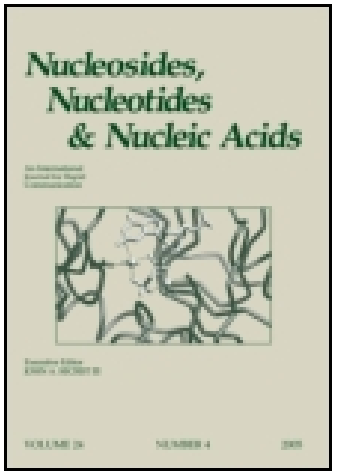


Nucleosides and Nucleotides
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lncn19

New Building Blocks for the Preparation of Branched Oligonucleotides

Alex Azhayev [a] , Seppo Auriola [a] & Jari Hovinen [b]
[a] Department of Pharmaceutical Chemistry , University of Kuopio , FIN-70211, Kuopio, Finland
[b] Department of Chemistry , University of Turku , FIN-20014, Turku, Finland
Published online: 22 Aug 2006.

To cite this article: Alex Azhayev , Seppo Auriola & Jari Hovinen (1998) New Building Blocks for the Preparation of Branched Oligonucleotides, Nucleosides and Nucleotides, 17:9-11, 1527-1537, DOI: 10.1080/07328319808004683

To link to this article: http://dx.doi.org/10.1080/07328319808004683

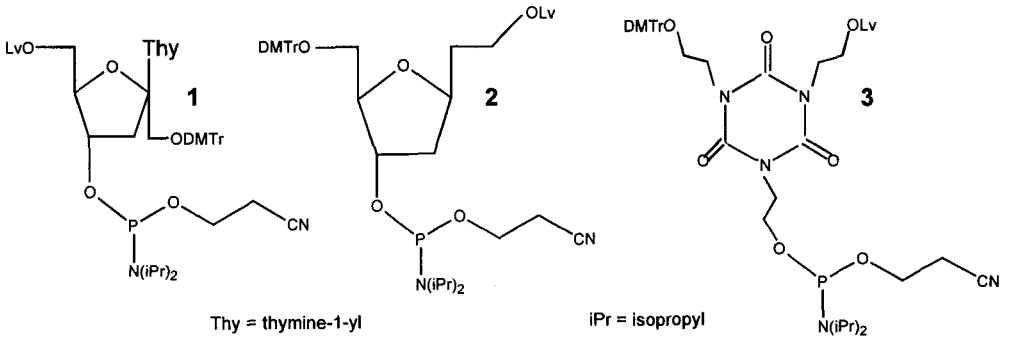
NUCLEOSIDES & NUCLEOTIDES, 17(9-11), 1527-1537 (1998)

NEW BUILDING BLOCKS FOR THE PREPARATION OF BRANCHED OLIGONUCLEOTIDES

Alex Azhayev*, Seppo Auriola and Jari Hovinen#
Department of Pharmaceutical Chemistry, University of Kuopio, FIN-70211 Kuopio, Finland
#Department of Chemistry, University of Turku, FIN-20014 Turku, Finland

ABSTRACT: New building blocks **2** and **3** were prepared and successfully employed for the synthesis of branched oligonucleotides **5** and **6**. The structure of oligomers obtained was confirmed by electrospray ionisation mass spectrometry.

Recently we have reported on the synthesis of branched and looped oligonucleotides,[1] some of which appeared to be powerful inhibitors of replication *in vitro*.[2] In that work we used a phosphoramidite building block[3] **1** at the point of branching. The monomer **1** was derived from 3'-deoxypsicothymidine, which in turn was obtained from **D**-fructose *via* a rather laborious procedure.[4] In the present communication we present more simple non-nucleoside building blocks **2** and **3**.



