Tetrahedron. Vol. 30, pp. 3073 to 3079. Pergamon Press 1974. Printed in Great Britain

ACTIVE MONOMERIC NUCLEOTIDE INTERMEDIATE IN THE OLIGONUCLEOTIDE SYNTHESIS

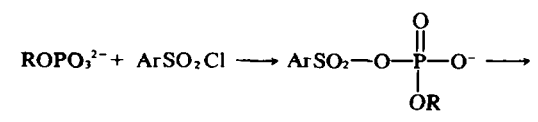
D. G. KNORRE, A. V. LEBEDEV, A. S. LEVINA, A. I. REZVUKHIN and V. F. ZARYTOVA

Institute of Organic Chemistry of Siberian Division of the Academy of Sciences of the USSR, Novosibirsk 630090, USSR

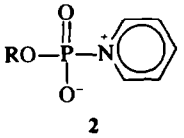
(*Received in the UK* 28 *January* 1974; *Accepted for publishing* 21 *February* 1974)

Abstract—The interaction of 3' - O - acetylthymidine - 5' - phosphate (*p*T–Ac) and *p* - nitrophenylphosphate (*p*PhNO$_2$) with triisopropylbenzenesulphonyl chloride (TPS) in pyridine was studied by the pulsed NMR spectroscopy on ^{31}P nuclei. The esters investigated were shown to convert to the corresponding disubstituted pyrophosphates, trisubstituted tripolyphosphates and compounds showing a singlet displaced 5–8 ppm from the original ester. The latter are the main final products of the interaction of nucleotide and *p*PhNO$_2$ with TPS. The investigation of the chemical conversion of a *p*T–Ac derivative with a 5·1 ppm shift showed that this product is a powerful phosphorylating reagent. It was shown that this active derivative contains one P atom in accordance with the structure of monomeric metaphosphoric acid ester proposed by Todd for these compounds. The reaction of monomeric metaphosphate with 5'-O-tritylthymidine results in formation of a compound, containing stoichiometric amounts of *p*T–Ac and 5' - O - tritylthymidilyl - (3'→5') - 3' - O - acetylthymidine residues. This compound was identified as a trisubstituted P^1 - 5' - O - trityl - P^1, P^2 - bis - (3' - O - acetylthymidine) - pyrophosphate.

The mechanism of the internucleotide bond formation in the course of the oligonucleotide synthesis is not yet established. Two main hypotheses have been proposed . According to Todd[1] treatment of nucleoside phosphate $ROPO_3^{2-}$ by dicyclohexylcarbodiimide (DCC) or arylsulphonyl chlorides results in monomeric nucleoside metaphosphate (**1**) formation which doubtless would be a powerful



1



2

R = nucleoside

phosphorylating agent for anions, amines and alcohols. The presence of metaphosphate in pyridine solution in the form of a pyridinium derivitave was also discussed.[2]

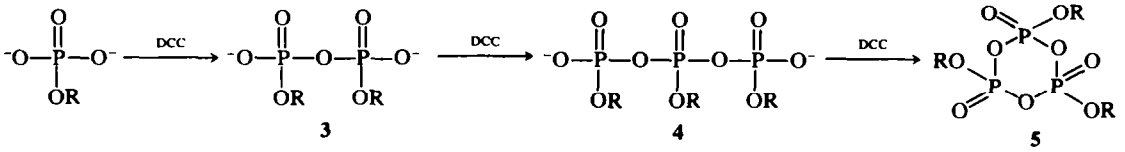
According to Weimann and Khorana[3] trinucleoside trimetaphosphates are formed as a result of stepwise reaction of nucleotide with DCC

Attempts to isolate trimetaphosphate **5** in the case of 3' - O - acetylthymidine - 5' - phosphate (*p*T–Ac) were unsuccessful, although inorganic trimetaphosphate (**5**, R=H) was obtained after heating of the reaction product of benzyl phosphate with DCC.[3]

The present paper deals with ^{31}P NMR investigation of the reaction of *p*T–Ac with triispropylbenzenesulphonyl chloride (TPS) and the reaction of the active phosphorylating compound formed with nucleoside component 5' - O - tritylthymidine (Tr–T). Due to significant differences between chemical shifts for ^{31}P of phosphoryl residues forming ester and anhydride bonds (Table 1) it was possible to identify compounds **3** and **4** in the reaction and to discriminate between proposed structures **1**, **2** and **5** of active nucleotide derivative in favour of Todd's proposal.

