

# Chemical ligation of DNA: the first non-enzymatic assembly of a biologically active gene

Zoe A. Shabarova, Irena N. Merenkova, Tatyana S. Oretskaya, Nadezhda I. Sokolova, Eugene A. Skripkin, Ekaterina V. Alexeyeva, Andrey G. Balakin and Alexey A. Bogdanov  
The Belozersky Laboratory of Molecular Biology and Bioorganic Chemistry and Chemistry Department, Moscow State University, Moscow 119899, USSR

Received March 5, 1991; Revised and Accepted July 16, 1991

## ABSTRACT

**An artificial gene comprising 183 base pairs has been assembled by template-directed condensation of 35- to 53-membered oligodeoxyribo nucleotides with cyanogen bromide as a condensing agent. The reaction is complete within several minutes at 0°C in buffer. The resulting mini-gene was cloned and expressed *in vivo* and *in vitro*. We have also found that the polymerase inhibition technique (toe-printing) is a good way to ascertain that translation initiation complexes form in the case of single-stranded DNAs as well. Thus, along with the fully chemical assembly of synthetic genes, a rapid and sufficiently reliable method for determining their ribosome-binding properties was developed.**

## INTRODUCTION

Until recently, the principal method for assembling DNA duplexes from synthetic oligodeoxyribonucleotides has been joining them within complementary complexes with the aid of T4 DNA ligase. This method was first elaborated in the late sixties by H. Khorana (1). Several other strategies combining chemical and enzymatical synthesis of extended DNA duplexes were reviewed in a recent paper (2).

However, polynucleotides are synthesized in the cell via using a handful of enzymes. The modern tendency of biochemistry is to mimic enzymatic processes with a non-catalytic chemical reactions. Oligonucleotides of the reliable length can be easily obtained by the use of automated synthesizers. However, to generate a relatively long DNA sequence one still needs to use enzymes for joining oligonucleotides together.

In recent years, interest in artificial analogues of nucleic acids has grown significantly. However, traditional enzymatic methods often fail to provide such structures. A few years ago we started developing a method as an alternative to the enzymatic one: fully chemical assembly of DNA duplexes. This approach, which we have called chemical ligation, consists in condensation of oligonucleotides on a complementary template with the aid of chemical reagents-water soluble carbodiimide (3) or cyanogen bromide (4). The main advantage of the latter is a high rate of reaction and the absence of by-products. Chemical ligation has been used to obtain a broad range of DNA duplexes with both natural (phosphodiester) and artificial (phosphamide,

pyrophosphate) internucleotide linkages (5). These investigations have demonstrated that in chemical ligation the requirements to the structure of the substrates are appreciably less stringent than in enzymatic ligation. In this paper we show that a relatively long gene can be synthesized without any enzymes at all.

## MATERIALS AND METHODS

Cyanogen bromide (BrCN), Tris, 2-morpholinoethane sulfonate (MES), acrylamide, N,N'-methylenebisacrylamide, formamide, yeast extract were purchased from Merck, bactotripton was from Difco, dNTP, ddNTP, tRNA<sup>Met</sup>, plasmid pT7-4 were obtained from Boehringer Mannheim,  $\gamma$ -<sup>32</sup>P-ATP, <sup>32</sup>P-orthophosphate, T4 DNA ligase, T4 polynucleotide kinase, T4 DNA polymerase, Taq DNA polymerase, AMV reverse transcriptase were from USSR, DNA polymerase I (Klenow fragment), restriction endonucleases Hind III, Xba I, Sma I, isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl  $\beta$ -D-galactopyranoside (Xgal) were purchased Pharmacia, RNase A was from Sigma. 30S ribosomal subunits of *E. coli* MRE 600 were isolated as described (6).

