High-quality oligo-RNA synthesis using the new 2′-*O*-TEM protecting group by selectively quenching the addition of *p*-tolyl vinyl sulphone to exocyclic amino functions[1]

Chuanzheng Zhou, Wimal Pathmasiri, Dmytro Honcharenko, Subhrangsu Chatterjee, Jharna Barman, and Jyoti Chattopadhyaya

Abstract: During the F^--promoted deprotection of the oligo–RNA, synthesized using our 2′-*O*-(4-tolylsulfonyl)ethoxymethyl (2′-*O*-TEM) group [Org. Biomol. Chem. **5**, 333 (2007)], *p*-tolyl vinyl sulphone (TVS) is formed as a by-product. The TVS formed has been shown to react with the exocyclic amino functions of adenosine (A), guanosine (G), and cytidine (C) of the fully deprotected oligo–RNA to give undesirable adducts, which are then purified by HPLC and unambiguously characterized by ^{1}H, ^{13}C Heteronuclear Multiple Bond Correlation (HMBC) NMR and mass spectroscopic analysis. The relative nucleophilic reactivities of the nucleobases toward TVS have been found to be the following: N^6–A > N^4–C > N^2–G > > N^3–U. This reactivity of TVS toward RNA nucleobases to give various Michael adducts could, however, be suppressed by using various amines as scavengers. Among all these amines, morpholine and piperidine are the most efficient scavenger for TVS, which gave highly pure oligo–RNA even in the crude form and can be used directly in RNA chemical biology studies.

Key words: RNA synthesis, RNA alkylation, *p*-tolyl vinyl sulphone, Michael addition.

Résumé : Au cours de la déprotection sous l'influence de l'ion fluorure de l'oligo–ARN synthétisé à l'aide de notre groupe protecteur 2′-*O*-(4-tolylsulfonyl)éthoxyméthyl (2′-*O*-TEM) [Org. Biomol. Chem. **5**, 333 (2007)], il y a formation de la *p*-tolyl vinyl sulfone (TVS) comme sous-produit. Il a été montré que la TVS réagit avec les fonctions aminées exocycliques de l'adénosine (A), de la guanosine (G) et de la cytidine (C) de l'oligo-ARN complètement déprotégé pour conduire à la formation de produits indésirables qui ont été purifiés et caractérisés sans ambiguïté par spectrométrie de masse et par corrélation des liaisons multiples hétéronucléaires (CLMH) des spectres RMN du ^{1}H et du ^{13}C. On a établi que les réactivités nucléophiles relatives des nucléobases vis-à-vis de la TVS sont dans l'ordre suivant: N^6–A > N^4–C > N^2–G > > N^3–U. En utilizant diverses amines comme pièges, il serait toutefois possible d'éliminer cette réactivité de la TVS vis-à-vis des nucléobases pour conduire à divers adduits de Michael. Parmi ces amines, la morpholine et la pipéridine sont les pièges les plus efficaces pour la TVS; leur utilization conduit, même dans le produit brut, à la formation d'oligo–ARN très pur qui peut être utilizé dans des études de chimie biologique de l'ARN

Mots-clés : synthèse d'ARN, alkylation d'ARN, *p*-tolyl vinyl sulfone, addition de Michael.

[Traduit par la Rédaction]

Introduction

The discovery of RNA interference (RNAi) (1) and especially the finding that small interfering RNA (siRNA) (2) can induce degradation of the target RNA have highlighted the need for facile chemical synthesis of oligo–RNA sequences (e.g., a synthetic 21 mer siRNA) (2) and have thereby stimulated work on the development of new methodology for the chemical synthesis of RNA (3–7). Recently, we have developed a new RNA synthesis strategy based on protecting 2′-OH with a 2-(4-tolylsulfonyl)ethoxymethyl (2′-*O*-TEM) (8) group and showed its successful use in the solid-phase synthesis of at least 12 different oligo–RNAs of varying sizes (14–38 mer long) and sequences (8). The main advantages of this 2′-*O*-TEM group based RNA synthesis strategy (8) are (*i*) the low coupling time (120s), (*ii*) high

Received 5 February 2007. Accepted 13 March 2007. Published on the NRC Research Press Web site at canjchem.nrc.ca on 27 April 2007.

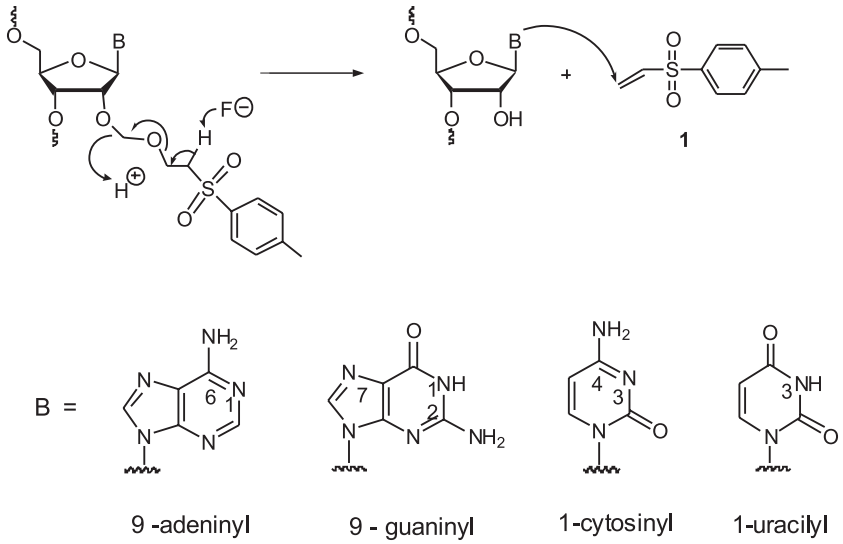
In honor of the career contributions of Professor Kelvin Kenneth Ogilvie and for his significant contributions in the field of nucleoside/nucleic acid chemistry and automated DNA and RNA synthesis.

C. Zhou, W. Pathmasiri, D. Honcharenko, S. Chatterjee, J. Barman, and J. Chattopadhyaya.[2] Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden.

[1]This article is part of a Special Issue dedicated to Professor Kelvin K. Ogilvie.
[2]Corresponding author (e-mail: jyoti@boc.uu.se).

 doi:10.1139/V07-025

Scheme 1. Mechanism of deprotection of the 2′-*O*-TEM group and the adduct formation with tolyl vinyl sulfone (**1**).



stepwise coupling yield (97%–99%), leading to (*iii*) the crude RNA, obtained after a simple deprotection step. The crude form RNA was sufficiently pure for chemical biology research without the need for any additional purification steps.

The 2′-*O*-TEM group is stable upon ammonia treatment and can be quickly cleaved by 1 mol/L tetrabutylammonium fluoride (TBAF) in THF through F^--promoted β elimination (Scheme 1). But the by-product *p*-tolyl vinyl sulphone (TVS) (**1**), which is generated during deprotection, attacks the nucleophilic sites of nucleobases of the oligo–RNA to form some undesirable side products that somewhat reduce the purity of the oligo–RNA. To obtain clean oligo–RNA at the end of the synthesis, we had earlier employed a mixture of 10% (*v*/*v*) *n*-propylamine and 1% (*v*/*v*) bis(2-mercaptoethyl)ether in 1 mol/L TBAF–THF as a scavenger for TVS during the deprotection of the 2′-*O*-TEM group. Clearly, we wanted to find other neutral alternative(s) to this scavenger because bis(2-mercaptoethyl)ether has a foul odour and it can also deactivate many enzymes, even in trace amounts (ppm).

