Pergamon

Bioorganic & Medicinal Chemistry Letters 10 (2000) 1795–1798

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Synthesis and Properties of 2′-*O*-Methyl-2-thiouridine and Oligoribonucleotides Containing 2′-*O*-Methyl-2-thiouridine

Koh-ichiroh Shohda, Itaru Okamoto, Takeshi Wada, Kohji Seio and Mitsuo Sekine*

Department of Life Science, Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuta, Midoriku, Yokohama 226-8501, Japan

Received 19 April 2000; accepted 7 June 2000

Abstract—A new method for the synthesis of 2′-*O*-methyl-2-thiouridine (s^2Um) found in thermophilic bacterial tRNA was developed. Structural properties of s^2Um and s^2Um_pU were studied by using ^{1}H NMR spectroscopy. A modified nonaribonucleotide (RNA*: 5′-CGUUs^2UmUUGC-3′) was synthesized to study the base-recognition ability of s^2Um in formation of RNA–RNA and RNA–DNA duplexes. The UV melting experiments revealed that RNA*–RNA and RNA*–DNA duplexes having an s^2U-A base pair are more stable than those having a U-A base pair. On the contrary, the thermal stability of RNA*–RNA and RNA*–DNA duplexes having an s^2U-G wobble base pair was much lower than that of the unmodified duplexes having a natural U-G base pair. It is concluded that s^2Um has higher selectivity toward A over G than unmodified U.

Introduction

2′-*O*-Methyl-2-thiouridine is a doubly modified nucleoside discovered from tRNAs of extremely thermophilic archaebacteria (*Sulfolobus solfataricus*, *Thermoproteus neutrophilus*, and *Pyrodictium occultum*) by Edmonds et al.[1] Both the ribose residues of 2-thiouridine (s^2U)[2] and 2′-*O*-methyluridine (Um)[3] have a C3′-*endo* conformation, since the 2-thiocarbonyl and 2′-*O*-methyl groups are more sterically hindered than the 2-carbonyl and 2′-hydroxyl groups, respectively.[4] Therefore, the ribose puckering of s^2Um is predicted to have preferentially the C3′-*endo* form because of the additive effect of the 2-thiolation and 2′-*O*-methylation. The C3′-*endo* predominance of nucleosides is entropically favorable for stable A-type RNA duplex formation, because all the ribose conformations of A-type RNA duplexes are C3′-*endo*.[5] It was reported that RNA duplexes containing s^2U or Um were thermodynamically more stable than unmodified natural RNA duplexes.[6,7] Therefore, an oligoribonucleotide bearing s^2Um, which was predicted to have a rigid C3′-*endo* form, was expected to form more stable A-type RNA duplexes. Moreover, the hydrophobic nature of the 2-thiocarbonyl and 2′-*O*-methyl groups may enhance permeability to the cell membrane, and the 2′-*O*-methylated oligoribonucleotides acquire resistance against nucleases.[7b] These properties of s^2Um are expected to be significantly useful for the antisense strategy.[8]

In addition to this expectation, oligonucleotides containing s^2Um are thought to form double helixes more selectively than those containing uridines. The word "selectively" means that the 2-thiocarbonyl group destabilizes an s^2U-G wobble base pair because of the weaker hydrogen bonding nature of the thiocarbonyl group and the steric repulsion derived from the C=S bond that is longer than the C=O bond. Although the U-G wobble base pair is a well-known base mismatch, it does not decrease the *T*m value so much as other base mismatches like the C-A base pair. This is because the U-G pair can form two hydrogen bonds (N^3 of U$\cdots O^6$ of G, and O^2 of U$\cdots N^1$ of G).[9] Quite recently, the thermodynamic parameters (in 1 M NaCl) of RNA duplexes having an s^2U-G wobble base pair were reported by Testa et al.[10] They described that the melting temperatures of RNA duplexes having the s^2U-G wobble base pair were considerably lower than those obtained in the case of the matched s^2U-A base pair.

In order to examine the effects of the 2-thiocarbonyl group on formation of a base pair, we performed ab initio calculation of U-G wobble base pairs (Fig. 1).[11] The hydrogen bonding energy of the base pair formed between 1-methyl-2-thiouracil and 9-methylguanine was −10.4 kcal/mol, while that of the natural U-G base was calculated to be −14.1 kcal/mol.[12] These results apparently suggest the weakened hydrogen bonding property of the 2-thiocarbonyl group as expected above. In addition, the optimized geometry of the 1-methyl-2-thiouracil and 9-methylguanine base pair was simultaneously calculated.

