Synthesis of Lactose-Operator Gene Fragments by the Improved Triester Method[1]

K. ITAKURA, N. KATAGIRI, AND S. A. NARANG

Division of Biological Sciences, National Research Council of Canada, Ottawa, Canada K1A 0R6

Received August 19, 1974

K. ITAKURA, N. KATAGIRI, and S. A. NARANG. Can. J. Chem. **52**, 3689 (1974).

The synthesis of lactose-operator gene fragments have been achieved by the improved triester methodology and new phosphorylating and condensing reagents.

K. ITAKURA, N. KATAGIRI, et S. A. NARANG. Can. J. Chem. **52**, 3689 (1974).

La synthèse des fragments du gêne opérant de la lactose a été réalisée par la méthode améliorée du triester en utilisant de nouveaux agents de phosphorylation et de condensation.

[Traduit par le journal]

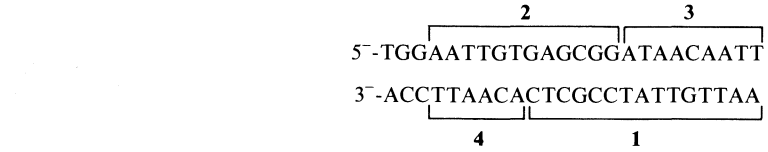
The lactose-operator gene constitutes a unique sequence of 21–24 bases in *E. coli* which has the remarkable property of regulating the biological function of the lactose-operon (1). Recently Gilbert and his co-worker (2) have presented a unique sequence of 24 base pairs (see Scheme 1) by determining the sequence of RNA transcription copies of this fragment. In this communication, we wish to report the application of our improved triester approach for the chemical syntheses of pentadeca- (**1**), dodeca- (**2**), nona- (**3**), and hexanucleotides (**4**) representing the twofold symmetrical region of the lac-operator gene sequence.

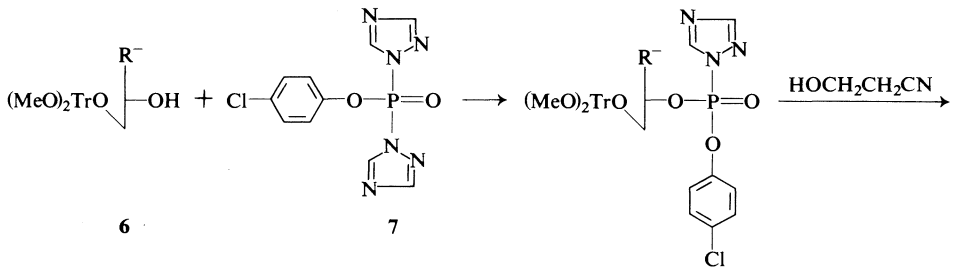
The basic feature of the triester approach as reported by most other laboratories (3–6), involved a "two-step", reaction, *i.e.* phosphorylation followed by subsequent coupling at each synthetic step. However, we observed that the initial phosphorylation did not give a pure 3′-phosphodiester component in quantitative yield thus leading to very complicated reaction mixtures on addition of the second component. We overcame this problem by using 5′-protected nucleosides containing fully protected 3′-phosphate groups **5** as starting material (7). Thus 5′-dimethoxytrityl nucleoside **6** was phosphorylated with bistriazolide of *p*-chlorophenyl phosphate **7**,[2] a new and highly efficient phosphorylating reagent (8). This reaction went to completion in 6 h at room temperature and on subsequent treatment with β-cyanoethyl alcohol, a fully protected mononucleotide **5** was obtained in quantitative yield after isolation by silica gel chromatography. Treatment of **5** with 80% acetic acid afforded 5′-hydroxy 3′-fully protected nucleotide **8**, whereas a very mild alkaline treatment (0.05 *N* sodium hydroxide, dioxane–water) selectively removed the β-cyanoethyl group to give a 5′-protected nucleoside 3′-phosphodiester product **9**. Using triisopropylbenzenesulfonyl chloride (TPS) as condensing reagent and the synthetic path outlined in the scheme, we observed that the yields of oligonucleotides containing purine bases (especially guanine) were consistently low (*ca.* 20%) (9, 10). By introducing two new condensing reagents, mesitylenesulfonyl triazolide (MST) (10) and *p*-nitrobenzenesulfonyl triazolide[3] (m.p. 147–149°) (8) (*p*-NBST),

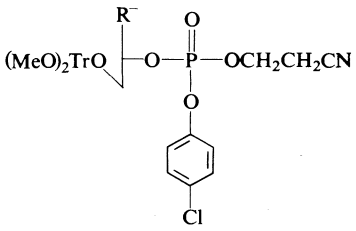
[1]NRCC No. 14252.

