

Mapping of T7 RNA polymerase active site with novel reagents – oligonucleotides with reactive dialdehyde groups

Vera L. Tunitskaya^a, Ekaterina E. Rusakova^a, Lyudmila V. Memelova^a,
Sergey N. Kochetkov^{a,*}, Arthur Van Aerschot^b, Piet Herdewijn^b, Ekaterina V. Efimtseva^a,
Boris S. Ermolinsky^a, Sergey N. Mikhailov^a

^aEngelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov str. 32, 117984 Moscow, Russia

^bRega Institute, Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Received 24 November 1998

Abstract Oligonucleotides of a novel type containing 2'-*O*-β-ribofuranosyl-cytidine were synthesized and further oxidized to yield T7 consensus promoters with dialdehyde groups. Both types of oligonucleotides were tested as templates, inhibitors, and affinity reagents for T7 RNA polymerase and its mutants. All oligonucleotides tested retained high affinity towards the enzyme. Wild-type T7 RNA polymerase and most of the mutants did not react irreversibly with oxidized oligonucleotides. Affinity labeling was observed only with the promoter-containing dialdehyde group in position (+2) of the coding chain and one of the mutants tested, namely Y639K. These results allowed us to propose the close proximity of residue 639 and the initiation region of the promoter within initiation complex. We suggest the oligonucleotides so modified may be of general value for the study of protein-nucleic acid interactions.

© 1999 Federation of European Biochemical Societies.

Key words: T7 RNA polymerase; Transcription; Promoter; Oligonucleotide; Dialdehyde group; Affinity labeling

1. Introduction

Transcription enzymes – DNA-dependent RNA polymerases – function as multicomponent complexes including the enzyme, promoter-containing DNA template, nucleotide substrates (rNTPs), and the synthesized RNA. In contrast to most RNA polymerases known, the bacteriophage T7 DNA-dependent RNA polymerase (T7 RNAP) consists of a single polypeptide chain and is able to perform the complete transcription cycle in the absence of any additional protein factors. Because of its molecular properties, this enzyme is widely used as a model in the study of transcription mechanisms [1]. Besides X-ray analysis [2], the most frequently used approaches for studying the enzyme are investigations of different mutants or otherwise modified forms of T7 RNAP [1,3]. These methods mostly give information about the enzyme alone, whereas there are only indirect data on its functioning within the active complex.

T7 RNAP-promoter interactions were studied in a number of recent works. Thus, Jorgensen et al. [4], using the method of chemical modification, demonstrated multiple contacts between the enzyme and nucleotide bases in the promoter

sequence (–5)–(–12). T7 RNAP was also shown to interact with the promoter at the transcription start region (coding chain) [5] and at the sequence (–10)–(–17) (non-coding chain) [6]. However, because of the lack of covalent bonds between template and enzyme, no amino acid (aa) residues contacting the promoter were founded. Mutagenesis studies have revealed contacts between aa residue 748 and nucleotides (–10) and (–11) (the so-called specificity triplet), as well as those of residue 758 and the nucleotide in position (–8) [7,8]. The involvement of aa sequence 563–571 in the enzyme-promoter interaction was also suggested [9].

As the active complex formed by T7 RNAP is rather unstable and cannot be isolated, the most suitable approach to map it is the formation of covalent bonds fixing the interaction between definite parts of the enzyme molecule and closely arranged areas of substrate, promoter and/or RNA product.

Periodate-oxidized nucleotides were extensively used for affinity modification of enzymes involved in nucleic acid metabolism [10]. This labeling involves the formation of an unstable dihydroxymorpholine derivative which may be reduced to stable morpholine compounds (Fig. 1).

Recently, we have developed for the first time the preparation of oligonucleotide (ON) derivatives with regiospecifically incorporated dialdehyde reactive groups and their successful use as affinity labels for restriction endonucleases and DNA methyltransferases [11,12].

This paper describes the simultaneous mapping of amino acids and nucleotides involved in the T7 RNAP-promoter interaction. We used synthetic promoters whose cytidine residues were modified so that periodate oxidation and following treatment with NaBH₄ made possible specific covalent binding of the oxidized group with the closely arranged lysine residues of the enzyme (Fig. 1), provided that such a proximity actually takes place in the native complex structure.

