Russian Chemical Reviews **67** (5) 435 – 450 (1998) © 1998 Russian Academy of Sciences and Turpion Ltd

UDC 577.213 : 577.215

Chemical approaches to the elucidation of template biosynthesis: study of replication and reverse transcription

D G Knorre, N V Kudryashova, O I Lavrik

Contents

I.	Introduction	435
II.	The main types of affinity reagents used for the study of replication and reverse transcription	436
III.	Study of the interaction of DNA polymerases with templates, primers, and nucleoside 5′-triphosphates using group-specific and affinity-based chemical modification. Structural aspects	439
IV.	Catalytically competent labelling of DNA polymerases	443
V.	Study of interactions of DNA polymerases with templates, primers and nucleoside 5′-triphosphates by affinity-based chemical modification. Quantitative aspects	444
VI.	Chemical methods for the study of replicative complexes	446
VII.	Conclusion	448

Abstract. Chemical approaches to the elucidation of DNA tem plate biosynthesis using DNA polymerases and reverse tran scriptases, predominantly DNA polymerase I from *E. coli* and reverse transcriptase from HIV-1, are considered. Special atten tion is paid to the comparison of chemical modification and X-ray structural analysis data as well as to quantitative aspects of the interaction of DNA polymerases with templates, primers, and nucleoside 5′-triphosphates. New versions of a catalytically competent labelling method developed in the course of investigation of these enzymes are described. The data obtained in the analysis of replication complexes using primer analogues with photoreactive groups are presented. The bibliography includes 98 references.

I. Introduction

Some general problems concerning the application of chemical approaches to the study of the mechanism of template biosynthesis of nucleic acids and the main results obtained in the study of enzymes that catalyse RNA biosynthesis (transcription) by chemical methods have been reviewed.[1] The present review is devoted to the application of chemical methods to the study of DNA biosynthesis, namely, its replication and reverse transcription.[2]

The main chemical event in this process, *viz*., the elongation of the polynucleotide chain, in the case of DNA polymerases is fully identical with that for RNA polymerases considered previously.[1] A nucleoside triphosphate ($pppN_i$) complementary to the translatable element $p\underline{N}_i$ of the template (its role is played by single-stranded DNA) phosphorylates the 3′-OH group of the 3′-terminal nucleotide residue of the growing DNA chain. It is in this way that the nucleotide residue is transferred from the triphosphate to the 3′-OH group of the growing polymeric chain; therefore the enzyme represents a nucleotidyl transferase.

$$(5')\ \ldots\ldots pN_{i-2}-pN_{i-1}-OH \quad (3')$$

$$(3')\ \ldots\ldots p\underline{N}_{i-2}-p\underline{N}_{i-1}-p\underline{N}_{i}-p\underline{N}_{i+1}-p\underline{N}_{i+2}\ldots\ldots\ (5')$$

$$\qquad\qquad\qquad\qquad pppN_i \downarrow$$

$$(5')\ \ldots\ldots pN_{i-2}-pN_{i-1}-pN_{i}-OH \quad (3')$$

$$(3')\ \ldots\ldots p\underline{N}_{i-2}-p\underline{N}_{i-1}-p\underline{N}_{i}-p\underline{N}_{i+1}-p\underline{N}_{i+2}\ldots\ldots\ (5') + pp,$$

where p is the orthophosphate residue; N_i and $\underline{N}_i$ are the nucleoside residues in the product and the template, respectively, and i is the ordinal number of the residue in the polynucleotide chain starting from the 5′-end of the product and the 3′-end of the template, respectively.

However, this is virtually the last point where the similarity between DNA polymerases and RNA polymerases ends. RNA polymerases interact with double-stranded DNA, which are temporarily unwound under the action of this very enzyme to form a non-extended unwound region (protrusion). One of the unwound strands (on which transcription occurs) becomes involved in the selection of a substrate in the active centre. If the replication is accomplished by DNA-dependent DNA polymerases, the double-stranded DNA is first unwound by a special system mentioned below. DNA polymerase is not directly involved in DNA unwinding.

Transcription begins at a predetermined site of the template and requires preliminary binding of the enzyme in the region where the transcription begins (the promoter region). Following the formation of a complex with an enzyme and specific transcription factors at the required region of DNA subject to transcription the process is initiated, *i.e*., the synthesis of a dinucleotide occurs. It should be noted that the first dinucleotide residue retains a triphosphate group. Initiation of DNA synthesis in living organisms always begins with the synthesis of a ribooligonucleotide primer the 3′-terminal OH group of which serves as an acceptor of deoxyribonucleotide 5′-phosphate. It is not until this process has been completed that the synthesis of a new DNA chain begins. The 3′-terminal deoxyribonucleotide fragment now acts as a primer. A special enzyme, primase, in prokaryotes and eukaryotes is responsible for the synthesis of a ribooligonucleotide

D G Knorre, N V Kudryashova, O I Lavrik Novosibirsk Institute of Bioorganic Chemistry, Siberian Branch of the Russian Academy of Sciences, prosp. Akad. Lavrent'eva 8, 630090 Novosibirsk, Russian Federation. Fax (7-383) 235 16 65. Tel. (7-383) 235 64 61 (D G Knorre)

Received 26 September 1997
Uspekhi Khimii **67** (5) 486 – 502 (1998); translated by R L Birnova

primer; therefore the primary product is a hybrid oligonucleotide, which has a ribooligonucleotide at its 5′-end.

The region where the direct chemical conversion takes place contains a complex of the 3′-terminal nucleotide of the primer with a complementary nucleotide of the template, the protruding single-stranded fragment of the template and the selected substrate. Therefore, it makes no difference whether we deal with an exogenous primer or the terminus of a newly synthesised strand.

In the case of RNA polymerases, the product is a single-stranded RNA fixed by complementary interactions to the strand of the template subjected to transcription only at its 3′-end. In the case of DNA polymerases, a duplex is formed, which is preserved in biological systems up to the beginning of the next replication cycle. In this context, the range of problems that can be solved by chemical methods in respect of DNA polymerases differs from that for RNA polymerases described in the review cited.[1] It is also noteworthy that X-ray data about the spatial structure of DNA polymerases and their complexes with templates and substrates are now more abundant. Therefore, it is expedient to consider the large body of experimental results concerning the chemical (including affinity-based) modification of DNA polymerases in relation to the X-ray data available. Much of the above can be said of the enzymes responsible for reverse transcription, *i.e.*, to RNA-dependent DNA polymerases (reverse transcriptases).

While considering DNA polymerases, special mention should be made of two groups of enzymes differing in their biological functions. One of them is responsible for the synthesis of new DNA molecules, which occurs *via* separation of strands of the parent DNA and the formation of daughter replicas on each of the separated strands. This results, with the accuracy of inevitable errors, in the formation of two double-stranded structures that are identical with the original one. However, this process requires the participation of several additional enzymes and proteins. The unwinding of the double helix is effected by enzymes called helicases. They function in the branching region of the original double helix, namely, in the replication fork. The recombination of the two strands is prevented owing to their binding to a special protein, most commonly referred to as the SSB (single strand binding) protein. Continuous synthesis of a complementary chain towards the replication fork takes place on one of the strands with a $5' \rightarrow 3'$ orientation from the replication fork. This strand is usually named the leading strand. The synthesis on the second strand can proceed only in the direction away from the replication fork, and a new initiation must take place at certain invervals of time with the liberation of the template. In other words, the synthesis of a new RNA primer, which is further elongated in the direction away from the replication fork, takes place. Thus, the fragments of the daughter strand with a ribonucleotide at the 5′-end (Okazaki fragments) are the products of synthesis on this lagging strand. A hybrid DNA-RNA duplex is formed at the primer site. Its RNA fragment must be removed from the final replication product. This is achieved, at least partly, with a special enzyme, RNAse H, which catalyses the hydrolysis of ribooligonucleotides within this hybrid duplex. The preceding Okazaki fragment can then be elongated with DNA polymerase to reach the 5′-end of the second chain. As a result, the lagging daughter strand is formed with ruptures called nicks. Linking of 5′-phosphate with the 3′-OH group of the preceding fragment requires yet another enzyme, DNA ligase, which catalyses the formation of a phosphodiester bond between these two groups. Thus, a minimum set of replication enzymes and auxiliary proteins includes DNA polymerase itself, helicases, the SSB protein, RNAse H, and DNA ligase.

As a matter of fact, this set is even greater. An important feature of the replication process is its processivity. After elongation, the 'template-elongated primer' system is not separated from the enzyme. Instead, a new encoding element of the template and the 3′-end of the elongated primer are transposed relative to the enzyme in such a way that they appear in the active centre of the enzyme. There is a large body of evidence that the processivity is provided by special auxiliary proteins that consolidate the enzyme complex with the template-primer duplex. These proteins are called 'clamps'.[3]

The second group of DNA polymerases cannot synthesise complete dauthter replicas, but they are necessary to fill short gaps that often appear on one of DNA strands. The classical example of such gaps is the excision of damaged regions of one of DNA strands by special enzymes that play an important role in restoring injured sites (DNA repair). Therefore, DNA polymerases pertaining to this group are often regarded as DNA repair enzymes, although their functions are hardly confined to repair. DNA polymerase I from *E. coli* and human DNA polymerase β are the most well-studied enzymes of this group. The spatial structure of these two enzymes has been established by X-ray structural analysis.

II. The main types of affinity reagents used for the study of replication and reverse transcription

Application of affinity-based modification to the study of enzymes is particularly efficient in two cases. First, when it is necessary to determine the position of a site(s) of modification of an enzyme by a reactive substrate or effector analogue, which points either to a subunit or a region in the protein molecule that interacts with the substrate or the effector. Second, when it is necessary to determine the affinity of a substrate or a cofactor for the protein from the data on the competition between the reagent and the substrate (or the cofactor).[4] The second case will be considered in more detail in Section V.

DNA polymerases and reverse transcriptases interact with the template, the primer, and deoxynucleoside 5′-triphosphate (further referred to as dNTP). In the case of DNA polymerases, the strands of a double-stranded DNA are first unwound by helicases. Thus, single-stranded DNA serves as a template for replication. There is every reason to believe that reactive derivatives of DNA modify the enzyme at the binding site of the template.

One of the strands in double-stranded structures corresponds to the template, and the other one to the primer. If a complex of two DNA fragments has both a double-stranded and a hanging one-stranded element and the shorter fragment at their junction has a free 3′-OH group, it is this fragment that functions as a primer. Complex **1** is an example which can be useful for further consideration:

5′ d(TCGCAGC CGTC CA)
3′ d(AGCGTC GGCAGGTTCCCAAA)

1

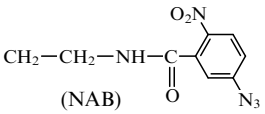
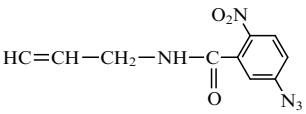
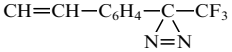
Apparently, it is the upper strand in this complex that plays the role of a primer. However, the situation with dNTP analogues is not so unambiguous. Even in the framework of a simplified scheme implying the existence of distinctly separated binding sites for a template, a 3′-terminal fragment of the template, and a dNTP, a dNTP analogue will not necessarily form a complex with the binding site of the substrate. It will be shown later (see Section V) that dNMP and dNDP also possess affinity for the binding site of the template. As a result, modification of an enzyme by such an analogue will not allow the unambiguous assignment of the modified region of the protein to the binding site of dNTP because it may be the binding site of the template as well.

Introduction of reactive groups into dNTP can be easily carried out at the γ-phosphate[5] *via* nucleoside 5′-trimetaphosphate. Using this route, derivatives containing 4-(*N*-2-chloroethyl-*N*-methylamino)benzylamine,[6,7] imidazole,[8,9] and *p*-azidoaniline residues[10] were obtained. In the case of alkylating and phosphorylating derivatives, the inactivation proceeded only in the presence of a template-primer complex. This can be regarded as evidence in favour of analogue binding in the substrate-binding site.[6–9] In the case of a photoaffinity reagent (*p*-azidoaniline derivative), inactivation was also observed upon irradiation in the absence of the complex, but its presence enhanced inactiva-

tion.[10] A wide variety of dNTP analogues with a photoreactive group in the heterocyclic base have been recently synthesised. Derivatives containing a photoreactive group at position 4 of cytosine, at position 5 of uracil (dTTP analogues), and at positions 2 and 8 of adenine have been described and tested.[11–20] Table 1 lists the presently known dNTP containing reactive groups in their heterocyclic bases. It is important to mention that some of these compounds can act as substrates for DNA polymerase. This property makes it possible to use these analogues for specific photoaffinity-based modification of DNA polymerases.

Photoreactive analogues must possess good photochemical characteristics that allow their irradiation at the wavelengths that cause no photoinactivation of proteins and nucleic acids.[11] At the same time, these analogues manifest good substrate properties in

Table 1. Derivatives of dNTP containing photoreactive groups in the heterocyclic base.

