Notes

[Chem. Pharm. Bull. 34(5)2202—2208(1986)]

Synthesis of a Gene Coding for Human Lysozyme

MORIO IKEHARA,*[,a] KAZUKO FUJIMOTO,[a] YOSHIYUKI AOYAMA,[a] KOICHI YANASE,[a] JITSUHIRO MATSUGI,[a] TETSUYA INAOKA,[a] TOMOKO TOKUNAGA,[a] SEIICHI UESUGI,[a] SHIGENORI IWAI,[b] and EIKO OHTSUKA[b]

Faculty of Pharmaceutical Sciences, Osaka University,[a] 1–6 Yamadaoka, Suita, Osaka 565, Japan and Faculty of Pharmaceutical Sciences, Hokkaido University,[b] Kita-12, Nishi-6 Kita-ku, Sapporo 060, Japan

(Received September 9, 1985)

A gene for human lysozyme with 130 amino acids has been synthesized by ligation of chemically synthesized oligodeoxyribonucleotides with chain lengths of 13—18. The duplex with 401 base pairs was divided into 52 fragments, which were prepared by the phosphotriester solid-phase method involving condensation of dinucleotide units. By-products obtained in the synthesis have been characterized.

Keywords——solid-phase phosphotriester method; oligodeoxynucleotides; polystyrene support; restriction sequence; DNA ligase

We have synthesized a gene coding for human growth hormone(hGH) and have shown that it is efficiently expressed in *Escherichia coli*.[1] This indicates that synthesis of any gene coding for a protein of up to about two hundred amino acids could be achieved in principle and modification of such a gene would lead to mutant proteins. Thus, we planned to synthesize a gene coding for human lysozyme, which has long been a target of enzymatic and crystallographical studies. In particular, the three dimensional structure of hen egg lysozyme has been elucidated.[2] Therefore, we can obtain much information about structure–activity relationships by modifying the lysozyme structure by using suitably altered genes, which could be synthesized chemically. In this paper we described the synthesis of the gene fragments for human lysozyme and also some side reactions observed during the synthesis.

Design and Synthesis of the Human Lysozyme Gene

We first designed the gene for human lysozyme, which has 130 amino acids as shown in Fig. 1. We selected codons commonly used for yeast proteins,[3] and added starting and stopping signals at the beginning and end of the structural gene, as well as protruding nucleotides for the restriction enzyme Xho I at both ends of the gene. This strategy is almost the same as that used taken in the case of the gene for hGH.[1] For construction of the lysozyme gene, we divided two strands of deoxyribonucleic acid (DNA) into 52 oligonucleotides of 13—18 nucleotides length as indicated by arrow heads in Fig. 1. By means of a computer search, fragments having self-complementary or pallindromic sequences were avoided by altering codon usage.