*Corresponding author. Tel.: +81-45-924-5706; fax: +84-45-924-5772; e-mail: msekine@bio.titech.ac.jp

0960-894X/00/$ - see front matter
PII: S0960-894X(00)00342-5

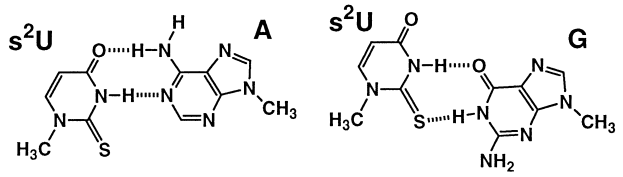


Figure 1. Match (left) and wobble (right) base pairs containing s^2U.

The distance between the two methyl groups of 1-methyl-2-thiouracil and 9-methylguanine was longer by ca. 1 Å than that of the corresponding U-G base pair. This implies that the steric repulsion also contributes to the duplex destabilization by s^2U.

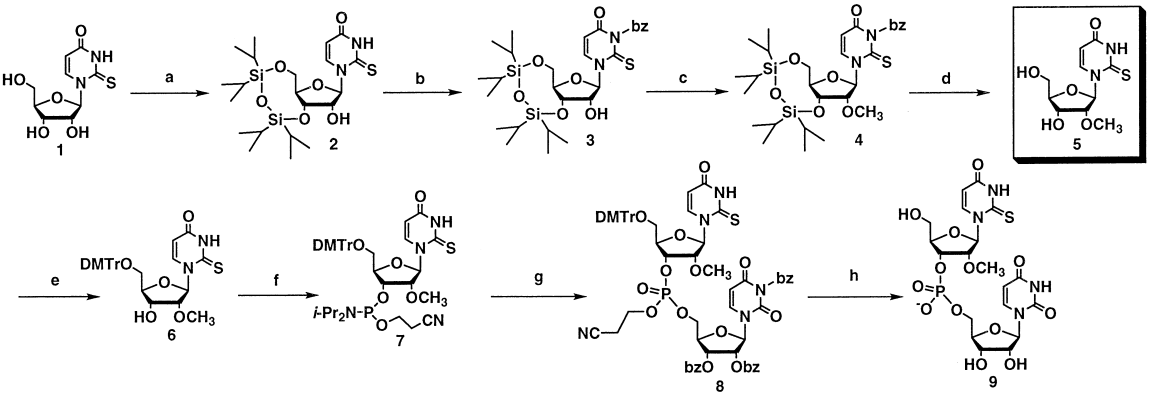
In this study, we developed a new route to s^2Um and a building block for incorporation of s^2Um into oligonucleotides. Furthermore, structural and thermodynamic properties of s^2Um and oligoribonucleotides having s^2Um were studied by means of ^1H NMR spectroscopy and *T*m experiments.

Results and Discussion

Edmonds et al. first reported that s^2Um was obtained by methylation of O^2,5′-anhydrouridine followed by a ring-opening reaction with H_2S.[1] However, no details of the 2′-*O*-regioselectivity of the methylation and the yield at each step were given. It seemed that it was difficult to separate the 2′-*O*-and 3′-*O*-methyl-O^2,5′-anhydrouridine derivatives unless HPLC was used. Therefore, we searched for a more practical route to s^2Um capable of large-scale synthesis of this material required for the antisense strategy. Consequently, all the reactive functional groups of s^2U were protected except for the 2′-hydroxyl group, and the successive 2′-*O*-methylation was performed, as depicted in Scheme 1.