Photoreactive group	Position of the group in the heterocycle	Substrate properties	Modification (inactivation)
N_3	Ade-8	–	Modification of Ile765 and Tyr766 in DNA polymerase I from *E. coli* in the absence of a template – primer duplex [17]
$SCH_2CC_6H_4N_3$-*p* (O)	Ade-2	Competitive inhibitor with respect to dTTP [19]	Inactivation of the Klenow fragment in the absence of a template – primer duplex [18] Inhibition of modification in the presence of poly(dA) – oligo(dT) [18] Labelling of Asp732 [18]
$CH_2-CH_2-NH-C(=O)-C_6F_4N_3$-*p* (FAB)	Cyt-4	Substrate of DNA polymerase from human α-primase instead of dTTP [21] Substrate of reverse transcriptase from HIV-I instead of dCTP [13] and dTTP [16]	Modification of the 165, 59, and 49 kDa subunits (the human complex) [23] Modification of the 165 and 49 kDa subunits (the human complex, competent labelling) [23] Modification of the 150, 59, and 48 kDa subunits (the calf thymus complex) [23] Modification of p66 and p51 [14] Modification of p66 (competent labelling) [14]
$CH=CHCH_2NHCC_6F_4N_3$-*p* (O) (AlFAB)	Cyt-4	–	Decrease in substrate properties and enhancement of terminating properties
 (NAB)	Cyt-4	Substrate of DNA polymerase from α-primase of calf thymus [15] Substrate of reverse transcriptase from HIV-1 instead of dCTP and dTTP [16]	Decrease in substrate properties and enhancement of terminating properties
$(CH_2)_2NHCCH_2ON=CHC_6F_4N_3$-*p* (O)	Cyt-4	Substrate of reverse transcriptase from HIV-1 instead of dCTP and dTTP [16]	Decrease in substrate properties and enhancement of terminating properties
$C(CH_2)_3ON=CHC_6F_4N_3$-*p* (O)	Cyt-4	Substrate of reverse transcriptase from HIV-1 instead of dTTP [16]	Decrease in substrate properties and enhancement of terminating properties
$NHC(CH_2)_4CNH-N=CHC_6F_4N_3$-*p* (O, O)	Cyt-4	Substrate of reverse transcriptase from HIV-1 instead of dCTP and dTTP [16]	Decrease in substrate properties and enhancement of terminating properties
$CH=CHCH_2NHCC_6F_4N_3$-*p* (O)	Cyt-5	Substrate of DNA polymerase from α-primase of calf thymus [23] Substrate of reverse transcriptase from HIV-1 [14]	Modification of p66 and p51 Modification of p66 (competent labelling) [14]
N_3	Ura-5	–	Modification of DNA polymerase I from *E. coli* in the absence of a template – primer duplex [24]
	Ura-5	Substrate of reverse transcriptase from HIV-1 instead of dTTP [15]	Modification of p66 and p51
	Ura-5 (ddUTP)	Substrate of reverse transcriptase from HIV-1 instead of dTTP [20]	Modification of p66

reactions catalysed by DNA polymerase a from human placenta,[21] reverse transcriptase from HIV,[13–16] and DNA polymerase β.[12] These substrates can act as terminators by preventing further elongation of α primer after being incorporated into its 3′-end. In other cases, the presence of a modified residue at the 3 -end does not prevent the addition of the next unit, *i.e.*, further elongation of the chain. These derivatives can be used for conducting the polymerase chain reaction (PCR), which yields a duplex with a large number of reactive groups.[22] The corresponding characteristics of reactive dNTP derivatives are listed in Table 1.

Despite the apparent similarity of analogues containing reactive groups at the same positions, their properties as affinity reagents may vary substantially. Thus three dATP analogues containing 4-$N_3C_6H_4CH_2S$, 4-$N_3C_6H_4SS$, and 2-NO_2-4-$N_3C_6H_3S$ groups at position 2 of adenine have been described.[19] It was found that only the first of them could inactivate the Klenow fragment of DNA polymerase I upon irradiation (see Section III.1). It should be noted that inactivation does not require the presence of a template–primer duplex. Modification reaches saturation when one residue is incorporated into the enzyme. Although the modifying reagent is based on dATP, its competitive substrate is dTTP. Moreover, this modification is inhibited by the poly(dA)–oligo(dT) duplex, which may attest to the binding of the analogue in the template region of the enzyme.[18]

Among other nucleotide analogues having a reactive group in their carbohydrate fragment, 2′,3′-epoxyATP has been studied in more detail. In the presence of the duplex **1**, the enzyme was inactivated in accordance with the classical law of affinity-based modification;[4] its K_d was 21 μM and the inactivation rate constant was 0.021 s^{-1}. The 2′,3′-epoxyAMP residue was incorporated into the primer simultaneously with inactivation. Naturally, the residue incorporated played the role of a terminator. It was shown[25] that the duplex carrying this residue at its terminus did not modify the enzyme; its inhibitory effect is due to the enhanced stability of a triple complex with the enzyme.

Sometimes, DNA derivatives designed for modification of DNA-binding sites of an enzyme can be obtained by direct introduction of a reactive group into the DNA molecule or its fragment. Thus the decanucleotide $d(pT)_2pC(pT)_7$ modified at the N(3) atom of the cytosine residue with *cis*-aquahydroxodiammineplatinum, *i.e.*, containing the *cis*-{$(NH_3)_2Pt([N(3)\text{-}dC](OH)$} fragment, was used as an affinity reagent for the matrix site of DNA polymerase α from human placenta and the Klenow fragment of DNA polymerase I.[26–28] Incubation of this decanucleotide with the enzyme results in gradual inactivation of the latter. This inactivation is competitively inhibited by deoxyoligonucleotides of various lengths and structures.

Since both the template and the primer have sufficiently extended contacts with the enzyme, when used as the affinity reagents (DNA analogues), it is especially desirable to obtain a series of derivatives with a reactive group at different predetermined positions of DNA. Two main approaches to the synthesis of these derivatives have been developed. One of them is based on the introduction of a reactive group during the synthesis of the corresponding oligonucleotide. Naturally, this approach can be used only for derivatives with a reactive group in their heterocycles. One can also synthesise DNA or an oligonucleotide derivative containing a group where the reactive residue will be introduced in the final stage of the synthesis. As an example, we can mention the synthesis of a series of primer analogues with a photoreactive group at different positions relative to the 3′-end of the growing chain.[29] Thus a thymidine analogue containing a $(CH_2)_3NHCOCF_3$ residue at position 5 of uracil was introduced instead of thymidine into a specific position of the primer during the synthesis.[29] After removal of the trifluoroacetyl group by mild alkaline hydrolysis, the primer was treated with 5-azido-2-nitrobenzoic acid *N*-hydroxysuccinimide ester. Then the modified primer was hybridised with DNA from phage M13 used as a template and elongated by one unit with DNA polymerase from phage T4 using [α-^{32}P]ddATP as a substrate. The final procedure sharply diminishes the risk of cleavage of the 3′-terminal residue due to 3′ → 5′ exonuclease activity inherent in the enzyme and allows simultaneous introduction of the radioactive label into the primer. This step is essential for the affinity-based modification with this compound. Using this approach, primers with a photoreactive arylazide group at positions 3, 9, 13, 18, 22, 27, and 46 relative to the 3′-terminal fragment have been obtained.[29] When the corresponding template–primer duplex is used, the reactive residues can be introduced into the specific positions of the primer by the enzymic method using dNTP analogues bearing a reactive group as DNA polymerase substrates.[12, 14, 15, 23]

The second approach to the synthesis of DNA with a predetermined position of the label can be exemplified by the synthesis of a series of derivatives of oligonucleotide primers with different positions of the fluorescent label relative to the 3′-end of the primer. A set of such derivatives were obtained[30] from one oligonucleotide with a thymidine analogue U* at a definite position; this analogue contained the fluorescent 5-dimethylaminonaphthalene-1-sulfonyl group at position 5 instead of the methyl group. The starting duplex had the structure **2**.

5′ d(TCGCAGCCGU*C)
3′ d(AGCGTCGGCA GGTTCCCAAA)
2

The structure of the template allowed the elongation of the primer in the presence of DNA polymerase, by one unit (compound **3**) with dCTP, by three units (compound **4**) with dCTP + dATP, and by six units (compound **5**) with dCTP + dATP + dGTP. The fluorescent label was translocated at a certain distance from the enzyme active centre.

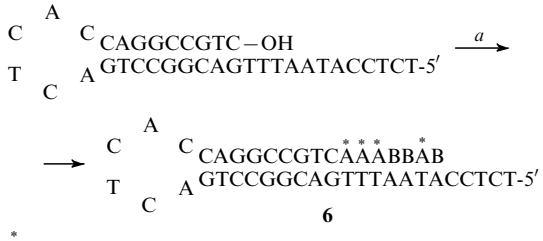
5′ d(TCGCAGCCGU*CC)
3′ d(AGCGTCGGCA GGTTCCCAAA)
3

5′ d(TCGCAGCCGU*CCAA)
3′ d(AGCGTCGGCA GGTTCCCAAA)
4

5′ d(TCGCAGCCGU*CCAAGGG)
3′ d(AGCGTCGGCA GGTTCCCAAA)
5

In some experiments, it was possible to use a hairpin structure with a hanging single-stranded fragment instead of a template–primer duplex. Such a structure mimics the template–primer duplex and the short fragment of the hairpin stem can be elongated in the presence of DNA polymerase. When photoreactive dNTP are used for elongation, a photoreactive stem mimicking the photoreactive duplex can be obtained.

This approach was used in particular for the synthesis of a hairpin containing several residues of bromodeoxyuridine (compound **6**).[31]



6

A* is the radiolabelled dAMP residue,
B is 5-bromodeoxyuridine;
(*a*) DNA polymerase, [α-^{32}P]dATP + BrdUTP (BrdUTP is 5-bromodeoxyuridine triphosphate).

This hairpin structure has been used for photoaffinity-based modification of DNA polymerase γ (mitochondrial DNA polymerase) from frog (*Xenopus*) and DNA polymerase ε from calf thymus.[32]

III. Study of the interaction of DNA polymerases with templates, primers, and nucleoside 5′-triphosphates using group-specific and affinity-based chemical modification. Structural aspects

1. DNA polymerase I from *E. coli* and the Klenow fragment

This enzyme is made up of one polypeptide chain (M_r is 103 kDa, 928 amino acid residues) and forms three domains. The N-terminal domain (1–330) comprises an active centre for $3' \rightarrow 5'$ exonuclease activity and catalyses stepwise hydrolysis of nucleotides from the 5′-end. The central domain (residues 331–514) contains an active centre for $5' \rightarrow 3'$ exonuclease activity and catalyses the cleavage of the 3′-terminal nucleotide from the growing polynucleotide chain. It is noteworthy that this cleavage is directed predominantly against the residue that is not complementary to the template and thus prevents preservation of an erroneously incorporated monomer. The 515–928 domain contains the DNA polymerase centre itself.[33] Mild treatment of the enzyme with a protease (subtilisin) results in the splitting of the $5' \rightarrow 3'$ exonuclease domain to form the so-called Klenow fragment, which retains only two activities and which is now the subject of the most active studies.[34]

X-Ray structural analyses of the Klenow fragment and of its complexes with a model DNA duplex and with dCTP have been carried out with a resolution of 3.9 Å (Fig. 1).[35,36] This figure clearly demonstrates the presence of both domains. The N-terminal $5' \rightarrow 3'$ exonuclease domain is formed by seven α-helices *A*–*G* and four β-structures *1*–*4*. The C-terminal polymerase domain of the Klenow fragment is formed by three subdomains, which have been termed the 'palm', the 'fingers', and the 'thumb' by anatomic analogy with the right hand (Fig. 1). The 'palm' domain consists of six β-structures *7*–*9* and *12*–*14*, the 'thumb' domain comprises four α-helices *H*–*K*, and the 'fingers' domain is made up of five α-helices *L*–*P*. The regions of the other two subdomains that are the most remote from the 'palm' come together to form a cleft where DNA binding takes place. The bottom of this cleft contains carboxy groups of two aspartic acid residues (Asp705, Asp882) and one glutamic acid residue (Glu883) (Fig. 2), which are supposed to coordinate the triphosphate group of the substrate through Mg^{2+} ions. Magnesium ions play an important role in the main catalytic process, being responsible for the transfer of the monophosphate residue to the 3′-OH group of the primer. Similar triads of dicarboxylic acid residues were also detected in other DNA polymerases the spatial structures of which have been studied. Special mention should be made of DNA polymerase β; its structure in the complex with a dCTP and dATP was established with high resolution both in the presence and in the absence of a template–primer complex.[37,38] The most detailed information about the contact sites of the template–primer complex with the enzyme was obtained by Modak *et al.*[39] These authors performed direct photo-cross-linking of two [5′-^{32}P]-labelled oligonucleotides, containing 25 and 37 monomeric units the 3′-termini of which formed a hairpin structure with the enzyme. The stem of the hairpin mimicked the template–primer complex. The main modification involved the peptide fragment corresponding to the DNA polymerase fragment 759–775 at Ile765, Tyr766, Ser769, and Phe771 residues in the *O* helix of the 'fingers' domain.