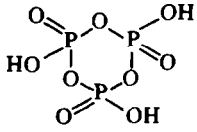
RESULTS

Stepwise addition of TPS to *p*T-Ac solution was used to simplify interpretation of the ^{31}P NMR spectra of the reaction mixture. Addition of 0·5 eqv. TPS to *p*T–Ac in pyridine solution results in the appearance in ^{31}P NMR spectrum of the reaction mixture of a new singlet signal with chemi-



3 **4** **5**

Table 1. Chemical shifts and Coupling constants of some phosphoric acid derivatives[4]

Compound	Chemical shift, ppm	Coupling constants, cps
$(O)P(OH)_3$	0·0	
$(HO)_2P(O)OP(O)(OH)_2$	10·6–11·8	
$(HO)_2P_\alpha(O)OP_\beta(O)(OH)OP_\alpha(O)(OH)_2$	$\alpha = 11·5$ $\beta = 23·9$	$P_\alpha P_\beta = 16·7$
$(HO)_2P_\alpha(O)OP_\beta(O)(OH)OP_\beta(O)(OH)OP_\alpha(O)(OH)_2$	$\alpha = 11·5$ $\beta = 23·9$	$P_\alpha P_\beta = 16·7$ $P_\beta P_\beta = 15·9$
	20·7	
$(OH)P(O)(OC_2H_5)_2$	0·0–1·3	
$(OC_2H_5)_2P(O)OP(O)(OC_2H_5)_2$	12·5–13·8	
$(OH)P(O)(OC_6H_5)_2$	9·4–12·7	
$(OC_6H_5)_2P(O)OP(O)(OC_6H_5)_2$	23·3	
$(OC_6H_5)_2P_\alpha(O)OP_\beta(O)(OC_6H_5)OP_\alpha(O)(OC_6H_5)_2$	$\alpha = 26·6$ $\beta = 35·6$	$P_\alpha P_\beta = 16$
pT	−3·35	
$p_\alpha Tp_\beta T$	$\alpha = -1·9$ $\beta = 1·8$	
$O(p\text{T–Ac})_2$ (this paper)	10·3	

cal shift $\delta = 10·3$ ppm (Fig 1a) identical with that of synthetically prepared P^1, P^2 - bis - (3′ - O - acetylthymidine - 5′) - pyrophosphate $O(p\text{T–Ac})_2$. A new compound with electrophoretic mobility $E_f = 0·76$ identical with that of $O(p\text{T–Ac})_2$ was found in the reaction mixture. Reaction proceeds for 60 min. The yield of $O(p\text{T–Ac})_2$ is not quantitative probably due to the presence of traces of water in the reaction mixture.

The next addition of 0·5 eqiv. TPS results in appearance in the ^{31}P NMR spectrum of the reaction mixture of additional signals—two nearly overlapping doublets with $\delta = 11·5$ ppm and one triplet (or double of doublets) with $\delta = 21·6$ ppm both with spin–spin coupling constant $J_{P\text{-}O\text{-}P} = 17$ Hz (Fig 1b). The intensities of multiplets (2:1) and this J value permit one to assign these signals to one compound **A**. The spectrum observed is in a good agreement with that expected for P^1, P^2, P^3 - tris - (3′ - O - acetylthymide - 5′) - tripolyphosphate (**4a**).

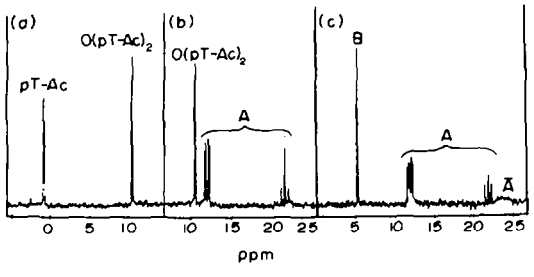


Fig 1.

Electrophoresis of the reaction products precipitated from the mixture with ether shows the presence of the nucleotide derivative with $E_f = 0·67$ in agreement with value obtained for **4a** in Ref 3. Treatment of the mixture containing A with water results in formation of a mixture of starting nucleotide and $O(p\text{T–Ac})_2$ in stoichiometric amounts. Including this stage reaction proceeds in a good agreement with the mechanism proposed in Ref 3.

