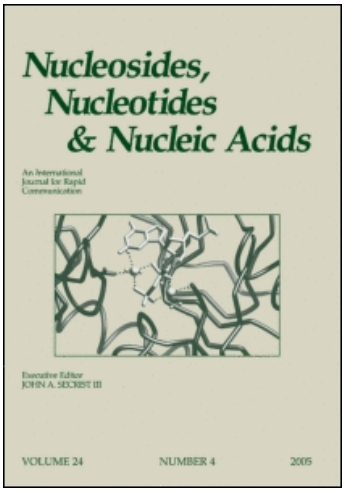


Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Covalent Interaction of Proteins and Nucleic Acids. Synthetic and Natural Nucleotide-Peptides

B. A. Juodka[a]
[a] Vilnius State University, Vilnius, USSR

To cite this Article Juodka, B. A.(1984) 'Covalent Interaction of Proteins and Nucleic Acids. Synthetic and Natural Nucleotide-Peptides', Nucleosides, Nucleotides and Nucleic Acids, 3: 5, 445 — 483
To link to this Article: DOI: 10.1080/07328318408081283
URL: http://dx.doi.org/10.1080/07328318408081283

NUCLEOSIDES & NUCLEOTIDES, 3(5), 445-483 (1984)

COVALENT INTERACTION OF PROTEINS AND NUCLEIC ACIDS. SYNTHETIC AND NATURAL NUCLEOTIDE-PEPTIDES

B. A. Juodka

Vilnius State University, Vilnius, Lithuanian SSR, USSR

Abstract. The review reports data on the occurrence, structure and functions of covalent nucleotide and NA-protein complexes. Methods of synthesis of model nucleotide-peptides with phosphoester and phosphoamide bonds and their hydrolytic properties are discussed. Separate sections are devoted to methods of specific cleavage of phosphoamide and phosphoester bonds in nucleotide-peptides and their application in biochemistry.

In recent years covalent NA-protein structures have been isolated from various prokaryotic and eukaryotic sources. On the functioning of certain enzymes, covalent nucleotide-protein complexes are formed. Researchers are interested in their structure, including the interbiopolymer bond type, and functions. Synthetic nucleotide-peptides with different covalent bonds between the components can play an important role in solving these problems. It has been found that in practically all natural nucleotide- and NA-protein structures, with the exception of compounds containing activated amino acid, the protein part of the molecules is linked to the NA or the nucleotide through the phosphate residue. Therefore, various derivatives of nucleotides and oligonucleotides were studied, where the amino acids (peptides) are linked just through the phosphoric acid residue. On the other hand, the protein molecule can be linked to the phosphate residue through any functional group of its own. Therefore, the properties of models with the phosphoester bond, in the formation of which amino acids containing the hydroxyl group are involved, have been studied, as well as those of

 0732-8311/84/0305-0445$3.50/0

models with the phosphoamide bond, in the formation of which the imidazole ring of histidine, the ϵ-amino group of lysine and the α-amino group of various amino acids are involved. Of particular importance was a detailed study of the effect of the structure of nucleotide and amino acid components and of the functional environment on the behavior of a particular nucleotide-peptide linkage in synthetic nucleotide-peptides. On isolation and various treatment of NA-protein structures, functional groups of the NA and the protein can find themselves near the interbiopolymer bond linking the individual components of covalent nucleic acid and nucleotide-protein complexes. It is necessary to know how they will affect the properties of this bond. Migration of the NA- or the nucleotide residue from one functional group of amino acids to another is especially dangerous. Such migration can lead to an erroneous conclusion about the nature of the nucleotide-peptide linkage in the native complex. A study of model compounds can help to avoid the danger. The purpose of the present review is to summarize data on the occurrence, structure and functions of natural covalent nucleotide- and NA-protein complexes, as well as data on the synthesis, the study of properties, and the application of synthetic nucleotide-peptides of the phosphoamide and the phosphoester type. The results concerning synthetic nucleotide-peptides, obtained before 1968, were published in Review.[1] Some data on natural covalent nucleotide- and NA-protein complexes were reported in Review.[2]

1. NATURAL NUCLEOTIDE- AND NA-PROTEIN COMPLEXES

Biological functions of nucleic acids manifest themselves in direct contact with proteins. Nearly all nucleic acids in vivo are found in the form of complexes with proteins. The interaction of nucleic acids and proteins in such complexes is mainly non-covalent. However, in the process of protein biosynthesis, under the action of a number of enzymes, as well as on the synthesis and functioning of certain nucleic acids, covalent nucleotide- and NA-protein structures (natural nucleotide-peptides) are also formed. In recent years a number of works dealing with the structure and functions of natural nucleotide-peptides have appeared. The aim of the present section of the review is to consider the occurrence, structure and functions of nucleotide- and NA-protein complexes.

1.1. Nucleotide-protein complexes

A number of enzymes forming covalent complexes with nucleotides are known at present. These are enzymes, catalyzing different chemical reactions

according to the "ping-pong" mechanism, such as DNA and RNA ligases, glutamine synthetase, aspartylkinase, galactose-1-phosphate uridylyltransferase, a number of phosphodiesterases, RNA-polymerase and, possibly, aminoacyl-tRNA synthetase.

1.1.1. Complexes of NA ligases and adenylic acid. In 1967 new enzymes, DNA ligases (EC 6.5.1), were isolated from Escherichia coli and E. coli infected with phages. The DNA ligases catalyze NAD^+ or ATP-dependent linking of oligonucleotides (polynucleotides) on the complementary matrix and play an important role in the reparation, replication and genetic recombination (see Reviews[3-5]).

The DNA ligase has been found in or isolated from various prokaryotic and eukaryotic sources (see Reviews[3-5]), which indicates the universality of the enzyme. Shortly after the discovery of the DNA ligase, there appeared works discussing the mechanism of its action. It was shown (see Review[5]) that one of the substrates of the E. coli DNA ligase (EC 6.5.1.2) is the coenzyme NAD^+, which breaks down into AMP and nicotinamide mononucleotide as a result of the reaction. The DNA ligase from E. coli infected with phages (see Review[5]), from animals and plants (see Review[4]) (EC 6.5.1.1) requires the presence of ATP to put the ligase in action. In the course of the reaction the ATP breaks down into AMP and pyrophosphate. In the first step of the DNA ligase reaction, the ligase-AMP complex, which was discovered in 1967, is formed in all cases.[6] It was found[7] that, in the complex from E. coli and E. coli infected with T4 phages, the AMP is linked to the enzyme through the ϵ-amino group of lysine by a phosphoamide bond. In the case of eukaryotes the fine structure of the adenylate-DNA ligase complex has not been determined.

In 1972 a new enzyme, RNA ligase (EC 6.5.1.3) (see Review[8]), catalyzing ATP-dependent cyclization and linking of oligonucleotides, was isolated from E. coli infected with T4 phages. It was shown (see Review[8]) that, in the first step of the RNA ligase reaction, the enzyme-AMP complex is also formed. Apparently,[9] in this case, the AMP is also linked to the ligase through the ϵ-amino group of lysine by a phosphoamide bond.

1.1.2. Nucleotide-peptides formed on regulating the action of glutamine synthetase. Glutamine synthetase (EC 6.3.1.2) catalyzes important metabolic amino acid and nitrogen reactions-biosynthesis of glutamine from glutaminic acid

and ammonia in the presence of ATP. The enzyme was found in most various prokaryotic and eukaryotic sources (see References[2]). Glutamine synthetase, which is a very complex enzyme, has been studied most thoroughly in the case of E. coli. It was found[10] that, under the action of adenylyltransferase, adenylic acid is linked to glutamine synthetase to form the AMP-glutamine synthetase complex, which lacks enzymatic activity. It was shown[11] that the AMP is linked to the tyrosine residue of glutamine synthetase by a phosphodiester bond. The activity of adenylyltransferase turned out to be regulated by protein P_{11} (M 44,000).[12] Uridylylation and deuridylylation of protein P_{11}, in their turn, regulate its activity. It was found[13] that in the UMP-protein P_{11} complex the nucleotide is linked to the tyrosine residue of the protein by a phosphodiester bond.

Thus, on regulating the action of glutamine synthetase, two covalent nucleotide-protein complexes are formed. In both cases the nucleotide is linked to the tyrosine residue of the proteins by a phosphodiester bond.

1.1.3. The AMP-aspartylkinase complex. Aspartylkinase (EC 2.7.2.4) catalyzes the ATP-dependent conversion of aspartic acid to 4-phosphoaspartic acid. In E. coli three aspartylkinases have been found.[14,15] It was shown[14,15] that aspartylkinase III in vivo is found both in the active and in the inactive forms. The inactive form of the enzyme is characterized by the presence of the AMP residue linked to it. The nature of the covalent bond by which the AMP is linked to the protein is unknown.

Adenylylation of aspartylkinase III is connected with the regulation of its action.[15]

1.1.4. UMP-galactose-1-phosphate uridylyltransferase. Galactose-1-phosphate uridylyltransferase (EC 2.7.7.12) catalyzes the reversible transformation of UDP-glucose into UDP-galactose. The enzyme was found in microorganisms, animals and plants (see References[2]). The enzyme from E. coli has been studied most extensively. It was shown[16] that the reaction proceeds through the UMP-galactose 1-phosphate uridylyltransferase complex (UMP-E). It has been found recently[17] that the UMP is linked to the protein by a phosphoamide bond through the imidazole ring of histidine.

1.1.5. Other nucleotide-enzyme complexes. From various sources a number of other enzymes have been isolated, the functioning of which leads to the formation of stable complexes with nucleotides. To such enzymes belong

nucleotide phosphodiesterases (EC 3.1.4.1). It has been shown or discussed that, at one of the stages of the action of 5'-nucleotide phosphodiesterase from bovine intestines,[18] snake venom,[19] certain plants,[20] guanylyltransferase from HeLa cells,[21] 3'-nucleotide phosphodiesterase from spleen,[22] exonuclease A5,[23] a covalent nucleotide-enzyme complex is formed. However, the nature of the bond has not been determined with certainty.

It is assumed[24] that, on the functioning of aminoacyl-tRNA-synthetases, a covalent complex with AMP is also formed. It has been found that adenylylation and deadenylylation is one of the ways of regulating the action of E. coli RNA polymerase.[25] It has been shown recently that the AMP is covalently bonded to the α-subunit of RNA polymerase.[26]

In conclusion it may be said, that under the action of a number of enzymes, the nucleotides are covalently bonded to the protein. In some cases nucleotide-protein complexes are formed as intermediates of enzymatic reactions, while in others the linking of the nucleotide to the enzyme plays the regulating role. However, the nature of the covalent bond in nucleotide-enzyme complexes has been determined only in a few cases.

1.2. DNA-protein structures

In cells of all organisms DNA are specifically linked to proteins, peptides and other natural compounds, forming a complex functioning genetic apparatus. The interaction of DNA and proteins is mainly non-covalent. However, it has long been found that DNA of various origin contain about 0.1-3% residual protein, the functions of which have not been found (see References[2]). The first attempt to determine the structure of this type of DNA-protein complex were made with DNA from rat liver nuclei.[27] The authors have shown that the bond between the DNA and the protein is labile in acid. They failed, however, to determine the nature of the bond. With DNA from bovine sperm and human leucocytes, the amino acid composition of the residual peptide has been determined,[28] aminohydroxy acids accounting for the greater part of them. It might be thought that it is the aminohydroxy acids that play a particular role in the formation of the DNA-peptide complex and are linked directly to the polynucleotide chain. The discovery of O-phosphoserine in acidic hydrolyzates of DNA and a change in the molecular weight of DNA under the action of hydroxylamine led to the assumption[28] that peptides or proteins bind individual single- and double-stranded fragments of DNA into a native molecule. Analogous protein linkers were also assumed in the case of DNA from E. coli.[29] A further study of the DNA-peptide

complex from E. coli has shown that under the action of pronase there is no decrease in the molecular weight of DNA. This fact allowed the authors to abandon the hypothesis about protein binding in the DNA and led them to assume that the peptides are linked to the DNA through the internucleotide phosphate residue by a phosphoamide bond.

Similar DNA-peptides have been isolated from Ehrlich ascites cells,[30-32] animals[33-35] and plants.[36] Under the action of pronase on DNA from Ehrlich ascites cells, a decrease in the sedimentation constant from 70 to 26 s was observed.[30] The authors assume that the DNA under consideration has protein linkers. Their molecular weight is 54,000-68,000. The molecular weight of the DNA, isolated after alkaline denaturation and treatment of the DNA-protein complex with proteases, is $8.5 \cdot 10^6$ [31]. It is interesting to note that incubation of the complex with nuclease S_1 has led to DNA with the same molecular weight. The whole complex of experimental data made it possible to conclude that the DNA-protein complexes concerned contain protein linkers in both chains of the DNA and that there is no overlap. Recent data on the nature of the chemical bonds between the individual fragments of the DNA suggest that they are linked by a phosphodiester bond.[37]

The functions of protein linkers are unknown. It is possible that they impart flexibility to the rigid DNA molecule and facilitate its packing in the virus particle[28] or in chromatin.[34] It is assumed[31] that peptide linkers are initiation points of replication of eukaryotic DNA.

In E. coli infected with phage P1, transducing particles consisting of phage P1, E. coli DNA and protein have been found.[38] They were labile in alkali, which led the authors to assume the presence of the ester bond.