2. Materials and methods

2.1. Chemical synthesis

*N*⁴-Benzoyl-1-[5-*O*-dimethoxytrityl-2-*O*-(2,3,5-tri-*O*-benzoyl-β-D-ribofuranosyl)-β-D-ribofuranosyl]cytosine (**2**) was prepared starting from *N*⁴-benzoyl-1-[2-*O*-(2,3,5-tri-*O*-benzoyl-β-D-ribofuranosyl)-β-D-ribofuranosyl]cytosine (**1**) [13]. The structure of **2** was confirmed by NMR and mass spectra.

*N*⁴-Benzoyl-1-[5-*O*-dimethoxytrityl-3-*O*-(*P*-β-cyanoethyl-*N,N*-diisopropylaminophosphinyl)-2-*O*-(2,3,5-tri-*O*-benzoyl-β-D-ribofuranosyl)-β-D-ribofuranosyl]cytosine (**3**) (the synthon for ON synthesis) was prepared by a standard procedure [13]. For C₇₂H₇₂N₅O₁₆P calc. 1293.4712. LSIMS (NBA) *m/z* 1294 (M+H⁺, 1), 303 (DMTr, 100). The structure of **3** was confirmed by ¹H NMR and ¹³C NMR spectra. ³¹P NMR: (ppm, external ref. = H₃PO₄ capil.) 150.53, 151.73.

*Corresponding author. Fax: (7) (095) 135 14 05.

E-mail: kochet@genome.eimb.rssi.ru

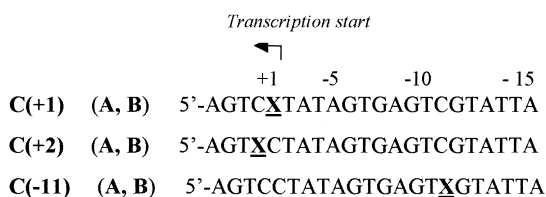
Abbreviations: T7 RNAP, bacteriophage T7 RNA polymerase; ON, oligonucleotide; aa, amino acid

2.2. Oligonucleotide synthesis

This was performed on an ABI 392 synthesizer at 1 mmol scale using commercial 2-cyanoethylphosphoramidites and standard methodology, but for a longer coupling time (80 s) and at a higher concentration (0.15 M) for the modified synthone to ensure high coupling yields. The ONs obtained were deprotected, removed from the solid support and purified as described [11,13]. The structure of the ONs (**4a**, **5a**, and **6a**) was proved by MALDI spectra: for $C_{222}H_{281}N_{80}O_{137}P_{21}$ Calc. 6912.58. Found ($M+H^+$): 6913.1, 6912.9, and 6913.4.

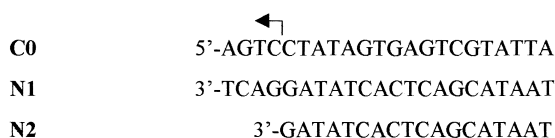
The following ONs were synthesized.

2.2.1. Modified ONs



In **C(+1)** X=4 (series A, see Fig. 1), in **C(+2)** X=5 (series B).

2.2.2. Natural ONs



2.3. Periodate oxidation

$NaIO_4$ treatment of modified ONs was carried out as described earlier [11]. An equal volume of 0.01 M aqueous solution of $NaIO_4$ was added to the ON solution (approximately 10 OD units/ml, probe volume 10 μ l) in TE buffer (10 mM Tris-HCl, pH 7.9, 1 mM EDTA). After incubation for 2 h at 20°C in the dark, oxidized ONs were precipitated by adding 10 volumes of 2% $LiClO_4$ in acetone.

2.4. ^{32}P -Labeling of the oligonucleotides

This was performed according to Promega protocols [14]. After labeling, ONs were precipitated with 2% $LiClO_4$ in acetone as described above.

After 30 min at $-20^\circ C$ the precipitate was centrifuged, washed twice with cold acetone and air-dried. When the periodate-oxidized ONs were used, the labeling preceded the oxidation.