### Synthesis of oligonucleotide

Oligodeoxyribonucleotides were synthesized by phosphoramidite chemistry on a 'Cyclon' Biosearch DNA synthesizer by running 0.2  $\mu$ mol programme. After deprotection, the products were purified by denaturing polyacrylamide gel electrophoresis (PAGE) as usual. The part of purified oligonucleotides were [<sup>5'</sup>-<sup>32</sup>P]-phosphorylated with ATP and T4 polynucleotide kinase. Oligomers *a* and *h* (see Fig.1) were radiolabelled with [ $\gamma$ -<sup>32</sup>P]-ATP by standart procedure (7).

### Chemical ligation

1. *One step reaction*: The mixture of 5'-phosphorylated oligonucleotides *b-g* (410pmol) and *a,h* (400pmol) in 9  $\mu$ l buffer (0.25M MES adjusted with Et<sub>3</sub>N to pH 7.5, containing 0.02M MgCl<sub>2</sub>) were heated to 97°C, slowly cooled to 0°C, then 1  $\mu$ l 10M BrCN in *abc*. acetonitrile was added. After 1 min the nucleotide material was precipitated with ethanol. 5'-<sup>32</sup>P-labelled oligomers *a* and *h* (10-15 nCi) were added previously to the reaction mixture.

**2. Reaction within single-stranded DNA (ssDNA):** The mixture of 5'-phosphorylated oligonucleotides *a* (400pmol), *b-d* (410pmol), non-phosphorylated oligonucleotides *e-h* (410pmol) and 5'-<sup>32</sup>P-labelled oligomer *a* (10–15nCi) in 9 μl buffer (0.25M MES adjusted with Et<sub>3</sub>N to pH 7.5, containing 0.02M MgCl<sub>2</sub>) were ligated with BrCN as described above. The oligonucleotides *e-h* were ligated analogously, using 5'-phosphorylated *e-h*, non-phosphorylated *a-d* and 5'-<sup>32</sup>P-labelled *h* oligomers. After purified synthetic ssDNA was isolated in a total yield of 5%.

**Enzymatic ligation**

Enzymatic ligation reaction was carried out with T4 DNA ligase at 10°C for 12 hours (7). The reaction mixture was analyzed in the same way as described above.

**Purification of the synthetic ssDNA**

The preparative denatured PAGE was conducted to purify the ssDNA (B,C, Fig.1). After autoradiography, the band containing the expected product was cut off and the radioactivity was measured. After elution from PAG, ssDNA was precipitated with ethanol as usual.