Synthesis of the fragments of DNA was performed by a solid-phase method using 1% crosslinked polystyrene.[4] Except for the first nucleosides, we used mainly dinucleotides[5] suitably protected (Chart 1). Starting with 20—30 mg (3—5 μmol) of nucleoside resin, the synthetic procedure illustrated in Table I was repeated using 3—4 eq amounts of the dimers,

```
TCGAGATG..          XhoI Met-
    CTAC..
                                        10                                  20
Lys Val Phe Glu Arg Cys Glu Leu Ala Arg Thr Leu Lys Arg Leu Gly Met Asp Gly Tyr
AAG GTT TTT GAG AGA TGC GAA TTA GCC AGA ACT TTG AAG AGA TTG GGT ATG GAC GGC TAC
TTC CAA AAA CTC TCT ACG CTT AAT CGG TCT TGA AAC TTC TCT AAC CCA TAC CTG CCG ATG
 * 10          * 20          * 30          * 40          * 50          * 60
                                        30                                  40
Arg Gly Ile Ser Leu Ala Asn Trp Met Cys Leu Ala Lys Trp Glu Ser Gly Tyr Asn Thr
CGT GGT ATT TCT TTA GCC AAC TGG ATG TGT CTT GCT AAG TGG GAA TCC GGC TAT AAC ACT
GCA CCA TAA AGA AAT CGG TTG ACC TAC ACA GAA CGA TTC ACC CTT AGG CCG ATA TTG TGA
 * 70          * 80          * 90          *100          *110          *120
                                        50                                  60
Arg Ala Thr Asn Tyr Asn Ala Gly Asp Arg Ser Thr Asp Tyr Gly Ile Phe Gln Ile Asn
AGA GCT ACC AAT TAC AAC GCT GGC GAC CGT TCT ACA GAC TAT GGT ATT TTC CAA ATT AAC
TCT CGA TGG TTA ATG TTG CGA CCG CTG GCA AGA TGT CTG ATA CCA TAA AAG GTT TAA TTG
*130          *140          *150          *160          *170          *180
                                        70                                  80
Ser Arg Tyr Trp Cys Asn Asp Gly Lys Thr Pro Gly Ala Val Asn Ala Cys His Leu Ser
TCT AGA TAT TGG TGT AAC GAT GGC AAG ACT CCA GGT GCC GTC AAC GCC TGT CAC TTA TCT
AGA TCT ATA ACC ACA TTG CTA CCG TTC TGA GGT CCA CGG CAG TTG CGG ACA GTG AAT AGA
*190          *200          *210          *220          *230          *240
                                        90                                 100
Cys Ser Ala Leu Leu Gln Asp Asn Ile Ala Asp Ala Val Ala Cys Ala Lys Arg Val Val
TGC TCA GCT TTG CTT CAG GAC AAC ATT GCT GAT GCT GTT GCC TGC GCT AAG AGA GTT GTC
ACG AGT CGA AAC GAA GTC CTG TTG TAA CGA CTA CGA CAA CGG ACG CGA TTC TCT CAA CAG
*250          *260          *270          *280          *290          *300
                                       110                                 120
Arg Asp Pro Gln Gly Ile Arg Ala Trp Val Ala Trp Arg Asn Arg Cys Gln Asn Arg Asp
CGT GAC CCA CAG GGT ATT AGA GCC TGG GTC GCT TGG AGA AAC AGA TGC CAA AAT AGA GAT
GCA CTG GGT GTC CCA TAA TCT CGG ACC CAG CGA ACC TCT TTG TCT ACG GTT TTA TCT CTA
*310          *320          *330          *340          *350          *360
                                       130
Val Arg Gln Tyr Val Gln Gly Cys Gly Val
GTC AGA CAA TAC GTT CAA GGT TGT GGT GTT
CAG TCT GTT ATG CAA GTT CCA ACA CCA CAA
*370          *380          *390          *400

..TAATAGC
,.ATTATCGAGCT   -End End XhoI
```