2.5. Oligonucleotide duplexes

Single-stranded ONs (both native and modified) were mixed in equimolar ratios and 0.1 volume of 10 \times buffer (200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 500 mM NaCl) was added. The mixture was heated for 1–2 min at 65°C, then slowly cooled and stored at $-20^\circ C$. With the periodate-oxidized ONs, the equimolar mixture in the same buffer was either heated at 37°C for 1 min or stored for 3 h at room temperature or overnight at $+4^\circ C$.

The duplexes formed by corresponding single-stranded ONs are denoted as their combinations (i.e. **C0/N1** is the non-modified promoter with blunt ends).

2.6. Enzyme purification and assays

The wild-type T7 RNAP and its mutants were isolated from *Escherichia coli* strains and purified as described previously [15]. The enzyme activity using synthetic promoter-containing ONs was determined as described in [16].

2.7. Binding parameters

The enzyme-promoter binding for unproductive duplexes was estimated by two methods: (a) measurement of inhibitory effect of the modified duplex on transcription with control (unmodified) duplexes for the wild-type enzyme (K_i); (b) direct measurement of promoter binding using sorption on nitrocellulose filters (K_s) [16]. The constant

of covalent binding of mutant enzyme to the modified promoter (K_i irr.) was determined by scanning the proper bands in PAGE under variable concentrations of the labeled promoter using the FAST SCAN 300A densitometer (Molecular Dynamics, USA).

2.8. Affinity labeling

Duplexes containing ^{32}P -labeled periodate-oxidized ONs were incubated with equimolar quantities of the enzyme (wild-type or mutant) in a solution containing 25 μ l of 50 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 5 mM β -mercaptoethanol, and, in some cases, 0.4 mM GTP or mixed rNTPs. After 30 min (37°C), 1 μ l of freshly prepared aqueous 0.01 M $NaBH_4$ was added. The reaction was carried out for 1 h in the dark at room temperature, then the probe was diluted with an equal volume of protein electrophoresis buffer (0.1 M Tris-HCl, pH 6.8, 5% SDS, 20 mM EDTA, 20% glycerol, 5% β -mercaptoethanol, 0.1% bromophenol blue). After electrophoresis (15% PAGE), the gels were autoradiographed at the X-ray film at $-70^\circ C$.

3. Results and discussion

Recently we used modified DNA duplexes prepared by sodium periodate oxidation starting with ONs containing 1-(β -D-galactopyranosyl)-thymine as affinity labels for restriction endonucleases and DNA methyltransferases [11,12]. Here we describe the preparation of ONs with dialdehyde groups starting from 2'-O- β -D-ribofuranosyl nucleosides [11,13] (**1**, Fig. 1). Compound **1** was converted via its dimethoxytrityl derivative to the corresponding phosphoramidite which was then used for preparation of ONs **C(+1)A**, **C(+2)A**, and **C(-11)A** using standard methodology. It should be noted that the model dinucleoside phosphates containing 2'-O- β -D-ribofuranosyl nucleosides were oxidized by sodium periodate much more readily than the dinucleoside phosphates containing 1-(β -D-galactopyranosyl)thymine [11]. The proposed method of introduction of dialdehyde function into ONs is of general value and has several advantages over the previously known schemes: simple preparation of ONs containing 2'-O- β -D-ribofuranosyl nucleosides with their following oxidation.

Synthetic double-stranded ONs containing the T7 consensus promoter and five bases in the message region were used as templates for T7 RNAP (Table 1). Non-coding chains of these duplexes were of two types: 18 nucleotides in length (promoter and one base in the message region) or 22 bases (promoter and five bases in the message). The coding chain of the template contained the modified cytidine residues in posi-

Table 1
Affinity constants (μ M) for synthetic promoter duplexes

Template	Affinity constant (K_m or K_i (K_s) ^a)	
	wild-type T7 RNAP	Tyr639Lys
C0/N1	320 \pm 160	–
C0/N2	630 \pm 270	–
C(+1)A/N1	90 \pm 10 ^a	160 \pm 30 ^a
C(+1)A/N2	90 \pm 10 ^a	260 \pm 110 ^a
C(+1)B/N1	130 \pm 70 ^a	130 \pm 80 ^a
C(+1)B/N2	180 \pm 80 ^a	230 \pm 110 ^a
C(+2)A/N1	150 \pm 50 ^a	160 \pm 70 ^a
C(+2)A/N2	non-linear	non-linear
C(+2)B/N1	140 \pm 80 ^a	460 \pm 120 ^b
C(+2)B/N2	non-linear	510 \pm 150 ^b
C(-11)A/N1	120 \pm 40	n.d.
C(-11)A/N2	660 \pm 200	n.d.