To obtain information about the binding site for the substrate, a series of investigations were carried out using reactive derivatives of dNTP. These studies began with direct photo-cross-linking of the enzyme with dTTP by irradiation of the enzyme–substrate complex with light at $\lambda = 254$ nm.[40] It was shown that photoaddition occurred at the His881 residue. As can be seen from

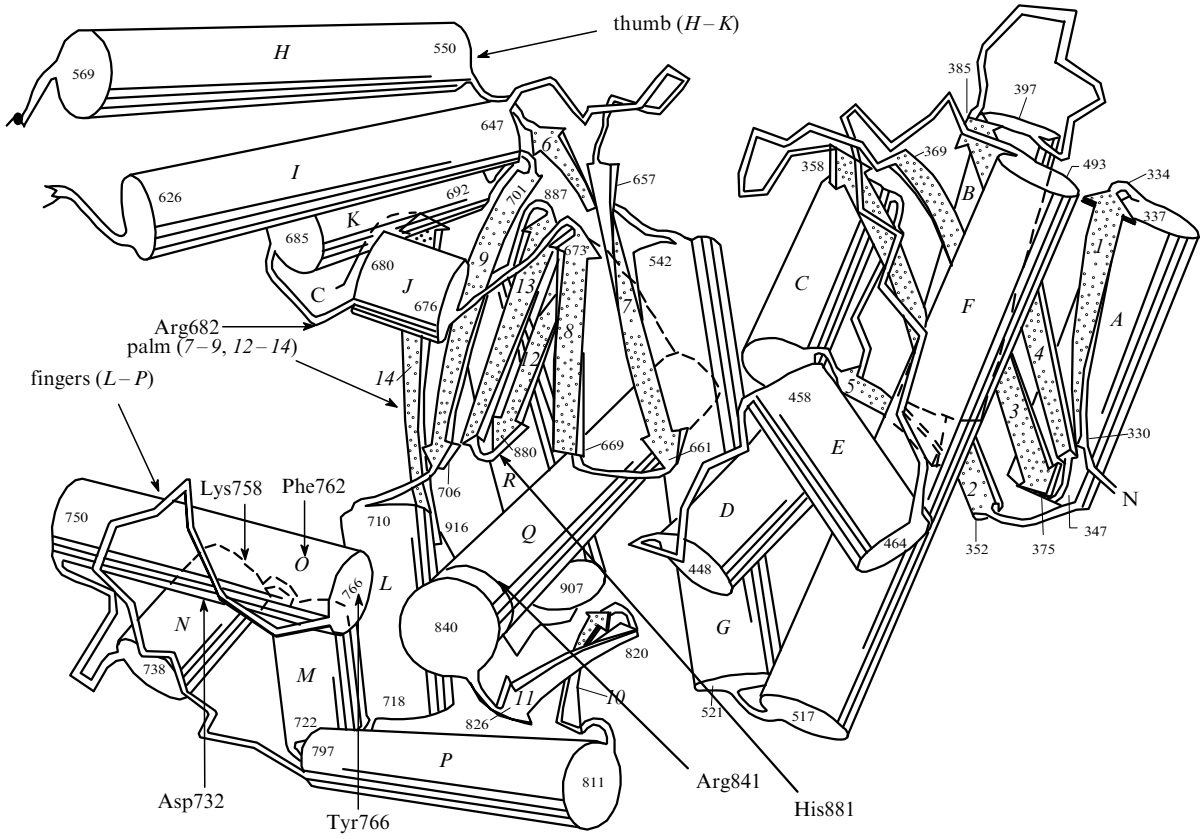


Figure 1. A scheme of the three-dimensional structure of the Klenow fragment of DNA polymerase I from *E. coli*. The α-helical regions of the polypeptide are shown as cylinders; the β-pleated sheet, as ribbons, with arrows indicating the direction of the chain in the folds from the N- to the C-end. The arrows indicate the amino acid residues detected by affinity-based modification.

Fig. 1, this residue is localised between the 12th and 13th β-structures of the 'palm' subdomain. When the Klenow fragment was modified with 8-azido-dAMP, photoaddition occurred predominantly at Tyr776 localised in the α-helix of the 'fingers' subdomain.[17] It is interesting to note that the same tyrosine residue was modified upon irradiation of the enzyme in the complex with the template and the primer analogue **7**.[41]

3′ d(AGCGTCGGCAGGTTCCCAAA)
5′ d(TCGCAGCCGU*)

7

U* is the residue of the dTMP analogue containing a 5-azido-2-nitrobenzoyl group linked *via* a spacer, $(CH_2)_3NH$, at position 5.

This coincidence shows that Tyr766 is localised directly in the catalytic centre. Evidence in favour of the presence of a tyrosine residue in the catalytic centre can also be derived from the results of chemical modification of the Klenow fragment by *N*-acetylimidazole, a specific reagent for the OH group of tyrosine. This modification was accompanied by a loss of the enzymic activity. The enzyme could be protected from inactivation only in the presence of a complete template – primer – dNTP complex. The latter must be complementary to the template used.[4, 42]

Even this example shows that chemical modification data do not fit into a simplified pattern of statistical distribution of substrate binding sites. Comparison of these data with the spatial structure of DNA polymerase obtained from X-ray analysis shows that the distance between the Tyr766 residue modified by 8-azido-dAMP and the His881 residue modified by dTTP is rather large (9 Å). This suggests that the orientation of the latter is different in different nucleotides, at least in the binary enzyme – dNTP complex.

In addition to dNTP analogues, other reagents, such as pyridoxal 5-phosphate (which often behaves as a nucleotide analogue) and 5′-*p*-fluorosulfonylbenzoyl adenosine (the size of the FSO_2Bz-fragment of which is close to that of the triphosphate group), were used for modification of DNA polymerase (the Klenow fragment). In the former case, the modification affected Lys758,[43] in the latter case, Arg682.[44] It should be noted that the latter modification was performed in the presence of a template – primer complex modelling the complex of the Klenow fragment with the template – primer and dNTP.

If one assumes that the residues thus revealed are directly related to the active centre of the enzyme, it would be reasonable to suppose that the positively charged residues of both Lys758 and Arg682 will interact with the triphosphate fragment of dNTP. However, this interpretation, too, is inconsistent with the enzyme structure obtained from X-ray analysis. Indeed, Lys758 is localised in the *O* helix pertaining to the 'fingers' domain, while Arg682 is located in the 'thumb' domain, *i.e.*, in the opposite side of the cleft (Fig. 1). At the same time, the triad of amino dicarboxylic acid residues localised in the 'palm' domain is the most probable participant in the binding of the triphosphate fragment. Moreover, these assumptions are at variance with the results obtained by photo-modification of the Klenow fragment with 2-(4-azidophenacylthio)-dATP.[18] It was found[18] that the modification affected the Asp732 residue localised in the helix *N*.

When conducting a simultaneous analysis of the X-ray and chemical or affinity-based modification data, one should bear in mind that these results may not be adequate from the theoretical standpoint. X-Ray analysis is carried out on crystalline preparations and gives an optimised structure. It is obvious, however, that the conformation of the protein backbone may fluctuate to some extent, and the crystal conformation deduced from X-ray analysis will not necessarily be optimum for the binding of dNTP or its triphosphate fragment. In an even greater degree this applies to the results of chemical modification. It was found[45] that the thermodynamically optimum structure of an affinity reagent – enzyme complex will hardly provide the optimum mutual orientation of the reactive fragment and the amino acid residue of the target. The use of the molecular mechanics method for the analysis of the process occurring in the duplex formed by complementary oligonucleotides one of which carries the *N*-2-chlorothylamino group showed[46] that the energetically optimum conformations do not provide sufficient approximation of the heterocycle to be alkylated and the intermediate ethyleneimmonium cation that formed from the original alkylating group. To provide the transition from the thermodynamically optimum to the kinetically optimum structure, it is necessary to overcome a certain energetic barrier. The term 'store-conformationt' has been introduced to describe energetically minimised structures able to pass (easily but not freely) to a kinetically optimum conformation.[46] Therefore, a small inconsistency between the X-ray and chemical data, which does not exceed several Angström, can only reflect the conformational differences between the structure of the enzyme with the optimum arrangement of the substrate and of the modifying agent.

It was shown[47] that the treatment of the Klenow fragment with [7-^{14}C]phenylglyoxal, which specifically reacts with the arginine residues, inactivates the enzyme. Incorporation of four residues of the reagent into the protein gives a 95% inactivation. Since each arginine residue can react with two phenylglyoxal molecules, this implies that the inactivation occurs upon modification of two arginine residues. Only the template – primer complex can protect the enzyme from inactivation by phenylglyoxal. The main site of modification is Arg841. This residue is localised in the *Q* helix near Tyr766, Lys758, and His881, which, as mentioned previously, take part in the interaction with dNTP or their analogues. The position of amino acid residues within the three-dimensional structure of the Klenow fragment of DNA polymerase I from *E. coli* revealed by affinity-based modification is shown in Fig. 1. The triad of dicarboxylic amino acid residues is marked by black circles in Fig. 2.

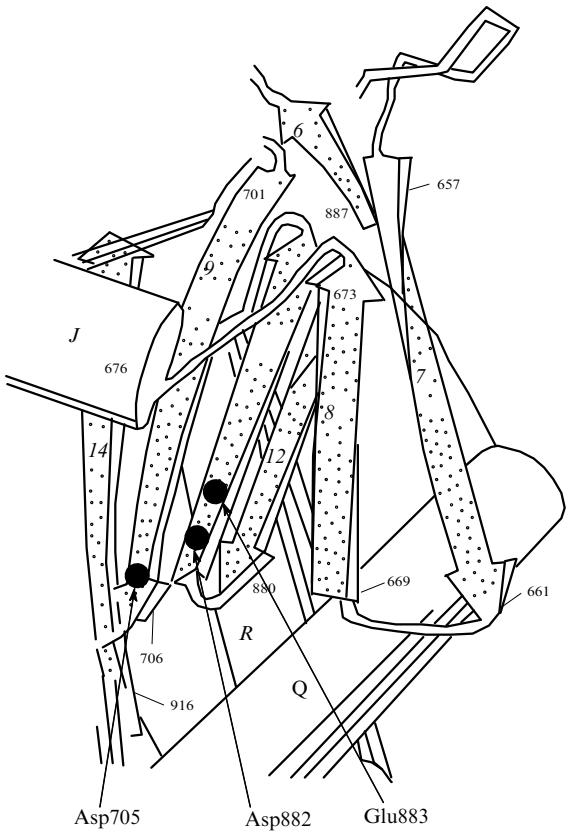
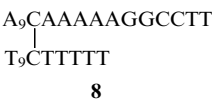


Figure 2. The position of the triad of dicarboxylic amino acids in a three-dimensional structure of the Klenow fragment of DNA polymerase I from *E. coli*.

An unexpected result was obtained in the study of the structure of the Klenow fragment of DNA polymerase I. It was found that the active centre where the elongation of the primer takes place and the site of the $3' \rightarrow 5'$ exonuclease activity are at a large (about 30 Å) distance from one another. According to the generally accepted and well-rationalised experimentally viewpoint concerning the centre of the $3' \rightarrow 5'$ exonuclease activity, the newly attached 3′-terminal fragment must undergo 'testing' in the exonuclease centre after each elongation act. This requires the translocation of the terminal fragment of the primer by several units. This problem was studied using a template – primer duplex **8** with covalently cross-linked chains as a model.[48]



8

Here, the amino groups of the two cytosine residues are linked through an ethylene bridge. Such a cross-linking is achieved by introducing a cytidine analogue (instead of cytidine) containing a highly reactive three-membered ethyleneimine ring at position 4 into a matrix oligonucleotide. This analogue easily alkylates an opposite amino group of the primer chain. It was found that only the 5th (starting from the cross-linking site) dTMP residue of the primer can be split off by exonuclease. Splitting of the 4th residue is strongly hindered (the rate constant for this reaction is two orders of magnitude lower than that for the 5th residue). At the same time, this shortened cross-linked primer, the 3′-OH group of which is at a distance of four nucleotides from the cross-linking site, is still a good substrate for elongation. Hence, the difficulty in separating the strands affects only one of the enzyme activities. This result can be regarded as an experimental proof that the 3′-terminal fragment of the strand that is synthesised should be translocated at a large distance for correction.[48]

Some data about the contact sites of the enzyme and nucleic acid components were obtained with the use of fluorescent methods. A series of duplexes **2 – 5** were used to study the contact sites of a template – primer duplex with the enzyme.[30] These duplexes contained a dTMP analogue with the fluorescent 5-dimethylaminonaphthalene-1-sulfonyl group (instead of the methyl group) at position 5 introduced into a predetermined position of the primer component. Thus the fluorescent label was set aside at a certain distance (which was different for different duplexes) from the enzyme active centre. The fluorescence lifetime and the inner rotation correlation time measured for all the samples showed that at least six primer residues interact with the enzyme. The label becomes more remote from the 3′-end of the primer, some changes in its fluorescent properties were observed, which reflect the changes in both the label accessibility for the solvent and its closeness with the protein.