Thy = thymine-1-yl

DMTr = 4,4'-dimethoxytrityl
Lv = levulinyl (4-oxopentanoyl)

Synthesis of Building Blocks. Compound **2**[5] was prepared starting from 2-deoxy-**D**-ribose using a relatively simple 5-step method. The intermediate ethyl[2-deoxy-5-*O*-(4,4'-dimethoxytrityl)-β-**D**-*erythro*-pentofuranosyl]acetate[6] was protected with *t*-buthyldimethylsilyl chloride (1.5 eq TBDMS-Cl in dry pyridine) and the ester function was reduced (3.5 eq LiBh4 in THF, 6h at 20 °C). The resulting alcohol was treated with levulinic anhydride as described previously[3] and finally desilylated under standard conditions to give Ethyl[2-deoxy-5-*O*-(4,4'-dimethoxytrityl)-β-**D**-*erythro*-pentofurano syl]levulinate.[7] Standard phosphitylation of this yielded monomer **2**.[5]

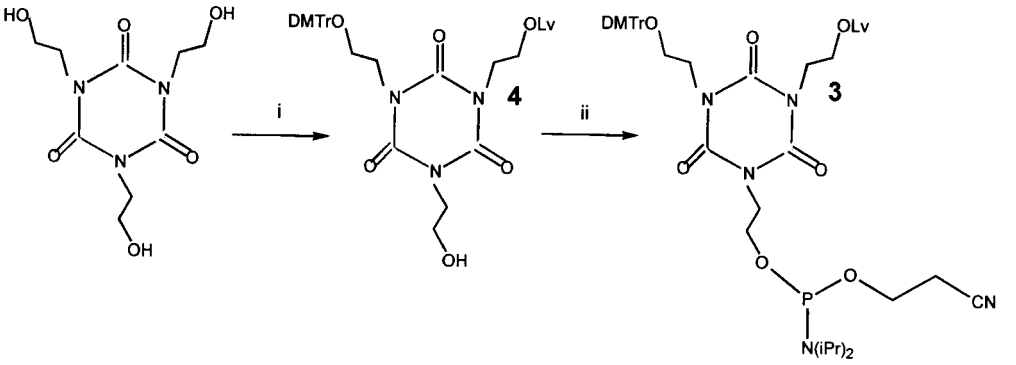
Phosphoroamidite **3** was synthesised from commercially available 1,3,5-tris(2-hydroxyethyl)cyanuric acid, employing a routine straightforward protocol (**SCHEME 1**).

The key intermediate **4**[8] was obtained in 87 % yield using the standard methods of introduction of dimethoxytrityl and levulinyl groups.[3] Finally, treatment of **4** with 2-cyanoethyl-*N,N,N',N'*-tetraisopropylphosphorodiamidite[9] (1.25 equiv.) in acetonitrile in the presence of 1H-tetrazole (1.0 equiv) followed by aqueous work up and precipitation from toluene with cold hexane, gave **3** in 72% overall yield as a white powder.[10]

