STRUCTURE AND BIOSYNTHESIS OF TRANSFERRIN

T. A. Salikhov

UDC 547.963.3

The review gives a literature analysis of the state of the question of the structure and biosynthesis of transferrin with the results of modern investigations.

Transferrin (TF) is the main carrier of iron in the vertebrate organism and is an iron donor for such vitally important hemoproteins as hemoglobin, myoglobin, cytochrome, catalase, etc. Genetic anomalies of the structure and function of TF may play an important role in the pathogenesis of hereditary diseases. Thus, in hemochromatosis a quantitative insufficiency of TF and a disturbance in the capacity of this protein for binding with iron ions is shown. It is assumed that the primary effect in this monogenic mutation of man is a mutational change in the structure of the transferrin-coding gene, although no direct proofs of this hypothesis have yet been obtained. Other, fairly rare, forms of inadequacy of the expression of the TF gene revealed phenotypically as a atransferrinemia (a marked decrease in the concentration of transferrin in the blood plasma, going as far as its complete absence) have also been described in the literature. The molecular defect lying at the basis of this form of hereditary pathology has not yet been found.

The advances in the molecular genetics of higher organisms connected primarily with the development of the methods of gene engineering and nucleic acid chemistry form the foundation of rapid progress in the study of the molecular nature of monogenic mutations of man and, in particular, those forms of hereditary diseases that are manifested in the form of an inadequacy of the expression of individual genes (thalassemia, hepatolenticular degeneration, some forms of sugar diabetes, immunodeficient states, etc.). Hereditary anomalies of the synthesis and structure of TF can also, apparently, be assigned to this group of diseases. Consequently, progress in their study must be connected above all with the existence of information on the structure of the normal gene of mammals, the sequence of the stages of expression of the TF gene, and the molecular organization and functional activity of transferrin mRNA (TF mRNA).

In the present review, literature information is generalized on the primary structure of TF, features of its biosynthesis and regulation, and the molecular organization of the chromosomal transferrin gene, and our results on the biosynthesis of TF in mammals, on the structural and functional features of the TF mRNA from rat liver, and the cloning of reverse transcripts of TF mRNA in bacterial vector systems are given.

Primary Structure of Transferrin and the Principles of Its Domain Organization

The primary structure of TF has been studied for more than 10 years, but only at the end of 1982 did the complete amino acid sequence of avian and human TFs become known. The primary sequence of ovotransferrin (ovotransferrin and conalbumin are avian transferrins) was determined independently of one another by Williams et al. [1] and by Chambon [2]. It is interesting to note that these authors used completely different approaches in the study of the structure of ovotransferrin. Williams' group studied the sequence of amino acid residues in the protein molecule directly, while Chambon investigated the primary structure of ovotransferrin mRNA, and then deduced the primary structure of this protein on the basis of the rules of the genetic code. The results of the investigation mentioned above coincided completely and showed the following features of the structure of the molecules of avian TFs. The ovotransferrin molecule consists of 686 amino acid residues. The accurate molecular weight of the protein taking is carbohydrate component into account is 77,770. Great interest is caused by the fact that a high degree of homology has been found within the TF molecule. The N-terminal and C-terminal sequences located at amino acid residues 1-228 and 334-622, respectively, proved to be homologous. These sequences have been called the N- and C-domains,

Institute of Bioorganic Chemistry, Academy of Sciences of the Uzbek SSR, Tashkent. Translated from Khimiya Prirodnykh Soedinenii, No. 3, pp. 291-305, May-June, 1985. Original article submitted January 11, 1984.

0009-3130/85/2103-0271$09.50

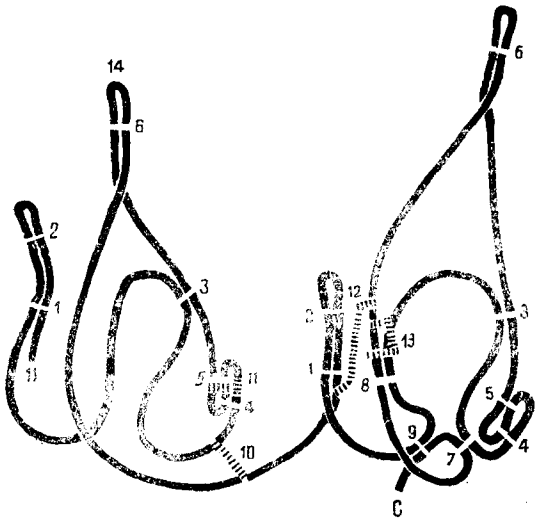

Fig. 1. Domain structure of avian and human transferrins [1, 7]. On the left, the N-domain; on the right, the C-domain: 1-6) homologous disulfide bonds of the N- and the C-domains; 7-9) additional disulfide bonds of the C-domain; 10-13) disulfide bonds specific for human TF; 14) specific cysteine residue in human lactoferrin.

since in the study of TF crystals it was found that the molecule of this protein consists of two globular structures (domains) [3, 4]. Furthermore, these domains can be isolated after partial trypsin hydrolysis [5]. The N- and C-domains have a 50% homology. Homology is also observed within each domain, although to a smaller degree. We have also detected N- and C-domains in rat TF [6, 15]. The limited proteolysis of rat TF molecules gave the N- and C-terminal fragments of the TF with molecular weights of 38 and 36 kD, respectively. With respect to their amino acid compositions, and immunochemical properties, these fragments possess a high degree of homology, i.e., they may be considered as domains. Williams ascribes great importance to the disulfide bonds in the TF molecule. He has proposed a structure of the ovotransferrin molecule (Fig. 1). As can be seen from Fig. 1, ovotransferrin consists of two globular domains. The N-domain has six disulfide bonds (1-6) and the C-domain has, in addition to these, another three such bonds (7-9).

The complete amino acid sequence of human TF is astonishingly similar to the sequence of ovotransferrin. Human TF contains 678 amino acid residues and its molecular weight with the carbohydrate moiety is 79,500. Likewise, two homologous domains are observed in positions 1-336 and 337-678. As compared with ovotransferrin, human TF has four more disulfide bonds (see Fig. 1). These bonds are located mainly in the region connecting the N- and C-domains [5, 8].

At the present time, the almost complete primary structure of human lactoferrin has been studied; here, again, a high degree of homology between the N- and C-domains is observed [9, 10]. In lactoferrin the same disulfide bonds have been detected as in human serum transferrin. It must be mentioned that the high degree of homology of TF was noticed long before its primary structure was established [11]. On the whole, the primary structures of the TFs are characterized by a high degree of homology. On analyzing information on the amino acid sequence of avian and human TFs, Williams put forward a hypothesis of the evolution of transferrin (Fig. 2), according to which the precursor of vertebrate transferrin was a protein X, which has been detected in chordates 500 million years old. It had a molecular weight of 40,000, one iron-binding center, and six disulfide bonds. In the process of evolution, the chain of protein X doubled and then underwent divergence into the N- and C-domains, which were subsequently converted into two types of domains, N' and N", and C' and C".

