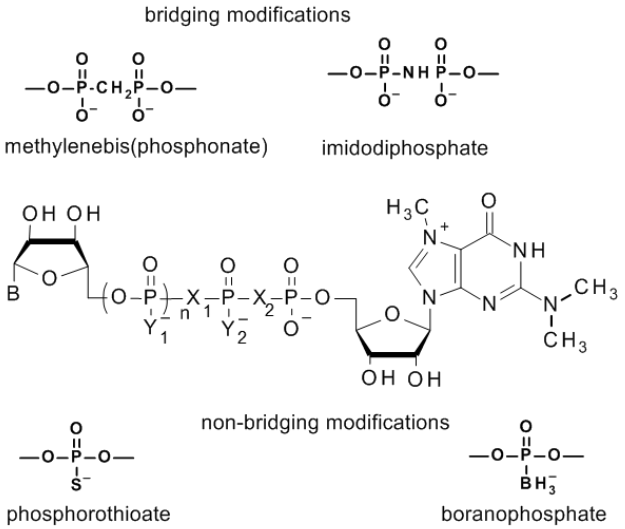


Fig. 1. Structure of monomethylguanosine (MMG) and trimethylguanosine TMG RNA caps.

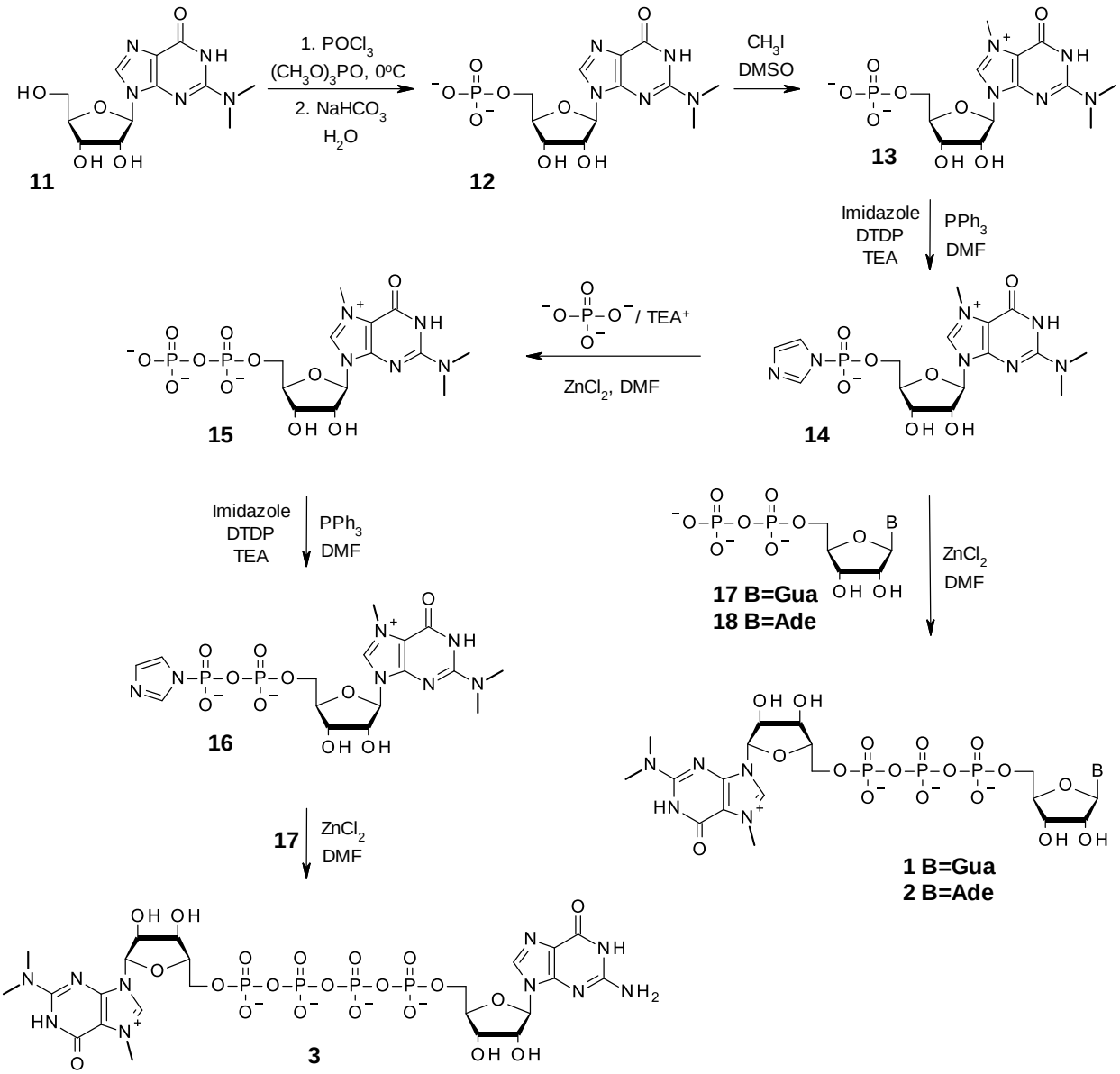
It has been shown that the addition of a synthetic TMG cap to model splice-correcting oligonucleotides increases their nuclear concentration and, in turn, results in higher levels of correct splicing. These findings made the TMG cap a promising candidate for a universal, small-size NLS for therapeutic (macro)molecules and bioconjugates. TMG cap addition should particularly benefit applications of splice-correcting oligonucleotides in experimental therapies for genetic disorders such as Duchenne Muscular Dystrophy[10] or Spinal Muscular Atrophy[11]. However, the influence of TMG cap size, attachment mode, and other modifications on affinity for snurportin, and efficiency as an NLS have not yet been examined. Moreover, when delivered into the human body as a part of a therapeutic agent, the TMG cap could be prone to degradation by extra- and intracellular pyrophosphatases. It is reasonable to assume that the effectiveness of a TMG cap as an NLS would be, among others, dependent on its sufficient stability under physiological conditions and its affinity to snurportin. A TMG cap analog with favorable properties should be characterized by increased resistance to enzymatic degradation without any loss of affinity to snurportin. Little is known about degradation pathways specific for TMG caps; however, it has been shown that TMG-capped RNAs are decapped by the NudT16 enzyme,[12] which cleaves the triphosphate bridge between α and β phosphates resulting in 5'-diphosphate N^2,N^2,N^7-trimethylguanosine ($m_3^{2,2,7}$GDP) and 5'-phosphorylated snoRNA. Numerous phosphate modifications have previously been shown to be useful for the physiological stabilization of MMG cap in the context of therapeutic applications.[13] For example, the half-life of exogenously delivered capped-mRNAs may be significantly improved by replacing the β-phosphate with a phosphorothioate[14] or boranophosphate[15] group or introducing methylenebisphosphonate[16] or imidodiphosphate[15] moieties at the α,β-position in the triphosphate bridge of the MMG cap.[17]

Table 1. TMG cap analogs synthesized in this study



No[a]	Abbreviation	X_1	X_2	Y_1	Y_2	n	B
1	$m_3^{2,2,7}$GpppG	O	O	O	O	1	Gua
2	$m_3^{2,2,7}$GpppA	O	O	O	O	1	Ade
3	$m_3^{2,2,7}$GppppG	O	O	O	O	2	Gua
4	$m_3^{2,2,7}$GppCH$_2$pG	CH$_2$	O	O	O	1	Gua
5	$m_3^{2,2,7}$GpCH$_2$ppG	O	CH$_2$	O	O	1	Gua
6	$m_3^{2,2,7}$GppNHpG	NH	O	O	O	1	Gua
7	$m_3^{2,2,7}$GpNHppG	O	NH	O	O	1	Gua
8a,b	$m_3^{2,2,7}$Gppp$_S$G (D1) and (D2)	O	O	S	O	1	Gua
9a,b	$m_3^{2,2,7}$Gpp$_S$pG (D1) and (D2)	O	O	O	S	1	Gua
10a,b	$m_3^{2,2,7}$Gpp$_{BH3}$pG (D1) and (D2)	O	O	O	BH$_3$	1	Gua

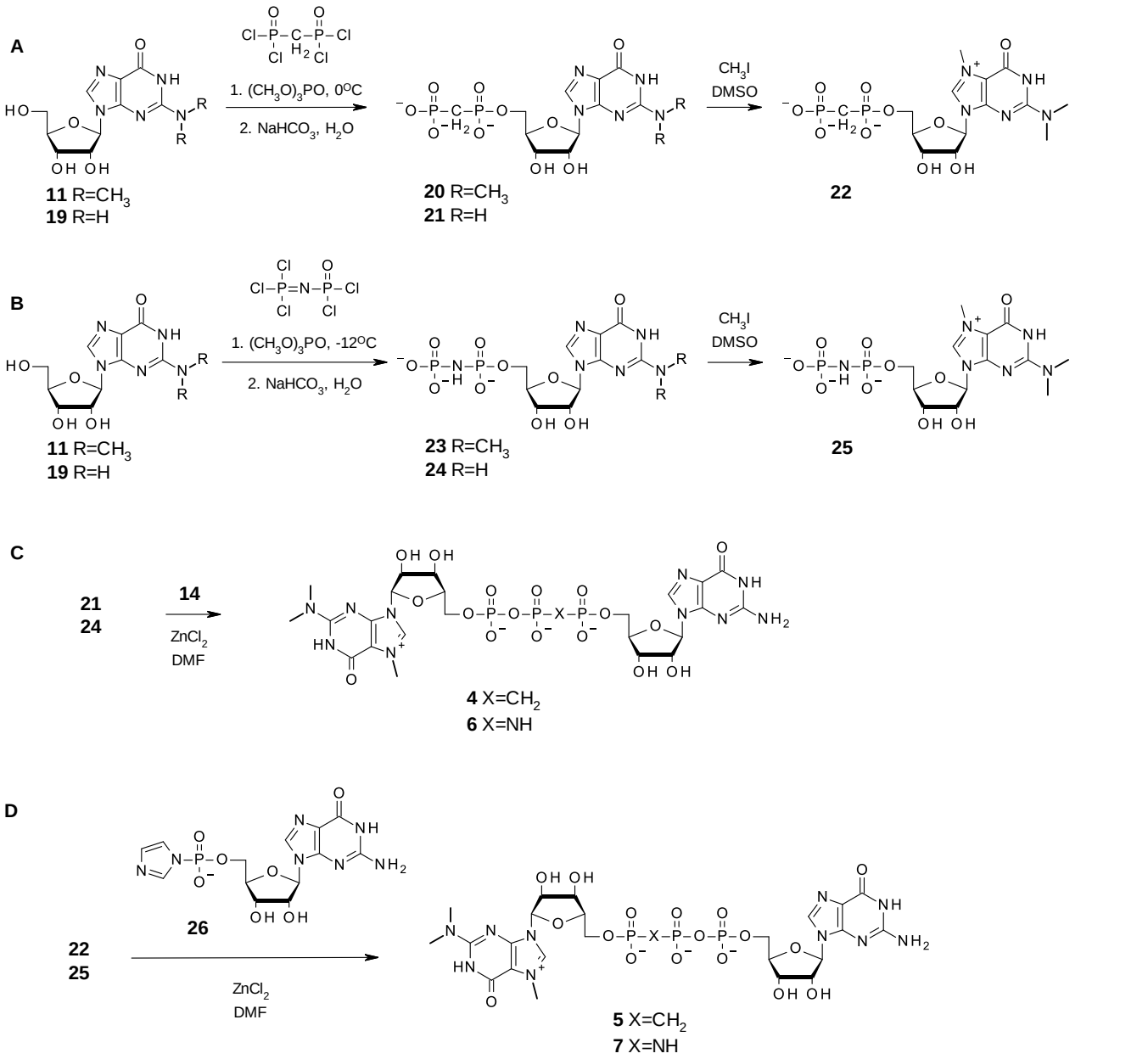
[a] a and b refer to either of the two P-diastereomers of a given compound, D1 and D2. D1 denotes the isomer eluting faster from a reversed-phased HPLC column.