Addition of a next 0·5 eqiv. TPS, the ^{31}P NMR spectrum of the mixture exhibited a new singlet signal with $\delta = 5·1$ ppm (compound B) (Fig 1c). No further reactions were observed in the absence of nucleoside component or other nucleophiles added. The chemical shift of **B** differs greatly from one expected for cyclic metaphosphate **5**. The latter has to be ~20 ppm since inorganic trimetaphosphate has $\delta = 20·7$ ppm and as it may be seen (Table 1) for corresponding orthophosphate and pyrophosphate thymidine esters the change of the chemical shift due to introduction of the thymidine residue never exceeds 3–4 ppm.

To determine the number of P atoms in compound **B** the mixture of pT–Ac and p-nitrophenyl phosphate ($pPhNO_2$) was treated with TPS. In a separate experiment it was shown that reaction of TPS with $pPhNO_2$ proceeds via the same steps as with pT–Ac, i.e. via pyrophosphate, tripolyphosphate and compound $\mathbf{B}_1$ with singlet ^{31}P NMR signal at $\delta = 12·2$ ppm being formed consecutively. The reaction time and, consequently, the reactivity of p-nitrophenyl derivatives are of the same order

of magnitude as those of 3′ - O - acetylthymide derivatives. Therefore it should be expected that mixed compounds with both 3′ - O - acetylthymidine and *p* - nitrophenyl residues would be present in measurable amounts in all types of compounds containing more than one P atom. Due to significant difference of δ values for 3′-O-acetylthymidine and *p* - nitrophenyl derivatives (5–7 ppm) P atoms of these mixed compounds should be unequivalent and splitting of the signals should be observed for all these derivatives.

As it may be seen (Figs 2a, 2b) such splitting does take place for the mixture of pyrophosphates and tripolyphosphates. After addition of 0·5 eqiv. TPS to the mixture of *p*PhNO$_2$ and *p*T–Ac two doublets with constant J = 17 Hz may be seen in pyrophosphate range of ^{31}P NMR spectrum besides singlet signals of symmetric pyrophosphates O(*p*PhNO$_2$)$_2$ and O(*p*T–Ac)$_2$. These signals may be related to P^1 - *p* - nitrophenyl - P^2 - (3′ - O - acetyltyhymidine - 5′) - pyrosphosphate O$\begin{smallmatrix}/\ p\text{PhNO}_2\\ \backslash\ p\text{T–Ac}\end{smallmatrix}$. The spectrum of a mixture of tripolyphosphates appeared after addition of a next 0·5 equiv. TPS to the same reaction mixture also has some signals additional to signals of **4a** and P^1, P^2, P^3 - tris - (*p* - nitrophenyl) - tripolyphosphate thus indicating accumulation of some amounts of mixed tripolyphosphates. However in the same reaction mixture the signals of compounds **B** and **B**$_1$ are singlets (Fig 2c). Consequently, both **B** and **B**$_1$ contain one P atom per molecule and therefore may be regarded as some phosphate ester derivatives presumably metaphosphate esters.

The spectra of **B** and **B**$_1$ recorded without heteronuclear spin-spin decoupling ^{31}P-{^{1}H} are represented in Fig 2d. It may be seen that spectrum of **B** is triplet with J = 8·5 Hz characteristic for spin–spin coupling P–O–C–H. Spectrum of **B**$_1$ remains unsplit. This fact disagrees with the structure **2** since the spin–spin interaction with J ~ 10 Hz should be expected for P–$\overset{+}{N}$ = C–H system isoelectronic with P–C=C–H system in the compounds containing P–Ph bond.[5] Therefore only structure **1** of metaphosphate ester agrees with ^{31}P NMR data for active nucleotide derivative. Cer-

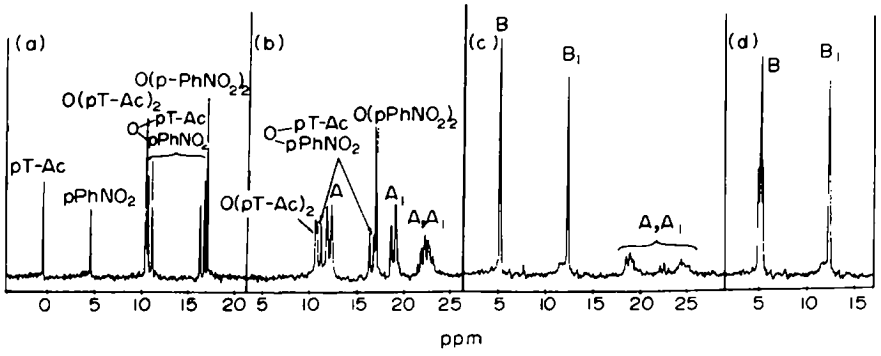


Fig 2.