Covalently bound proteins have recently been found in DNA from human and simian adenoviruses (see References[2,39]), bacteriophages ϕ29 (see References[39,40]), ϕ15, M2,[42] GA-1.[43,44] These DNA, isolated by extraction in the presence of proteolytic enzymes, have a linear double-stranded structure with a molecular weight of about $1\text{-}1.2 \cdot 10^7$ in the case of bacteriophages and $2.3 \cdot 10^7$ in the case of adenoviruses. The DNA isolated in the absence of proteases was shown to have a circular multistrand focus or, with a high virus concentration, an oligomeric structure (see References[2]). Most of such structures are resistant to various chemical agents.[2] However, proteases or sodium dodecylsulphate readily convert all these forms of DNA to linear bihelical molecules.[2] The formation of these DNAs was found to be determined by non-specific interaction of the proteins linked to both ends of linear bihelical DNA (see References[2]).

DNA-protein structures have been studied more extensively in the case of DNA from type 2 and type 5 human adenoviruses (Ad2, Ad5) and from bacteriophage ϕ29. It has been found[45] that the amount of protein in such DNA-protein complexes does not exceed 2% of the total weight of the complex. The molecular weight of the protein fragments is about 55,000 with adenoviruses[45] and in the range of 28,000-31,000 with bacteriophage ϕ29.[39] The stability of the complexes to a number of chemical agents made it possible to conclude that the protein is linked to the DNA by a covalent bond (see References[2]). Recent research has shown (see References[2]) that the protein is linked to both 5' ends of the DNA. It has been found[46] that the 5' ends of Ad5 DNA, to which the proteins are linked, are non-hybridized. The length of the single-stranded fragment is less than 10 nucleotides. Elucidation of the nucleotide sequence of the terminal fragments of DNA from Ad2,[47,48] Ad5[49] and phage ϕ29[50,51] has shown that DNA from adenoviruses have dCMP at the 5' ends of both chains, while ϕ29 DNA has dAMP at these sites. In both cases the protein was found to be linked to the DNA by a phosphoester bond through the hydroxyl group of serine.[40,52]

The functions of the protein in the DNA-protein complex are unknown. It is assumed that the protein can protect the DNA from the action of exonucleases and participate in transfection (see References[2]). There exists a view[48] that the protein can play a structural role in DNA packing in the virus particle. The functions of this protein are most probably connected with DNA replication.[39] A model has been proposed,[45] that involves the DNA-protein complex in the replication mechanism and does not require cyclization and concatamerization of DNA. It has been shown recently (see Review[53]) that adenoviruses have an enzymatic system catalyzing the formation of a covalent deoxycytidylate - protein complex. The molecular weight of the protein is not 55,000 but 80,000. The linkage in the deoxycytidylate-protein complex is analogous to that in the DNA-protein complex. All this implies that the linking of the protein with a molecular weight of 80,000 to dCMP is the first step in adenovirus DNA replication. An analogous mechanism of initiation of replication has been proposed in the case of phage ϕ29.[53]

The above examples of covalent DNA-protein structures are not isolated cases. It was found in 1975[55] that on treatment of DNA from SV 40 or from virus-infected cells with 1M NaCl, two or three proteins remain firmly bound to it. It has been shown recently[56] that, on isolation of SV 40 DNA in the presence of sodium dodecylsulphate, one of two complementary strands of circular DNA is broken, and a covalent DNA-protein complex is formed. The protein was found to

be linked to one end of the broken strand of the circular DNA. However, it is not known to which end of the polynucleotide chain and by what bond it is linked. Practically nothing is known about the functions of the protein complex. It is possible that the protein plays a certain role in the initiation of DNA replication, as well as at early stages of transfection.[55,56]

A number of other covalent DNA-protein complexes formed under the action of protein on supercoiled DNA have been described. One of such proteins is specific endonuclease in the case of the relaxation complex of Helinsky.[57,58] The relaxation complex from plasmids Col E1 was shown to consist of a supercoiled circular DNA and three proteins with molecular weights of 60,000, 16,000 and 11,000. In the complex these proteins are linked to the DNA by non-covalent bonds. The action of proteases or sodium dodecylsulphate on the complex induces DNA relaxation,[57] which results in the breakage of its heavy strand. The DNA molecule is converted to its open-circular form, and the protein with a molecular weight of 60,000 remains linked to the broken heavy strand by a covalent bond.[57] The protein was found to be linked to the 5'-terminal nucleotide of the broken strand of the DNA.[58] Neither the nature of the chemical bond between the DNA and the protein nor the biological functions of the DNA-protein complex are known. It is assumed[58] that the protein of the complex may function as nicase and be a signal for the beginning of replication. The first experimental evidence has recently been obtained[59] confirming that the DNA-protein complex from the plasmids Col E1 is closely connected with their transport during conjugation.

Another group of proteins cleaving supercoiled DNA are relaxing proteins. In the literature they are also referred to as ω-proteins, nicking-closing proteins, swivelases, DNA topoisomerases. They have been isolated from a number of prokaryotes and eukaryotes (see Reviews[60-65]). All topoisomerases may be divided into two groups: type I topoisomerases, which break only one strand of supercoiled double-stranded DNA and do not require any cofactors, and type II topoisomerases, which break both chains of supercoiled DNA. DNA gyrase also belongs to the latter group. Topoisomerases of both types catalyze the breaking and rejoining of DNA backbone bonds. Thus, topoisomerases act both as endonucleases and as DNA ligases. A detailed study has been made of E. coli topoisomerases I, which can also, under certain conditions, interact with single-stranded DNA, forming a covalent complex. Topoisomerase I also catalyzes the formation of duplex circular DNA from complementary single-stranded rings (see Reviews[61-63]). Eukaryotic topoisomerases also interact with single-stranded

DNA.[66] Prokaryotic topoisomerases are very much alike in their effect and partly in their structure. All of them partially eliminate only negative supercoils of DNA, require Mg^{2+}, are inhibited by 0.2M NaCl or KCl. However, the degree of relaxation and the molecular weight of proteins varies with proteins from different sources.

All eukaryotic type I topoisomerases relax both positively and negatively supercoiled DNA and, unlike the analogous proteins from prokaryotes, require monovalent ions but not Mg^{2+}. The only exception is rat liver mitochondria topoisomerase,[67] which relaxes only positively supercoiled DNA. It has been found (see Reviews[60-63]) that DNA relaxation with the participation of proteins from any source proceeds through the formation of a DNA-protein structure, where the protein is covalently bonded to the cleavage site, as was the case with certain plasmids. It has been determined that in prokaryotic DNA-protein complexes the protein is linked to the 5' end of the broken DNA strand, whereas in the case of eukaryotes it is linked to the 3' end. It has been found[68,69] that in all cases the DNA and topoisomerases are linked by a phosphoester bond, in the formation of which the hydroxyl group of tyrosine is involved.

The functions of topoisomerases are connected with replication, transcription, recombination and packing of virus particles (see Reviews[60-63]) and other biochemical processes in which DNA are involved.

It has been shown recently[70-72] that $\phi X174$ gene A product shows a property similar to that of topoisomerase I. It breaks the viral (+) strand of a superhelical RF DNA molecule in cistron A between nucleotides 4305 (G) and 4306 (A) and is linked to 5'-deoxyadenylic acid.[71] It is interesting to note that protein A also cleaves the single-stranded $\phi X174$ DNA at the same site as in the case of RF I.[73] It has been found recently[73-75] that the $\phi X174$ gene A* protein also shows endonuclease activity. Unlike protein A, it cleaves only single-stranded DNA at a great number of sites. After the cleavage protein A* also remains covalently bonded to the 5'-terminal nucleotide. The nature of the chemical bond between the DNA and protein A is unknown. Apparently, proteins A and A* from bacteriophage G4 also have similar properties.[76] It has been found that protein A performs a number of functions in DNA replication. By contrast, protein A* acts as an inhibitor of $\phi X174$ DNA replication[77] and is apparently necessary for DNA packing in virus particles.[74]

The discovery of topoisomerases and nucleases similar to them in microorganisms, yeast, animals and plants indicates that they occupy an important place in the biochemistry of DNA. However, in many cases neither the

functions of these proteins nor the fine structure of the DNA-protein complexes formed in the course of their action are known.

1.3. RNA-protein structure

In most of the known ribonucleoproteins the proteins are linked to the RNA by non-covalent bonds. It has been found, however, that RNA from yeast, calf pancreas, E. coli, rat liver and other sources (see References[2]) contains firmly bound peptides.

RNA-protein complexes from yeast, E. coli and a number of viruses have been studied most extensively. It has been shown[78] that in yeast the protein is apparently linked to the ribose residue of RNA by an ester bond. Neither the exact structure of the RNA-protein complex nor the functions of the covalently bonded protein have been determined.

It has been found (see References[2]) that, in RNA-peptides from pancreas and E. coli, the peptides are linked to the internucleotide phosphorus by a phosphoamide bond. The functions of the peptides are unknown.

RNA-protein complexes have recently been isolated from type I and type II[85,86] polioviruses, foot-and-mouth disease virus,[87,88] encephalomyocarditis virus,[86,89-91] cowpea mosaic virus,[92,93] caliciviruses,[94,95] tobacco ringspot virus,[96] southern bean mosaic virus,[97] pea enation mosaic viruses,[98] tomato black ring viruses,[99] potato leafroll viruses,[100] pancreatic necrosis viruses,[101] tobacco etch virus,[102] nepoviruses[103] and other viruses.[104] It has been suggested[86] that analogous complexes are also formed in rhinoviruses. Apparently, such covalent complexes are characteristic of all picornaviruses (see Reviews[104-107]). Treatment of all these RNA protein complexes with a number of denaturing agents did not cause their breakage. Pronase treatment resulted in a complete breakdown of the protein part of the complex. All this suggests that the protein is linked to the RNA by a covalent bond in all the above-mentioned viruses. Interestingly, both cowpea mosaic virus and tobacco ringspot virus have two different RNAs, each of them containing covalently bound proteins. In foot-and-mouth disease[108] and encephalomyocarditis[91] viruses, two different proteins covalently bonded to RNA have been found. In all the above-mentioned viruses the proteins were found to be linked to the 5'-terminal nucleotide residue.[87-95,104] In most cases the nucleotide is uridine-5'-monophosphate. It has been found recently that, in the case of polioviruses,[82,84] encephalomyocarditis[86,90,91] and, apparently, foot-and-mouth disease[108] viruses, the

protein is linked to the RNA by a phosphodiester bond, in the formation of which the hydroxyl group of tyrosine is involved. Apparently,[83,84,108] there is only one tyrosine residue in the protein. Recent research has shown that the protein in the RNA-protein complex from poliovirus I consists of 22 amino acids,[109] and the RNA contains 7,433 nucleotides.[110] At the 3' end of RNA there is a poly A sequence averaging 60 residues of adenylic acid.[110]

Interestingly, poliovirus mRNA has no covalently bonded protein,[79-82,111] though during replication each newly formed RNA and even the (-) virus RNA strand contains a protein. It was assumed[80,81] that, before the RNA gets into the ribosome, the protein or nucleotidyl (oligonucleotidyl)-protein should be split off the RNA-protein complex of the virion. It has been shown recently[80,81,110] that the RNA of the virion and its mRNA have the same 5'-terminal sequence. This means it is only the protein that is split off the RNA-protein complex during mRNA formation. Enzymatic activity cleaving the bond between the protein and the poliovirus RNA has been revealed recently.[112] This is a new type of processing of virus macromolecules. The presence of protein in the poliovirus genome is apparently its only difference from mRNA. Interestingly, in the case of poliovirus[113] and foot-and-mouth disease virus[114] in vitro, both mRNA and the RNA-protein complex can serve as a matrix on translation. However, mRNA is incapable of forming the virus particle. It follows that the functions of the protein linked to the RNA of viruses are connected with morphogenesis of virus particles.[80,82,83,111] It has been suggested (see Review[106]) that the protein also takes part in the initiation of poliovirus RNA replication. It has been found[82,111] that it is not required for transfection. With RNA-protein complexes from other viruses, the functions of proteins are unknown. It has been shown[93] that, in the case of cowpea mosaic virus, the protein is not necessary for infection and for translation in vitro. Conversely, with tobacco ringspot virus,[96,97] calici-viruses[94,95] and southern bean mosaic virus,[104] the protein of the complex is necessary for infection. Reovirus has recently been found to contain polyadenylic acids containing a covalently bonded protein with a molecular weight of 10,000.[115] Neither the nature of the chemical bond between the components nor the functions of such complexes have been determined.

In conclusion of the present part it may be said that covalent RNA-protein complexes are widely distributed in nature. They represent a new phenomenon in the molecular biology of viruses.

2. SYNTHESIS AND PROPERTIES OF MODEL NUCLEOTIDE-PEPTIDES

It follows from the first part of the review that covalent nucleotide- and NA-protein structures are common in nature. However, their fine structure, including the covalent bond type, has not been studied sufficiently. Model nucleotide-peptides with particular chemical bonds between the nucleotide and the peptide fragments can play an important role in solving this problem. In studying all model nucleotide-peptides, of particular interest was the question about the effect of the nature of amino acids and nucleotides directly forming the nucleotide-peptide bond, of the length and composition of peptide and oligonucleotide chains, the position of the phosphoric acid residue in the nucleotide (oligonucleotide), i.e. the functional environment of the nucleotide-peptide, on the properties and mechanism of nucleotide-peptide bond cleavage. These questions are of concern not only from the biochemical but also from the chemical point of view. The mechanism and stereochemistry of nucleophilic substitution at the tetrahedral phosphorus atom have not been studied sufficiently. Nucleotide-peptides are also interesting compounds in this respect. Having functional groups near the centre linking the individual components, nucleotide-peptides must be of interest to enzymologists, too. Enzymes speed up chemical reactions partly due to the fact that, with the help of enzymes, approach and certain orientation of the substrates takes place, while the functional groups of active centres participate directly in acid-base catalysis. Making use of model structures, in which the approach and orientation of the reacting groups takes place as a result of covalent bonding of the components containing these groups within one molecule, is a convenient approach to solving these problems. Nucleotide-peptides containing free carboxyl, hydroxyl and amino groups can also serve as such models.