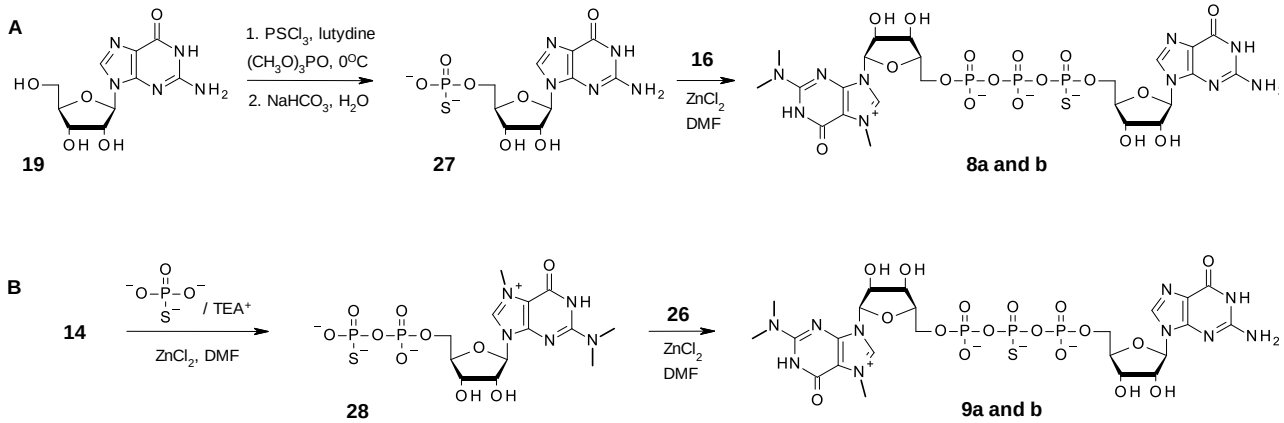
Scheme 1. The synthesis of non-modified TMG cap analogs $m_3^{2,2,7}$GpppG (**1**), $m_3^{2,2,7}$GpppA (**2**), and TMG cap analog with extended 5',5'-phosphate bridge $m_3^{2,2,7}$GppppG (**3**)

The β-phosphorothioate and β-boranophosphate MMG cap analogs had higher affinity to the eIF4E translation initiation factor, which resulted in the higher translational efficiency of mRNAs capped with those analogs, making them beneficial for therapeutic applications such as mRNA-based anticancer immunization[18]. On the other hand, it has been found that the modification of γ-phosphate or β,γ-pyrophosphate bond in free MMG cap analogs has the greatest influence on their stability under physiological conditions[19,20]. Recent studies indicate that phosphate-modifications may also be beneficial for augmenting the stability of TMG caps. Honcharenko *et al.*[21] found that α-β methylenebis(phosphonate) modification at the 5',5'-triphosphate bridge of a free TMG cap ($m_3^{2,2,7}$GppCH$_2$p$^{2'OMe}$A) increases its stability in HeLa cell lysates in contrast to the same modification at the β,γ-position. However, to optimize TMG cap analog properties as an NLS, a more thorough examination of degradation under various conditions is required. It is currently not clear whether there are other enzymes, besides hNudt16, which are responsible for TMG cap degradation. Also, different pyrophosphatase activities may target TMG capped-oligonucleotides, TMG cap-bioconjugates

and free TMG cap-structures, both intra- and extracellularly. Finally, the influence of phosphate modifications on interaction with snurportin1 should also be evaluated to eliminate modifications that would impair TMG efficiency as an NLS. Studies on the interaction of MMG cap analogs with specific proteins have revealed that the same chemical modification may have a strikingly different influence on cap-protein complex stability depending on the structural details of the cap-binding site.[13, 22]



Scheme 2. The synthesis of TMG cap analogs modified with methylenebisphosphonate or imidodiphosphate in a 5',5'-triphosphate bridge. (A) Synthesis of methylenebisphosphonate modified nucleosides; (B) Synthesis of imidodiphosphate modified nucleosides; (C) Synthesis of TMG cap analogs modified in the 5',5'-triphosphate bridge at α-β position; (D) Synthesis of TMG cap analogs modified at β-γ position of the 5',5'-triphosphate bridge.



Scheme 3. The synthesis of TMG cap analogs possessing a phosphorothioate modification in α (**8a** and **b**) or β (**9a** and **b**) positions in the 5',5'triphosphate bridge.

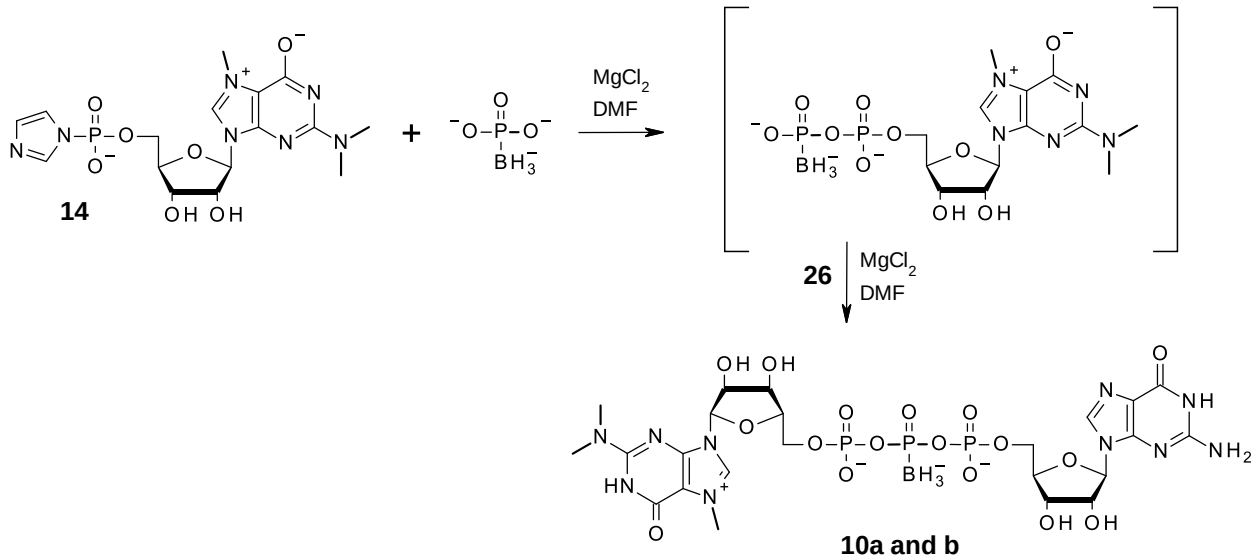
Here, we created a set of novel synthetic TMG cap analogs as a toolbox that can be utilized for addressing the pending questions regarding TMG cap degradation. Several phosphate-modified TMG cap analogs were synthesized and preliminarily characterized by biochemical and biophysical assays. The novel analogs were modified at various positions of the 5',5'-triphosphate bridge by replacing either one of the bridging oxygens with CH_2 or NH moieties, or one of the non-bridging oxygens with S or BH_3 (Table 1). All synthesized analogs possessed guanosine as a second nucleoside to enable their efficient incorporation into RNA transcripts by *in vitro* transcription with SP6 RNA Polymerase. Selected dinucleotide TMG cap analogs and TMG-capped transcripts were then used to study TMG cap degradation to determine which positions of the triphosphate bridge in the TMG cap structure are crucial for its physiological stability, either related to the activity of the specific decapping protein hNudt16 or to unspecific degradation in blood serum. Two tested modifications of the triphosphate bridge increased the resistance against decapping enzyme hNudt16; however, the resistance was strictly dependent on the position of the modification. The influence of selected modifications on interaction with snurportin1 was determined by fluorescence spectroscopy. No destabilizing effect and even enhancement of the interaction with snurportin1 was observed for the tested compounds. Hence, the results suggest that triphosphate-bridge modifications may confer increased stability on TMG cap analogs, while maintaining their affinity to nucleus-importing machinery and, thus, may benefit delivery strategies for nucleus-targeting therapeutics.

Results and discussion

Chemical synthesis

Several synthetic approaches have been developed for the synthesis of mono-, di- and oligonucleotide TMG cap analogs.[9,23-28] Among these, there are only two examples of TMG cap analogs bearing a modified triphosphate bridge, either by a methylenebisphosphonate group[21] or by the presence of an ethylene glycol linkage between two phosphate subunits.[29] The dinucleotide TMG cap analogs **1-10** synthesized in this work include phosphate-unmodified analogs **1** and **2** (reference compounds), analog **3** bearing a tetraphosphate instead of triphosphate bridge and analogs modified with either a methylenebis(phosphonate) (**4, 5**), imidodiphosphate (**6, 7**), phosphorothioate (**8, 9**) or boranophosphate (**10**) moiety (Table 1). The synthesis of analogs **1**–**10** was accomplished based on phosphorimidazolide chemistry (Schemes 1–4), in which one nucleotide subunit is converted into an electrophilic P-imidazolide and then reacted with a second subunit acting as a nucleophile. The reaction is accelerated by addition of divalent metal chloride excess, usually zinc or magnesium. The metal chlorides increase the solubility of nucleotides in DMF, activate the imidazole as a leaving group and presumably also serve as a template for the reacting nucleotides. The key starting material in the synthesis of all TMG cap analogs was N^2,N^2-dimethylguanosine (**11**), for which several synthetic routes have been described, amongst which we selected a three-step route based on reductive methylation. [24a,30] Compound **11** was then efficiently converted into either dimethylguanosine 5'-monophosphate (**12**), 5'-imidodiphosphate (**23**) or 5'-methylenebisphosphonate (**20**) using an appropriate phosphorylating reagent, $POCl_3$, $Cl_2(PO)CH_2(PO)Cl_2$ or $Cl_2(PO)NPCl_2$, respectively, in trimethylphosphate at 0 or -12 ºC (Scheme 1 and Scheme 2).[31] To obtain phosphate-

unmodified analogs **1-3** (Scheme 1), N^7,N^2,N^2-trimethylguanosine 5'-monophosphate (**13**) was synthesized by N^7-methylation of compound **12** with a methyl iodide. Next, compound **13** was converted into its P-imidazolide (**14**) by condensation with imidazole using the Mukaiyama and Hashimoto oxidation-reduction conditions.[32] Compound **14** was coupled to triethylammonium salts of either GDP (**17**) or ADP (**18**) in DMF in the presence of an 8-fold $ZnCl_2$ excess to afford analogs **1** and **2** in, respectively, 83% and 36% isolated yields after ion-exchange chromatography. To obtain the tetraphosphate cap analog **3,** compound **14** was first coupled with triethylammonium orthophosphate, and the resulting N^7,N^2,N^2-trimethylguanosine 5'-diphosphate (**15**) was converted into P-imidazolide (**16**) using the same method as for compound **13**. A $ZnCl_2$-mediated coupling of **16** with GDP **17** gave analog **3** with a 41% yield.



Scheme 4. The synthesis of a TMG cap analog containing a boranophosphate group at the β position of the 5',5' triphosphate bridge.

To obtain analogs containing CH_2 or NH substitutions at either the α-β (**4**, **6**) or β-γ (**5**, **7**) positions of the 5',5'-triphosphate bridge (Scheme 2), guanosine or $m_3^{2,2,7}$-trimethylguanosine 5'-methylenebis(phosphonate) (**21**, **22**) or 5'-imidodiphosphate (**24**, **25**) were coupled with appropriate P-imidazolides (**14**, **26**) in the presence of excess $ZnCl_2$. The preparative yields ranged from 36 to 51% after ion-exchange purification; however, the NH-substituted analogs (**6, 7**) required additional purification on semi-preparative HPLC. The synthesis of analogs bearing a non-bridging phosphorothioate modification at either α- or β-position of the triphosphate bridge is depicted in Scheme 3. Guanosine 5'-thiophosphate (**27**) was synthesized by thiophosphorylation of guanosine (**19**) by $PSCl_3$ in trimethylphosphate in the presence of 2,6-lutidine.[33] Trimethylguanosine 5'-β-thiodiphosphate (**28**) was obtained by coupling **14** with triethylammonium thiophosphate in the presence of zinc chloride in DMF.[34] In this case, it was necessary to methylate the N^7 position before introducing a phosphorothioate moiety, since otherwise the sulfur atom would preferably undergo reaction with methyl iodide. Coupling compounds **27** and **16** gave a mixture of two P-diastereoisomers of analog **8** (**8a** – isomer D1 and **8b** – isomer D2), whereas coupling compounds **26** and **28** led to a mixture of two diastereoisomers of compound **9**. Finally, two diastereoisomers of analog **10** (**10a** and **10b**) containing a boranophosphate group at the β position of the 5',5'-triphosphate bridge were synthesized (Scheme 4). Because of the instability of nucleoside β-boranodiphosphates,[18b] the synthesis of compound **10** was performed by a one-pot two step procedure. First, compound **14** was coupled with triethylammonium boranophosphate in DMF in the presence of excess $MgCl_2$ to yield N^7,N^2,N^2-trimethylguanosine β-boranodiphosphate, to which, without isolation, compound **26** and a new portion of magnesium chloride in DMF were added, finally producing two diastereoisomers of compound **10** (**a** and **b** mixture) along with a side product resulting from the self-coupling reactions. The isolated product was obtained with a satisfactory yield of 43%. All the compounds that were isolated as P-diastereoisomeric mixtures (**8**, **9** and **10**) were subsequently separated by semi-preparative RP HPLC and thus could be subjected to further studies as diastereomerically pure samples.

The structures and homogeneities of all the synthesized compounds were confirmed by HRMS, ^{1}H NMR and ^{31}P NMR.

The presence of each phosphate modification results in a characteristic δ_P shift relative to δ_P of the unmodified cap analog (**1**) (Fig. 2, Table 2) and, thus, both the type and position of the modification can be easily confirmed by ^{31}P NMR.

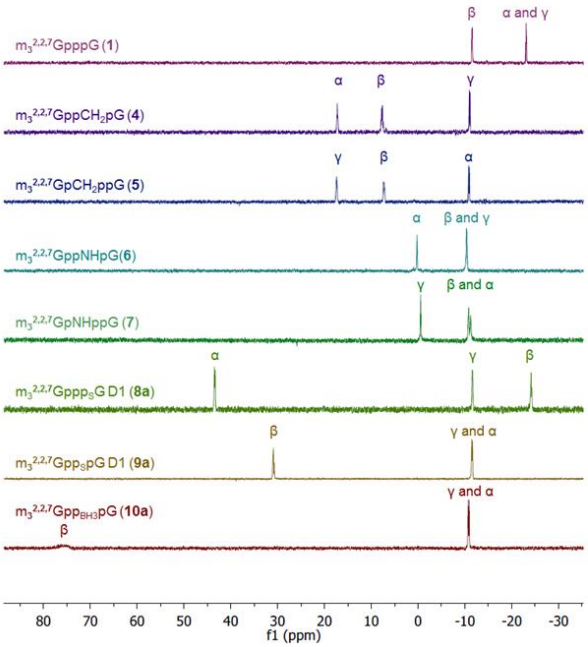


Fig. 2. ^{31}P NMR spectra of phosphate-modified TMG cap analogs.

