Organic & Biomolecular Chemistry

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2014, **12**, 1052

Received 19th November 2013,
Accepted 17th December 2013

DOI: 10.1039/c3ob42302f

www.rsc.org/obc

Site-selected incorporation of 5-carboxymethylaminomethyl(-2-thio)uridine into RNA sequences by phosphoramidite chemistry†

Grazyna Leszczynska,* Jakub Pięta, Karolina Wozniak and Andrzej Malkiewicz

5-Carboxymethylaminomethyluridine (cmnm^{5}U) and 5-carboxymethylaminomethyl-2-thiouridine (cmnm^{5}s^{2}U) are located at the wobble position in several cytosolic and mitochondrial tRNA sequences. In this paper, we report the first site-selected incorporation of cmnm^{5}U and cmnm^{5}s^{2}U into RNA sequences by phosphoramidite chemistry on a CPG solid support. Trifluoroacetyl and 2-(trimethylsilyl)ethyl were selected for the protection of the amine and carboxyl functions, respectively.

Several C-5 substituted (2-thio)uridines are native components of tRNAs isolated from prokaryotic and eukaryotic cells, including mitochondrial tRNAs (mt-tRNAs).[1] 5-Carboxymethylaminomethyluridine (cmnm^{5}U, **1**, Fig. 1a) and 5-carboxymethylaminomethyl-2-thiouridine (cmnm^{5}s^{2}U, **2**, Fig. 1a) are located at position 34 ("wobble" position, the first anticodon letter) of *Saccharomyces cerevisiae* mt-tRNAs specific for Leu and Lys, respectively (Fig. 1c).[1,2] Interestingly, the structurally related pair of "wobble" (2-thio)uridines modified at C-5 with the methylene–taurine residue, that is, 5-taurinomethyluridine (τm^{5}U, **3**, Fig. 1b) and its 2-thio analogue (τm^{5}s^{2}U, **4**, Fig. 1b), are characteristic of human mt-tRNA (hmt-tRNA) sequences specific for Leu(UUR) and Lys, respectively.[1,3]

Nucleosides **1–4** play a crucial role in the translation processes that restrict mt-tRNA decoding capacity to A and G as the third codon letter.[2,4,5] The presence of C-5 glycine and taurine residues in the structure of nucleosides **1–4** seems to be important for decoding G,[4–6] while the lack of both 5-substituents and the 2-thio group totally stalls the translation machinery.[4–7]

It is noteworthy that the point mutations 3243(A → G) in the hmt-tRNA^{Leu} gene and 8344(A → G) in the hmt-tRNA^{Lys} gene induce the loss of τm^{5}U and τm^{5}s^{2}U, respectively, which results in incurable mitochondrial diseases (MELAS, MERRF).[5]

Recently, yeast cells have been suggested as a useful model for studies of the molecular and cellular effects related to human mitochondrial diseases.[8] Partial correction of the MERRF and MELAS syndromes was observed after targeting import of the "chimeric" yeast tRNA^{Lys} into human

