REGIOSPECIFIC SYNTHESIS OF BRANCHED TETRANUCLEOTIDES:


X.-X. ZHOU$^1$, A. NVILAS$^{1,2}$, G REMAUD$^1$ and J. CHATTOPADHYAYA$^1$

(1) Department of Bioorganic Chemistry, Box 581, Biomedical Center, University of Uppsala, S-751 23 Uppsala, Sweden
(2) Center for Agricultural Biotechnology, 2101 Gödöllö, Hungary

(Received in UK 21 July 1987)

Abstract: An efficient general strategy for the synthesis of branched tetranucleotides $U4$, $U5$, $A4$ and $A5$ is described using key intermediates $U2$ and $A2$ to give protected tetranucleotides $U2$ and $A2$ which could be specifically deprotected to give either $U4$ or $U5$ and $A4$ or $A5$ in good yields.

Precision in the chemical processing (splicing) of eukaryotic pre-mRNA, involving the excision of correct sequence of introns and ligation of exons is absolutely essential for the biological activity$^1$ of the resulting product. In group II and nuclear mRNA splicing, the scission of the intron and the subsequent ligation of exons is processed through the formation of a branched structure (lariat) with the $3'$-exon attached giving a circular intron (lariat) and the ligated $5'$- and $3'$-exons$^2$. In such lariat structures, adenosine residue forms the branch point with an additional $2'$-$5'$ nucleobase, which is invariably a guanine residue, while the $3'$-$5'$ nucleobase is either a uracil or a cytosine residue$^2,3$. We have recently reported an unambiguous regiospecific synthesis of simple trimeric branched structures$^4$ $U1$ and $A1$, which allowed us to carry out their 270 MHz and 500 MHz $^1$H-NMR studies. Such conformational studies have revealed an unique structural feature of these branched trinucleotides $U1$ and $A1$ consisting of preferential $2'$-$5'$ stacking$^5$ (free-energy minimum) and a complete absence of $3'$-$5'$ stacking.

We now report an expedient synthesis of four branched tetranucleotides $U4$, $U5$, $A4$ and $A5$ comprising of an additional either uridine or adenosine residue at the $5'$-end of the branched molecules $U1$ and $A1$, in order to address the conformational influence of the fourth base residue at the $5'$-terminus on the overall branched structures of $U1$ and $A1$. These branched tetranucleotides are naturally occurring and have been actually isolated from yeast cells$^2,3$. In this new procedure, we introduced 2-phenylsulfonylethyl (PSE)$^6$ or 9-fluorenylmethyl (FM)$^7$ phosphiteamidite function selectively at the $2'$-OH of the building block $U2$ to give compound $U4$ in 86 % yield ($^3$P-NMR: 149.9 and 149.7 ppm) and $A4$ in 55 % yield ($^3$P-NMR: 149.8 ppm) respectively which we coupled to an appropriate $5'$-hydroxy block $U5$, in presence of tetrazole$^8$, to give the fully protected dinucleotide $U6$ in 79 % yield ($^3$P-NMR: -3.6 and -4.3 ppm) or $A7$ in 56 % yield ($^3$P-NMR: -3.29 and -3.88 ppm). The $5'$-hydroxyl functions of these dimers $U6$ and $A7$ were regiospecifically released$^9$, using 0.2 M aqueous HCl in dioxane, to give compound $U2$ ($^3$P-NMR: -3.25 and -3.9 ppm) in 86 % yield and compound $A8$ ($^3$P-NMR: -3.27 ppm) in 70 % yield.
Regiospecific synthesis of branched tetranucleotides

11: \( B = U; R = \text{DMT} \)

22: \( B = \text{Abz}; R = \text{Tol} \)

13: \( B = U; R = \text{DMT} \)

23: \( B = \text{Abz}; R = \text{Tol} \)

14: \( B = U \)

24: \( B = \text{A} \)

15: \( B = U \)

25: \( B = \text{A} \)

(MDMP) (PSE) (FM) (2-ClPh) (Tol)