Fig. 1. Human Lysozyme

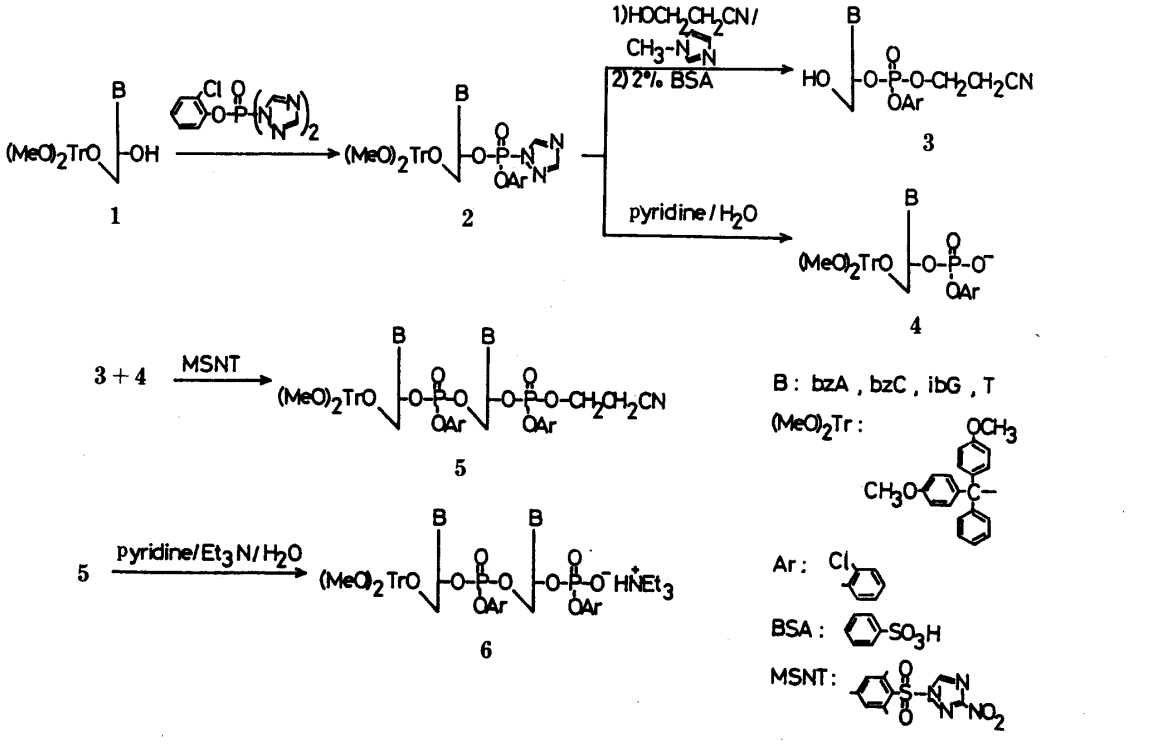

Chart 1

Table I. Synthetic Procedures

Step	Reagent or solvent	Amount	Reaction time	No. of operations
1	CH_2Cl_2–MeOH (7:3)	2 ml	0.5 min	3
2	2% BSA in CH_2Cl_2–MeOH (7:3)	2 ml	1 min	1
3	CH_2Cl_2–MeOH (7:3)	2 ml	0.5 min	1
4	2% BSA in CH_2Cl_2–MeOH (7:3)	2 ml	1 min	1
5	CH_2Cl_2–MeOH (7:3)	2 ml	0.5 min	2
6	Pyridine	2 ml	0.5 min	3
7	Pyridine	0.3 ml	Co-evaporation	1
8	Dimer block (3—4 eq) in pyridine	20—30 mg/0.3 ml	Co-evaporation	1
9	MSNT (15—20 eq) in pyridine	20—30 mg/0.3 ml	20 min (40 °C)	1
10	Pyridine	2 ml	0.5 min	2
11	0.1 M DMAP in pyridine AC_2O	1.8 ml 0.2 ml	3 min	1
12	Pyridine	2 ml	0.5 min	3

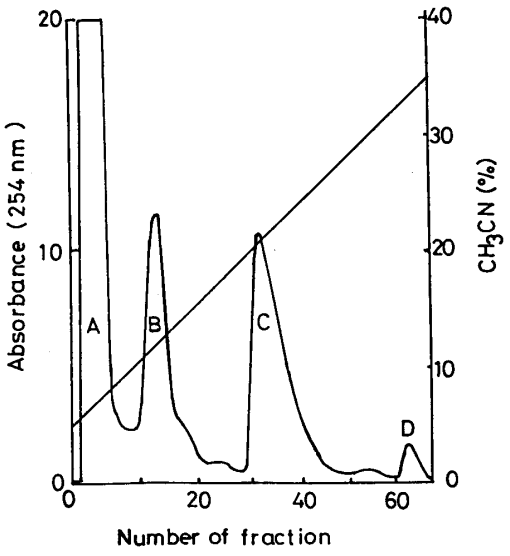

Fig. 2. C18 Silica Gel Column Chromatography

Sample: d(TAAAGAAATACCACG). Column: C18 (0.7 × 7.5 cm i.d.). Solvent: A, 50 mM TEAA, 5% CH_3CN; B, 50 mM TEAA, 35% CH_3CN. About 3 ml/fraction.