[2]It was prepared by condensing *p*-chlorophenyl phosphodichloridate (1 equiv.) with 1*H*-1,2,4-triazole (2 equiv.) in the presence of triethylamine (2 equiv.) in dioxane. Due to extremely labile nature of bistriazolide of *p*-chlorophenyl phosphate, the reaction mixture was used immediately after the removal of triethylammonium hydrochloride.

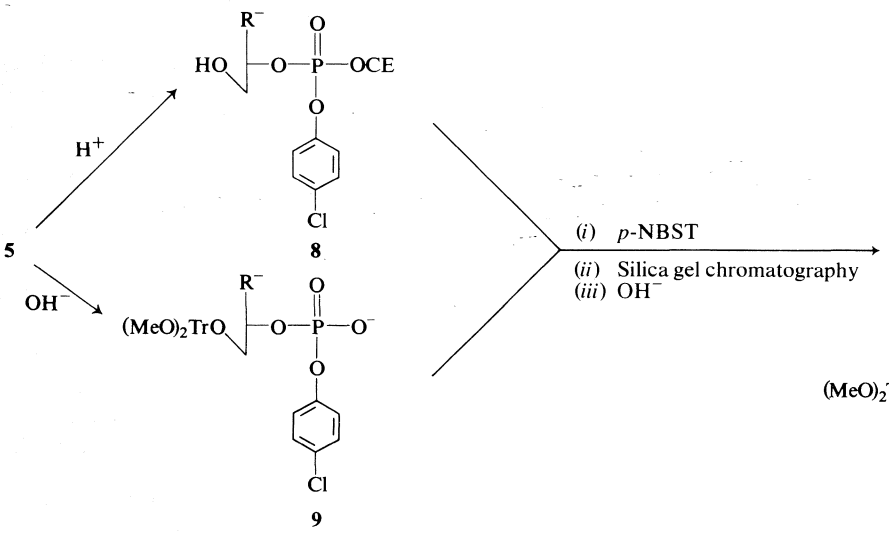
[3]It was prepared in the quantitative yield according to the published procedure (10). It was characterized satisfactorily by elemental and spectral analyses.