\((\text{U}^{\text{NP}})\) \((\text{GTBB})\) \((\text{Abz})\) (DMT)
respectively. Chain elongation at the 5'-end was accomplished, using methodologies of the phosphotriester approach\textsuperscript{10}, by coupling with an appropriate 5'-protected phosphodiester block \textsuperscript{8} or \textsuperscript{12} in presence of 1-mesitylenesulfonyl-(3-nitro-1,2,4-triazole) [MSNIT]\textsuperscript{11} to give the protected trimer \textsuperscript{2} in 87 % yield (\textsuperscript{31}P-NMR: -3.4 and -7.4 ppm) or \textsuperscript{20} in 78 % yield (\textsuperscript{31}P-NMR: -2.61, -2.66, -6.98 and -7.34 ppm). Then the 2-phenylsulphonyl (PSE) group\textsuperscript{6} from compound \textsuperscript{2} and the 9-fluorenethylmethyl (FM) group\textsuperscript{7} from compound \textsuperscript{20} were selectively removed [Et\textsubscript{3}N (10 equiv.) in dry pyridine at 20 °C for 2 h] in almost quantitative yields to give compound \textsuperscript{10} (\textsuperscript{31}P-NMR: -2.64 and -7.5 ppm) and compound \textsuperscript{22} (\textsuperscript{31}P-NMR: -2.34, -7.03 and -7.57 ppm), respectively. Both 3'-silyl and 2-chlorophenyl groups were then removed in one step using n-tetrabutylammonium fluoride\textsuperscript{12} in moist tetrahydrofuran at 20 °C to give pure partially protected trimer \textsuperscript{11} (\textsuperscript{31}P-NMR: -0.7 and -0.8 ppm) in 87 % yield or \textsuperscript{22} (\textsuperscript{31}P-NMR: -0.59 and -0.68 ppm) in 90 % yield. Finally, compound \textsuperscript{22} and \textsuperscript{11} were coupled to a 5'-protected phosphoramidite uridine derivative \textsuperscript{12}, in presence of tetrazole, to give tetranucleotides \textsuperscript{14} (96 %) and \textsuperscript{24} (89 %) which were deprotected in two different ways\textsuperscript{13} to give \textsuperscript{14} in 46 % yield (\textsuperscript{31}P-NMR: -0.8, -1.0 and -1.2 ppm) and \textsuperscript{24} in 41 % yield (\textsuperscript{31}P-NMR: -0.83, -0.98 and -1.27 ppm) [liquid
Regiospecific synthesis of branched tetranucleotides

NH₃ for 48 h, aqueous ammonia for 120 h, 80 % aqueous acetic acid for 5 h at 20 °C] or to give 15 in 49 % yield (³¹P-NMR: -0.9, -1.0 and -1.3 ppm) and 25 in 39 % yield (³¹P-NMR: -0.83, -0.90 and -1.29 ppm) [4-nitrobenzaloximate for 24 h, aq. NH₃ for 120 h, 80 % aqueous acetic acid for 5 h]. The 270 MHz ¹H-NMR spectra, 2D ¹H correlations and ³¹P/¹H correlation spectra of the tetranucleotides in D₂O are shown in Figures 1 - 12 which show that the desired regiospecificities of the phosphodiester linkages have indeed been achieved. Assignment of resonances with practical details will be described elsewhere.

**EXPERIMENTAL**

¹H-NMR spectra were recorded, in δ scale, with Jeol 90 Q and JNM GX 270 spectrometers at 90 and 270 MHz respectively, using TMS or acetonitrile (set at 2.0 ppm) as internal standards. ³¹P-NMR spectra were recorded at 36 MHz in the same solvent using phosphoric acid as an external standard and chemical shifts quoted are in ppm. TLC was carried out using pre-coated silica gel F₂₅₄ plate in following solvent systems: (A) methanol-dichloromethane [9.5:0.5 (v/v)], (B) methanol-dichloromethane [9:1 (v/v)], (C) methanol-dichloromethane [8:2 (v/v)], (D) ethylacetate-hexane-triethylamine [4:4:2 (v/v/v)]. The short column chromatographic separations were carried out using Merck G60 silica gel.

Fig. 3: 270 MHz ¹H-NMR spectrum of A₃'p₅'A₅'p₅'G₃₇p₅'U in D₂O at 298 K.

