Nucleic Acids

Genetic Engineering is a separate keyword.

HELMUT BURTSCHER, Boehringer Mannheim GmbH, Penzberg, Federal Republic of Germany (Chaps. 1 – 6)

SIBYLLE BERNER, Boehringer Mannheim GmbH, Tutzing, Federal Republic of Germany (Chap. 7)

RUDOLF SEIBL, Boehringer Mannheim GmbH, Penzberg, Federal Republic of Germany (Chap. 8)

KLAUS MÜHLEGGER, Boehringer Mannheim GmbH, Tutzing, Federal Republic of Germany (Chap. 9)

1. Introduction

Nucleic acids are high molecular mass compounds found in all living cells and viruses. Their name originates from their discovery in the nuclei of eucaryotic cells. They can be chemically degraded to yield phosphoric acid, pentoses, and nitrogen-containing heterocycles (bases). Nucleic acids can be divided into two main classes depending on the sugar they contain: *deoxyribonucleic acids* (DNA) contain 2-deoxy-D-ribose and *ribonucleic acids* (RNA) contain D-ribose.

Nucleic acids are long, unbranched chains of sugar and phosphate (Fig. 1, see next page): the C-3′ atom of each sugar is linked by a phosphodiester bond to the C-5′ atom of the neighboring sugar. Either a purine (adenine, guanine) or a pyrimidine (cytosine and thymine in DNA; cytosine and uracil in RNA) is attached to C-1′ of the sugar by a β-glycosidic bond.

Although nucleic acids have been known since the second half of the nineteenth century it was only in the 1940s that their importance as the carrier of genetic information became clear. Genetic engineering and improved physical and biochemical methods of analysis have led to enormous progress in the understanding of the structure of DNA, DNA – protein interactions, and gene organization, expression, regulation, and transfer. The importance of nucleic acids became even more obvious after the discovery that they can have other functions in addition to their ability to store and transfer genetic information. It is widely assumed that in the course

Ullmann's Biotechnology and Biochemical Engineering, Vol. 1

ISBN: 978-3-527-31603-8

of evolution first RNA and then DNA came into being [36, 37].

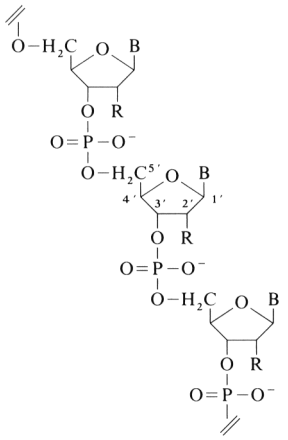


Figure 1. Structure of DNA (R = H) and RNA (R = OH)
B = base (adenine, guanine, thymine or uracil, cytosine)

2. Structure

2.1. Structure of DNA

The joining of the DNA building blocks by 5′- and 3′-phosphodiester bonds gives the molecule polarity (Fig. 1); base sequences are always written starting with the 5′-terminus, i.e., in the $5' \rightarrow 3'$ direction. The specific base sequence of DNA and its ability to form double-stranded structures according to precisely defined rules are of utmost importance for the storage of genetic information and for interactions with other nucleic acids and proteins.

From X-ray analysis data, CRICK and WATSON proposed a double-stranded structure for DNA in 1953 in which two antiparallel (i.e., $5' \rightarrow 3'$ and $3' \rightarrow 5'$) polynucleotide chains form a right-handed helix (i.e., looking along the axis of the helix, the strands are coiled clockwise). Naturally occurring DNA usually consists of right-handed helices with a major and a minor groove (Fig. 2). The hydrophobic bases are located inside the helix and the sugar – phosphate "backbone" on the outside [38]. Bases that are opposite each other are paired according to defined rules as a result of hydrogen bond formation: adenine always pairs with thymine or uracil and guanine with cytosine. Complementary bases can be bound by the more common Watson – Crick pairing (Fig. 3 A) or by Hoogsteen base pairing (Fig. 3 B). The double-stranded structure is further stabilized by hydration of the phosphate groups and hydrophobic interactions between the aromatic ring systems that result in stacking of the bases.

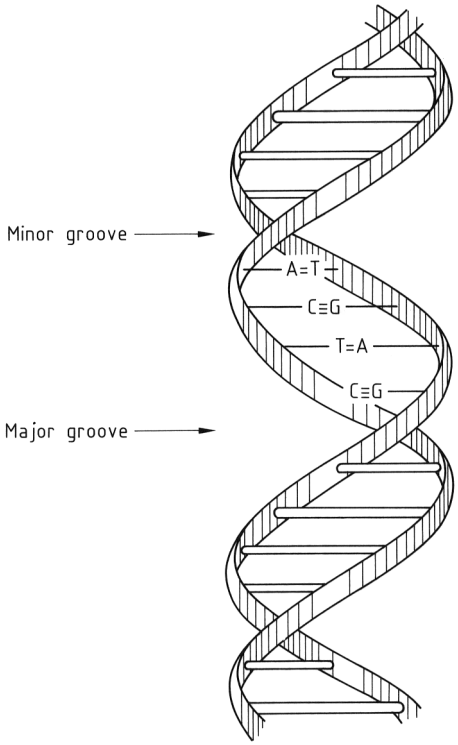


Figure 2. Right-handed double-helix of DNA
A = adenine, C = cytosine, G = guanine, T = thymine

In double-stranded DNA, the bases are densely stacked and there is a cooperative effect between hydrogen bonding and stacking. Internal bases can be continuously paired and unpaired; double-stranded regions open and form single-stranded "bubbles" ("breathing" of DNA). Breathing is more frequent in regions rich in A – T pairs and could be important for interactions with proteins. DNA helices can exist in various forms (A, B, C, D, and Z) [39] some of which are interconvertible depending on the concentration and type of salts present. The helices always exhibit a degree of microheterogeneity that plays an important part in genetic regulation mechanisms.