Introduction of fluoresecent labels into the protein, DNA, and dNTP was used to calculate the distances between the specific site in the enzyme, DNA, and dNTP using the Förster theory based on the resonance energy transfer. 4-Amino-7-nitrobenzo-2,1,3-oxadiazole can be introduced into the Klenow fragment as an energy acceptor at Cys907 through the $CH_2C(O)OCH_2CH_2$ spacer. A similar acceptor can be introduced into position 5 of uracil through the trimethylene spacer, and a DNA fragment can be synthesised on the basis of the thymidine fragment thus obtained. The fluorescent 5-aminonaphthalene-1-sulfonate group can be introduced as an energy donor into position 5 of uracil through the $SCH_2C(O)NHCH_2CH_2$ spacer. This can serve as a basis for obtaining a mononucleotide, a nucleoside triphosphate, and a DNA fragment carrying a donor group. This study made it possible to determine the distances between dNMP, dNTP, DNA, and Cys907 as well as the distance between DNA and dNMP.[49]

Studies of the active centres of other DNA polymerases by affinity-based modification are still in their infancy. Human DNA polymerase β for which high resolution X-ray data are now available seems to be a tempting object of investigation.[37,38] An attempt was made to localise the dNTP-binding site by modification with 8-azido-ATP and 5-azido-dUTP.[50] However, these experiments were carried out in the absence of the template – primer complex. This may account for the modification of the 8 kDa subdomain involved in the interaction with single-stranded DNA.[50] Recently, catalytically competent labelling of DNA polymerase β was performed using FAB-dCTP.[12] Given that the site(s) of modification are established, they can be more reliably associated with the dNTP-binding site.

2. Reverse transcriptases

Enzymes responsible for reverse transcription, *i.e.*, catalysing template biosynthesis of DNA on RNA templates, are characteristic of a large group of viruses called retroviruses. Some investigations based on the use of chemical methods were carried out with reverse transcriptases of the avian myeloblastosis virus (AMV) and the Moloney murine leukemia virus (MuLV). Identification of reverse transcriptase in the human immunodeficiency virus (HIV-1) has given a strong impetus to intensive studies of these enzymes. Reverse transcriptase is regarded as a promising target in the fight against AIDS; therefore this section will largely deal with the HIV-1 enzyme.

Normally, reverse transcriptase is made up of two subunits designated according to their molecular weights (in kDa) as p66 (555 amino acid residues) and p51 (446 amino acid residues). The smaller subunit (p51) represents a shortened subunit of p66 devoid of the fragment 447 – 555.[51 – 53] This enzyme catalyses several reactions. Its functioning begins with reverse transcription of genomic viral RNA expressed in the virion by two molecules, 9750 nucleotides each.[54] Like all other DNA polymerases, this enzyme requires the presence of a primer. Host-cell tRNA is a natural primer and human $tRNA^{Lys,3}$ is a primer for the HIV-1 enzyme. In a definite (final) stage of reverse transcription, which results in the formation of a minus-replica of DNA, the enzyme begins to function as an ordinary DNA polymerase by forming a plus-strand of DNA from the information contained in the minus-strand. The formation of the minus-strand is accompanied by gradual destruction of the genomic viral RNA under the action of another activity inherent in the enzyme, namely, that of RNAse H. It is this enzymic activity that is present in the 447 – 555 domain, which is absent from the p51 subunit.[53,55] In principle, homodimers made up of identical subunits can be obtained under special conditions instead of the p66/p51 heterodimer.

X-Ray structural analysis of the heterodimer revealed [56] that the p66 subunit consisted of several subdomains. They form a structure similar to that of the Klenow fragment of DNA polymerase I represented by the 'fingers', 'palm', and 'thumb' subdomains. The residues corresponding to these subdomains, which are linked to the C-terminal subdomain responsible for the RNAse H activity through a linker subdomain, were identified in the primary sequence of reverse transcriptase.[56] The 'palm' subdomain contains three Asp residues (110, 185, 186) that seem to play the same role as the corresponding triads in DNA polymerases I and β.

The first three domains form a polymerase domain which, like the Klenow fragment, has a cleft for the binding of the template – primer duplex. However, unlike the Klenow fragment, the polymerase domain is formed by the N-terminal part of the polypeptide chain but not by its C-terminal part.[57]

The contact site of the duplex with the enzyme was studied in most detail using footprinting with the $^{\bullet}OH$ radicals.[58] Several probes were prepared on an artificial template – primer duplex (its central part is shown below) by elongation of the primer with the corresponding dNTPs and one ddNTP.

....d(AAACCCTTCTGGTTCCCCTTTCCCC....) (3′) (78 units)
d(GGGG....) (5′) (35 units)
ddAGGGG....
ddGAAAGGGG....
ddCAAGGGGAAAGGGG....
ddTGGGAAGACCAAGGGGAAAGGGG....

Using this approach, a series of duplexes with the terminating fragment at a distance of 1, 4, 10, and 19 units from the original fragment were obtained. For example, the 10-nucleotide-shift was achieved by adding dATP, dGTP, and ddCTP to the duplex–enzyme mixture. Since the incorporation of the terminator occurs in the active centre of the enzyme, complexes with the duplex at the same position relative to the active centre but with different nucleotide environment were formed in all the cases studied. In all these cases the enzyme protected the sites from +3 to −15 on the template strand and the primer residues from −1 to −15 (the numbering is given relative to the terminator incorporation site). At the same time, the sites from −9 to −11 of the template and the sites from −8 to −10 of the primer are not protected. In addition, a pair with a strongly enhanced sensitivity to modification neighbours (with respect to the 5′-end of the primer) the sequence protected from the $^{\bullet}OH$ radicals (Fig. 3). The dark area designates the regions in the primer and the template between positions +3 and −15 protected by the enzyme from free radicals. The hypersensitive spot is the site of enhanced cleavage of the template chain induced by changes in the template structure owing to the interaction of the HIV-1 reverse transcriptase heterodimer with the RNAse H domain. The region of the duplex where this conformational change takes place is near the contact site for RNAse H. It is therefore probable that it is the RNAse H domain that induces a conformational transition in this region of the duplex. Similar, although not identical results were obtained in enzymic footprinting studies. Irrespective of the site where the synthesis is stopped by the terminator, the template residues from +7 to −23 and the primer residues from −1 to −25 are protected from DNAse I. This site becomes accessible for the DNAse attack at the junction of the regions forming the polymerase and RNAse H centres (nucleotides −20 of both the template and the primer).[59]

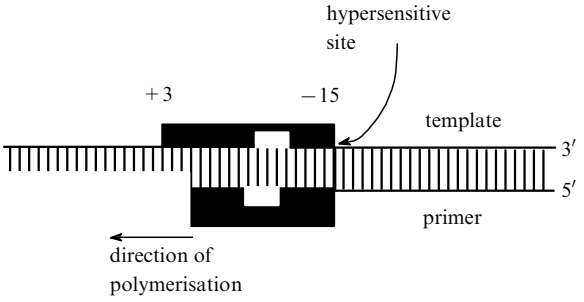


Figure 3. A scheme of footprinting with the hydroxyl radicals of the template and the primer chains.

One of the most intriguing properties of reverse transcriptases is the use of definite host-cell tRNAs as primers for reverse transcription. The 3′-terminal sequences of these tRNAs are always complementary to the site of the primer binding with genomic RNA. However, some other domains of this tRNA are also involved in the binding with both genomic RNA and the enzyme itself. Studies with *trans*-Pt[$(NH_3)_2Cl_2$] as a bifunctional cross-linking reagent[60] revealed that the anticodon loop of $tRNA^{Lys,3}$ appears to be linked to reverse transcriptase of HIV-1. Using the same reagent, it was shown that cross-linking of the anticodon loop with reverse transcriptase of HIV-1 also takes place when tRNA is synthesised by transcription of the encoding gene by RNA polymerase from phage T7, which does not contain modified heterocycles, *viz*., 5-methoxycarbonylmethyl-2-thiouracil and *N*-(2-methylthiopurin-6-yl)threonine. This implies that this interaction is not due to the presence of specifically modified heterocycles in the anticodon loop.[61]

In order to identify $tRNA^{Lys}$ fragments that make contact with the enzyme, tRNA analogues with specifically incorporated 4-thiouridine residues were used.[62] In most cases, irradiation of the enzyme–tRNA complex with light at $\lambda = 365$ nm resulted in the cross-linking of tRNA with the protein. It was found that in the case of tRNAs containing a photoactive residue at positions −1 and +16, *i.e*., within the acceptor stem and the dihydrouridil loop, the cross-linking occurred exclusively with the p66 subunit. If the thiouridine residue was at position 36 (the anticodon loop), the cross-linking occurred with both subunits, although the reaction with p66 was predominant. When the thiouridine residue was at position 43 (the anticodon stem), no photo-cross-linking of tRNA with the enzyme took place. These results give an idea of which regions in the tRNA primer are in a direct contact with the enzyme.

Chemical modification with reagents specifically interacting with single-stranded sites of nucleic acids, *viz*., dimethyl sulfate, *N*-cyclohexyl-*N*′-[2-(*N*″-methylmorpholinium)ethyl]-carbodiimide, diethyl pyrocarbonate,[63] made it possible to carry out a detailed analysis of the interaction of the primer binding site (PBS) of genomic RNA of HIV-1 and $tRNA^{Lys}$. The absence of or a drastic decrease in the level of modification of some particular sites of genomic RNA suggest their involvement in the formation of complementary nucleotide pairs. It was found that 18 nucleotide residues at the 3′-end of tRNA form a double helix with the 179–196 site of genomic RNA. The results obtained with the use of simpler unnatural systems, *e.g*., poly(rA)–oligo(dT) duplexes, indicate that reverse transcription does not require such a complicated and specific primer. Nevertheless, the 3′-end of this primer not only corresponds to a definite site in viral genomic RNA, *i.e*., fulfils the function of a primer, but also forms a complex with other fragments of tRNA and the template. The hexanucleotide fragment 33–39 of this primer, which forms an anticodon loop, is linked to the complementary 162–167 site of the template with a fall-out of one minor nucleotide from the two-helical site of the loop. The 40–47 site forms duplexes with the 140–143 and 158–160 sites of viral RNA.[63]

Thus, the use of chemical methods for the study of reverse transcriptase of HIV-1 has made it possible to identify the regions of the template and the product located on the enzyme in the vicinity of its catalytic centre as well as the sites of the natural primer interacting with the enzyme and genomic RNA. The regions interacting with the primer were also established for genomic RNA.

The information obtained from the analysis of functionally important protein regions by chemical methods was more scarce and sometimes controversial. Of undoubted interest are the results obtained with affinity reagents and related to the distribution of modification sites between the subunits. Direct photo-cross-linking of dTTP with the enzyme occurs only on the p66 subunit.[64]

The same results were obtained by photoaffinity labelling of the enzyme in the template (18-, 21-meric oligomers)–primer (12 units) complexes elongated with 4-thiouridine triphosphate[65] and with poly(dA)–oligo(dT) used as the primer–template and 2′,3′-dideoxy-(5*E*)-[4-(3-trifluoromethyl-3*H*-diazirin-3-yl)styryl]UTP as a dTTP analogue.[20] Direct photo-cross-linking was performed[66] for the duplex formed by a 40-meric 5′-terminal fragment of genomic RNA and a 20-meric DNA primer. In this case, the primer was added exclusively to the p66 subunit. At the same time, after elongation of the primer to 40 nucleotides, the cross-linking occurred also with the p51 subunit. Presumably, this indicates the presence of a common pocket, which binds the primer and which is formed by both subunits of the enzyme.[66] In any case, it is obvious that the p51 subunit begins to interact with the primer at a sufficiently large distance from the enzyme active centre.

Similar results were obtained by Lavrik *et al.*[67] In these studies, the enzyme was irradiated with light at $\lambda = 266$ nm in the presence of [^{32}P]d(pT)$_{10}$ with or without a complementary template [poly(rA)] added; radioactivity was found only in the p66 subunit. At the same time, when the d(pT)$_{10}$ derivative containing a photoreactive residue of 2-nitro-5-azidobenzamide or β-alanylamidoethidium linked by a phosphamide bond with the 5′-terminal phosphate was used, the photoinduced binding occurred with the both subunits. This can be interpreted as a result of remoteness of the photoreactive group from the enzyme active centre where the opposite 3′-terminal residue of the nucleotide has to be localised. At the same time, introduction of a dCTP analogue residue with a photoreactive azidoperfluorobenzoyl group linked with C(5) of cytosine through the spacer CH=CH−CH$_2$−NH (AlFAB-dCTP), or with the amino group of cytosine through the spacer CH$_2$−CH$_2$−NH(FAB-dCTP) to the 3′-end of the primer also results in cross-linking of the primer with both subunits upon irradiation.[14]

This change in the distribution pattern upon introduction of a relatively small fragment separating the photoreactive group and the primer shows that the p51 subunit is located near the catalytic site of p66. X-Ray data for an HIV-1 reverse transcriptase – DNA complex together with computer simulation suggest the involvement of the p51 subunit in the interaction between revertase and the template – primer complex.[56, 57, 68]

Studies of reverse transciptase of the avian myeloblastosis virus, which is also made up of two subunits, show their non-equivalence. This enzyme is used as a natural primer for tRNA^{Trp}. It was found[69] that tRNA^{Trp} oxidised with periodate at the 3′-terminal *cis*-diol group inactivates the enzyme. When the modification is not brought to completion, all the tRNA^{Trp} appears to be bound with only one of the two subunits designated as α. The latter is probably a p66 analogue of the HIV-1 enzyme because it contains all the activities inherent in this enzyme including that of RNAse H.