and mesitylenesulfonyl-3-nitrotriazolide[6] (MSNT) as the condensing reagent. The reaction was performed either at 40 °C for 20 min or at 20—25 °C for 40 min. After the condensation, removal of the protecting groups was performed by pyridine aldoxamate[6] and ammonia treatments, and the 5′-dimethoxytrityl protected oligomers were purified on a column of Silica gel C18 as shown in Fig. 2. The product was eluted in peak C, while the following peak D containing by-products. The structure of materials in the peak D is discussed later. When oligomers obtained from peak C were not pure enough, chromatography was repeated. The final product was analyzed by both high pressure liquid chromatography (HPLC) and mobility shift analyses.[7]

The oligonucleotides obtained as above were grouped in five and joined separately as indicated in Fig. 3. The oligomers were 5′-phosphorylated by using [γ-^{32}P]adenosine triphosphate (ATP) and polynucleotide kinase[8] except for fragments U1 and L26. Ligation was carried out with DNA ligase as described earlier[1] and 10—15 μg of fragments I—V were obtained. These segments were ligated to each other to obtain a gene for lysozyme in *ca.* 1 μg yield.

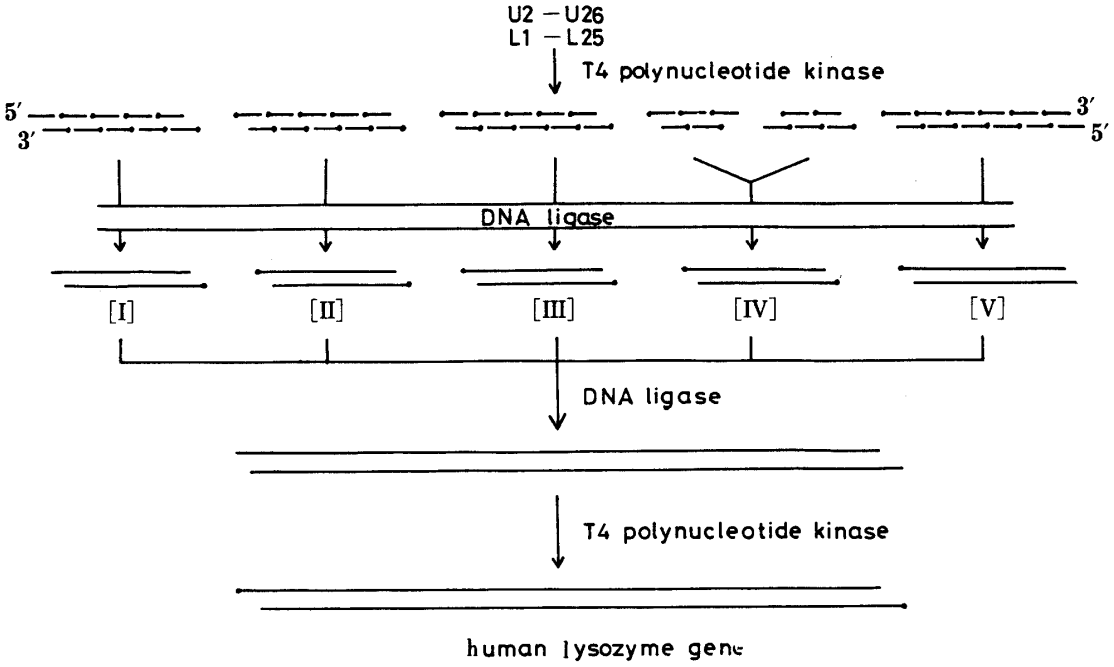

Fig. 3. Ligation of Fragments to Prepare Lysozyme Gene
●, means phosphate.