Earlier work (8) has excluded N^3 of 1-uracilyl as a reactive centre for the side product formation. To this point, the potential nucleophilic sites have been known to be N^1 of 9-adeninyl (9, 10), N^7 and N^1 of 9-guaninyl (11, 12), and N^3 of 1-cytosinyl (13, 14). In this study, through isolation of the pure adducts formed with TVS and by identifying them unambiguously by MS and NMR, we have unexpectedly found that N^2 of 9-guaninyl, N^4 of 1-cytosinyl, and especially N^6 of 9-adeninyl are the reactive nucleophilic sites for adduct formation with TVS, and more powerful scavengers were therefore developed.

Results and discussion

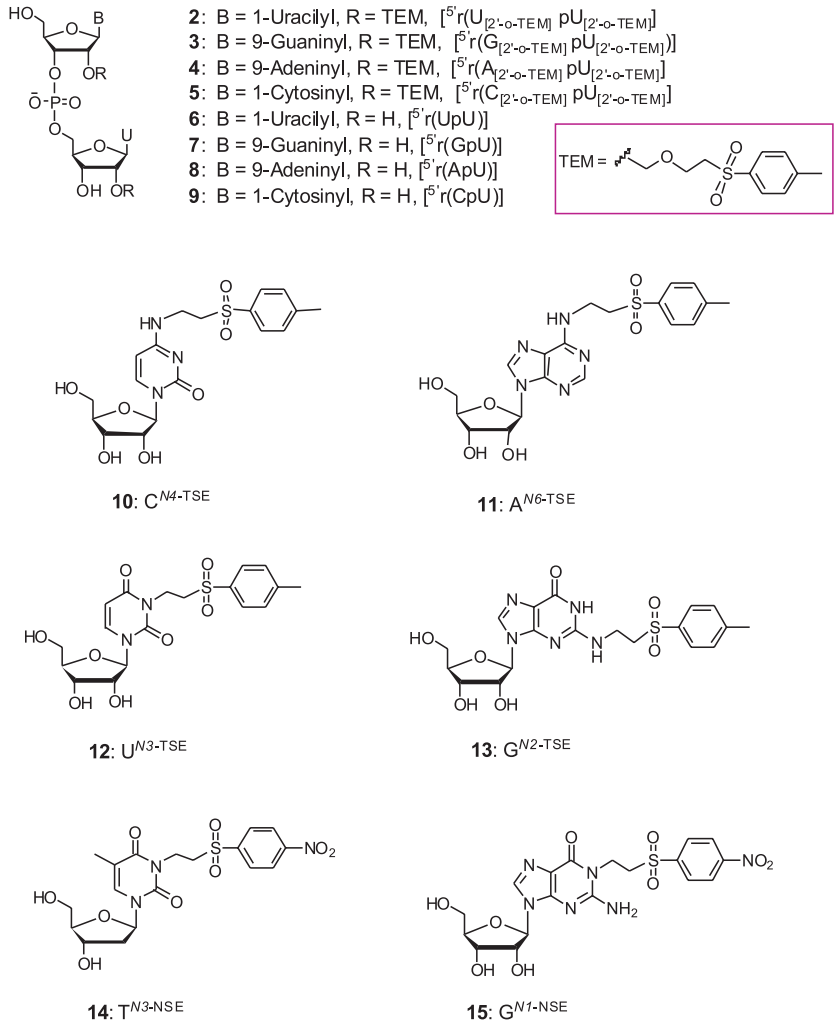
Deprotection of the 2′-*O*-TEM group from dimeric RNAs

It was known (8) that when deprotecting with 1 mol/L TBAF–THF, pure poly-U can be easily obtained as the sole product, hence the lactam function of 1-uracilyl and the internucleotidic phosphate backbone are not amenable to adduct formation with TVS. We anticipated that the possible nucleobases that could potentially be involved in the adduct formation with TVS are 9-guaninyl, 9-adeninyl, and 1-cytosinyl moieties in oligo–RNA. Keeping this in mind, four dimeric ribonucleotides, $^{5'}r(U_{[2'\text{-}o\text{-TEM}]}\ pU_{[2'\text{-}o\text{-TEM}]})$ (**2**), $^{5'}r(G_{[2'\text{-}o\text{-TEM}]}\ pU_{[2'\text{-}o\text{-TEM}]})$ (**3**), $^{5'}r(A_{[2'\text{-}o\text{-TEM}]}\ pU_{[2'\text{-}o\text{-TEM}]})$ (**4**), $^{5'}r(C_{[2'\text{-}o\text{-TEM}]}\ pU_{[2'\text{-}o\text{-TEM}]})$ (**5**), (Scheme 2), were synthesized. They were treated with 1 mol/L TBAF–THF for 4 h at RT, then the reaction mixtures were analyzed by reverse phase HPLC (RP-HPLC). The RP-HPLC profiles are shown in Fig. 1.

As expected, $^{5'}r(U_{[2'\text{-}o\text{-TEM}]}\ pU_{[2'\text{-}o\text{-TEM}]})$ (**2**) upon treatment with 1 mol/L TBAF–THF gave a single product (Fig. 1*A*, R_T = 5.89 min), which proved to be the desired product $^{5'}r(UpU)$ (**6**) by MALDI-TOF MS (*m*/*z* calcd.: 551.10 $[MH]^+$; found: 551.11). Treatment of $^{5'}r(G_{[2'\text{-}o\text{-TEM}]}\ pU_{[2'\text{-}o\text{-TEM}]})$ (**3**) with 1 mol/L TBAF–THF also gave the desired product $^{5'}r(GpU)$ (**7**). (Fig. 1*B*, R_T = 6.56 min, *m*/*z* calcd.: 590.12 $[MH]^+$; found: 590.11). We therefore could safely exclude 1-uracilyl and 9-guaninyl aglycons as substrates for adduct formation with TVS.

The RP-HPLC analysis of the reaction of $^{5'}r(C_{[2'\text{-}o\text{-TEM}]}\ pU_{[2'\text{-}o\text{-TEM}]})$ (**5**) is shown in Fig. 1*C*. The product with R_T = 4.75 min was proved by MS to be the fully deprotected product $^{5'}r(CpU)$ (**9**) (*m*/*z* calcd.: 550.12 $[M]^-$; found: 550.12). However, the peak appearing at 18.43 min (Fig. 1*C*) was the major product, which was analyzed by MS (*m*/*z* found: 730.32 $[M]^-$). This suggested that this product was formed by the reaction of one molecule of PVS with $^{5'}r(CpU)$ (**9**). This major product was separated and its base composition was analyzed by digesting it with mixture of phosphodiesterase I and shrimp alkaline phosphatase (see experimental section). The digestion mixture was applied to RP-HPLC analysis. As shown in the Fig. 1*C* inset, the digestion gave two products. One appears at 4.06 min, corresponding to uridine. The second peak appears at 17.01 min. It was probably formed by the addition of TVS to cytidine and clearly needed further characterization (represented by **X** in Fig. 1*C*).

Scheme 2. **10**: N^4-(4-tolylsulfonylethyl)cytidine ($C^{N4\text{-}TSE}$); **11**: N^6-(4-tolylsulfonylethyl)adenosine ($A^{N6\text{-}TSE}$); **12**: N^3-(4-tolylsulfonylethyl)uridine ($U^{N3\text{-}TSE}$); **13**: N^2-(4-tolylsulfonylethyl)guanosine ($G^{N2\text{-}TSE}$); **14**: N^3-(4-nitrophenylsulfonylethyl)thymidine ($T^{N3\text{-}NSE}$); **15**: N^1-(4-nitrophenylsulfonylethyl)guanosine ($G^{N1\text{-}NSE}$).