Table 2. The composition of the reaction mixtures after nucleophilic agents addition to the monomeric *p*T–Ac derivative solution (compound **B**). The data are represented in mole per cent. The analysis was carried out by ^{31}P NMR spectra registered 2–5 min after nucleophile addition

reagents	H_2O	MeOH	*p*T–Ac		O(*p*T–Ac)$_2$	*p*PhNO$_2$
mole ratio reagent: B	200:1	100:1	3:2	3:1	0·6:1	1:1
compounds						
*p*T–Ac	95	—	26	71	—	—
O(*p*T–Ac)$_2$	5	5	52	29	25	1·5
A	—	—	22	—	75	11*
Me–*p*T–Ac	—	95	—	—	—	—
O(*p*PhNO$_2$)$_2$	—	—	—	—	—	8·5
*p*PhNO$_2$	—	—	—	—	—	21
O$\begin{smallmatrix}/\ p\text{PhNO}_2\\ \backslash\ p\text{T–Ac}\end{smallmatrix}$	—	—	—	—	—	58

*mixture of tripolyphosphates.

tainly it cannot be excluded that pyridine molecules take part in some specific stabilization of structure without covalent binding to metaphosphate ester.

The reactivity of compound **B** towards nucleophiles investigated is in full accordance with the proposed monomeric structure. The addition of the excess of water or MeOH results in a quantitative formation of the *p*T–Ac or of its methyl ester. In the presence of the excess of *p*T–AcO(*p*T–Ac)$_2$ is obtained in a nearly quantitative yield. With a slight excess of *p*T–Ac a mixture of O(*p*T–Ac)$_2$ and tripolyphosphate **a** is formed, the latter presumably being a product of a secondary reaction of initially formed O(*p*T–Ac)$_2$ with unreacted **B**. The addition of *p*PhNO$_2$ results in formation of unsymmetric pyrophosphate $O\begin{matrix}\diagup\, p\text{hNO}_2\\ \diagdown\, p\text{T–Ac}\end{matrix}$ as a main product (^{31}P NMR data). The results of the **B** reactivity investigation are presented in Table 2.

All reactions with strong nucleophiles proceed immediately and several minutes after addition of a nucleophile no traces of **B** signal is observed in the ^{31}P NMR spectra of the reaction mixtures.

The addition of nucleoside component Tr–T to pyridine solution of **B** results in a relatively slow transformation of the compound **B** with δ = 5·1 ppm to a compound **C** which shows multiplet in ^{31}P NMR spectrum with δ = 11–13 ppm (pyrophosphate range) (Fig 3a). The same multiplet is formed immediately after the addition of 5′ - O - trithylthymidilyl - (3′ → 5′) - 3′ - O - acetylthymidine (Tr–T*p*T–Ac) to the solution of **B** (Fig 3b). The addition of the excess of water to the solution of **C** results in the formation of stoichiometric amounts of Tr–T*p*T–Ac and *p*T–AC. Therefore the structure of the trisubstituted pyrophosphate P^1, P^2 - bis - (3′ - O - acetylthymidine - 5′) - P^2 - (5′ - O - tritylthymidine) - pyrophosphate may be attributed to the compound **C** which is formed in the first case due to two consecutive reactions the first being a rate determining step. Due to the small difference in the shielding of P atoms in the compound **C** it