Some studies provide evidence about the sites or regions of reverse transcriptase of HIV-1 that either interact with affinity reagents or are essential for manifestation of catalytic functions. In the above experiment with direct photo-cross-linking of the enzyme with dTPP, the main modification point was determined to be a Lys73 residue localised in the 'fingers' subdomain.[61] In contrast, in the template (18 – 21 units) – primer (12 units) complex elongated with 4-thiouridine triphosphate, photoaffinity-based labelling was restricted to the peptide fragment 313 – 423.[62] This fragment corresponded to the lower part of the 'thumb' subdomain and the linker domain. It was found also that modification with phenylglyoxal, which is specific for arginine, results in complete inactivation of the enzyme.[70] The presence of a template – primer complex prevents the inactivation; the Arg277 residue appears to be protected from modification. This residue is also present in the 'thumb' subdomain, but at its opposite end (relative to the peptide), and is modified with 4-thiouridine.

To establish the binding site for the primer, (pdT)$_{16}$ was directly cross-linked with a homodimer made up of two p66 subunits.[71] Analysis of tryptic hydrolysates revealed that a peptide 288 – 307 localised in the upper part of the 'thumb' subdomain is the main binding site for oligo(dT). Fragments Leu289 – Thr290 and Leu295 – Thr296 were identified as the most probable sites of modification. This region also comprises a polynucleotide-binding domain for the homodimer of reverse transcriptase of HIV-1.[72]

It should be noted that some photoactivated groups in the primer modify both the protein and the template DNA, which results in the cross-linking of the primer and template chains. This phenomenon was observed in enzyme – template – primer systems of reverse transcriptase of HIV-1 and of a DNA polymerase a complex with primase when primers with a FAB-dCTP residue were used.[13, 14, 21]

IV. Catalytically competent labelling of DNA polymerases

The method of catalytically competent labelling of template enzymes was first actively applied to the study of DNA-dependent RNA polymerases. The main principle of this method consists in chemical modification of the enzyme by some unlabelled reactive analogue followed by elongation of a covalently attached substrate with a second substrate bearing a radioactive label. In this case only those residues of the first substrate analogue will contain the radioactive label the orientation of which allows them to take part in subsequent catalysis due to covalent binding. These residues are usually referred to as 'catalytically competent'.[1, 73]

For the first time, this principle was applied to DNA polymerase I.[74] Oligothymidylate analogues with an alkylating group linked to the 5′-terminal phosphate were used as affinity reagents (Fig. 4). Affinity-based modification of DNA polymerase I was carried out by covalent binding of a reactive thymidylate analogue in the presence of a poly(dA) template. After covalent binding, a radioactive label was introduced by elongation of oligothymidylate with [α-^{32}P]dTTP. This approach made use of the catalytic competence of the oligonucleotide primer covalently attached at its 5′-end. The most intense labelling was observed in the peptide Asn832 – His928 located in the polymerase region of the enzyme. As the 3′-OH group of the primer must be in the active centre of the enzyme during catalytically competent labelling, more detailed localisation of the binding site at the 5′-end of the primer might allow establishing of the orientation of the primer relative to the enzyme. The orientation of the primer complex with DNA polymerases is currently a matter of vivid discussions.[38, 75]

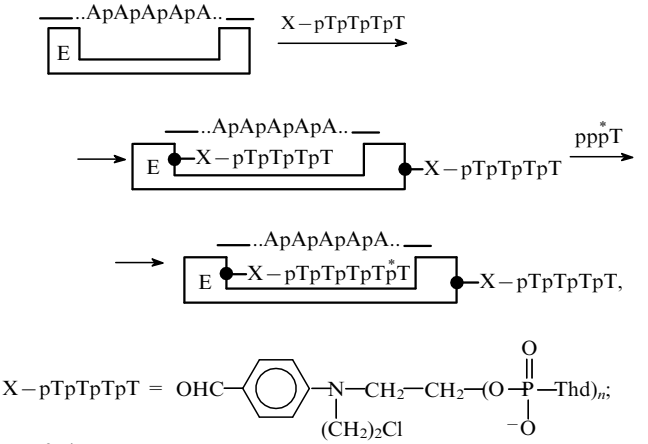


Figure 4. A scheme of highly efficient affinity-based modification of DNA polymerase I from *E. coli*: $\overset{*}{p}$ is the radiolabelled phosphorus atom; X – pTpTpTpT is the affinity reagent; X is the affinity reagent residue covalently bound with the enzyme; E is the enzyme.

An analogous approach was used[67] in a study of the contribution of revertase subunits to the interaction with the template – primer complex. It is interesting to note that covalent binding of the primer at its 5′-end did not prevent subsequent attachment of at least four thymidine residues to the growing chain. Therefore, one should not rule out the possibility of translocation of the enzyme relative to the template – primer duplex. This system appears to be an interesting model for translocation studies. The same result was obtained for primase from *E. coli*[76] where the covalent attachment of adenosine trimetaphosphate followed by elongation with [α-^{32}P]GTP was used for competent labelling. Elongation with [α-^{32}P]GTP, together with other dNTP, yielded tetra- and pentanucleotides.

In this case, too, the covalent binding did not prevent several elongation acts to be accomplished.

A new version of catalytically competent labelling was carried out using photoreactive analogues of dNTP with arylazido groups linked with a heterocyclic base of the nucleotide through spacers. In contrast with the above-described versions of catalytically competent labelling where immobilisation was carried out at the 5′-end of a monomer or an oligonucleotide, the triphosphate group of dNTP does not take part in substrate immobilisation on the enzyme and thus remains unmodified. Hence, it can participate in the reaction with the 3′-OH group of the primer added.

This possibility was studied using a series of photoreactive analogues the substrate activity of which had been tested beforehand (see Table 1, Refs 11, 12, 14–16, 23). It was found that in all the cases studied the dNTP analogues covalently linked to DNA polymerase through a heterocyclic base could interact with the template–primer complex. In the presence of [^{32}P] phosphate at the 5′-end of the primer, this resulted in competent labelling of the binding site of the dNTP analogue. In other studies, a non-labelled template–primer complex was added to a DNA polymerase conjugate with a dNTP analogue followed by addition of [α-^{32}P]dNTP corresponding to the template. In this case labelling of the conjugate also took place. Thus, the template–primer complex immobilised through a heterocyclic base can participate in the elongation of the primer.[12, 14, 21, 23]

According to some data, binding of such reagents can increase the selectivity of modification. Modification of HIV-1 reverse transcriptase with a template–primer complex carrying a FAB-dCTP or AlFAB-dCTP residue at the 3′-end of the primer is described in Section III.[14] In these studies, the photoreactive group was introduced into the primer enzymically, *i.e.*, with the help of a free substrate. This resulted in modification of both subunits. In analogous experiments, the same photoreactive groups were introduced into the primer with the help of preimmobilised analogues of dNTP. In this case, modification was more selective and affected only the p66 subunit. In the case of NAB-dUTP, the use of this approach did not increase the selectivity. The label was detected in the subunits both after preliminary immobilisation and without it.[15] This provides additional evidence for the close localisation of the p51 subunit and the active centres of the p66 subunit.

That the triphosphate group of an immobilised dNTP retains substrate activity may be of considerable interest in the study of an enzyme centre involved in the catalysis of the nucleotidyltransferase reaction. It is therefore important to preserve at a maximum the ability of the immobilised dNTP to take part in the elongation of the primer. One of the approaches has been checked for the example of FAB-dCTP and reverse transcriptase of HIV-1.[12] To this end, photoaddition of the reagent was carried out in the presence of a template–primer complex with the corresponding dideoxynucleotide at the 3′-end of the primer component. A fragment of this system is shown below:

(5′) dA-dG- -dA-dA-dT-dG.....(3′) (template)

pppdC-FAB ddT-dT- dA-dC.....(5′) (primer).

The duplex used could not undergo elongation due to the absence of the OH-group at the 3′-end of the primer. In other features, it fully mimicked the functional duplex. After irradiation, the dideoxy analogue of the primer was replaced with an excess of an ordinary ^{32}P-labelled primer. The level of catalytically competent binding of the enzyme in the presence of a template and a primer was determined by the amount of the label bound with the enzyme. It was found that the level of labelling of the p66 subunit increased substantially (severalfold) in comparison with that observed upon photo-cross-linking of the same analogue in the absence of the other components. It is interesting to note that analogous experiments with human DNA polymerase β did not reveal any increase in the level of FAB-dCTP binding with the enzyme. This may indicate a lesser dependence of the interaction of dNTP with the active centre of this enzyme on the template.[12] The primer (both covalently attached to the 3′-end by catalytically competent elongation and containing a photoreactive group in the 3′-terminal unit after covalent attachment) can be further elongated with an [α-^{32}P]dNTP.[12, 14] This approach can also be used for the analysis of the translocation mechanism of DNA polymerases.

V. Study of interactions of DNA polymerases with templates, primers and nucleoside 5′-triphosphates by affinity-based chemical modification. Quantitative aspects

As mentioned above, affinity-based modification of biopolymers, *i.e.*, modification by reagents with a specific structure that is recognised by a biopolymer and carrying a reactive group, is widely used for selective modification of those groups of a biopolymer that are localised at the site of its interaction with the ligand. In other words, affinity-based modification is used as a tool for the study of the structural organisation of active centres of enzymes. The kinetics of affinity-based modification can be studied by measuring either the degree of the reagent binding with the biopolymer or a decrease in the biopolymer activity in the course of the reaction. Kinetic data provide information about the reagent affinity for the biopolymer. The measurements performed in the presence of a ligand competing with the reagent for the binding to the enzyme active centre allow one to determine the affinity for the competing ligand. The methods used to estimate the dissociation rate constants (K_d) for the ligands competing with affinity reagents have been generalised in the reviews.[4, 77] In the simplest form, the affinity-based modification of an enzyme (E) by a reagent (X) in the presence of a competing ligand (Y) can be represented as follows:

$$E + X \overset{K_X}{\rightleftharpoons} EX,$$

$$EX \xrightarrow{k} EZ,$$

$$E + Y \overset{K_Y}{\rightleftharpoons} EY,$$

where EX is the enzyme–reagent complex; EZ is the modification product; EY is the enzyme–Y complex; k is the rate constant for the conversion of EX into EZ; K_X is the rate constant for the dissociation of the EX complex, and K_Y is the rate constant for the dissociation of the EY complex.

At a sufficiently large excess of the affinity reagent X, modification occurs as a first-order reaction. The values of apparent rate constants for the modification (k_{app}) at various concentrations of the reagent and the competing ligand can be determined from the dependence of the degree of modification or the residual activity of DNA polymerase on the time of its incubation with a reactive analogue of an oligonucleotide. The dependence of k_{app} on the initial concentration of the reagent X_0 and the concentration of the competing ligand Y_0 can be written as equation (1):

$$\frac{1}{k_{app}} = \frac{1}{k} + \frac{K_X}{[kX_0(1 + Y_0/K_Y)]}, \quad (1)$$

where k_{app} is the apparent rate constant for the modification, k is the rate constant for the conversion of EX into EZ, K_X is the rate constant for the dissociation of the EX complex, and K_Y is the rate constant for the dissociation of the EY complex.

Using this anamorphosis, it is not difficult to obtain the constant K_Y from the dependence of k_{app} on Y_0 at constant X_0.