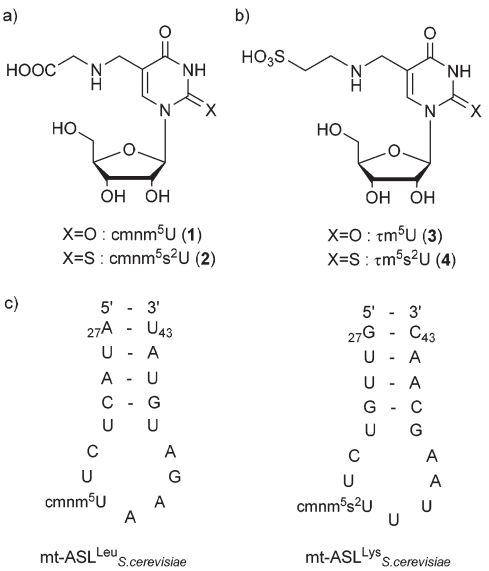


Fig. 1 (a) The structure of 5-carboxymethylaminomethyluridine (cmnm^{5}U, **1**) and 5-carboxymethylaminomethyl-2-thiouridine (cmnm^{5}s^{2}U, **2**). (b) The structure of 5-taurinomethyluridine (τm^{5}U, **3**) and 5-taurinomethyl-2-thiouridine (τm^{5}s^{2}U, **4**). (c) The sequence and secondary structure of the *Saccharomyces cerevisiae* mitochondrial tRNA^{Leu} and tRNA^{Lys} anticodon stem and loop (mt-ASL$^{Leu}_{S.\ cerevisiae}$, mt-ASL$^{Lys}_{S.\ cerevisiae}$) modified with cmnm^{5}U and cmnm^{5}s^{2}U, respectively. The native sequence of mt-ASL$^{Leu}_{S.\ cerevisiae}$ has three additional modified nucleosides: pseudouridines (ψ) at positions 31 and 39 and 1-methylguanosine at position 37. In the native sequence of mt-ASL$^{Lys}_{S.\ cerevisiae}$ there are two pseudouridines (ψ) at positions 28 and 31, and *N*-[(9-β-D-ribofuranosyl-9*H*-purin-6-yl)carbamoyl]-L-threonine (t^{6}A) at position 37.

Institute of Organic Chemistry, Zeromskiego St. 116, Lodz, Poland.
E-mail: grazyna.leszczynska@p.lodz.pl; Fax: +4842 636 55 30; Tel: +4842 631 31 50

† Electronic supplementary information (ESI) available: General information, experimental details, spectral characterisation of **11a**/**11b**, **12a**/**12b**, **13a**/**13b**, oligoribonucleotide synthesis, MALDI–TOF spectra. See DOI: 10.1039/c3ob42302f

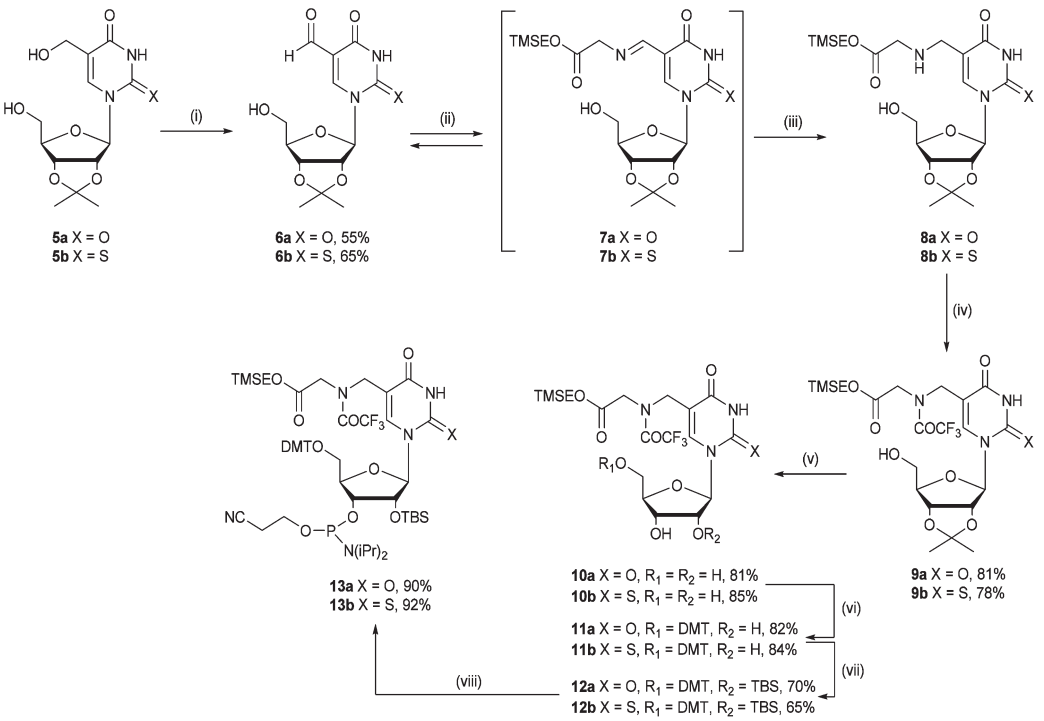
mitochondrial cybrid cells.[9] It was also demonstrated that hmt-tRNA^{Lys,Leu} precursors, equipped with suitable carriers, are effectively imported to the organelle and processed to mature tRNAs, in this way rescuing mitochondrial protein biosynthesis.[10]

