Pergamon

Tetrahedron Letters 40 (1999) 4673–4676

TETRAHEDRON LETTERS

Synthesis and enzymatic incorporation of Morpholino thymidine-5'-triphosphate in DNA fragments

Florence Marciacq[+], Sylvie Sauvaigo[+], Jean-Paul Issartel[++], Jean-François Mouret[+++], and Didier Molko[+]*

[+]*Département de Recherche Fondamentale sur la Matière Condensée - Service de Chimie Inorganique & Biologique Laboratoire des Lésions des Acides Nucléiques.*
[++]*Département de Biologie Moléculaire & Structurale - Laboratoire de BioEnergétique Cellulaire et Pathologique.*
[+++]*Société Génome Express S.A.*
Commissariat à l'Energie Atomique, 17, rue des Martyrs, 38054 Grenoble cedex 9, France

Received 18 March 1999; accepted 27 April 1999

Abstract: 4-(carboxymethyl)-2-(thymidin-9-yl)-6-(hydroxymethyl)morpholine-6-triphosphate (morpholino thymidine-5'-triphosphate) was synthesised from 1-(β-D-*erythro*-pentofuranosyl) thymine. It was fully characterised by NMR, UV and mass spectrometry. Taq polymerase enzymatic incorporation of this nucleotide analogue into DNA fragments was investigated. Morpholino thymidine-5'-triphosphate was incorporated in a base-specific process and acted as a novel chain terminator in DNA sequencing, similarly to the corresponding dideoxynucleotide.

DNA Sequence analysis is one of the most expanding technique of molecular biology nowadays. The most popular is the chain termination method developed by Sanger *et al.*[1]. In this method a short synthetic oligonucleotide, the primer (usually a 17 to 25-base oligonucleotide) is annealed to a DNA template, of unknown sequence. In the presence of deoxynucleotides and chain terminators, a DNA polymerase catalyses the elongation of the primer complementary to the template sequence. DNA elongation proceeds by incorporation of natural deoxynucleotides, whereas incorporation of a chain terminator stops the elongation. The result of these reactions is a mixture of randomly terminated chains. Size-analysis of these fragments allows to decipher the sequence of the DNA template.

Present limitations in DNA sequencing rely on the difficulties to read long sequence stretches in a short time and on the cost of a sequencing reaction. Originally, the reaction developed by Sanger *et al.* involved the use of dideoxynucleotides as chain-terminators. Whereas important efforts have been made towards the reduction of the cost of sequencing reactions, few concern the replacement of expensive dideoxynucleotides by new chain terminators[2,3]. Improvements deal with the use of fluorescent labels[4-6] or development of novel methods for automated DNA sequencing[7-9]. We report herein the synthesis of morpholino thymidine-5'-triphosphate, the first compound of a novel kit of

Fax : 33 (0)4 76 88 50 90 - E-mail : molko@drfmc.ceng.cea.fr

0040-4039/99/$ - see front matter
PII: S0040-4039(99)00842-4

chain terminators for DNA sequencing reactions. We show that it was recognised by the Taq Polymerase and specifically stopped DNA-elongation at complementary adenine positions of the matrix.

The chemical structure of morpholino nucleotides (Figure 1) is closed to that of deoxynucleotides. The nucleobase is conserved and most atoms of the deoxy sugar ring are present. Similarly to the dideoxynucleotides, the absence of a C3' hydroxyl function rules out the possibility of grafting another nucleotide during the polymerisation process. Morpholino nucleotides are characterised by the presence of an organic function on the nitrogen atom of morpholine which offers the possibility of linking fluorophores at the ribose-equivalent moiety of the molecule. Such a chemical modification of morpholine avoids the usual derivatization of the nucleobase, a process that may alter enzymatic affinity or binding selectivity.

To test whether this class of compounds could be used in sequencing reactions, morpholino thymidine-5'-triphosphate was synthesised according to Figure 1.

1-(β-D-*erythro*-pentofuranosyl)thymine **1** was synthesised as previously reported by Vorbrüggen *et al.*[10-12] and the synthesis of the morpholine ring leading to 4-(carboxymethyl)-2-(thymidin-9-yl) morpholine was inspired from the method of Erlanger and Beiser[13].