The present part of the review deals with the synthesis, the study of properties of nucleotide-peptides with phosphoamide and the phosphoester bonds between the components and with their application in biochemistry.

2.1. Synthesis of nucleotide-peptides

2.1.1 Synthesis of nucleotidyl-(P→O)-amino acids (peptides) and their esters. Model nucleotide-peptides of the phosphoester type are of interest from the point of view of organic chemistry of phosphorus and from the biological standpoint. To diesters of phosphoric acid belong a number of lipids and the nucleic acids. It has been found recently (see the first part of the present review)

that in some nucleotide- and NA-protein complexes individual components are also linked by a phosphoester bond.

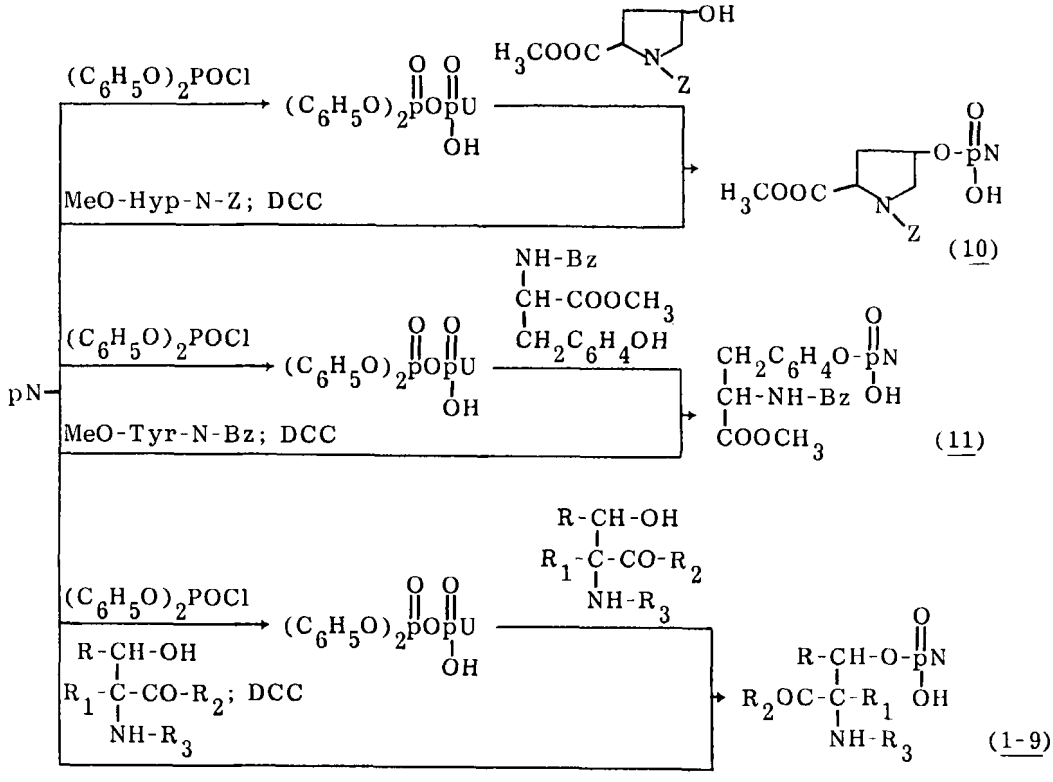
The synthesis of nucleotidyl-(P→O)-N-acyl amino acid (dipeptide) esters was carried out by the DCC and pyrophosphate methods (Scheme 1).[116-120]

In the case of the pyrophosphate method, the phosphoric acid residue of the nucleotide is activated by diphenylchlorophosphate.[121] Addition of N-acyl amino acid (dipeptide) esters with a free hydroxyl group to P^1-nucleoside-5'-P^2-diphenyl-pyrophosphate leads to the formation of nucleotidyl-(5'→O)-N-acyl amino acids (dipeptides). The yields of reaction products did not exceed 20%. Addition of pyridine or trialkylamine increased the yield up to 60%. However, the pyrophosphate method is labor-consuming and is not always reproducible. The DCC method turned out to be the best method of synthesis of nucleotidyl-(5'→O)-amino acid esters. Boiling of the reaction mixture for 1 h in abs. pyridine in the presence of DCC leads to the formation of reaction products in yields up to 80%. Esters of uridylyl-(5'→O)-N-Z-DL-serine (**1**), -DL-threonine (**2**), -2-methyl-DL-serine (**3**), -DL-seryl-glycine (**4**), -DL-threonyl-glycine (**5**), -DL-alanyl-DL-serine (**6**), N-Z-DL-serine phosphoesters of 5'-deoxythymidylic (**7**), 5'-deoxycytidylic (**8**), 5'-deoxyadenylic (**9**) acids, methyl esters of uridylyl-(5'→O)-N-Z-L-oxyproline (**10**) and -N-Bz-L-tyrosine (**11**) have been synthesized by the pyrophosphate and the DCC methods. The DCC method was also used for the synthesis of the methyl ester of deoxythymidylyl-(3'→O)-N-Z-DL-serine (**12**).

The benzyloxycarbonyl group in uridylyl-(5'→O)-N-Z-amino acid (dipeptide) esters was eliminated by reduction over platinum oxide[116] or by HBr in abs. dioxane.[117,119] The latter method is especially good and gives 80-95% yields of reaction products. Esters of uridylyl-(5'→O)-DL-serine (**13**), -DL-threonine (**14**), -DL-seryl-glycine (**15**) and -DL-threonyl-glycine (**16**) were obtained in this way.

Uridylyl-(5'→O)-N-Z-amino acids (dipeptides) with a free carboxyl group were prepared by soft saponification of the ester analogues. Optimal reaction conditions have been determined.[119,120] Uridylyl-(5'→O)-N-Z-DL-serine (**17**), -DL-threonine (**18**), -2-methyl-DL-serine (**19**), -DL-seryl-glycine (**20**), -DL-threonyl-glycine (**21**), -DL-alanyl-DL-serine (**22**) and -L-oxyproline (**23**) were obtained in this way.

The synthesis of nucleotide-peptides with a longer peptide chain was carried out by stepwise elongation of the peptide chain.[120] Here the elongation of the peptide chain may proceed from the C- or N-terminals of aminohydroxy acids. In the former case, N-Z-Ala-Ser(OpU)-Gly-OEt (**24**) was obtained from H-Ser(OpU)-



1, R = H; R_1 = H; R_2 = OCH_3; R_3 = Z; N = U
2, R = CH_3; R_1 = H; R_2 = OCH_3; R_3 = Z; N = U
3, R = H; R_1 = CH_3; R_2 = OCH_3; R_3 = Z; N = U
4, R = H, R_1 = H; R_2 = $NHCH_2COOC_2H_5$; R_3 = Z; N = U
5, R = CH_3; R_1 = H; R_2 = $NHCH_2COOC_2H_5$; R_3 = Z; N = U
6, R = H; R_1 = H; R_2 = OCH_3; R_3 = Z-Ala; N = U
7, R = H; R_1 = H; R_2 = OCH_3; R_3 = Z; N = dT
8, R = H; R_1 = H; R_2 = OCH_3; R_3 = Z; N = dC
9, R = H; R_1 = H; R_2 = OCH_3; R_3 = Z; N = dA

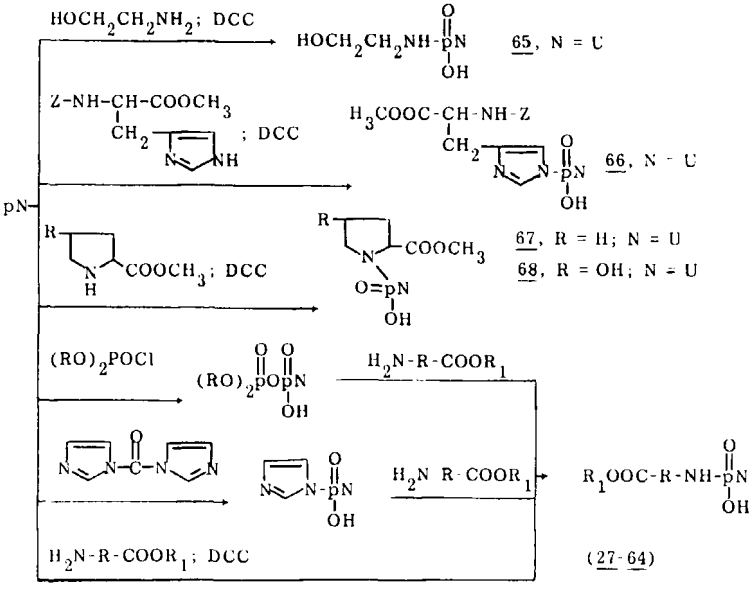
(12) dTp-OCH_2-CH(COOCH$_3$)-NH-Z

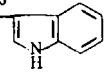
Scheme 1

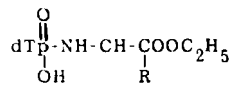
Gly-OEt (**15**) and N-Z-Ala-OH, while N-Z-Ala-Thr(OpU)-Gly-OEt (**25**) was synthesized from H-Thr(OpU)-Gly-OEt (**16**) and N-Z-Ala-OH. The yield was 80%. In the latter case, N-Z-Ser(OpU)-Gly-OEt (**4**) was obtained from N-Z-Ser(OpU)-OH and H-Gly-OEt, while N-Z-Thr(OpU)-Gly-Leu-OEt (**26**) was synthesized from N-Z-Thr(OpU)-Gly-OH (**21**) and H-Leu-OEt. The yield was 60%. In both cases DCC was used as activator of the carboxyl group, and the reaction was carried out in abs. dimethylformamide. Higher yields of reaction products and smaller losses during the preparation of original substances permit giving preference to peptide chain elongation from the C-terminal.

2.1.2. Synthesis of nucleotidyl-(5'→N)- and (3'→N)-amino acids (peptides) and their esters. The phosphite method has been used for the first time for the synthesis of nucleotidyl-(P→N)-amino acid esters.[122] However, it is a multistage method giving low yields of reaction products. At present the pyrophosphate, DCC and carbonyldiimidazole (CDI) methods are employed for the synthesis of nucleotide-peptides of the phosphoamide type (Scheme 2) (see Review[1,123-127]).

In the case of the pyrophosphate method, of diphenyl-, di (p-nitrophenyl) and di (trichloroethyl) chlorophosphates, it is more rational to use the latter, since this dialkylchlorophosphate is more stable and the yields of reaction products are higher.[125] The carbonyldiimidazole method and the DCC method in organic solvents are suitable only for readily soluble pyrimidine nucleotides.[126] The best method for the synthesis of nucleotidyl-(P→N)-amino acid (peptide) esters is the DCC method in a tert-butyl alcohol-water (3:1) mixture. On boiling the reaction mixture for 1 h the reaction goes to completion. These conditions are unsuitable for the synthesis of amino acid derivatives of purine deoxynucleotide, since partial cleavage of the glycoside bond takes place. In this way were prepared: esters of uridylyl-(5'→N)-glycine (**27**), -alanine (**28**), -valine (**29**), -leucine (**30**), -β-alanine (**31**), aspartic acid (**32**), -phenylalanine (**33**), -tyrosine (**34**), -histidine (**35**), -serine (**36**), -threonine (**37**), -glycyl-serine (**38**), -glycyl-threonine (**39**), -seryl-glycine (**40**), -alanyl-glycine (**41**), -alanyl-alanine (**42**), -alanyl-leucine (**43**), -glycyl-glycine (**44**), -glycyl-phenylalanine (**45**), -α-lysine (**46**), -tryptophan (**47**), -methionine (**48**), -α-arginine (**49**), phenylalanine derivatives of adenylic (**50**), cytidylic (**51**), guanylic (**52**) acids, histidine derivatives of adenylic (**53**) and cytidylic (**54**) acids, adenylyl-(5'→N^{α})-lysine (**55**), -arginine (**56**), deoxythymidylyl-(5'→N)-glycine (**57**), -phenylalanine (**58**), -alanine (**59**), -glycyl-phenylalanine (**60**), deoxyadenylyl-(5'→N)-phenylalanine (**61**), deoxyuridylyl-(5'→N)-phenylalanine