The DNA helices can exist in linear form (e.g., in the chromosomes of higher organisms)

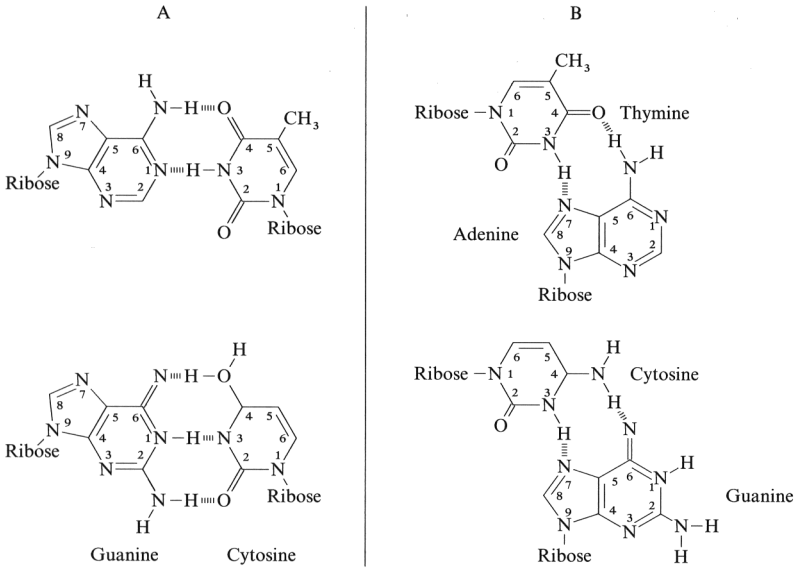


Figure 3. Watson – Crick base-pairs (A), and Hoogsteen base-pairs (B)

or as closed rings (e.g., in *Escherichia coli*); the molecules can also be twisted (superhelicity or supercoiling). In order to accommodate the large amount of DNA present in living cells, it must be packaged as compactly as possible with the help of proteins and RNA. Proteins can recognize specific binding sites on the DNA. The grooves of the DNA helix are large enough to allow proteins to come into contact with the bases [40]. Defined regions in DNA can also be recognized with the help of the methylation pattern of the bases (see Section 4.3).

Forms of DNA.

A-DNA can be observed in X-ray analyses at 66 % relative humidity. It has 11 base pairs per turn of the helix, the planes of the base pairs are tilted away from the vertical helical axis (19°), the helix is right-handed and has a diameter of ca. 2.3 nm.

B-DNA is the classical Watson – Crick form. It represents the structure of DNA at a relative humidity of > 92 % and largely corresponds to that found under physiological conditions. The helix is also right-handed with about 10.2 – 10.4 base pairs per turn and a diameter of ca. 2 nm. Single unpaired bases can be "looped out" of the helix and barely disturb the rest of the structure [41,42]. Protein – DNA interactions usually require recognition of nucleotide sequences in the major groove of the B-DNA double helix.

C-DNA helices can be observed at a relative humidity of 44 – 66 % in the presence of lithium salts. The helix is also right-handed and similar to the B form, but with 9.3 base pairs per turn.

D-DNA occurs in nature only in sequences with alternating adenine and thymine residues and in the DNA of the bacteriophage T 2 (T-DNA). The helix is also right-handed and has 8 base pairs per turn.

The left-handed conformation of *Z-DNA* has an alternating sequence of pyrimidines and purines and is formed in vitro at high salt concentrations (> 2 mol/L NaCl) or in the presence of divalent cations ($Mg^{2+} > 0.7$ mol/L). Unlike the right-handed helices (which have two grooves), this structure forms a single, very deep groove that penetrates the helix axis. The sugar – phosphate backbone assumes a zig-zag arrangement (therefore Z-DNA) with 12 base pairs per turn of the helix. Z-helices can form in vivo at physiological salt concentrations. They are less stable than B-DNA, but are stabilized by supercoiling, proteins, special ions, and methylation [43]. Torsional stress of DNA in vivo can favor the formation of Z-DNA [44]. Z-DNA and B-DNA are interconvertible; part of a DNA molecule may exist in the B form and another part in the Z form.

Supercoiling. Circular DNA and DNA between fixed sites can be twisted to supercoils. The term *supercoiling* refers to the curvature of the double helix axis. Supercoiled (superhelical) DNA was discovered in the 1960s in polyoma virus [45]. Rotation in the direction of winding is called *positive supercoiling* and rotation in the opposite direction is called *negative supercoiling*.

Torsional stress due to negative supercoiling can be overcome by the formation of DNA structures other than the B form. Negative supercoiling is a strong driving force for the stabilization of Z-DNA. Supercoiling makes DNA more compact, which is very important in DNA packaging. Almost all naturally occurring superhelical DNAs are underwound (i.e., have negative superhelices) but overwound DNAs also exist [46]. The strain produced by over- or underwinding can be accommodated by the formation of local single-stranded regions which tends to increase with increasing temperature. "Breathing" of the DNA (see third paragraph in Section 2.1) plays an important part here too. A sequence with $> 90\,\%$ A – T can exist permanently unpaired in a superhelical molecule. This is important for many reactions of DNA. Supercoiling influences transcription (see Section) and vice versa. Positive supercoils are formed in front of the transcription apparatus and negative supercoils behind it; these supercoils are controlled by enzymes [47].

Bending. The base sequence of DNA is of tremendous importance for its structure [44]. In a right-handed helix, the twist angle between two bases changes depending on the sequence. This may result in the bending of a linear double helix. Bending can also be caused by proteins. Bending is of significance for the packaging of DNA and for many of its biological reactions [48, 49].

Intrinsically bent DNA is formed when special base sequences or structural motives are repeated in phase with the DNA helical repeat; homopolymeric A tracts being the best example [50, 51]. Protein-induced DNA bending plays an important role in recombination, initiation of transcription, and replication [52, 53]. Bends are also important structural features; indeed, regulatory protein binding sites can be replaced by an intrinsic bend [49, 54].