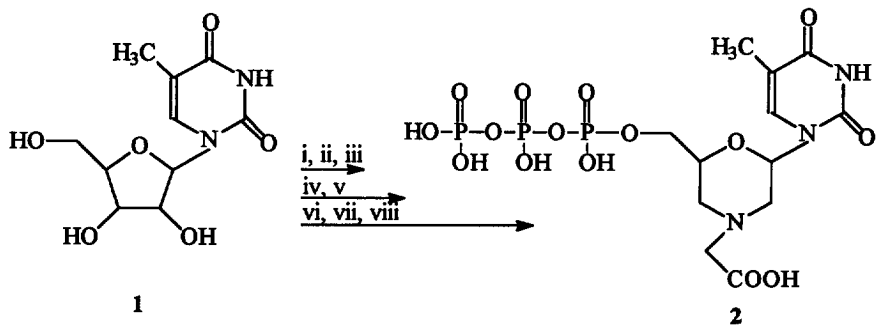

Figure 1: Reaction scheme of the synthesis of the morpholino thymidine-5'-triphosphate i) $NaIO_4$. ii) H_2N-CH_2-COOH, pH 9-10. iii) $NaBH_4$. iv) $POCl_3$/Imidazole. v) H_2O. vi) Carbonyldiimidazole. vii) $P_2O_7(Bu_3NH)_4$. viii) H_2O

In a second step, the corresponding 5'-monophosphate was prepared using 1.5 eq. of phosphorous oxychloride as a phosphorylating reagent in the presence of imidazole. This compound was purified by HPLC in a 57% yield. This product was further characterised by ^1H and ^{31}P NMR. The latter analysis showed a singlet at 1.74 ppm in agreement with the expected chemical shift of a monophosphate. Finally the triphosphate **2** was prepared according to Hoard and Ott[14]: The tributylammonium salt of the monophosphate was treated with 5 eq. of carbonyldiimidazole in anhydrous dimethylformamide, methanol was added to destroy the excess of carbonyldiimidazole and 5 eq. of tributylammonium pyrophosphate were added. The expected triphosphate **2** was purified in a 27% yield using flash chromatography on an ion-exchange phase (DEAE Sepharose Fast Flow, Pharmacia Biotech). Its purity was checked using HPLC (Figure 2) Its ^1H NMR

spectrum[15] was consistent with the structure of a morpholino-nucleoside[16] and ^{31}P NMR spectrum revealed three signals corresponding to the triphosphate chain[17]. Finally, mass spectrometry (MH^+ at 554.48 $g.mol^{-1}$) confirmed the structure of the morpholino thymidine-5'-triphosphate **2**.

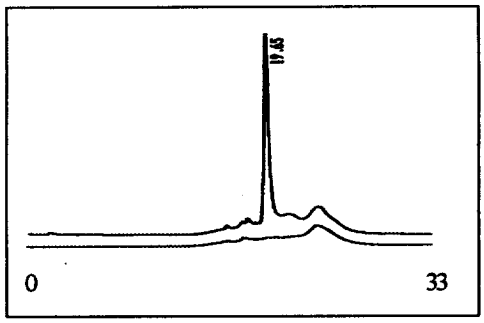

Figure 2: Ion exchange HPLC purity control of the morpholino thymidine-5'-triphosphate **2**.
Column: Polyvinyldiimidazole (PVDI 1000/5 - Société Française Chromato. Colonne) 100 x 4.6 mm, 5 µm particles. Solvent: 25 to 330 mM triethylammonium bicarbonate at 1 mL/min. (20 min. linear gradient).

Compound **2** was tested in sequencing reactions as chain terminator with fluorescent primers and Taq Polymerase (Applied Biosystems, Perkin-Elmer, Foster City, CA, USA) using Bluescript plasmid DNA as a matrix (Stratagene, La Jolla, CA, USA). Sequencing results were analysed with an Applied Biosystems 377 DNA sequencer on a sieving denaturing (urea 7M) matrix of cross-linked polyacrylamide. Two sets of reactions were carried out: one with **2**, the other one with the commercially available dideoxythymidine triphosphate (ddTTP), as the standard reaction. Two concentrations of the morpholinothymidine-5'-triphosphate (500 and 200 µM) were checked. Specific incorporation of **2** was obtained even at the lower concentration (Figure 3).