Structure of the By-product

The structure of the by-product obtained in peak D (Fig. 2) was analyzed as follows. This material contains the dimethoxytrityl group as revealed by the development of yellow color on acid treatment. After detritylation with acetic acid, the 5′-OH was phosphorylated with ^{32}P as described above. The product was then treated with nuclease P1 or snake venom phos-

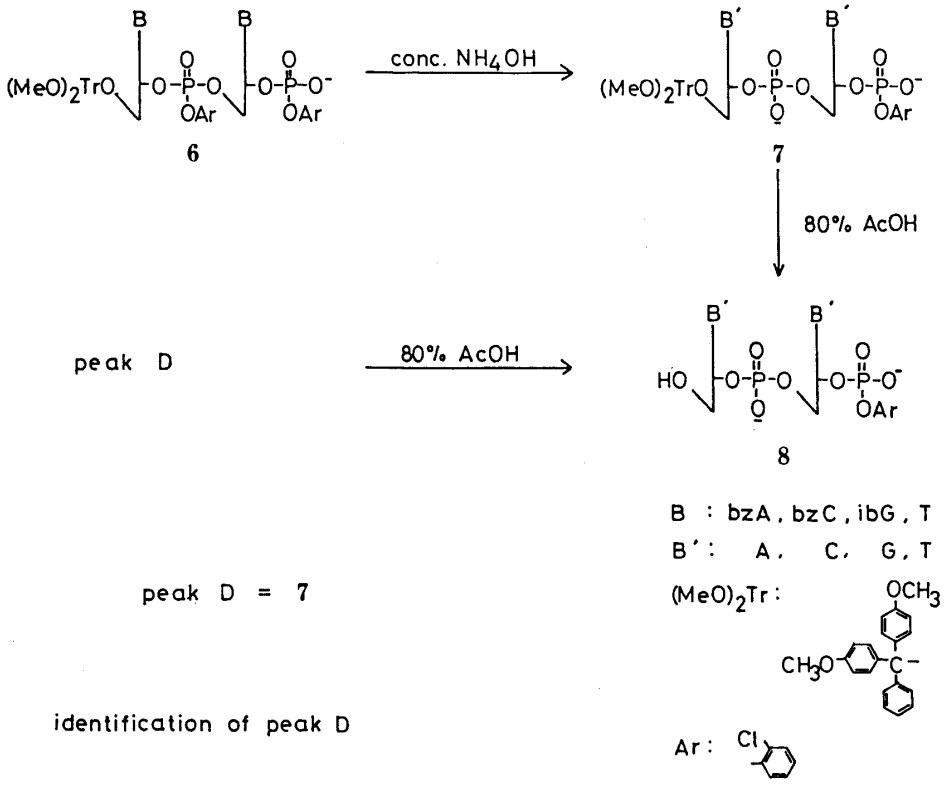

Chart 2

phodiesterase. Although this material resisted digestion with these enzymes, prolonged treatment with nuclease P1 gave products, which were analyzed on a nucleic acid analyzer and by paper electrophoresis at pH 3.5. It was found that the product was a dinucleotide bearing only the terminal *o*-chlorophenyl protection by comparison with authentic dinucleotide derived from the dinucleotide **6** (Chart 2) by ammonia and acid treatments. The mechanism leading to the *o*-chlorophenyl compound **8** is presumably as follows. In the condensation of dinucleotides to the 5′-end of preformed oligomers on the polymer support, the O^6 or O^4 position of guanine or thymine was attacked by entering 3′-phosphate to give the O-phosphorylated compound (Chart 3). On treatment with acid for elongation of the chain,