Special Structural Elements. Short sequences are frequently repeated in regulatory regions. *Repeats* can be recognized by DNA-binding proteins. Owing to DNA breathing, double-stranded regions (hairpins or stem-loops and cruciform structures) can be formed at repeats within a single strand (Fig. 4). This rarely happens in double-stranded DNA because stem loops are energetically less favorable than linear double strands. However, it is encountered frequently in single-stranded DNA and RNA. Supercoiling can promote the formation of cruciform structures, whereas transcription inhibits it [55]. Hairpins can play a part in replication, transcription, and RNA processing [56].

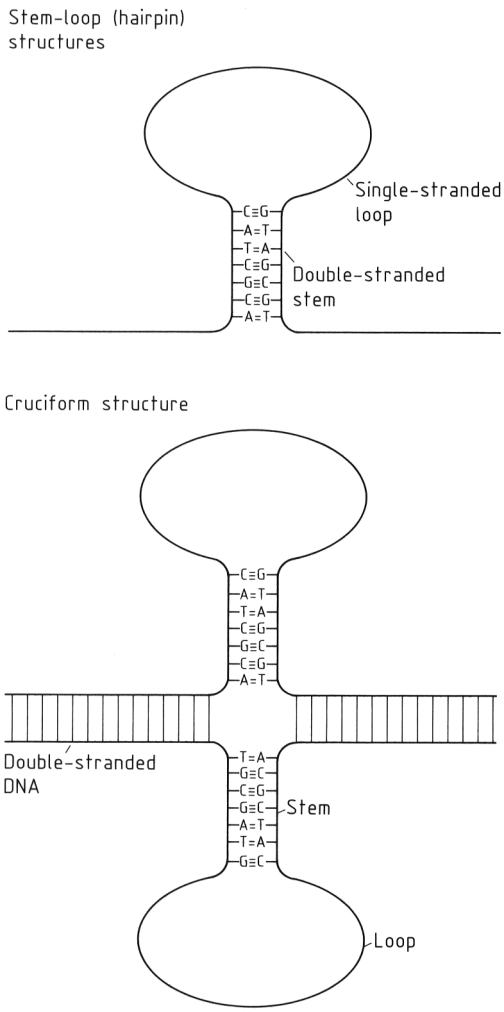


Figure 4. Special structural elements in nucleic acids

Homopyrimidine – homopurine runs are frequently found in regulatory regions of eucaryotic genes and are especially sensitive to nucleases [57]. There is a high tendency to form right-handed structures other than B-DNA in such regions [58].

Although the chains of double-stranded DNA are normally antiparallel, *parallel double-stranded oligomers* have also been found in vitro [59]. They form a right-handed helix and are even recognized by several enzymes. They are less temperature-stable than the corresponding oligomers.

At homopurine – homopyrimidine sections of the DNA, a homopyrimidine oligonucleotide can attach itself parallel to the homopurine strand in the major groove and form a *triple helix* [58]. Structures of this type can be used for specific strand cleavage with the help of coupled ellipticine derivatives or metal chelates [60,61].

Centromeres are important compact DNA structures of the eucaryotic chromosome that are rich in adenine and thymine. Their exact structure is not known but they are important for the attachment of the spindle fibers during mitosis.

The ends (*telomeres*) of linear chromosomes (as in eucaryotic DNA) pose a special problem. DNA polymerases synthesize DNA from a DNA template and always require an RNA primer to start replication. Cleavage of this primer then results in a small 5′-gap which cannot be closed by the polymerase. Under normal replication conditions, the ends should therefore beocme shorter with every cycle of DNA replication (see Section 4.1). Special enzymes (telomerases) are responsible for adding telomere repeats to the chromosome ends to maintain constant length (Fig. 5): repeats can fold back and provide a 3′-OH group which serves as a primer for copying the last segment of a linear DNA molecule. Disturbances in telomeres can lead to aging phenomena [62,63] and a role in carcinogenesis is also being discussed. Broken ends of chromosomes that are no longer protected by telomeres are very susceptible to fusion with other DNA ends and to degradation by nucleases [64]. The antiparallel structure and function of telomeres are highly conserved in all eucaryotes and are species specific. They consist of simple, tandemly repeated sequences with clusters of G residues [65,66]. The G-rich strand is aligned in the $5' \rightarrow 3'$ direction towards the end of the chromosome and has a single-stranded 3′-end containing 12 – 16 nucleotides. Telomeres can associate to form stable, parallel, four-stranded structures (G 4-DNA) [67].

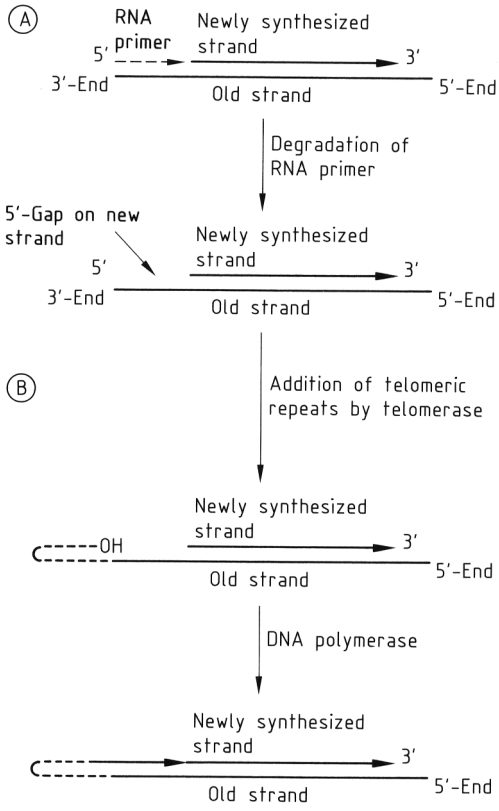


Figure 5. The importance of telomeres
Without telomere addition (A) newly synthesized DNA strands become shorter; with telomere addition by telomerase (B) constant length can be maintained.