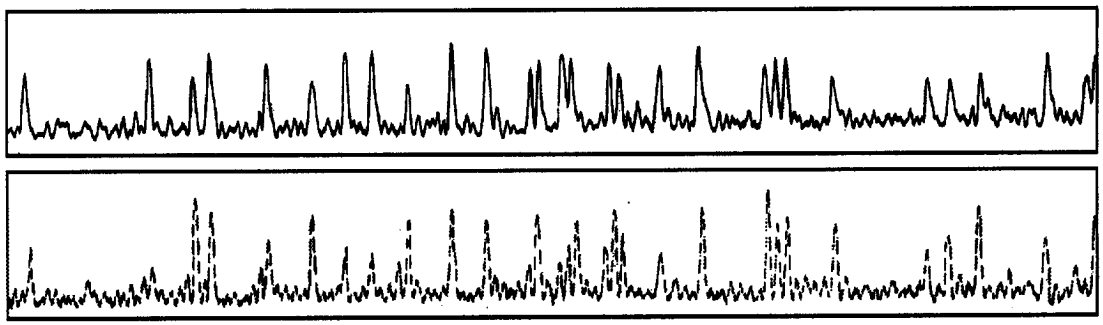

Figure 3: Electropherogramms of the sequencing reactions. Peaks correspond to the sequencing fragments of various lengths. The specific incorporation of 200 µM morpholino thymidine-5'-triphosphate (lower curve) and 250 µM ddTTP (upper curve) was analysed in the 250 to 400 bases region of the template.

In conclusion, as compared with ddTTP, morpholino thymidine-5'-triphosphate was specifically incorporated by the Taq polymerase complementary to adenine and stopped the strand elongation. Morpholino thymidine-5'-triphosphate appears as a suitable substrate for a novel DNA sequencing kit. The three other bases are already under characterisation in a full DNA sequencing reaction.

Acknowledgements:

This work was supported by "Le Ministère de l'Education Nationale de la Recherche & de la Technologie": A.C.C. S.V. n°14.

References and Notes:

1. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. Natl. Acad. Sci. U.S.A.*, **1977**, *74*, 5463-5467.
2. Marx, A., Mc Williams, M. P., Bickle, T. A., Schwitter, U. & Giese, B. *J. Am. Chem. Soc.*,**1997**, *119*, 1131-1132.
3. Wojczewski, C., Faulstich, K & Engels, J. W. *Nucleosides & Nucleotides*, **1997**, *16*, 751-754.
4. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H. & Hood, L. E. *Nature*, **1986**, *321*, 674-679.
5. Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., Zagursky, R. J., Cocuzza, A. J., Jensen, M. A. & Baumeister, K. *Science*, **1987**, *238*, 336-341.
6. Rosenblum, B. B., Lee, L. G., Spurgeon, S. L., Kahn, S. H., Menchen, S. M., Heiner, C. R. & Chen, S. M. *Nucl. Ac. Res.*, **1997**, *25*, 4500-4504.
7. Maeda, M., Patel, A. D. & Hampton, A. *Nucl. Ac. Res.*,**1977**, *4*, 2843-2853.5.
8. Swerdlow, H. & Gesteland, R. *Nucl. Ac. Res.*, **1990**, *18*, 1415-1419.
9. Carrilho, E., Ruiz-Martinez, M. C., Berka, J., Smirnov, I., Goetzinger, W., Miller, A. W., Brady, D. & Karger, B. L. *Anal. Chem.*, **1996**, *68*, 3305-3313.
10. Vorbrüggen, H., Krolikiewicz, K. & Bennua, B. *Chem. Ber.*, **1981**, *114*, 1234-1255.
11. Vorbrüggen, H. & Hofle, G. *Chem. Ber.*, **1981**, *114*, 1256-1268.
12. Vorbrüggen, H. & Bennua, B. *Chem. Ber.*, **1981**, *114*, 1279-1286.
13. Erlanger, B. F. & Beiser, S. M. *Proc. Natl. Acad. Sci. U. S. A.*, **1964**, *52*, 68-74.
14. Hoard, D. E. & Ott, D. G. *J. Am. Chem. Soc.*, **1965**, *87*, 1785-1788.
15. ^{1}H NMR (D_2O) data for **2**: δ (ppm) 7.74 (s, 1H, H6), 5.92 (dd, 1H, H1'), 4.25 (m, 1H, H4'), 4.15 (m, 2H, H5', H5"), 3.81 (s, 2H, CH_2 glycine), 3.54 (dd, 1H, H2"), 3.10 (dd, 1H, H3"), 2.56 (t, 1H, H2'), 2.45 (t, 1H, H3'), 1.95 (s, 3H, CH_3 thymine).
16. Girault, I., Shuker, D. E. G., Cadet, J. & Molko, D. *Bioconjugate Chem.*, **1996**,7,445-450.
17. ^{31}P NMR (D_2O) data for **2**: δ (ppm) -10.03 (d, 1P, γP), -10.88 (d, 1P, αP), -22.65 (t, 1P, βP).