Williams ascribes a special position to the precursor X. His experimental results show that protein X had small dimensions for existence in the free state in the bloodstream. It is known that proteins with molecular weights smaller than 60,000 are rapidly eliminated from the blood into the urine. In order to test the possibility of the existence of protein X in the bloodstream in the free state, Williams and his coworkers isolated the N-domain in the

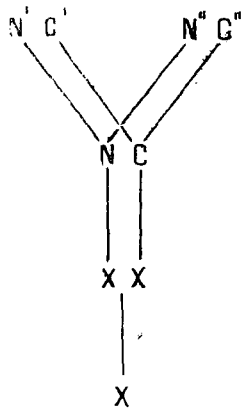

Fig. 2. Scheme of evolution of the domain of transferrin [7]: X: protein precursor; N and C: corresponding N- and C-domains of transferrin.

free form and introduced it into the bloodstream [12]. It was found that the N-domain was rapidly excreted through the kidneys into the urine. Williams assumes that X was not a serum protein but most likely existed in membrane-bound form. It is interesting to note that a membrane-bound tumor antigen known as protein P97 has been found which reveals itself in a melanoma [13].

According to Brown et al. [13], the P97 antigen has a molecular weight of 97 kD and possesses numerous properties characteristic for the TF molecule. In particular, antigen P97 effectively binds Fe^{3+} ions and has immunochemical cross-reactivity with TF. In a later investigation, Brown et al. determined the human chromosome in which the gene for protein P97 was localized. They obtained mouse–man somatic hybrids having different sets of human chromosomes. The expression of the P97 protein was determined with the aid of monoclonal antibodies to P97. In a comparison of the expression of P97 in hybrids with different sets of human chromosomes it was found that the genes for the P97 protein was located in chromosome 3. At the same time, it is known that human TF is likewise coded by chromosome 3 [60]. Brown et al. assumed the possibility that TF and P97 are coded by the same gene. Since, at least in birds, the chromosomal TF gene contains a number of introns, in such a case the differences between TF and P97 can be explained by different mechanisms of the splicing of the primary product of the transcription of a single chromosomal gene.

Returning to Williams' evolutionary hypothesis, we may note that the chromosomal gene of the P97 antigen is probably the product of the evolution of a precursor. However, unlike TF, it has not lost its capacity for binding with membranes.

In Williams' hypothesis, attention is attracted by the fact that in the course of evolution, as the TF molecule became more complicated, additional disulfide bonds appeared. Williams considers that the number of S–S bonds can serve as a marker characteristic for the TFs of different animals [12]. On the whole Williams' hypothesis is of great interest, although a definitive answer to the question of the correctness of this hypothesis can be given only after primary structures of all TFs become known.

Features of the Biosynthesis of Transferrin

Transferrin is a blood plasma protein. Its concentration in human blood is 2.3 mg/ml, and in the rat this level reaches 4.8 mg/ml [16]. As with many serum proteins, TF is synthesized in the liver. This has been shown by numerous experiments, among which the work of Morgan on rats must be mentioned, where the site of the synthesis was determined from the inclusion of [^{14}C]leucine in this protein [17]. Twenty minutes after the introduction of the label radioactivity was detected only in the liver TF while after 2 h ^{14}C-radioactive TF appeared in other organs, as well. Small amounts of TF are also synthesized in the spleen, kidneys, lungs, and brain. In birds, in addition to the liver, TF (conalbumin) is synthesized in the oviduct.

MacKnight et al. have studied the distribution of the sequences of TF RNA in various tissues of birds by hybridizing the RNA with transferrin cDNA [18, 19]. It has been shown that only the liver and the oviduct contain a substantial number of TF mRNA molecules per cell (700 and 65 molecules, respectively). A certain amount of TF mRNA (45 molecules per cell) is present in the brain. The authors found no TF mRNA sequences in other organs. Transferrin belongs to those proteins, which, after their synthesis, are exported from the cells into the bloodstream. At the present time, an enormous number of publications devoted to the biosynthesis and post-translational maturation of export proteins exist. On analyzing these publications it is possible to see the following features of the biosynthesis of the secretory proteins [20]. A long reaction pathway exists which leads from the growing polypeptide chain to the mature protein. It involves the reactions of limited proteolysis, glycosylation, acetylation, phosphorylation, etc. Furthermore, the newly formed proteins pass through a complicated pathway of intracellular transport.

Together with albumin, transferrin forms a convenient model for the detailed study of the mechanisms of the biosynthesis and posttranslational modification of secretory proteins in the liver. In 1971 a paper appeared by Morgan and Peters who had studied the biosynthesis and secretion of these basic serum proteins [16]. [^{14}C]Leucine was administered to rats and the inclusion of the label in the TF and the albumin of various fractions of the cells and serum was measured. The results showed that the albumin and the TF were bound to endoplasmatic membranes during the process of synthesis. However, the secretion of TF took place more slowly than that of albumin. Transferrin appeared in the blood 30 min, and albumin 16 min, after the administration of the label. The transit times of the synthesized TF and albumin were 80 and 23 min, respectively. During these periods the rat liver synthesized 173 μg of TF and 270 μg of albumin. The authors assumed that the process of secreting TF is accompanied by reactions of limited proteolysis and glycosylation. The conclusion was drawn that albumin and TF have a secretory route: rough-surfaced reticulum—smooth-surfaced reticulum—Golgi apparatus—bloodstream. A more detailed investigation of the biosynthesis and secretion of TF in the rat was made by Schreiber et al. [21] and by Gaitskhoki et al., [22]. Schreiber et al. showed that intracellular TF is represented by three molecular forms which differ in their sialic acid contents, the dynamics of the appearance of these forms in different subcellular fractions being different. Furthermore, in this investigation the influence of the protease inhibitor N-α-tosyl-L-lysylchloromethane and of the glycosylation inhibitor tunicamycin on the secretion of TF was considered. It was found that the protease inhibitor suppressed the secretion of TF into the bloodstream almost completely, and the glycosylation inhibitor did not affect the secretion of TF. These facts permitted the conclusion that protolytic maturation was necessary for the secretion of TF while glycosylation was not so necessary for this process.

The transferrin-synthesizing polyribsosomes of the rat have been studied [21]. As was found, the biosynthesis of TF takes place in polysomes associated with the membranes of the endoplasmatic reticulum. In a determination of the dimensions of the TF-synthesizing polysomes it was found that TF mRNA is translated simultaneously by 11-13 ribosomes, and the addition of the elongation inhibitor cycloheximide led to an increase in the load to 16 ribosomes per mRNA molecule. The biosynthesis and posttranslational maturation of TF in the subcellular fractions of the rat liver have also been investigated [22]. After synthesis, transferrin migrates from the membranes of the rough-surfaced reticulum to the membranes of the smooth-surfaced reticulum and into the Golgi apparatus, and it is then secreted into the bloodstream. A retarded secretion of TF into the bloodstream was also established in later investigations, in which a prolonged retention of the TF in the Golgi apparatus was observed.