First, the 3′- and 5′-hydroxyl groups of 2-thiouridine[13] (**1**) were protected by use of the 1,1,3,3-tetraisopropyldisiloxane-1,3-diyl (TIPDS) group.[14] Subsequently, the base moiety of the resulting product **2** was acylated via transient protection of the 2′-hydroxyl group with the trimethylsilyl group.[15] The 2′-*O*-methylation of the *N*-benzoylated uridine derivative **3** was examined by using CH_3I–NaH or CH_3I–Ag_2O.[16] When we used the former conditions, 2′-*O*-methylation of **3** was accomplished, but the latter caused complex side reactions because Ag_2O activated simultaneously the 2-thiocarbonyl group of **3**. Finally, both the TIPDS and benzoyl groups of the 2′-*O*-methylated product **4** were removed in the usual manner. Thus, the desired product **5** could be successfully synthesized in an overall yield of 49% from **1**.[17] The 5′-*O*-dimethoxytritylation of **5** followed by 3′-*O*-phosphitylation gave the building block **7** required for incorporation of s^2Um into oligonucleotides.[18] By using this building block, the fully protected dimer **8** bearing s^2Um as a 5′-upstream nucleoside was synthesized (Scheme 1). Since it was reported that the thiocarbonyl group of 2-thiothymidine reacted with I_2,[19] the phosphite intermediate was oxidized with *t*-butyl hydroperoxide in a manner similar to that described by Kumar et al.[6,20] The successive treatments of **8** with ammonium hydroxide and 80% AcOH gave the dinucleoside monophosphate **9** (s^2Um_pU) after reversed-phase HPLC purification. In order to investigate the conformational properties of s^2Um, the ^1H NMR spectra of **5** and **9** were measured in sodium phosphate buffer (pH 7.0). The ratio of the C3′-*endo* and C2′-*endo* conformers of s^2Um was calculated by using the coupling constant between the vicinal protons of **5** and **9**, and these results are listed in Table 1.[21]

The %N (C3′-*endo*) value of **5** was calculated to be 70%, while that of 2-thiouridine (s^2U) reported by Sierzputowska–Gracz was 78%.[2] In spite of the presence of the bulky 2′-*O*-methyl group, the predominance of the C3′-*endo* conformer of **5** was unexpectedly almost equal to that of s^2U. On the other hand, the ribose puckering of s^2Um in the dimer **9** was almost C3′-*endo* only ($^3J_{1'H\text{-}2'H} < 1$ Hz), although the predominance of the C3′-*endo* conformer of the 5′-upstream nucleoside in s^2UpU was at most 90%.[22] These results indicated that the steric repulsion between the 2-thiocarbonyl and 2′-*O*-methyl groups of s^2Um was more effective in the



Scheme 1. Synthesis of s^2Um and s^2Um_pU. Reagents and conditions: (a) $TIPDSCl_2$ in pyridine, 99%; (b) i. HMDS in CH_3CN, ii. BzCl, Bu_4NBr in CH_2Cl_2/Na_2CO_3 aq, iii. TFA in MeOH-CH_2Cl_2 (1:1, v/v), 76%; (c) NaH, CH_3I in DMF, 79%; (d) i. Bu_4NF, AcOH in THF, ii. Concd NH_3–pyridine (1:9, v/v), 83%; (e) DMTrCl in pyridine, 85%; (f) Chloro(2-cyanoethoxy)(*N*, *N*-diisopropylamino)phosphine, diisopropylethylamine in CH_2Cl_2, 81% (purity 89%); (g) i. N^3,2′,3′-*O*-Tribenzoyluridine, 1*H*-tetrazole in CH_3CN, ii) *t*-BuOOH, 38%; (h) i. Concd NH_3–pyridine (9:1, v/v), ii. 80% AcOH, 67%.

Table 1. Vicinal coupling constants ($^3J_{1'H\text{-}2'H}$) and C3′-*endo* preference of nucleosides

				s²UmpU		s²UpU		UpU	
	s²Um	s²U	U	5′-Nu	Nu-3′	5′-Nu	Nu-3′	5′-Nu	Nu-3′
$^3J_{1'H\text{-}2'H}$ (Hz)	3.1	2.5[a]	4.6[a]	⩽1.0	2.8	1.6[b]	3.5[b]	4.4[b]	3.7[b]
%N (C3′-*endo*)[c]	70	78	48	100	74	91	64	51	61

[a]These coupling constants were reported in ref. 2.
[b]These data were taken from ref. 22.
[c]These values were calculated according to the equation: %N (C3′-*endo*) = {(7.9−$^3J_{1'H\text{-}2'H}$ (Hz))/6.9}×100.[21]

presence of the 3′-phosphate group because of its steric hindrance. Therefore, it was expected that the ribose puckering of s²Um in oligonucleotides would also be C3′-*endo*. Interestingly, the C3′-*endo* preference of s²Um influenced the sugar conformation of the 3′-downstream nucleoside, which became C3′-*endo*. The %N (C3′-*endo*) value of the 3′-nucleoside in the s²Um$_p$U calculated was 74%, i.e., greater by 10% than that of s²U$_p$U which showed 64% C3′-*endo* predominance.