Synthetic modified tRNA fragment(s) have become useful tools for probing the structural and mechanistic aspects of biopolymer activities, particularly in model studies on the structural restraints of "wobble" pairing.[11]

Modified oligomers are also used as substrates for enzymatic ligation of RNA fragments to generate longer RNA constructs.[12] The synthesis of cmnm^{5}U and cmnm^{5}s^{2}U, as well as their conformations in solution and the solid state, have been reported.[13] Using the classical triester approach in solution, both modified units were built in anticodon triplets related to tRNA$^{Lys}_{B.\ subtilis}$ and tRNA$^{Gly}_{B.\ subtilis}$.[14] Recently, 5′,3′-*O*-bisphosphate of cmnm^{5}U has been used to incorporate this modification into position 34 of the tRNA$^{Leu(UUR)}$ sequence in *E. coli via* semienzymatic ligation of RNA fragments.[4] In this paper, we present CPG-supported (phosphoramidite chemistry) synthesis of 17-mer long RNAs related to the sequence of the anticodon arm of *S. cerevisiae* mt-tRNA^{Leu} and mt-tRNA^{Lys}.

Synthesis of RNA fragments modified with cmnm^{5}U and cmnm^{5}s^{2}U required several interventions in the standard protocol of phosphoramidite chemistry on a CPG support. In the chemical structure of nucleosides **1** and **2**, there are highly reactive aliphatic amine and carboxyl functions which should be protected in a manner compatible with the protection of the commercially available canonical monomeric units. The aliphatic amine function was protected with a base labile trifluoroacetyl group, following the methodology elaborated previously, which was used for incorporation of 5-methylaminomethyl(-2-thio)uridine (mnm^5(s^2)U) into RNA fragments.[15–18] For the protection of the carboxyl function, removable under neutral conditions, 2-(trimethylsilyl)ethyl (TMSE) was selected. The usefulness of this strategy was previously confirmed during the synthesis of RNAs modified with (2-methylthio)adenosine bearing a N^6-treonylcarbamoyl residue (t^{6}A, ms^{2}t^{6}A).[18,19] It is noteworthy that the TMSE ester is stable under prolonged treatment with ethanolic ammonia and offers a simple and effective way for simultaneous removal of alkaline labile protecting groups and cleavage of oligomers from the CPG support without the risk of amide formation.

The multistep synthesis of the fully protected 5-carboxymethylaminomethyl(-2-thio)uridine 3′-*O*-phosphoramidite **13a**/**13b** is presented in Scheme 1. For the synthesis of TMSE-protected cmnm^5(s^2)U **8a**/**8b** we employed an effective procedure of reductive amination including 5-formyl(-2-thio)uridine **6a**/**6b** and 2-(trimethylsilyl)ethyl glycine ester, in the presence of sodium triacetoxyborohydride (NaBH(OAc)$_3$).[20] For this purpose, 5-hydroxymethyl-2′,3′-*O*-isopropylidene(-2-thio)uridine **5a**/**5b** was selectively oxidized with activated MnO_2 in acetone–DCM solution, giving 5-formyl(-2-thio)uridine **6a**/**6b**, which was used in a reaction with glycine ester in DCM–DMF solution. The imine **7a**/**7b**, without isolation, was reduced with NaBH(OAc)$_3$. The crude secondary amine **8a**/**8b** was protected with a



Scheme 1 Reagents and conditions: (i) activated MnO_2, acetone–DCM (4 : 1, v/v), 2–7 h, 55 °C; (ii) HCl × $NH_2CH_2COOTMSE$, TEA, DCM–DMF, 1.5 h, rt; (iii) NaBH(OAc)$_3$, 2 h, rt; (iv) TFAA, py, 1.5 h, 0 °C; (v) 25% aq. AcOH, 1 h, 90 °C; (vi) DMT-Cl, py, 16 h, rt; (vii) TBS-Cl, imidazole, py, 4 h, rt; (viii) $[(CH_3)_2CH]_2NP(Cl)OCH_2CH_2CN$, $[(CH_3)_2CH]_2NCH_2CH_3$, DCM, 5–7 h, rt.