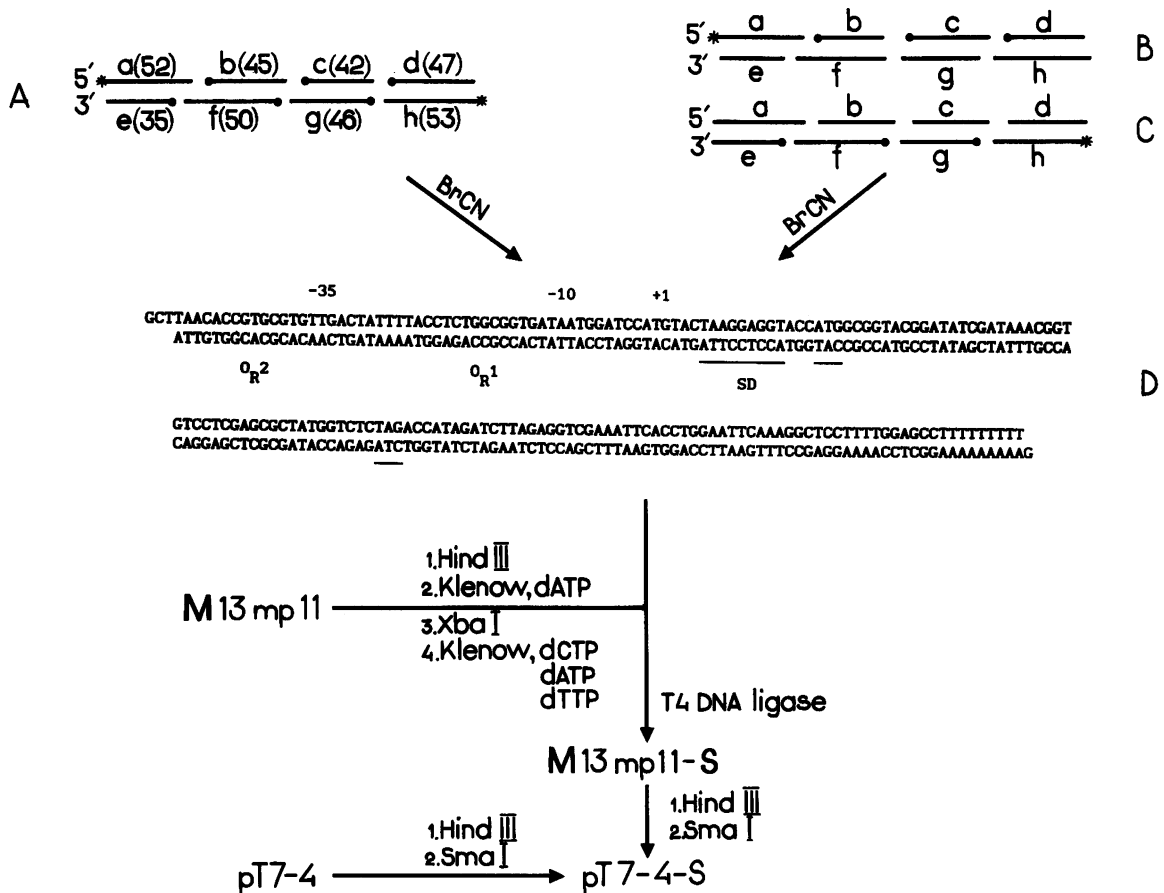
**Molecular cloning of the synthetic DNA**

After elution from PAG, both chains were annealed by cooling from 97°C to 0°C, and the DNA duplex was ligated with RF DNA phage M13 mp11 cut with restriction endonucleases Hind III and Xba I and partially built up with DNA polymerase I (Klenow fragment). Thus, during hydrolysis with restriction endonuclease Hind III, the building procedure involved dATP and resulted in formation of a three-membered end, complementary to the synthetic fragment. The Xba I site was built up in the presence of three nucleoside triphosphates: dCTP, dATP and dTTP. For growth in the presence of IPTG and Xgal, the *E.coli* XLBlue cells infected with phage M13 with inserts gave colorless plaques that were picked up and screened by sequencing. A Hind III-Sma I fragment from M13 mp11-S was cloned to the corresponding restriction sites in plasmid pT7-4.

All genetic engineering manipulations such as transformation of competent cells, phage M13 and plasmid DNA isolation, screening of recombinant clones were routine (7).

**RNA isolation**

20 μCi <sup>32</sup>P-orthophosphate was added to the *E.coli* PR13 suspension growing on log phase (0.2 OD) in 1 ml LB-broth and



**Figure 1.** Strategy of mini-gene assembly with the use of BrCN and cloning of a synthetic duplex into M13 mp11 and pT7-4. Ligation of DNA duplex (A) and separate ligation of upper (B) and lower (C) chains. — 5'-phosphorylated oligonucleotide, \* — 5'-<sup>32</sup>P-labelled oligonucleotide, consecutive lower-case letters denote the oligonucleotide with the number of monomer units given in parentheses.

incubated for 15 min at 37°C. RNA was isolated from cells by boiling (1% SDS, 8 mM EDTA, pH 7.0) for 2 min (8). For primer extension analysis the same procedure was used without <sup>32</sup>P-label addition.

#### Sequence analysis of synthetic gene

A M13 ssDNA (5 pmol) was mixed with 5'-<sup>32</sup>P-labelled primer (5 pmol) in 10 µl buffer containing 50 mM Tris-HCl (pH 8.3), 540 mM KCl 1 mM EDTA. Sequence procedure with AMV reverse transcriptase was performed as described (9).