For denominations of templates see the legend to Fig. 2.

^aBinding constants for unproductive duplexes were measured as K_i or K_s (see Section 2); the difference between the values did not exceed 15%.

^bIrreversible inhibition (see text).

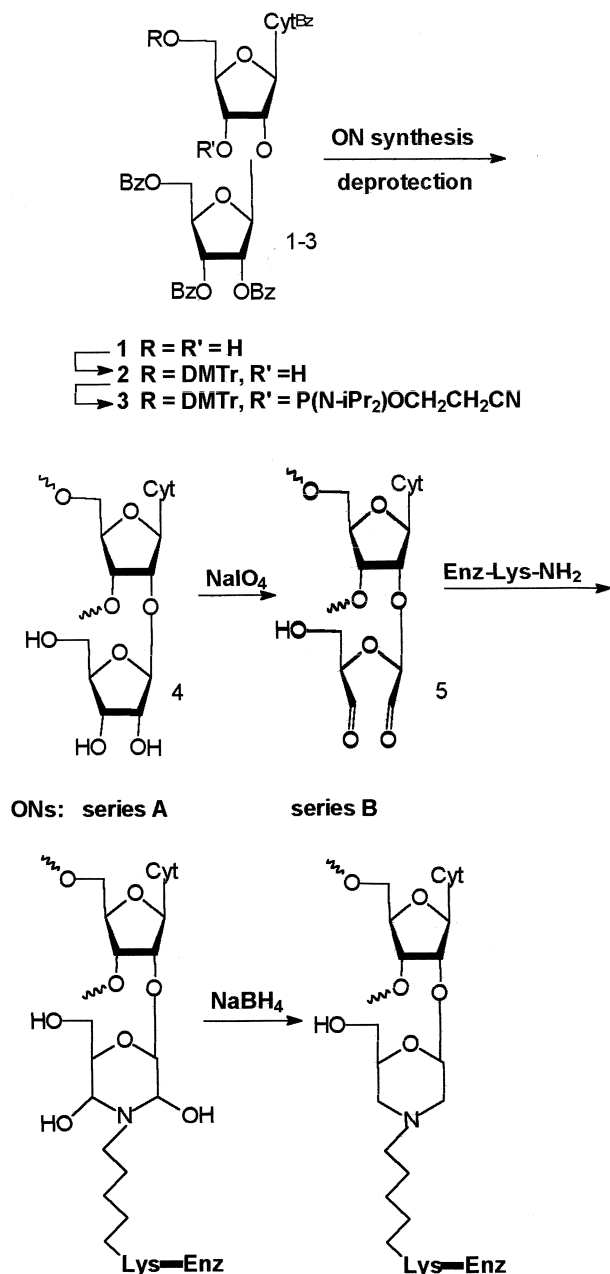


Fig. 1. Synthesis of modified ONs and affinity labeling of T7 RNAP.

tions (+1), (+2) and (-11) relative to the transcription start. Two first positions correspond to the region of RNA synthesis initiation. The last position resides within the so-called specificity triplet involved in promoter recognition [7]. Both control (unmodified) and three modified ONs were annealed with two types of different length non-coding chains forming either blunt- or sticky-ended templates. In addition, the modified duplexes may be presented in both the original and periodate-oxidized forms (series A and B, Fig. 1).

Fig. 2 demonstrates PAGE separation of the products of the wild-type T7 RNAP transcription for native and modified templates. As can be seen, no correct RNA synthesis occurs with modified templates. This observation correlates with the data reported in [17] that modifications in position (+1) resulted in the lack of transcription. As can be seen from Fig. 2,

insertion of the modified cytidine residue in position (+2) (C(+2)) also leads to the disappearance of the enzyme activity, whereas in the case of modification in (+1), the effect depends on the type of template ends. When the duplex with blunt ends C(+1)A/N1 was used, no transcripts were observed, while in the case of the sticky-ended duplex C(+1)A/N2 non-complementary dinucleotides were synthesized instead of extended products. Since the labeled NTP used in the experiment was [³²P]ATP, it should be incorporated into position (+3) of the transcript. Nevertheless, the main reaction products were labeled dinucleotides (Fig. 2, lanes 11, 12). So, with the modified duplex C(+1)A/N2, the incorporation of AMP instead of GMP happens in the (+2) position. Promoters modified near the initiation site were earlier shown to facilitate the misincorporation [17].