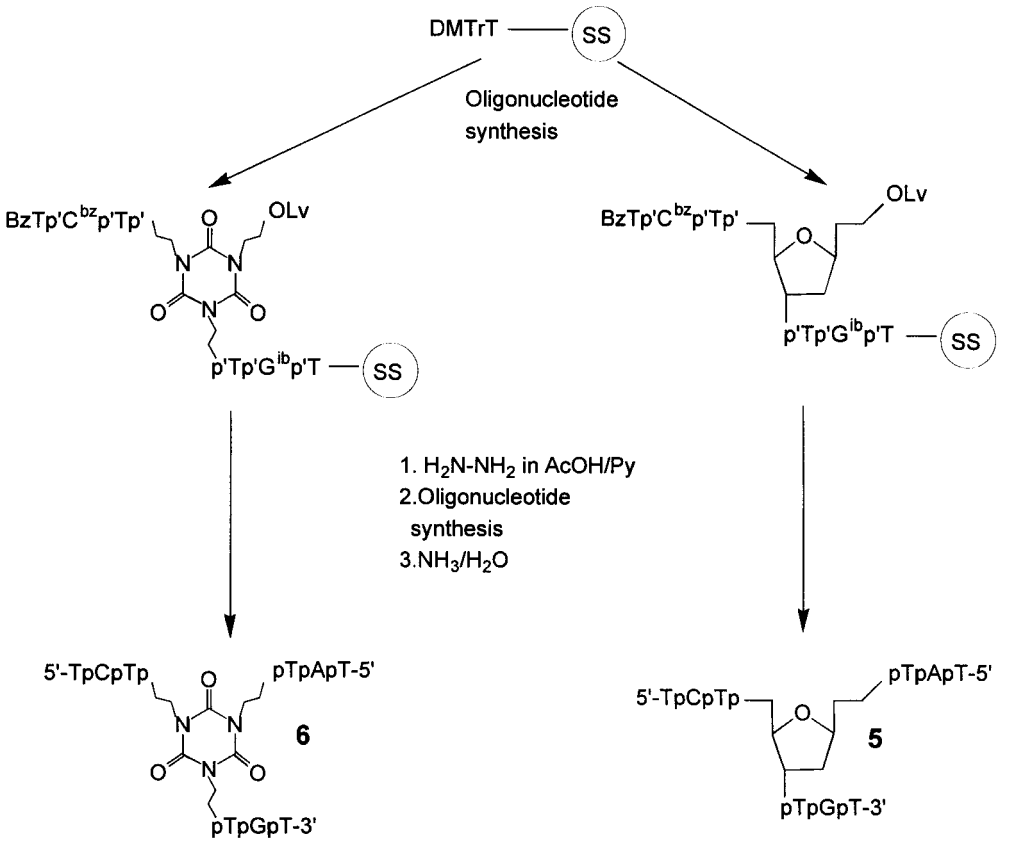
Synthesis of Branched Oligonucleotides

In order to demonstrate the applicability of **2** and **3** in the synthesis of branched oligonucleotides the short sequences **5** and **6** were prepared (**SCHEME 2**). The oligomers were assembled on a PE Applied Biosystems 392 Synthesiser (0.2 μmol scale) according to the standard recommended protocol. No difference in coupling efficiency (>98% as determined by trityl assay) was observed between coupling of **2** or **3** and commercial phosphoroamidites of 2'-deoxynucleosides. After assembling the chain 5'-BzTpC^{bz}p'Tp'X^{lv}p'Tp'G^{ib}p'T-3', employing 5'-*O*-benzoylated thymidine building block to terminate the elongation of this chain,[3] the column was removed from the synthesiser.

The levulinyl protection was then cleaved by treating the solid support with 0.5 M hydrazine hydrate in pyridine/acetic acid (4:1 v/v, 5 ml) for 10 min., followed by



SCHEME 1. i: 1. DMTrCl; 2. Lv_2O. ii: 2-cyanoethy- *N,N,N',N'*-tetraisopropylphosphorodiamidite / TetrH / MeCN.



SCHEME 2. SS: solid support; p': 2-cyanoethyl protected phosphodiester group; p: phosphodiester moiety.

washing the column with ethanol and acetonitrile. After the column was reinstalled to the synthesiser, the second branch, 5'-Tp'A^{bz}p'T, was assembled starting from the liberated hydroxyls of residues, derived from either **2** or **3**. After the completion of the synthesis, the standard ammoniolytic step was used. The branched oligonucleotides **5** and **6** were finally isolated by a combination of anion exchange and RP HPLC, followed by desalting by gel filtration.[3] The yields of **5** and **6** were 78% and 86%, correspondingly.

Characterisation of Branched Oligonucleotides. Electrospray ionisation mass spectrometry (ESI-MS) has been proven to be a gentle and sensitive method for analysis of natural and modified oligonucleotides.[11] The negative ion ESI mass spectra of oligonucleotides show multiply deprotonated molecules as major ions, from which the molecular mass can be deconvoluted. Mass measurements can be complicated by sodium adduct ion formation, which can be reduced by sample desalting or addition of organic bases to the mobile phase. Typically a good solvent system for ESI analysis of oligonucleotides consist of high content organic modifier (50 -80 % of acetonitrile or isopropanol) and 1-20 mM addition of triethylamine, piperidine or imidazole.[12] Tandem mass spectrometry has been widely used for structural analysis of synthetic oligonucleotides and lately also for oligonucleotide metabolites.[13,14] The nomenclature scheme for describing oligonucleotide fragmentation patterns has been described by McLuckey et al.[15,16] These papers describe also the unique feature of an ion trap mass spectrometer to use multiple stage mass spectrometry for sequencing of oligonucleotides. In this method the daughter ions formed after the first dissociation can be further collided with the helium target gas and the resulting ions can be measured.[15,16]