A determination of the molecular weights of the TF in the subcellular fractions permitted the detection of a high-molecular-weight TF precursor. Thus, in the membranes of the smooth reticulum a TF was detected with a molecular weight of 88 kD, which is substantially higher than the molecular weight of the mature TF molecule (78 kD). It was concluded that in the process of maturation of the TF molecule not only a signal peptide but also other superfluous sequences localized at the C-end are split off, since, according to Gaitskhoki et al. [22] the amino acid sequence of the mature TF follows after the signal peptide. This conclusion agrees will with the results of Schreiber et al. according to which a protease inhibitor completely suppresses the secretion of TF [21]. Later [24], Schreiber et al. investigated features of the biosynthesis of a number of secretory proteins, including TF, in the human liver. They showed that the acidic α_1-glycoprotein, α_1-antitrypsin, and albumin are intracellular precursors of TF. It was established that two forms of the TF exist in the liver which have

the same molecular weight (76,500) and the same N-end: Val-Pro-Asp-Lys-Try-Val... They differ by the fact that one form, like the TF from serum, contains in its C-terminal region 4.4 sialic acid residues per protein molecule, while the other does not contain them at all. Differences in the sialic acid content have also been found for intracellular forms of other secretory proteins. In view of the fact that amino acid sequence at the N-terminal end of the intracellular forms and of serum TF are the same, the authors came to the conclusion that TF undergoes intracellular processing from the C-end. Summarizing the facts presented above, it can be stated that rat and human TFs are synthesized in the form of a high-molecular-weight precursor, and that the maturation of the TFs has a complex nature.

It is known that secretory proteins are synthesized on polyribosomes bound to the membranes of the endoplasmatic reticulum. For a long time it was not known how such a topography of the biosynthesis of proteins is programmed. Having analyzed information on the amino acid sequence of the superfluous peptides that appear under the conditions of cell-less translation, Blobel [25] formulated a "signal" hypothesis explaining the specificity of the synthesis of the secretory protiens in the cell and the mechanism of their vectorial transmembrane passage in the course of synthesis.

According to this hypothesis, the translation of the mRNA begins on the free polyribosomes, but the NH_2 end of the growing chain acquires hydrophobic properties and is bound to the receptor proteins in the composition of the membranes. A channel is formed in the membrane through which the polypeptide chain passes to the systems of the endoplasmatic reticulum. After this, the "signal" peptide is split off by "signalase." The "signal" hypothesis has been confirmed by a number of experiments. "Signal" peptides have been found for all secretory proteins [20, 26-29]. On the translation of the mRNA of rat TF in a cell-free system from wheat embryo, a product was found with an excess of 20 amino acid residues [21]. A "signal" sequence has also been found for avian TF [30].

Various aspects of the biosynthesis of avian TF and the expression of the gene of this protein have been most studied at the present time. This is due to the fact that the conalbumin (transferrin) of the oviduct and the serum TF in birds are the products of the expression of one and the same gene, although the mechanisms of the TF-synthesizing reaction in these two organs differ. Conalbumin and TF differ only in respect to the carbohydrate component. Conalbumin contains mannose and glucosamine, and TF contains mannose, glucosamine, two glactose residues and one or two sialic acid residues per molecule of protein. No differences exist between TF and conalbumin with respect to other criteria, including immunoelectrophoresis, iron-binding capacity, and amino acid composition [31].

The first information on the different mechanisms of the regulation of the expression of the TF gene in the oviduct and in the liver appeared in 1978 [31]. The action of an estrogen and of progesterone on the expression of the transferrin gene was studied. A complete coincidence of the curves of the hybridization of conalbumin cDNA with the total RNA fractions from the liver and from the oviduct was shown, which indicated the presence of one and the same form of RNA in the two tissues. This conclusion was also confirmed by the melting of cDNA–RNA hybrids. The melting curves were also identical for the hybrids of cDNA with the RNAs of the two tissues.

These results demonstrate the complete homology of the two mRNAs. Subsequently, the authors determined and compared the molecular weights of the mRNAs of conalbumin nd TF. The results of electrophoresis in 2.5% polyacrylamide gel showed that the dimensions of the two mRNAs were the same and corresponded to a molecular weight of 1,033,000, or 3200 nucleotides. Thus, the complete identity of the transferrin and conalbumin mRNAs was established. Furthermore, a later investigation also showed that both proteins were synthesized as precursors that had the same signal sequence [30]. The TF mRNA from avian liver and conalbumin mRNA from the oviduct were translated in a cell-free system in the presence of labeled amino acids. The labeled conalbumin and labeled TF were precipitated with antibodies and were studied by automatic Edman stepwise degradation from the N-end. It was found that both proteins were synthesized with a superfluous "signal" sequence consisting of 19 amino acids: Met-Lys-Leu-Ile-Leu-Cys-Try-Val-Lau-Ser-Leu-Gly-Ile-Ala-Ala-Val-Cys-Phe-Ala-. The "signal" peptide attracts attention by virtue of the fact that the majority of its amino acid residues are hydrophobic, as is also the case in the "signal" peptides of ovomucoid [32] and of lysozyme [33].

The avian oviduct is an organ that is sensitive to the action of steroid hormones such as estradiol, progesterone, and testosterone. In response to the introduction of estradiol into the oviduct, the synthesis of the main proteins – ovalbumin, ovomucoid, conalbumin, and

lysozyme – increases substantially, while the elimination of the hormone lowers the biosynthesis of these proteins [34, 35]. In order to reveal the intimate mechanisms of the regulation of the biosynthesis of proteins in the oviduct by estrogens, Lee et al. [31] determined the number of conalbumin mRNA molecules in the oviduct at various stages of induction. The synthesis of TF mRNA in the liver was followed in parallel. In the norm, the amount of TF mRNA per cell amounts to 500-600 molecules. For the conalbumin mRNA this magnitude is somewhat lower. Hormonal induction by an estrogen did not appreciably affect the synthesis of TF in the liver. The stimulation of the biosynthesis of this protein was insignificant: from 1.7% of the total synthesis to 2.5%, while the concentration of the TF mRNA molecules rose to 1600. Progesterone had no effect whatever on the synthesis of TF in the liver. So far as concerns conalbumin, steroid hormones sharply increased its synthesis, the amount of the protein rising approximately 10-fold under the action of progesterone and 50-fold under the action of an estrogen. The concentration of the conalbumin and mRNA rose in proportion to this. Thus, information has been presented which shows different mechanisms of the regulation of the expression of the transferrin gene in the avian oviduct and liver. These results have been confirmed by other authors [36, 37].