On the other hand, duplexes of all types inhibited the reaction due to competition with unmodified templates (Table 1). For several modified duplexes K_i values were even lower than those for control templates. The same regularities were also shown when the binding constants (K_s) were measured by direct sorption on nitrocellulose filters. One of the studied duplexes, C(+2)/N2, in both initial (A) and oxidized (B) forms, showed non-linear kinetics of binding, thus creating obstacles to the determination of binding constant. Thus, all duplexes tested completely retain the affinity to the enzyme and may be of general value for the study of protein-nucleic acid interactions.

This fact allowed us to study the possibility of affinity labeling of the enzyme with the modified promoters after NaBH₄ reduction. However, PAGE did not reveal any covalent complexes of the wild-type T7RNAP and ³²P-labeled modified promoters (Fig. 3).

The X-ray structure of T7 RNAP suggests the binding of the DNA template within the cleft formed by 'palm', 'thumb' and 'fingers' subdomains of the enzyme. Inside this cleft, structural motifs A, B, and C [18], conservative for most single-subunit nucleotide polymerases and containing functionally essential aa residues, are located. These residues form a putative active site of T7 RNAP. Among these motifs, motif B

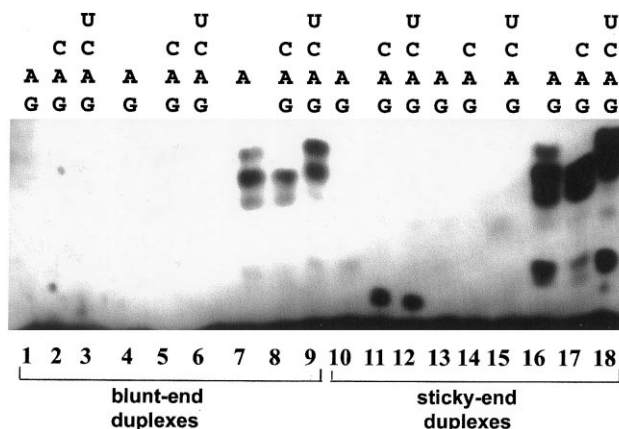


Fig. 2. RNA synthesis by wild-type T7 RNAP on native and modified promoter duplexes. [³²P]GTP was used as labeled NTP. Lanes 1–9 and 10–18 correspond to duplexes with blunt (non-coding chain, N1) and sticky (non-coding chain, N2) ends. Lanes: 1–3, duplex C(+2)/N1; 4–6, C(+1)/N1; 7–9, C0/N1 (control); 10–12, C(+2)/N2; 13–15, C(+1)/N2; 16–18, C0/N2 (control). NTPs added to the incubation mixtures are indicated in the upper part of the figure.

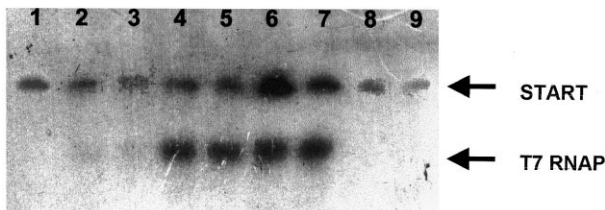


Fig. 3. Affinity labeling of T7 RNAP with periodate-oxidized modified promoter duplexes with subsequent NaBH_4 reduction. Lanes: 1, wild-type T7 RNAP, duplex C(+2)/N1; 2, Tyr639Lys T7 RNAP, C(+1)/N1; 3, Tyr639Lys T7 RNAP, C(+1)/N2; 4, Tyr639Lys T7 RNAP, C(+2)/N1; 5, Tyr639Lys T7 RNAP, C(+1)/N1, 4NTPs (400 μM); 6, Tyr639Lys T7 RNAP, C(+1)/N2; 7, Tyr639Lys T7 RNAP, C(+1)/N2, 4NTPs (400 μM); 8, Tyr639Asn T7 RNAP, C(+2)/N2; 9, Tyr639Phe T7 RNAP, C(+2)/N2.