27, R = CH_2; R_1 = CH_3; N = U. 28, R = $CH(CH_3)$; R_1 = CH_3; N = U. 29, R = $CH[CH(CH_3)_2]$; R_1 = CH_3; N = U. 30, R = $CH[CH_2CH(CH_3)_2]$; R_1 = CH_3; N = U. 31, R = $(CH_2)_2$; R_1 = CH_3; N = U. 32, R = $CH(CH_2COOCH_3)$; R_1 = CH_3; N = U. 33, R = $CH(CH_2C_6H_5)$; R_1 = CH_3; N = U. 34, R = $CH(CH_2C_6H_4OH)$; R_1 = CH_3; N = U. 35, R = $CH(CH_2$ ); R_1 = CH_3; N = U. 36, R = $CH(CH_2OH)$; R_1 = CH_3; N = U. 37, R = $CH[CH(CH_3)OH]$; R_1 = CH_3; N = U. 38, R = $CH_2CONHCH(CH_2OH)$; R_1 = CH_3; N = U. 39, R = $CH_2CONHCH[CH(CH_3)OH]$; R_1 = CH_3; N = U. 40, R = $CH(CH_2OH)CONHCH_2$; R_1 = CH_3; N = U. 41, R = $CH(CH_3)CONHCH_2$; R_1 = CH_3; N = U. 42, R = $CH(CH_3)CONHCH(CH_3)$; R_1 = CH_3; N = U. 43, R = $CH(CH_3)CONHCH[CH_2CH(CH_3)_2]$; R_1 = CH_3; N = U. 44, R = $CH_2CONHCH_2$; R_1 = CH_3; N = U. 45, $CH_2CONHCH(CH_2C_6H_5)$; R_1 = CH_3; N = U. 46, R = $CH[(CH_2)_4NH_2]$; R_1 = CH_3; N = U. 47, R = $CH(CH_2$ ); R_1 = CH_3; N = U. 48, R = $CH(CH_2CH_2SCH_3)$; R_1 = CH_3; N = U. 49, R = $CH[(CH_2)_3NHC(NH)NH_2]$; R_1 = CH_3; N = U. 50, R = $CH(CH_2C_6H_5)$; R_1 = CH_3; N = A. 51, R = $CH(CH_2C_6H_5)$; R_1 = CH_3; N = C. 52, R = $CH(CH_2C_6H_5)$; R_1 = CH_3; N = G. 53, R = $CH(CH_2$ ); R_1 = CH_3; N = A. 54, R = $CH(CH_2$ ); R_1 = CH_3; N = C. 55, R = $CH[(CH_2)_4NH_2]$; R_1 = CH_3; N = A. 56, R = $CH[(CH_2)_3NHC(NH)NH_2]$; R_1 = CH_3; N = A. 57, R = CH_2; R_1 = CH_3; N = dT. 58, R = $CH(CH_2C_6H_5)$; R_1 = CH_3; N = dT. 59, R = $CH(CH_3)$; R_1 = CH_3; N = dT. 60, R = $CH_2CONHCH(CH_2C_6H_5)$; R_1 = CH_3; N = dT. 61, R = $CH(CH_2C_6H_5)$; R_1 = CH_3; N = dA. 62, R = $CH(CH_2C_6H_5)$; R_1 = CH_3; N = dU. 63, R = $(CH_2)_4CH(NH_2)$; R_1 = CH_3; N = U. 64, R = $(CH_2)_4CH(NH_2)$; R_1 = CH_3; N = A.



69, R = CH_3
70, R = $CH(CH_3)OH$
71, R = $CH_2C_6H_5$

Scheme 2

(**62**), uridylyl-(5'→N$^{\epsilon}$)-lysine (**63**), adenylyl-(5'→N$^{\epsilon}$)-lysine (**64**), uridylyl-(5'→N)-ethanolamine (**65**), -N-Z-histidine (**66**), -proline (**67**), -hydroxyproline (**68**), deoxythymidylyl-(3'→N)-alanine (**69**), -threonine (**70**), -phenylalanine (**71**).

Nucleotidyl-(5'→N)-peptides with the peptide fragment consisting of more than two amino acids have been synthesized by stepwise or block elongation of the peptide chain.[128] The carboxyl group of nucleotidyl-(5'→N)-amino acids (dipeptides) was activated by DCC and their reaction with amino acid (dipeptide) esters was carried out in the abs. dimethylformamide. It has been found that it is more expedient to use dipeptide derivatives of the nucleotide as the starting compound, as the α-carboxyl group of the monoamino acid analogue causes secondary processes leading to partial decomposition of the starting compound. This method was used to prepare MeO-Phe-Gly-Thr-pU (**72**) from HO-Gly-Thr-pU and H-Phe-OMe, MeO-Gly-Phe-Gly-Thr-pU (**73**) from HO-Gly-Thr-pU and H-Phe-Gly-OMe, EtO-β-Ala-Leu-pU (**74**) from HO-Leu-pU and H-β-Ala-OEt, EtO-β-Ala-Gly-pU (**75**) from HO-Gly-pU and H-β-Ala-OEt, EtO-β-Ala-Phe-Gly-pU (**76**) from HO-Phe-Gly-pU and H-β-Ala-OEt, EtO-β-Ala-Gly-Gly-pU (**77**) from HO-Gly-Gly-pU and H-β-Ala-OEt, EtO-Ala-Ala-Ala-Ala-pU (**78**) from HO-Ala-Ala-pU and H-Ala-Ala-OEt, EtO-Val-Ala-Leu-Ala-pU (**79**) from HO-Leu-Ala-pU and H-Ala-Val-OEt, and EtO-Phe-Gly-Gly-pdT (**80**) from HO-Gly-pdT and H-Gly-Phe-OEt.

Deoxyadenylyl-(3'→N$^{\alpha}$)-lysine (**81**) and deoxyadenylyl-(3'→N$^{\epsilon}$)-lysine (**82**) were synthesized from the anhydride of mesitylene carboxylic acid and nucleotide and lysine at pH 10 as described.[129] HO-Gly-pU (**83**), HO-Ala-pU (**84**), HO-Val-pU (**85**), HO-Leu-pU (**86**), HO-β-Ala-pU (**87**), HO-Phe-pU (**88**), HO-Tyr-pU (**89**), HO-His-pU (**90**), HO-Gly-Gly-pU (**91**), HO-Phe-Gly-pU (**92**), HO-Ala-Ala-pU (**93**), HO-Leu-Ala-pU (**94**), HO-Gly-pdA (**95**), HO-Phe-pA (**96**), HO-Phe-pG (**97**), HO-Phe-pC (**98**), HO-Pro-pU (**99**), HO-Hyp-pU (**100**), dTp-Phe-OH (**101**), dTp-Ala-OH (**102**), dTp-Thr-OH (**103**) were prepared by saponification of their ester analogues.[130,131]

2.1.3. Synthesis of oligonucleotidyl-(P→N)-amino acids. The synthesis of oligonucleotidyl-(P→N)-amino acids can be carried out in two ways: by linking the amino acid component to the prepared oligonucleotide or by oligonucleotide chain elongation in the mononucleotide derivative of amino acid. In the synthesis of oligonucleotidyl-(P→N)-amino acids by the former method, the phosphoric acid residue of the oligonucleotide was activated by DCC.[126,128] The reaction was carried out on boiling in a tert-butyl alcohol-water mixture. EtO-Phe-pdTpdT (**104**) and EtO-Phe-pdTpdTpdT (**105**) were obtained in 60-80% yield. The method is

not applicable for the synthesis of purine deoxyribooligonucleotide-peptides. Using the chloroanhydride of mesitylene carboxylic acid[128] or the DCC method in organic solvents[132] turned out to be the best method of activation of the terminal phosphate in purine oligonucleotides. This method was used to prepare dApdApdAp-α-Lys (**106**) and dApdApdAp-ϵ-Lys (**107**) (the yields were 12 and 40% respectively) as well as some other compounds.

In the synthesis of oligonucleotidyl-(P → N)-amino acids by stepwise elongation of the oligonucleotide chain, chemical and enzymatic methods have been employed.[128] Incubation of EtO-Phe-pdT and pdT(Ac) in the presence of triisopropyl benzosulphochloride (TPS) gives EtO-Phe-pdTpdT(Ac) (**108**) in 65% yield. Treatment of the latter compound with acid results in pdTpdT(Ac), which is subsequently used for the next step in chain elongation. EtO-Phe-pdTpdTpdT(Ac) (**109**) was obtained in 40% yield. The method of stepwise elongation of the oligonucleotide chain apparently has an advantage over direct linking of amino acids to the prepared oligonucleotide.

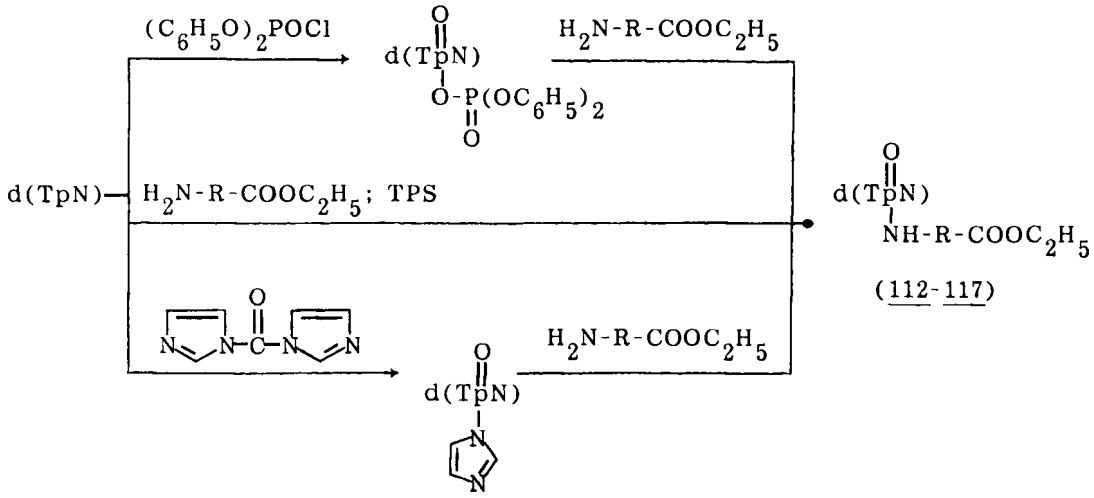
Pancreatic ribonuclease and polynucleotidephosphorylase from E. coli have been used for the synthesis of amino acid derivatives of oligoribonucleotides.[128] By means of pancreatic ribonuclease EtO-Phe-pUpC (**110**) was synthesized from EtO-Phe-pU>p and cytidine, and by means of polynucleotidephosphorylase EtO-His-pApU (**111**) was prepared from EtO-His-pA and UDP.

2.1.4. Synthesis of oligonucleotidyl-($P_m \rightarrow N$)-amino acids. For activation of internucleotide phosphate were used: diphenylchlorophosphate, TPS and carbonyldiimidazole (Scheme 3).[133-135]

The carbonyldiimidazole method appears to be the best method, as it makes possible to obtain oligonucleotidyl-($P_m \rightarrow N$)-amino acids in up to 80% yield.

The above methods were used to synthesize EtO-Phe-($P_m \rightarrow N$)-d(TpT) (112); EtO-Ser-($P_m \rightarrow N$)-d(TpT) (**113**); EtO-Ser-($P_m \rightarrow N$)-d(TpC) (**114**); EtO-Lys-N^{α}-($P_m \rightarrow N$)-d(TpT) (**115**); EtO-Lys-N^{ϵ}-($P_m \rightarrow N$)-d(TpT) (**116**) and EtO-Phe-($P_m \rightarrow N$)-d(TpA) (**117**). HO-Phe-($P_m \rightarrow N$)-d(TpT) (**118**) and HO-Phe-($P_m \rightarrow N$)-d(TpA) (**119**) were prepared by saponification of their ester analogues.[135]

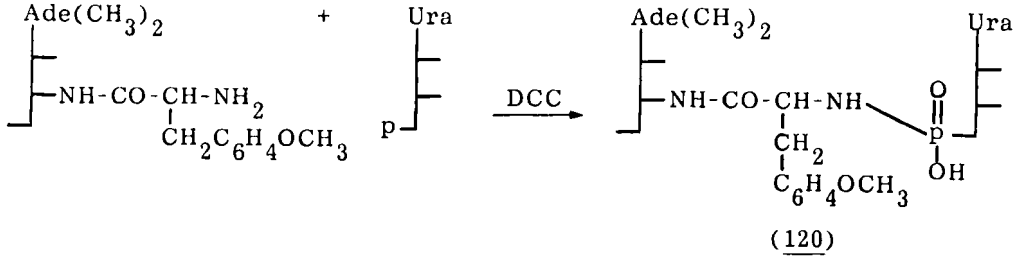
2.1.5. Modeling of protein linkers in nucleic acids. Recent research has shown[30-36] that proteins can link individual fragments of nucleic acids by covalent bonds. The type of chemical bond in such mixed biopolymers is unknown; however, they were stable in alkali. For modeling protein linkers, puromycin has



112, R = $CH(CH_2C_6H_5)$; N = dT
113, R = $CH(CH_2OH)$; N = dT
114, R = $CH(CH_2OH)$; N = dC
115, R = $CH(CH_2)_4NH_2$; N = dT
116, R = $(CH_2)_4CH(NH_2)$; N = dT
117, R = $CH(CH_2C_6H_5)$; N = dA

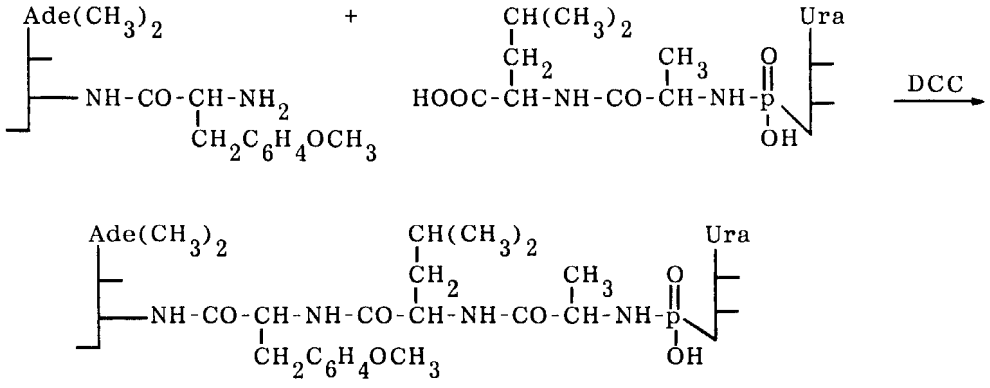
Scheme 3

been used (unpublished data by B. A. Juodka). On incubation of puromycin with UMP in a tert-butyl alcohol-water mixture in the presence of DCC, uridylyl-(5'→N)-puromycin (**120**) was formed in 75% yield (Scheme 4).