Stability in Fetal Bovine Serum

To determine the influence of various phosphate-modifications on the general physiological stability of TMG cap analogs, we studied their degradation in fetal bovine serum (FBS). FBS contains various nucleases capable of hydrolyzing pyrophosphate and phosphodiester bonds and has been previously employed to study the general stability of modified nucleotides and oligonucleotides under physiological conditions.[35] To study the serum stability, each TMG cap analog was incubated in 1% FBS at 37 ºC at a final concentration of 10 µM. The samples were analyzed at different time points by RP HPLC (Fig. 3 A,B). Based on the chromatographic integration, a percentage of the remaining product was estimated and plotted against time. The experimental data were fitted to an exponential decay model to determine half-life values. The phosphate-unmodified analogs **1** ($m_3^{2,2,7}$GpppG) and **2** ($m_3^{2,2,7}$GpppA) had half-lives of 71 ± 4 min and 105 ± 2 min, respectively.

Table 2. Influence of various modifications on δ_P chemical shifts in TMG cap analogs

Type of modification	$\Delta\delta_P$ (ppm) modified phosphate *versus* unmodified (ppm)
methylenebisphosphonate (P-CH_2-P)	+22
imidodiphosphate (P-NH-P)	+11
phosphorothioate (–O-)$_2$(PO)-S^-)	+55
Boranophosphate (–O-)$_2$(PO)-BH_3^-)	+95

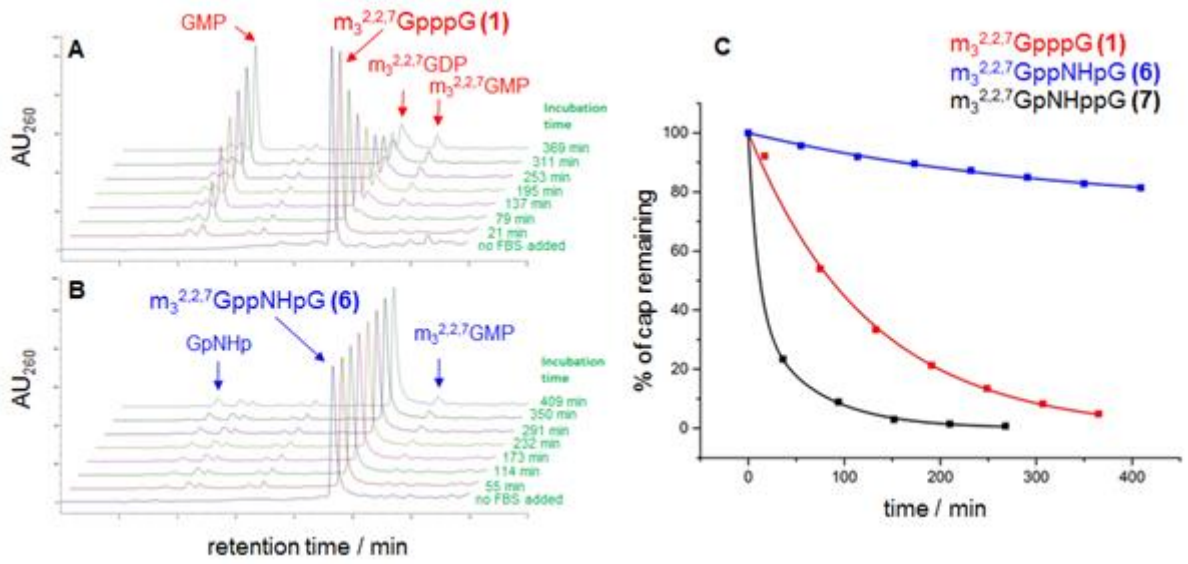


Fig. 3. Stability of TMG cap analogs in FBS studied by an HPLC-based assay. (A) Decay of cap analogs **1** (B) and 6 in 1% FBS monitored by RP HPLC. (C) Comparison of the stability of cap analogs **1**, **6** and **7** based on the HPLC assay.

Table 3. Stability of phosphate-modified TMG cap analogs in FBS. The stability was tested for 10 µM cap analogs in 1% FBS diluted with PBS (pH 7.2).

	Compound	Half-life (min)
1	$m_3^{2,2,7}$GpppG	70.6 ± 3.5
2	$m_3^{2,2,7}$GpppA	105 ± 2
3	$m_3^{2,2,7}$GppppG	46 ± 1
4	$m_3^{2,2,7}$GppCH$_2$pG	40 ± 2
5	$m_3^{2,2,7}$GpCH$_2$ppG	29 ± 2
6	$m_3^{2,2,7}$GppNHpG	>400
7	$m_3^{2,2,7}$GpNHppG	12 ± 1
8a	$m_3^{2,2,7}$Gppp$_S$G (D1)	200 ± 4
8b	$m_3^{2,2,7}$Gppp$_S$G (D2)	154 ± 8
9a	$m_3^{2,2,7}$Gpp$_S$pG (D1)	53 ± 1
9b	$m_3^{2,2,7}$Gpp$_S$pG (D2)	50 ± 6
10a	$m_3^{2,2,7}$Gpp$_{BH3}$pG (D1)	45 ± 2
10b	$m_3^{2,2,7}$Gpp$_{BH3}$pG (D2)	75 ± 10

The studied phosphate modifications influenced the cap stability to various extents, wherein both the modification type as well as its position in the triphosphate bridge was important (Table 3).

Cap analogs **9a** ($m_3^{2,2,7}$Gpp$_S$pG D1), **9b** ($m_3^{2,2,7}$Gpp$_S$pG D2), **10a** ($m_3^{2,2,7}$Gpp$_{BH3}$pG D1) and **10b** ($m_3^{2,2,7}$Gpp$_{BH3}$pG D2) had half-lives comparable to compound **1**, indicating that non-bridging modifications at the β-phosphate moiety have little influence on the stability in FBS. Cap analogs **8a** ($m_3^{2,2,7}$Gppp$_S$G D1), **8b** ($m_3^{2,2,7}$Gppp$_S$G D2) and **6** ($m_3^{2,2,7}$GppNHpG) had notably longer half-lives of 200 ± 4 min, 154 ± 8 and above 400 min, respectively, indicating that the bridging of the α,β-bridging modifications and α-non-bridging modifications of the triphosphate bridge may increase the stability in FBS. The significant stabilization observed for $m_3^{2,2,7}$GppNHpG (**6**) and the analysis of the degradation products (Fig. 3) suggest that the degradation of $m_3^{2,2,7}$GpppG-type nucleotides in FBS occurs mainly through the cleavage of α,β-pyrophosphate bonds. Surprisingly however, the compound **4** ($m_3^{2,2,7}$GppCH$_2$pG)**,** which is modified at the α,β-bridging position similarly to compound **6** but contains the CH_2 bridge instead of the NH bridge, did not show increased stability ($t_{1/2}$ = 40 ± 2 min). Furthermore, both O to NH and O to CH_2 substitutions at the β,γ-position have shown rather a destabilizing effect, which was most pronounced for NH-substituted compound **7** ($t_{1/2}$ = 12 ± 1 min).

Honcharenko *et al.* previously studied two TMG analogs containing 2'-O-methyladenosine as the second nucleoside and modified with a CH_2 bridge at either α,β-position ($m_3^{2,2,7}$GppCH$_2$pA$_{2'\text{-O-Me}}$) or β,γ-position ($m_3^{2,2,7}$GpCH$_2$ppA$_{2'\text{-O-Me}}$) under similar conditions[21]. Interestingly, they found that the α,β-substitution with CH_2 highly stabilized the compound in FBS, which is in contrast to our observations for analogs of $m_3^{2,2,7}$GpppG. This may indicate that the degradation pathways for different dinucleotide cap analogs in serum are also strongly dependent on the structure of the second nucleoside.

Susceptibility to hNudt16

Next, we studied the influence of selected phosphate-modified TMG cap analogs on susceptibility to the human decapping enzyme, hNudt16.[12] hNudt16, similar to the mRNA decapping enzyme - Dcp2,[36] belongs to the NUDIX family of phosphohydrolases,[37] which contain acidic, metal binding amino acids in the catalytic site and utilize substrates composed of a Nucleoside Diphosphate linked to a moiety X.

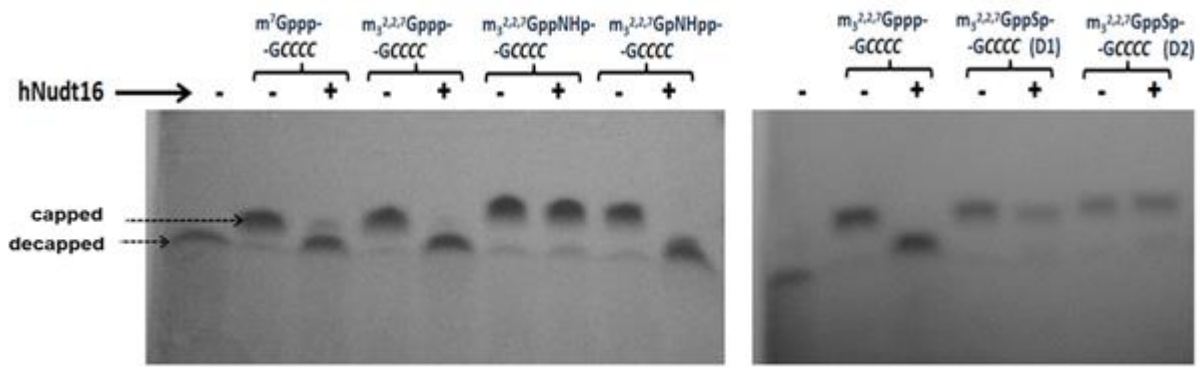


Fig. 4. Susceptibility of short RNA transcripts capped with either TMG cap analogs **1, 6, 7, 9a** and **9b** or monomethylguanosine cap analog m^7GpppG to the Nudt16 decapping enzyme. Electrophoretic mobility of a control 5 nt capped transcript (m^7GpppGCCCC) and uncapped transcript (pppGCCCC) are indicated with arrows. $m_3^{2,2,7}$GpppG- (**1**), $m_3^{2,2,7}$GpNHppG- (**7**), $m_3^{2,2,7}$GppNHpG- (**6**) and $m_3^{2,2,7}$Gpp$_S$pG- (D1 and D2, **9a** and **9b**) capped transcripts migrate virtually to the same level as m^7GpppGCCCC under the conditions used. Active hNudt16 enzyme (lanes indicated with [+]), but not thermally inactivated hNudt16 (lanes indicated with [-]), shows decapping activity on transcripts capped with m^7GpppG, **1** and **7** as the band corresponding to the decapped transcript appears as a sole product after 60 min of incubation. Transcripts capped with **6**, **9a** and **9b** are significantly less susceptible to hNudt16 decapping activity under the same conditions.

Journal Name

RSCPublishing

ARTICLE

It has been previously shown that hNudt16 cleaves both MMG and TMG capped RNAs between the α and β phosphates to produce m^{7}GDP or m$_3^{7,2,2}$GDP and 5'-phosphorylated RNA.[12] Our previous studies on Dcp2, which also cleaves at the α,β-position, have revealed that the α,β-bridging substitutions and β-non-bridging substitutions may produce mRNAs resistant to decapping.[15] To verify whether analogous modifications may confer TMG capped RNAs with resistance to hNudt16, we tested RNAs capped with the phosphate-unmodified analog, m$_3^{2,2,7}$GpppG (**1**), and four phosphate-modified analogs, the α,β-imidodiphosphate containing analog (**6**), the β,γ-imidodiphosphate containing analog (**7**) and the two β-phosphorothioate containing analogs (**9a**; diastereoisomer D1 and **9b**; diastereoisomer D2). The RNA capped with an MMG cap analog (m^{7}GpppG) was also used as a reference. The selected cap analogs were incorporated into short (5 nt) RNAs by *in vitro* transcription and the resulting 5' capped transcripts were subjected to degradation by hNudt16 and the reaction products were analyzed on 20% denaturating polyacrylamide gel (Fig. 4). Under these conditions, capped transcripts (5'-MMGpppRNA or 5'-TMGpppRNA) migrate more slowly than decapped transcripts (5'-pRNA). The position at which the decapped transcripts should appear is the same as the position of a control uncapped transcript (*i.e.* GTP-initiated, 5'-pppRNA). As shown in Figure 4, the transcripts capped with the mono-methylated cap m^{7}GpppG- and tri-methylated caps m$_3^{2,2,7}$GpppG- or m$_3^{2,2,7}$GpNHppG- are specifically decapped by hNudt16, since the band corresponding to the decapped RNA is observed as a sole product after incubation with the enzyme. The transcripts capped with the m$_3^{2,2,7}$GppNHpG and m$_3^{2,2,7}$Gpp$_S$pG D2 cap analogs are resistant to hNudt16 decapping activity under the reaction conditions, since the electrophoretic band corresponding to the capped transcript remains intact in the presence of active hNudt16 (Fig. 4). For m$_3^{2,2,7}$Gpp$_S$pG D1, the degradation was observed; however, the band corresponding to the reaction substrate was still present after 1 h; therefore, the cap was assigned as being partially resistant.

These results indicate that hNudt16 is active on short transcripts possessing both MMG and TMG cap structures. The O to NH substitution at α,β-position and O to S at the β-position of the TMG cap result in resistance to hNudt16, which suggests a similar specificity of the enzyme to that observed for Dcp2.[14, 16, 38] The α,β-imidodiphosphate containing analog (**6**) is a particularly interesting candidate for further studies due to its very low susceptibility to both of the studied types of degradation. However, other phosphate-modifications may also be useful for protection of TMG-capped oligos and TMG capped-bioconjugates against enzymatic decapping. Therefore, more detailed studies are required to assess which type of degradation (specific *versus* non-specific) is more critical for the physiological stability of TMG caps and their conjugates of therapeutic potential.