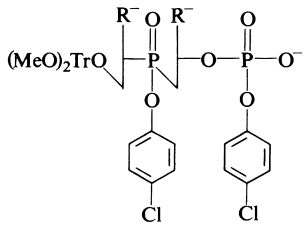






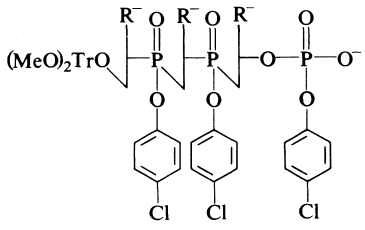
5



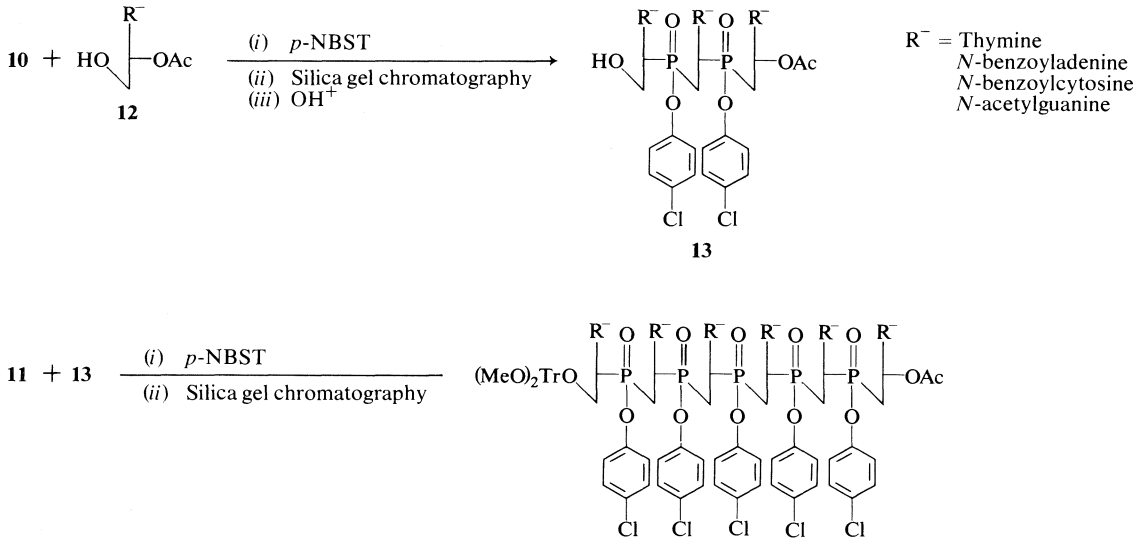


10

10 + 8 → (i) p-NBST; (ii) Silica gel chromatography; (iii) OH^-



11



SCHEME 1. Lactose-operator gene sequence

substantially higher yields were realized at each step. For example, the protected dodecanucleotide [$(MeO)_2$Tr]dbzA-bzA-T-T-acG-T-acG-bzA-acG-bzC-acG-acG(Ac) was synthesized by condensing [$(MeO)_2$Tr]dbzA-bzA-T-T-acG-Tp(Cl-Ph) (0.5 mmol) with dacG-bzA-acG-bzC-acG-acG(Ac) (0.4 mmol) in the presence of *p*-nitrobenzenesulfonyl triazolide[4] (*p*-NBST) (2.5 mmol) for 2 days. The reaction was then taken up in chloroform (25 ml) which was washed with 0.1 *M* triethylammoniumbicarbonate buffer (5 ml × 3) and finally with water (10 ml). The organic layer was dried over anhydrous sodium sulphate and evaporated. The resulting residue was co-evaporated with toluene three times to remove pyridine and applied to silica gel column[5] (10 × 5 cm) in chloroform (10 ml). The column was eluted with a mixture of chloroform–methanol (10%, 1 l). Fractions of 2 ml/5 min were collected at room temperature.[6] The fully protected dodecanucleotide (0.98 g) was isolated from fraction 36–44 in 41%. By using the similar experimental conditions, we obtained the following isolated yields of hexa- (60–65%), nona- (55–60%), and pentadecamer (30–40%) (see Table 1).

The complete deprotection of fully protected dodecanucleotide (0.1 g) was carried out with 0.1 *N* sodium hydroxide – dioxane (100 ml) at room temperature for 48 h. After neutralization with Dowex-50 (pyridinium form), the concentrated residue was treated with concentrated ammonia (50 ml) for 2 days at room temperature and finally with 80% acetic acid (50 ml) for 20 min at room temperature. The deprotected compound was isolated by gel filtration on Sephadex G-25 (SF).

The homogeneity of each unprotected compound was checked rigorously by (*i*) t.l.c. on silica gel using aqueous solvent systems (11); (*ii*) homochromatography of 5′-^{32}P-labeled compound according to Sanger (12) and Wu (13) techniques; (*iii*) complete digestion with snake venom or spleen phosphodiesterase enzymes to their expected ratios of nucleoside to nucleotides. Finally the sequence of each fragment was confirmed by two-dimensional chromatography (electrophoresis on cellulose acetate strip at p*H* 3.5, pyridine–acetate buffer followed by homochromatography on DEAE cellulose thin-layer plates in 2% partially hydrolyzed RNA containing 7 *M* urea) of 5′-labeled oligomers partially digested with snake venom phosphodiesterase enzyme. Each oligomer gave the expected finger-printing pattern (14).

In conclusion, with this improved triester approach and the new phosphorylating and condensing reagents, we have been able to achieve large scale syntheses of longer oligonucleotides of defined sequence in high yields using approximately equimolar ratios of each component.

[4]It did not cause any sulfonation of 5′-hydroxyl group of oligonucleotide. Moreover the condensation reaction using (*p*-NBST) afforded the very clean reaction as compared with TPS.

[5]Thin-layer chromatographic silica gel was used for column chromatography.

[6]Every second fraction was checked on silica gel t.l.c. plates in chloroform–methanol (9:1) solvent.