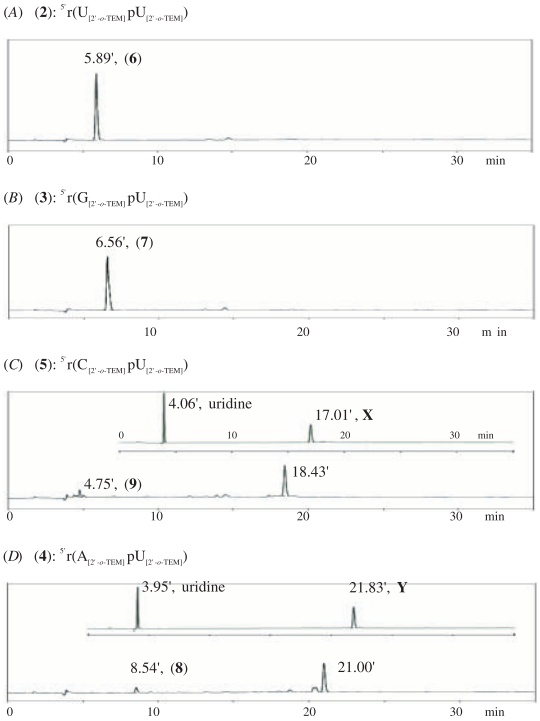
Dimer $^{5'}r(A_{[2'\text{-}o\text{-}TEM]}\ pU_{[2'\text{-}o\text{-}TEM]})$ (**4**) was treated with 1 mol/L TBAF–THF and gave a very small amount of the desired product $^{5'}r(ApU)$ (**8**) (Fig. 1*D*, R_T = 8.54 min, *m*/*z* calcd.: 572.41[M]$^-$; found: 572.31). The predominant reaction product (*m*/*z* found: 754.34 [M]$^-$) appears at 21.00 min. Base composition analysis of this major product by enzymatic digestion, as described earlier, gave uridine (Fig. 1*D* inset, R_T = 3.95 min) and another product (Fig. 1*D*, represented by Y, R_T = 21.83 min). This result suggested that both the cytidine and adenosine were reactive toward adduct formation with TVS.

Characterization of TVS adducts with aglycones of RNA

To determine the chemical nature of the adducts (**X** and **Y** in the HPLC elution profile in Figs. 1*C* and 1*D*), either cytidine or adenosine was treated with 3 equiv. of TVS at RT in 1 mol/L TBAF–THF for 24 h. The major product of the reaction between cytidine and TVS was separated by silica gel chromatography. This compound showed $[MH]^+$ at *m*/*z* 426.17, which suggested that it was an unimolecular adduct of TVS and cytidine. The chemical shift of β-C ($\delta_{\beta\text{-}C}$, 34.69 ppm, see Fig. 2 and Table 1) (12) suggested that the TVS is attached to the N atom, and the potential site could be either N^4 or N^3 of cytidine. The assignment of the covalent binding site of TVS could easily be achieved by ^1H, ^{13}C Heteronuclear Multiple Bond Correlation (HMBC) NMR experiment (12). As shown in Fig. 2*A*, the β-H shows a strong cross peak with C4 (shown by red line) but no cross-correlation with the C2, which strongly supports the explanation that the TVS has been formed as an adduct to N^4 of cytidine to give N^4-(4-tolylsulfonylethyl)cytidine ($C^{N4\text{-}TSE}$, **10**).

The reaction between adenosine and TVS gave a major product with $[MH]^+$ at *m*/*z* 450.18, which also indicated that a unimolecular adduct had formed between TVS and adenosine. Two unambiguous NMR proofs were used to confirm that TVS is indeed covalently bonded to the N^6 of adenosine to give product N^6-(4-tolylsulfonylethyl)adenosine ($A^{N6\text{-}TSE}$, **11**): (*i*) in the ^1H,^{13}C HMBC spectrum (Fig. 2*B*), the HN^6 shows cross coupling with C5, C6, and β-C, and (*ii*) more direct proof is the fact that the HN^6 of adenosine shows a strong cross peak with β-H (Fig. S2.1) in the ^1H, ^1H

Fig. 1. RP-HPLC analysis of the reaction products upon treatment of compounds **2–5** with 1 mol/L TBAF–THF at RT for 4 h. The products (R_T = 18.43 min in *C* and R_T = 21.00 min in *B* were digested respectively by mixture of phosphodiesterase and alkaline phosphatase. HPLC analysis of the digestion reaction is shown in the inset. RP-HPLC conditions: 0′→10′→35′, A →A–B 9:1, *v*/*v*→A–B 2:8, RT, flow rate 1 mL/min. Buffer A: 0.1 mol/L TEAA in H_2O–CH_3CN (*v*/*v*, 95:5). Buffer B: 0.1 mol/L TEAA in H_2O–CH_3CN (*v*/*v* 50:50).



COSY NMR spectrum, thereby confirming that the N^6 of adenosine has undergone a conjugate Michael addition to TVS.[3]

Co-injection of compound **10** with **X** on RP-HPLC has also shown that they are the same compound. In the same way, **Y** was shown to be compound **11** (Fig. S5.1 and S5.2).[3]

It has been reported that thymidine and guanosine can react with with acrylonitrile or phenylvinyl sulfone in the presence of a base such as tetra-*n*-butylammonium hydroxide (12), TBAF–THF (15), or ammonia (16, 17). These α,β-unsaturated compounds were adducted to the N^3 of thymidine (12, 15–17) and the N^1 of guanosine (12), respectively. Here, we have also studied the reaction between uridine or guanosine with TVS in 1 mol/L TBAF–THF. As reported (12), TVS is adducted to the N^3 of uridine (NMR spectra shown in Figs. 2*C* and S3) to give the product N^3-(4-tolylsulfonylethyl)uridine ($U^{N3\text{-}TSE}$, **12**).[3] For guanosine, the predominant product was proved to be N^2-(4-tolylsulfonylethyl)guanosine ($G^{N2\text{-}TSE}$, **13**) in an unambiguous manner by NMR and not $G^{N1\text{-}TSE}$, as earlier believed (12). Thus, (*i*) the imino proton is present ($\delta_{N1\text{-}H}$ = 10.94 ppm, see Fig. S4.1) in the ^1H NMR spectrum of **13**; (*ii*) the $\delta_{\beta\text{-}C}$ and $\delta_{\beta\text{-}H}$ (see Table 1) in **13** are much different than in compound N^1-(4-nitrophenylsulfonylethyl)guanosine ($G^{N1\text{-}NSE}$, **15**, Scheme 2) (12); (*iii*) in the ^1H,^{13}C HMBC spectrum (Fig. 2*D*), β-*H* shows a strong cross peak with C2 but none with C5, C6, or C8 in **13**.[3]

The Michael addition between the α,β-unsaturated compound, such as acrylonitrile with DNA or RNA under different conditions, leads to cyanoethylation of nucleobases, which has been studied extensively. Because of the natural electronic structure of nucleobases (18), the cyanoethylation generally occurs at N^3 of uridine (19) and cytidine (9), N^1 of adenosine (9, 20), N^7, and N^1 of guanosine (9, 12). Though cyanoethylation at N^6 of adenosine (20) was also observed, it was in fact obtained by Dimroth rearrangement (21) of N^1 of cyanoethylated adenosine in the presence of a strong base such as aqueous sodium hydroxide. Alkylation of nucleosides with other reagents such as diazomethane and dimethyl sulfate also occurs at the ring nitrogen atom such as N^7 and N^1 of purine (10, 11) and N^3 of pyrimidine (11, 14).

The present unexpected findings may have important applications in directly alkylating exocyclic amines of nucleosides as probes to determine the structure and (or) function of a nucleic acid (22, 23). Many methods have been developed for the alkylation of exocyclic amino groups (24–26), but these methods (24–26) are multi-step reactions and are time-consuming and laborious. To our knowledge, no direct alkylation method has been developed until now. The present finding that TVS can indeed be added selectively to an exocyclic amino of cytidine, adenosine, and guanosine provides an easier way to the direct alkylation of exocyclic amino groups of nucleobases if they exist as a part of a non-hydrogen-bonded loop or hairpin structure in the RNA scaffolds. Moreover, the phenyl moiety of TVS is relatively easy to derivatize for fluorescence detection purposes. Alternatively, by attaching other functional groups to the phenyl moiety, such as some specific cross-linking agents, we can conveniently introduce various functional groups to the non-hydrogen-bonded exocyclic amino of the nucleobases to map the folding in the RNA scaffolds by estimating spatial proximities.