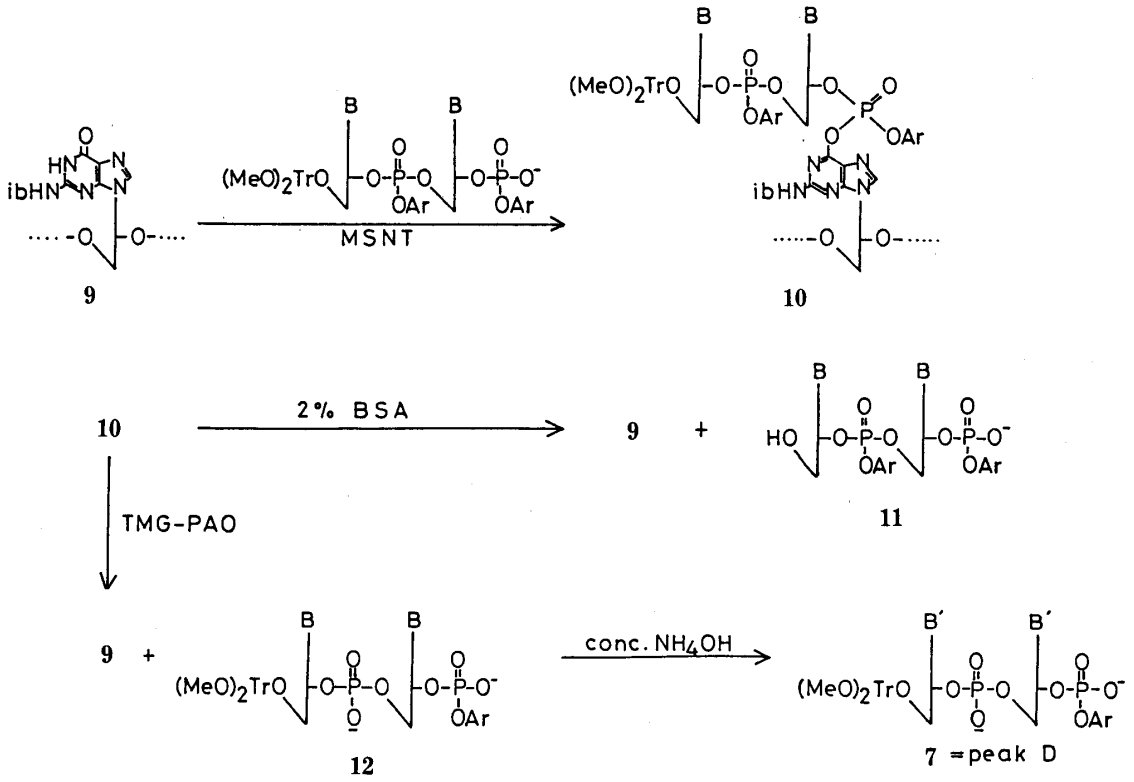

Chart 3

such by-products are degraded to give normal oligomers. However, for the final elongation step, no acid treatment was needed and the usual procedures to remove internucleotide phosphate and acyl protections should lead to compound **14** which has been detected in peak D. This point was confirmed by a model experiment as follows. When a fully acylated deoxyguanosine was treated with 5′-dimethoxytrityl-N^6-benzoyldeoxyadenosine 3′-(2-chlorophenyl) phosphate in the presence of MSNT, we obtained a compound which could be hydrolyzed to N^2-acetylguanosine and N^6-benzoyldeoxyadenosine 3′-(2-chlorophenyl) phosphate. Similar experiments with thymidine and deoxyadenosine derivatives showed that only in the case of thymidine was a by-product obtained. It appears that the compounds obtained as by-products in the condensation reactions must be O-phosphorylated compounds, as suggested by Reese and Ubasawa.[9]

Experimental

Reagents and Enzymes——DNA ligase was obtained from Takara Shuzo Co. and Boehringer Mannheim GmbH. Restriction endonuclease Xho I was purchased from Boehringer Mannheim, GmBb. Other enzymes including restriction endonucleases were obtained from Takara Shuzo Co.