Subsequently, MacKnight et al. studied the influence of an iron-deficient diet on the expression of the transferrin gene in the oviduct and the liver [18, 19]. Chicks were kept for three weeks on a diet with a very low iron content nd the biosynthesis of TF and the concentration of TF mRNA were measured at various times. Under these conditions the synthesis of TF in the liver increased by a factor of 3-4. The increase in the biosynthesis of TF was due to a rise in the concentration of TF mRNA to 3500 molecules per cell. The efficacy of translation did not change, and the TF mRNA was associated with ribosomes identically in the control rats and in the rats with the iron deficiency. This shown that the regulation of the biosynthesis of TF in the liver is effected at the transcription level.

In the oviduct, a deficiency of iron caused no changes either in the biosynthesis of of conalbumin or in the concentration of the conalbumin mRNA. In the investigations mentioned above, it was shown unambiguously that the biosynthesis of conalbumin in the oviduct under the action of hormones and the biosynthesis of TF in the liver under the action of a deficiency of iron rise as a result of an increase in the concentration of mRNA. At the same time, the accumulation of TF mRNA is a consequence of the activation of its synthesis and is not connected with an increase in the stability of the mRNA [18, 19, 34].

Since the conalbumin of the oviduct and the TF of the liver are the products of a single gene, the differences in the expression of the TF gene under the action of various agents could be explained by the assumption that the tissue-specific regulation of the transferrin gene is its property of promoting the transcription of the gene by RNA polymerase in different ways. In such a case, the fundamental question of the architecture of the gene and its flanking sequences arises.

Another explanation of the differences in the expression of the TF gene consists in the assumption that the nuclei of the oviduct differ by a greater sensitivity to the action of hormones as the result of the presence of a large number of hormonal receptors. Palmiter et al. have studied the relationship of the progesterone receptors to the induction of conalbumin mRNA in the oviduct [38]. In this investigation they showed that 30 min after the injection of the optimum dose of progesterone the number of nuclear receptors for the hormone had increased 5 times (to 1500 molecules per cell) and it remained at the same level for 8 hours. The accumulation of the conalbumin mRNA was directly proportional to the increase in the number of receptors. In other words, 1/2 of the maximum level of conalbumin mRNA was found when the number of receptors amounts to half the maximum level. For the conalbumin mRNA a lag in the synthesis of RNA was observed, 1/2 of the maximum level being reached at a level of the nuclear receptors of about 80% of the maximum. This lag was explained by the authors in terms of intermediate reactions between the receptors and the activation of transcription. Palmiter's results can be explained to some degree by differences in the expression of the TF gene in the liver and in the oviduct. Nevertheless, this does not exclude differences in the nucleotide sequences of the 5'-terminal regions of the transferrin gene.

Mammalian TF mRNA has been studied to a smaller degree than ovian TF mRNA. In this connection it is possible to refer to investigations by N. A. Timchenko et al. They isolated and characterized rat liver TF mRNA [39, 40]. According to their results, rat TF mRNA has a molecular weight of 0.925 MD, which corresponds to ∿2800 nucleotides. This size of the TF mRNA agrees well with the size of avian TF mRNA [2, 18, 19]. In the translation of rat

TF mRNA in a lysate from rabbit reticulocytes it was established that the TF mRNA programs the synthesis of TF precursors with a molecular weight of 82 kD. A study of the structural organization of rat TF mRNA has shown that it contains a 3'-terminal poly(A) sequence heterogeneous with respect to the length. At the 5'-end of the TF mRNA molecule there is a structure of the cap type playing an important role in the translation of transferrin mRNA [40]. In experiments on the kinetics of the hybridization of TF cDNA with individual TF mRNA and the total polysomal RNA it was found that the steady-state concentration of TF mRNA in hepatocytes was ~6000 molecules per cell. This is substantially greater than the concentration of TF mRNA in the avian liver.

Molecular Organization of the Avian Transferrin Gene and the Regulation of Its Expression

The gene of avian conalbumin (transferrin) is one of the individual genes of higher organisms that has been studied in most detail. Many fundamental laws of the structural and functional organization of eukaryote genes have been elucidated with the TF gene as model. In the original experiments on the structure of the TF (conalbumin) gene the preparative isolation of the conalbumin mRNA from the avian oviduct was undertaken and attempts were made at its reverse transcription. This showed that under the standard conditions of reverse transcription that are the optimum for the synthesis of the full-length cDNAs of ovalbumin, of ovimucoid, and of lysozyme, only partial cDNA copies of the conalbumin mRNA were formed [31]. It is interesting to note that the complexity of the production of full-length transcripts of TF mRNA has also been established for rat transferrin mRNA (Ryskov et al.). In this work, double-stranded cDNA was synthesized on a TF mRNA template and was integrated into the pBR 322 plasmid at a PstI site. The set of pRTf recombinant plasmids obtained was characterized by heterogeneous dimensions of the insert of the transferrin cDNA of from 150 to 1500 n.p. The authors did not succeed in obtaining the full-length double-stranded copy of rat TF mRNA. Nevertheless, the results of restriction analysis permitted the assumption that the bulk of the TF mRNA molecule was represented in the set of cloned DNA sequences [43]. Subsequently, the conditions for the reverse transcription of this mRNA and the synthesis of its full-sized transcript were optimized [2, 41, 42].

The main contribution to the study of the structural organization of the TF gene had been made by Chambon's group in France and Palmiter's group in the USA. In the combined work of these authors [42], full-length cDNA was synthesized on a conalbumin mRNA template. Then the second strand of the DNA was constructed in a polymerase reaction and the two-stranded structural gene was cloned in pBR 322. After the preliminary selection of the transformed bacterial clones, the recombinant plasmid pBR322—Con I bearing the sequence of transferrin cDNA was identified. The DNA of this plasmid was selectively hybridized with mRNA, which programmed the synthesis of TF in a cell-free system [42]. The size of the insert was determined by the electron microscopy of the heteroduplexes formed on the hybridization of pBR322 (wild type) with the recombinant plasmid pBR322—Con I, and proved to be 2350 n.p. [42]. This value corresponds to the refined length of the TF mRNA molecule, which, according to the results of electrophoresis in agarose gel with methylmercury hydroxide, is 2400 nucleotides. Thus, the recombinant plasmid obtained carried an almost full-length DNA copy of the mRNA and, in any case, the complete sequence that is translated. This permitted the cloned "synthetic" gene to be used as a hybridization probe for the study of the structure of the natural avian TF gene.

In the first place, a compartive restriction analysis was made of the cloned synthetic gene and the chromosomal TF gene in avian DNA [44]. It was found that the TF gene has a complex structural organization and the combination of its coding segments is located within a fragment of the chromosomal DNA with a length of not less than 10 thousand n.p. Restriction analysis also showed that the coding sequences of this gene (exons) were interrupted at several points (not less than six) by noncoding inserts (introns) containing additional recognition sections for a number of restrictases. Thus, the sequence of the synthetic TF gene contained only one recognition section for restrictase EcoRI, while in the chromosomal TF gene three EcoRI fragments having lengths of 12.0, 4.0, and 2.6 thousand n.p. (fragments *a*, b, and c, respectively) were detected. No sections of attack for restrictase PstI were detected within the synthetic structural TF gene, but the chromosomal gene after treatment with this nuclease gave six fragments with different molecular weights containing coding sequences of the TF gene. The restrictases HpaI and KpnI cleaved neither the synthetic nor the chromosomal gene.