Relative nucleophilic reactivities of the exocyclic amino groups of RNA toward TVS

Earlier we showed that in dimeric RNAs neither uridine nor guanosine participates in the adduct formation with TVS. Subsequently we showed that all of the four monomeric ribonucleosides are found to react with TVS. This apparent contradiction may be explained by comparing the nucleophilic reactivities of different ribonucleoside aglycons in an oligo–RNA toward TVS. We therefore chose a tetrameric RNA, 5′r(UpApCpG) (**16**, Scheme 3), for a model study; 0.5 OD (12.5 μmol) of **16** was treated with 0, 1/4, 1/2, 1, 2, 4, and 20 equiv. of TVS, respectively, in 1 mol/L TBAF–THF (total reaction volume 40 μL) at RT for 24 h. The aliquots of the reactions were analyzed by RP-HPLC (Fig. 3). The maldi-tof MS spectra showed that each of the

[3] Supplementary data for this article are available on the journal Web site (http://canjchem.nrc.ca) or may be purchased from the Depository of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Ottawa, ON K1A 0R6, Canada. DUD 5155. For more information on obtaining material refer to http://cisti-icist.nrc-cnrc.gc.ca/irm/unpub_e.shtml.

Fig. 2. ^{1}H, ^{13}C HMBC NMR spectrum of compounds **10** (panel *A*, crosspeaks at δ 3.81), **11** (panel *B*, crosspeaks at δ 6.9), **12** (panel *C*, crosspeaks at δ 4.36), and **13** (panel *D*, crosspeaks at δ 3.51, 3.58).

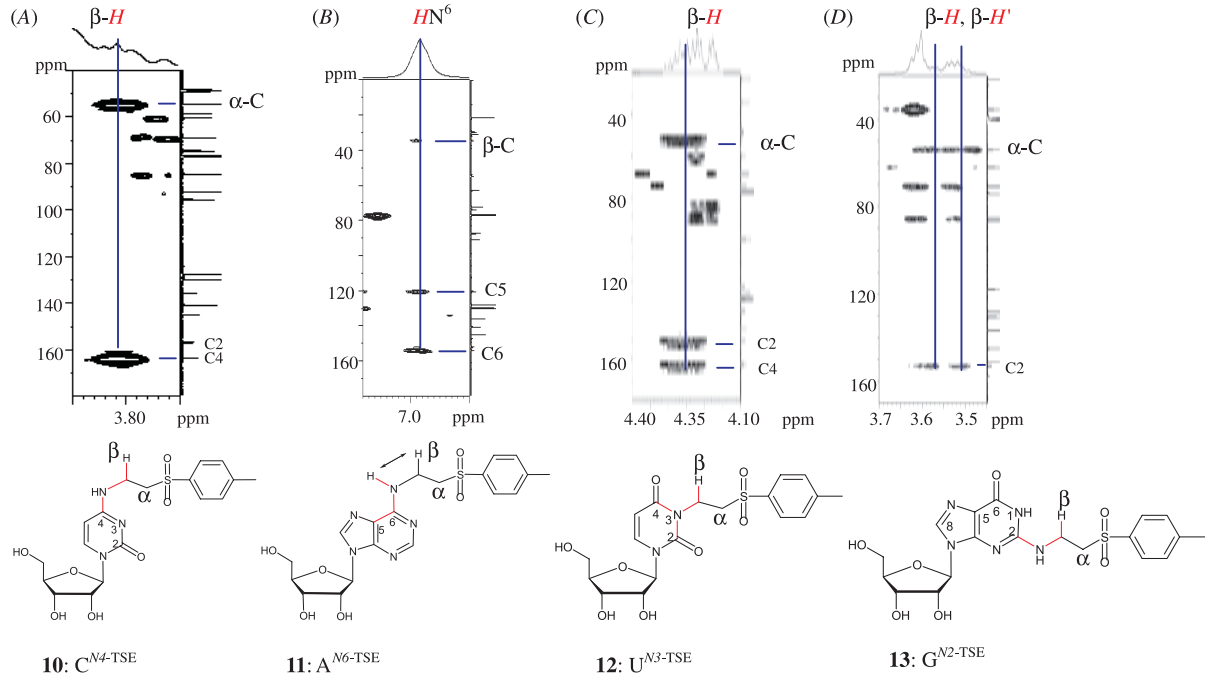


Table 1. Comparison of the chemical shifts of β-H and β-C of compounds **10**–**15**.

Compounds	$\delta_{\beta\text{-C}}$ (ppm)	$\delta_{\beta\text{-H}}$ (ppm)
10	34.69	3.80
11	34.65	3.89
12	34.95	4.25
13	35.66	3.61, 3.50
14[a]	34.60	4.12
15[a]	34.70	4.31

[a]The NMR data are taken from Mag and Engels (12).

three products **17** (*m/z* found: 1406.23 $[MH]^+$), **18** (*m/z* found: 1406.23 $[MH]^+$), and **19** (*m/z* found: 1406.26 $[MH]^+$) has only one molecule of TVS attached, thereby suggesting that the products **17**, **18**, and **19** are isomeric in nature. Furthermore, base composition analysis of these products by enzyme digestion showed that **17** is $^{5'}$r(UpApCpG$^{N2\text{-TSE}}$), **18** is $^{5'}$r(UpAp$C^{N4\text{-TSE}}$ pG), and **19** is $^{5'}$r(Up$A^{N6\text{-TSE}}$ pCpG) (see Fig. 4).

The yields of the three products in the reactions between $^{5'}$r(UpApCpG) (**16**) and different stoichiometries of TVS are listed in Table 2. In the presence of a limited amount of TVS, (1/4, 1/2 equiv.), only N^6 of adenosine was alkylated to give $^{5'}$r(Up$A^{N6\text{-TSE}}$ pCpG) (**19**). With the increase in concentration of TVS (1 or 2 equiv.), N^4 of cytidine was also alkylated to give $^{5'}$r(UpAp$C^{N4\text{-TSE}}$ pG) (**18**). By increasing the amount of TVS further to 20 equiv., a small amount of product by the alkylation of N^2 of guanosine was also observed to give $^{5'}$r(UpApCpG$^{N2\text{-TSE}}$) (**17**). This proves that the alkylation tendency of ribonucleotide by TVS in oligo–RNA in 1 mol/L TBAF–THF follows the sequence, N^6 – A > N^4 – C > N^2 – G > > N^3 – U. By controlling the concentration of TVS, selective alkylation of N^6 of adenosine, or both N^6 of adenosine and N^4 of cytidine, or all the exocyclic amino functions of ribonucleobases is possible. Since all the exocyclic amino functions participate in H-bond formation in the tertiary structure of nucleic acid, their selective alkylation by TVS gives the possibility of identifying those nucleobases that are not hydrogen bonded, and thus TVS can be used as a chemical probe for the structure and function of the nucleic acid.

Morpholine and piperidine as scavenger for TVS to avoid adduct formation

In the 2′-*O*-TEM group based solid-phase RNA synthesis (8), a mixture of 10% *n*-propylamine and 1% bis(2-mercaptoethyl) ether was used as the scavenger for TVS. Two disadvantages of employing bis(2-mercaptoethyl)ether as the deprotecting agent are its foul odour and its ability, in trace amounts, to possibly cause enzyme deactivatation in biological assays. In this study we wanted to use an appropriate amine as the sole scavenger to examine if the quality of the crude RNA could be further improved.