#### Toe-printing analysis of initiation translation complexes

6 pmol of ssDNA were annealed with an equimolar amount of a 5'-end labelled oligonucleotide primer in 12 µl of a buffer containing 10 mM Tris-HCl (pH 7.4), 60 mM NH<sub>4</sub>Cl, 6 mM β-mercaptoethanol for 2 min at 80°C. After cooling in ice, 8 µl of the same buffer containing 60 mM MgAc<sub>2</sub> were added to this mixture. Then 1 µl of each of the four dNTP (each in 4 mM concentration) and 2 µl of 30S subunits (2 pmol/µl) were added to 2 µl of an annealing mixture which was preincubated at 37°C for 5 min. Following this procedure, 2 µl tRNA (20 pmol/µl) were added to the mixture, and it was preincubated for another 5 min at 37°C. Then, added to the mixture were 2 units of AMV reverse transcriptase, 0.1–0.2 units of DNA polymerase I (Klenow fragment) or T4 DNA polymerase, and the primer extension reaction was conducted for 15 min at 37°C. It was stopped by heating for 3 min at 100°C. RNase A was added to conc. 0.2 mg/ml, after 10–20 min the mixture was diluted to 50 µl, extracted with phenol:chloroform (1:1 v/v) and precipitated with ethanol. Patterns of primer extension were analyzed on sequencing PAG.

#### Determination of radioactivity

An automatic counter Delta-300 (Tracor, Netherland) was used for the measurement.

## RESULTS AND DISCUSSION

### Chemical ligation

In spite of its small size the synthesized DNA duplex, whose structure is presented in Fig.1, is a real gene: it comprises a promoter (an analogue of strong promoter lambda P<sub>R</sub>), including operator segments O<sub>R</sub>1 and O<sub>R</sub>2; a translation initiation segment with an extended Shine-Dalgarno sequence (9 b.p.), followed by an 18 codon open reading frame starting with the ATG initiation codon and ending with the TAG termination codon; and also a rho-independent transcription terminator (an analogue of t<sub>td</sub>). Transcription of this mini-gene should give a mRNA, 130 nt long. Each of the strands of the DNA duplex was subdivided into four oligomers having a length of 35 to 53 nucleotides (denoted in Fig.1 by the letters a–h) in such a manner that in the duplex these oligomers overlap by 5 to 14 units. The synthetic DNA duplex was assembled in two ways: by simultaneous chemical ligation of both strands (Fig.1A) and by ligation of each strand separately with non-phosphorylated oligonucleotides being used as the template (Fig.1B,C). As our results showed, the second method is practically better with view of easiness of product isolation. In parallel, ligation with T4 DNA ligase was carried out (Fig.2). In this experiment, one major advantage of the chemical ligation method became immediately apparent, namely, a high rate of the reaction, which took only one minute, as compared to 12 hours in the case of T4 DNA

ligase. We should note that the bases are not modified under these conditions. The comparison UV-spectra and HPLC data of nucleotides and oligonucleotides before and after BrCN-treatment proves this fact (data not shown).

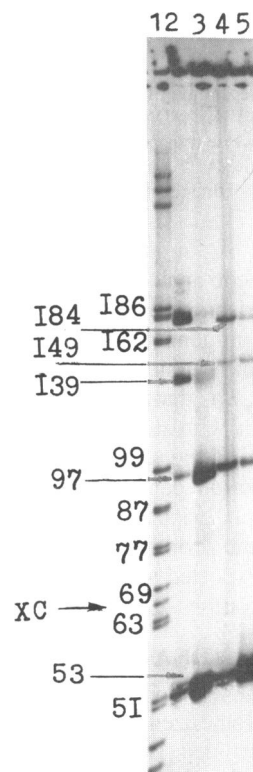
### Molecular cloning of the synthetic DNA