(120)

Scheme 4



(121)

Scheme 5

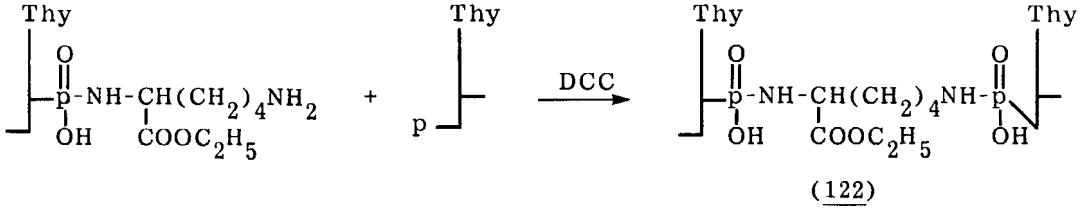
Compound **(121)**, where the linking peptide consists of three amino acids, was synthesized analogously from puromycin and HO-Leu-Ala-pU in 32% yield (Scheme 5).

The nucleotide-peptides, in which lysine and the peptide link two nucleotides by phosphoamide bonds, are also stable in alkali. The synthesis of such compounds was carried out by the DCC method in abs. DMF (unpublished data by B. A. Juodka) (Scheme 6)

The yield of compound **(122)** was up to 50%. Compound **(123)** was synthesized analogously from deoxythymidylyl-(3'→N)-alanyl-leucine and deoxythymidylyl-(5'→N$^{\epsilon}$)-lysine, ethyl ester. The yield was 60-70% (Scheme 7).

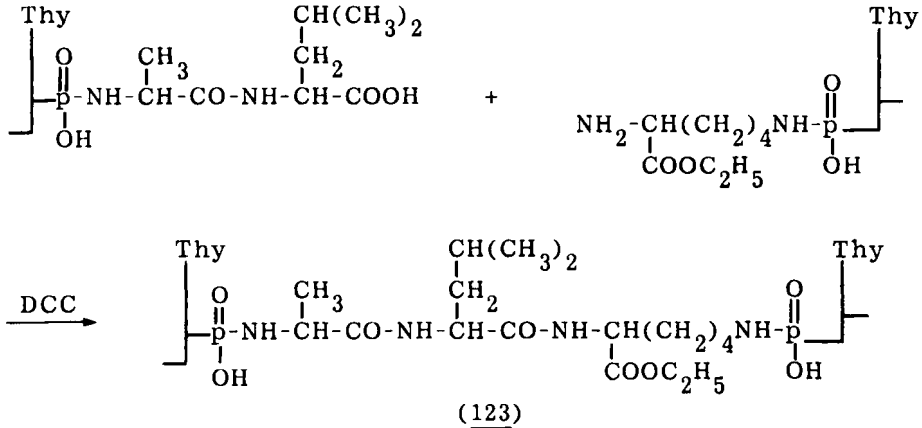
2.2 Hydrolytic stability of nucleotide-peptides

2.2.1. Hydrolytic stability of nucleotidyl-(P→O)-amino acids (peptides). Nucleotide-peptides of the phosphoester type have practically not been investigated. The problem became particularly urgent in view of the discovery of



(122)

Scheme 6



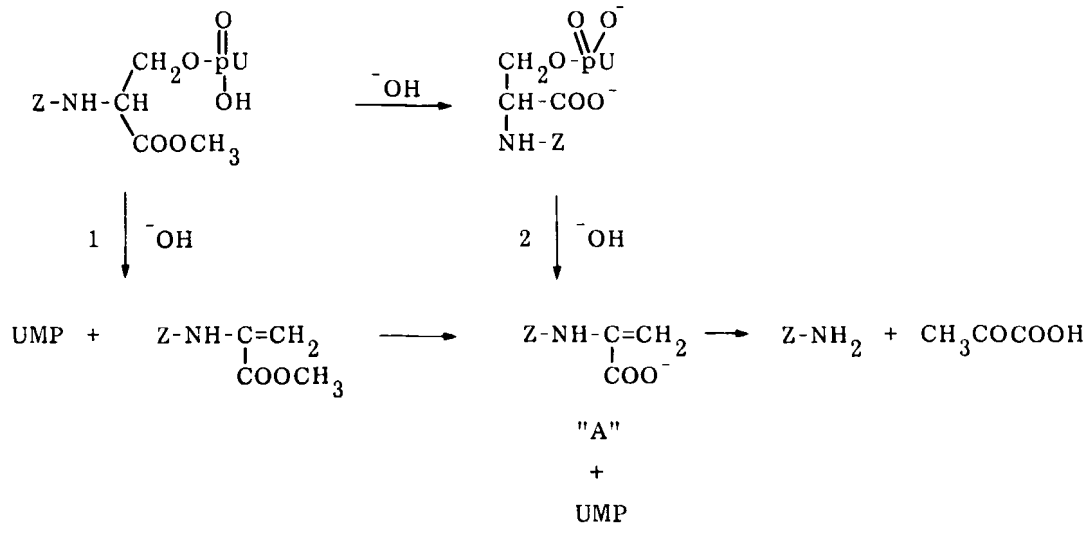
(123)

Scheme 7

the phosphoester bond in some natural nucleotide- and NA-protein structures. A study of the hydrolytic stability of some uridylyl-(5'→O)-N-acylamino acids (peptide) esters has shown that their phosphoester bond is stable in acid. The behavior of this bond in alkali depends greatly on the nature of amino acids containing the hydroxyl group. The hydroxyproline and tyrosine derivatives of uridylic acid are stable even in 2N NaOH. The serine (**1, 4, 6, 24**) and threonine (**2, 5**) analogues are readily cleaved in alkali. However, the efficiency of their cleavage depends greatly on the state of the carboxyl group of serine and threonine.

The difference in the behavior of the tyrosine and the hydroxyproline derivatives of uridylic acid in alkali as compared to that of the β-aminohydroxy acid analogues is accounted for by different mechanisms of phosphoester bond cleavage.[118-120] It has been found that in the case of the former two compounds, under rigid alkaline conditions (2N NaOH, 100°C, 1 h), phosphoester bond cleavage takes place by the mechanism of nucleophilic substitution at the tetrahedral phosphorus atom with the participation of the hydroxyl anion, whereas the serine and threonine derivatives of nucleotides are cleaved by the β-elimination mechanism (Scheme 8)

The presence of intermediate "A" (λ_{max} = 241 nm) as well as of the final product of cleavage, pyruvic acid, was evidence in favor of such a reaction mechanism. The total amount of these compounds was always equal to the amount of uridylic acid, which indicates that β-elimination is the only mechanism of cleavage of nucleotidyl-(5'→O)-N-Z-β-aminohydroxy acid (peptide) esters in



Scheme 8

alkali. The appearance of a free carboxyl group of amino acids inhibits β-elimination to a great extent. It is interesting to note that the threonine derivatives are slightly more stable than their serine analogues. Recent research has shown[119,120] that the nature of the nucleotide has no influence on the β-elimination efficiency. The presence of a blocked amino group of aminohydroxy acids is an indispensable condition of β-elimination of the nucleotide in nucleotidyl-(5'→O)-N-acylamino acids (peptides). If the amino group is free, in alkali, O→N migration of the nucleotide residue takes place.[117] Further evidence for the β-elimination mechanism was obtained on studying N-Z-Ser(Me)(OpU)-OMe (**3**) lacking hydrogen at the α-carbon atom of serine. The compound is stable in alkali and behaves analogously to the tyrosine and hydroxyproline analogues.

A study of the properties of model nucleotide-peptides of the phosphoester type suggests that if the functional groups (-COOH, $-NH_2$, etc.) find themselves in the vicinity of the phosphoester bond, conversions will take place, which may hinder the identification of bond type.

2.2.2. Hydrolytic stability of nucleotidyl(oligonucleotidyl)-(P→N)-amino acids (peptides) and their esters. It has been shown earlier[1] that nucleotidyl-(P→N)-amino acids are stable in neutral and alkaline media and are cleaved in acid. The isotope effect and linear dependence of the hydrolysis rate constant of nucleotidyl-(5'→N)-amino acids on acid concentration suggest that, in acid,

protonation of the phosphoamide centre takes place.[127] It is assumed[127] that phosphoamide nitrogen is protonated and the amino group ready to leave is formed. The reaction of hydrolysis of the phosphoamide bond in nucleotide-peptides was found to be a bimolecular first-order reaction.

2.2.2.1. Influence of the nature of amino acid and nucleotide on the hydrolytic stability of the phosphoamide bond in nucleotidyl-(5'→N)-amino acid esters. A study of methyl esters of uridylyl-(5'→N)-amino acids differing in the nature of amino acid has shown that the methyl ester of uridylyl-(5'→N_{im})-N-Z-histidine (**66**) (unpublished data by B. A. Juodka) is the most labile of them. Even on elution with water from the chromatogram it is cleaved completely to the nucleotide and its symmetric pyrophosphate. This implies that if the nucleotide or the nucleic acid is linked to the protein by such a phosphoamide bond, compounds of this type must be unstable. Nucleotide-peptides in which uridylic acid is linked to the α-amino group of various amino acids have been studied more extensively.[127] The rate of phosphoamide bond cleavage in nucleotide-peptides has been shown to be dependent on the structure of the amino acid residue. The glycine derivative of uridylic acid is the most labile, the alanine, valine, leucine, phenylalanine and tyrosine derivatives belong to the average group, and the N^{α}-histidine, N^{α}-lysine and N^{α}-arginine analogues turned out to be the most stable. It is assumed[127] that the efficiency of phosphoamide bond cleavage determines the induction and steric effects of amino acid residues. It is interesting to note that it is the compounds having in their amino acid residues functional groups which can be protonated (the imidazole ring of histidine, the ϵ-amino group of lysine and the guanidine group of arginine) that turned out to be stable. Possibly, in the present case, in addition to the induction and steric effect of the amino acid residues, there is competition between the above functional groups for protons of the medium as well as electrostatic repulsion of protons of the medium as well as electrostatic repulsion of protons from the phosphoamide centre.

In natural nucleotide- and NA-protein complexes any of the four main nucleotides can be involved in the formation of the bond with the protein. The influence of the nature of heterocyclic base on the efficiency of hydrolysis of the phosphoamide bond in methyl esters of nucleotidyl-(5'→N)-phenylalanine has been studied.[1,127] It has been shown that at pH≤1 (0.1-0.5N HCl) all the compounds studied have the same hydrolytic stability. At pH≥1.5 (0.05-0.005N HCl) some differences are observed. The guanylic and uridylic derivatives are the most

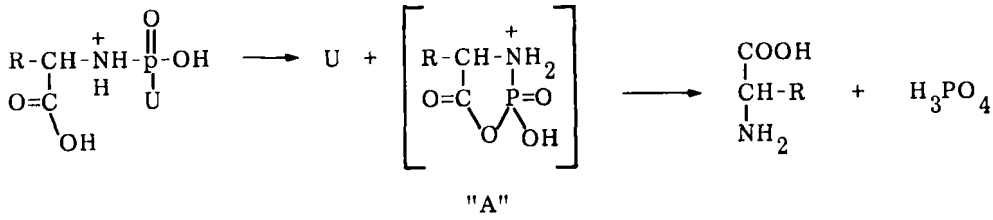
labile. The influence of the added nucleosides on the efficiency of phosphoamide bond cleavage has been studied.[127] It is assumed[127] that the influence of heterocyclic bases on the hydrolytic stability of the compounds in a weakly acidic medium manifests itself in the competition between P→N nitrogen and heterocyclic bases for protons of the medium. The appearance of the cationoid centre in the derivatives of adenylic and cytidylic acids possibly hinders the access of protons to the phosphoamide nitrogen, which accounts for the increased stability of these compounds in a weakly acidic medium.

2.2.2.2. Influence of elongation of the peptide and the oligonucleotide fragments on the efficiency of phosphoamide bond cleavage in nucleotide-peptides. Covalent NA-protein complexes have complex polypeptide and polynucleotide components. It was interesting to determine the effect of the elongation of the peptide and the nucleotide chains in the hydrolytic stability of the phosphoamide bond in model nucleotide-peptides. To elucidate these problems, the hydrolytic stability of a number of model compounds has been studied.[1,128]

It was shown that peptide and oligonucleotide chain elongation stabilizes the phosphoamide bond. The stability of the phosphoamide bond depends not only on the peptide and oligonucleotide chain length but also on its amino acid or nucleotide composition. It is assumed[128] that the stabilizing influence is accounted for by the steric effect of the peptide and oligonucleotide chains.

A study of the hydrolytic stability of adenylyl-(5'→N^{ϵ})-lysine and adenylate - DNA and RNA ligase complexes (unpublished data by B. A. Juodka), in which adenylic acid is linked to the protein through the ϵ-amino group of lysine, in acid has shown that the protein molecule stabilizes the phosphoamide bond only half as much again. This indicates that the elongation of the peptide chain in nucleotide-peptides shows the stabilizing effect only up to a certain length of the peptide chain.