2.2. Structure of RNA

RNA is an unbranched single-stranded polymer with many intramolecular double-stranded sections that may account for 50 – 67 % of the molecule. As in DNA, the backbone of RNA consists of 3′,5′-phosphodiester bonds (Fig. 1); however the sugar is ribose (and not deoxyribose) and uracil replaces thymine. Double-stranded RNA cannot form a B-helix because of steric hindrance caused by the 2′-OH groups of ribose; helices of the A type are, however, possible.

The functional groups of the nucleotides in the major groove of the A type of double he-

lix found in RNA are not easily accessible to proteins [68]. Protein binding to RNA probably occurs via interaction with single-stranded regions.

Four functional RNA families exist: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and small nuclear RNA (only in eucaryotes). The structure of tRNAs has been studied most extensively; about half of the ca. 75 – 90 nucleotides within the tRNA molecule are paired, resulting in a secondary structure with a stem and three loops similar to that of a cloverleaf [69].

RNA has many different biological functions and exhibits a spectrum of flexible structures that more closely resemble those of proteins rather than those of the chemically related DNA [70]. RNAs have secondary structures–double-stranded sections, hairpins, internal loops, and bulged bases. With unpaired nucleotides pronounced tertiary structures are formed in addition to the secondary structure. Examples of tertiary structure motives are pseudo knots [71], produced by folding back in a hairpin and formation of a second stem – loop structure [72].

Formation of DNA – RNA hybrids is of importance in the replication and transcription of DNA and in the reverse transcription of viral RNA. Such hybrids can form secondary structures but they are considerably more polymorphous than DNA alone [73].

3. Properties

3.1. Physical and Chemical Properties

The size of naturally occurring DNA varies from a few thousand to 10^9 base pairs. The length of such molecules (micro- to centimeter range) can easily be measured under the electron microscope.

DNA absorbs UV light at 260 – 280 nm due to its bases. Aqueous DNA solutions are very viscous; viscosity depends on DNA length, DNA concentration, and temperature. Heating to a critical temperature is accompanied by a decrease in viscosity because the hydrogen bonds responsible for base pairing are disrupted and the helix structure collapses. This process is called *thermal denaturation or melting of DNA*. The temperature at which one-half of the base pairs is disrupted is denoted the *melting temperature*. It depends on the base composition (G – C pairs are more stable than A – T pairs). Double-stranded DNA ranging in size from 100 to > 100 000 base pairs melts at ca. 90 °C. In shorter double strands a gradual decrease in the melting temperature is observed. The melting temperature increases with increasing salt concentrations because the solubility of the bases decreases and hydrophobic interactions are increased. Chemicals that compete with hydrogen bond formation, such as urea or formamide, lower the melting temperature of DNA. Methanol has a similar effect; it increases the solubility of the bases and increases the interaction with water. The "melting" of double-stranded DNA is also facilitated by solvents such as ethylene glycol, dimethylformamide, dimethyl sulfoxide; low ionic strength; or extreme pH values. DNA can be denatured at an alkaline pH because the keto – enol equilibria of the bases are shifted preventing these groups from participating in hydrogen bonding.

Since the stacked bases in the double-stranded helix are not as easily excited by UV light as in single strands, absorption at 260 nm is lower for double-stranded DNA than for single strands. Increase in UV absorption can thus be used to measure DNA denaturation. At 260 nm solutions containing 50 μg/mL of double-stranded DNA, 50 μg/mL of single-stranded DNA, and 50 μg/mL of free bases have absorptions of ca. 1.00, 1.37, and 1.60, respectively.

Denaturation can also occur in the presence of proteins that destabilize the helix (melting proteins). Such proteins are required to unwind the helix during replication and to facilitate interaction between single strands during genetic recombination.

The reassociation (renaturation) of thermally denatured DNA is a spontaneous process but only occurs if the solution is cooled slowly below the melting temperature. Renaturation can take several hours, depending on the size of the molecule, because it initially relies on random base pairing (hybridization); it is, however, a cooperative process. Rapid cooling of denatured DNA at salt concentrations > 50 mmol/L produces a very compact molecule in which about two-thirds of the bases are hydrogen bonded or stacked. At salt concentrations below

10 mmol/L the DNA remains denatured even after cooling.

The length of RNA varies greatly: tRNA has a length of 75 – 90 nucleotides and mRNA can be up to several thousand nucleotides long. Denaturation effects are rarely observed because RNA has few truely double-stranded regions; it is most likely to be observed in tRNA.

Because they are extremely long, DNA molecules are extremely sensitive to mechanical influences (shearing forces, e.g., vigorous stirring) and easily break into small fragments (ca. 1000 base pairs). Ultrasonic treatment of DNA in solution produces fragments of ca. 100 – 500 base pairs owing to disruption of hydrogen bonds and single-strand and double-strand breaks in the sugar – phosphate backbone [74]. Nucleic acids are sparingly soluble in water (depending on the molecular mass). They are negatively charged and acidic at physiological pH and form water-soluble alkali and ammonium salts that can be precipitated with ethanol.

RNA and DNA are insoluble in cold acid. DNA is more sensitive to acid hydrolysis than RNA. At pH < 1, however, both DNA and RNA break down into the free bases, phosphoric acid, and (deoxy)ribose. Acid hydrolysis can be used to determine the base composition of nucleic acids (e.g., total hydrolysis can be achieved by heating DNA in 90 % formic acid at 180 °C for 30 min). The β-glycosidic linkage between the N-9 of purines and the C-1 of deoxyribose is selectively cleaved at ca. pH 4, resulting in apurinic sites. Anhydrous hydrazine cleaves the pyrimidine residues.

DNA is stable at pH 13, only 0.2 of 10^6 phosphodiester bonds are broken per minute at 37 °C. In contrast, RNA is rapidly hydrolyzed at alkaline pH.