The Michael addition of amine to TVS (27–29) has shown that primary amines are notably less reactive than secondary amines because the former binds to two moles of solvent through H-bonds, compared with the latter that binds only to one mole of the solvent. It was also shown (27–29) that large differences in the basicity of amines has relatively little effect on the reactivity. Thus morpholine (pK_a = 8.4) is about one seventh as reactive as piperidine (pK_a = 11.1), but the former is > 100 times less basic. In our case, the ideal amine should have better nucleophilicity toward TVS than the exocyclic amino groups of the nucleobases of RNA, and at the same time the basicity of the amine cannot compromise any oligo–RNA chain cleavage. Keeping this in mind, several types of amines with very different pK_a values were se-

Scheme 3.

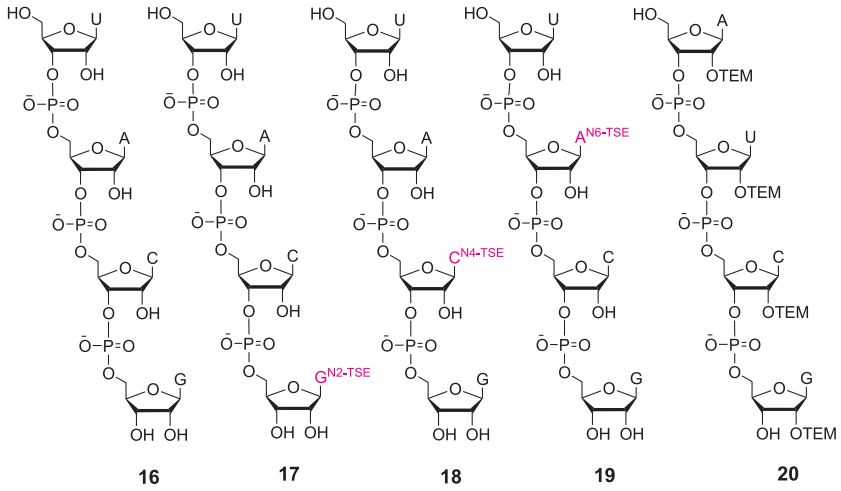
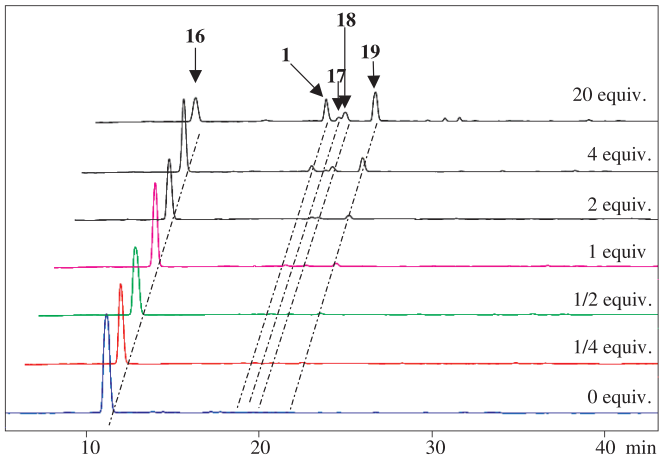


Fig. 3. RP-HPLC analysis of the reactions between $^{5'}$r(UpApCpG) (**16**) (0.5 OD, 12.5 µmol) with 0, 1/4, 1/2, 1, 2, 4, and 20 equiv. of TVS in 1 mol/L TBAF–THF at RT for 24 h. Each reaction has the same total volume of 40 µL. RP-HPLC conditions: 0′→10′→42′, A→A–B 9:1 →B, RT, flow rate 1 mL/min. Buffer A: 0.1 mol/L TEAA in H_2O–CH_3CN (*v*/*v*, 95:5). Buffer B: 0.1 mol/L TEAA in H_2O–CH_3CN (*v*/*v* 50:50).

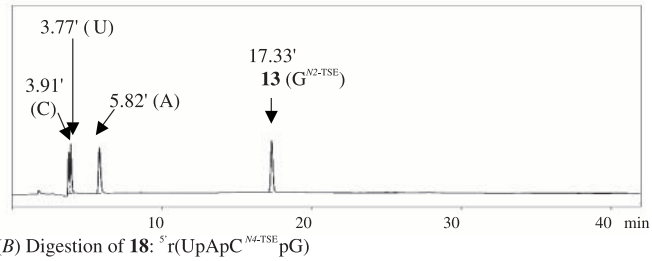


lected to test their abilities as scavengers of TVS in 1 mol/L TBAF–THF solution. The $^{5'}$r(A$_{[2'\text{-}o\text{-TEM}]}$ pU$_{[2'\text{-}o\text{-TEM}]}$ pC$_{[2'\text{-}o\text{-TEM}]}$ pG$_{[2'\text{-}o\text{-TEM}]}$ (**20**, Scheme 3) was used as the model compound and was treated with 1 mol/L TBAF–THF containing 10% (*v*/*v*) of different amines at RT for 24 h, and the reaction mixtures were analyzed by RP-HPLC (Fig. S7).[3] The percentage of chain cleavage product and side products formed through the reaction of TVS to oligo–RNA were obtained by integrating the HPLC profile (see Fig. S7) and they are given in Table 3.[3]

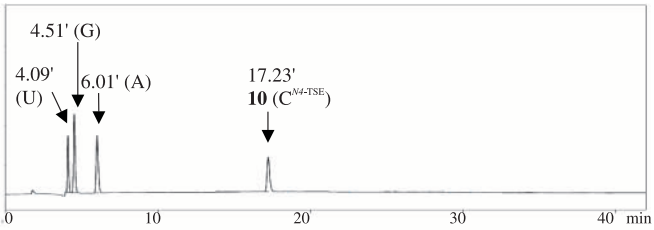
Table 3 shows that a primary or secondary amine, such as *n*-propylamine, piperidine, or morpholine, when used alone can scavenge TVS more efficiently than the standard deprotection condition (8) [a mixture of *n*-propylamine and bis(2-mercaptoethyl)ether] employed earlier for the deprotection of the 2′-*O*-TEM group. Among these amines, the morpholine and piperidine are especially attractive.

Fig. 4. The compounds **17**, **18**, and **19** were digested by a mixture of phosphodiesterase and alkaline phosphatase. The HPLC profiles of digestion mixture are shown in (*A*) **17**, (*B*) **18**, and (*C*) **19.** RP-HPLC conditions: 0′→10′→42′, A→A–B 9:1, *v*/*v* →B, RT, flow rate 1 mL/min. Buffer A: 0.1 mol/L TEAA in H_2O–CH_3CN (*v*/*v*, 95:5). Buffer B: 0.1 mol/L TEAA in H_2O–CH_3CN (*v*/*v* 50:50). Abbreviations are A = adenosine, G = guanosine, C = cytidine, U = uridine, $A^{N6\text{-TSE}}$ = N^6-(4-tolylsulfonylethyl)adenosine, $C^{N4\text{-TSE}}$ = N^4-(4-tolylsulfonylethyl)cytidine, $G^{N2\text{-TSE}}$ = N^2-(4-tolylsulfonylethyl)guanosine.

(*A*) Digestion of **17**: $^{5'}$r(UpApCpG$^{N2\text{-TSE}}$)



(*B*) Digestion of **18**: $^{5'}$r(UpApC$^{N4\text{-TSE}}$pG)



(*C*) Digestion of **19**: $^{5'}$r(UpA$^{N6\text{-TSE}}$pCpG)

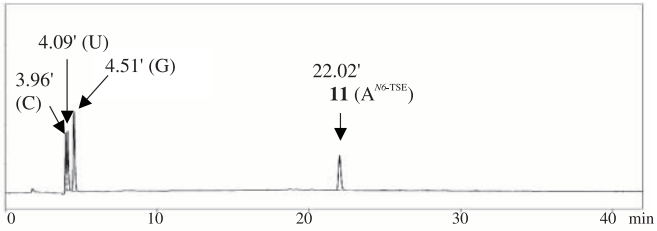


Table 2. The yields (%) of three isomeric products **17**, **18**, and **19** in the reactions between $^{5'}$r(UpApCpG) (**16**) and different stoichiometries of TVS (**1**).