In this study we have used negative ion ESI-MS for structural characterisation of branched oligonucleotides. The purity of the compounds was characterised using the instrument in the full scan mode. The molecular weight of the compounds was verified by measuring the monoisotopic masses using the zoom scan feature, which allows separation of carbon-13 isotope peaks. The structure of the branched

oligonucleotides was characterised using tandem mass spectrometry, which allows differentiation of bases in the 3'- and 5' ends of the nucleotide. The sequential tandem mass spectrometry in the ion trap by using MS3 option was used for structural analysis of the fragment ions. Mass spectra were acquired using an LCQ quadrupole ion trap mass spectrometer equipped with an electrospray ionisation source (Finnigan MAT). The spray needle was set to - 3 kV. The spray was stabilised using a nitrogen sheath flow, the value was set to 70. The stainless-steel inlet capillary was heated to 200 °C. The full scan mass spectra were measured using 200 ms for collection of the ions in the trap. When the narrow range (10 mass units) zoom scan spectra were measured, the control was set to fill the trap with less than 10^6 ions. Typically 5 to 10 scans were averaged during measurement of each injection. The eluent consisted of 80 % isopropanol containing 25 mM triethylamine. The flow was set to 10 µl/min and 5 µl samples containing 100 pmol of oligonucleotide in 60% isopropanol and 25mM triethylamine were injected to the system.

The structure of oligonucleotides was studied using sequential tandem mass spectrometry by using collision energy value 25%. In the first experiment the doubly deprotonated molecule was subjected to collisional activation with helium target gas in the trap (MS2). The resulting daughter ions were measured with full scan for mass range m/z 400-2000. In the following experiments the major fragment ions were collected and further collided with the target gas (MS3) and the full scan mass spectra were measured. The molecular weights of the ions were calculated using computer program Gretas Carbos, created by W. Hines and B.W. Gibson (UCSF). The full scan mass spectra of the branched oligonucleotides show doubly deprotonated molecules as the base peaks at m/z 1508 for **6**, and at m/z 1458 for **5**. The spectra show also a series of sodium adduct ions, despite of the desalting of the samples, and despite of the use of triethylamine in the ESI-background solvent (**FIG. 1**). The measurement of zoom scan mass spectra of compounds show, that resolution of the ion trap instrument is good enough to separate the carbon-13 isotope peaks of oligonucleotides (**FIG. 2**). The resolution value 3000 was achieved by limiting the number of ions collected in the trap to 10^6.