is assumed to be of great importance for initiation of RNA synthesis [1,3,19]. The Lys-631, Tyr-639 and Gly-640 residues, which are invariant in all single-subunit RNA polymerases and most DNA polymerases, belong to this motif. Since the wild-type enzyme was not labeled with ONs containing a periodate-oxidized ribose group which is highly specific for lysine residues, Lys-631 is evidently located neither near the 'specificity triplet' (position -11) nor near the site of initiation of RNA synthesis (positions $+1$ and $+2$). The lack of labeling may be due to the absence of such residues in the contact area.

At the same time, another conserved residue of motif B, Tyr-639, is apparently involved in the NTP binding [19,20] and therefore should be located in close proximity to the initiation site. To test this hypothesis, we have taken the opportunity of using the recently obtained T7 RNAP mutant containing the Tyr \rightarrow Lys substitution in position 639 [21]. The latter mutant is partially active (about 10% of wild-type T7 RNAP), and completely retains the affinity to the promoter (Table 1). Fig. 3 demonstrates that the duplexes modified in the $(+1)$ position did not react covalently with this mutant whereas those modified in $(+2)$ gave the ^{32}P -labeled covalent adduct upon NaBH_4 reduction. Neither the duplex type (blunt or sticky ends) nor the addition of NTPs affected the labeling. The affinity constant for irreversible binding (K_i irr.) (see Section 2) was $460 \pm 120 \mu\text{M}$ (Table 1).

The data obtained allow us to suppose that Lys-639 is labeled with the modified promoter and, consequently, the residue in position 639 is located near the initiation site of T7 RNAP. However, another explanation of these results could be also proposed: Tyr \rightarrow Lys substitution could result in alterations of the enzyme tertiary structure allowing labeling of any other amino acid residue. Taking this into consideration, we used as controls other T7 RNAP mutants containing substitutions in position 639, namely Tyr639Asn, Tyr639Phe [21], and the 'double' mutant Tyr639Phe,Ser641Ala [22]. As is seen from Fig. 3, none of these mutants demonstrated the cross-linking with the modified promoter. Thus, we have good reason to suppose that the mutations in this region per se do not lead to perturbations in T7 RNAP structure and the affinity labeling is indeed in the lysine residue in position 639 of the protein molecule.

As follows from the obtained data, the affinity labeling happens with promoter duplex bearing the modified group in position $+2$, but not in position $+1$. This result suggests rather strict geometry of the reaction complex and argues in

favor of the labeling specificity. Probably, the bulky ribose moiety of modified cytidine is located near the putative site of the second incoming NTP. Tyr-639 was shown to participate in such interactions possibly playing a certain role in ribo/deoxy-NTP discrimination.

The data on X-ray structure of T7 RNAP [2] at the relatively low resolution of 3.3 Å contain no information concerning the bound ligands. Nevertheless, the model studies suggest that the distance between C α atoms of residue 758 (which was shown to interact with the nucleotides in position -8 of the promoter) [8] and residue 639 $- 36 \text{ \AA}$ \approx roughly corresponds to the distance between the plane of the heterocyclic bases in position -8 and oxygen atoms of the 'additional' ribose of the $+2$ cytidine of duplex C(+2)A/N2. This distance is approximately 31–35 Å. Thus, the modified duplex may, in principle, fit correctly in the template binding cleft. Certainly, the determination of the three-dimensional structure of T7 RNAP complexed with promoter and some NTP derivative could clarify this issue.

Acknowledgements: The authors are grateful to Dr. S.V. Kochetkova for the synthesis of the unmodified ONs. The research was supported by Grants of the Howard Hughes Medical Institute (75195-545003), INTAS-RFBR (IR-97-1635), NATO (HTECH.LG 961324) and Russian Foundation for Basic Research (97-04-49376 and 98-04-48194). A.V.A. is a research associate of the Fund for Scientific Research of Flanders.