DNA can be both specifically and nonspecifically cleaved by a variety of enzymes [deoxyribonucleases (DNases)]. RNA is cleaved by ribonucleases (RNases). Some of these cleavage reactions are exploited for sequencing RNA [75, 76].

3.2. Interaction with Proteins

In bacteria, DNA occurs as a complex with RNA and proteins that is bound to but not surrounded by a membrane. The DNA often has a closed circular form and is organized in a series of superhelical loops.

The DNA of higher cells is enclosed within the nuclear membrane as morphologically distinct units of varying size (chromosomes); it is associated with basic proteins called histones. The number and size of the chromosomes are species specific (karyotype). Two full turns of the DNA double helix (146 base pairs) are wound around a histone octamer (diameter ca. 8.6 nm) to form a nucleosome. The width of the grooves varies due to the periodic arrangement of A – T trinucleotides on the inside and G – C trinucleotides on the outside of the nucleosome at intervals of about ten base pairs [53]. Nucleosomes can become condensed into fibers of 10 or 30 nm (super superhelices, solenoids).

Eucaryotic cellular organelles (e.g., mitochondria, chloroplasts) possess closed circular DNA that is not associated with histones.

4. Biosynthesis and Biological Function

4.1. DNA Replication

The genetic information of all cellular organisms is stored in double-stranded DNA (viruses may, however, also have single-stranded DNA or RNA, as well as double-stranded RNA). It is extremely important that the transfer of biological information in DNA (i.e., its base sequence) occurs with a very high degree of accuracy. Because of perfected proofreading and repair mechanisms (see Section 4.5) DNA replication has an error level of $10^{-8} - 10^{-11}$ [77], i.e., for every $10^8 - 10^{11}$ bases in newly synthesized DNA only one is incorrectly incorporated. The replication of DNA is carried out by DNA polymerases which require a single strand of DNA as a template and a short double-stranded piece of nucleic acid (formed with the help of a primer) for initiation (Fig. 6; see also → Enzymes, Chap. 6.2.).

The DNA is synthesized from deoxyribonucleotide triphosphates which are polymerized on the single-stranded DNA template with the release of pyrophosphate; the cleavage of pyrophosphate by a pyrophosphatase provides the energy required for DNA biosynthesis. The addition of new nucleotides always takes place

at the 3′-OH group of the sugar, therefore all biologically synthesized nucleic acids grow in the $5' \rightarrow 3'$ direction. Divalent cations such as Mg^{2+} or Mn^{2+} are important cofactors. Some polymerases contain $3' \rightarrow 5'$ exonuclease activity and can remove bases that have been incorrectly incorporated. The primer in vivo is usually RNA, which is synthesized by a special RNA polymerase (primase) and removed later by an exonuclease (e.g., DNA polymerase with $5' \rightarrow 3'$ exonuclease activity). The various activities are combined in a multienzyme complex.

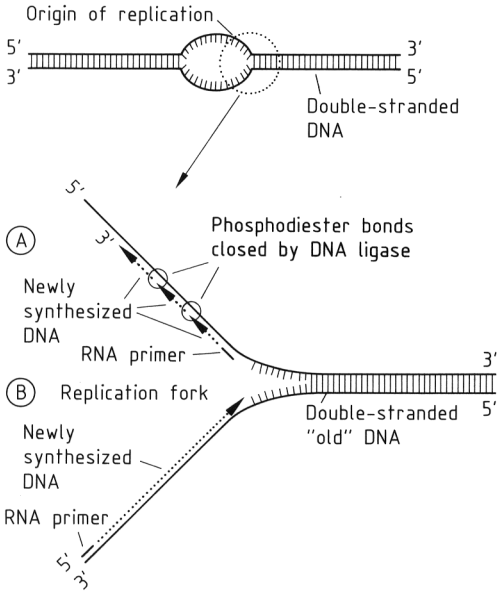


Figure 6. Simplified scheme of DNA replication
A) Lagging strand synthesis: after primer addition by primase a stretch of DNA is synthesized by DNA polymerase, the primer is removed and after synthesis of the adjacent stretch of DNA the phosphodiester bond between the two stretches is closed by a DNA ligase.
B) Leading strand synthesis: after primer addition by primase the new strand is synthesized by DNA polymerase. Arrows denote direction of synthesis.

The replication of DNA is a semiconservative process, i.e., one strand of each of the two new daughter molecules of DNA is an old strand and the other a newly synthesized one [78]. Replication starts with the creation of replication forks at the "origin of replication", and proceeds in opposite directions along the DNA (see Fig. 6).

The DNA of procaryotes is replicated from a single replication origin. The main replication enzyme in *Escherichia coli*, DNA polymerase III, consists of at least ten different subunits. Replication is probably a membrane-bound process [79].

Because DNA strands are antiparallel, one strand (lagging strand) must be synthesized "backwards" to keep the replication complex together. The direction of synthesis of this strand is contrary to the direction of movement of the replication fork and occurs discontinuously. Segments (Okazaki fragments) are formed which are then joined by a DNA ligase. The strand, that is synthesized "forward" is called the leading strand. Replication in eucaryotes starts at several points along the DNA and the replication complexes are fixed to the nuclear matrix [80]. Viroids are the smallest self-replicating structures known. These single-stranded, circular, protein-free RNAs are a few hundred nucleotides long and occur as pathogens in higher plants [81].

4.2. Gene Expression

The base sequence of DNA constitutes the genetic information in the form of discrete units (genes). A sequence of three successive bases (codon) acts as a code for a certain amino acid (e.g., GCT, GCC, GCA, or GCG code for alanine) or as a stop signal (TAG, TGA, or TAA) in protein synthesis. Of the 64 ($=4^3$) possible triplets, 61 code for amino acids and 3 specify stop signals. Since this genetic code is almost universal, eucaryotic proteins can be synthesized in procaryotes from eucaryotic DNA ($\rightarrow$ Genetic Engineering). Some organisms and organelles differ in a few codons from the universal code [82]. The genetic information contained in DNA is normally used as a template for the synthesis of mRNA which itself then acts as a template for protein synthesis.