DNA duplex D, prepared by the chemical ligation method (Fig.1B,C), was cloned into Hind III and Xba I sites of phage M13. Cohesive ends of the synthetic fragment have been complementary to the corresponding cohesive ends of the vector, formed after cutting with restriction endonucleases and partial build up with the DNA polymerase I (Klenow fragment), thus prevents multiplication of the fragment being cloned and self-ligation of the vector. As a result of cloning, recombinant bacteriophage M13 mp11-S was obtained, carrying a synthetic insert orientated in the same direction as the lacZ gene (11), (see Fig.3). Notably, this orientation of the inserted sequence is identical with the primary structure of mRNA.

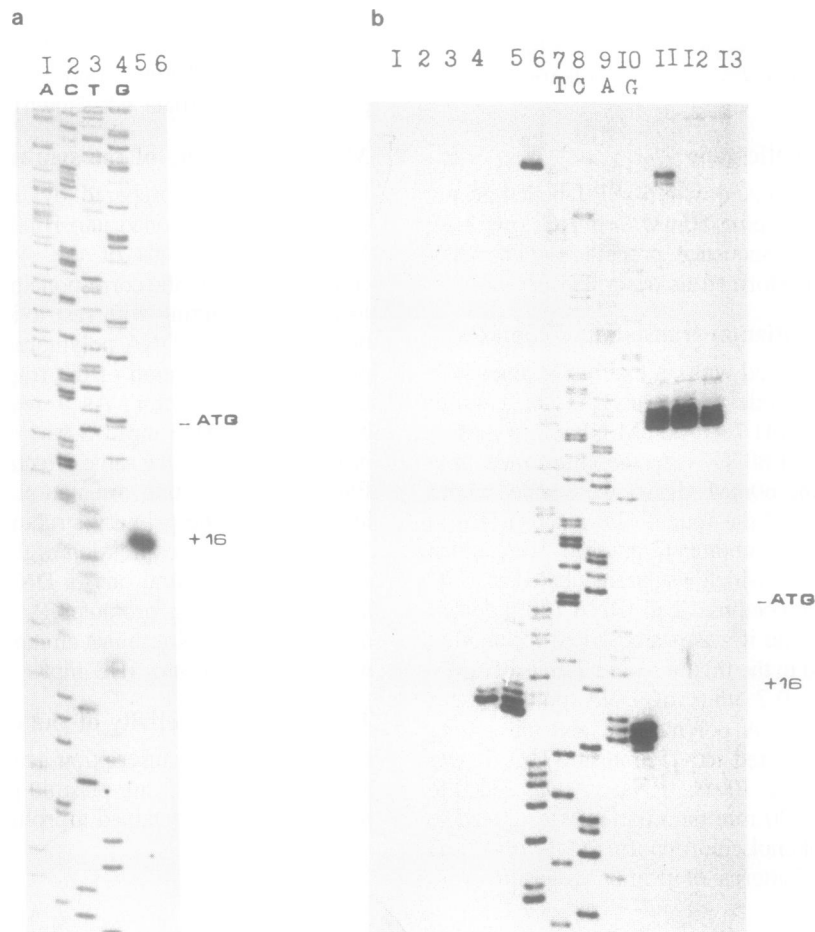
To compare the expression of the synthetic structure with the corresponding natural one, a DNA fragment of bacteriophage lambda, containing promoter P<sub>R</sub> and the gene cro translation initiation site (22 N-terminal amino acid residues of protein Cro), was also cloned into M13 mp8.