Affinity to snurportin

It is also crucial to check how the various phosphate chain modifications influence molecular recognition of the TMG cap by snurportin1, to exclude modifications which could impair this interaction and decrease the efficiency of the TMG-mediated nuclear import.

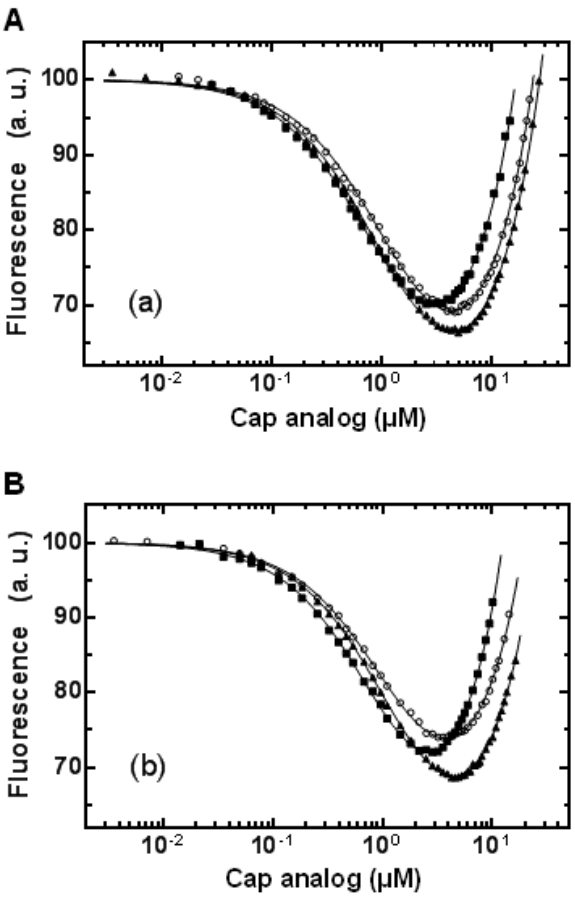


Fig. 5. Example fluorescence quenching curves from time-synchronized titration of snurportin with TMG cap analogs. Binding isotherms were determined from intrinsic protein fluorescence quenching upon titration of wild-type snurportin1 (A) and N-truncated snurportin1 (66-360) (B) with m$_3^{2,2,7}$GpppG (**1**) (○),

$m_3^{2,2,7}$GppNHpG (**6**) (■), $m_3^{2,2,7}$GppspG (**9a**) (▲) at 20 °C, in HEPES/NaOH 50 mM pH 7.2, 150 mM NaCl, 0.5 mM EDTA, 2 mM DTT. The increase in the fluorescence signal towards the end of the titrations comes from the emission of free ligand molecules present in the solution. The K_{as} values for complexes of snurportin with different cap analogs were determined by fitting the experimental data to a theoretical model, as described in Experimental Section, and the results are collated in Table 5.

To preliminarily assess the effect of phosphate-modifications on interaction with snurportin1, selected analogs were tested for their binding affinity to the human protein. The equilibrium binding constants for the cap analog complexes with snurportin were determined by a titration assay based on the intrinsic protein fluorescence quenching upon the complex formation. The affinities of natural TMG 5' cap analogs ($m_3^{2,2,7}$GpppG, $m_3^{2,2,7}$GpppA, m^7GpppA and $m_3^{2,2,7}$GpppAmpUmpA) to N-truncated snurportin1 have been previously shown to fall in the micromolar range.[39] In this work, the affinity was determined for two forms of snurportin1: the same as previously used for the crystal structure determination, N-truncated form (residues 66-360) and the full-length form (residues 1-360). Two newly synthesized phosphate-modified analogs, the α,β-imidodiphosphate containing analog (**6**), and the β-phosphorothioate containing analog (**9a**; diastereoisomer D1) were selected for this study, as well as the unmodified TMG cap analog, $m_3^{2,2,7}$GpppG (**1**) as the control. Examples of the binding isotherms leading to determination of the association constants, K_{as}, are shown in Figure 5, and the results are gathered in Table 5.

Table 5. Equilibrium association constants for the binding of the trimethylguanosine snRNA 5' cap analogs to snurportin1 wild-type and its truncated form, at 20 °C, in HEPES/NaOH 50 mM pH 7.2, 150 mM NaCl, 0.5 mM EDTA, 2 mM DTT.

No	Abbreviation	K_{as} (μM^{-1})	$\Delta G°$ (kcal/mol)
	full length snurportin1 (1-360)		
1	$m_3^{2,2,7}$GpppG	0.90 ± 0.06	-7.98 ± 0.01
6	$m_3^{2,2,7}$GppNHpG	1.37 ± 0.14	-8.23 ± 0.01
9a	$m_3^{2,2,7}$GppspG D1	1.16 ± 0.12	-8.13 ± 0.01
	truncated snurportin1 (66-360)		
1	$m_3^{2,2,7}$GpppG	0.91 ± 0.08	-8.00 ± 0.01
6	$m_3^{2,2,7}$GppNHpG	1.20 ± 0.09	-8.16 ± 0.01
9a	$m_3^{2,2,7}$GppspG D1	0.94 ± 0.07	-8.01 ± 0.01

The K_{as} value of 0.91 ± 0.08 μM^{-1} determined herein for $m_3^{2,2,7}$GpppG (**1**) binding to snurportin at 20 °C, in HEPES/NaOH 50 mM pH 7.2, 150 mM NaCl, 0.5 mM EDTA, 2 mM DTT corresponds to quite efficient binding free energy of about -8 kcal/mol, and is very close to the published result obtained at pH 7.5 and at lower salt concentration of 100 mM NaCl (K_d= 1 ± 0.003 μM which corresponds to K_{as} ~ 1 μM^{-1}).[39] The analogous value for the full-length protein is virtually the same, which indicates that the truncated form of snurportin used for crystal structure determination[39] is entirely sufficient for effective TMG cap binding (Table 5). The affinities of analogs **6** and **9** to both forms of snurportin were either equal to or higher than that of analog **1**. In particular, the compound **6** containing the O to NH substitution at the α,β-position in the 5',5'-triphosphate bridge has the affinity by 30-40% higher than the compound **1**, which makes it a promising candidate for future research. There is still an unresolved issue as to how the efficiency of the TMG cap-mediated nuclear import is dependent on the affinity of the TMG cap to snurportin, and thus the TMG cap analogs with improved affinity to snurportin may help addressing the question. It is also interesting to compare the effects of the studied phosphate-modifications on the cap-snurportin interaction with other cap-binding protein targets. The influence of various hydrolytically resistant modifications of MMG cap analogs has been determined for eIF4E.[13] For example, it has been found that the phosphorothioate or boranophosphate groups do not disrupt the interaction with eIF4E, and in some cases (depending on the position within the triphosphate bridge and absolute configuration) even enhance the binding constant up to 5-fold, which corresponds to a more negative binding energy of about -1 kcal/mol.[33] The imidodiphosphate modification also had a rather stabilizing effect on the interaction of MMG cap analogs with eIF4E, whereas the methylenebisphosphonate was rather destabilizing.[40]

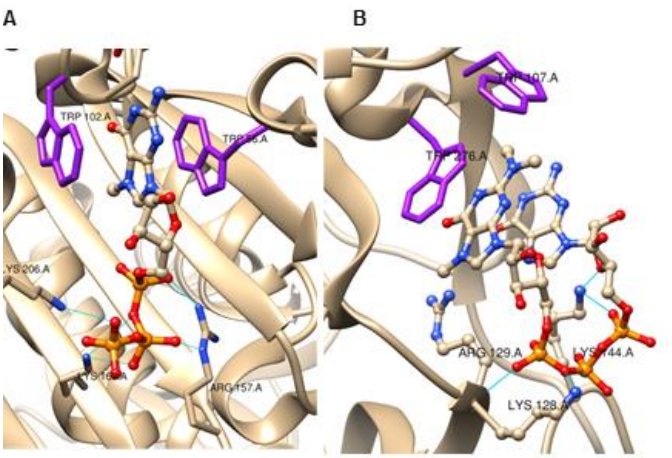


Fig. 6. Hydrophilic and hydrophobic interactions crucial for the recognition of MMG cap and TMG cap structures by specific proteins: a comparison between (A) m^7Gppp-eIF4E and (B) $m_3^{2,2,7}$GpppG-snurportin1 complexes (pdb entries: 1L8B[22,41] and 1XK5[39], respectively). Base recognition: eIF4E binds the monomethylguanine moiety by stacking between two aromatic tryptophan residues (Trp102 and Trp56) to form a sandwich complex; a similar binding mode is used by many other evolutionarily unrelated proteins (VP39,[42] CBP20,[43] PB2,[44] DcpS,[45] TGS1[43]). Snurportin1 utilizes a unique mode to form the sandwich complex by stacking a single tryptophan (Trp276) moiety with a self-stacked dinucleotide. A second tryptophan (Trp107) moiety is involved in the specific recognition of the N,N-dimethylamino group. Interaction of the triphosphate bridge of m^7GpppG with eIF4E: the triphosphate bridge forms hydrogen bonds and salt bridges with the side chains of Lys162, Lys206 and Arg157, either directly or through the involvement of water molecules. Interaction of the triphosphate bridge of $m_3^{2,2,7}$GpppG cap dinucleotide with snurportin 1: the triphosphate bridge forms hydrogen bonds and salt bridges with the side chains of Lys128 and Lys144 as well as with the main chain of Arg129.

Among cap-binding proteins, snurportin1 is unique both due to the self-stacked structure of the UsnRNA 5' cap necessary for

the productive complex formation (Fig. 6) and due to the requirement of the trimethylated cap structure as a part of the nuclear localization signal[4] to be bound (the affinity of snurportin to the monomethylguanosine cap is two orders of magnitude weaker). However, the fact that the studied phosphorothioate and imidodiphosphate TMG cap analogs have equal or slightly better binding properties to snurportin than reference TMG cap analog **1** (Table 5) suggests that the general requirements for the triphosphate chain binding of the two proteins are similar. This appears to be in agreement with the known co-crystal structures of eIF4E and snurportin1 bound, respectively, to m^7GpppG and $m_3^{2,2,7}$GpppG. Despite significant differences in the recognition of the methylated nucleoside moiety, the binding mode of the triphosphate bridge is in both cases based mainly on salt bridges and hydrogen bonds between the negatively charged triphosphate bridge and the basic amino acids and water molecules present in the cap binding pockets (Fig. 6). [22, 39, 41]

Conclusions

We described the synthesis and properties of a series of novel TMG cap ($m_3^{2,2,7}$GpppG) analogs containing various modifications in the 5',5'-triphosphate bridge as tools to study TMG cap degradation, and binding by specific proteins, including snurportin1. Preliminary studies revealed that some of the phosphate modifications increase the stability of the TMG cap in blood serum, confer TMG-capped RNAs resistance to enzymatic cleavage by hNudt16 decapping enzyme and stabilize the interaction of TMG cap with snurpotin1. TMG cap analogs with such properties may have potential clinical applications as nucleus localization signals, when attached to nucleus-targeting drugs.[36] Cap analog **6**, containing an imidodiphosphate group at the α,β-position ($m_3^{2,2,7}$GppNHpG), was characterized by particularly favorable properties and therefore was selected as a promising candidate for future applications. Further studies on the synthesized compounds directed towards a more thorough characterization of their biophysical and biological properties, as well as their structure-activity relationship for TMG as a nuclear localization signal are being currently performed.

Experimental

Materials and Methods

Guanosine was purchase from US Biological. Methylenebis(phosphonic dichloride) was purchased from Sigma Aldrich, guanosine diphosphate sodium salt and adenosine diphosphate sodium salt were purchased from Sigma-Aldrich and were converted into triethylammonium salts (**17**-**18**) using Dowex 50 × 8 (200-400 mesh) ion-exchange resin. N^2,N^2-dimethylguanosine (DMG) **11** was prepared as described previously.[23,24] N^2,N^2-dimethylguanosine monophosphate ($m_2^{2,2}$GMP) **12,** N^2,N^2,N^7-trimethylguanosine monophosphate ($m_3^{2,2,7}$GMP) **13** and N^2,N^2,N^7-trimethylguanosine 5'-monophosphate P-imidazolide ($m_3^{2,2,7}$GMP-Im) **14** were prepared as described by Honcharenko *et al.*[14]. Thiophosphate and boranophosphate triethylammonium salts were obtained as described earlier.[18, 23] Dichlorophosphoroimidoyltrichloride (PCl_3NPOCl_2) was prepared as descried earlier.[20] Guanosine phosphorimidazolide sodium salt **26** was prepared as described previously.[19] Guanosine 5'-methylene(bisphosphonate) (GpCH$_2$p) **21**, guanosine 5'-imidodiphosphate (GpNHp) **24** and guanosine 5'-thiophosphate **27** (Gp$_S$) were synthesized as described previously. [20,34,40]

Ion-exchange chromatography was performed on a DEAE-Sephadex A-25 column. Buffers for elution were as follows: triethylammonium bicarbonate (TEAB) in deionized water with a linear gradient of 0-0.7 M TEAB for nucleoside monophosphates, a linear gradient of 0-0.9 M TEAB for nucleoside diphosphate and a linear gradient of 0-1.2 M TEAB for dinucleoside triphosphates. Collected fractions were evaporated under reduced pressure with the addition of ethanol and acetonitrile. Compounds were isolated as TEA salts.