Fig. 4: 270 MHz ¹H-NMR spectrum of A₃'p₅'A₅'p₅'G₃₇p₅'C in D₂O at 298 K.
To a suspension of 2-phenylsulfonyylethylphosphorodichloridite (2.29 g, 8 mmol) and 1,2,4-triazole (1.93 g, 28 mmol) in dry THF (20 ml) at -25 °C, was added N,N-diisopropylethylamine (4.89 ml, 28 mmol). After stirring the reaction mixture for 12 min at -25 °C, a solution of compound 3 (2.4 g, 4 mmol) in dry THF (20 ml) was added dropwise within 1 h and kept it stirring for 40 min, trimethylsilyl-N,N-dimethylamine (4.4 ml, 28 mmol) was then added. After 10 min, the reaction mixture was warmed up to room temperature, and kept stirring for a further period of 15 min. The reaction mixture was subsequently poured into saturated aqueous sodium chloride solution (150 ml), and extracted with ethylacetate (100 ml). The organic layer was washed with sodium chloride solution and water successively, and dried in vacuo. The residue was purified by short silica gel column chromatography, [hexane: dichloromethane:ethylacetate:triethylamine, 3:3:1:1 (v/v/v/v)], the product was precipitated from cooled hexane (-70 °C). Yield: 2.95 g (86 %) Rf = 0.6 (solvent D). 1H-NMR (mixture of two diastereomers) (CDCl3 + pyridine-d5): 8.74 and 8.73 (2 x s, 1H) H-8; 8.33 and 8.32 (2 x s, 1H) H-2; 8.08-7.19 (m, 10H) BZ and PSE; 6.05 (m, 1H) H-1'; 4.66 (m, 2H) H-2', H-3'; 4.15 (m, 5H) H-4', H-5', PSE; 3.42 (m, 2H) PSE; 2.61, 2.58, 2.51 and 2.48 (4 x s, 6H) (CH3)2N-; 1.05 (m, 28H) TIPDSi. 31P-NMR (CDCl3): 149.9 and 149.7.
Synthesis of fully protected dinucleoside monophosphate (6)

A mixture of compounds 4 (2.3 g, 3 mmol), 5 (527 mg, 1 mmol) and 1,2,3,4-tetrazole (1.4 g, 20 mmol) was dissolved in dry acetonitrile (20 ml). After 40 min, iodine solution (0.1 M) in pyridine: THF: water mixture (1:8:1, v/v/v) was added until the iodine color was not further discharged. After 20 min, the reaction mixture was worked up in the usual way. The product was purified by short silica gel column chromatography. Yield 1.07 g (79 %) $R_f = 0.4$ (solvent B). $^{31}P$-NMR (CDCl$_3$): -3.56 and -4.27.

![Double Quantum Filter COSY of U3'p5'S2'p5'G 3'p5'C](image)

Fig. 6

Synthesis of partially protected dinucleoside monophosphate (7)

Compound 6 was dissolved in 0.2 M aqueous HCl in dioxane (14 ml) under stirring at 20 °C. A few drops of 0.2 M aqueous HCl were subsequently added until the reaction mixture became opalescent. After 1 h, the mixture was poured into saturated sodium hydrogen carbonate solution (100 ml) and extracted with dichloromethane (3 x 40 ml). The extract was concentrated and chromatographed by short silica gel column chromatography. Yield: 708 mg (66 %) $R_f = 0.35$ and 0.37 (solvent B). $^{31}P$-NMR. (CDCl$_3$): -3.25 and -3.88.
Synthesis of fully protected trimer (9)

Compound 7 (487 mg, 0.35 mmol) and 8 (707 mg, 0.7 mmol) were dissolved in dry pyridine (10 ml). To this solution was added 1-mesitylenesulfonyl-3-nitro-1,2,4-triazole (MSNT) (1.6 g, 4.9 mmol) at 20 °C. After 20 min, the reaction mixture was poured into a saturated sodium hydrogen carbonate solution (80 ml), and extracted with dichloromethane (3 x 30 ml). The product was subsequently isolated by short silica gel column chromatography. Yield: 692 mg (87 %). 


Synthesis of compound (10)

Compound 9 (680 mg, 0.3 mmol) was dissolved in pyridine (10.5 ml), followed by triethylamine (0.83 ml, 6 mmol). After 90 min at 20 °C, the reaction mixture was dried in vacuo, and purified by short silica gel column. Yield: 614 mg (99 %). 

31P-NMR (CDC13 + pyridine-d9): -2.64 and -7.49.

Synthesis of compound (11)

Compound 10 (600 mg, 0.29 mmol) was dissolved in THF (10 ml) and, subsequently, 1 M tetrabutylammonium fluoride in moist THF (1.74 ml) was added at 20 °C. After 1 h, the reaction mixture was dried in vacuo and purified by silica gel column chromatography. Yield: 480 mg (86 %). 