Oligonucleotides——Deoxyoligonucleotides with the chain length of 15—24 bases were synthesized by the phosphotriester solid-phase method, using 1% cross-linked polystyrene as the support.[4] Fifty-two oligonucleotides were prepared by addition of dinucleotide blocks to nucleosides linked through a 3′-succinyl group to the resin. Sixteen dinucleotide blocks were synthesized with mesitylenesulfonyl-3-nitrotriazolide[6] essentially as reported[10] and purified mostly by reversed-phase chromatography[11] under the conditions described.[5] A typical synthesis of a 15-mer started with treatment of 5′-dimethoxytritylated resin (20—30 mg, 3—5 μmol) with benzenesulfonic acid (BSA) (2%, 2 ml) for 1 min twice in a small vessel with a sintered glass filter and a stopcock at the bottom. Dimers (3—4 eq) were added with MSNT (15—20 eq) and the reaction was performed either for 20 min at 40 °C or 30 min at 30 °C. The resin was treated with acetic anhydride (0.2 ml) in 4-dimethylaminopyridine–pyridine (0.1 M, 1.8 ml) for 3 min. The product was deblocked with N^1,N^1,N^3,N^3-tetramethylguanidinium 2-pyridinealdoximate in dioxane–H_2O (0.5 M, 90% dioxane) at 30 °C overnight and with conc. ammonia at 50—60 °C for 5 h. The dimethoxytritylated 15-mer was separated on a column (0.7 × 7.5 cm) of C-18 Silica gel (Waters, 35—100μm) with a gradient of acetonitrile (5—35%, 200 ml) in 50 mM triethylammonium acetate (TEAA). The material thus obtained (peak C in Fig. 2) was treated with 80% acetic acid. The product was applied to a column (0.7 × 21 cm) of DEAE-cellulose 650S (Toyo Soda Co.) with a gradient of NaCl (0.1—0.3 M, total 200 ml) in 7 M urea/20 mM Tris–HCl, pH 7.5. The purity of the 15-mer was checked by HPLC on C-18 Silica gel (Toyo Soda, TSK-410 AK) and ion-exchange resin (TSK gel, DEAE 2SW). Impurities were removed by HPLC, and the purity and base sequence of the product were confirmed by mobility shift analysis.[7]

Ligation of Oligomers for Lysozyme Gene——Fragment oligonucleotides were combined in groups of five containing 10—12 oligomers each (as shown in Fig. 3). In each group the oligomers (0.1 A_{260}) were separately kinated at 5′-OH (except for U1 and L26) using kination buffer (2 μl containing 50 mM Tris–HCl (pH 8.0), 10 mM $MgCl_2$, 10 mM β-EtSH and 1 mM spermine, [γ-^{32}P]ATP (1 μl, 10 mM), polynucleotide kinase (6U/μl) 0.5 μl and water (6.5 μl). The reaction mixture was incubated at 37 °C for 90 min. After deactivation of the kinase at 90 °C for 5 min, the oligonucleotides were ligated under the conditions described in Table I. For the ligation, each oligomer mixture, which was kinated as before, was mixed with ligation buffer and heated at 90 °C for 3 min, cooled rapidly in a bath to 75 °C, and then gradually cooled to room temperature. The mixture was kept at 15 °C for 10 min. β-EtSH and ligase were added and the mixture was incubated at 15 °C overnight. The protein was removed with phenol and each segment was obtained by ethanol precipitation. The pure material was obtained by 10% polyacrylamide gel electrophoresis (PAGE). Elution from the gel was performed by electroelution with 45 mM Tris–borate buffer (pH 8.4) and the yield was 10—15 μg (0.2—0.3 A_{260} units).

Each segment was ligated at 20 °C, and the gene for lysozyme was obtained in a yield of 1 μg.