2.2.2.3. Influence of the carboxyl group of amino acids on the mechanism of cleavage of the phosphoamide centre in nucleotide-peptides. It has been shown earlier that nucleotidyl-(5'→N)-amino acid esters are cleaved in acid to nucleotides and amino acid esters. A study of the analogues with a free carboxyl group has shown[130,131] that, in addition to nucleotides and amino acids, there are nucleosides and an inorganic phosphate in their acid hydrolyzates. Thus, nucleotidyl-(5'→N)-amino acids in acid are cleaved in two directions. The first direction is phosphoamide bond cleavage by the $S_N2(P)$ mechanism, which also



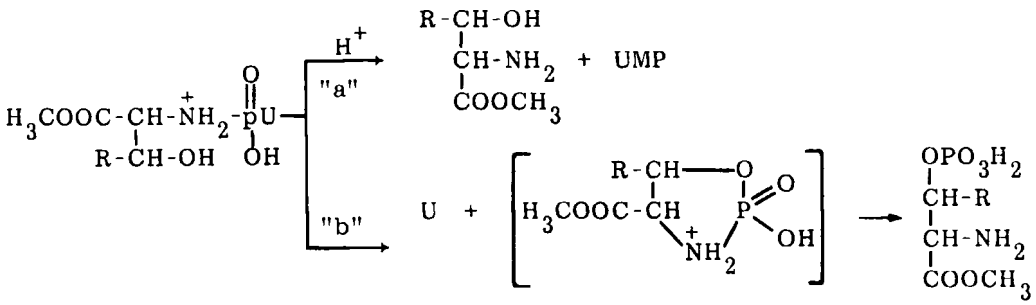
"A"

Scheme 9

took place in the case of the ester analogues. The second direction, leading to the appearance of nucleosides, amino acids and phosphoric acid, is determined by the free carboxyl group of amino acids. A study of the reaction kinetics and carrying out experiments on the phosphoamide centre in a water-ethyl alcohol mixture and in $H_2{}^{18}O$ made it possible to conclude[131] that the effect of the carboxyl group on the mechanism of cleavage of nucleotidyl-(5'→N)-amino acids manifests itself through intramolecular nucleophilic catalysis (Scheme 9). It should be noted that the α-carboxyl group of glycine, proline, hydroxyproline, as well as the remote carboxyl groups of β-alanine and the peptide analogues have no effect on the mechanism of cleavage of nucleotide-peptides, with the exception of nucleotidyl-(5'→N^{ϵ})-lysines, whose acid hydrolyzates were found to contain small amounts of nucleoside.[136] All these data indicate that the distance of the carboxyl group is not the only factor determining its intramolecular effect.

It has been found[130,131] that the efficiency of intramolecular nucleophilic catalysis with the participation of the carboxyl group of amino acids depends on the nature of amino acids and heterocyclic bases. It is possible that in the former case the amino acid residues affect the orientation of the free carboxyl group with respect to the phosphoamide centre. On the other hand, the induction effects of amino acid residues may have a direct influence on the nucleophilic capacity of the carboxyl group of amino acids and the electrophilic capacity of phosphorus.

The efficiency of intramolecular catalysis with the participation of the carboxyl group of amino acids was found to depend greatly on the position of the amino acid residue.[137] It was shown that, in acid, deoxythymidylyl-(5'→N)-phenylalanine is cleaved at the phosphoamide bond twice as readily as the 3'-analogue. A conformational study of 3'- and 5'-phenylalanine derivatives of deoxythymidylic acid by n.m.r. spectroscopy has shown (unpublished data by B. A. Juodka) that in the case of HO-Phe-pdT there is interaction between the plane of



Scheme 10

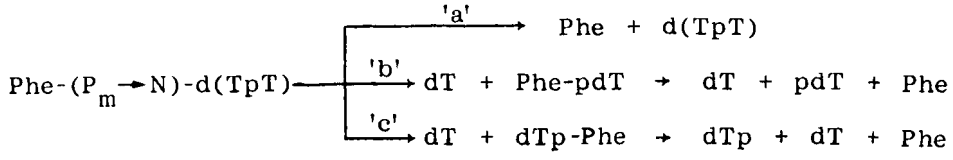
thymine and that of the phenylalanine residue. Possibly, this accounts for a convenient position of the carboxyl group at the phosphorus atom.

2.2.2.4. Influence of the hydroxyl group of aminohydroxy acids on the mechanism of cleavage of nucleotidyl-(5'→N)-amino acids (peptides). A study of the hydrolytic stability of uridylyl-(5'→N)-aminohydroxy acid (peptide) esters has shown that the hydroxyl groups of serine and threonine take part in the intramolecular catalysis in acid and in alkali.[138,139] While esters of nucleotidyl-(5'→N)-amino acids (peptides) having no functional groups in the amino acid fragment were hydrolyzed in acid only at the phosphoamide bond, the hydrolyzates of esters of uridylyl-(5'→N)-β-aminohydroxy acids (peptides) were found to contain uridine and esters of O-phosphoamino acids (peptides) besides UMP and esters of β-aminohydroxy acids (peptides). The nature of the cleavage products suggests that, in acid, the compounds investigated are cleaved simultaneously in two directions (Scheme 10).

The first direction is phosphoamide bond cleavage by the $S_N2(P)$ mechanism analogous to that for acid hydrolysis of nucleotidyl-(5'→N)-amino acids. The cleavage of the phosphoamide centre in uridylyl-(5'→N)-β-aminohydroxy acid (peptide) esters in the second direction proceeds without the participation of water at the first stage and results from intramolecular nucleophilic substitution involving the hydroxyl groups of β-aminohydroxy acids.

Uridine, amino acids (peptides) and inorganic phosphate have been found in alkaline hydrolyzates of uridylyl-(5'→N)-aminohydroxy acid (peptide) esters.[138] The cleavage of the phosphoamide centre in alkali is caused by the hydroxyl group of β-aminohydroxy acids. It is assumed that the influence of the hydroxyl group manifests itself through intramolecular nucleophilic catalysis.

The remote hydroxyl groups in esters of uridylyl-(5'→N)-tyrosine,



Scheme 11

-hydroxyproline and -glycyl-serine have no effect on the mechanism of phosphoamide centre cleavage.[138]

2.2.2.5. Oligonucleotidyl-(P_m→N)-amino acids. Oligonucleotidyl-(P_m→N)-amino acid esters are known to be far more stable than their mononucleotide analogues and to be hydrolyzed only in acid to oligonucleotides and amino acid esters.[1] In recent years the main attention has been given to studying the influence of the functional environment on the mechanism of phosphoamide centre cleavage.[133-136] A study of the hydrolytic stability of HO-Phe-(P_m→N)-d(TpN) (where N is deoxythymidine or deoxyadenosine) has shown that their acid hydrolyzates contain dT, dN, 3'- and 5'-deoxynucleotide, deoxynucleotidyl-(5'→N)- and -(3'→N)-phenylalanine, d(TpN) and phenylalanine. The reaction products suggest that three parallel processes take place (Scheme 11).

Route 'a' is hydrolysis of the phosphoamide bond, which also takes place in the case of the ester analogue. This reaction proceeds only 5-8%. The main way in acid is phosphoester bond cleavage (routes 'b' and 'c'). As a result of these reactions, deoxynucleotides and deoxynucleotidyl-(5'→N)- and -(3'→N)-phenylalanine are formed. The compounds are subsequently cleaved to nucleosides and nucleotides. The cleavage of HO-Phe-(P_m→N)-d(TpN) in the 'b' and 'c' directions is caused by the free carboxyl group of phenylalanine. It is assumed[133,135] that intramolecular nucleophilic catalysis takes place, as it was proved in the case of mononucleotidyl-(P→N)-amino acids.[131]

It has been found[134] that the hydroxyl group of serine in EtO-Ser(P_m→N)-d(TpN) (N-deoxythymidine or deoxycytidine) also has the intramolecular effect. It turned out that, in acid, N→O migration of the oligonucleotide residue takes place, and EtO-Ser(NH_2)-(P_m→O)-d(TpN) is formed from EtO-Ser(P_m→N)-d(TpN). The hydroxyl group of serine also causes the cleavage of the phosphoamide centre in alkali. HO-Ser(P_m→N)-d(TpN), d(TpN), dT, HO-Ser-pdN, ammonia and pyruvic acid have been found in alkaline hydrolyzates of EtO-Ser(P_m→N)-d(TpN). The main process in alkali is phosphoamide bond cleavage.

This is a very rare case in the chemistry of phosphoamides. It is assumed[134] that, in alkali, β-elimination takes place.

The data on the hydrolytic stability of nucleotide-peptides of the phosphoamide type suggest that the behavior of such a bond in natural nucleotide- and NA-protein complexes will depend on the nature of the N-terminal amino acid, the nucleotide, on the length and composition of the peptide and the oligonucleotide constituents of molecules, on the position of the protein component of the complex (terminal or internucleotide phosphate of NA) and the functional environment of the phosphoamide centre.

A study of the hydrolytic stability of model nucleotide-peptides of all types makes it possible to conclude that, in solving structural problems of mixed biopolymers, one cannot use their chemical treatment and, the more so, judge of the bond type between the components by the products of chemical decomposition. It follows from the studies presented above that it is necessary to cleave polymer parts of a complex by enzymes, to isolate the nucleotide-peptide unit and then to determine the bond type in it. At this level data on the reactivity of model nucleotide-peptides must be most useful. Methods of specific cleavage of particular nucleotide-peptide bonds can play a prominent role in the investigation.

2.3. Cleavage of nucleotide-peptides by enzymes.

Natural nucleotide- and NA-protein complexes can be cleaved by nucleases and proteases. The question arises how the nucleotide fragment will affect the action of proteases and how the polypeptide fragment will affect the action of nucleases. The hydrolytic activity of pronase B, carboxypeptidase A, pancreatic ribonuclease and phosphodiesterases with respect to some nucleotidyl (oligonucleotidyl)-(P→N)-peptides and their esters has been studied recently.[128] The binding of amino acids to the 3'- or 5'-terminal phosphoric acid residues of oligonucleotides was found to have no influence on the efficiency of the action of nucleases and phosphodiesterases.[128,140] The linking of nucleotide to the amino group of N-terminal amino acid of peptide does not affect the efficiency of the action of pronase B but hinders the action of carboxypeptidase A to some extent. The data obtained suggest that treatment of covalent nucleotide- and NA-protein complexes with nucleases and proteases is a mild method of cleavage of polymeric fragments of the complex and it can be used for isolation of the nucleotide-peptide unit. This approach has already found a wide application (see Review[2]).

2.4. Specific methods of cleavage of nucleotide-peptides.

In determining the bond type in nucleotide-peptides isolated after treatment of nucleotide- or NA-protein complexes with enzymes, methods of specific cleavage of particular bonds can play an important role. Hydroxylamine (pH 4.5-5) has been proposed earlier for phosphoamide bond cleavage (see Review[1]). Data on the influence of the structure of the nucleotide and peptide fragments on the efficiency of the reaction concerned and on its specificity have been obtained.[128,140] It has been shown that, of hydroxylamine, N-methyl- and O-methylhydroxylamines, N-methylhydroxylamine is the most effective. It is also evident that the elongation of the peptide and the oligonucleotide chains in nucleotide-peptides stabilizes the phosphoamide bond. It should be noted that hydroxylamine also cleaves the phosphoamide bond, in the formation of which the internucleotide phosphate is involved, but it does not cleave the phosphoester bond in nucleotidyl-(P → O)-amino acids. Phosphodiesterases turned out to be specific agents of phosphoester bond cleavage.[119,120] Snake venom phosphodiesterase cleaves the phosphoester bond in nucleotidyl-(5' → O)-amino acids, and spleen phosphodiesterase cleaves the phosphoester bond in nucleotidyl-(3' → O)-amino acids. Thus, the cleavage of nucleotide-peptides by specific phosphodiesterases not only suggests the bond type but also shows whether the protein is linked to the 3' end or to the 5' end of NA. A study of the behavior of nucleotidyl-(P → O)-amino acids (peptides) with respect to chemical agents makes it possible to determine the nature of aminohydroxy acid to which the nucleotide is linked. If the nucleotide-peptide is cleaved by diesterases, its stability in 2N NaOH (1 h, 37°C) will indicate that the nucleotide is linked through the hydroxyl group of tyrosine or hydroxyproline. The lability of nucleotide-peptide in alkali makes it possible to conclude that serine or threonine is involved in the formation of the phosphoester bond. The presence of pyruvic acid in alkaline hydrolyzates suggests that the N-terminal amino acid is serine.

3. DETERMINATION OF BOND TYPE IN THE RNA LIGASE-ADENYLATE COMPLEX

For the approbation of the applicability of all recommendations concerning ways of determining the nature of the interbiopolymer bond in natural nucleotide- and NA-protein structures and the use of data on the properties of synthetic nucleotide-peptides, the nature of the nucleotide-peptide bond in the intermediate of RNA ligase reaction has been revealed.[9] At the first stage the RNA ligase reacts with ATP to form a covalent RNA ligase-adenylate complex (see Review[8]).

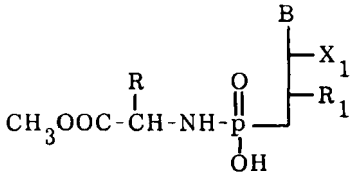
Optimal conditions of the formation of the complex have been determined.[141] The labelled [^{14}C] AMP-ligase complex was isolated by gel filtration, cleaved by proteinase K, and adenylyl-peptide was isolated. It was stable with respect to snake venom phosphodiesterase but was readily cleaved by hydroxylamine. All these findings make it possible to conclude that adenylic acid is linked to RNA ligase by a phosphoamide bond. The application of the method of modification of RNA ligase and its adenylate complex by means of 2,4-pentanedione made it possible to suggest[9] that adenylic acid is linked to the protein through the ϵ-amino group of lysine.

4. THE APPLICATION OF SYNTHETIC NUCLEOTIDE-PEPTIDES

Some synthetic nucleotidyl-(P→N)-amino acids have been employed for modeling enzymatic reactions. Adenylyl-(5'→N$^{\epsilon}$)-lysine was used for modeling three steps of the DNA ligase reaction,[142] and methyl ester of uridylyl-(5'→N$_{im}$)-N-Z-histidine was used for modeling two steps of the galactose-1-phosphate uridylyltransferase reaction (unpublished data by B. A. Juodka). The data obtained suggest that in some cases model nucleotide-peptides can help to identify intermediate enzyme-substrate complexes.