Stoichiometry of TVS (**1**)	$^{5'}$r(UpApCpG$^{N2\text{-}TSE}$) (**17**)	$^{5'}$r(UpApC$^{N4\text{-}TSE}$ pG) (**18**)	$^{5'}$r(UpA$^{N6\text{-}TSE}$ pCpG) (**19**)
0	0	0	0
1/4	0	0	1.4
1/2	0	0	1.5
1	0	1.0	3.4
2	0	1.9	5.1
4	0.8	5.3	15.1
20	4.5	13.2	42.0

Table 3. Deprotection[a] of **20** with 1 mol/L TBAF–THF containing different amines.

Amine[a]	pK_a[c]	Oligo–RNA chain cleavage (%)	TVS adduct with oligo–RNA (%)
Standard condition[b]	—	9.1	17.0
n-Propylamine	10.71	1.6	8.1
Piperidine	11.12	8.3	0
Morpholine	8.33	0	7.2

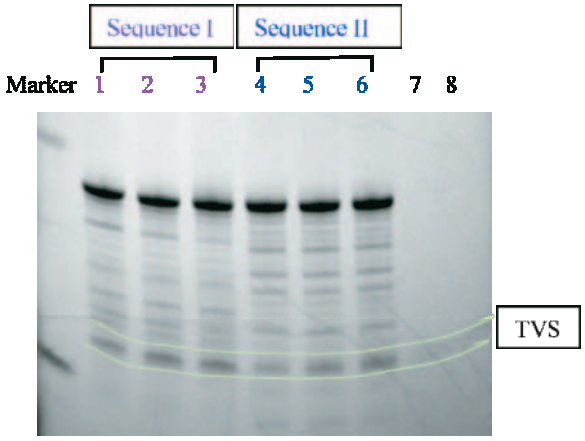
[a]Deprotection reagents: a solution of 10% (*v*/*v*) of each amine in 1 mol/L TBAF–THF, RT, 24 h.
[b]A mixture of 10% *n*-propylamine and 1% bis(2-mercaptoethyl)ether in 1 mol/L TBAF–THF (8).
[c]The pK_a values are taken from Perrin (30).

Morpholine, benefiting form its lower pK_a (8.4), does not cause any chain cleavage but allows the addition of TVS to oligo–RNA to some extent (7.2%) in 1 mol/L TBAF–THF solution. On the other hand, piperidine can completely quench the reaction between TVS and oligo–RNA, but it leads to a small amount (8.3%) of chain cleavage (Table 3).

To examine the potential of morpholine and piperidine as candidates as scavengers of the undesirable byproduct TVS, we have deprotected two 21-mer RNAs using 1 mol/L TBAF–THF, containing either a mixture of *n*-propylamine and bis(2-mercaptoethyl)ether (8), piperidine, or morpholine. Thus, two 21-mer RNAs were synthesized on a solid support with the 2′-*O*-TEM group using our published procedure (8) (sequence I: 5′-GAC GUA AAC GGC CUC AAG UUC; sequence II: 5′-ACU UGU GGC CGU UUA CGU CGC). After treatment with 25% NH_3–MeOH for 20 h at RT and then for 4 h at 40 °C, they were divided into three parts and incubated with (*a*) the standard deprotecting reagent (8) (1 mol/L TBAF–THF containing 10% *n*-propylamine and 1% bis(2-mercaptoethyl) ether), (*b*) 1 mol/L TBAF–THF containing 10% of piperidine, or (*c*) 1 mol/L TBAF–THF containing 10% of morpholine at RT for 20 h. After evaporation of the volatile materials, the products were first passed through a Sep-Pak® Plus (C18, reverse phase) (31) column to remove hydrophilic matter such as excess of F^-, and then NAP^{TM}-10 (Sephadex G25) column (32) to separate the smaller molecules from the larger ones on the basis of size (thus, the full-size oligo–RNA elutes from the truncated sequences and hydrophobic protecting group residues) to give the crude products, which are of high purity and can be used directly for various biological experiments. These crude products were analyzed by 20% denatured PAGE and gel picture obtained by UV–visualization (Fig. 5).

As shown in Fig. 5, all three deprotecting conditions give highly pure crude RNA. Use of morpholine or piperidine,

Fig. 5. 20% denatured PAGE picture of crude RNAs obtained from the deprotection of RNA sequence I (5′-GAC GUA AAC GGC CUC AAG UUC) and RNA sequence II (5′-ACU UGU GGC CGU UUA CGU CGC) with different 2′-*O*-TEM deprotecting conditions. The 2′-*O*-TEM group was deprotected with 1 mol/L TBAF–THF containing different scavengers at RT for 20 h. Lanes 1 and 4: 10% *n*-propylamine and 1% bis(2-mercaptoethyl)ether. Lanes 2 and 5: 10% piperidine. Lanes 3 and 6: 10% morphholine. Marker, upper band: xylene cyanol blue; bottom band: bromophenol blue. Gel picture obtained by UV–vis (254 nm) showing all UV absorbing compounds both nucleic acids (full product and n-1 failure sequences) as well as the cleaved protecting groups. Note that the noncharged TVS (**1**) (λ_{max} = 233 nm) also runs on the gel and is visible by 254 nm light shadowing if in high concentration (Fig. S8).[3] The bands between dashed lines correspond to TVS. Lanes 7 and 8 were run together as a control. Lane 7 product obtained by incubating compound TVS with 1 mol/L TBAF–TMF. Lane 8 product obtained by incubating compound TVS with 1 mol/L TBAF–TMF containing 10% morpholine.



however, gives only a very minor improvement of RNA purity (~84%) compared with that of the standard condition (8) (~82%) [a mixture of *n*-propylamine and bis(2-mercaptoethyl)ether], which is in contrast to the result obtained from the deprotection of the tetrameric RNA (**20**) (see Table 3) with the above three reagents.

Conclusion

In this study, we found that N^6 of adenosine, N^4 of cytidine, and N^2 of guanosine, just as N^3 of thymidine (12), can be alkylated by the α,β-unsaturated compound TVS in 1 mol/L TBAF–THF solution through Michael addition. In oligo–RNA, their nucleophilicities toward TVS are different: N^6 – A > N^4 – C > N^2 – G >> N^3 – U. Hence, by controlling the concentration of TVS, selective alkylation of these nucleophilic sites can be easily achieved. This discovery not only provides a new approach for the alkylation of non-hydrogen-bonded exocyclic amino groups of nucleobases in RNA, but also provides a new strategy for the potential application of the reagent TVS as a chemical probe for detecting the structure and (or) the function of nucleic acid. To avoid adduct formation and chain cleavage during deprotection of 2′-*O*-TEM group with 1 mol/L TBAF–THF, both morpholine and piperidine have been proved to be good scavengers of TVS, which can give high quality crude RNA, bypassing the practical problems of employing the mixture of *n*-propylamine and bis(2-mercaptoethyl)ether (8).

Experimental section

General experimental methods

Chromatographic separations were performed on Merck G60 silica gel. Thin layer chromatography (TLC) was performed on Merck pre-coated silica gel 60 F_{254} glass-backed plates. ^{1}H NMR spectra were recorded at 270.1 MHz and 500 MHz, respectively. ^{13}C NMR spectra were recorded at 67.9 MHz, 125.7 MHz, and 150.9 MHz, respectively. Chemical shifts are reported in ppm (δ scale). Reversed-phase (RP) HPLC: Kromasil 100, C18, 5μ, 100*4.6 mm. Flow 1 mL/min at RT, UV detector with detecting wavelength of 260 nm. Buffer A: 0.1 mol/L TEAA in H_2O–CH_3CN (95:5). Buffer B: 0.1 mol/L TEAA in H_2O–CH_3CN (50:50). Polyacrylamide gel electrophoresis (PAGE): 20% acrylamide (acrylamide/bisacrylamide 29:1), 7 mol/L urea, TBE buffer.