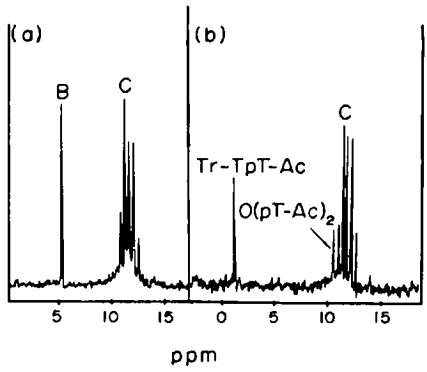


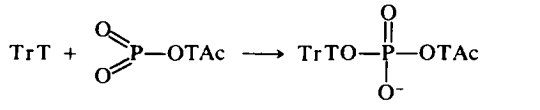
Fig 3.

should be regarded as AB-system. More complicated picture of the multiplet may be a result of the existence of two diastereoisomers due to the asymmetric P atom. The existence of two diastereoisomers for dinucleosidealkylphosphotriesters was demonstrated in Refs 6 and 7. The intensities of multiplet lines are calculated under the assumption that the compound **C** is an equimolar mixture of two isomers. These intensities are in satisfactory agreement with experimental values.

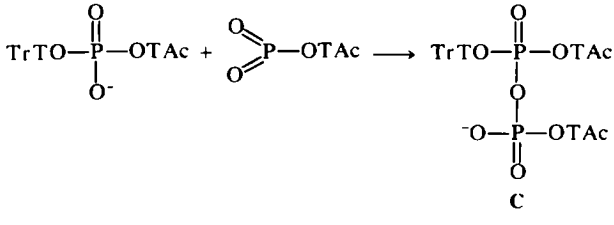
Table 3. Experimental and calculated multiplet lines intensities of compound **C**

Lines	1	2	3	4	5
Chemical shifts, ppm	10·7	11·1	11·5	12·0	12·5
Exp. intensities, %	10·7	35·6	21·9	23·9	7·8
Calc. intensities, %	12·2	37·3	20·1	22·2	7·8

To elucidate whether **C** may be regarded as an active phosphorylating reagent, an excess of Tr–T was added to the solution containing 70 μmole C. After 20 h 110 μmole Tr–T*p*T–Ac was found in the reaction mixture. Since **C** contains one Tr–T*p*T–Ac residue per molecule, additional 40 μmoles of Tr–T*p*T–Ac was formed as a result of the interac-



B



C

Table 4. ^{31}P Chemical shifts of the investigated compounds related to 85% H_3PO_4

Compound	Chemical shift, ppm
1. 3′ - 0 - Acetylthymidine - 5′ - phosphate (*p*T–Ac)	− 1·2
P^1, P^2 - bis - (3′ - 0 - acetylthymidine - 5′) - pyro-Phosphate (O(*p*T–Ac)$_2$)	10·3
P^1, P^2, P^3 - tris - (3′ - 0 - acetylthymidine - 5′) - tri-Polyphosphate **A**	11·6
	21·6
Compound **B**	5·1
2. *p*-Nitrophenylphosphate (*p*PhNO$_2$)	4·2
P^1, P^2 - bis - (*p*-nitrophenyl) - pyrophosphate (O(*p*PhNO$_2$)$_2$)	17·0
P^1, P^2, P^3 - tri - (*p*-nitrophenyl)	19·2
- tripolyphosphate (A_1)	29·4
Compound B_1	12·2
3. 5′-O-Tritylthymidilyl-(3′ → 5′)-3′-O-acetylthymidine (Tr-TpT-Ac)	1·0
4. 3′ - 0 - Acetylthymidine - 5′ - 0 - methylphosphate (Me–*p*T–Ac)	− 1·0

tion of *p*T–Ac fragment of **C** with Tr–T. Chemical shifts of the investigated compounds are represented in Table 4.

DISCUSSION

The data obtained demonstrate that the main product of the interaction of TPS with *p*T–Ac is a very reactive phosphorylating reagent having one P atom per molecule. These data exclude the trimetaphosphate **5** as a main intermediate in the oligonucleotide synthesis. Certainly, this does not prove that the phosphorylating compound **B** has structure **1** of monomeric metaphosphate but the proposal seems to be most reasonable. Compound **B** does not contain an arylsulphonyl residue since it was shown in our work with V. K. Potatov and Z. A. Shabarova (the data will be published separately) that the same compound is formed in the course of reaction of *p*T–Ac with polymeric cross-linked arylsulphonyl chloride. The formation of metaphosphate in the course of interaction of TPS with trisubstituted tripolyphosphate should most probably be regarded as a result of a rapid elimination of the product of arylsulphonyl residue addition to tripolyphosphate **8** (see Scheme below, reaction 7, 8). The anion of a strong acid **7** is formed identical with the expected product of the interaction of pyrophosphate with TPS (reaction 4 of the Scheme below). This intermediate should undergo similar elimination (reaction 5) with production of the anion **6**, also capable of elimination with metaphosphate formation. Anion **6** is identical with the expected product of the interaction of the starting nucleotide with TPS.