The regulation of gene expression in living organisms is based primarily on the recognition of nucleic acid sequences by proteins. However, nucleic acids can also be involved. "Antisense RNA" for instance, can hybridize with mRNA, leading to a block of translation and to the degradation of the mRNA by RNase. The termination of transcription can occur in a similar way. This principle may possibly be exploited for therapeutic applications [83].

4.2.1. Transcription

The process by which genetic information is transferred from DNA to mRNA is called transcription. This process is catalyzed by DNA-dependent RNA polymerases and involves the synthesis of mRNA from ribonucleotide triphosphates in the presence of Mg^{2+} or Mn^{2+} and a single- or double-stranded DNA template. The direction of synthesis is $5' \rightarrow 3'$ and no primer is required. As in DNA synthesis, energy is obtained from the cleavage of pyrophosphate. The accuracy of transcription is lower than that of DNA replication (ca. 10^{-4}) because there are no repair processes [77]. Most RNA polymerases are complex enzymes consisting of several subunits. Bacteria generally have one, but eucaryotes have three.

Due to the specific base pairing in the double strand, the base sequence information of DNA determines the base sequence of the mRNA synthesized by RNA polymerases. Messenger RNA acts as an intermediate for conveying the information required for protein synthesis from the DNA to the protein-synthesizing structures, i.e., the ribosomes (see below). For interaction with the ribosome, mRNA possesses a ribosomal binding site shortly before the translation initiation sequence. This site is complementary to rRNA sequences [84]. The average length of *Escherichia coli* mRNA is about 1200 base pairs, eucaryotic mRNA can be much longer. Some types of mRNA are very rapidly degraded and exist in relatively low concentrations. Other types are more stable and accumulate in considerable amounts in the cell. The ends of mRNA are important for its stability.

Modifications in the form of loop structures [85] in procaryotes, and poly(A) at the $3'$-terminus and a cap (formed by addition of 7-methyl-GTP followed by methylation) at the $5'$-terminus in eucaryotes protect the ends from nonspecific degradation by nucleases (Fig. 7). Capping and the addition of poly(A) usually occur before splicing. Random degradation of mRNA is also prevented by cellular ribonuclease inhibitors. The mRNA transcript is finally exported to the cytoplasm where it acts as a template for protein synthesis.

Transcription is increased or decreased by sequence elements on the DNA that act as binding sites for protein factors [86–88]. Transcription also involves initiation and termination factors, some of which are ribonucleoproteins. Structural features also affect transcription, e.g., bent DNA can activate the process [52].

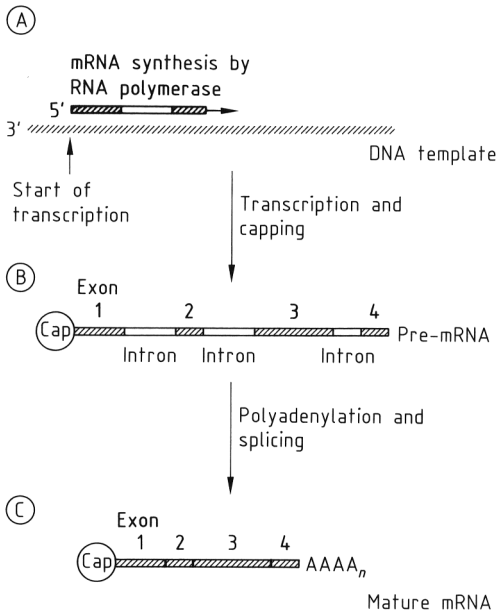


Figure 7. Simplified scheme of transcription in eucaryotes
A DNA template is transcribed into pre-mRNA by RNA polymerase (A) and a cap is added to the $5'$-end of the mRNA (B). A stretch of A-residues is then added to the $3'$-end. During splicing intervening sequences (introns) are removed and the sequences carrying the information for protein synthesis (exons) are joined to form mature mRNA (C).

RNA polymerase binds to a promoter region in the DNA template and then transcribes the DNA into mRNA. In most eucaryotic genes protein-coding sequences (exons) are interrupted by intervening noncoding sequences (introns). Many exons code for protein domains; exon rearrangements in DNA can facilitate the exchange of domains (exon shuffling). Lower eucaryotes have a higher content of intron-free genes than higher eucaryotes. Introns are rarely found in bacterial genes [89]. Introns are removed from eucaryotic precursor mRNA and the exons are then joined to give mature mRNA in a process known as splicing (see Fig. 7). The length of introns varies from 30 to 100 000 base pairs, the average exon length is 50 – 300 base pairs. Splicing demands a high degree of accuracy and is carried out by a multicomponent complex known as a spliceosome that consists

of protein and small RNAs (snRNAs = small nuclear RNAs). The main components of the spliceosome are ribonucleoproteins (snRNPs = small nuclear ribonucleoproteins) [90,91]. Introns usually start with GU (5′-splice site) and end with a pyrimidine-rich section followed by AG (3′-splice site) [92]. The half-life of introns in mRNA varies from a few seconds to 10–20 min. Splicing always takes place in the nucleus. Unspliced RNA remains in the nucleus and is degraded.

The splicing mechanism is highly conserved: a mammalian cell extract can process yeast RNA. The number of possible transcription products can be increased by alternative splicing (i.e., not all available exons of a gene are used to create one mature mRNA).

In protozoa RNA is also "edited", the primary transcript is converted into its functional form not only by post-transcriptional removal of certain bases, but also by insertion of others [93,94].

In some RNA viruses (retroviruses), replication of the genetic material proceeds via a DNA intermediate in a process known as reverse transcription. The information in RNA is transcribed back to DNA with suitable enzymes (reverse transcriptases). The DNA is then used as a template for RNA synthesis. Reverse transcription can also be used in vitro to produce intron-free complementary DNA (cDNA) from eucaryotic mature mRNA. Telomerases are special reverse transcriptases with an internal RNA template [95,96].