### Transcriptional activity of the cloned mini-gene

For studying the transcription activity of the synthesized mini-gene, it was cloned into plasmid pT7-4. The resulting plasmid pT7-4-S (Fig.1) contained a promoter of T7 RNA polymerase



**Figure 2.** Electrophoretic separation (autoradiographs 8% PAG with 7M urea) of reaction mixtures after T4 DNA ligation (lanes 2,4), BrCN-induced ligation (lanes 3,5) upper and lower chains of mini-gene, respectively. 1-fragments obtained by cleavage with Taq I restriction endonuclease of pHSB plasmid (10). The chain length of oligonucleotides are indicated to the left.

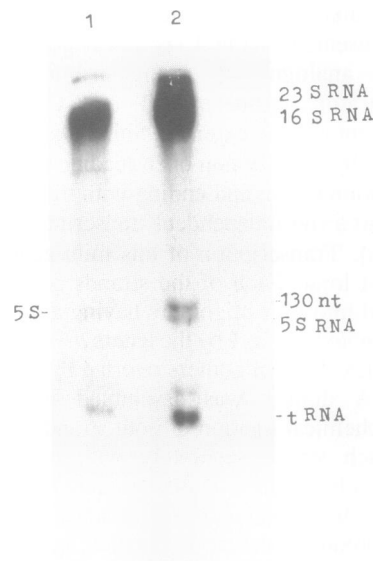


**Figure 3.** Electrophoretic separation of primer extension products (autoradiographs 8% PAG with 7M urea) a. ssDNA phage M13 mp11-S: A, C, T, G-dideoxy sequencing reactions (lanes 1-4); ssDNA + 30S subunits + tRNA<sup>Met</sup> + reverse transcriptase (lane 5); ssDNA + reverse transcriptase (lane 6). b. ssDNA phage M13 mp8-cro:ssDNA + reverse transcriptase (lane 1); ssDNA + Klenow fragment (lane 2); same as lane 1 + 30S subunits (lane 3); same as lane 2 + 30S subunits (lane 4); same as lane 3 + tRNA<sup>Met</sup> (lane 5); same as lane 4 + tRNA<sup>Met</sup> (lane 6); T, G, C, A-dideoxy sequencing reactions (lanes 7-10); ssDNA + 30S subunits + tRNA<sup>Met</sup> + T4 DNA polymerase (lane 11); same as lane 11 without tRNA<sup>Met</sup> (lane 12); same as lane 12 without 30S subunits (lane 13).

in tandem with a synthetic analogue of promoter lambda P<sub>R</sub> and the mini-gene. Comparison with similar experiments on plasmid constructs, carrying mini-genes with a natural promoter lambda P<sub>R</sub>, terminator t<sub>fd</sub> and the beginning of ORF Cro (plasmid series pMM (12)), has shown that the synthesized promoter is comparable in strength with the natural one (13). As can be inferred from the size of the synthesized RNA, 130 nt long, (Fig.4, lane 2), the rho-independent transcription terminator is also active. Reverse transcriptase mapping has shown that the main point at which transcription begins (the end of full-size cDNA) coincides, as expected with position +1, the additional transcript begins at position -2 (not shown).

**Toe-printing assay on ssDNA**

Along with effective transcription, gene expression is determined by the capacity of the synthesized mRNA for translation; the level of translation is determined largely already at the stage of formation of the initiation complex from small ribosome subunits and initiator tRNA. In this paper we found that the initiation complex has formed in ssDNA itself in presence 30S and tRNA<sup>Met</sup>, using the so called toe-printing technique (14). To compare the ribosome-binding activity of the artificial DNA with the corresponding natural one, a M13 ssDNA with the fragment



**Figure 4.** Electrophoretic separation of the total cellular RNAs from host *E.coli* strain XL-1Blue (lane 1) and transformed by plasmid pT7-4-S (lane 2).