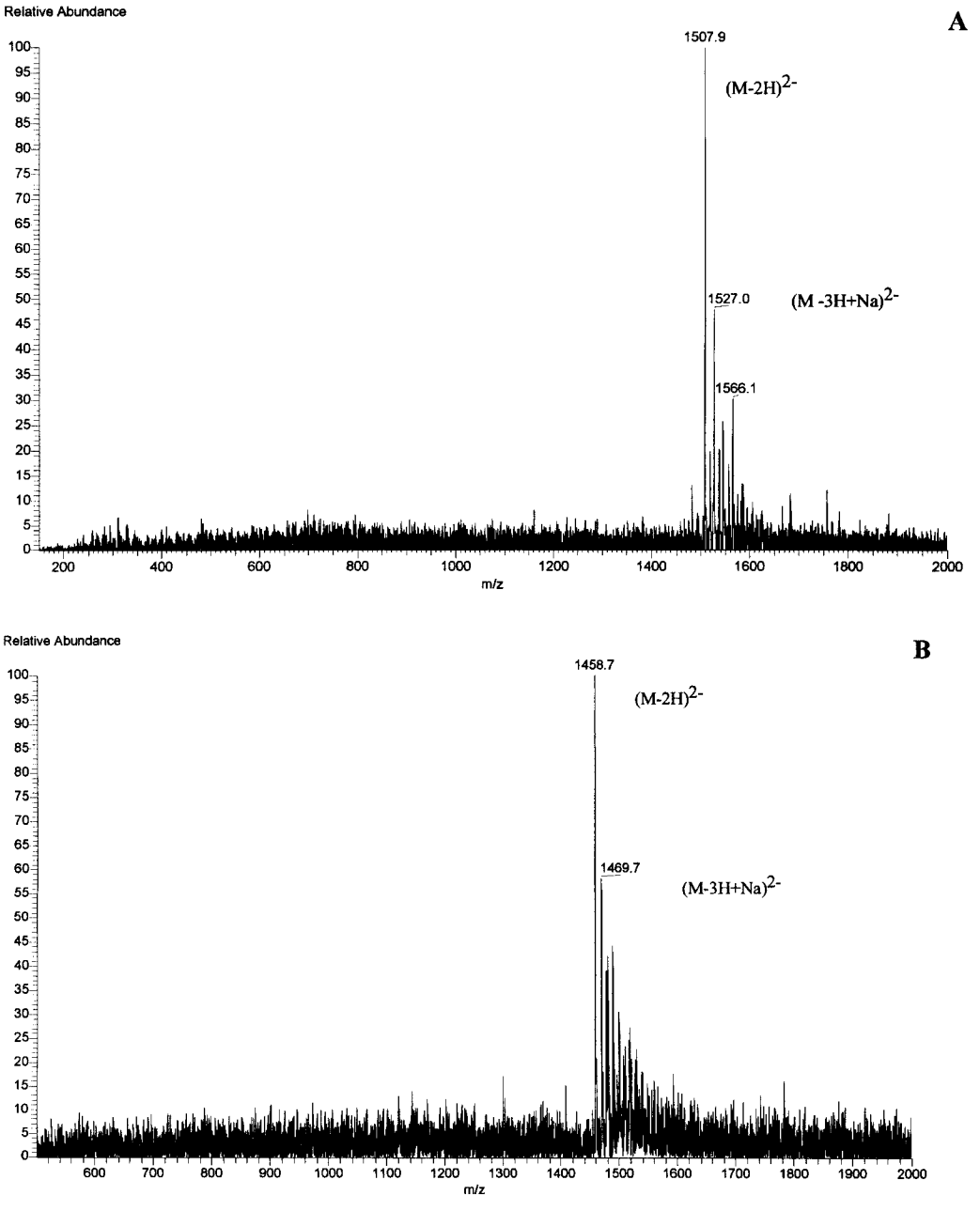


FIG. 1. Full scan negative ESI mass spectra of A) **6** and B) **5**.

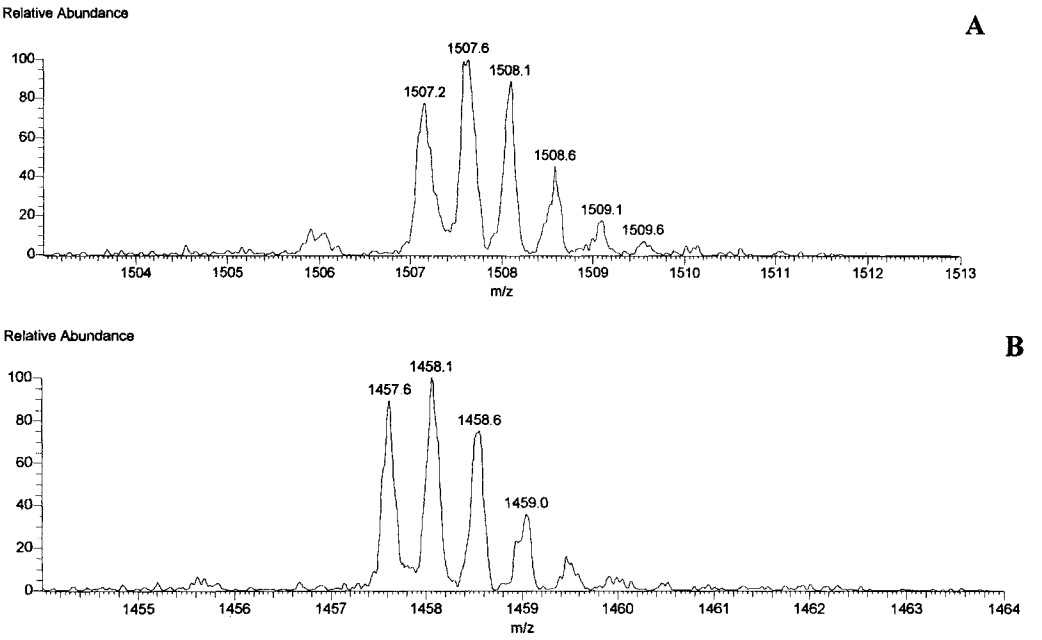


FIG. 2. Zoom scan negative ESI mass spectra of A) **6** and B) **5**, showing the isotopic pattern of the oligonucleotides.