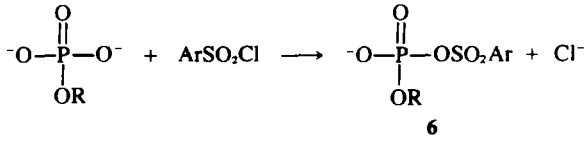
Therefore it is reasonable to suggest that the formation of pyrophosphate and tripolyphosphate proceeds via the formation of metaphosphate which reacts then either with correspondingly starting nucleotide or pyrophosphate. Due to high nucleophility of the latter species metaphosphate reacts rapidly and is not accumulated during the first period of the reaction in measurable amounts. The total reaction scheme may be represented as follows.

Tripolyphosphate is probably a much poorer nucleophile than pyrophosphate to a marked extent. However, formation of amounts of some products with δ in the region 22–24 ppm, probably longer polyphosphates or cyclic polymetaphosphates, was noticed in some experiments (Fig 1, signals Ā).

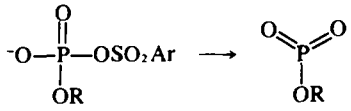
In the most simple case of the oligonucleotide synthesis through the use of protected nucleoside and mononucleotide, metaphosphate may be regarded as an immediate precursor of dinucleoside-phosphate. The latter reacts immediately with a next molecule of metaphosphate forming compound **C**. This compound is also phosphorylating reagent and may react with the excess of nucleoside to give additional amounts of dinucleoside phosphate. Compound **C** is similar to P^1,P^1-diphenyl-P^2 - nucleoside pyrophosphates widely used in nucleotide chemistry.[8] Our data do not allow one to discriminate between two possible mechanisms of this reaction—via direct attack of OH group of the nucleoside component at P atom forming anhydride bond or via reversible dissociation of **C** to Tr–TpT–Ac and metaphosphate ester reacting irreversible with mucleoside.

In any case, in the most typical oligonucleotide synthesis where one or both components contain phosphodiester residues no accumulation of a free metaphosphate should be expected and trisubstituted pyrophosphates must be the main active intermediates presented in the reaction mixture in significant amounts.

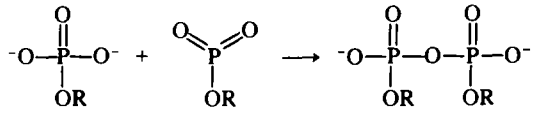
Reaction scheme



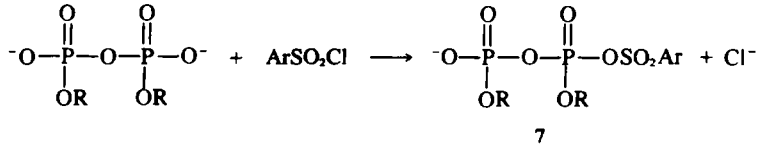
(1)



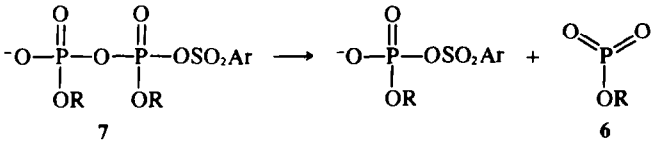
(2)



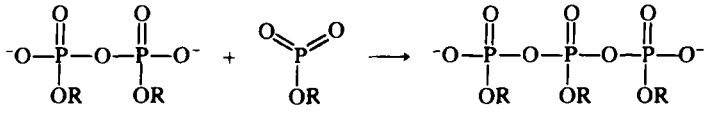
(3)



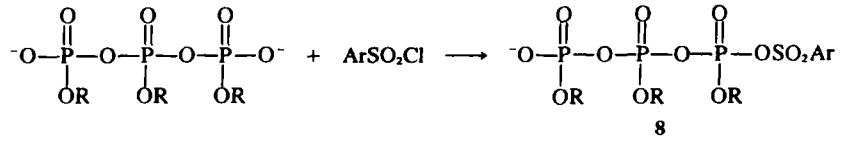
(4)