Since the C3′-*endo* preference of s²Um was clearly confirmed, the stability of duplexes containing s²Um was evaluated by the melting temperature (*T*m) measurement. An oligoribonucleotide nonamer (5′-CGUUs²UmUU GC-3′) was synthesized by the usual phosphoramidite method, except for the oxidation, similar to that described in the synthesis of dimer **8**. Subsequent deprotection was performed by the routine procedure, and the resulting mixture was separated by anion exchange HPLC.[23] For the control experiments, an oligoribonucleotide having s²U instead of s²Um was obtained by Kumar's method.[6]

UV melting curves of RNA–RNA and RNA–DNA duplexes were measured in 10 mM sodium phosphate buffer (pH 7.0) containing 150 mM NaCl, and these profiles and *T*m values are shown in Fig. 2 and Table 2, respectively. In the case of RNA–RNA duplexes, the *T*m value of the duplex containing either s²Um or s²U was higher than that of the wild-type matched duplex (ΔTm = +5.5 °C (s²Um), and +6.9 °C (s²U)). Such a tendency was also found in the case of an RNA–DNA matched duplex, but the ΔTm value became larger than those of the RNA–RNA series (ΔTm = +6.9 °C (s²Um), and +9.4 °C (s²U)). Although oligonucleotides modified with s²U or Um form thermodynamically stable A-type RNA duplexes,[6,7,10] the *T*m values of both RNA–RNA and RNA–DNA duplexes modified by s²Um were somewhat lower than those of the duplexes modified by s²U, and no additive effect of the double modification was observed.

The selectivity toward the target sequence is also an important property of antisense agents. Particularly, the discrimination of the U-G wobble base pair from the U-A Watson–Crick base pair is essential for a highly specific antisense method. Therefore, the thermal stability of RNA–RNA duplexes containing an s²Um-G wobble pair was examined.

In order to evaluate the selectivity toward the target sequence, difference in the *T*m value between the duplex having a s²Um-G and the duplex having a s²Um-A base pair (= ΔTm*, see Table 2) was calculated. These ΔTm* values were found to be −15.7 and −14.7 °C, for s²Um and s²U, respectively. However, the ΔTm* value between the unmodified duplexes was only −3.9 °C. Testa et al. also reported a similar trend for U-G and s²U-G, but the absolute values of ΔTm* are smaller than those in this study.[10] Similar experiments were performed in a series of RNA–DNA duplexes, and the ΔTm* values were determined to be −20.4, −18.5, and −11.2 °C for the duplexes having s²Um, s²U, and U, respectively. In both the RNA–RNA and RNA–DNA

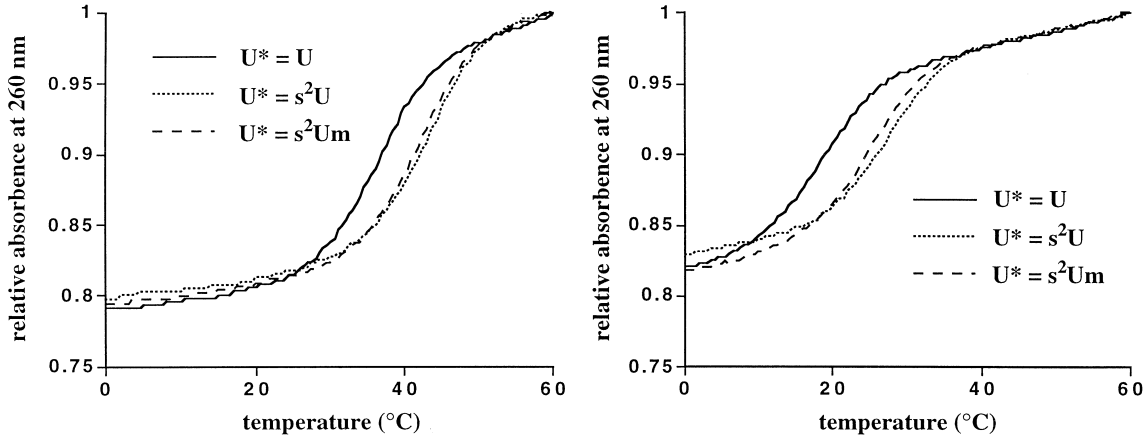


Figure 2. Melting curves of CGUUU*UUGC/GCAAAAACG (RNA–RNA series, left) and CGUUU*UUGC/d(GCAAAAACG) (RNA–DNA series, right).