The measured monoisotopic molecular weights of **6** and **5** were 3016.3 (calculated 3015.6) and 2917.2 (calculated 2916.6) correspondingly. The sequence of the branched oligonucleotides **6** and **5** was partially confirmed using sequential MS/MS analysis. The CID (collision induced dissociation) mass spectra of the branched oligonucleotides show three types of fragment ions: 1) ions generated by loss of a neutral base; 2) ions formed with a single cleavage from either 3'- or 5'-end of the oligonucleotide; and 3) ions formed by a double or triple cleavage from two or three ends of the branched oligonucleotide. Following the nomenclature of McLuckey,[14] the most diagnostic ions formed are the w-series ions (used to determine the sequence in the 3' to 5' direction) and a-Base series ions (used to determine sequence in 5' to 3' direction). In the MS2 mass spectrum of **6** (**FIG. 3 A**), the ion at m/z 1432 is generated by loss of guanine from the doubly deprotonated oligonucleotide ion (m/z 1508). Ions from loss of adenine and cytosine are also observed at m/z 1440 and m/z

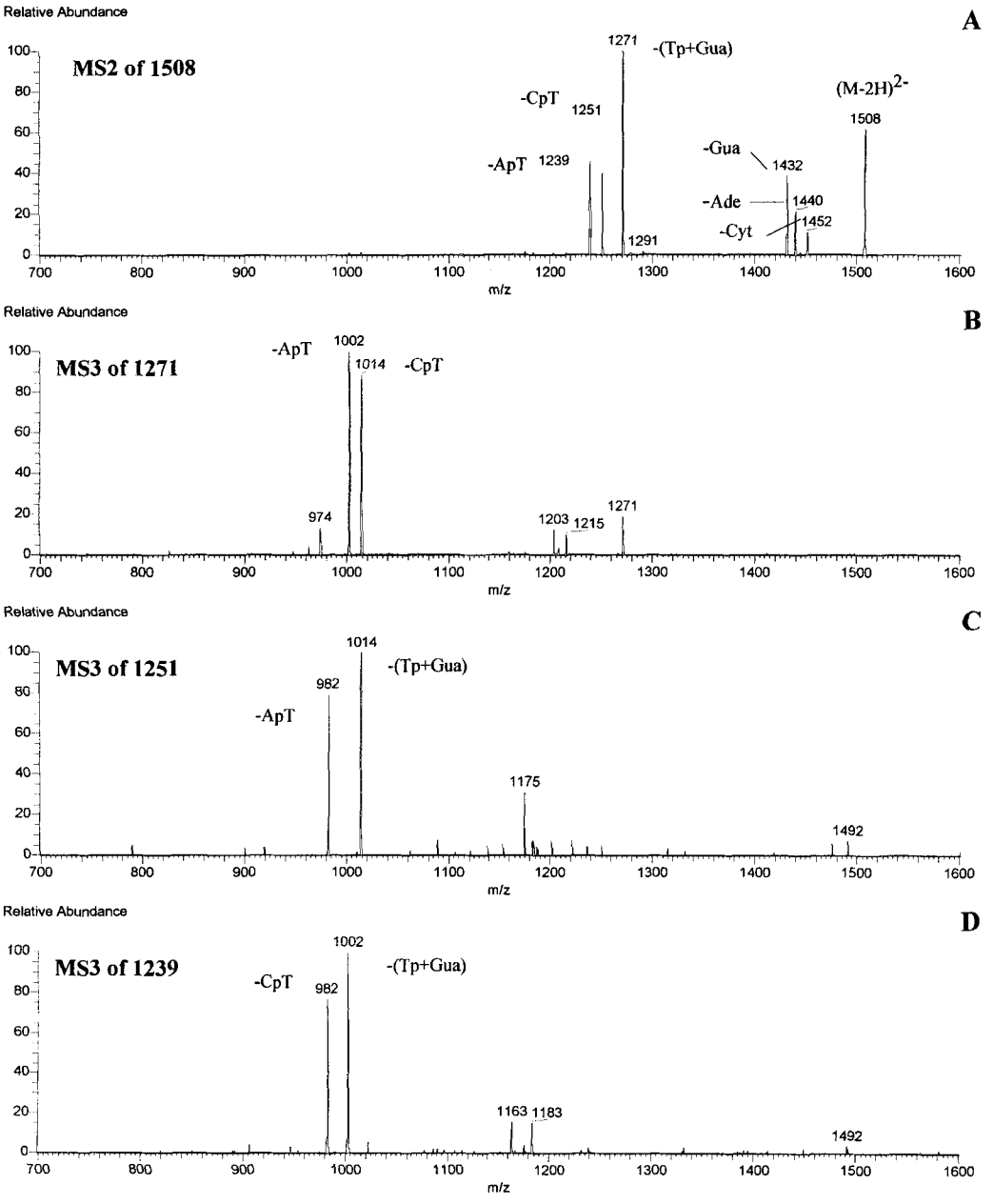


FIG. 3. A) MS2 spectrum of the doubly deprotonated molecule **6**, precursor ion m/z 1508. B) MS3 spectrum resulting from the MS2 fragment ion m/z 1271. C) MS3 spectrum resulting from the MS2 fragment ion m/z 1251. D) MS3 spectrum resulting from the MS2 fragment ion m/z 1239.

TABLE 1. Collision induced dissociation fragments generated by MS2 and MS3 in the ion trap mass analyser.

Branched oligonucleotide **6**:			
MS2		**MS3**	
1508 - ApT	=1239	1239 - CpT	=982
		1239 - (Tp + Gua)	=1002
1508 - CpT	=1251	1251 - ApT	=982
		1251 - (Tp + Gua)	=1014
1508 -(Tp + Gua)	=1271	1271 - ApT	=1002
		1271 - CpT	=1014

Branched oligonucleotide **5**:			
MS2		**MS3**	
1458 - ApT	=1189	1189 - CpT	=932
		1189 -(Tp + Gua)	=952
1458 - CpT	=1201	1201 - ApT	=932
		1201 -(Tp + Gua)	=964
1458 -(Tp + Gua)	=1221	1221 - ApT	=952
		1221 - CpT	=964

1452, respectively. The most abundant ion at m/z 1271 is the a-Base series ion, which formed by loss of one T, one phosphodiester group (p) and a guanine base from the 3' end.