31P-NMR (CDC13 + pyridine-d9): -0.71 and -0.78.
Synthesis of compound (12)

$\text{ONP}_2\text{Ac}_2\text{U}$ (1.35 g, 3 mmol) was dissolved in dichloromethane (25 ml) and subsequently N,N-diisopropylethylamine (2.1 ml, 12 mmol) and methoxy diisopropylaminochlorophosphine (1.14 ml, 6 mmol) were successively added. After 50 min, the reaction mixture was poured into a saturated sodium chloride solution (100 ml) and extracted with ethyl acetate (4 x 50 ml). The product was subsequently purified by silica gel column chromatography. Yield: 1.61 g (90%). $R_f = 0.65$ (solvent D). $^31\text{P-NMR (CDCl}_3$): 150.5 and 150.4.

Synthesis of compound (13)

Compound 11 (350 mg, 0.2 mmol) and 12 (596 mg, 1 mmol) and tetrazole (700 mg, 10 mmol) were dissolved in dry acetonitrile (7 ml). After 30 min, the reaction mixture was oxidized with 0.1 M iodine solution in THF-pyridine-water mixture (1:8:1, v/v/v) and then worked up in the usual way. The column chromatographic separation was carried out, first, by washing the column with 4% methanol in dichloromethane and then the product was eluted out by 30% methanol in dichloromethane. Yield: 428 mg (96%). $R_f = 0.2$ (solvent C). $^31\text{P-NMR (CDCl}_3 + \text{pyridine-d}_5$): -0.78, -0.85, -1.32, -1.42, -2.22, -2.78.

Deprotection of compound (13)

Procedure A: Compound 12 (200 mg, 0.09 mmol) was dissolved in dioxane-water mixture (30 ml, 8:2 v/v) and then N,N,N,N-tetramethylguanidine (0.36 ml, 2.79 mmol)
and \( \text{syn-} p \)-nitrobenzaldoxime (448 mg, 2.7 mmol) were added. After 24 h at 20 °C, concentrated ammonia (70 ml, \( d = 0.9 \)) was added. The reaction mixture was kept stirring for 6 days at 20 °C, and then dried in vacuo. The residue was coevaporated with water once, and then treated with 80 % aqueous acetic acid (50 ml) for 5 h. The reaction mixture was dried in vacuo, residue dissolved in water and extracted with dichloromethane (6 x 20 ml). The aqueous layer was dried in vacuo and subjected to DEAE-Sephadex A25 column chromatographic separation using a linear gradient (0 to 0.5 M) of triethylammonium bicarbonate (TEAB) (pH 7.4) as an eluent. Appropriate fractions were pooled and concentrated. The residue was subsequently coevaporated with distilled water several times to remove traces of buffer to give pure compound \( z \). Yield: 1940 \( A_{260} \) o.d. units (49%). \( ^{31} \text{P} \)-NMR (D2O): -0.88, -0.98, -1.30.

Procedure B: Compound \( u \) (200 mg, 0.09 mmol) was treated with liquid ammonia (30 ml) for 2 days at 20 °C, then the volatile matters were evaporated. The residue was then dissolved in concentrated ammonia (70 ml, \( d = 0.9 \)). After 6 days, the solution was dried in vacuo. After a few coevaporations with water, it was treated with 80 % aqueous acetic acid (50 ml) for 5 h at 20 °C. The reaction mixture was dried in vacuo, the residue was partitioned between water (50 ml) and dichloromethane (50 ml). The aqueous layer was further extracted with dichloromethane (3 x 40 ml) and subsequently aqueous layer was concentrated and applied on a DEAE-Sephadex A25 column using a linear gradient of 0 to 0.5 M TEAB solution (pH 7.4) as eluent. Appropriate fractions were pooled, concentrated, coevaporated with distilled water several times to give compound \( u \). Yield: 1718 \( A_{260} \) o.d. units (46 %). \( ^{31} \text{P} \)-NMR (D2O): -0.83, -0.98, -1.20.