Optical unit measurements were performed in 0.1 M phosphate buffer, pH = 7 or pH = 6 (for $m_3^{2,2,7}$G nucleotides), and at 260 nm. Analytical HPLC was performed on an Agilent Tech. Series 1200 using a Supelcosil LC-18-T reverse phase column (4.6 × 250 mm, flow rate 1.3 mL/min, UV-detection at 260 nm, fluorescence detection - excitation at 260 nm and detection at 370 nm). Buffers for analytical reversed-phase chromatography were as follows: A: 50 mM ammonium acetate (CH_3COONH_4), pH 5.9; B: 50 mM CH_3COONH_4, pH 5.9 in 50% MeOH (linear gradient from 0 to 50% of buffer B in 15 min for compounds **4**-**9**, and linear gradient from 0 to 100% B in 15 min for compounds **2** and **3**).

Semi-Preparative HPLC was performed on the same apparatus equipped with a Discovery RP Amide C-16 reverse phase column (25 cm × 21.2 mm, 5 μm, flow rate 5.0 mL/min, UV-detection at 260 nm). Buffers for semi-preparative HPLC were as follows: A: 50 mM ammonium acetate (CH_3COONH_4), pH 5.9; B: 50 mM CH_3COONH_4, pH 5.9 in 30% MeCN (from 0 to 50% of buffer B in 90 min).

HPLC for FBS assay was performed on an Agilent Tech 1260 series with a thermostatic autosampler and a column compartment and equipped with a Supelcosil LC-18-T reverse phase column (4.6 × 250 mm, flow rate 1.3 mL/min, UV-detection at 260 nm, fluorescence detection - excitation at 260 nm and detection at 370 nm). Buffers for stability assays were as follows: A: 0.1 M potassium phosphate buffer (KH_2PO_4/K_2HPO_4), pH 7.0; B: 0.1 M potassium phosphate buffer (KH_2PO_4/K_2HPO_4), pH 7.0 in 50% MeOH (linear gradient from 0 to 50% of buffer B in 15 min).

High-resolution mass spectra were recorded on a Thermo Scientific LTQ Orbitrap Velos. Low-resolution mass spectra were recorded on an AB Sciex API 3200. NMR spectra were recorded at 25° ºC on a Varian UNITY-plus spectrometer at 399.94 MHz or 500 MHz. Chemical shifts were calibrated to sodium 3-trimethylsilyl-[2,2,3,3-D4]-propionate (TSP) in D_2O

as an internal standard. ^{31}P NMR chemical shifts were reported to 20% phosphorus acid in D_2O as an external standard.
Concentrations and molar extinction coefficients of the newly synthesized chemical cap analogs were determined from weighed amounts, after at least four-fold lyophilization. Absorption spectra were measured on a Shimadzu UV-1800 spectrophotometer.

Chemical syntheses

P1-(N^2,N^2,N^7-trimetylguanosine -5'-yl)- P3-guanosin-5'-yl triphosphate $m_3^{2,2,7}$GpppG (1)

Compounds **14** ($m_3^{2,2,7}$GMP-Im, Li salt, 55 mg, 1360 opt.mu., 0.12 mmol) and **17** (GDP, TEA salt, 77 mg, 1450 opt.mu., 0.12 mmol) were mixed in anhydrous DMF (3 mL) followed by the addition of anhydrous $ZnCl_2$ (275 mg, 1.92 mmol). After 24 h, the reaction was completed and quenched by the addition of a solution of EDTA (756 mg, 1.92 mmol) and $NaHCO_3$ (378 mg, 0.96 mmol) in water (30 mL). Product **1** was purified on DEAE-Sephadex, and after evaporation obtained as a glassy solid (TEA salt, 917 opt.mu., 0.10 mmol, 83%). Further purification was performed using preparative RP HPLC yielding 30 mg of **1** (NH_4^+ salt, 0.040 mmol, 40%). ^{1}H NMR δ_H (400 MHz, D_2O, TSP), 3.17 (6 H, s), 4.05 (3 H, s), 4.22-4.34 (3H, m), 4.39-4.44 (4H, m), 4.56 (1H, dd, *J* 6.0, 5.0), 4.57 (1H, dd, *J* 3.2, 5.0), 5.75 (1 H, d, *J* 6.0), 5.92 (1H, d, *J* 3.2), 7.95 (1 H, s), 8.95* (1H, s), *slowly exchanging in D_2O.
^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4): −10.73 (2P, d), 22.3 (1P, t). HRMS (ES^+): calcd for $C_{23}H_{34}N_{10}O_{18}P_3$ $[M+H]^+$ 831.1260, found 831.1274.

P1-(N^2,N^2,N^7-trimethylguanosin-5′-yl) P3-adenosin-5′-yl triphosphate $m_3^{2,2,7}$GpppA (2)

Compound **18** (ADP, TEA salt, 100 mg, 0.17 mmol) was dissolved in dry DMF (1 mL). After 5 min, compound **14** ($m_3^{2,2,7}$GMP-Im, Li salt, 110 mg, 0.26 mmol, 1.5 eq) and $ZnCl_2$ (7.18 mg, 1.8 mmol, 7 eq) were added. After 48 h, the reaction was completed and quenched by the addition of an aqueous solution (15 mL) EDTA (672.9 mg, 1.8 mmol, 7 eq) and $NaHCO_3$ (338.4 mg, 1.8 mmol, 7 eq). The crude product was purified using ion-exchange chromatography (Sephadex, TEAB 0-1.2 M). Further purification on semi-preparative RP HPLC gave 51 mg of compound **2** (36% yield).
^{1}H NMR δ_H (500 MHz, D_2O, TSP), 3.15 (6 H, s), 4.03 (3 H, s), 4.26-4.43 (7H, m), 4.45 (1H, dd, *J* 3.9, 5.2), 4.52 (1H, dd, *J* 3.5, 3.9), 4.55 (1H, dd, *J* 5.2, 5.8), 5.84 (1 H, d, *J* 5.8), 5.90 (1H, d, *J* 3.5), 5.96 (1H, d, *J* 5.84), 8.19 (1H, s)*, 8.36 (1 H, s), 8.92* (1H, s), *slowly exchanging in D_2O. ^{31}P NMR δ_P (202.5 MHz, D_2O, 25 °C, H_3PO_4), −23.08 (1P, t, *J* 19.4), -11.53 (2P, d, *J* 19.4).
HRMS (ES^+): calcd for $C_{23}H_{34}N_{10}O_{17}P_3$ $[M+H]^+$ 815.1322, found 815.1400

P1-(N^2,N^2,N^7-trimetylguanosine -5'-yl)- P4-guanosin-5'-yl tetraphosphate $m_3^{2,2,7}$GppppG (3)

Compounds **16** ($m_3^{2,2,7}$GDP-Im, TEA salt, 97 mg, 1744 opt.mu., 0.15 mmol) and **17** (GDP, TEA salt, 98 mg, 1850 opt.mu., 0.15 mmol) were mixed in anhydrous DMF (3 mL) followed by the addition of anhydrous $ZnCl_2$ (160 mg, 1.78 mmol). After 48 h, the reaction was completed and quenched by the addition of a solution of EDTA (438 mg, 1.78 mmol) and $NaHCO_3$ (219 mg, 0.89 mmol) in water (30 mL). The product was purified on DEAE-Sephadex, obtained as a white solid (TEA salt, 1923 opt.mu., 0.12 mmol, 77%). Further purification was performed using preparative HPLC yielding 56 mg of the final compound (NH_4^+ salt, 0.06 mmol, 41%).
^{1}H NMR δ_H (500 MHz, D_2O, TSP), 3.21 (6 H, s), 4.12 (3 H, s), 4.33-4.29 (3H, m), 4.37-4.40 (1H, m), 4.41-4.46 (2H, m), 4.51 (1H, dd, *J* 5.0, 5.1), 4.54 (1H, dd, *J* 3.3, 4.9), 4.69 (1H, dd, *J* 3.6, 5.1), 4.69 (1H, dd, *J* 3.6, 5.1), 4.70 (1H, dd, *J* 4.9, 6.3), 5.84 (1H, d, *J* 6.3), 6.04 (1H, d, *J* 3,6), 8.06 (1 H, s), 9.14* (1H, s), *slowly exchanging in D_2O. ^{31}P NMR δ_P (202.5 MHz, D_2O, H_3PO_4), −22.98 (2P, m), -11.17 (2P, m). HRMS (ES^+) calcd for $C_{23}H_{35}N_{10}O_{21}P_4$ $[M+H]^+$ 911.0923 found 911.0922.

P1-(N^2,N^2,N^7-trimetylguanosine -5'-yl)- P3-guanosin-5'-yl 2,3 methylenetriphosphate $m_3^{2,2,7}$GppCH$_2$pG/ Na^+ (4)

Compounds **14** ($m_3^{2,2,7}$GMP-Im, Na salt, 33 mg, 830 opt.mu., 0.073 mmol) and **21** ($GpCH_2p$, TEA salt, 70 mg, 1230 opt.mu., 0.11 mmol) were mixed in anhydrous DMF (3 mL) followed by the addition of anhydrous $ZnCl_2$ (160 mg, 1.2 mmol). After 5 h, the reaction was completed and quenched by the addition of a solution of EDTA (436 mg, 1.2 mmol) and $NaHCO_3$ (220 mg, 0.60 mmol) in water (40 mL). Product **5** was isolated as TEA salt by DEAE-Sephadex chromatography and then was converted into Na salt using a solution of sodium perchlorate $NaClO_4$ (84 mg, 0.37 mmol) in acetone (15 mL) (sodium salt, 39 mg, 1060 opt. mu., 0.047 mmol 64% yield). ^{1}H NMR δ_H (500 MHz, D_2O, TSP), 3.22 (6H, s), 4.12 (3H, s), 4.20-4.32 (4H, m), 4.40-4.45 (2H, m), 4.45-4.51 (2H, m), 4.64 (1H, dd, *J* 3.3, 3.4), 4.66 (1H, dd, *J* 4.3, 5.8), 5.82 (1 H, d, *J* 5.8), 6.02 (1H, d, *J* 3.3), 8.04 (1H, s), 9.14* (1H, s), *slowly exchanging in D_2O. ^{31}P NMR δ_P (202.5 MHz, D_2O, H_3PO_4), -11.02 (1P, d, *J* 25.7), 7.67 (1P, m), 17.21 (1P, m). HRMS (ES^+) calcd m/z for $C_{24}H_{36}N_{10}O_{17}P_3$ $[M+H]^+$ 829.1467 found 829.1538.

P1-(N^2,N^2,N^7-trimetylguanosine -5'-yl)- P3-guanosin-5'-yl 1,2- methylenetriphosphate $m_3^{2,2,7}$GpCH$_2$ppG (5)

Compounds **26** (GMP-Im, Na salt, 32 mg, 737 opt.mu. 0.060 mmol) and **22** ($m_3^{2,2,7}GpCH_2p$, TEA salt, 28 mg, 614 opt.mu., 0.054 mmol) were mixed in anhydrous DMF (2 mL) followed by the addition of anhydrous $ZnCl_2$ (132 mg, 0.98 mmol). After 5 h, the reaction was completed and quenched by the addition of an aqueous solution (40 mL) containing EDTA (363 mg, 0.98 mmol) and $NaHCO_3$ (182 mg, 0.7 mmol). Product **4** was isolated by DEAE-Sephadex chromatography as a white solid. To convert the TEA salt into sodium salt, the product was dissolved in 1 mL of water and then precipitated using an acetone (15 mL)/$NaClO_4$ (56 mg, 0.46 mmol) solution (sodium salt, 32 mg, 680 opt. mu., 0.031 mmol 42% yield). ^{1}H NMR δ_H

(400 MHz, D_2O, TSP), 2.43 (2H, dd, *J* 3.9, 20.6), 3.22 (6H, s), 4.11 (3H, s), 4.19-4.32 (3H, m), 4.34-4.37 1H, m), 4.39-4.44 (2H,m), 4.46 (1H, dd, *J* 3.3, 6.9), 4.49 (1H, dd, *J* 3.8, 5.8), 4.64 (1H, dd, *J* 3.3, 4.6), 4.66 (1H, dd, *J* 5.8, 5.6), 5.80 (1H, d, *J* 5.8 Hz), 6.01 (1H, d, *J* 3.3), 8.03 (1H, s), 9.15* (1H, s), *slowly exchanging in D_2O. ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4), -11.2 (1P, d, *J* 26.4), 7.71 (1P, m), 17.21 (P_2, m). HRMS (ES^+) calcd for $C_{24}H_{36}N_{10}O_{17}P_3$ $[M+H]^+$ 829.1467 found 829.1538.

P1-(N^2,N^2,N^7-trimethylguanosin-5′-yl) P3-guanosin-5′-yl 2,3-imidotriphosphate $m_3^{2,2,7}$GppNHpG (6)

Compounds **14** ($m_3^{2,2,7}$GMP-Im, Li salt, 34 mg, 764 opt.mu., 0.067 mmol) and **24** (GpNHp, TEA salt, 40 mg, 809 opt.mu., 0.067 mmol) were mixed in anhydrous DMF (1 mL) followed by the addition of anhydrous $ZnCl_2$ (63 mg, 0.51 mmol). After 24 h, the reaction was completed and quenched by the addition of a solution of EDTA (115.3 mg, 0.67 mmol) in water and adjusted to pH 6 with solid $NaHCO_3$. Product **6** was purified on DEAE-Sephadex, obtained as a glassy solid (TEA salt, 405 opt.mu., 0.036 mmol, 53%). Further purification was performed using preparative HPLC yielding 21.0 mg of the final compound (NH_4^+ salt, 0.024 mmol, 36%). ^1H NMR δ_H (400 MHz, D_2O, TSP) 3.19 (6 H, s), 4.07 (3 H, s). 4.17-4.31 (4H, m), 4.35-4.46 (4H, m), 4.59 -4.65- (2H, m), 5.79 (1H, d, *J* 6,1), 5.96 (1H, d, *J* 2,8), 8.08 (1 H, s), 9.05* (1H, s); ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4), −10.32 (2P, m), 0.21 (1P, d, *J* 5.6), *slowly exchanging in D_2O. HRMS (ES^+) calcd m/z for $C_{23}H_{35}N_{11}O_{17}P_3$ $[M+H]^+$ 830.1420, found 830.1414.