The ions at m/z 1250.7 and m/z 1239 indicated losses of CpT and ApT from the 5' and 5''- ends, respectively. Further fragmentation can be achieved either by using higher collision energy in the trap (data not shown), or by isolating the primary fragment ions and subjecting them again to collisional activation. **FIG. 3 B** shows the full scan MS3 mass spectrum resulting from dissociation of ion m/z 1271. The two major peaks at m/z 1014 and m/z 1002 indicate losses of CpT and ApT, respectively. The MS3 mass spectrum of ion m/z 1251 (**FIG. 3 C**) shows fragment ions at m/z 982 and at m/z 1014, which indicate losses of ApT and Tp+guanine. In the similar manner, the MS3 mass spectrum of ion m/z 1239 shows ions at m/z 982 and m/z 1002, losses of CpT and Tp+guanine, respectively (**FIG. 3 D**). All these ions formed

by cleavages from two branches are relatively stable, and the ions that indicates fragmentation of all three branches at m/z 745 (doubly charged) and at 1492 (singly charged) are of very low abundance. The fragmentation pathways for doubly deprotonated ions of **5** and **6** are described in **TABLE 1**.

As a conclusion it is noteworthy to note that building blocks **2** and **3** may be easily prepared and successfully employed for the synthesis of branched oligonucleotides. The ion trap mass spectrometer is well suited for characterisation of branched oligomers **5** and **6**. By using the zoom scan option, the monoisotopic molecular weight of the compounds can be verified with an error less than 1 mass unit. When the doubly deprotonated oligonucleotides are subjected to collisional activation, the resulting mass spectra show w- and a-Base type ions, same type which have been described to be the most abundant for linear oligonucleotides.[15] The presence of these ions makes it possible to identify the base sequence in the 3', 5' and 5'' ends of compounds . However, the nucleotides closest to the branching point structure are not cleaved from the ions by further collision. This phenomena hampers the complete sequencing of branched oligonucleotides.