trifluoroacetyl group.[21] 2′,3′-Acetonide **9a**/**9b** was removed under mild acidic conditions to give the N-protected nucleoside **10a**/**10b**. The synthesis of 5′-*O*-DMT-2′-*O*-TBS-phosphoramidite **13a**/**13b** followed typical procedures.[16] For more details of the procedures for preparation of **11–13** and their characterization, see the ESI.†

Cmnm^5(s^2)U-modified anticodon arm domains (Fig. 1c) were synthesized manually on a 5 μmol scale using commercial tac-protected phosphoramidites of the canonical units and a slightly modified version of Sproat's protocol (ESI†).[22] For improving the overall yield of the oligomers, the couplings were conducted in 8 molar excess of A, U, C, and G amidites for 8 min, while the modified units were used in 12 molar excess and coupled twice, each time using 6 molar excess of an amidite and 12 min coupling time. Coupling yields were in the range of 90–95%.

It is generally known that the 2-thiocarbonyl group of 2-thiouridine derivatives reacts with various oxidizing reagents used in the oligoribonucleotide synthesis giving products of 2-thio → 2-oxo transformation and/or oxidative desulfurization (s^2 → H^2).[17,23,24] The scope and nature of the above-mentioned side-processes strongly depend on C-5 substituents and the oxidant nature.[17,24] Consequently, an effective and universal protocol for the P^{III} → P^V oxidation cycle has not been published so far.

We tested several oxidizing reagents for the synthesis of the cmnm^{5}s^{2}U-modified RNA oligomer using 2-thiouridine **12b** and a previously published methodology.[17,18] We have found that the 2-thiocarbonyl group of cmnm^{5}s^{2}U **12b** remains intact in a 0.02 M iodine (8 equiv.) solution in THF–H_2O–pyridine,[23] in contrast to its easy oxidation with standard oxidizing reagents (0.16 M I_2–water, 1 M *tert*-BuOOH–ACN) or partial loss of s^2 with a diluted solution of *tert*-BuOOH (0.25 M) in anhydrous toluene or acetonitrile.

In the first deprotection step the DMT group was removed from the 5′-terminal residue and CPG-bound RNAs were treated with TEA–ACN (1 : 1 v/v), and then with 8 M ethanolic ammonia. The two-step deprotection procedure made it possible to avoid the reaction of heterobase residues with the acrylonitrile generated during the deprotection of phosphate residues.[25] To remove the TBS and TMSE protecting groups, the tetraalkylammonium salts were used. The 1 M solution of Bu_4NF gave the best result in desilylation of cmnm^{5}U-modified RNA, while 1 M Et_4NF turned out to be the most effective desilylating reagent for cmnm^{5}s^{2}U-modified RNA. The use of TEA × 3HF gave only partial deprotection of the TMSE group in accordance with previous observations.[19] The fully deprotected oligomers were quenched with phosphate buffer, desalted, and then purified by preparative IE-HPLC (Fig. 2a, ESI†). The desalted RNAs were lyophilized to yield 8.5 mg of mt-ASL^{Leu}(cmnm^{5}U$_{34}$) and 5 mg of mt-ASL^{Lys}(cmnm^{5}s^{2}U$_{34}$). The homogeneity and chemical structure of the synthetic oligoribonucleotides were verified by MALDI-TOF data (ESI†) as well as RNA enzymatic digestion (ESI†) to the expected mixture of nucleosides, whose composition was tested by RP HPLC, and the data were compared with those registered