4.2.2. Translation

Protein synthesis takes place on the ribosomes, which are complexes of rRNAs and more than 50 different proteins. Protein factors are involved in all phases of translation. Various types of RNA are involved in protein synthesis: mRNA, the carrier of genetic information from the DNA; rRNA, component of ribosomes and directly involved in almost all stages of protein synthesis [97]; and tRNA, which makes activated amino acids available for protein synthesis. Translation is an irreversible process with error rates of 10^{-3}–10^{-4} [77].

The initiation of protein synthesis is a very intricate multistep process [98,99]. A translation initiation complex is formed from ribosomes, mRNA, a special tRNA, initiation factors, and GTP. The mRNA start codon for protein biosynthesis is usually AUG but GUG, UUG, or AUU may also be used [89]. After initiation of translation, the protein is synthesized step by step (elongation) from amino acids that are delivered and activated by tRNA. Each amino acid has a specific tRNA.

The tRNAs act as adapters between the codons of mRNA and the corresponding amino acids. Apart from the "usual" bases, tRNA contains several rare bases (e.g., 5-hydroxymethylcytosine). Transfer RNAs have an amino acid attachment site and a template recognition site (anticodon) for the mRNA. Amino acids are attached by an ester bond either to the 3′-or 2′-OH group of the terminal adenosine of tRNAs. Loading the tRNAs with the correct amino acids is crucial for exact translation of genetic information. This step is catalyzed by specific enzymes (aminoacyl-tRNA synthetases). Each of these enzymes must be able to recognize a specific amino acid and its associated tRNA.

The amino acid sequence of the protein product is specified by the mRNA codons and thus ultimately by the base sequence of DNA. A stop codon (TAA, TAG, TGA) on the mRNA leads to termination of protein synthesis and the release of the protein. Differences between translation complexes in procaryotes and eucaryotes permit the selective interruption of procaryotic protein synthesis (e.g., by some antibiotics).

4.3. Modification and Degradation

DNA is exposed to many damaging influences and can be modified and cross-linked by radiation and chemicals. Spontaneous hydrolysis of DNA bases may also occur.

Mutations change the normal base sequence of DNA. Erroneous bases can appear as a result of replication errors or chemical modification (→ Mutagenic Agents). Intercalating agents such as benzopyrenes, aflatoxins, or ethidium bromide can cause deletions. Insertion of individual bases is also possible. Insertion of larger DNA segments can occur through mobile DNA elements (transposable elements, jumping genes). The insertion site is usually arbitrary.

This process (transposition) can lead to gene activation or inactivation.

Methylation of DNA is used by cells in different ways [100, 101], e.g., to distinguish between newly synthesized strands (unmethylated) and old strands (methylated). Methylation occurs at defined sites: in bacteria at the N-6 position of adenosine and the C-5 position of cytosine, and in eucaryotes, physiologically only at cytosine. The methylation pattern also differentiates between endogenous and foreign DNA and is recognized by restriction – modification systems (in procaryotes): endogenous DNA is methylated at specific sites and foreign DNA is digested. Methylation can be used to switch genes on or off in eucaryotes. C-Methylation influences the equilibrium between B- and Z-DNA [102] and can lead to altered DNA – protein interactions by the modulation of DNA topology. Methylation patterns are stable and hereditary. Methylation can have a mutagenic effect: the deamination of 5-methylcytidine to thymine occurs spontaneously.

Degradation of nucleic acids can be effected in vivo and in vitro by a series of enzymes. Some DNases are specific for single strands and others for double strands, some cleave bonds located within a double strand (endonucleases) and others require a free DNA end (exonucleases). Exonucleases act either in the $3' \rightarrow 5'$ or the $5' \rightarrow 3'$ direction; exonuclease activity is often associated with DNA polymerases for proofreading in DNA replication. Restriction endonucleases are a special class of nucleases. RNases can also be divided into exo- and endonucleases. RNase II is a bacterial exonuclease; RNase H is an endonuclease that digests RNA in DNA – RNA hybrids.

4.4. Recombination

In recombination two DNA molecules interact and exchange segments to produce two new molecules that contain genetic information from both parental DNAs. DNA recombination is by no means limited to genetic engineering, but occurs widely in nature. This process is catalyzed by enzymes and usually requires the presence of homologous regions (homologous recombination). In rarer cases, nonhomologous, illegitimate recombination can take place. Transposition (see Section 4.3) is also a recombination phenomenon [103].

4.5. DNA Repair

Although changes in base sequences create the genetic variety required for evolution, these changes are only rarely beneficial for individual organisms. Despite many damaging influences (see Section 4.3), nucleic acids are replicated with amazing accuracy. This exact conservation of genetic material is one of the fundamental requirements for genetic success. Damaged DNA that is not repaired can lead to mutations and cell death, and give rise to tumors. Living organisms have therefore developed a variety of enzymatic mechanisms for eliminating DNA damage. Most damage is recognized by repair systems in the cell and repaired either by directly reversing the damage or by using the undamaged second strand as a template. *Escherichia coli* has been used as a model for studying repair mechanisms [104]. The best studied mechanism is nucleotide excision repair. This multistep process involves recognition of the error, nicking of the damaged strand, removal of nucleotides at the damaged site, repair DNA synthesis, and ligation. It requires the participation of multienzyme complexes [105, 106].

4.6. Nucleic Acids as Enzymes

Ribozymes are RNA molecules that act as catalysts: sequence-specific RNA endonucleases, nucleotidyl transferases, phosphatases, kinases, and glucanotransferases [37, 107]. The sequence-specific cleavage of oligodeoxyribonucleotides and single-stranded DNA can also be catalyzed by ribozymes [108]. Previously, only proteins (enzymes) were thought to have activities of this kind. Ribozymal properties were discovered in the early 1980s in the self-splicing introns of *Tetrahymena* [109] and ribonuclease P [110], which cleaves tRNA precursors and contains RNA as the catalytic subunit. In eucaryotes, small nuclear RNAs (snRNAs) are required for the splicing of pre-mRNA; they are part of the spliceosome, an RNA – protein complex [90].