These authors also made a comparative study of the nature of the cleavage by a number of restrictases of the chromosomal TF genes in DNA preparations from individual birds. Substantial interallele differences were found in the structures of the sections recognized by the restrictases EcoRI, HindIII, and BamHI. It was also established that the DNAs of the liver and oviduct each contained one copy of the TF gene in the haploid genome and that, according to the results of restriction analysis, the TF gene in these two tissues has the same molecular organizations.

The results of the work of Lee et al. [44] were confirmed by the experiments of Perrin et al. [45]. These workers showed that the sequences coding the TF mRNA are contained in the three EcoRI fragments (a, b, and c) of avian genomic DNA, and within a transcription unit they were oriented in the following way: 5'-b-c-a-3'. By the hybridization of colonies with radioactive DNA of the pBR322-Con1 plasmid from three avian gene libraries obtained independently in the λ-vector clones bearing fragments b and c and part of fragment a of the TF gene were isolated. In one of these clones (λC4-con1), an insert was detected that contained the 5'-terminal sequences of the mRNA with a total length of 940 nucleotides. The DNA of this clone was hybridized with TF mRNA under conditions for the formation of R loops, which were analyzed by electron microscopy. It was found that the hybrid of DNA of phage λC4-con1 with TF mRNA contained six introns with lengths of 1306 ± 125, 266 ± 61, 1108 ± 128, 134 ± 35, 333 ± 47, and 457 ± 44 n.p. Seven exons forming hybrids with the cloned fragment of the chromosomal TF gene had lengths of 113 ± 22, 174 ± 20, 133 ± 31, 197 ± 26, 145 ± 21, 84 ± 19, and 107 ± 12 n.p. Thus, the region of the natural gene determining the structure of less than half of the TF mRNA molecule from its 5'-end exceeds this length by a factor of five. The same relationships between the length of the exons and of the introns were found for a series of other individual genes.

An exhaustive pattern of the molecular organization of the avian TF gene was created after the cloning of the complete EcoRI - a fragment of the chromosomal gene [46], which, from the results of the electron microscopy of DNA-mRNA hybrids, contained 14 exons and, correspondingly, 13 introns separating these exons. The combination of all the results of the study of the cloning of the fragments of the TF genes indicates that the TF gene is represented by 17 exons and contains 16 introns (Fig. 3). The total length of the exons of the TF gene, according to Chambron's groups, is 2508 ± 164 n.p. This magnitude agrees well with the results of a determination of the complete primary structure of the TF mRNA for the avian oviduct published in 1982 [2]. According to these results, the TF mRNA chain consists of 2376 nucleotides (not counting the 3'-terminal polyadenylate sequence).

The combination of the coding segements of the TF gene was localized within a segment of the chromosomal DNA with a length of 10.3 thousand n.p. (between the 5'-terminal exon 1 and the and 3'-terminal exon 17). This conclusion was confirmed by the results of a determination of the dimensions of the nuclear precursor of TF mRNA, which showed that the longest precursor has a chain length of 9900-10,100 nucleotides [46].

The study of the primary structure of the chromosomal avian TF gene has only begun at the present time [2, 46]. There is information in the literature on the nucleotide sequence of exon 1, which codes the 5'-untranslated zone of the TF mRNA (from the cap to the initiatory AUG codon) having a length of 76 nucleotides, and also the section of the transcribed zone of the TF mRNA determining almost all the signal pres-sequence (14 codons out of 19). In general form, this result confirms the hypothesis that each exon codes a definite functional or structural domain of the protein product of the gene [46-50]. In particular, exon 1 codes the domain of the pre-TF (signal sequence) that is necessary for the vectorial transport of this protein in the cell. A similar situation has also been described for the genes of the light chains of the immunoglobulins - the first exon likewise codes 15 out of the 19 amino acids of the signal pre-sequence [49, 50].

Thus, the avian TF gene is one of the most split genes that has been investigated up to the present time. It contains 17 exons characterizing very short lengths (50-200 n.p.). It is interesting to note that in intron C has been detected one of the moderately repeating sequences dispersed through the avian genome to which possible regulatory functions in the expression of the genes are assigned [46].

Figure 4 shows a scheme of the molecular organization of the mature TF mRNA constructed from the results of the determination of its nucleotide sequences [2]. The TF mRNA has a chain length of 2376 nucleotides not counting the poly(A) tract and the inverted m^7G component of the 5'-terminal cap. At the 3'-end of the TF mRNA chain there is a short untrans-

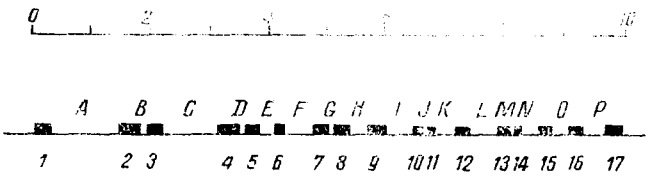

Fig. 3. Molecular organization of the avian transferrin gene [46]: scale at the top – thousands of n.p.; 1-17) exons; A-P) introns.

lated sequence (182 nucleotides) which, on direct structural analysis, proved to be substantially shorter than was assumed on the basis of a comparison of the length of the TF mRNA chain and the molecular weight of the pre-TF [31, 41]. In the 3'-untranslated zone of TF mRNA the canonical sequence AAUAAA localized at a distance of 18 nucleotides from poly(A) and playing, as is assumed, the role of the recognition section for poly(A) polymerase is found [2]. The length of the 5'-untranslated region of the TF mRNA amounts to 76 nucleotides.

Thus, the results of a study of the molecular architecture of the avian chromosomal TF gene show that it, like the majority of eukaryotic genes investigated, has a linear-mosaic structure. Attention is attracted by the multiplicity of introns and their considerable length, several times exceeding the total length of the coding segments of the TF gene. This type of structural organization is also characteristic for a number of other structural genes. Thus, the pro-α2-collagen gene of the chick has a total length of 42 thousand n.p. and contains 50 introns. The coding sequence of this gene is represented by short exons (with lengths of 45-108 n.p., most frequently 54 n.p.) [51]. It is assumed that such a complexly organized procollagen gene was formed in the course of evolution as the result of numerous duplications of an elementary coding unit with a length of 54 n.p. The chromosomal vitellogenin gene of birds and amphibia is also characterized by a large number of introns [52]. A comparison of the mosaic structure of the TF gene and the domain organization of TF itself also permits the assumption that the TF gene was formed as the result of the duplication of a precursor gene in the course of evolution [1, 2, 7] or even by the quadruplication of an early gene.