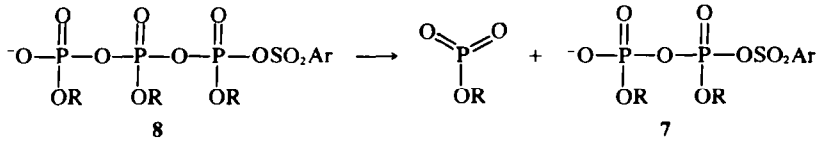
(5)



(6)



(7)



(8)

EXPERIMENTAL

3′ - O - acetylthymidine - 5′ - phosphate (*p*T–Ac), 5′ - O - tritylthymidine (Tr–T), P^1, P^2 - bis - (3′ - O - acetylthymidine - 5′) - pyrophosphate (O(*p*T–Ac)$_2$) and 5′ - O - tritylthymidilyl - (3′→5′) - O - acetylthymidine (Tr–T*p*T–Ac) were prepared according to Ref 9. Triisopropylbenzenesulphonyl chloride was prepared according to Ref 10. Pyridine with water content less than 0·05% stored above molecular sieves 4A was used as a solvent. The reactions were carried out in pyridine soln at 28° in 10 mm tubes (1·5 ml of the reaction mixture). The ^{31}P NMR spectra were taken with a Bruker HX-90 pulse spectrometer operating at 36·43 MHz. Fourier transform was performed using Bruker B-NC 12-FFT computer after 100–500 accumulations. Pulse width was 15 μsec and the time between pulses was 0·7 sec. D_2O was used as an external standard for the stabilization of the resonance conditions. The most spectra are recorded with heteronuclear spin–spin decoupling ^{31}P-{^1H}. The chemical shifts are reported in ppm related to external 85% H_3PO_4 and are judged to be accurate ±0·1 ppm. Initial concentration of P atoms was 0·15 M in all experiments.

Electrophoretic mobilities E_f of the compounds were measured relative *p*T–Ac in 0·05 M triethylammonium bicarbonate at pH 7·5 using apparatus for high voltage electrophoresis VEFA-5-0·35 (SKD BPA, USSR).

*p*T–Ac, O(*p*T–Ac)$_2$, Tr–T*p*T–Ac, 3′ - O - acetylthymidine - 5′ - methylphosphate were identified in the reaction according to their ^{31}P NMR spectra (δ values), E_f and R_f in the system ethyl alcohol-1M ammonium acetate, pH 7·5 (5:2, v/v). P^1, P^2 - bis - (*p* - nitrophenyl) - pyrophosphate and P^1, P^2, P^3 - tris - (*p* - nitrophenyl) - tripolyphosphate were identified according to their ^{31}P NMR spectra similar to corresponding 3′ - O - acetylthymidine derivatives.

REFERENCES

[1] A. Todd, *Proc. Natl. Acad. Sci. U S A* **45**, 1389 (1959)
[2] A. M. Michelson, *The chemistry of nucleosides and nucleotides.* Academic Press, London and New York (1963)
[3] G. Weimann and H. G. Khorana, *J. Am. Chem. Soc.* **84**, 4329 (1962)
[4] M. M. Crutchfield, C. H. Dungan, J. H. Letcher *et al.*, *Topics in phosphorus chemistry* Vol 5, Interscience p, New York-London-Sydney (1967)
[5] J. W. Emsley, J. Feency and L. H. Sutcliffe, *Progress in nuclear magnetic resonance spectroscopy* Vol 1, Pergamon Press, London-New York (1966)
[6] P. S. Miller, K. N. Fang, N. S. Kondo and Paul O. P. Ts'O, *J. Am. Chem. Soc.* **93**, 6657 (1971)
[7] E. S. Werstiuk and T. Neilson, *Can. J. Chem.* **50**, 1283 (1972)
[8] A. M. Michelson, *Chem. & Ind.* 1267 (1960)
[9] P. T. Gilham and H. G. Khorana, *J. Am. Chem. Soc.* **80**, 6212 (1958)
[10] R. Lohrmann and H. G. Khorana, *Ibid.* **88**, 829 (1966)