This method was used to study the interaction of DNA polymerases with templates, primers, and deoxynucleoside 5′-triphosphates. The affinity reagent $(pdT)_2p\{(NH_3)_2Pt[N(3)$-

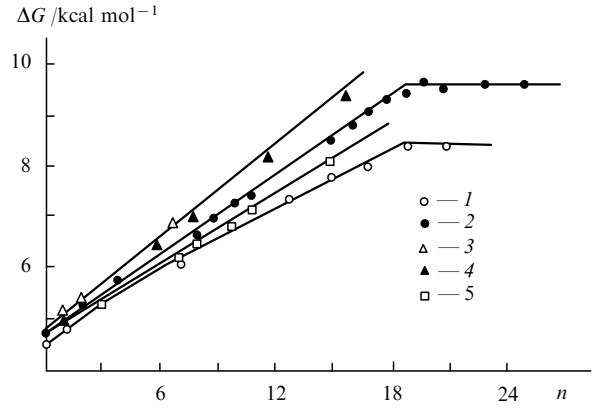


Figure 5. Dependence of ΔG^0 for the complex of the Klenow fragment with oligocytidylate (*1*), oligothymidylate (*2*), oligoguanylate (*3*), and oligoadenylate (*4*), and for the complex of DNA polymerase α from human placenta with oligothymidylate (*5*) on the length of the oligonucleotide.

dC](OH)}$(pdT)_7$ was used to study the binding site of the template. Its reaction with the enzyme was followed by monitoring the changes in the activity of DNA polymerase.[74] A broad array of oligonucleotides of various lengths and composition were used as competitive inhibitors. The values of the dissociation constants for the DNA polymerase α complex and the Klenow fragment with oligonucleotides and the corresponding standard free energies of binding (ΔG^0) were established.[28, 78, 79] The dependence of ΔG^0 on the length n was determined for a series of homooligonucleotides (Fig. 5). This dependence is linear up to about $n = 20$. This suggests that the length of the region in which the template binds with the enzyme is about 20 nucleotide residues. For decathymidylates this dependence is described by equation (2).

$$\Delta G^0 = -6.0 - 0.32n \text{ (kcal mol}^{-1}\text{)}. \quad (2)$$

This means that the contribution of each nucleotide unit is 0.32 kcal mol^{-1}. For other types of nucleotide units these values are different. Thus for the most hydrophilic pdC, this value is -0.28 kcal mol^{-1}, whereas for more hydrophobic pdA and pdG it is -0.40 kcal mol^{-1} (see Ref. 74). The linear dependence (2) makes it possible to calculate the value of ΔG^0 for the interaction of DNA polymerase with an oligonucleotide with a random sequence of the monomers (Table 2). The calculated values are in good agreement with those obtained experimentally.

Table 2. K_d values for the complexes of the Klenow fragment of DNA polymerase I from *E. coli* with oligonucleotides of various lengths and structures.[28]

Oligonucleotide	K_d /μM	
	calculated	experimental
d(pTpApApTpApCpGpA)	0.140	0.170
d(pCpTpGpCpGpTpTpApC)	0.120	0.130
d(pApApTpCpApTpGpGpTpCpApT)	0.013	0.015
$d(pCpA)_5$	0.070	0.060
$d(pGpA)_5$	0.024	0.015
$d(pGpA)_3$	0.400	0.450
$d(pGpT)_6$	0.011	0.009

The correlation between the nature of the oligonucleotide and its role in the interaction with the enzyme suggests that this interaction is hydrophobic rather than electrostatic. It is interesting to note that according to direct photo-cross-linking data, the 759–774 region of the Klenow fragment, which interacts directly with the DNA duplex, does not contain charged amino acid residues,[39] but has eight highly hydrophobic residues (Ile760, Phe762, Leu764, Ile765, Tyr766, Met768, Phe771, and Leu773).

This finding was confirmed in experiments with oligothymidylates esterified (ethylated) at internucleotide phosphates, *i.e.*, devoid of the negative charge.[26] The dissociation constants (K_d) for the complex of human placenta DNA polymerase α with the octanucleotide $d(Tp)_7T$ and the octanucleotide $d[Tp(C_2H_5)Tp]_3Tp(C_2H_5)T$ ethylated at four internucleotide phosphates are 0.4 and 0.33–0.39 μM, respectively (depending on the configuration of the chiral phosphorus atoms). In other words, elimination of four charges in the oligonucleotide molecule does not affect its affinity for the enzyme. At the same time, on complete esterification of the octanucleotide $d[Tp(C_2H_5)]_7T$ the affinity drops by one order of magnitude ($K_d = 2.8$ μM).[26]

It is noteworthy that this dependence is not extrapolated to zero at $n = 0$. This implies that the free energy ΔG^0 of the interaction of the enzyme with orthophosphate is -6.0 kcal mol^{-1}. This was also confirmed by direct experiments where enzyme inactivation by the reagent was inhibited in the presence of orthophosphate. Moreover, neutral triethyl phosphate ($\Delta G^0 = -4.5$ kcal mol^{-1}) acted as a competitive inhibitor of the inactivation reaction. This fact illustrates the role of hydrogen bonds in the interaction of the phosphate fragment with DNA. The experimental data altogether allow a conclusion to be made about the significant contribution of only one internucleotide phosphate in the formation of a template–DNA polymerase complex.

In studies of the interaction of DNA polymerase with a primer to form a complex, the Michaelis constant, K_m, was used to characterise the primer affinity for the complex. In preliminary affinity-based modification experiments, it was shown that K_m differed slightly from the dissociation constant of the primer complex. A series of experiments was performed aimed at the enzyme inactivation by 2′,3′-anhydroATP.[80] This was shown [80] to inactivate the enzyme in the presence of a template–primer [(poly(dT)–oligo(dA)] complex, the rate of inactivation being dependent on the concentration of this complex. Since the latter depends on the concentration of the primer, one may calculate the value of the dissociation constant for the primer complex. The equation used for calculation (3) was obtained on the assumption [80] that the primer and 2′,3′-anhydroATP bind independently with the dissociation constants K_Y and K_X, respectively, but the reaction proceeds only when both the primer and the reagent are present simultaneously:

$$\frac{1}{k_{app}} = \frac{(1 + K_X/X_0)(1 + K_Y/Y_0)}{k}. \quad (3)$$

Table 3 gives the values for several oligodeoxyadenylates and human placenta DNA polymerase α. As is seen, the differences between these constants are small. This suggests that the Michaelis constant can be used as a measure of the primer affinity. For the Klenow fragment this coincidence is even better. Therefore, the main regularities of the interaction of DNA polymerases with the

Table 3. The K_d and K_m values for the complexes of human DNA polymerase α and the Klenow fragment of DNA polymerase I from *E. coli* with primers.

Primer	DNA polymerase α		Klenow fragment	
	K_d /μM	K_m /μM	K_d /μM	K_m /μM
$d(pA)_2$	70	–	120	–
$d(pA)_5$	9	32	30	34
$d(pA)_6$	10	20	15	16
$d(pA)_7$	1.8	6.5	–	–
$d(pA)_8$	1.5	4.7	7	8
$d(pA)_{10}$	0.33	0.44	1.2	1.3
$d(pA)_{12}$	0.2	0.37	–	–

primer were analysed preferentially for the poly(dA)-oligo(dT) system using the K_m values.[78,81] The latter were determined conventionally from the dependence of the rate of the primer elongation on its concentration. The dependence of $-\log K_m$ on the primer length (the number n of the nucleotide residues) was found to be linear up to $n = 9-12$. The K_m values for a series of oligothymidylates ethylated at internucleotide phosphates were compared. It was shown that the primer affinity for the enzyme–template complex strongly depends on the presence of a charged phosphate group at the 3′-terminal nucleotide residue. Ethylation of all the other internucleotide phosphates has practically no effect on K_m. The presence of the OH group at the 3′-end of the primer is a crucial factor. The oligonucleotide d(pT)$_9$p(ddT), which carries a 2′,3′-dideoxythymidylate residue at its 3′-end and is devoid of the OH group, is a noncompetitive inhibitor of the primer. In these experiments with variable concentrations of the template, the oligonucleotide d(pT)$_9$p(ddT) behaved as a typical competitive inhibitor. This result testifies to its binding to the matrix site of the enzyme.[81,82]

The dependence of $-\log K_m$ on the length of the primer is extrapolated to ~ 4 at $n = 1$. This suggests that the mononucleotide, too, can serve as a primer for DNA polymerase. This was demonstrated in direct experiments on human placenta DNA polymerase α where poly(dT) was used as a template, dATP as a substrate, and [^{3}H]dAMP as a primer. As was shown by reversed phase chromatography of the reaction mixture, the dinucleotide d(pA)$_2$ was predominantly formed together with smaller analogues of longer oligomers. The substantial predominance of the dimer over longer oligomers attests to the distributive character of the synthesis. Similar results were obtained when [^{3}H]dATP was used as a primer. This implies in fact that DNA polymerases can initiate DNA synthesis *in vitro*, although as has already been mentioned, in all cases studied the initiation of DNA synthesis *in vivo* begins with the synthesis of a ribooligonucleotide primer.[81] These results are in favour of general regularities in the interaction of a DNA duplex with prokaryotic and eukaryotic DNA polymerases and fit well into a nonspecific sequence-independent model of DNA recognition. Indeed, the contribution of hydrophobic interactions of bases is crucial for single-stranded sites, whereas the interaction of the sugar–phosphate backbone is a critical factor in the case of DNA duplexes. It is noteworthy that the number of electrostatic interactions of both template and primer chains with the enzyme is limited. X-Ray data for the complexes of the Klenow fragment of DNA polymerase I with DNA[36] and of DNA polymerase β with DNA[38] are consistent with the conclusion about the limited number of contacts between the charged amino acid residues and DNA. For example, only one charged residue, *viz*., Arg254, was found in DNA polymerase β. Presumably, this residue forms a hydrogen bond with the phosphate of a 3′-terminal dNMP newly introduced into the primer.[38] Simulation of the interaction of a DNA duplex with reverse transcriptase of HIV-1 carried out on the basis of X-ray data also revealed a limited number of electrostatic contacts, *viz*., one or two lysine residues in the binding site of the 3′-end of the primer and several (but no more than 5–6) residues in the template region.[68] These data are consonant with the idea about the electrostatic nature of the DNA-binding site of DNA polymerases and the presence of a strong electrostatic field in this region.[83] This structure may be necessary for primary diffusion of DNA; however, the number of genuine 'strong' contacts can be restricted in order to provide the 'mobility' of DNA during the translocation and dissociation of the products.

The mechanism of interaction of dNTP with DNA polymerases was analysed by affinity-based modification of enzymes with γ-P-imidazolides of deoxynucleoside 5′-triphosphates and P-imidazolides of deoxynucleoside 5′-monophosphates.[79,84,85] It is noteworthy that these reagents inactivated DNA polymerases only in the presence of both the template and the primer and only when the template was complementary to the reagent base. This suggests that the reagents bind to the site where the substrate is selected for subsequent transfer of the nucleotide residue to the 3′-terminal hydroxy group of the primer. Quantitative estimation of nucleotide binding with a different number of phosphate groups was carried out on the basis of the protective effect of poly(dA) and oligo(dT)$_{15}$ on enzyme inactivation by dTMP-imidazolide. The values of dissociation constants for the complexes of dNMP, dNDP, and dNTP with human DNA polymerase α determined by this method are given in Table 4.[85] According to these data, the K_d values for different dNMP and dNDP are very similar, which suggests that β-phosphate hardly contributes much to the formation of dNTP complexes with DNA polymerase.

Table 4. K_d Values for the complexes of human DNA polymerase α with nucleotides in the presence of a template [poly(dA)] and a primer [d(pT)$_{15}$].[84]

Base	K_d (K_i or K_m)/μM				
	dNMP	dNDP	dNTP	NTP	γ-methylamide dNTP
C	15	11	7(15)	3	–
T	9 (28)	7	0.06(3.1)	0.05	0.03
G	7 (6)	6.0	0.3(0.64)	0.3	–
A	3 (7.8)	–	0.8	0.7	0.3

The affinity of dNMP for the enzyme increases in the following order: dCMP < dTMP < dGMP < dAMP. There was a linear correlation between ΔG^0 for dNMP binding and the retention time of the corresponding nucleosides on a reversed-phase adsorbent. Hence, the α-phosphate group of various dNMP makes an equal contribution to the total efficiency of complex formation with the enzyme. The difference in the affinity of both dNMP and dNDP for the active centre of DNA polymerase is largely due to the difference in the efficiency of hydrophobic interactions of their bases with the enzyme and does not reflect their complementariness to the template. At the same time, the situation is quite different for triphosphates. The base of dTTP is complementary to that of the template in the enzyme–poly(dA)–oligo(dT) complex and this nucleotide possesses the highest affinity for the complex. The affinity of all of the dNTP is higher than that of di- and mono-phosphates. This leads to a conclusion that the integrity of the triphosphate group of dNTP is required to enable the complementary interaction of nucleotides with the template.

Similar results were obtained for the Klenow fragment.[86] It should be emphasised, however, that the integrity of the triphosphate group is also important for the formation of a nonspecific complex between dNTP and DNA polymerases; this follows from the above data, especially for dATP and dGTP. The conclusion about the essential role of the triphosphate group for the recognition of dNT was confirmed by X-ray data, according to which the catalytic centres of DNA polymerases contain three carboxy groups that determine the Mg^{2+}-mediated orientation of the triphosphate fragment of dNTP.[37,87]

VI. Chemical methods for the study of replicative complexes

As mentioned earlier, the functioning of the enzymes catalysing replication occurs within complicated multisubunit structures. The two main problems currently arising in this respect are concerned with the detection of contacts between the polynucleotide components and individual polypeptides of the complex, on the one hand, and between individual polypeptides, on the other hand.