Oligoribonucleotide synthesis

Oligo–RNAs **2–5** and **20** were synthesized through our published procedure (8). After the completion of the automated solid-phase synthesis, the solid supports were treated with 25% NH_3–MeOH at RT overnight, filtered, and purified by HPLC. Oligo–RNA **14** was synthesized by a TBDMS approach according to reported procedure (33) and purified by HPLC. Their structural integrity were verified by MALDI-TOF MS: (**2**) calcd. 974.92 $[MH]^+$, found 975.29; (**3**) calcd. 1013.94 $[MH]^+$, found 1014.24; (**4**) calcd. 997.95 $[MH]^+$, found 998.25. (**5**) cal. 973.94 $[MH]^+$, found 974.29; (**20**) calcd. 2071.85 $[MH]^+$, found 2070.15; (**14**) calcd. 1224.21 $[MH]^+$, found 1224.19.

The 21-mer RNAs, sequence I and sequence II, were synthesized according to our published procedure (8). After completion of the synthesis, the solid supports were treated with 25% NH_3–MeOH at RT for 20 h and then at 40 °C for 4 h. After filtration, each sample was separated into three parts and the volatile materials were evaporated. The residues were co-evaporated with dry methanol twice and dried in vacuum. Then they were treated with 1 mol/L TBAF–THF (dried over 4 Å molecular sieves overnight) containing different scavengers at RT for 20 h. After evaporation of the volatile matter, the residues were diluted with double-distilled water and applied to a NAP-10 column (Amersham Pharmacia Biotech, Sweden). Loading volume is 1 mL and elution was done with 1.2 mL of double-distilled water. The collected fractions were subjected to Sep-Pak® Plus C18 cartridges (Waters) according to manufacturer instructions, to give the crude products. The crude RNAs were subjected to 20% denatured PAGE analysis without any purification step.

General procedure of reaction between ribonucleosides with *p*-tolyl vinyl sulphone (1)

Ribonucleoside (1 mmol) was dissolved in dry 1 mol/L TBAF–THF (5ml) and *p*-tolyl vinyl sulphone **1** (3 mmol for adenosine, cytidine, uridine and 1 mmol for guaonsine) was added. The reaction was stirred at RT for 24 h and analyzed by HPLC. Then the reaction mixture was evaporated to dryness. The residue was diluted with dichloromethane (DCM) and washed with satd. $NaHCO_3$ and dried over $MgSO_4$. The major product was obtained by short column chromatography.

N^4-(4-tolylsulfonylethyl) cytidine (10) ($C^{N4\text{-}TSE}$)

HPLC analysis of the reaction mixture showed that 38.6% of the cytidine was transformed into product **10** and 38.3% of the cytidine was unreacted. CC (DCM, MeOH from 2% to 10%). ^{1}H NMR (270 MHz, $CDCl_3$ + CD_3OD) δ: 2.45 (s, 3H, Ph*Me*), 3.46 (t, *J* = 5.8 Hz, 2H, SO_2CH_2), 3.75–3.96 (m, 4H, $SO_2CH_2CH_2$, 5′,5″-*H*), 4.09 (broad, 1H, 4′-*H*), 4.18 (broad, 2H, 2′3′-*H*), 5.69 (d, *J* = 1.5 Hz, 1H, 1′-*H*), 5.76 (d, *J* = 7.4 Hz, 1H, 5-*H*), 7.37 (d, *J* = 7.7 Hz, 2H, ph*H*), 7.77 (d, *J* = 8.3 Hz, 2H, ph*H*), 7.87 (d, *J* = 7.4 Hz, 1H, 6-*H*). ^{13}C NMR (67.9 MHz, $CDCl_3$ + CD_3OD) δ: 23.94 (Ph*Me*), 34.69 ($SO_2CH_2CH_2$) 54.79 ($SO_2CH_2CH_2$), 60.88 (5′-C), 69.31 (3′-C), 74.94 (2′-C), 84.88 (4′-C), 92.53 (1′-C), 95.81 (5-C), 127.98 (ph), 130.24 (ph), 136.10 (ph), 141.11 (C6), 145.47 (ph), 157.21 (2-C), 164,05 (C4). MALDI-TOF MS: calcd.: 426.13 $[MH]^+$; found: 426.17.

N^6-(4-tolylsulfonylethyl) adenosine (11) ($A^{N6\text{-}TSE}$)

HPLC analysis of the reaction mixture showed that 35.6% of the adenosine was transformed into product **11** and 56.8% of the adenosine was unreacted. CC (DCM, Acetone from 20% to 50%). ^{1}H NMR (270 MHz, $CDCl_3$ + CD_3OD) δ: 2.36 (s, 3H, Ph*Me*), 3.50 (t, 2H, SO_2CH_2), 3.69 (t, 1H, 5′-*H*), 3.84–4.01 (m, 3H, 5″-*H*, $SO_2CH_2CH_2$), 4.26 (broad, 1H, 4′-*H*), 4.44 (broad, 1H, 3′-*H*), 5.50 (broad, 1H, 2′-O*H*), 5.84 (d, *J* = 6.7 Hz, 1H, 1′-H), 6.62 (broad, 1H, 5′-O*H*), 6.97 (t, *J* = 6.1 Hz, 1H, N^6H), 7.25 (d, *J* = 7.9 Hz, 2H, ph*H*), 7.74 (d, *J* = 8.2 Hz, 2H, ph*H*), 7.90 (s, 1H, 8-*H*), 8.03 (s, 1H, 2-*H*). ^{13}C NMR (67.9 MHz, $CDCl_3$ + CD_3OD) δ: 21.60 (Ph*Me*), 34.65 ($SO_2CH_2CH_2$) 54.91 ($SO_2CH_2CH_2$), 62.99 (5′-C), 72.37 (3′-C), 74.01 (2′-C), 87.45 (4′-C), 90.86 (1′-

C), 120.46 (5-C), 127.90 (ph), 130.00 (ph), 135.95 (ph), 140.67 (8-C), 145.07 (ph), 147.49 (4-C), 152.06 (2-C), 154.08 (6-C). MALDI-TOF MS: calcd.: 450.14 $[MH]^+$; found: 450.19.

N^3-(4-tolylsulfonylethyl) uridine (12) ($U^{N3\text{-}TSE}$)

HPLC analysis of the reaction mixture showed that 57.2% of the uridine was transformed to product **12** and 34.1% of the uridine was unreacted. CC (DCM, MeOH from 5% to 7%). ^{1}H NMR (270 MHz, $CDCl_3$ + CD_3OD) δ: 2.39 (s, 3H, Ph*Me*), 3.42 (t, *J* = 7.0 Hz, 2H, SO_2CH_2), 3.76 (dd, *J* = 12.2 Hz, 2H, 5′,5″-*H*), 3.99 (d, *J* = 1.5 Hz, 1H, 4′-*H*), 4.10–4.21 (m, 4H, 2′, 3′, -*H*, $SO_2CH_2CH_2$), 5.65 (d, *J* = 8.1 Hz, 1H, 5-*H*), 5.72 (d, *J* = 2.3 Hz, 1H, 1-*H*), 7.32 (d, *J* = 7.8 Hz, 2H, ph*H*), 7.74 (d, *J* = 8.2 Hz, 2H, ph*H*), 7.85 (d, *J* = 8.1 Hz, 1H, 6-*H*). ^{13}C NMR (67.9 MHz, $CDCl_3$ + CD_3OD) δ: 23.88 (Ph*Me*), 34.95 ($SO_2CH_2CH_2$) 52.56 ($SO_2CH_2CH_2$), 60.95 (5′-C), 69.41 (3′-C), 74.73 (2′-C), 85.16 (4′-C), 91.13 (1′-C), 101.40 (5-C), 128.04 (ph), 130.11 (ph), 135.67 (ph), 139.81 (6-C), 145.36 (ph), 150.92 (2-C), 162.60 (4-C). MALDI-TOF MS: calcd.: 427.11 $[MH]^+$; found: 427.15.