REFERENCES

1. Azhayeva, E., Azhayev, A., Guzaev, A., Hovinen, J., Lönnberg, H. *Nucleic Acids Res.* **1995**, **23**, 1170 - 1176.
2. Azhayeva, E., Azhayev, A., Guzaev, A., Lönnberg, H. *Nucleic Acids Res.* **1995**, **23**, 4255 - 4261.
3. Azhayev,A., Guzaev, Hovinen, J., Azhayeva, E., Lönnberg, H. *Tetrahedron Lett.* **1993**, **34**, 6435 - 6438.
4. Azhayev, A., Guzaev, A., Hovinen, J., Mattinen, J., Siillanpää, R., Lönnberg, H. *Synthesis* **1994**, 396 - 400.
5. Compound **2**: ^{31}P NMR ($CDCl_3$): 148.94 (0.5 P); 148.71 (0.5 P).
6. Hovinen, J., Salo, H. *J. Chem. Soc. Perkin Trans.I* **1997**, in press.
7. Ethyl[2-deoxy-5-*O*-(4,4'-dimethoxytrityl)-β-**D**-*erythro*-pentofuranosyl] levulinate: ^{1}H NMR ($CDCl_3$):7.47 -8.80 (13H, m, arom.); 4.20 (4H, m, H-1, H-3, -$\underline{CH_2}$OCO-), 3.90 (1H, m, H-4), 3.78 (6H, s, -O$\underline{CH_3}$); 3.20 (1H, dd, $J_{4,5a}$ =4.6 Hz, $J_{5a,5b}$=9.8 Hz, H-5a); 3.07 (1H, dd, $J_{4,5b}$=5.6 Hz, $J_{5a,5b}$=9.8 Hz, H-5b); 2.73 (2H, t, *J*=6.3 Hz, -$\underline{CH_2}$COO-); 2.55 (2H, t, *J*=6.3 Hz, -$\underline{CH_2}$CO-); 2.20 (1H, d, *J*=3.6 Hz, 3-OH); 2.17 (3H, s, $COCH_3$); 1.98 (1H, ddd, $J_{2a,3}$=1.9 Hz, $J_{2a,CH2}$=3.6 Hz, $J_{2a,2b}$=10.7 Hz, H-2a); 1.89 (2H, m, -CH_2-); 1.75 1H, m, H-2b).

8. Compound **4**: ^{1}H NMR ($CDCl_3$): 7.43 - 6.73 (13H, m arom.); 4.32 (2H, t, *J*=5.4 Hz, LvO$\underline{CH_2}$-); 4.16 - 4.06 (3H, m, -$\underline{CH_2}$N=); 3.83 (2H, t, *J*=4.9 Hz, HO$\underline{CH_2}$-); 3.78 (6H, s, -O$\underline{CH_3}$); 3.41 (2H, t, *J*=5.1 Hz, DMTrO$\underline{CH_2}$-); 2.65 (2H, t, *J*=6.6 Hz, -OCO$\underline{CH_2}$); 2.43 (2H, t, *J*=6.9 Hz, $CH_3CO\underline{CH_2}$-); 2.13 (3H, s, $\underline{CH_3}$-).
9. Nielsen, J., Dahl, O. *Nucleic Acids Res.* **1987**, **15**, 3626.
10. Compound **3**: ^{31}P NMR ($CDCl_3$): 149.17 (0.5 P); 149.05 (0.5 P).
11. Potier, N., Van Dorsselaer, A., Cordier, Y., Roch, O., Boschoff, R. *Nucleic Acids Res.* **1994**, **22**, 3895 - 3903.
12. Greig, M., Griffey, R. *Rapid Commun. Mass Spectrom.* **1995**, **9**, 97 - 102.
13. Barry, J., Vorous, P., Van Schepdael, A., Law, S.-J. *J. Mass Spectrom.*, **1995**, **30**, 993 - 1006.
14. Griffey, R., Greig, M., Gaus, H., Liu, K., Monteith, D., Winniman, M., Cummins, L. *J. Mass Spectrom.*, **1997**, **32**, 305 - 313.
15. McLuckey, S., Berkel, G., Glish, G. *J.Am Soc. Mass Spectrom.* **1992**, **3**, 60 - 70.
16. McLuckey, S., Habibi-Goudarzi, S. *J.Am. Chem. Soc.* **1993**, **115**, 12085 - 12095.