The combined length of the exons of the TF gene amounts to about 25% of its total length. For some other genes the proportion of exons is only 5-10% of the total length of the transcription unit, and the bulk of the DNA of these units is due to very long introns. Such genes include [53] ovalbumin gene and the gene coding dihydrofolate reductase [53].

The study of the molecular organization of the TF gene is of interest in connection with the fact that in birds this gene is actively expressed in two organs – the liver and the oviduct. However, these two organs differ from one another both in the degree of expression of the TF gene and in the regulation of its functional activity (see above). In the oviduct the expression of the TF gene is under the strict control of estrogenic hormones which only slightly stimulate the synthesis of TF in the liver. At the same time, a deficiency of iron induces the synthesis of TF and the accumulation of TF mRNA in the hepatocytes but scarcely affects the expression of the TF gene in the oviduct [18, 19]. Therefore, a determination of the features of the molecular organization of the TF gene in the two tissues would be of considerable interest for an understanding of the molecular mechanisms effecting the tissue-specific positive control of the expression of the TF gene. However, at the present time no features whatever of the molecular organization of the TF genes in the genomes of the liver and oviduct have been found from the results of the restriction mapping of the DNAs of these organs or a comparative analysis of the molecular dimensions of the TF mRNA.

In a series of investigations by Chambon's group, the cloned avian TF gene was used as a model object for the study of the features of the structural organization of the 5'-flanking regions of eukaryotic genes and for the identification of the functionally important sequences of these regions that are necessary for the adequate initiation of transcription.

It is known that the initiation of the transcription of prokaryotic genes takes place as the result of the binding of RNA polymerase to promotor sequences. One of them (the Pribnow box) is at a distance of 10 nucleotides from the point of the beginning of the synthesis of RNA, and another is 35 nucleotides away [54]. Apparently, the RNA polymerase is bound successively to these sections, and its complex with the Pribnow box is stronger and ensures accuracy of initiation.

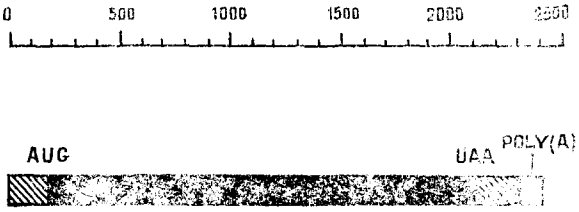

Fig. 4. Scheme of the structure of mature avian transferrin mRNA [2]: scale at the top – number of nucleotides; the translated zone is black and the 5'- and 3'-untranslated zones are hatched.

In recent years, substantial advances have been achieved in the study of the structure of eukaryotic promotors and fine mechanisms ensuring the specificity of the initiation of transcription. This progress is due, in the first place, to the possibility of cloning individual genes and their flanking regions of the genomic DNA and, in the second place, to the creation of highly effective systems for cell-free transcription with the participation of RNA polymerases II and III. Above all, it has been shown that in the 5'-flanking regions of the majority of eukaryotic genes that code proteins and are transcribed by RNA polymerase II there is a universal AT-rich sequence (prototypical sequence TATAAAA for the coding strand), which has acquired the name of the TATA block or the Hogness–Goldberg box [53, 55, 57]. The TATA block has been found in more than 60 genes that have been studied. A strictly conservative element of this block is the tetranucleotide TATA, but in a pairwise comparison of a number of eukaryotic genes having nothing in common from the evolutionary or functional points of view even longer sections of homology can be found. Thus, in the 5'-flanking region of the avian TF gene and the late transcription unit of adenovirus 2 a common 12-nucleotide sequence 5'-CTATAAAAGGGG-3' has been found [57]. This sequence, just like the TTA blocks of other genes, is found at a distance of 25-30 n.p. in the 5'-direction from the start point of transcription, which, to a first approximation, coincides with the beginning of the first exon.

The second conservative element of the primary structure of the 5'-flanking regions of the majority of eukaryotic genes studied is the so-called CAT block which is 70 nucleotides away from the start of transcription. It is assumed that the CAT and TATA blocks are spatially separated elements of the promotor for RNA polymerase II. The CAT block ensures the effective binding of RNA polymerase II and the TATA block is necessary for the accurate initiation of transcription. Basic inflormation showing the necessity of the TATA sequence for the accurate initiation of transcription has been obtained in experiments on the cell-free transcription of the avian TF gene and some other cloned genes. Wasylyk et al. [57] studied the comparative efficacy of the cell-free transcription and accurate initiation of the transciption of the genes for TF, ovalbumin, and the early and late regions of the genome of adenovirus 2 (Ad-2). It was found that the TATA sequence is necessary for specific initiation, i.e., for the correct choice by the RNA polymerase II of the starting point for the synthesis of RNA. The TF gene and the late region of Ad-2 containing the 12-nucleotide common TATA sequence were transcribed with the same efficiency, while the ovalbumin gene and the early region of Ad-2, having no such extended zone of homology in the TATA block were transcribed less effectively. The relatively low efficiency of transcription with the ovalbumin gene in the cell-free system is of special interest, since in the cells of the avian oviduct on induction by estrogenes it is precisely this gene that is transcribed with the maximum efficiency.

With this aim of a further analysis of the role of the TATA block in the proper initiation of the transcription of the TF gene, mutants of this gene with changes in the structure of this zone have been obtained. In the first place, a mutant was constructed with the point substitution T → G, a so-called TAGA mutant [56]. For this purpose, the restriction fragment Pst5–Pst6 of the TF chromosomal gene, bearing the TATA sequence, was subcloned in the single-stranded vector jd 103 and was then hybridized with the synthetic 11-membered oligodeoxynucleotide 5'-CTTTTCTAGAG, which played the role of primer in the reaction for constructing the second strand with DNA polymerase I. This primer is complementary to the TATA sequence of the cloned fragment of the TF gene with the exception of the G residue in position 9 of the primer (underlined). The recombinant molecule obtained in this way containing the T → G substitution in one of the strands was cloned in E. coli. The clones bearing the mutant TAGA sequence were selected from the nature of the cleavage of the DNA by the restric-

tase XbaI, since as the result of the T → G substitution in the TATA sequence a section of recognition for this endonuclease was formed. In a comparative analysis of the products of cell-free transcription of a normal fragment of the TF gene and its TAGA mutant in a cell-free system with RNA polymerase II it was found that the point substitution T → G in the TATA block sharply decreased the specificity of the initiation of transcription, which for the TAGA mutant was only 5% of the control level.

The well-defined influence of the point substitution T → G on the interaction of the TATA block with the RNA polymerase can be explained by two alternative mechanisms of the action of this mutation: 1) The mutation changes the primary structure of the TATA block, the specificity of which is necessary for the appropriate initiation of transcription; and 2) the mutation is accompanied by a change in the nucleotide composition of the TATA block — in particular, the replacement of a AT pair by a GC pair — and by an increase, because of this, in the stability of the DNA duplex which may interfere with its local untwisting on binding with the DNA polymerase.