N^2-(4-tolylsulfonylethyl) guanosine (13) ($G^{N2\text{-}TSE}$)

HPLC analysis of the reaction mixture showed that 50.3% of the guanosine was transformed to product **13** and 23.8% of the guanosine was unreacted. The product **13** was too polar to purify by silica gel column chromatography, hence preparative HPLC was used to purify a small amount of product **13** for NMR and MS analysis. ^{1}H NMR (270 MHz, DMSO-d_6) δ: 2.38 (s, 3H, Ph*Me*), 3.51–3.60 (m, 6H, SO_2CH_2, $SO_2CH_2CH_2$, 5′,5″-*H*), 3.99 (t, *J* = 1.5 Hz, 1H, 4′-*H*), 4.11 (t, 1H, 3′-*H*), 4.47 (t, *J* = 5.3 Hz, 1H, 2′-*H*), 5.68 (d, *J* = 5.8 Hz, 1H, 1′-*H*), 6.88 (br., 1H, N^2H), 7.42 (d, *J* = 7.9 Hz, 2H, ph*H*), 7.77 (d, *J* = 8.2 Hz, 2H, ph*H*), 7.93 (s, 1H, 8-*H*), 10.94 (br., 1H, N^1H). ^{13}C NMR (67.9 MHz, DMSO-d_6) δ: 21.64 (Ph*Me*), 34.95 ($SO_2CH_2CH_2$) 54.11 ($SO_2CH_2CH_2$), 62.11 (5′-C), 70.99 (3′-C), 74.27 (2′-C), 85.87 (4′-C), 87.21 (1′-C), 117.70 (5-C), 128.17 (ph), 130.43 (ph), 136.46 (8-C), 136.70 (ph), 144.96 (ph), 151.06 (4-C), 152.89 (2-C), 157.67 (6-C). MALDI-TOF MS: calcd.: 466.23 $[MH]^+$; found: 466.13.

Base composition analysis

Shrimp alkaline phosphatase and phosphodiesterase I (*Crotalus adamanteus* venom) were purchased from Amersham Pharmacia Biotech, Sweden. A reaction mixture containing 20 mmol/L Tris-HCl (pH 8.0), 10 mmol/L $MgCl_2$, shrimp alkaline phosphatase (0.5 unit), and phosphodiesterase I (*Crotalus adamanteus* venom) (0.4 unit) was added to oligoribonucleotide (0.1 OD unit at 260 nm) with a total reaction volume of 30 µL. The reaction mixture was incubated at 37 °C for 12 h and subjected to HPLC analysis directly.

Acknowledgements

Generous financial support from the Swedish Natural Science Research Council (Vetenskapsr et), the Swedish Foundation for Strategic Research (Stiftelsenf Strategisk Forskning), and the EU-FP6 funded RIGHT project (Project no. LSHB-CT-2004–005276) is gratefully acknowledged.

References

1. A. Fire, S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello, Nature (London), **391**, 808 (1998).
2. S.M. Elbashir, J. Harborth, W. Lendechel, A. Yalcin, K. Weber, and T. Tuschl. Nature (London), **411**, 494 (2001).
3. C.B. Reese. Org. Biomol. Chem. **3**, 3851 (2005).
4. S. Pitsch, P.A. Weiss, L. Jenny, A. Stutz, X. Wu, and H.C. Acta. Helv. Chim. Acta, **84**, 3773 (2001).
5. S.A. Caringe, F.E. Wincott, and M.H. Caruthers. J. Am. Chem. Soc. **120**, 11820 (1998).
6. A. Semenyuk, A. Foldesi, T. Johansson, C. Estmer-Nilsson, P. Blomgren, M. Brannvall, L.A. Kirsebom, and M. Kwiatkowski. J. Am. Chem. Soc. **128**, 12356 (2006).
7. T. Ohgi, Y. Masutomi, K. Ishiyama, H. Kitagawa, Y. Shiba, and J. Yano. Org. Lett. **7**, 3477 (2005).
8. C.Z. Zhou, D. Honcharenko, and J. Chattopadhyaya. Org. Biomol. Chem. **5**, 333 (2007).
9. J.J. Solomon, I.L. Cote, M. Wortman, K. Decker, and A. Segal. Chem. Biol. Interactions, **51**, 167 (1984).
10. J.W. Jones and R.K. Robins. J. Am. Chem. Soc. **85**, 193 (1963).
11. P.D. Lawley and P. Brookes. Biochem. J. 127 (1963).
12. M. Mag and J.W. Engels. Nucleic Acids Res. **16**, 3525 (1988).
13. P. Brookes, and P.D. Lawley. J. Chem. Soc. 1348 (1962).
14. R.L.C. Brimacombe, B.E. Griffin, J.A. Haines, W.J. Haslam, and C.B. Reese. Biochemistry, **4**, 2452 (1965).
15. K.K. Ogilvie and S.L. Beaucage. Nucleic Acids Res. **7**, 805 (1979).
16. T. Umemoto and T. Wada. Tetrahedron Lett. **46**, 4521 (2005).
17. D.C. Capaldi, H. Gaus, A.H. Krotz, J. Arnold, R.L. Carty, M.N. Moore, A.N. Scozzari, K. Lowery, D.L. Cole, and V.T. Ravikumar. Org. Process Res. Dev. 832 (2003).
18. B. J. Pullman. Chem. Soc. B, 1621 (1959).
19. J. Ofengand. Methods Enzymol. **20**, 162 (1971).
20. A. Segal, U. Mate, and J.J. Solomon. Chem. Biol. Interactions, **28**, 333 (1979).
21. J.D. Engel. Biochem. Biophys. Res. Comm. **64**, 581 (1975).
22. R. Micura, W. Pils, C. Höbartner, K. Grubmayr, M.O. Ebert, and B. Jaun. Nucleic Acids Res. **29**, 3997 (2001).
23. H.M. Kong, L.F. Lin, N. Porter, S. Stickel, D.P.J. Byrd, and R.J. Roberts. Nucleic Acids Res. **28**, 3216 (2000).
24. O. Gorchs, M. Mernandez, L. Garriga, E. Pedroso, A. Grandas, and J. Farras. Org. Lett. **4**, 1827 (2002).
25. C. Höbartner, C. Kreutz, E. Flecker, E. Ottenschlager, W. Pils, K. Grubmayr, and R. Micura. Monatsh. Chem. **134**, 851 (2003).
26. M. Beier, and W. Pfleiderer. Helv. Chim. Acta, **86**, 2533 (2003).
27. S.T. McDowell, C.J.M. Stirling. J. Chem. Soc. B, 343 (1967).
28. S.T. McDowell and C.J.M. Stirling. J. Chem. Soc. B, 348 (1967).
29. S.T. McDowell and C.J.M. Stirling. J. Chem. Soc. B, 351 (1967).
30. D.D. Perrin. Dissociation constants of organic bases in aqueous solution. Supplement. Butterworths, London, UK. 1972.
31. F. Wincott, A. DiRenzo, C. Shaffer, S. Grimm, D. Tracz, C. Workman, D. Sweedler, C. Gonzalez, S. Scaringe, and N. Usman. Nucleic Acids Res. **23**, 2677 (1995).
32. Handbook of gel filtration, principles and methods. *Edited by* Amersham Biosciences.
33. J. Milecki, E. Zamaratski, T.V. Maltseva, A. Földesi, R. Adamiak, and J. Chattopadhyaya. Tetrahedron Lett. **55**, 6603 (1999).