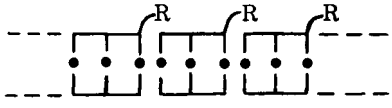
Some nucleotidyl- and oligonucleotidyl-(P→N)-lysines were immobilized on sepharose activated by cyanogen bromide.[143] Immobilization of oligonucleotides through the internucleotide phosphate is of particular interest. This principle of immobilization has the advantage over other methods that heterocyclic bases and the hydroxyl groups of monosaccharide remain free and can interact with certain parts of proteins or NA without steric hindrance.

It has been found[144] that some nucleotidyl-(5'→N)-amino acids and their esters show inhibiting activity with respect to lipoamide dehydrogenase from yeast Saccharomyces cerevisiae. pI$_{50}$ ranges from 2.5 to 3.5, depending on the structure of inhibitor. It makes the impression that the inhibiting properties of nucleotidyl-(P→N)-amino acids depend greatly on the nature of heterocyclic base and on the state of the carboxyl group of amino acid. However, no clear-cut correlation has been found between the structure of the compounds studied and their inhibiting function. MeO-Phe-pU and HO-β-Ala-pU have been studied more extensively. The former compound has a blocked carboxyl group, and the latter has a free carboxyl group. A detailed study of inhibition kinetics, its dependence on the concentration of inhibitors, substrates, activator of the lipoamide dehydrogenase reaction NAD^+ made it possible to conclude that the inhibiting

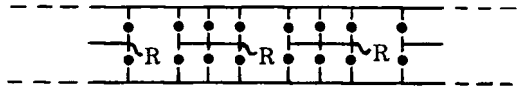


I

X_1 - H or OH;
R_1 - OH or nucleotide residue
R - amino acid residue



II



III

Scheme 12

effect of MeO-Phe-pU manifests itself at the level of complex formation between lipoamide dehydrogenase and NAD^+, whereas that of HO-β-Ala-pU manifests itself at the level of interaction of the enzyme and lipoic acid.[144] Interestingly, these compounds are quite ineffective with respect to glucose-6-phosphate dehydrogenase and alcohol dehydrogenase. This suggests that the inhibitors concerned are specific for lipoamide dehydrogenase.

The amino acid derivatives of mono- and oligonucleotides (Scheme 12) are especially convenient model systems for elucidating the elementary processes of protein-nucleic acid interaction ('single' interactions of protein and nucleic acid fragments), which are of great interest in connection with creating a general theory of protein-nucleic acid interaction and recognition.

In studying the secondary structure of these compounds or their complexes with complementary oligo(poly) nucleotide matrices in solution by physical methods, one can obtain information on noncovalent interactions of the residues of individual amino acids with certain heterocyclic bases or nucleic acid helices.

Using models (I), where R = H, containing 4 natural heterocyclic bases and $1,N^6$-ethenoadenine (εAde) and all aromatic amino acids, and employing methods of circular dichroism (CD), optical rotatory dispersion, n.m.r. spectroscopy and fluorescence, the existence of a secondary structure, stabilized by hydrophobic interaction of the aromatic systems of the amino acid and the nucleotide fragments of the molecule, has been shown.[145,146] Four conformational models

have been proposed, showing the relative position of the aromatic systems in compounds (I), which allowed to present a geometric pattern of interaction of nucleotide residues and aromatic amino acid residues.[147,148]

The main regularities of interaction of aromatic amino acids and heterocyclic bases of nucleotides remain on transition to amino acid derivatives of oligonucleotides (I), where R' is the oligonucleotide residue.[145,149-152] In amino acid (Phe, Tyr and Trp) derivatives of d(pA)n, d(pG)n (n = 2-4), $d(pC)_2$, d(pGGT) and d(pεApT) the aromatic amino acid residue interacts with the nearest base of the oligomer without disturbing the "stacking" of bases in it, the interaction being independent of oligomer chain length.[149-152] It has been shown by the example of Trp-d(pTpεA) that the possibility of interaction of the aromatic amino acid residue and the remote heterocyclic bases is not ruled out either. If the amino acid is linked to the oligomer with a weak "stacking" of bases[141,151] or bound to the internucleotide phosphate,[145] the interplanar interactions of bases are disturbed.

Important information was obtained in studying trihelical complexes (III) of amino acid (Phe, Tyr and Trp) amides of homogeneous oligonucleotides $d(pA)_n$ and $d(pG)_n$ with poly (U) and poly (C) respectively[149-151] as well as in studying bihelical (II) complexes of tryptophan derivatives of heterogeneous oligomers d(pGGT) and d(pTGG) with $d(pCCA)_2$ and $d(pACC)_3$.[151] It was found that the Phe, Tyr and Trp residues are fixed in a narrow grove of the Watson-Crick part of the trihelix and interact in it with bases without changing the geometry of the helix and without intercalation. In addition, they destabilize complexes containing (A,U).[152] The destabilizing effect of Trp without introducing it into the helix was also observed in the case of heterogeneous double-stranded complexes (II).

Thus, characteristics have been obtained concerning elementary interactions of all possible pairs of aromatic amino acids and heterocyclic bases, which form a constituent part of gradually complicated structures: mononucleotides, oligonucleotides and nucleic acid helices. These data have been successfully used for determining the position of the Tyr and Phe residues of phage f1 gene 5 protein in its complex with oligonucleotides.[153,154] In addition, they formed the basis of the hypothesis about the participation of aromatic amino acid residues in reducing T_m of DNA by phage f1 gene 5 protein and by other DNA-linking (helix-destabilizing) proteins.[153,154]

The obtained data on synthetic nucleotide-peptides suggest that they are

interesting new organic compounds, which have found direct and indirect application in a number of fields of molecular biology.

REFERENCES

1. Z. A. Shabarova, in "Prog. Nucleic Acid Res. Mol. Biol., J. N. Davidson, W. E. Cohn, eds., Vol. X, Academic Press, New York, 1970, 145-182.
2. B. A. Juodka, Bioorgan. Khimiya, **6**, 1445-1465 (1980).
3. A. I. Gaziev, Uspekhi sovr. biologiyi, **78**, 171-187 (1974).
4. S. Soderhall, T. Lindahl, FEBS Lett., **67**, 1-8 (1976).
5. I. R. Lehman, Science, **186**, 790-797 (1974).
6. J. W. Little, S. R. Zimmerman, C. K. Oshinsky, M. Gellert, Proc. Nat. Acad. Sci. USA, **58**, 2004-2011 (1967).
7. R. I. Gumport, I. R. Lehman, Proc. Nat. Acad. Sci. USA, **68**, 2559-2563 (1971).
8. R. I. Gumport, O. C. Uhlenbeck, in "Gene Amplification and Analysis", Vol. II, Analysis of Nucleic Acid Structure by Enzymatic Methods, Elsevier, North Holland, Inc., New York (1981).
9. B. A. Juodka, A. Ya. Markuckas, M. A. Snechkute, V. J. Žilinskiene, Yu. Drigin, Bioorgan. Khimiya, **6**, 1733-1734 (1980).
10. E. R. Stadtman, A. Ginsburg, in "The Enzymes", P. D. Boyer, ed., Vol. X, Academic Press, New York and London, 1974, p. 755-807.
11. B. M. Shapiro, E. R. Stadtman, J. Biol. Chem., **243**, 3769-3771 (1968).
12. J. H. Mangum, G. Magni, E. R. Stadtman, Arch. Biochem. Biophys., **158**, 514-525 (1973).
13. S. P. Adler, D. Purich, E. R. Stadtman, J. Biol. Chem., **250**, 6264-6272 (1975).
14. E. G. Niles, E. W. Westhead, Biochemistry, **12**, 1715-1722 (1973).
15. E. G. Niles, E. W. Westhead, Biochemistry, **12**, 1723-1729 (1973).
16. L.-J. Wong, P. A. Frey, J. Biol. Chem., **249**, 2322-2324 (1974).
17. D.-L. L. Yang, P. A. Frey, Biochemistry, **18**, 2980-2984 (1979).
18. M. Landt, L. G. Butler, Biochemistry, **17**, 4130-4135 (1979).
19. P. M. J. Burgers, F. Eckstein, J. Biol. Chem., **254**, 7476-7478 (1979).
20. Chr.-U. Rugevics, H. Witzel, Biochem. Biophys. Res. Commun., **106**, 72-78 (1982).
21. D. Wang, Y. Furuichi, A. J. Shatkin, Molec. & Cell. Biol., **2**, 993-1001 (1982).
22. W. E. Razzelt, H. G. Khorana, J. Biol. Chem., **236**, 1144-1149 (1961).

23. I. A. Kozlov, R. K. Ledneva, Biokhimiya, **38**, 1115-1125 (1973).
24. A. A. Krajevsky, L. L. Kiseliov, B. P. Gottikh, Mol. Biol., **7**, 769-775 (1973).
25. C. A. Chelala, L. Hirschbein, H. N. Torres, Proc. Nat. Acad. Sci. USA, **68**, 152-154 (1971).
26. W. Seifert, D. Rabussay, W. Zillig, FEBS Lett., **16**, 175-179 (1971).
27. K. J. Monty, A. L. Dounce, J. Gen. Physiol., **41**, 595-608 (1958).
28. A. Bendich, H. S. Rosenkranz, in "Nucleinovie kisloti", A. N. Beloserskii, ed., Mir, Moscow, 1965, p. 245-257.
29. Yu. F. Drigin, A. A. Bogdanov, M. A. Prokofjev, Khim. Prirod. Soed., No. 1, 218 (1966).
30. H. V. Hershey, D. Werner, Nature, **262**, 148-150 (1976).
31. W. Krauth, D. Werner, Biochim. Biophys. Acta, **564**, 390-401 (1979).
32. D. Werner, W. Krauth, H. V. Hershey, Biochim. Biophys. Acta, **608**, 243-258 (1980).
33. R. Welsh, K. Vyska, Biochim. Biophys. Acta, **655**, 291-306 (1981).
34. D. Werner, D. Hadjiolov, B. Neuer, Biochem. Biophys. Res. Commun., **100**, 1047-1054 (1981).
35. D. Werner, H.-P. Zimmermann, E. Rauterberg, J. Spalinger, Exp. Cell Res., **133**, 149-157 (1981).
36. J. Capesius, W. Krauth, D. Werner, FEBS Lett., **110**, 184-186 (1980).
37. B. Neuer, U. Plagens, D. Werner, J. Mol. Biol., **164**, 213-235 (1983).
38. H. Ikeda, J. Tomizawa, J. Mol. Biol., **14**, 110-119 (1965).
39. M. Salas, E. Vinuela, Trends Biochem. Sci., **5**, 191-193 (1980).
40. J. M. Hermoso, M. Salas, Proc. Nat. Acad. Sci. USA, **77**, 6425-6428 (1980).
41. J. Ito, F. Kawamura, S. Yanofsky, Virology, **70**, 37-51 (1976).
42. F. Kawamura, J. Ito, Virology, **83**, 233-245 (1977).
43. F. Arwert, G. Venema, J. Virology, **13**, 584-589 (1974).
44. A. C. Arnberg, F. Arwert, J. Virology, **18**, 783-784 (1976).
45. D. M. K. Rekosh, W. C. Russell, A. J. D. Bellet, A. J. Robinson, Cell, **11**, 283-295 (1977).
46. H. Ariga, H. Shimojo, S. Hidaka, K.-I. Miura, FEBS Lett., **107**, 355-358 (1979).
47. M. Shinagawa, R. Padmanabhan, Biochem. Biophys. Res. Commun., **87**, 671-678 (1979).
48. J. R. Arrand, R. J. Roberts, J. Mol. Biol., **128**, 577-594 (1979).
49. P. H. Steenbergh, J. S. Sussenbach, Gene, **6**, 307-319 (1979).
50. C. Escarmis, M. Salas, Proc. Nat. Acad. Sci. USA, **78**, 1446-1450 (1981).