P1-(N^2,N^2,N^7-trimethylguanosin-5′-yl) P3-guanosin-5′-yl 1,2-imidotri-phosphate $m_3^{2,2,7}$GpNHppG (7)

Compounds **26** (GMP-Im, Na salt, 27 mg, 750 opt.mu., 0.062 mmol) and **25** ($m_3^{2,2,7}$GpNHp, TEA salt, 32 mg, 533 opt.mu., 0.047 mmol) were mixed in anhydrous DMF (1 mL) followed by the addition of anhydrous $ZnCl_2$ (68 mg, 0.50 mmol). After 24 h, the reaction was completed and quenched by the addition of a solution of EDTA (150 mg, 0.40 mmol) in water and adjusted to pH 6 with solid $NaHCO_3$. The product was purified on DEAE-Sephadex, obtained as a glassy solid (TEA salt, 471 opt.mu., 0.041 mmol, 88%). Further purification was performed using preparative RP HPLC yielding 21 mg of compound **7** (NH_4^+ salt, 0.024 mmol, 51%). ^1H NMR δ_H (500MHz, D_2O, TSP) 3.21 (6H,s), 4.12 (3H, s), 4.19-4.30 (2H, m), 4.31-4.37 (2H, m), 4.37-4.40- (1H, m), 4.42-4.44 (1H, m), 4.50 (1H, dd, *J* 4.5, 5.3), 4.51 (1H, dd, *J* 4.8, 5.7), 4.63 (1H, dd, *J* 5.2, 5.3), 4.67 (1H, dd, , *J* 3.6, 4.8), 5.86 (1H, d, *J* 5.2), 6.02 (1H, d, *J* 3.6), 8.29 (1H, s), 9.20 (1H, s)*, *slowly exchanging in D_2O. ^{31}P NMR δ_P (202.5 MHz, D_2O, H_3PO_4) -11.28 (1P, m). -10.82 (1P, m), 0.55 (1P, m), HRMS (ES^+) calcd m/z for $C_{23}H_{35}N_{11}O_{17}P_3$ $[M+H]^+$ 830.1420, found 830.1407.

P1-(N^2,N^2,N^7-trimetylguanosine -5'-yl)- P3-guanosin-5'-yl 3-thiotriphosphate $m_3^{2,2,7}$Gppp$_S$G D1 and D2 (8a and 8b)

Compound **16** ($m_3^{2,2,7}$GDP-Im, Li salt, 25 mg, 513 opt.mu., 0.045 mmol) and compound **27** (GpS^{19} , TEA salt, 20.2 mg, 423 opt. mu., 0.045 mmol) were mixed in anhydrous DMF (2 mL) followed by the addition of anhydrous $ZnCl_2$ (61 mg, 0.45 mmol). After 48 h, the reaction was completed and quenched by the addition of a solution of EDTA (161 mg, 0.45 mmol) and $NaHCO_3$ (80 mg, 0.23 mmol) in water (30 mL). The product was isolated by DEAE-Sephadex chromatography as a mixture of two diastereoisomers of compound 8 (0.040 mmol, 840 opt. mu.) which were then separated using RP HPLC. The obtained fractions were repeatedly lyophilized to give 12.5 mg of diastereoisomer 1 (**8a**) and 20 mg of diastereoisomer 2 (**8b**) (NH_4^+ salt, 932 opt. mu., 0.04 mmol, 37%).

D1: ^1H NMR δ_H (400MHz, D_2O, TSP), 3.19 (6H, s), 4.08 (3H,s), 4.25-4.36 (3H, m), 4.36-4.46 (3H, m), 4.46-4.50 (2H,m), 4.62 (2H,m), 5.80 (1H, d, *J* 5.9), 5.96 (1H, d, *J* 3.2), 8.12 (1H, s), 9.02 (1H, s)*, *slowly exchanging in D_2O. ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4), -24.22 (1P, t, *J* 21.9,), -11.61 (1P, d, *J* 16.1), 43.38 (1P, m, *J* 25.3).

D2: ^1H NMR δ_H (400MHz, D_2O, TSP), 3.19 (6H, s), 4.07 (3H, s), 4.24-4.32 (2H, m), 4.34-4.42 (3H, m), 4.43-4.52 (3H, m), 4.60 (1H, dd, *J* 3,1, 6.3), 4.60 (1H, dd, *J* 4.5, 5.9, 4.5), 5.78 (1H, d, *J* 5.8), 5.94 (1,H. d, *J* 3.1), 8.07 (1H,s), 8.99 (1H,s)*, *slowly exchanging in D2O. ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4) -24.14 (1P,m). -11.60 (1P,m), 43.67 (1P,m), HRMS (ES^-) calcd m/z for $C_{23}H_{34}N_{10}O_{17}P_3S$ $[M-H]^-$ 847.1031, found 847.1024 (D1) and 847.1033 (D2).

P1-(N^2,N^2,N^7-trimetylguanosine -5'-yl)- P3-guanosin-5'-yl 2-thiotriphosphate $m_3^{2,2,7}$GppSpG D1 and D2 (9a and 9b)

Compound **28** ($m_3^{2,2,7}$GDPβS, TEA salt, 140 mg, 2 330 opt. mu., 0.21 mmol) and compound **26** (GMP-Im, Na salt, 128 mg, 960 opt.mu., 0.24 mmol, 1.24 eq) were mixed in anhydrous DMF (3 mL) followed by the addition of anhydrous $ZnCl_2$ (290 mg, 2.13 mmol). After 24 h, the reaction was completed and quenched by the addition of a solution of EDTA (790 mg, 2.13 mmol) and $NaHCO_3$ (396 mg, 1.06 mmol) in water (25 mL). The product was isolated by DEAE-Sephadex chromatography as a mixture of two diastereoisomers of compound **9** which were then separated using RP HPLC. The obtained fractions were repeatedly lyophilized to give 20.3 mg of diastereoisomer 1 (**9a**) and 16.1 mg of diastereoisomer 2 (**9b**) (ammonium salt, 861 opt. mu., 0.041 mmol, 37%**).**

D1: ^1H NMR δ_H (400MHz, D_2O, TSP), 3.17 (6H,s), 4.05 (3H,s), 4.24-4.36 (4H, m), 4.36-4.40- (1H,m), 4.42-4.44 (1H,m), 4.46 (1H, dd, *J* 5.0, 5.3), 4.49 (1H, dd, *J* 3.7, 5.1), 4.60(1H, dd, *J* ,3.4, 5.3), 4.62 (1H, dd, *J* 5.1, 6.1), 5.78 (1H, d, *J* 6.1), 5.96 (1H, d, *J* 3.4), 7.98 (1H, s), 8.98 (1H, s)*, *slowly exchanging in D_2O. ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4) -11.56 (2P, m). 30.87 (1P, m),

D2: ^1H NMR δ_H (400MHz, D_2O, TSP) 8.97(1H,s)*, 7.97 (1H, s), 5.94 (1H, d, *J* 3.4), 5.76 (1H, d, *J* 6.1), 4.61 (1H, dd, *J* 3.5, 6.1), 4.59 (1H, dd, *J* 3.4, 4.4), 4.44-4.47 (2H, m), 4.38-4.43 (2H, m), 4.21-4.37 (4H, m), 4.07 (3H, s), 3.17 (6H, s) *slowly exchanging in D_2O. ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4) 11.60 (2P, m), 30.74 (1P, t, *J* 26.0)-.HRMS (ES^-) calcd m/z for $C_{23}H_{34}N_{10}O_{17}P_3S$ $[M-H]^-$ 847.1031, found 847.1081 (D1) and 847.1058 (D2).

P1-(N^2,N^2,N^7-trimetylguanosine -5'-yl)- P3-guanosin-5'-yl 2-boranetriphosphate $m_3^{2,2,7}$GppBH$_3$pG D1 and D2 (10a and 10b)

Compound **14** ($m_3^{2,2,7}$GMP-Im, Na salt, 48 mg, 1250 opt.mu., 0.11 mmol) and boranophosphate (90 mg, 0.33 mmol) were mixed in anhydrous DMF (2 mL) followed by the addition of anhydrous $MgCl_2$ (74 mg, 0.77 mmol). After 10 minutes, to the reaction mixture was added a suspension of compound **26** (GMP-Im Na salt, 151 mg, 3990 opt. mu., 0.33 mmol) and $MgCl_2$ (74 mg, 0.77 mmol) in anhydrous DMF (1.5 mL) (one pot reaction). After 8h, the reaction was completed and quenched by the addition of 50 mL of water. The product was isolated by DEAE-Sephadex chromatography as a mixture of two diastereoisomers of compound **10** which were then separated using preparative RP HPLC. The obtained fractions were repeatedly lyophilized to give 17 mg of diastereoisomer 1 D1 and 22 mg of diastereoisomer 2 D2 (ammonium salt, 982 opt. mu., 0.045 mmol, 43%).

D1: ^{1}H NMR δ_H (400 MHz, D_2O, TSP), 3.14 (6 H, s), 4.05 (3 H, s), 4.20-4.34 (1 H, m), 4.20-4.38 (1 H, m), 4.38 (2H, m, *J*3.7, 6.5,), 4.42 (1H, dd, *J*3.2, 6.5), 4.45 (1H, dd, *J*3.7, 5.0), 4.57 (1H, dd, *J* 5.0, 6.0), 4.59 (1H, dd, *J* 3,2, 5.0), 5.76 (1H, d, *J* 6.0), 5.96 (1H, d, *J* 3,2) ,7.98 (1 H, s), ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4), −10.82 (2P, d, *J*26.9), 75.89 (1P, m).

D2: ^{1}H NMR δ_H (400 MHz, D_2O, TSP), 3.14 (6 H, s), 4.06 (3 H, s), 4.19-4.44 (6H, m), 4.56 (1H, dd, *J* 2.7, 3.7), 4.59 (1H, dd, *J*5.5, 6.0), 5.75 (1H, d, *J* 6,0), 5.94 (1H, d, *J* 2,7), 7.96 (1 H, s), 8.87 (1 H, s).^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4), −10.85 (2P, d, *J* 29.8), 75.75 (1P, m).

HRMS (ES^-) calcd for $C_{23}H_{37}BN_{10}O_{17}P_3$ $[M-H]^-$ 829.1633, found 829.1621 (D1) and 829.1617 (D2).

***N*^{2},*N*^{2},*N*^{7}-trimethyl-guanosine5'-diphosphate imidazolide $m_3^{2,2,7}$GDP-Im (16)**

N^2,N^2,N^7-trimethylguanosine diphosphate **15** (TEA salt, 52 mg, 0.07 mmol, 800 opt.mu.), imidazole (49 mg, 0.7 mmol) and 2,2'-dithiopyridine (47 mg, 0.21 mmol) were suspended in anhydrous DMF (2 mL). Subsequently, triethylamine (30 µL, 0.21 mmol) and triphenylphosphine (57 mg, 0.21 mmol) were added and the mixture was stirred at RT for 7 h. The resulting clear solution was poured into a flask containing 41 mg of anhydrous sodium perchlorate in 20 mL acetone. After cooling at 4^{o}C for 1 h, the resulting precipitate was filtered, washed with dry acetone and dried overnight in a vacuum desiccator over P_4O_{10} to give 32 mg (0.05 mmol) of compound 1**7** sodium salt yield 77%. ^{1}H NMR δ_H (400 MHz, D_2O, TSP), 3.12 (6 H, s), 4.04 (3 H, s), 4.09 (1 H, m), 4.24 (1H, m), 4.34 (1H, m, *J* 5.2), 4.37 (1H, dd, *J* 4,7, 5,2), 4.65 (1H, dd, *J*3,7, 4,7), 6.10 (1H, d, *J*3,7), 6.99 (1H, d), 7.31 (1H, d), 7.93 (1 H, s), *slowly exchanging in D_2O ^{31}P NMR δ_P (162 MHz, D_2O, H_3PO_4), −19.06 (1P, d, *J* 21.5) -11.01 (1P, d, *J* 21.4), HRMS (ES^-) calcd m/z for $C_{16}H_{24}N_7O_{10}P_2$ $[M-H]^-$ 536.1065, found 536.1051.