For an unambiguous answer to these questions, the TAAA mutant of the Pst5—Pst6 fragment of the cloned TF gene was constructed and its transcription was also studied in a cell-free system [58]. The authors concerned started from the assumption that the TAAA mutation is a transversion which is not accompanied by changes in the nucleotide composition of the corresponding block or by any shifts whatever in the stability of the duplex DAN with respect to untwisting action. It was found that the accuracy of the initiation of the transcription of the TAAA mutant was half that for the TAGA mutant. Thus, RNA polymerase recognizes precisely the primary structure of the TATA blocks and not AT-rich sections, and point nucleotide substitutions in this block sharply affect the specificity of transcription. The results of experiments with various mutants bearing local deletions (1-4 nucleotides) in the TATA block have confirmed this conclusion [58].

Mutant	Yield of specific transcripts, %
TATA (wild type)	100
TAGA	5
TAAA	2.5
ΔTA*	3.25
ΔTAA*	1.5
ΔTAAA*	0.75

Proofs of the functioning of the TATA block as an element of the promotor of RNA polymerase II were also obtained in experiments with the excision of the TATA block in a fragment of the adenovirus DNA from positions −32 to −12 and its introduction into pBR322, which led to the appearance of a new initiation section in the plasmid DNA [59].

In addition to the TATA and CAT blocks identified as elements of the promotors of eukaryotic genes transcribed by RNA polymerase II, the efficacy of the transcription of a number of genes is controlled by elements located 100 n.p. and more from the start point [53]. It is assumed that these regulatory elements function as sections binding the regulator proteins of transcription and this binding in some way changes the conformation of the gene in the chromatin and its accessibility for RNA polymerase II.

Thus, a considerable amount of material relative to the molcular organization of the avian TF gene and the routes of its expression has accumulated in the literature. Furthermore, using the avian TF gene as model, a number of universal laws of the structure and functioning of eukaryotic genes have been revealed. Nevertheless, a number of unanswered fundamental questions remain (the sequence of the reactions in the processing of the primary transcript and the maturing of the TF mRNA, the organ-specific regulation of the expression of the gene by estrogens and iron ions, etc.). The structure and expression of the mammalian TF gene has remained completely unstudied, although investigations in this direction should be of interest from the point of view of the pathways of evolution of the genes for TF and other metalloproteins and of the molecular mechanisms of the pathology of the synthesis of TF in hereditary and acquired diseases.

*Deletional mutants.

LITERATURE CITED

1. J. Williams, T. C. Elleman, Y. B. Kingston, A. G. Wilkins, and K. H. Kuhn, Eur. J. Biochem., 122, No. 2, 297 (1982).
2. J. M. Jeltch and P. Chambon, Eur. J. Biochem., 122, No. 2, 291 (1982).
3. L. Y. DeLucas, F. L. Suddath, R. A. Gams, and C. E. Bugg, J. Biol. Chem., 123, No. 2, 285 (1978).
4. B. Gorinsky, C. Horsburgh, D. F. Lindley, D. S. Moss, M. Parkar, and Y. L. Watson, 5th Eur. Crystallogr. Meet. Copenhagen, Collect. Abstr., s. 1, 162 (1979).
5. Y. E. L. Brune, S. R. Martin, B. S. Boyd, R. M. Palmour, and H. E. Sutton, Tex. Rept. Biol. Med., 36, No. 1, 47 (1978).
6. T. A. Salikhov and A. A. Buganov, Khim. Prir. Soedin., 786 (1983).
7. J. Williams, Trends Biochem. Sci., 7, No. 11, 394 (1982).
8. R. T. A. MacGillivray, E. Mendez, S. K. Sinha, M. K. Sutton, J. Lineback-Zins, and K. Brew, Proc. Nat. Acad. Sci. USA, 79, No. 8, 2504 (1982).
9. M. H. Metz-Boutigue, J. Jollès, J. Magurier, G. Spik, J. Montreuil, and P. Jollès, Biochemie, 60, No. 5, 557 (1978).
10. M. H. Metz-Boutigue, J. Mazurier, J. Jollès, G. Spik, J. Montreuil, and P. Jollès, Biochem. Biophys. Acta, 670, No. 2, 234 (1981).
11. A. Bezkorovainy, D. Grohlich, Comp. Biochem. Physiol., 47, No. 4, 787 (1974).
12. J. Williams, S. A. Grace, and J. M. Williams, Biochem. J., 201, No. 2, 417 (1982).
13. J. P. Brown, R. M. Hewick, J. Hellstrom, R. F. Doolittle and W. J. Dereyer, Nature (London), 296, No. 5853, 171 (1982).
14. H. F. Lodish and J. Rose, J. Biol. Chem., 252, No. 4, 1181 (1977).
15. T. A. Salikhov and A. A. Buglanov, Khim. Prir. Soedin., 139 (1984).
16. E. H. Morgan and T. Y. Peters, J. Biol. Chem., 246, No. 11, 3500 (1971).
17. E. H. Morgan, J. Biol. Chem., 244, No. 15, 4153 (1969).
18. S. G. MacKnight, D. C. Lee, D. Hemmaplardh, C. A. Finch, and R. D. Palmiter, J. Biol. Chem., 255, No. 1, 144 (1980).
19. S. G. MacKnight, D. C. Lee, and R. D. Palmiter, J. Biol. Chem., 25, No. 1, 148 (1980).
20. V. S. Gaitskhokhi, The Messenger RNAs of Animal Cells [in Russian], Moscow (1980).
21. G. Schreiber, H. Dryburg, A. Millership, J. Matsuda, A. Inglis, J. Phillips, K. Edwards, and J. Maggs, J. Biol. Chem., 254, No. 23, 1203 (1979).
22. V. S. Gaitskhoki, N. A. Timchenko, L. T. Timchenko, L. V. Puchkova, Kh. A. Aslanov, and T. A. Salikhov, Biokhimiya, 46, No. 8, 1426 (1981).
23. V. S. Gaitskhoki, N. A. Timchenko, L. T. Timchenko, L. V. Puchkova, Kh.A. Aslanov, and T. A. Salikhov, Biokhimiya, 47, No. 1, 13 (1982).
24. G. Schriber, H. Dryburg, K. Weigand, M. Schreiver, J. Witt, H. Seydewitz, and G. Howlett, Arch. Biochem. Biophys., 212, No. 2, 319 (1981).
25. G. Blobel and B. Dobbersteine, J. Cell. Biol., 57, No. 3, 835 (1975).
26. J. Garnon, R. D. Palmiter, and K. A. Walch, J. Biol. Chem., 253, No. 20, 7664 (1978).
27. G. Kreil, Ann. Rev. Biochem., 50, 317 (1981).
28. R. D. Palmiter, J. M. Davidson, J. Garnon, D. W. Rowe, and P. Bornstein, J. Biol. Chem., 254, No. 5, p. 1433 (1979).
29. A. W. Strauss, C. D. Bennett, and A. M. Donohue, J. Biol. Chem., 252, 6946 (1977).
30. S. N. Thibodeau, D. C. Lee, and R. D. Palmiter, J. Biol. Chem., 253, No. 11, 3771 (1978).
31. D. C. Lee, G. S. MacKnight, and R. D. Palmiter, J. Biol. Chem., 253, No. 10, 3494 (1978).
32. S. N. Thibodeau, R. D. Palmiter, and K. A. Walch, J. Biol. Chem., 253, No. 24, 9018 (1978).
33. R. D. Palmiter, J. Garnon, L. Ericsson, and R. Walch, J. Biol. Chem., 252, 6386 (1977).
34. G. S. MacKnight and R. D. Palmiter, J. Biol. Chem., 254, No. 18, 9050 (1979).
35. R. D. Palmiter and L. T. Smith, Nature New Biology, 246, No. 151, 74 (1973).
36. N. E. Hynes, B. Groner, A. E. Sippel, S. Jeep. T. Wurtz, M. C. Nguen Hun, K. Giesecke, and G. Schutz, Biochemistry, 18, No. 4, 616 (1979).
37. G. Schutz, M. C. Nguen Hun, A. E. Sippel, H. Land, W. Lindenmarier, T. Wurtz, K. Giesecke, and H. Y. Hauser, Z. Physiol. Chem., 358, No. 9, 1147 (1978).
38. E. R. Mulvihill and R. D. Palmiter J. Biol. Chem., 255, No. 5, 2085 (1980).
39. N. A. Timchenko, L. T. Timchenko, A. L. Shvartsman, T. A. Salikov, and V. S. Gaitskhoki, Mol. Biol., 17, No. 2, 322 (1983).
40. N. A. Timchenko, Al. L. Shvartsman, T. A. Salikhov, and V. S. Gaitskhoki, Mol. Biol. Mikrobiol. Virusologiya, 1, No. 2, 33 (1983).
41. G. L. Buell, M. P. Wickens, F. Payar, and R. T. Schimke, J. Biol. Chem., 253, No. 7, 2471 (1978).