In this context, replicating DNA polymerase from *E. coli* (DNA polymerase III), which actually represents a replicative complex, has been studied in most detail. This enzyme is made up of ten subunits, *viz*., α, ε, β, τ, γ, δ, δ′, χ, ψ, and θ.[88] Partial

dissociation of the total complex usually takes place during its isolation; therefore, studies are carried out with intermediate incomplete complexes, *viz*., complex III, which includes α, ε, and θ subunits; complex III′, which contains additionally the subunit τ, and complex III*, which contains all subunits except for β. Complex III represents the nucleus of the total DNA polymerase. It comprises DNA polymerase α, 3′,5′-exonuclease ε, and the subunit θ the function of which has not been finally established. These subunits are responsible for the two most important functions, *i.e*., the binding of complex III′ at the initiation site including the 3′-end of the primer bound with the template, and subsequent migration of the enzyme along the template without any dissociation of the enzyme–template complex.

In order to determine the subunits to which DNA binds in the course of elongation, a series of primers with a photoreactive group at different positions relative to the 3′-end of the growing chain were synthesised and assayed.[29] Their synthesis is described in Section II. Primers with a photoreactive arylazide group at positions 3, 9, 13, 18, 22, 27 and 46 relative to the 3′-terminal fragment were used for this purpose. After formation of the enzyme–template–primer complex, this was irradiated. This resulted in the cross-linking of a definite nucleotide residue with the enzyme. Analysis of subunits of the irradiated complex showed that the presence of a photoreactive group at positions 3 and 9 resulted in labelling of the α-subunit; at position 13, of the α- and γ-subunits; at position 18, of the γ-subunit alone, and at position 22, of the γ- and β-subunits. Primers with a photoreactive group at positions 27 and 46 from the 3′-end did not cross-link with any of the subunits. Assuming that in all these complexes the 3′-end of the primer is at the same position relative to the enzyme, the results obtained can be interpreted as evidence in favour of translocation of the product relative to the components of the complete enzyme.

The second system more or less thoroughly characterised was DNA polymerase in a complex with auxiliary proteins of bacteriophage T4. Bacteriophage T4 contains about 150 genes including those encoding proteins of the replicative complex. They are usually marked by numbers corresponding to the gene number and by the symbol gp (gene product). An efficiently functioning system contains DNA polymerase gp43 itself and the auxiliary proteins gp44, gp62, and gp45. Moreover, it requires the presence of a protein that binds single-stranded DNA (gp32). The auxiliary proteins fulfil the function of a clamp. The prerequisite for the formation of a stable complex between the 3′-end of the template–primer complex and three auxiliary proteins is the presence of ATPγS, a non-hydrolysable analogue of ATP that contains the S atom at the γ-atom of P. It should be noted that DNA polymerase does not bind with this complex; its binding takes place only when a normal, hydrolysable ATP is added to the system.[89]

This system was studied by determining direct photo-cross-links upon powerful laser irradiation (one or several 8 ns pulses, irradiation power was $5-8\times10^{16}$ photons). The most intense cross-linking was observed for the protein gp32. In this case, $(dT)_{10}$ is cross-inked with a far greater efficiency than $(dC)_{10}$. Direct photo-cross-linking of $(dT)_{10}$ occurs also with the individual proteins gp43 and gp45, although less intensely; no modification of gp44 and gp62 takes place. At the same time, in a system comprising an ensemble of all the three auxiliary proteins, photomodification affects only gp62. It is worth noting that its modification requires the presence of ATP.[90]

A strong effect of ATP was also established in studies using a template–primer complex with a photoactive analogue of thymidine (NAB-dTTP) at several predetermined positions relative to the 3′-end of the primer.[91] Addition of ATP and DNA polymerase resulted in significant changes in the contacts between the auxiliary proteins and the primer. For example, the binding of gp62 with the primer in the presence of all of the other auxiliary proteins was much stronger in the presence of ATPγS, but not of ATP. This suggests that the addition of ATP results in its hydrolysis, which initiates a conformational transition, which either displaces gp62 to a larger distance from the primer or completely displaces it from the system.

In the case of eukaryotes, the fragment of a replicative complex containing DNA polymerase α and primase was studied first. Each of these enzymes consists of two subunits, which can be easily isolated as a single complex composed of four subunits, *viz*., a 180 kDa subunit (DNA polymerase itself, or sometimes as a 165 kDa subunit formed upon partial proteolysis), a 72 kDa subunit (its function is unknown), and 60 and 50 kDa subunits (γ- and δ-subunits of primase). After direct photo-cross-linking of [α-^{32}P]ATP with a complex from yeast, the label is detected only in the δ-subunit.[92] After modification of the complex with adenosine triphosphate *P*-*p*-formylphenyl ester followed by identification of modification points by catalytically competent labelling, the radioactive label was found in both subunits. Modification of an analogous complex from human placenta with adenylate *P*-*p*-formylphenyl esters with subsequent introduction of the radioactive label using [α-^{32}P]ATP revealed that the ATP ester was bound exclusively with the δ-subunit in the presence of poly(dT) and with both subunits of primase in the absence of poly(dT).[93]

In studies of a DNA polymerase–primase complex from human placenta, a template–primer duplex (25 and 10 units, respectively) was elongated with FAB-dCTP. It was found that the primer was photoadded not only to the 165 kDa subunit of DNA polymerase but also to primase subunits. However, the 72 kDa polypeptide was not labelled under these conditions. Photoaddition of FAB-dCTP to the complex with subsequent attachment of the primer (catalytically competent labelling) resulted in modification of only the 165 kDa subunit of DNA polymerase and the minor (50 kDa) subunit of primase.[21] Similar results were obtained for a DNA polymerase–primase complex from calf thymus. However, in the latter case the use of catalytically competent labelling did not result in labelling of the primase subunits at all.[23] These results suggest that photoaffinity modification can be used for gaining information about the role of various subunits of the complex in the formation of active structures and the change-over from the primase function to elongation of a newly synthesized DNA.[21]

Photoaffinity modification was also used to study the interaction of a eukaryotic replication factor (RPA) with DNA. RPA is made up of three subunits with molecular weights of 70, 32, and 14 kDa. Its functions are to stabilise the configuration of denatured DNA and to establish specific interactions of DNA with other proteins involved in replication, repair, and recombination of DNA.[94] The interaction of RPA subunits with a DNA duplex having an extended template chain, was studied by introducing a photoreactive group into the 3′-end of a radiolabelled primer using DNA polymerases α or β with 5-(5-azido-2-nitrobenzoyl)-deoxyuridine 5′-triphosphate as a substrate.[95] The p32 and p70 subunits underwent the most intense labelling upon UV irradiation. The minor (p14) subunit was not modified under these conditions. Based on the results obtained, a model for the interaction of RPA subunits with the DNA duplex was proposed (Fig. 6). It is supposed that the p70 subunit interacts with the single-stranded site of the DNA duplex, the p32 subunit is near the 3′-end of the primer, while the p14 subunit is remote. This scheme is very efficient for the studies of DNA interaction with other replication factors, such as PCNA and RFC, as well as for the study of formation of the whole replicative ensemble in the course of DNA synthesis.

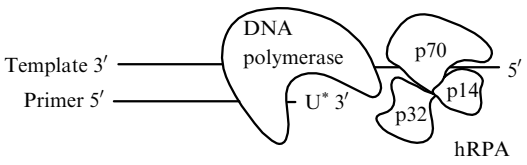


Figure 6. The interaction of the eukaryotic replication factor (hRPA) with a DNA duplex containing a photoreactive residue, dNMP(U*), introduced into the 3′-end of the primer by DNA polymerase; p70, p32, and p14 are the RPA subunits.

VII. Conclusion

Chemical approaches to the study of the mechanism of action of enzymes are not confined to the use of chemical modification for the analysis of template biosynthesis of DNA. Various substrate analogues obtained mostly by chemical methods have found wide use in enzymological studies. With respect to DNA polymerases, this approach has been successfully developed in studies by Kraevsky.[96,97]

The data presented in this review demonstrate that the use of chemical (particularly, affinity-based) modification to elucidate the nature of active centres of DNA polymerases is a very practicable approach. However, the results obtained by this method must be analysed together with X-ray data. Analysis of these data together with the results obtained by site-directed mutagenesis is no less important. Most recent studies resort to methods of computer simulation more and more often because they allow a more detailed interpretation of information obtained by experimental methods.

The processes catalysed by DNA polymerases are characterised by a high degree of dynamism. The same can be said of the stage of the monomer selection where the monomer is tested for complementariness to the encoding element of the template prior to the nucleotidyl transferase reaction), to the translocation stage, where the template-elongation product complex migrates relative to the active centre of the enzyme, and to the repair stage involving $3' \rightarrow 5'$-exonuclease where the 3′-end of the synthesised chain is translocated from the polymerase to the exonuclease centre. These events occur over milliseconds; one of the few possible ways to monitor these transient changes in the contact sites of enzymes, templates, and products is fast photochemical labelling of the enzyme by photoreactive analogues of the substrate or the primer. This can be done for a millisecond interval using pulse irradiation. Although these experiments have just begun, a considerable chemical ground for this method including the synthesis of photoreactive analogues of dNTP and introduction of photoreactive nucleotide residues into definite positions of the primer chain has been created.

An important and very promising result of the most recent investigations is the elaboration of new techniques for catalytically competent labelling of enzymes with dNTP derivatives carrying a reactive group in their heterocyclic base. These derivatives retain their triphosphate fragment after covalent cross-linking with the enzyme and can sometimes act as donors of the nucleotide residue. This opens up interesting possibilities for the study of the mechanism of the nucleotidyl transferase reaction catalysed by the enzyme. Studies of elongation of primers covalently immobilised on the enzyme also seem to hold much promise. Since this system allows simultaneous realisation of several elongation acts and, as a consequence, of several translocation acts, it remains to be hoped that a convenient model for the study of the translocation mechanism will be developed despite limited mobility of the primer.

One of the peculiarities of studies of nucleic acid polymerases by the chemical modification method is the simultaneous presence in the system of single- and double-stranded nucleic acids and a protein (or a group of proteins). As mentioned above, simultaneous modification of the enzyme and the template was observed for primers carrying an azidoperfluorobenzoyl group linked to a heterocyclic base.[14,21] In some cases, this can impede the interpretation of results of affinity-based modification, *e.g.*, due to the impossibility of separation of DNA chains or because simultaneous binding of two or more photoreactive groups of the primer with the template and the enzyme yields products that cannot be easily identified. Therefore, preference may be given to reagents endowed with the ability to selectively modify proteins without any cross-linking of the primer with the matrix, *e.g.*, when the primer contains NAB-dUMP residues.[15] Studies on a model system in which the amino acid residue and the photoreactive group (azidoperfluorobenzoyl, 5-azido-2-nitrobenzoyl or *p*-azidoanilide) are artificially converged by addition to the 3′- and 5′-terminal phosphates of complementary oligonucleotides, respectively, showed that only in the former case can the cross-linking also take place in the absence of the amino acid residue at the 3′-end.[98] Therefore, among other arylazides used for protein modification, derivatives with a 5-azido-2-nitrobenzoyl or *p*-azidoanilide groups are more preferable.

Chemical approaches, in the first place, covalent cross-linking of reactive primer-template complexes with reactive fragments at predetermined positions of these complexes are an indispensable tool in the study of both RNA polymerases and transcription complexes, on the one hand, and replicative complexes, on the other hand. In particular, they are used for the study of the mutual arrangement of individual components of the above complexes (*e.g.*, eukaryotic DNA polymerase and auxiliary proteins, such as SSB, PCNA, RFC) and their translocation relative to the template – primer complex in the course of elongation.

This work has been supported by the Russian Foundation for Basic Research (Project No. 96-04-50114).