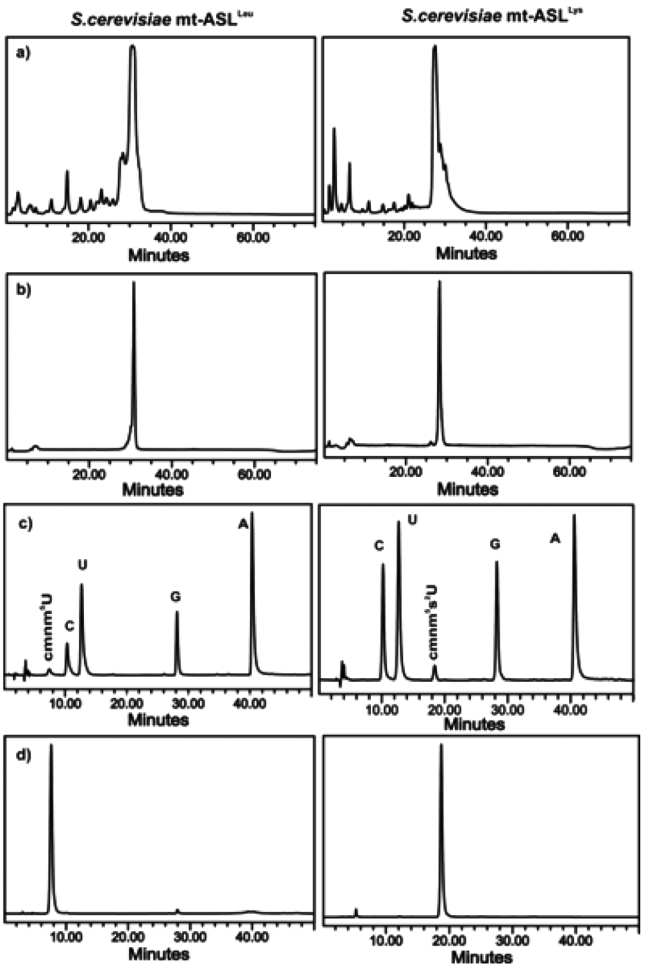


Fig. 2 (a) Anion-exchange HPLC of crude, deprotected mt-ASL$^{Leu}_{S.\ cerevisiae}$-(cmnm^{5}U$_{34}$) and mt-ASL$^{Lys}_{S.\ cerevisiae}$(cmnm^{5}s^{2}U$_{34}$). (b) Analytical injection of desalted and lyophilized oligoribonucleotides. (c) HPLC nucleoside composition of mt-ASLs. (d) References of cmnm^{5}U and cmnm^{5}s^{2}U injected in control experiments.

under identical conditions for modified nucleosides[20] as a reference (Fig. 2c and d). As shown in Fig. 2c, HPLC elutions monitored at 264 nm indicate that mt-ASL^{Leu} and mt-ASL^{Lys} contain cmnm^{5}U (7.59 min) and cmnm^{5}s^{2}U (18.80 min), respectively.

For the synthesis of cmnm^5(s^2)U-modified 17-mers (Fig. 1), 2-(*p*-nitrophenyl)ethyl (NPE)[26] protection of the cmnm^5(s^2)U carboxyl function was also used (experimental data not shown). The RNA synthesis was conducted using the same protocol with the exception of oligomer deprotection (ESI†). The use of NPE-protected cmnm^5(s^2)U amidites caused a significant decrease of the oligomer efficiency in comparison with the incorporation of TMSE-protected analogues (shorter and partially deprotected oligomers were observed in HPLC). The utilization of the NPE-strategy for the synthesis of the cmnm^{5}s^{2}U-modified oligomer resulted in a complete desulphurization. It is likely that the strong basic DBU solution required for NPE deprotection causes the s^2 → o^2 transformation in the cmnm^{5}s^{2}U structure.

Conclusions

The 2-(trimethylsilyl)ethyl group (TMSE) was selected for protection of carboxyl acid residue of glycine and used for the preparation of cmnm^{5}U and cmnm^{5}s^{2}U phosphoramidites (**13a**, **13b**). TMSE blockage is compatible with the protection of commercially available canonical monomeric units as well as t^{6}A, which is a natural modification located at position 37 of yeast mt-ASL^{Lys}. Hypermodified monomers **13a** and **13b** were effectively incorporated into the anticodon arm sequences of *S. cerevisiae* mt-tRNA^{Leu} and mt-tRNA^{Lys}, respectively. The reported procedure enables elimination of the unfavorable desulphurization process in the case of cmnm^{5}s^{2}U incorporation which could significantly decrease the yield and purity of the final product. Both oligomers have been obtained in sufficient purity and quantity for NMR structural studies. Future work will aim at the synthesis of the double modified native sequence of mt-ASL^{Lys}(cmnm^{5}s^{2}U$_{34}$, t^{6}A$_{37}$).