42. M. Cochet, F. Perrin, F. Gannon, A. Krust, P. Chambon, G. S. MacKnight, D. C. Lee, K. E. Mayo, and R. D. Palmiter, Nucl. Acids Res., 6, No. 7, 2435 (1979).
43. A. P. Ryskov, N. A. Timchenko, L. T. Timchenko, T. A. Salikhov, and V. S. Gaitskhoki, Mol. Biol., 18, No. 1 (1984).
44. D. S. Lee. G. S. MacKnight, and R. D. Palmiter, J. Biol. Chem., 255, No. 4, 1442 (1980).
45. F. Perrin, M. Cochet, P. Gerlinger, B. Cami, J. P. LePennec, and P. Chambon, Nucl. Acids. Res., 6, No. 8, 2731 (1979).
46. M. Cochet, F. Gannon, R. Hen, L. Moroteaux, F. Perrin, and P. Chambon, Nature (London), 282, No. 5739, 567 (1979).
47. A. L. Shvartsman, Usp. Sovr. Biol., 90, No. 1, 3 (1980).
48. F. Crick, Science, 204, 264 (1979).
49. H. Sakano, J. Rogers, R. Muppi, C. Brack, A. Trauneckok, R. Maki, R. Wall, and S. Tonegawa, Nature (London), 277, No. 5698, 627 (1979).
50. S. Tonegawa, A. M. Maxam, K. Tizard, O. Berhard, and W. Gilbert, Proc. Nat. Acad. Sci. USA, 75, No. 3, 1485 (1978).
51. J. Wozney, D. Hanahan, R. Morimoto, H. Boedtker, and P. Doty, Proc. Nat. Acad. Sci. USA, 78, No. 2, 712 (1981).
52. W. Wahli, J. B. Dawid, T. Wyler, R. Jaggi, and G. U. Ryffel, Cell, 16, No. 3, 535 (1979).
53. E. Bresnick, J. Levy, R. M. Hines, W. Lewin, and P. E. Thomas, Arch. Biochem. Biophys., 212, No. 2, 501 (1981).
54. D. Pribnow, J. Mol. Biol., 99, 419 (1975).
55. S. J. Tsai, M. Y. Tsai, and B. W. O'Malley, Proc. Nat. Acad. Sci. USA, 78, 879 (1981).
56. B. Wasylyk, R. Derbyshire, A. Guy, D. Molko, A. Roget, T. Teonle, and P. Chambon, Proc. Nat. Acad. Sci. USA, 77, No. 12, 7024 (1980).
57. B. Wasylyk, C. Kedinger, J. Corden, O. Brison, and P. Chambon, Nature (London), 285, No. 5764, 367 (1980).
58. B. Wasylyk and P. Chambon, Nucl. Acids Res., 9, No. 8, 1813 (1981).
59. P. Sassone-Gorsi, J. Cordon, C. Kedinger, and P. Chambon, Nucl. Acids Res., 9, No. 19, 3914 (1981).
60. R. Newman, C. Schneider, R. Sutherland, L. Vodinelich, and M. Greaves, Trends Biochem. Sci., 7, No. 11, 397 (1982).

ANALYSIS OF α-GLYCOLIC COMPOUNDS USING THE POTENTIOMETRIC DIFFERENTIATED TITRATION OF PERIODATE AND FORMALDEHYDE

B. A. Spintse and A. Ya. Veveris

UDC 543.242.3:547.426.1+547.426.24

The conditions of the analysis of α-glycolic compoounds (ethylene glycol, glycerol, mannitol, D-glucose, D-ribose, and dihydroxyacetone) using the potentiometric differentiated titration of periodate and formaldehyde have been studied. A procedure for quantitative determination is proposed and the possibilities have been shown of using it for evaluating the stoichiometry of oxidation and the separate analysis of two-component mixtures. The performance of the determinations is distinguished by simplicity and adequate reliability.

In the quantitative analysis of compounds containing α-glycol groupings, use is frequently made of oxidation with periodate followed by the determination if its excess or of one of the reactions products – iodate, an aldehyde, or formic acid [1-3]. In the study of the structure and kinetics of the oxidation of polyhydroxy compounds, in the analysis of mixtures, and in other cases the necessity arises for performing parallel determinations both of the consumption of periodate and of the amount of formic acid, formaldehyde, and other products that have been formed. A combination of two or more methods is not infrequently used for these purposes [4, 5], but the performance of the analyses is then an extremely complicated and laborious process.

All-Union Scientific-Research Institute of Applied Biochemistry. "Biokhimreaktiv" Scientific-Production Association, Olaine. Translated from Khimiya Prirodnykh Soedinenii, No. 3, pp. 305-309. May-June, 1985. Original article submitted July 3, 1984.