References

1. D G Knorre, N V Kudryashova, O I Lavrik *Usp. Khim.* **66** 401 (1997) [*Russ. Chem. Rev.* **66** 363 (1997)]
2. D G Knorre, S D Myzina *Biologicheskaya Khimiya* (Biological Chemistry) (Moscow: Vysshaya Shkola, 1992)
3. J Kuriyan, M O'Donnell *J. Mol. Biol.* **234** 915 (1993)
4. G A Nevinskii, O I Lavrik, in *Bioorganicheskaya Khimiya (Itogi Nauki i Tekhniki)* [Bioorganic Chemistry (Advances in Science and Engineering Series)] (Moscow: Izd. VINITI, 1988) Vol. 13, p. 4
5. M A Grachev, D G Knorre, O I Lavrik, in *Uspekhi Biologicheskoi Khimii* (Progress in Biological Chemistry) (Moscow: Nauka, 1986) Vol. 27, p. 30
6. V N Buneva, T V Demidova, D G Knorre, N V Kudryashova, A G Romashchenko, M G Staroobrazova *Mol. Biol.* **14** 1080 (1980) [a]
7. V N Buneva, N V Kudryashova, L T Nebrat, A G Romashchenko, T A Chimitova, L F Yushkova *Dokl. Akad. Nauk SSSR* **268** 243 (1983) [b]
8. G A Nevinskii, S V Doronin, O I Lavrik *Biopolim. Kletka* **1** 247 (1985)
9. S V Doronin, G A Nevinskii, V V Khomov, O I Lavrik *Mol. Biol.* **25** 358 (1991) [a]
10. N V Kudryashova, M Yu Shamanina, T S Godovikova, E A Anan'ko, F F Akhmadieva, A G Romashchenko *Biokhimiya* **58** 224 (1993) [c]
11. W A Wlassoff, M I Dobrikov, I V Safronov, R Y Dudko, V S Bogachev, V V Kandaurova, G V Shishkin, G M Dymshits, O I Lavrik *Bioconjug. Chem.* **6** 352 (1995)
12. O I Lavrik, R Prasad, W A Beard, I V Safronov, M I Dobrikov, D K Srivastava, G V Shishkin, T G Wood, S H Wilson *J. Biol. Chem.* **271** 21891 (1996)
13. M I Dobrikov, S V Doronin, I V Safronov, G V Shishkin, O I Lavrik *Khimiya v Interesakh Ustoichivogo Razvitiya* (The Chemistry for Stable Progress) **2** 529 (1994)
14. S V Doronin, M I Dobrikov, M Buckle, P Roux, H Buc, O I Lavrik *FEBS Lett.* **354** 200 (1994)

15. N V Shcherbik, S N Khodyreva, V A Vlasov, M I Dobrikov, G M Dymshits, O I Lavrik *Mol. Biol.* **31** 344 (1997) [a]
16. I V Safronov, N V Shcherbik, S N Khodyreva, V A Vlasov, M I Dobrikov, G V Shishkin, O I Lavrik *Bioorg. Khim.* **23** 576 (1997) [d]
17. J Rush, W H Konigsberg *J. Biol. Chem.* **265** 4821 (1990)
18. B M Moore II, R K Jalluri, M B Doughty *Biochemistry* **35** 11 642 (1996)
19. B M Moore II, K Li, M B Doughty *Biochemistry* **35** 11634 (1996)
20. T Yamaguchi, M Saneyoshi *Nucl. Acids Res.* **24** 3364 (1996)
21. S V Doronin, M I Dobrikov, O I Lavrik *FEBS Lett.* **313** 31 (1992)
22. T S Godovikova, D M Kolpashchikov, T N Orlova, V A Richter *FASEB J.* **11** 1367 (1997)
23. I A Potapova, S N Khodyreva, O V Vladimirova, T V Safronov, M I Dobrikov, W A Wlassoff, G V Shishkin, O I Lavrik *Bioconjug. Chem.* (1998) (in the press)
24. R K Evans, J D Johnson, B E Haley *Proc. Natl. Acad. Sci. USA* **83** 5382 (1986)
25. C E Catalano, S J Benkovic *Biochemistry* **28** 4374 (1989)
26. G A Nevinskii, A S Levina, V N Podust, O I Lavrik *Bioorg. Khim.* **13** 58 (1987) [d]
27. T I Kolocheva, G A Nevinsky, V A Volchkova, A S Levina, V V Khomov, O I Lavrik *FEBS Lett.* **248** 97 (1989)
28. T I Kolocheva, G A Nevinsky, A S Levina, V V Khomov, O I Lavrik *J. Biomol. Struct. Dyn.* **9** 169 (1991)
29. J A Reems, S Wood, C S McHenry *J. Biol. Chem.* **270** 5606 (1995)
30. C R Guest, R A Hochstrasser, C G Dupuy, D J Allen, S J Benkovic, D P Millar *Biochemistry* **30** 8759 (1991)
31. N F Insdorf, D F Bogenhagen *J. Biol. Chem.* **264** 21491 (1989)
32. R A Bambara, C B Jessee *Biochim. Biophys. Acta* **1088** 11 (1991)
33. C M Joyce, T A Steitz *Trends Biochem. Sci.* **12** 288 (1987)
34. H Klenow, I Henningsen *Proc. Natl. Acad. Sci. USA* **65** 168 (1970)
35. L S Beese, J M Friedman, T A Steitz *Biochemistry* **32** 14095 (1993)
36. L S Beese, V Derbyshire, T A Steitz *Science* **260** 352 (1993)
37. M R Sawaya, H Pelletier, A Kumar, S H Wilson, J Kraut *Science* **264** 1930 (1994)
38. H Pelletier, M R Sawaya, A Kumar, S H Wilson, J Kraut *Science* **264** 1891 (1994)
39. V N Pandey, N Kaushik, M J Modak *J. Biol. Chem.* **269** 21828 (1994)
40. V N Pandey, K R Williams, K L Stone, M J Modak *Biochemistry* **26** 7744 (1987)
41. C E Catalano, D J Allen, S J Benkovic *Biochemistry* **29** 3612 (1990)
42. O I Lavrik, G A Nevinskii, I A Potapova, V V Khomov *Mol. Biol.* **22** 485 (1988) [a]
43. A Basu, M J Modak *Biochemistry* **26** 1704 (1987)
44. V N Pandey, M J Modak *J. Biol. Chem.* **263** 6068 (1988)
45. D G Knorre, V N Buneva, G I Baram, T S Godovikova, V F Zarytova *FEBS Lett.* **194** 64 (1986)
46. Yu N Vorob'ev *Bioorg. Khim.* **17** 211 (1991) [d]
47. P M Mohan, A Basu, S Basu, K I Abraham, M J Modak *Biochemistry* **27** 226 (1988)
48. M Cowart, K J Gibson, D J Allen, S J Benkovic *Biochemistry* **28** 1975 (1989)
49. D J Allen, S J Benkovic *Biochemistry* **28** 9586 (1989)
50. D K Srivastava, R K Evans, A Kumar, W A Beard, S H Wilson *Biochemistry* **35** 3728 (1996)
51. F di Marzo Veronese, T D Copeland, A L DeVico, R Rahman, S Oroszlan, R C Gallo, M G Sarngadharan *Science* **231** 1289 (1986)
52. J Hansen, T Schulze, K Moelling *J. Biol. Chem.* **262** 12393 (1987)
53. J Hansen, T Schulze, W Mellert, K Moelling *Embo J.* **7** 239 (1988)
54. L Ratner, W Haseltine, R Patarca, K J Livak, B Starcich, S F Josephs, E R Doran, J A Rafalski, E A Whitehorn, K Baumeister, L Ivanoff, S R Pettewing, M K Pearson, J A Lautenberger, T S Papas, J Ghrayeb, N I Chong , R C Gallo, F Wong-Staal *Nature* (*London*) **313** 277 (1985)
55. S P Becerra, G M Clore, A M Gronenborn, A R Karlstrom, S J Stahl, S H Wilson, P T Wingfield *FEBS Lett.* **270** 76 (1990)
56. L A Kohlstaedt, J Wang, J M Friedman, P A Rice, T A Steitz *Science* **256** 1783 (1992)
57. A Jacobo-Molina, J Ding, R G Nanni, A D Clark Jr , X Lu, C Tantillo, R L Williams, G Kamer, A L Ferris, P Clark, A Hizi, S H Hughes, E Arnold *Proc. Natl. Acad. Sci. USA* **90** 6320 (1993)
58. W Metzger, T Hermann, O Schatz, S F Le Grice, H Heumann *Proc. Natl. Acad. Sci. USA* **90** 5909 (1993)
59. B M Wohrl, C Tantillo, E Arnold, S F Le Grice *Biochemistry* **34** 5343 (1995)
60. C Barat, V Lullien, O Schatz, G Keith, M T Nugeyre, F Gruninger-Leitch, F Barre-Sinoussi, S F Le Grice, J L Darlix *Embo J.* **8** 3279 (1989)
61. C Barat, S F Le Grice, J L Darlix *Nucl. Acids Res.* **19** 751 (1991)
62. Y Mishima, J A Steitz *Embo J.* **14** 2679 (1995)
63. C Isel, C Ehresmann, G Keith, B Ehresmann, R Marquet *J. Mol. Biol.* **247** 236 (1995)
64. N Cheng, B M Merrill, G R Painter, L W Frick, P A Furman *Biochemistry* **32** 7630 (1993)
65. N Sheng, D Dennis *Biochemistry* **32** 4938 (1993)
66. J A Peliska, S J Benkovic *Science* **258** 1112 (1992)
67. S N Khodyreva, A S Levina, M I Dobrikov, A A Koshkin, O I Lavrik, M Ricchetti, P Roux, H Buc *Eur. J. Biochem.* (1998) (in the press)
68. T Hermann, T Meier, M Gotte, H Heumann *Nucl. Acids Res.* **22** 4625 (1994)
69. A Araya, E Hevia, S Litvak *Nucl. Acids Res.* **8** 4009 (1980)
70. L L W Mitchell, B S Cooperman *Biochemistry* **31** 7707 (1992)
71. A Basu, K K Ahluwalia, S Basu, M J Modak *Biochemistry* **31** 616 (1992)
72. R W Sobol, R J Suhadolnik, A Kumar, B J Lee, D L Hatfield, S H Wilson *Biochemistry* **30** 10623 (1991)
73. M A Grachev, T I Kolocheva, E A Lukhtanov, A A Mustaev *Eur. J. Biochem.* **163** 113 (1987)
74. R L Mitina, A A Mustaev, E F Zaychikov, V V Khomov, O I Lavrik *FEBS Lett.* **272** 181 (1990)
75. T A Steitz, S J Smerdon, J Jager, C M Joyce *Science* **266** 2022 (1994)
76. A A Mustaev, G N Godson *J. Biol. Chem.* **270** 15711 (1995)
77. D G Knorre (Ed.) *Affinnaya Modifikatsiya Biopolimerov* (Affine Modification of Biopolymers) (Novosibirsk: Nauka, 1983)
78. O I Lavrik, G A Nevinskii *Biokhimiya* **54** 757 (1989) [c]
79. D G Knorre, O I Lavrik, G A Nevinsky *Biochimie* **70** 655 (1988)
80. V N Podust, T O Korobeinicheva, G A Nevinsky, A S Levina, O I Lavrik *Mol. Biol. Rep.* **14** 247 (1990)
81. G A Nevinsky, A G Veniaminova, A S Levina, V N Podust, O I Lavrik *Biochemistry* **29** 1200 (1990)
82. I Sh Bukhrashvili, D Z Chinchaladze, O I Lavrik, A S Levina, G A Nevinsky, D A Prangishvili *Biochim. Biophys. Acta* **1008** 102 (1989)
83. P N Yadav, J S Yadav, E Arnold, M J Modak *J. Biol. Chem.* **269** 716 (1994)
84. S V Doronin, O I Lavrik, G A Nevinsky, V N Podust *FEBS Lett.* **216** 221 (1987)
85. G A Nevinskii, S V Doronin, V N Podust, O I Lavrik *Mol. Biol.* **21** 1070 (1987) [a]
86. S V Doronin, G A Nevinsky, T O Malygina, V N Podust, V V Khomov, O I Lavrik *FEBS Lett.* **259** 83 (1989)
87. H Pelletier, M R Sawaya, W Wolfle, S H Wilson, J Kraut *Biochemistry* **35** 12 762 (1966)
88. C S McHenry *J. Biol. Chem.* **266** 19 127 (1991)
89. M C Young, M K Reddy, P H Von Hippel *Biochemistry* **31** 8675 (1992)
90. J W Hockensmith, W L Kubasek, T A Cross, M K Dolejsi, W R Vorachek, P H Von Hippel, in *Proceedings of UCLA Symposium on Molecular and Cellular Biology. DNA Replication and Recombination* (Eds R McMacken, T J Kelly) (New York: Alan R Liss, 1987) Vol. 47, p. 111
91. T L Capson, S J Benkovic, N G Nossal *Cell* **65** 249 (1991)
92. M Foiani, A J Lindner, G R Hartmann, G Lucchini, P Plevani *J. Biol. Chem.* **264** 2189 (1989)
93. O D Zakharova, V N Podust, A A Mustayev, R O Anarbaev, O I Lavrik *Biochimie* **77** 699 (1995)
94. M S Wold *Annu. Rev. Biochem.* **66** 61 (1997)

95. O I Lavrik, H P Nasheuer, K Weisshart, M S Wold, R Prasad, W A Beard, S H Wilson, A Favre *Nucl. Acids Res.* (1998) (in the press)
96. A A Kraevskii *Mol. Biol.* **26** 725 (1992) [a]
97. A A Kraevsky, K A Watanabe *Nucleosides Nucleotides* **12** 649 (1993)
98. T S Godovikova, M V Berezovskii, D G Knorre *Bioorg. Khim.* **21** 858 (1995) [d]

[a] — *Mol. Biol. (Engl. Transl.)*
[b] — *Dokl. Chem. Technol., Dokl. Chem. (Engl. Transl.)*
[c] — *Biochemistry (Engl. Transl.)*
[d] — *Russ. J. Biorg. Chem. (Engl. Transl.)*