Table 2. Melting temperature (*T*m Value, °C) of RNA–RNA and RNA–DNA duplexes containing an s^2Um-A (match) or s^2Um-G (wobble) base pair

	GCAAXAACG		d(GCAAXAACG)	
CGUUU*UUGC	X = A (match *T*m value (ΔTm*)[a]	X = G (wobble) *T*m value (ΔTm*)[b]	X = A (match) *T*m Value (ΔTm)[a]	X = G (wobble) *T*m value (ΔTm*)[b]
U* = U (control)	36.7	32.8 (−3.9)	18.6	7.4 (−11.2)
U* = s^2Um	42.2 (+5.5)	26.5 (−15.7)	25.5 (+6.9)	5.1 (−20.4)
U* = s^2U	43.6 (+6.9)	28.9 (−14.7)	28.0 (+9.4)	9.5 (−18.5)

[a]ΔTm = (*T*m value of a modified duplex) − (that of a wild-type duplex).
[b]ΔTm* = (*T*m value of a duplex having a U*-G wobble base pair) − (that of a duplex having a matched U*-A base pair in place of a U*-G base pair).

duplexes having s^2U the ΔTm* values of the ones modified by s^2U derivatives were always larger than those of the unmodified, and the ΔTm* value of s^2Um-containing duplexes was fortunately somewhat larger than those having s^2U.

The duplex destabilization by the s^2Um (or s^2U)-G base pair was a consequence of both the weaker hydrogen bonding nature and the steric hindrance of the 2-thiocarbonyl group, as expected by ab initio calculation. However, the reason why the ΔTm* values of the duplexes containing s^2Um were larger than those of the duplex containing s^2U is not clear. Since the effect of the uridine-2-thiolation on the stabilization and selectivity enhancement of RNA duplexes is excellent, we propose that several or all uracil residues of the antisense agents should be replaced by 2-thiouracil, if such replacement is chemically and economically possible.[24] Therefore, antisense agents having 2-thiouracil residues instead of uracil will be one of the second-generation antisense drugs. Further studies are under way in this direction.

Acknowledgements

This work was supported by a Grant from ''Research for the Future'' Program of the Japan Society for the Promotion of Science (JSPS-RFTF97I00301) and a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan.