References

- [1] Chamberlin, M.J. and Ryan, T. (1982) *Enzymes* 15, 87–108.
- [2] Sousa, R., Chung, Y.J., Rose, J.P. and Wang, B.-C. (1993) *Nature* 364, 593–599.
- [3] Kostyuk, D.A., Lyakhov, D.L., Memelova, L.V., Rechinsky, V.O., Tunitskaya, V.L. and Kochetkov, S.N. (1996) in: *Chemical Modification of Enzymes* (Kurganov, B.I., Nagradova, N.K. and Lavrik, O.I., Eds.), pp. 347–387, Nova Science, New York.
- [4] Jorgensen, E.D., Durbin, R.K., Risman S, S.S. and McAllister, W.T. (1991) *J. Biol. Chem.* 266, 645–651.
- [5] Muller, D.K., Martin, C.T. and Coleman, J.E. (1989) *Biochemistry* 28, 3306–3313.
- [6] Martin, C.T. and Coleman, J.E. (1987) *Biochemistry* 26, 2690–2696.
- [7] McAllister, W.T. (1993) *Cell. Mol. Biol. Res.* 39, 385–391.
- [8] Rong, M., He, B., McAllister, W.T. and Durbin, R.K. (1998) *Proc. Natl. Acad. Sci. USA* 95, 515–519.
- [9] Rechinsky, V.O., Chernov, B.K., Dragan, S.M., Kostyuk, D.A., Tunitskaya, V.L. and Kochetkov, S.N. (1995) *Mol. Gen. Genet.* 247, 110–113.
- [10] Colman, R.F. (1983) *Annu. Rev. Biochem.* 52, 67–91.
- [11] Brevnov, M.G., Gritsenko, O.M., Mikhailov, S.N., Efimtseva, E.V., Ermolinsky, B.S., Van Aerschot, A., Herdewijn, P., Repyk, A.V. and Gromova, E.S. (1997) *Nucleic Acids Res.* 25, 3302–3309.
- [12] Efimtseva, E.V., Ermolinsky, B.S., Fomitcheva, M.V., Meshkov, S.V., Padyukova, N.Sh., Mikhailov, S.N., Van Aerschot, A., Rozenski, J. and Herdewijn, P. (1996) *Coll. Czech. Chem. Commun.* 61, (Special Issue) 206–209.
- [13] Mikhailov, S.N., Efimtseva, E.V., Gurskaya, G.V., Zavodnik, V.E., De Bruyn, A., Rozenski, J. and Herdewijn, P. (1997) *J. Carbohydr. Chem.* 16, 75–92.
- [14] Promega (1991) *Protocol and Application Guide*, 2nd edn., Promega Corporation, Madison, WI.
- [15] Tunitskaya, V.L., Luchin, S.V., Memelova, L.V., Lyakhov, D.L., Rechinsky, V.O. and Kochetkov, S.N. (1989) *Mol. Biol. (Moscow)* 23, 1273–1278.
- [16] Rusakova, E.E., Tunitskaya, V.L., Memelova, L.V., Kochetkova, S.V., Kostyuk, D.A. and Kochetkov, S.N. (1998) *FEBS Lett.* 423, 189–192.

- [17] Weston, B.F., Kuzmine, I. and Martin, C.T. (1997) *J. Mol. Biol.* 272, 21–30.
- [18] Delarue, M., Poch, O., Tordo, N., Moras, D. and Argos, P. (1990) *Protein Eng.* 3, 461–467.
- [19] Sousa, R. and Padilla, R. (1995) *EMBO J.* 14, 4609–4621.
- [20] Tunitskaya, V.L., Memelova, L.V., Kostyuk, D.A., Gudima, S.O. and Kochetkov, S.N. (1997) *Mol. Biol. (Moscow)* 31, 298–302.
- [21] Tunitskaya, V.L., Dragan, S.M., Kostyuk, D.A., Lyakhov, D.L., Memelova, L.V., Rechinsky, V.O. and Kochetkov, S.N. (1994) *Biochemistry (Moscow)* 59, 355–361.
- [22] Kostyuk, D.A., Dragan, S.M., Lyakhov, D.L., Rechinsky, V.O., Tunitskaya, V.L., Chernov, B.K. and Kochetkov, S.N. (1995) *FEBS Lett.* 369, 165–168.