Acknowledgements

This work was financially supported by the National Science Centre of Poland (no. 1306/B/H03/2011/40 for G.L.).

Notes and references

1 A. Czerwoniec, S. Dunin-Horkawicz, E. Purta, K. H. Kaminska, J. M. Kasprzak, J. M. Bujnicki, H. Grosjean and K. Rother, *Nucleic Acids Res.*, 2009, **37**, D118.

2 R. Martin, A. P. Sibler, Ch. Gehrke, K. Kuo, Ch. Edmonds, J. McCloskey and G. Dirheimer, *Biochemistry*, 1990, **29**, 956.

3 T. Suzuki, T. Suzuki, T. Wada, K. Saigo and K. Watanabe, *Nucleic Acids Res. Suppl.*, 2001, **1**, 257.

4 S. Kurata, A. Weixlbaumer, T. Ohtsuki, T. Shimazaki, T. Wada, Y. Kirino, K. Takai, K. Watanabe, V. Ramakrishnan and T. Suzuki, *J. Biol. Chem.*, 2008, **283**, 18801.

5 (*a*) T. Suzuki, A. Nagao and T. Suzuki, *WIREs RNA*, 2011, **2**, 376; (*b*) S. W. Schaffer, Ch. J. Jong, T. Ito and J. Azuma, *Amino Acids*, 2012, DOI: 10.1007/s00726-012-1414-8.

6 A. Rodriguez-Hernandez, J. L. Spears, P. A. Limbach, H. Gamper, Y. M. Hou, R. Kaiser, P. F. Agris and J. J. Perona, *J. Mol. Biol.*, 2013, **425**, 3888.

7 X. Wang, Q. Yan and M-X. Guam, *J. Mol. Biol.*, 2010, **395**, 1038.

8 (*a*) C. De Luca, Y. Zhou, A. Montanari, V. Morea, R. Oliva, C. Besagni, M. Bolotin-Fukuhara, L. Frontali and S. Francisci, *Mitochondrion*, 2007, **9**, 408; (*b*) A. Montanari, C. Besagni, C. De Luca, V. Morea, R. Oliva, A. Tramontono, M. Bolotin-Fukuhara, L. Frontali and S. Francisci, *RNA*, 2012, **14**, 275; (*c*) N. B. V. Sepuri, M. Gorla and M. P. King, *PLoS One*, 2012, **7**, e35321.

9 (*a*) O. A. Kolesnikova, N. S. Entelis, H. Mireau, T. D. Fox, R. P. Martin and I. Tarassov, *Science*, 2000, **289**, 1931; (*b*) O. A. Kolesnikova, N. S. Entelis, C. Jacquin-Becker, F. Goltzene, Z. M. Chrzanowska-Lightowlers, R. P. Martin and I. Tarassov, *Hum. Mol. Genet.*, 2004, **13**, 2519; (*c*) O. Z. Karicheva, O. A. Kolesnikova, T. Schirtz, M. Y. Vysokikh, A. M. Mager-Heckel, A. Lombés, A. Boucheham, I. A. Krasheninnikov, R. P. Martin, N. Entelis and I. Tarassov, *Nucleic Acids Res.*, 2011, **39**, 8173.

10 G. Wang, E. Shimada, J. Zhang, J. S. Hong, G. M. Smith, M. A. Teitell and C. M. Koehler, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 4840.