TABLE 1. Reaction conditions of various condensation steps and the yields

3′-Phosphodiester component		5′-Hydroxyl component			Product			
							Completely deblocked compound	
Identity	Amount (mmol (g))	Identity	Amount (mmol (g))	*p*-NBST[a] (mmol)	Identity	Yield (% (g))	Identity	Yield (%)
$[(MeO)_2Tr]$[b]dbzA-bzC[c]-bzAp(ClPh)	1.1 (2.1)	dbzA-T-T(Ac)	1 (1.3)	3	$[(MeO)_2Tr]$dbzA-bzC-bzA-bzA-T-T(Ac)	62 (2)	A-C-A-A-T-T **4**	71
$[(MeO)_2Tr]$dbzA-T-bzAp(ClPh)	0.5 (0.95)	dbzA-bzC-bzA-bzA-T-T(Ac)	0.48 (1.40)	2	$[(MeO)_2Tr]$dbzA-T-bzA-bzA-bzC-bzA-bzA-T-T(Ac)	57 (1.30)	A-T-A-A-C-A-A-T-T **3**	64
$[(MeO)_2Tr]$bzA-bzA-T-T-acG-Tp(ClPh)	0.5 (1.62)	dacG-bzA-acG-bzC-acG-acG(Ac)	0.4 (1.10)	2.5	$[(MeO)_2Tr]$dbzA-bzA-T-T-acG-T-acG-bzA-acG-bzC-acG-acG(Ac)	41 (0.98)	A-A-T-T-G-T-G-A-G-C-G-G **2**	52
$[(MeO)_2Tr]$dbzA-bzA-T-T-acG-Tp(ClPh)	0.15 (0.49)	dT-bzA-T-bzC-bzC-acG-bzC-T-bzC(Ac)	0.1 (0.43)	0.65	$[(MeO)_2Tr]$dbzA-bzA-T-T-acG-T-T-bzA-T-bzC-bzC-acG-bzC-T-bzC(Ac)	35 (0.26)	A-A-T-T-G-T-T-A-T-C-C-G-C-T-C **1**	45

[a]*p*-NBST signifies *p*-nitrobenzenesulfonyl triazolide, a new condensing reagent used in the present studies.
[b]Abbreviation are as suggested by the IUPAC–IUB Commission (15).
[c]Benzoyl protecting group was found to be stable under the reaction condition using *p*-nitrobenzenesulfonyl triazolide as condensing reagent.

1. F. JACOB and J. MONOD. J. Mol. Biol. **3**, 318 (1961).
2. W. GILBERT and A. MAXAM. Proc. Natl. Acad. Sci. **70**, 3581 (1973).
3. R. L. LETSINGER and K. K. OGILVIE. J. Am. Chem. Soc. **89**, 4801 (1967); R. L. LETSINGER and K. K. OGILVIE. J. Am. Chem. Soc. **91**, 3350 (1969).
4. C. B. REESE and R. SAFFHILL. Chem. Commun. 767 (1968); C. B. REESE and R. SAFFHILL. Chem. Commun. 869 (1971); N. J. CUSACK, C. B. REESE, and J. H. VAN BOOM. Tetrahedron Lett. 2209 (1973).
5. F. ECKSTEIN and I. RIZK. Chem. Ber. **102**, 2326 (1969).
6. T. NEILSON and E. S. WERSTIUK. Can. J. Chem. **49**, 3004 (1971); T. NEILSON and E. S. WERSTIUK. J. Am. Chem. Soc. **96**, 2295 (1974).
7. K. ITAKURA, C. P. BAHL, N. KATAGIRI, J. J. MICHNIEWICZ, R. H. WIGHTMAN, and S. A. NARANG. Can. J. Chem. **51**, 3649 (1973).
8. N. KATAGIRI, K. ITAKURA, and S. A. NARANG. Unpublished work.
9. T. C. CATLIN and F. CRAMER. J. Org. Chem. **38**, 245 (1973).
10. N. KATAGIRI, K. ITAKURA, and S. A. NARANG. Chem. Commun. 325 (1974).
11. J. J. MICHNIEWICZ, C. P. BAHL, K. ITAKURA, N. KATAGIRI, and S. A. NARANG. J. Chromatogr. **85**, 159 (1973).
12. G. G. BROWNLEE and F. SANGER. Europ. J. Biochem. **11**, 395 (1969).
13. E. JAY, R. BAMBARA, P. PADMANABHAN, and R. WU. Nucleic Acids Res. **1**, 331 (1974).
14. C. P. BAHL, K. J. MARIANS, R. WU, S. A. NARANG, K. ITAKURA, and N. KATAGIRI. In preparation.
15. IUPAC–IUB. Biochemistry, **9**, 4022 (1970).