51. H. Yoshikawa, T. Friedman, J. Ito, Proc. Nat. Acad. Sci. USA, **78**, 1336-1340 (1981).
52. S. V. Desiderio, T. J. Kelly, J. Mol. Biol., **145**, 319-338 (1981).
53. M. A. Penalva, M. Salas, Proc. Nat. Acad. Sci. USA, **79**, 5522-5526 (1982).
54. M. D. Challberg, T. J. Kelly, Ann. Rev. Biochem., **51**, 901-934 (1982).
55. J. Griffith, M. Dieckmann, P. J. Berg, Virology, **15**, 167-172 (1975).
56. H. Kasamatsu, M. Wu, Proc. Nat. Acad. Sci. USA, **73**, 1945-1949 (1976).
57. M. A. Lovett, D. R. Helinski, J. Biol. Chem., **250**, 8790-8795 (1975).
58. D. G. Gwiney, D. R. Helinski, J. Biol. Chem., **250**, 8796-8803 (1975).
59. G. J. Warren, A. J. Twigg, D. J. Sherratt, Nature, **274**, 259-261 (1978).
60. J. J. Champoux, Ann. Rev. Biochem., **47**, 449-479 (1978).
61. J. C. Wang, L. F. Liu, in "Molecular Genetics", Part III, J. H. Taylor, ed., Academic Press, Inc., 1979, p. 65-88.
62. N. R. Cozzarelli, Science, **207**, 953-960 (1980).
63. M. Gellert, Ann. Rev. Biochem., **50**, 879-910 (1981).
64. J. Siedlecki, W. Zimmermann, A. Weissbach, Nucl. Acids Res., **11**, 1523-1536 (1983).
65. J. V. Tricoli, D. Kowalski, Biochemistry, **22**, 2025-2031 (1983).
66. B. Prell, H.-P. Vosberg, Eur. J. Biochem., **108**, 389-398 (1980).
67. F. R. Fairfield, W. R. Bauer, M. V. Simson, J. Biol. Chem., **254**, 9352-9354 (1979).
68. Y.-C. Tse, K. Kirkegaard, J. C. Wang, J. Biol. Chem., **255**, 5560-5565 (1980).
69. J. J. Champoux, J. Biol. Chem., **256**, 4805-4809 (1981).
70. J.-E. Ikeda, A. Yudelevich, J. Hurwitz, Proc. Nat. Acad. Sci. USA, **73**, 2669-2673 (1976).
71. S. A. Langeveld, A. D. M. van Mansfeld, P. D. Baas, H. S. Jansz, G. A. van Arkel, P. J. Weisbeek, Nature, **271**, 417-420 (1978).
72. S. Eisenberg, A. Kornberg, J. Biol. Chem., **254**, 5328-5332 (1979).
73. S. A. Langeveld, A. D. M. van Mansfeld, J. M. de Winter, P. J. Weisbeek, Nucl. Acids. Res., **7**, 2177-2188 (1979).
74. L. Dubeau, D. T. Denhardt, Biochim. Biophys. Acta, **653**, 52-60 (1981).
75. A. D. M. van Mansfeld, H. A. A. M. van Taeffelen, J. Zandberg, P. D. Baas, H. S. Jansz, G. H. Veeneman, J. H. van Boom, FEBS Lett., **150**, 103-108 (1982).
76. P. Weisbeek, F. van Mansfeld, C. Kuhlemeier, G. van Arkel, S. Langeveld, Eur. J. Biochem., **114**, 501-507 (1981).
77. S. Eisenberg, R. Ascarelli, Nucl. Acids. Res., **9**, 1991-2001 (1981).

78. S. Akashi, T. Murachi, H. Ishihara, H. Goto, J. Biochem., **58**, 162-167 (1965).
79. Y. F. Lee, A. Nomoto, E. Wimmer, in "Prog. Nucleic Acid Res. Mol. Biol.", J. N. Davidson, W. E. Cohn, eds., Vol. XIX, Academic Press, New York, 1976, p. 89-96.
80. Y. F. Lee, A. Nomoto, B. M. Detjen, E. Wimmer, Proc. Nat. Acad. Sci. USA, **74**, 59-63 (1977).
81. A. Nomoto, R. Detjen, R. Pazzati, E. Wimmer, Nature, **268**, 208-213 (1977).
82. J. B. Flanegan, R. F. Pettersson, V. Ambros, M. J. Hewlet, D. Baltimore, Proc. Nat. Acad. Sci. USA, **74**, 961-965 (1977).
83. V. Ambros, D. Baltimore, J. Biol. Chem., **253**, 5263-5266 (1978).
84. P. G. Rothberg, T. J. B. Harris, A. Nomoto, E. Wimmer, Proc. Nat. Acad. Sci. USA, **75**, 4868-4872 (1978).
85. A. Babich, E. Wimmer, H. Toyoda, A. Nomoto, Intervirology, **1**, 192-199 (1980).
86. F. Golini, A. Nomoto, E. Wimmer, Virology, **89**, 112-118 (1978).
87. D. V. Sangar, D. J. Rowlands, T. J. R. Harris, F. Brown, Nature, **268**, 648-650 (1977).
88. M. J. Gruhman, Arch. Virology, **63**, 311-315 (1980).
89. D. E. Hruby, W. K. Roberts, Virology, **25**, 413-415 (1978).
90. Yu. F. Drygin, A. B. Vartapetjan, K. M. Chumakov, Mol. Biol., **13**, 777-787 (1979).
91. A. B. Vartapetjan, Yu. F. Drygin, K. M. Chumakov, A. A. Bogdanov, Nucl. Acids Res., **8**, 3729-3741 (1980).
92. S. D. Daubert, G. Bruening, R. C. Najarian, Eur. J. Biochem., **92**, 45-51 (1978).
93. J. Stanley, P. Rottier, J. W. Davies, P. Zabel, A. van Kammen, Nucl. Acids. Res., **5**, 4505-4522 (1978).
94. D. V. Sangar, D. J. Rowlands, T. J. R. Harris, F. Brown, Nature, **274**, 614-615 (1978).
95. F. L. Schaffer, D. W. Ehresmann, M. K. Pretz, M. E. Soergel, J. Gen. Virology, **47**, 215-221 (1980).
96. B. D. Harrison, H. Barker, J. Gen. Virology, **40**, 711-715 (1978).
97. A. Ghosh, R. Dasgupta, T. Salerno-Rife, T. Rutgers, P. Kaesberg, Nucl. Acids Res., **7**, 2137-2146 (1979).
98. D. Reisman, G. A. de Zoeten, J. Gen. Virology, **62**, 187-190 (1982).
99. J. Koenig, C. Fritsch, J. Gen. Virology, **60**, 343-354 (1982).

100. M. A. Mayo, H. Barker, D. J. Robinson, T. Tamada, B. D. Harrison, J. Gen. Virology, **59**, 163-168 (1982).
101. R. H. Persson, R. D. Macdonald, J. Virology, **44**, 437-443 (1982).
102. V. Hari, Virology, **112**, 391-399 (1981).
103. M. A. Mayo, H. Barker, B. D. Harrison, J. Gen. Virology, **59**, 149-162 (1982).
104. E. Wimmer, Cell, **28**, 199-201 (1982).
105. D. V. Sangar, J. Gen. Virology, **45**, 1-15 (1979).
106. N. Kitamura, C. Adler, E. Wimmer, Ann. N. Y. Acad. Sci., **354**, 183-201 (1980).
107. A. B. Vartapetjan, Bioorgan. Khimiya, **8**, 1581-1606 (1982).
108. A. M. Q. King, D. V. Sangar, T. J. R. Harris, F. Brown, J. Virology, **34**, 627-634 (1980).
109. C. J. Adler, M. Elzinga, E. Wimmer, J. Gen. Virology, **64**, 349-355 (1983).
110. N. Kitamura, B. L. Semler, P. G. Rothberg, G. R. Larsen, C. J. Adler, A. J. Dorner, E. A. Emini, R. Hanecak, J. J. Lee, S. van der Werf, C. W. Anderson, E. Wimmer, Nature, **291**, 547-553 (1981).
111. A. Nomoto, N. Kitamura, F. Golini, E. Wimmer, Proc. Nat. Acad. Sci. USA, **74**, 5345-5349 (1977).
112. V. Ambros, D. Baltimore, J. Biol. Chem., **255**, 6739-6744 (1980).
113. F. Golini, B. L. Semler, A. J. Dorner, E. Wimmer, Nature, **287**, 600-603 (1980).
114. D. V. Sangar, D. N. Black, D. J. Rowlands, T. J. R. Harris, F. Brown, J. Virology, **33**, 59-68 (1980).
115. C. A. Carter, B. Y. Lin, M. Metlay, J. Biol. Chem., **255**, 6479-6485 (1980).
116. W. E. Berkowitz, A. Bendich, Biochemistry, **4**, 1979-1983 (1965).
117. B. A. Juodka, E. P. Saveljev, Z. A. Shabarova, M. A. Prokofjev, Biokhimiya, **33**, 907-915 (1968).
118. B. A. Juodka, S. I. Sasnauskiene, Khim. Prirod. Soed., No. 2, 216-220 (1974).
119. B. Juodka, V. Kirveliene, P. Povilionis, Bioorgan. Khimiya, **8**, 326-333 (1982).
120. B. Juodka, V. Kirveliene, P. Povilionis, Nucleosides & Nucleotides, **1**, 111-125 (1982).
121. A. M. Michelson, J. Chem. Soc., 1957-1963 (1958).
122. Z. A. Shabarova, L. G. Satarova, M. A. Prokofjev, Dokl. Akad. Nauk SSSR, **123**, 864-867 (1958).
123. S. V. Zenin, B. V. Tyaglov, G. B. Sergeev, Z. A. Shabarova, E. S. Gromova, Dokl. Akad. Nauk SSSR, **207**, 861-864 (1972).

124. B. A. Juodka, S. Sasnauskiene, L. Liorančaite, Zh. Org. Khim., **9**, 1630-1634 (1973).

125. B. A. Juodka, Zh. Obshch. Khim., **44**, 2592-2593 (1974).

126. B. Juodka, L. Liorančaite, L. Janušonyte, S. Sasnauskiene, Bioorgan. Khimiya, **2**, 1513-1519 (1976).

127. B. Juodka, S. Sasnauskiene, J. Carbohydr. Nucleosides, Nucleotides, **8**, 19-39 (1981).

128. B. Juodka, V. Kirveliene, L. Liorančaite, J. Carbohydr. Nucleosides, Nucleotides, **6**, 333-357 (1979).

129. V. V. Shumyantzeva, N. J. Sokolova, Z. A. Shabarova, Nucl. Acids Res., **3**, 903-916 (1976).

130. B. A. Juodka, S. I. Sasnauskiene, S. A. Kazlauskaite, V. A. Kirveliene, Z. A. Shabarova, Bioorgan. Khimiya, **7**, 240-247 (1981).

131. B. Juodka, S. Sasnauskiene, Z. Shabarova, J. Carbohydr. Nucleosides, Nucleotides, **8**, 519-536 (1981).

132. V. D. Smirnov, M. G. Ivanovskaiya, E. V. Iljina, Z. A. Shabarova, M. A. Prokofjev, Dokl. Akad. Nauk SSSR, **206**, 1133-1136 (1972).

133. N. I. Sokolova, G. I. Gurova, L. G. Gatinskaiya, Z. A. Shabarova, M. A. Prokofjev, Mol. Biol., **3**, 837-845 (1969).

134. L. E. Liorančaite, B. A. Juodka, Nucl. Acids. Res., Symp. Ser., No. 9, 215-218 (1981).

135. B. A. Juodka, L. E. Liorančaite, S. I. Sasnauskiene, in "International Symposium on Physical Organic Chemistry", Chem. Inst. of Canada, 1979, p. 18.

136. O. E. Vorobjev, Z. A. Shabarova, M. A. Prokofjev, Dokl. Akad. Nauk SSSR, **190**, 842-845 (1970).

137. B. A. Juodka, L. E. Liorančaite, V. P. Baltenas, Khim. Prirod. Soed., No. 6, 740-746 (1982).

138. B. A. Juodka, I. V. Obruchnikov, V. K. Nedby, Z. A. Shabarova, M. A. Prokofjev, Biokhimiya, **34**, 647-654 (1969).

139. E. P. Saveljev, E. S. Gromova, Z. A. Shabarova, M. A. Prokofjev, Izvestiya Akad. Nauk SSSR, **8**, 1817-1822 (1970).

140. B. A. Juodka, L. E. Liorančaite, V. A. Kirveliene, R. S. Ferencaite, Bioorgan. Khimiya, **5**, 1623-1632 (1979).

141. B. A. Juodka, M. A. Sniečkute, L. M. Janušonyte, A. J. Markuckas, Biokhimiya, **44**, 599-604 (1979).

142. B. A. Juodka, S. I. Sasnauskiene, V. J. Meškenaite, K. V. Kadziauskiene, Zh. Obshch. Khim., **46**, 585-590 (1976).

143. B. Juodka, L. Liorančaite, R. Leimontaite, A. Malinauskas, J. Kulys, N. I. Sokolova, Z. A. Shabarova, Khim. Prirod. Soed., No. 3, 435-436 (1977).

144. A. Glemzha, B. Juodka, Nucleosides & Nucleotides, **2**, 33-40 (1983).

145. E. S. Gromova, B. V. Tyaglov, Z. A. Shabarova, Biochim. Biophys. Acta, **240**, 1-11 (1971).

146. B. V. Tyaglov, S. V. Zenin, E. S. Gromova, G. B. Sergeev, Z. A. Shabarova, M. A. Prokofjev, Mol. Biol., **8**, 595-602 (1974).

147. B. V. Tyaglov, E. S. Gromova, S. V. Zenin, G. B. Sergeev, Z. A. Shabarova, M. A. Prokofjev, Mol. Biol., **9**, 652-666 (1975).

148. B. V. Tyaglov, S. V. Zenin, E. S. Gromova, G. B. Sergeev, Z. A. Shabarova, Mol. Biol., **10**, 347-359 (1976).

149. V. D. Smirnov, E. S. Gromova, A. P. Razjivin, Z. A. Shabarova, FEBS Lett., **51**, 211-214 (1975).

150. E. S. Gromova, N. G. Dolinaiya, V. D. Smirnov, Z. A. Shabarova, M. A. Prokofjev, Bioorgan. Khimiya, **1**, 1716-1727 (1975).

151. N. G. Dolinaiya, E. S. Gromova, A. P. Razjivin, Z. A. Shabarova, M. A. Prokofjev, Dokl. Akad. Nauk. SSSR, **239**, 603-606 (1978).

152. V. D. Smirnov, E. S. Gromova, N. N. Nefedjeva, Z. A. Shabarova, M. A. Prokofjev, Mol. Biol., **12**, 429-442 (1978).

153. N. N. Veiko, E. S. Gromova, A. B. Sigalov, Z. A. Shabarova, Mol. Biol., **15**, 1385-1396 (1981).

154. N. N. Veiko, E. S. Gromova, Z. A. Shabarova, Mol. Biol., **13**, 1136-1146 (1979).

Received February 13, 1984.