of bacteriophage lambda containing promoter P<sub>R</sub> and the gene cro translation initiation region was used. In the synthesis of the complementary strand of the DNA, reverse transcriptase (Fig. 3a) or DNA polymerase I (Klenow fragment) stop at position +16 relative to the initiator ATG-codon; in this respect, DNA does not differ from the corresponding mRNA (15). It was found that the polymerase reaction in the case of the synthetic mini-gene is inhibited in exactly the same way as in natural DNA fragment of phage lambda. Interestingly, just as DNA polymerase I (Klenow fragment) or reverse transcriptase, the 30S initiator complexes are capable of inhibiting the reactions involving T4 DNA polymerase (Fig. 3b) and Taq DNA polymerase (not shown). It should be pointed out that in spite of the presence of several ATG-codons in the primary structure of the examined polynucleotides (see Fig. 1), the small ribosome subunit recognizes the right initiator codon in the appropriate context (the Shine-Dalgarno sequence) of the single-stranded DNA itself, in the presence of initiator tRNA.

## CONCLUSION

The present study has demonstrated that the method of chemical ligation is applicable to assembly of biological active genes. Moreover, this approach has a number of advantages over enzymatic one: a high reaction rate (one or two minutes versus 12–14 h with DNA ligase), possibility of introduction various modification and the low cost of the reagent as compared with the enzymes. We believe these advantages allow to totally automatize the procedure of obtaining any DNA sequences in the nearest future.

Using toe-printing technique on ssDNA allows to omit the isolation of the individual mRNA for primary screening of ribosome-binding capacity of the cloned genes.

## ACKNOWLEDGEMENTS

The authors would like to express their gratitude to S. Dokudovskaya for her assistance in the cloning and cell-transformation procedures and also to Dr. N. Tarasova for critical remarks.

## REFERENCES

1. Agarwal, K.L., Büchi, H., Caruthers, M.H., Gupta, N., Khorana, H.G., Kleppe, K., Kumar, A., Ohtsuka, E., RajBhandary, U.L., van de Sande, J.H., Sgarbetta, V., Weber, H. & Yamada, T. (1970) *Nature* **227**, 27–34.
2. Engels, I.W. & Uhlmann, E. (1989) *Angew. Chem.* **28**, 716–734.
3. Dolinnaya, N.G., Sokolova, N.I., Gryaznova, O.I. & Shabarova, Z.A. (1988) *Nucleic Acids Res.* **16**, 3721–3737.
4. Sokolova, N.I., Ashirbekova, D.T., Dolinnaya, N.G. & Shabarova, Z.A. (1988) *FEBS Lett.* **232**, 153–155.
5. Shabarova, Z.A. (1988) *Biochimie* **70**, 1323–1334.
6. Moore, P.B., Traut, R.R., Noller, H.F., Pearson, P., Delius, N. (1968) *J. Mol. Biol.* **31**, 441–461.
7. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor University Press, Cold Spring Harbor.
8. Jinks-Robertson, S., Gourse, R.L. & Nomura, M. (1983) *Cell* **33**, 865–876.
9. Shelness, G.S. and Williams, D.L. (1985) *J. Biol. Chem.* **260**, 8637–8646.
10. Kagramanova, V.K., Derchacheva, N.I., Mankin, A.S. (1988) *Nucleic Acids Res.* **16**, 4158.
11. Messing, J. & Vieira, J. (1982) *Gene* **19**, 269–276.
12. Skripkin, E.A., Bogdanova, S.L., Kopylov, A.M. & Bogdanov, A.A. (1986) *Dokl. Akad. Nauk SSSR* **287**, 237–241.
13. Skripkin, E.A., Mankin, A.S., Kopylov, A.M. & Khudyakov, Y.E. (1987) *Biopolymers & Cell (Russ.)* **3**, 263–269.

14. Hartz, D., McPheeters, D.S., Traut, R. & Gold, L. (1988) in *Methods in Enzymology* (Noller H.F. and Moldave K., eds.) **164**, 419–425, Academic Press, N.-Y.
15. Balakin, A.G., Skripkin, E.A., Shatsky, I.N., van Duin, J. & Bogdanov, A.A. (1989) *Dokl. Akad. Nauk SSSR* **305**, 1000–1003.