N2,N2-dimethylguanosine 5'-methylene(bisphos-phonate) m22,2GpCH2p (20)

A solution of methylenebis(phosphonic dichloride) (883 mg, 3.5 mmol) in trimethyl phosphate (5 mL), cooled to 0^{o}C, was added to a suspension of N^2,N^2-dimethylguanosine **11** (500 mg, 1.6 mmol) in trimethyl phosphate at 0^{o}C. The reaction mixture was stirred at 0^{o}C and samples were analyzed by HPLC after 15 min intervals. After 1.5h, on the disappearance of N^2,N^2-dimethylguanosine, 10% aqueous $NaHCO_3$ solution was added to pH 7. The product **20** was isolated by DEAE-Sephadex chromatography as a glassy solid (TEA salt, 790 mg, 1.4 mmol, 16 800 opt.mu., 87%). ^{1}H NMR δ_H (400 MHz, D_2O, 25 °C, TSP), 3.20 (6 H, s), 4.12 (3 H, s), 4.12-4.20 (1H, m), 4.20-4.28 (1H, m), 4.34-4.38 (1H, m), 4.52 (1H, dd, *J* 3,8, 5.9), 4.76-4.81 (1H, m), 6.08 (1H, d, *J* 3,8), 8.78* (1 H,s), *slowly exchanging in D_2O. HRMS (ES+) calcd m/z for $C_{13}H_{22}N_5O_{10}P_2$ $[M+H]^+$ 470.0836 found 470.0869.

***N*^{2},*N*^{2},*N*^{7}-trimethylguanosine 5'-methylene(bis-phosphonate) $m_3^{2,2,7}$GpCH$_2$p (22)**

To compound **20** (DMGpCH$_2$p, 327 mg, 0.53 mmol) dissolved in DMSO (3.7 mL), CH_3I was added (105 µL, 1.68 mmol) and the mixture was stirred for 2h at RT. Then, 50 mL of cold water was added and the solution was extracted 4 times with 20 mL portions of diethyl ether. The aqueous phase was neutralized with 1M $NaHCO_3$ and purified by ion-exchange chromatography (DEAE-Sephadex resin). Product **22** was obtained as TEA salt (234 mg, 0.37 mmol, 4290 opt. mu., 71% yield). ^{1}H NMR δ_H (400 MHz, D_2O, 25 °C, TSP), 3.18 (6 H, s), 4.17 (1H, m), 4.29 (1H, m), 4.40 (1H, m), 4.51 (1H, dd, *J* 4.0, 5.0), 4.75 (1H, dd, *J* 3.2, 4.0), 6.14 (1H, d, *J* 3.2), 9.32 (1 H, s), HRMS (ES-) calcd m/z for $C_{14}H_{22}N_5O_{10}P_2$ $[M-H]^-$ 482.0847 found 481.9975.

P1-(N^2, N^2-dimethyl-guanosine-5'-yl)imido-diphosphate DMGpNHp (23)

N^2,N^2-dimethylguanosine **11** (150 mg, 0.48 mmol) was suspended in trimethyl phosphate (2 mL) and cooled to -12^{o}C. Then, $Cl_3PNP(O)Cl_2$ was added under vigorous stirring. The reaction was maintained at -12^{o}C and after 2h was stopped by the addition 10% aqueous $NaHCO_3$ solution to pH 7.0. Product **23** was purified on DEAE-Sephadex and isolated as TEA salt (197 mg, 4741 opt. mu., 0.39 mmol, 81 % yield). ^{1}H NMR δ_H (400 MHz, D_2O, 25 °C, TSP), 3.18 (6 H, s), 4.12-4.16 (2H, m), 4.32-4.36 (1H, m), 4.57 (1H, dd, *J* 4.2, 5.0), 4.76-4.82 (1H, m), 6.04 (1H, d, *J* 5.7), 8.16 (1 H, s). ^{31}P NMR δ_P (162 MHz, D_2O), -1.39 (1P, m), 0.11 (1P, d, *J* 29.8). HRMS (ES-) calcd m/z for $C_{12}H_{19}N_6O_{10}P_2$ $[M-H]^-$ 469.0643, found 469.0607.

P1-(N^2,N^2-dimethylguanosine-5'-yl)imido-diphosphate $m_3^{2,2,7}$GpNHp (25)

To compound **23** (DMGpNHp, 150 mg, 0.25 mmol) dissolved in DMSO (3.7 ml), CH_3I was added (196 µL, 3.14 mmol) and the mixture was stirred for 12h at RT. Then, 50 mL of cold water was added and the solution was extracted 4 times with 20 mL portions of diethyl ether. The aqueous phase was

neutralized with 1M $NaHCO_3$ and purified by ion-exchange chromatography (DEAE-Sephadex resin). Product **25** was obtained as TEA salt (25 mg, 0.042 mmol, 480 opt. mu., 25% yield). ^{1}H NMR δ_H (400 MHz, D_2O, 25 °C, TSP), 3.19 (6 H, s), 4.12 (3 H, s), 4.16 (1H, m), 4.27 (1H, m), 4.41 (1H, m), 4.54 (1H, m), 4.73 (1H, m), 6.14 (1H, d, *J* 3.6). ^{31}P NMR δ_P (162 MHz, D_2O) -1.24 (1P, d, *J* 4.6), 0.54 (1P, s). HRMS (ES-) calcd m/z for $C_{13}H_{21}N_6O_{10}P_2$ $[M-H]^-$ 483.0799, found 483.0794.

N^2,N^2,N^7-trimethylguanosine 5'-(2-thiodiphosphate), $m_3^{2,2,7}$GDPβS (28)

N^2,N^2,N^7-trimethylguanosine 5'-monophosphate imidazolide sodium salt **14** (100 mg, 0.21 mmol) was mixed with thiophosphate triethylammonium salt (100 mg, 0.5 mmol). The resultant mixture was suspended in 2.5 mL of DMF. Subsequently, anhydrous $ZnCl_2$ (280 mg, 2.1 mmol) was added and the rapid dissolution of the reagents was observed. After 45 min, the reaction was quenched by the addition of a solution of EDTA (765 mg, 2.1 mmol) and $NaHCO_3$ (383, 1.1 mmol) in water (25 mL). Chromatographic isolation was performed on a DEAE-Sephadex column with a linear gradient of triethylammonium bicarbonate (from 0 to 1 M TEAB in deionized water). The collected fractions containing product **28** were evaporated with the addition of ethanol to decompose the TEAB buffer and dried over P_4O_{10} overnight (144 mg, 2344 opt. mu., 90% yield). ^{1}H NMR δ_H (400 MHz, D_2O, 25 °C, TSP), 3.12 (6 H, s), 4.10 (3 H, s), 4.22-4.43 (3H, m), 4.59 (1H, dd, *J* 4.8, 5.7), 4.73 (1H, dd, *J* 3.4, 4.8), 6.14 (1H, d, *J* 3.4), 9.08 (1 H, s), 9.20 (1H, s) ^{31}P NMR δ_P (162 MHz, D_2O) -11.30 (1P, m), 40.46 (1P, d, *J* 29.8). HRMS (ES^-) calcd m/z for $C_{13}H_{20}N_5O_{10}P_2S$ $[M-H]^-$ 500.0412, found 500.0797.

FBS stability assay

For the serum stability assay, 100 μM stock solutions of compounds (**1**–**10**) in MilliQ water were prepared. Then, 100 μL of respective stock solution, 890 μL of PBS buffer (pH 7.2, 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) and 10 μL of FBS were mixed in 1.5 mL HPLC vials, so each solution contained 1% of FBS and the final compound concentration was 10 μM. Immediately after FBS addition, the reaction mixtures were set up in a thermostatted compartment maintained at 37ºC, and 100 μL injections were made with an automated sample injector at various time points. Usually, a single sample set covered unmodified cap analog **1** and a pair of α/β and β/γ isomers, or D1 and D2 isomers bearing the same modification, so that each compound could be injected at 8 different time points at 60 min intervals. Afterwards, each assay column was washed with methanol for at least 1h before the next analysis. Based on the integration of the remaining cap analog and resultant products, substrate conversion for each compound was estimated using the equation: $\Sigma(AUCp/\varepsilon)/\Sigma[(AUCp/\varepsilon)+AUCs/\varepsilon]$, where AUC(p,s) – area under curve of products (p) and starting material (s), ε-molar extinction coefficient. Extinction coefficient values used for calculations were as follows: for trimethylated products ε = 12600 $M^{-1}cm^{-1}$, for guanosine analogs ε = 11400 $M^{-1}cm^{-1}$, for adenosine analogs ε = 15400 $M^{-1}cm^{-1}$ and for the cap analogs ε = 24600 $M^{-1}cm^{-1}$. The obtained data were analyzed using OriginPro software v. 9.1

Protein expression and purification

Wild-type human surportin 1 (wtSPN1, residues 1-360, MW 44 kDa) was expressed as an N-terminally His-tagged (MGHHHHHHHHHHSSGHIEGRH) protein in an *E. coli* Rosetta2 (DE3) strain transformed with a pET16b_wtSPN1 vector, where the snurportin coding sequence was cloned at *Nde*I-*XhoI* restriction sites (verified further by sequencing). wtSPN1 was expressed in LB medium supplemented with ampicillin (100 μg/mL). Expression was induced at OD600 ~ 0.6 with 0.5 mM IPTG and after 4 hours bacterial cells were pelleted, washed in PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) and the bacterial pellet was stored at -80°C for up to one week. Subsequently, pelleted cells were resuspended in a lysis buffer (20 mM HEPES-KOH pH 7.5, 300 mM NaCl, 300 mM Urea, 10 mM imidazole, 10% glycerol, 1% Triton 100 and 0.5 mM PMSF) supplemented with lysozyme, incubated on ice for 30 minutes and then disrupted by sonication (12 cycles for 35 sec, Bondelin Sonopuls). After centrifugation (17 500g for 30 min), the supernatant was added to NTA-agarose equilibrated with 20 mM HEPES buffer (20 mM HEPES-KOH pH 7.5, 300 mM NaCl, 10% Glycerol and 5 mM imidazole) and incubated for 1 hour in a cold room with gentle stirring. Unbound proteins were removed by washing with 20mM Tris HCl pH 7.5, 300 mM NaCl (10 x 2mL). His-tagged wtSPN1 was eluted from the resin using increasing concentrations of imidazole (20 mM to 300 mM) in wash buffer (20 mM Tris HCl pH 7.5, 300 mM NaCl). The collected fractions of His-tagged wtSPN1 were further purified (at 10 ºC) by applying the samples onto a Superdex 200 gel filtration column (GE Healthcare Life Science) equilibrated with 50 mM HEPES-KOH, 150 mM NaCl, 0.5 mM EDTA and 2 mM DTT, using an AKTA FPLC system (Amersham Parmacia). Aliquots of purified wtSPN1 (supplemented with 10% glycerol) were frozen in liquid nitrogen and stored at -80 ºC.

A truncated version of human SPN1 (MW 37 kDa) was expressed as a N-terminally His-tagged protein using a pET16b_shSNP1 vector, where the PCR product with a sequence encoding residues 66-360 of wtSNP1 was cloned at the *Nde*I-*XhoI* restriction sites of pET16b (verified further by sequencing). The purification steps for the shSPN1 protein were the same as those described for wtSPN1, except that the IPTG-induced *E. coli* culture was incubated at 18 ºC for 18 hours (which increased the expression of shSPN1 in the soluble fraction).

Human Nudt16 protein (hNudt16, 1-199 aa, MW 22 kDa) was expressed in *E. coli* Rosetta2 (DE3) as a C-terminally His-tagged protein using a pET16b_Nudt16 vector, where the coding sequence of Nudt16 was cloned at the *Nco*I-*BamH*I restriction sites (verified further by sequencing). A His-tag

composed from a sequence encoding four additional histidines was introduced just after the two terminal histidines of Nudt16 which resulted in a 6xHis C-terminal His-tag. Expression was induced at OD600 ~ 0.6 with 0.2 mM IPTG and after 4 hours incubation at 37°C bacterial cells were pelleted and washed in PBS buffer. Subsequently, the pelleted cells were resuspended in a lysis buffer (20 mM HEPES-KOH pH 8.0, 300 mM NaCl, 300 mM Urea, 10 mM imidazole, 10% glycerol, 1% Triton 100) supplemented with lysozyme, incubated on ice for 30 minutes and then disrupted by sonication. After centrifugation (17 500g for 30 min), the supernatant was added to NTA-agarose equilibrated with 20 mM HEPES buffer (20 mM HEPES-KOH pH 8.0, 300 mM NaCl, 10% glycerol and 5 mM imidazole) and incubated for 1 hour in a cold room with gentle stirring. Unbound proteins were removed by washing with 20 mM Tris HCl pH 8.0, 300 mM NaCl (10 x 2mL). The His-tagged hNudt16 protein was eluted form the resin using increasing concentrations of imidazole (20 mM to 300 mM) in wash buffer (20 mM Tris HCl pH 8.0, 300 mM NaCl). The collected fractions of pure hNudt16 protein were dialyzed against 50 mM Tris-HCl pH 8.0, 150 mM KCl and 20% glycerol. Aliquots of purified hNudt16 were supplemented with 1 mM DTT, frozen in liquid nitrogen and stored at -80ºC.

Analysis of capped ribooligonucleotide susceptibility towards hNudt16 decapping

Short capped RNAs (cap-GCCCC) carrying various cap structures (**1**, **2**, **6**, **7**, **9a** or **9b**) were synthesized by transcription *in vitro* as described previously[19]. Decapping reactions were performed with recombinant hNudt16 (final concentration of 0.2 μg/ml) on cap-GCCCC RNAs at 100 μM concentration in transcription buffer (40 mM Tris-HCl, 6 mM $MgCl_2$, 10 mM DTT, 10 mM NaCl, 2 mM spermidine, pH 7.9) for 1 hour at 37°C in a final volume of 10 μl. Inactivated hNudt16 (100 °C for 5 minutes) was used in control reactions. The reaction products (mixed with formazol 1:1, FORMAzol, Molecular Research Center, Inc.) were separated on 20% denaturating polyacrylamide gel with 7 M urea in 1xTBE buffer and visualized by UV shadowing.