11 (*a*) P. F. Agris, F. A. P. Vendeix and W. D. Graham, *J. Mol. Biol.*, 2007, **366**, 1; (*b*) Y. Bilbille, F. A. P. Vendeix, R. Guenther, A. Malkiewicz, X. Ariza, J. Vilarrasa and P. F. Agris, *Nucleic Acids Res.*, 2009, **37**, 3342; (*c*) W. D. Graham, L. Barley-Maloney, C. J. Stark, A. Kaur, K. Stolyarchuk, B. Sproat, G. Leszczynska, A. Malkiewicz, N. Safwat, P. Mucha, R. Guenther and P. F. Agris, *J. Mol. Biol.*, 2011, **410**, 698.

12 W. C. Kurschat, J. Muller, R. Wombacher and M. Helm, *RNA*, 2005, **11**, 1909.

13 (*a*) A. Malkiewicz, E. Sochacka, S. Ahmed and S. Yassin, *Tetrahedron Lett.*, 1983, **24**, 5395; (*b*) H. Sierzputowska-Gracz, E. Sochacka, A. Malkiewicz, K. Kuo, Ch. W. Gehrke and P. F. Agris, *J. Am. Chem. Soc.*, 1987, **109**, 7171; (*c*) M. Wieczorek, G. D. Bujacz, D. A. Adamiak, A. Malkiewicz and B. Nawrot, *Heteroat. Chem.*, 1994, **5**, 375; (*d*) Z. Gałdecki, A. Luciak, A. Małkiewicz and B. Nawrot, *Monatsh. Chem.*, 1991, **122**, 487.

14 A. Małkiewicz and E. Sochacka, in *Biophosphates and Their Analogues – Synthesis, Structure, Metabolism and Activity*, ed. K. S. Bruzik, W. J. Stec, Elsevier Science Publishers, Amsterdam, 1987, p. 205.

15 (*a*) A. Malkiewicz and E. Sochacka, *Tetrahedron Lett.*, 1983, **24**, 5387; (*b*) A. Malkiewicz and E. Sochacka, *Tetrahedron Lett.*, 1983, **24**, 5395.

16 P. F. Agris, A. Malkiewicz, A. Kraszewski, K. Everett, B. Nawrot, E. Sochacka, J. Jankowska and R. Guenther, *Biochimie*, 1995, **77**, 125.

17 G. Leszczynska, J. Pięta, P. Leonczak, A. Tomaszewska and A. Malkiewicz, *Tetrahedron Lett.*, 2012, **53**, 1214.

18 G. Leszczynska, J. Pięta, B. Sproat and A. Malkiewicz, *Tetrahedron Lett.*, 2011, **52**, 4443.

19 (*a*) M. Sundaram, P. F. Crain and D. R. Davis, *J. Org. Chem.*, 2000, **65**, 5609; (*b*) A. C. Bajji, M. Sundaram, D. G. Myszka and D. R. Davis, *J. Am. Chem. Soc.*, 2002, **124**, 14302.

20 G. Leszczynska, P. Leonczak, A. Dziergowska and A. Malkiewicz, *Nucleosides, Nucleotides Nucleic Acids*, 2013, **32**, 599.

21 A. Malkiewicz, B. Nawrot and E. Sochacka, *Z. Naturforsch.*, 1987, **42b**, 360.

22 B. S. Sproat, *Methods Mol. Biol.*, 2005, **288**, 17.

23 I. Okamoto, K. Seio and M. Sekine, *Tetrahedron Lett.*, 2006, **47**, 583.

24 (*a*) E. Sochacka, *Nucleosides, Nucleotides Nucleic Acids*, 2001, **20**, 1871; (*b*) E. Sochacka and I. Fratczak, *Tetrahedron Lett.*, 2004, **45**, 6729.
25 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.*, 2003, **7**, 832.
26 (*a*) V. Boudou, J. Langridge, A. Van Aerschot, C. Hendix, A. Millar, P. Weiss and P. Herdewijn, *Helv. Chim. Acta*, 2000, **83**, 152; (*b*) M. Eshete, M. T. Marchbank, S. L. Deutscher, B. Sproat, G. Leszczynska, A. Malkiewicz and P. F. Agris, *Protein J.*, 2007, **26**, 61; (*c*) A. C. Bajji and D. R. Davis, *J. Org. Chem.*, 2002, **67**, 5352.