Experiments for Structure Elucidation of the By-product——i) Synthesis of dNpNpo$^-$ TEA$^+$: d(DMTr)^{ib} GpTpo$^-$ TEA$^+$ (10 mg, 7.4 μmol) was dissolved in a small amount of pyridine, and conc. ammonia (2 ml) was added. The mixture was heated at 55 °C for 2 h. After confirmation of disappearance of the starting materials, the solvent was evaporated off, and coevaporation of the product with toluene gave a residue, which was treated with 80% acetic acid for 20 min. Dimethoxytrityl alcohol was removed by AcOEt extraction, and the aqueous solution was evaporated to give dGpTp-*o*-Cl-phenyl (*p*).

ii) The peak D fractions (Fr. No. 49—50) of L9 oligomer were collected and the solvent was removed *in vacuo*. The residue was treated with acetic acid as above to give the sample for correlation.

iii) When the samples obtained in i) and ii) were analyzed by HPLC (column, M & S Pack 6 × 150 nm; solvents, 5% CH_3CN+0.1 M TEAA, 25% CH_3CN+0.1 M TEAA, gradient of 28—80% in 24 min), both were eluted at 16.5 min and when they were mixed, only one peak was detected.

iv) dAcG^{iBu}Ac (21 mg, 50 μmol) and d(DMTr)A^{Bz}pO$^-$TEA$^+$ (95 mg, 100 μmol) were coevaporated with pyridine and condensed by using MSNT (148 mg, 0.5 mmol) in pyridine (3 ml). On thin layer chromatography (TLC) a spot appearing at higher *Rf* was extracted with MeOH. When this material was hydrolyzed with 2% BSA (2 ml) at room temperature for 5 min, dAcG^{iBu}Ac and dA^{Bz}p-O-phenyl were detected on TLC.

The same reactions with dAcTAc and aAcA^{Bz}Ac showed that only in the case of thymidine was this type of by-product formed.

Acknowledgement The authors wish to thank Dr. Don Dennis for reading the manuscript.

References

1) M. Ikehara, E. Ohtsuka, T. Tokunaga, Y. Taniyama, S. Iwai, K. Kitano, S. Miyamoto, T. Ohgi, Y. Sakuragawa, K. Fujiyama, T. Ikari, M. Kobayashi, T. Miyake, S. Shibahara, A. Ono, T. Ueda, T. Tanaka, H. Baba, T. Miki, A. Sakurai, T. Oishi, O. Chisaka, and K. Matsubara, *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 5956 (1984).
2) T. Imoto, L. N. Johnson, A. C. T. North, D. C. Phillips, and J. A. Rupley, "Enzymes," IIIrd Ed., Vol. 7, Academic Press, New York, 1972, p. 665.
3) T. Ikemura, *J. Mol. Biol.*, **158**, 573 (1982).
4) H. Ito, Y. Ikemura, S. Ikuta, and K. Itakura, *Nucleic Acids Res.*, **10**, 1755 (1982).

5) E. Ohtsuka, S. Iwai, T. Tokunaga, and M. Ikehara, *Chem. Pharm. Bull.*, **33**, 3153 (1985).
6) C. B. Reese, R. Titmas, and L. Yau, *Tetrahedron Lett.*, **1978**, 2727.
7) F. Sanger, J. E. Donelson, A. R. Coulson, H. Kössel, and D. Fischer, *Proc. Natl. Acad. Sci. U.S.A.*, **20**, 1209 (1973); E. Ohtsuka, S. Nishikawa, A. F. Markham, S. Tanaka, T. Miyake, T. Wakabayashi, M. Ikehara, and M. Sugiura, *Biochemistry*, **17**, 4894 (1978).
8) C. C. Richardson, *Proc. Natl. Acad. Sci. U.S.A.*, **54**, 158 (1965).
9) C. B. Reese and A. Ubasawa, *Tetrahedron Lett.*, **21**, 2265 (1980).
10) C. Broka, T. Hozumi, R. Arenzene, and K. Itakura, *Nucleic Acids Res.*, **8**, 5461 (1980).
11) H.-J. Fritz, R. Belagaje, E. Brown, R. H. Fritz, R. A. Jones, R. G. Lees, and H. G. Khorana, *Biochemistry*, **17**, 1257 (1978).