Spectrofluorimetric measurements

Fluorescence spectra were recorded on an LS-55 fluorometer (Perkin-Elmer, USA) in a thermostatted, quartz semi-micro cuvette (Hellma, Germany) with a magnetic stirrer, and the temperature of 20.0 ± 0.3 °C was controlled with a thermocouple inside the cuvette. The optical path lengths were 4 mm and 10 mm for absorption and emission, respectively. The fluorescence intensity was monitored by the registration of entire spectra and during continuous titrations of snurportin by cap analogs at a single wavelength set. The excitation wavelength of 280 nm (slit 2.5 nm, auto cut-off filter), and the emission wavelength of 345 nm (slit 4 nm, 290 nm cut-off filter) were applied, with an automatic correction for photomultiplier sensitivity, with the integration time of 30 seconds and a gap of 30 seconds for the addition of the ligand. During the gap, the UV xenon flash lamp was switched off to avoid the photobleaching of the sample. Fluorescence quenching experiments were performed in 50 mM HEPES/KOH pH 7.20, 150 mM NaCl, 2 mM DTT and 0.5 mM EDTA. Buffer solutions were filtered through 0.22 μm and degassed. The concentration of snurportin wild-type and its truncated form was determined from the absorbance, using ε_{280} = 58 500 $cm^{-1}M^{-1}$ and 54 100 $cm^{-1}M^{-1}$, respectively. Titrations were performed at several snurportin concentrations (0.08 to 0.4 μM) under steady-state conditions. Aliquots (1 μl) of solutions with increasing ligand concentrations (5 μM to 2 mM) were injected manually to 1400 μl of snurportin solution. The fluorescence intensities were corrected for the *inner filter* effect and for the sample dilution (≤ 3 %).

Fluorescence data analysis

The equilibrium association constants (K_{as}) and the corresponding standard molar Gibbs free energies of binding ($\Delta G°$) were determined by fitting the theoretical dependence of the fluorescence intensity on the total ligand concentration to the experimental data points, as described previously.[46] Numerical analyses were performed by non-linear, least-squares regression using ORIGIN 9.0 (OriginLab, USA) and Prism 6.01 (GraphPad, USA). The final K_{as} values were calculated as weighted geometric means[46] from three to five independent titrations selected on the basis of statistical criteria: run tests ($P > 0.1$) and goodness of fit ($R^2 > 0.99$). The final measurement uncertainties given in Table 5 provide for the propagation of both the error of the weighted means and the statistical standard error of the means (SEM).

Acknowledgments

We would like to thank Roger Strömberg, Malgorzata Honcharenko and Edvard Smith (Karolinska Institute, Stockholm) for inspiring discussions and to Natalia Stelmaszuk (University of Warsaw) for technical assistance with RNA gels. This project was supported by a grant from the Polish National Center of Research and Development (02/EuroNanoMed/2011). M.Z. is supported by the Foundation for Polish Science International Ph.D. Projects Programme, co-financed by the EU's European Regional Development Fund.

Notes and references

[a]Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Zwirki i Wigury 93, 02-089 Warsaw, Poland
[b]Centre of New Technologies, University of Warsaw, Banacha 2c, 02-089 Warsaw, Poland;
[c] *Laboratory of Biological Physics, Institute of Physics, Polish Academy of Sciences, 32/46 Lotnikow Ave., 02-668 Warsaw, Poland*
E-mail: jacekj@biogeo.uw.edu.pl, tel: +48 22 55 40 774, fax: +48 22 55 40 771

† Electronic Supplementary Information (ESI) available: RP HPLC profiles, HRMS and NMR spectra of the novel compounds. See DOI: 10.1039/b000000x/

1. Y. Furuichi, A. J. Shatkin, *Adv. Virus Res.*, 2000, ***55***, 135-184.
2. Y. Motorin, M. Helm, *Wiley Interdisciplinary Reviews-RNA*, 2011, ***2***, 611-631.

3. R. E. Rhoads, *J Biol Chem*, 2009, ***284,*** 16711-16715.
4. A. G. Matera, R. M. Terns, M. P. Terns, *Nat. Rev. Mol. Cell Biol.* 2007, ***8***, 209-220.
5. R. F. Liou, T. Blumenthal, *Mol. Cell. Biol.* 1990, ***10***, 1764-1768.
6. J. Mouaikel, C. Verheggen, E. Bertrand, J. Tazi, R. Bordonne, *Mol. Cell* 2002, ***9***, 891-901.
7. J. Hamm, E. Darzynkiewicz, S. M. Tahara, I. W. Mattaj, *Cell* **1990,** *62*, 569-577.
8. J. Huber, U. Cronshagen, M. Kadokura, C. Marshallsay, T. Wada, M. Sekine, R. Luhrmann, *EMBO J*, 1998, ***17***, 4114-4126.
9. P. M. D. Moreno, M. Wenska, K. E. Lundin, O. Wrange, R. Stromberg, C. I. E. Smith, *Nucleic Acids Res.* 2009, ***37,*** 1925-1935.
10. J. C. van Deutekom, A. A. Janson, I. B. Ginjaar, W. S. Frankhuizen, A. Aartsma-Rus, M. Bremmer-Bout, J. T. den Dunnen, K. Koop, A. J. van der Kooi, N. M. Goemans, S. J. de Kimpe, P. F. Ekhart, E. H. Venneker, G. J. Platenburg, J. J. Verschuuren, G. J. B. van Ommen, *New Engl. J. Med.*, 2007, ***357,*** 2677-2686.
11. A. H. M. Burghes, V. L. McGovern, *Gene. Dev.*, 2010, ***24,*** 1574-1579.
12. G. W. Lu, J. Zhang, Y. Li, Z. X. Li, N. Zhang, X. Xu, T. T. Wang, Z. H. Guan, G. F. Gao, J. H. Yan, *Protein Cell*, 2011, ***2*** , 64-73.
13. J. Jemielity, J. Kowalska, A. M. Rydzik, E. Darzynkiewicz, *New J Chem*, 2010, ***34,*** 829-844.
14. E. Grudzien-Nogalska, J. Jemielity, J. Kowalska, E. Darzynkiewicz, R. E. Rhoads, *RNA*, 2007, ***13***, 1745-1755.
15. W. Su, S. Slepenkov, E. Grudzien-Nogalska, J. Kowalska, M. Kulis, J. Zuberek, M. Lukaszewicz, E. Darzynkiewicz, J. Jemielity, R. E. Rhoads, *RNA* 2011, ***17,*** 978-988.
16. E. Grudzien, M. Kalek, J. Jemielity, E. Darzynkiewicz, R. E. Rhoads, *J. Biol. Chem.*, 2006, ***281,*** 1857-1867.
17. M. Ziemniak, M. Strenkowska, J. Kowalska, J. Jemielity, *Future Med. Chem.*, 2013, ***5***, 1141-1172.
18. (a) A. N. Kuhn, M. Diken, S. Kreiter, A. Selmi, J. Kowalska, J. Jemielity, E. Darzynkiewicz, C. Huber, O. Tureci, U. Sahin, *Gene Ther* 2010, ***17***, 961-971; (b) J. Kowalska, A. Wypijewska del Nogal, Z. M. Darzynkiewicz, J. Buck, C. Nicola, A. Kuhn, M. Lukaszewicz, J. Zuberek, M. Strenkowska, M. Ziemniak, M. Maciejczyk, E. Bojarska, R. E. Rhoads, E. Darzynkiewicz, U. Sahin, and J. Jemielity submitted
19. J. Kowalska, M. Lukaszewicz, J. Zuberek, M. Ziemniak, E. Darzynkiewicz, E.; J. Jemielity, *Bioorg Med Chem Lett* 2009, ***19***, 1921-1925.
20. A. M. Rydzik, M. Kulis, M. Lukaszewicz, J. Kowalska, J. Zuberek, Z. M. Darzynkiewicz, E. Darzynkiewicz, J. Jemielity, *Bioorg.Med. Chem.* 2012, ***20,*** 1699-1710.
21. M. Honcharenko, M. Zytek, B. Bestas, P. Moreno, J. Jemielity, E. Darzynkiewicz, C. I. E. Smith, R. Stromberg, *Bioorg. Med. Chem.* 2013, ***21***, 7921-7928.
22. A. Niedzwiecka, J. Marcotrigiano, J. Stepinski, M. Jankowska-Anyszka, A. Wyslouch-Cieszynska, M. Dadlez, A. C. Gingras, P Mak, E. Darzynkiewicz, N. Sonenberg, S. K. Burley, R. Stolarski, *J. Mol. Biol.* 2002, ***319,*** 615-635.
23. (a) E. Darzynkiewicz, J. Stepinski, I. Ekiel, Y. Jin, D. Haber, T. Sijuwade, S. M. Tahara,. *Nucleic Acids Res.* 1988, ***16***, 8953-8962; (b) R. Iwase, M. Sekine, Y. Tokumoto,Y. Ohshima, T. Hata, *Nucleic Acids Res.* 1989, ***17***, 8979-8989.
24. (a) M. Sekine, M. Kadokura, T. Satoh, K. Seio, T. Wada, U. Fischer,V. Sumpter, R. Luhrmann, R., *J. Org. Chem.* 1996, ***61***, 4412-4422; (b) A. Ohkubo, K. Sasaki, Y. Noma, H. Tsunoda, K. Seio, M. Sekine, *Bioorg. Med. Chem.* 2009, ***17*** , 4819-4824.
25. A. Ohkubo, N. Tago, A. Yokouchi, Y. Nishino, K. Yamada, H. Tsunoda, K. Seio, M. Sekine, *Org. Lett.*, 2012, ***14,*** 10-13.
26. J. Stepinski, M. Bretner, M. Jankowska, K. Felczak, R. Stolarski, Z. Wieczorek, A. L. Cai,R. E. Rhoads, A. Temeriusz, D. Haber, E. Darzynkiewicz, *Nucleosides & Nucleotides* 1995, ***14***, 717-721.
27. M. Jankowska-Anyszka, K. Piecyk, J. Samonina-Kosicka, *Org. Biomol. Chem.*, 2011, ***9*** , 5564-5572.
28. M. Honcharenko ,J. Romanowska, M. Alvira, M. Jezowska, M. Kjellgren, C. I. E. Smith, R. Stromberg, *RSC Advances*, 2012**,** ***2,*** 12949-12962.
29. A. Ohkubo, Y. Kondo, M. Suzuki, H. Kobayashi, T. Kanamori, Y. Masaki, K. Seio, K. Nagai, M. Sekine, *Org. Lett.*, 2013, ***15,*** 4386-4389.
30. A. Matsuda, M. Shinozaki, M. Suzuki, K. Watanabe, T. Miyasaka, *Synthesis-Stuttgart* 1986, **5,** 385-386.
31. M. Yoshikawa, T. Kato, T. Takenish, *Tetrahedron Lett.* 1967, **50**, 5065-5068.
32. T. Mukaiyama, M. Hashimoto, *Bull. Chem. Soc. Jpn.*, 1971**,** ***44,*** 2284-2284.
33. J. Kowalska, M. Lewdorowicz, J. Zuberek, E. Grudzien-Nogalska, E. Bojarska, J. Stepinski, R. E. Rhoads, E. Darzynkiewicz, R. E. Davis, J. Jemielity. *RNA*, 2008**,** ***14*** , 1119-1131.
34. J. Kowalska, M. Lewdorowicz, E. Darzynkiewicz, J. Jemielity, *Tetrahedron Lett.* 2007, ***48***, 5475-5479.
35. J. W. Conway, C. K. McLaughlin, K. J. Castor, H. Sleiman, *Chem. Comm.*, 2013**,** ***49***, 1172-1174.
36. E. van Dijk, N. Cougot, S. Meyer, S. Babajko, E. Wahle, B. Seraphin, *EMBO J*, 2002, ***21***, 6915-6924.
37. M. J. Bessman, D. N. Frick, S. F. Ohandley, *J. Biol. Chem.*, 1996, ***271,*** 25059-25062.
38. C. Piccirillo, R. Khanna, M. Kiledjian, *RNA* 2003**,** ***9*** , 1138-1147.
39. A. Strasser, A. Dickmanns, R. Luhrmann, R. Ficner, *EMBO J*, 2005, ***24***, 2235-2243.
40. M. Kalek, J. Jemielity, E. Grudzien, J. Zuberek, E. Bojarska, L. S. Cohen, J. Stepinski, R. Stolarski, R. E. Davis, R. E. Rhoads, E. Darzynkiewicz, *Nucleos Nucleot Nucl*, 2005, ***24*** , 615-621.
41. J. Marcotrigiano, A. C. Gingras, N. Sonenberg, S. K. Burley, *Cell*, 1997, ***89***, 951-961.
42. A. E. Hodel, P. D. Gershon, F. A. Quiocho, *Mol. Cell*, 1998, ***1***, 443-447.
43. T. Monecke, A, Dickmanns, R. Ficner, R, *Nucleic Acids Res.* 2009**,** ***37*** **,** 3865-3877.
44. D. Guilligay, F. Tarendeau, P. Resa-Infante, R. Coloma, T. Crepin, P. Sehr, J. Lewis, R. W. H. Ruigrok, J. Ortin, D. J. Hart, S. Cusack, *Nat. Struct. Mol. Biol.* 2008, ***15*** **,** 500-506.
45. M. G. Gu, C. Fabrega, S. W. Liu, H. D. Liu, M. Kiledjian, C. D. Lima, *Mol Cell.*, 2004, ***14***, 67-80.

46. A. Niedzwiecka, J. Stepinski, J. M. Antosiewicz, E. Darzynkiewicz,R. Stolarski, *Methods Enzymol* 2007, ***430,*** 209-245.