Preparation of compound (16)

To the mixture of 9-fluorenylethylphosphorodichloridite (1.78 g, 6 mmol) and 1,2,4-triazole (1.45 g, 21 mmol) in dry THF at -25 °C, was added N,N-diisopropylethylamine (3.66 ml, 21 mmol). After stirring for 10 min, a solution of compound 1

\[ \text{31P-}^1 \text{H CORRELATION SPECTRUM OF U3'p5'U} \]
(1.84 g, 3 mmol) in THF (10 ml) was added dropwise. The mixture was kept for 30 min and then triethylamyl-N,N-dimethylamine (3.3 ml, 21 mmol) was added. After 10 min, the reaction mixture was warmed up to room temperature, and kept stirring for a further period of 10 min. It was then worked up and separated by the procedure described for the preparation of compound 4. Yield: 1.46 g (55 %). 31P-NMR (CDCl3): 149.8.

A3'p5'C  U3'p5'A  A2'p5'G

31P-1H CORRELATION SPECTRUM OF U3'p5'A2'p5'G3'p5'C

Fig. 10

Preparation of compound (17)

Compound 16 (880 mg, 1 mmol) and 5 (220 mg, 0.5 mmol) were coupled in presence of 1,2,3,4-tetrazole (0.7 g, 2.4 mmol) in dry THF as described for the preparation of compound 6. Yield: 500 mg (56 %). 31P-NMR (CDCl3): -3.29 and -3.88.

Preparation of compound (18)

Compound 17 (500 mg, 0.33 mmol) was treated with 0.2 M aqueous HCl in dioxane using a procedure as described for the preparation of compound 7. Yield: 350 mg (70 %). 31P-NMR (CDCl3): -3.27.

Preparation of compound (20)

Compound 18 (500 mg, 0.37 mmol) and 12 (725 mg, 0.74 mmol) and MSNT (746 mg, 2.2 mmol) were reacted in pyridine solution according to the method described for the preparation of compound 2. Yield: 634 mg (77.5 %). Rf = 0.38 (solvent B). 31P-NMR (CDCl3): -2.61, -2.66, -6.98, -7.34.
Preparation of compound (21)

Compound 20 (630 mg, 0.28 mmol) was treated with triethylamine in dry pyridine in a similar way as described for the preparation of compound 12. Yield: 601 mg (100 %). \( E_F = 0.27 \) (solvent R). \(^{31}\text{P-NMR (CDCl}_{3}\)): -2.34, -7.03, -7.57.

\[ \text{PPH} \]

\[ \text{Correlation Spectrum of A3'p5'U A3'p5'A* A2'p5'G} \]

\[ \text{PPM} \]

\[ \text{-0.5} \]

\[ \text{-1.0} \]

\[ \text{-1.5} \]

\[ \text{-2.0} \]

\[ \text{31P-1H CORRELATION SPECTRUM OF A3'p5'U A3'p5'G} \]

\[ \text{Fig. 11} \]

Preparation of compound (22)

Compound 22 (598 mg, 0.28 mmol) was treated with tetrabutylammonium fluoride in dry THF in a similar way as described for the preparation of compound 11. Yield: 449 mg (90 %). \( E_F = 0.20 \) (solvent C). \(^{31}\text{P-NMR (CDCl}_{3}\)): -0.59 and -0.68.

Preparation of compound (23)

Compound 22 (390 mg, 0.22 mmol) was condensed with compound 12 (655 mg, 1.1 mmol) in acetonitrile in presence of tetrazole (770 mg, 11 mmol) in the usual way. Yield: 472 mg (89 %). \( E_F = 0.27 \) (solvent C). \(^{31}\text{P-NMR (CDCl}_{3}\)): -0.68, -0.78, -1.00, -1.17, -2.61, -2.71.

Deprotection of compound (23)

Procedure A: Compound 23 (204 mg, 0.085 mmol) was deprotected and purified in a similar way as described for the preparation of compound 15 to give compound 25. Yield: 1641 \( A_{260} \) o.d. units (39 %). \(^{31}\text{P-NMR (D}_{2}\text{O}): -0.83, -0.90, -1.29.\)

Procedure B: Compound 23 (204 mg, 0.085 mmol) was deprotected and purified in the similar way as preparation of compound 14 to get compound 24. Yield: 1636 \( A_{260} \) units (41 %). \(^{31}\text{P-NMR (D}_{2}\text{O): -0.83, -0.98, -1.27.}\)
Regiospecific synthesis of branched tetranucleotides

The authors gratefully acknowledge financial support from the Swedish Board for Technical Development and the Natural Science Research Council and thank Ms. Ingegärd Schiller for her excellent secretarial assistance.

REFERENCES