References and Notes

1. Edmonds, C. G.; Crain, P. F.; Hashizume, T.; Gupta, R.; Stetter, K. O.; McCloskey, J. A. *J. Chem. Soc., Chem. Commun*. **1987**, 909.
2. Sierzputowska-Gracz, H.; Sochacka, E.; Malkiewicz, A.; Kuo, K.; Gehrke, C. W.; Agris, P. F. *J. Am. Chem. Soc*. **1987**, *109*, 7171.
3. Davies, D. B. In *Progress in NMR Spectroscopy*; Emsley, J. W.; Feenly, J.; Sutcliffe, L. H., Eds.; Pergamon: New York, 1978; pp 135–225.
4. Sakamoto, K.; Kawai, G.; Watanabe, S.; Niimi, T.; Hayashi, N.; Muto, Y.; Watanabe, K.; Satoh, T.; Sekine, M.; Yokoyama, S. *Biochemisty* **1996**, *35*, 6533.
5. Seanger, W. In *Principles of Nucleic Acid Structure* Springer-Verlag: Berlin, 1981; Chapter 10.
6. (a) Kumar, R. K.; Davis, D. R. *Nucleic Acids Res*. **1997**, *25*, 1272. (b) Idem. *Nucleosides Nucleotides* **1997**, *16*, 1469.
7. (a) Lesnik, E. A.; Guinosso, C. J.; Kawasaki, A. M.; Sasmor, H.; Zounes, M.; Cummins, L. L.; Ecker, D. J.; Cook, P. D.; Freier, S. M. *Biochemistry* **1993**, *32*, 7832. (b) Cummins, L. L.; Owen, S. R.; Risen, L. M.; Lesnik, E. A.; Freier, S. M.; McGee, D.; Guinosso, C. J.; Cook, P. D. *Nucleic Acids Res*. **1995**, *23*, 2019.
8. (a) Crooke, S. T., Lebleu, B., Eds., *Antisense Research and Applications*; CRC Press: Boca Raton, 1993. (b) Crooke, S. T., Ed.; *Antisense Research and Application*; Springer-Verlag: Berlin, 1998.
9. Sugimoto, N.; Kierzek, R.; Freier, S. M.; Turner, D. H. *Biochemistry* **1986**, *25*, 5755.
10. Testa, S. M.; Disney, M. D.; Turner, D. H.; Kierzek, R. *Biochemistry* **1999**, *38*, 16655.
11. (a) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn*. **1996**, *14*, 117. (b) Hobza, P.; Sponer, J. *Chem. Rev*. **1999**, *99*, 3247.
12. The hydrogen bonding energy was estimated as the difference between the energy of complex (base pair) and energies of isolated nucleobases. The geometries and energies of all the systems were calculated at the HF/6-31G* level.
13. (a) Vorbrüggen, H.; Strehlke, P. *Chem. Ber*. **1973**, *106*, 3039. (b) Niedballa, U.; Vorbrüggen, H. *J. Org. Chem*., **1974**, *39*, 3654.
14. Markiewics, W. T. *J. Chem. Research (S)* **1979**, 24.; *J. Chem. Research (M)* **1979**, 0181.
15. Sekine, M. *J. Org. Chem*. **1989**, *54*, 2321.
16. (a) Kamimura, T.; Masegi, T.; Hata, T. *Chem. Lett*. **1982**, 965. (b) Inoue, H.; Hayase, Y.; Imura, A.; Iwai, S.; Miura, K., Ohtsuka, E. *Nucleic Acids Res*. **1987**, 6131
17. Spectral data of **5**: ^{1}H NMR (400 MHz, D_2O) δ 3.60 (3H, s, 2′-*O*-CH_3), 3.83 (1H, dd, $J_{4'H\text{-}5''H}$ = 3.8 Hz, $J_{5'H\text{-}5''H}$ = 13.1 Hz, 5′-H), 3.96 (1H, dd, $J_{4'H\text{-}5''H}$ = 2.5 Hz, 5″-H), 4.06 (1H, dd, $J_{1'H\text{-}2'H}$ = 3.1″ Hz, $J_{2'H\text{-}3'H}$ = 5.0″ Hz, 2′-H), 4.12 (1H, m, 4′-H), 4.30 (1H, dd, $J_{3'H\text{-}4'H}$ = 6.9 Hz, 3′-H), 6.15 (1H, d, $J_{5H\text{-}6H}$ = 8.1 Hz, 5-H), 6.76 (1H, d, 1′-H), 8.08 (1H, d, 6-H).
18. (a) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett*. **1981**, *22*, 1859. (b) Sinha, N. D.; Biernat, J.; Koster, H. *Tetrahedron Lett*. **1983**, *24*, 5843.
19. Kuimelis, R. G.; Nambiar, K. P. *Tetrahedron Lett*. **1993**, *34*, 3813.
20. Kumar, R. K.; Davis, D. R. *J. Org. Chem*. **1995**, *60*, 7726.
21. (a) de Leeuw, F. A. A. M.; Altona, C. *J. Chem. Soc., Perkin Trans. 2* **1982**, 375. (b) Altona, C. *Recl. Trav. Chim. Pays-Bas*, **1982**, *101*, 413. (c) Rousse, B.; Puri, N.; Viswanadham, P.; Agback, P.; Glemarec, C.; Sandstrom, A; Sund, C.; Chattopadhyaya, J. *Tetrahedron* **1994**, *50*, 1777.
22. Smith, W.; Sierzputowska-Gracz, H.; Sochacka, E.; Malkiewicz, A.; Agris, P. F. *J. Am. Chem. Soc*. **1992**, *114*, 7989.
23. The RNA nonamer (5′-CGUUs^2UmUUGC-3′) was characterized by MALDI-TOF-Mass spectroscopy (calcd. for $C_{84}H_{105}N_{26}O_{65}P_8S$ ($M-H^-$): 2797.33; Found: 2797.20).
24. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11804.