The smallest RNA structure that exhibits catalytic (cleavage) activity is 19 nucleotides long

and forms a hammer head structure with its target [111]. No pure DNA catalysts have been discovered yet [112]. However, since a few of the ribonucleotides in ribozymes can be replaced by deoxyribonucleotides, the three-dimensional structure that permits nucleic acid catalysis may not necessarily be limited to RNA [113, 114].

In all proven examples of RNA catalysis, the substrate attacked by the ribozyme is itself a nucleic acid. However, there are indications that rRNA also attacks other molecules (i.e., is an RNA enzyme). RNase P is the only known naturally occurring ribozyme that does not cleave itself, but attacks other molecules [72, 115].

5. Isolation, Purification, and Transfer

Chromosomal DNA from bacteria or eucaryotes has a very high molecular mass and is extremely fragile. Thus to isolate DNA, cells must be gently lysed with enzymes (pronase, protease K, lysozyme) and mild detergents. Accompanying RNA can be removed by treatment with enzymes which digest RNA or by gel filtration [116]. Proteins associated with DNA are denatured with phenol and extracted [117].

Plasmid DNA is relatively small and has a closed circular and supercoiled form. It is more stable than chromosomal DNA when subjected to shearing forces and can be renatured much more easily after denaturation; isolation and purification of plasmid DNA is achieved relatively easily (e.g., by equilibrium centrifugation in cesium chloride – ethidium bromide gradients at 150 000 – 400 000 g for 5 – 48 h). Kits are also available which permit the isolation of pure plasmid DNA in a short time without ultracentrifugation using binding to glass powder (glass milk) and/or gel filtration. Similar methods are used for the *isolation of RNA*. However, degradation by ribonucleases must be prevented (RNA is much more sensitive to nucleases than DNA). Either proteases or chloroform are employed to remove protein; with chloroform protein precipitates at the phase boundary. Phenol is not normally used because a small part of the RNA also goes into the phenol phase. Oligo(dT) cellulose (cellulose-bound oligonucleotides containing only thymine) can be used for the purification of eucaryotic mRNA [118] taking advantage of its poly(A) tails.

Nucleic acids can also be purified by ion-exchange chromatography (e.g., by HPLC) [119]. Suitable separating media are hydroxyapatite and anion exchangers. Reversed phase methods and gel filtration are also used.

Microgram amounts of DNA fragments can easily be isolated and purified by techniques such as electroelution from agarose gels, binding to anion-exchange paper, dialysis, or binding to glass milk (powdered glass suspension). After purification, DNA can easily be concentrated by precipitation with ethanol or 2-propanol.

DNA can be transferred relatively easily in vivo between related organisms (especially in bacteria) via endogenous mechanisms (e.g., conjugation, some viruses). Several other methods are available for the transfer of DNA in the laboratory: application as a $CaCl_2$ precipitate, exposure to an electric field (electroporation), and the use of viral vectors (see also → Genetic Engineering).

6. Analysis of Nucleic Acids

Sequencing. Computer data banks such as EMBL [120] and GenBank [121] contain sequence information in the order of millions of nucleotides.

Single or double-stranded DNA can be sequenced [122]. The most widely used methods are the Sanger method [123] and the Maxam and Gilbert method [124] (→ Genetic Engineering, Chap. 3.3.). In the Maxam and Gilbert method the DNA is cleaved with chemicals statistically and specifically at one of the four bases. The Sanger method employs enzymatic synthesis of a complementary copy of a single-stranded DNA template and random chain termination with dideoxynucleotides. Both methods produce DNA fragments that are separated by gel electrophoresis.

Both methods initially made use of radioactively labeled nucleotides and subsequent exposure on X-ray films for identification of the fragments. More recent variations use fluorophores [fluorescein, tetramethylrhodamine, Texas red, 4-chloro-7-nitrobenzo-2-oxa-1-diazole (NBD)] which are more stable, easier to use, less dangerous, and permit the automatic detection of

all four nucleotides on a single track. Sequences of > 450 bases can be determined per gel run and track [125–128]. An automatic sequencing machine allows determination of sequences of up to 10 000 bases per day. Further acceleration by combination with automatic DNA preparation in robotic workstations may be possible [129].

More recent techniques aim at faster sequence determinations (several hundred bases per second [130]) with smaller amounts of DNA. Variations of the Sanger method permit determination of sequences up to 10 000 bases without subcloning [131].

The Sanger method can also be used for sequencing RNA [132] and sequences up to 150 bases can be determined in a single run. Reverse transcriptase is used to synthesize DNA with RNA as a template. The RNA sequence is then deduced from the DNA sequence. Another method for RNA sequencing involves the use of specific RNases.

X-Ray Analysis. X-ray analysis is of considerable importance for elucidating the structure of nucleic acids and their building blocks [44]. Useable crystals can be obtained from aqueous solutions for molecules up to the size of tRNA (75 – 90 base pairs). Nucleic acids of higher molecular mass are too heterogeneous to produce crystals, quasi-crystalline fibers drawn from concentrated aqueous solutions are used instead.

Nuclear Magnetic Resonance (NMR). The best method for determining the structure of relatively short (10 – 20 bases or base pairs) DNA or RNA fragments in solution is ^{1}H-, ^{13}C-, ^{15}N-, or ^{31}P-NMR (one- or two-dimensional) [133, 134]. NMR is also used to study short triple helices [135], DNA – RNA hybrids in solution [136], DNA – drug complexes, mismatches, and intercalating agents. Torsion angles and the mobility of base pairs and the sugar phosphate backbone can be determined.

Electron Microscopy. Electron microscopy permits the determination of the length of DNA. DNA heteroduplexes, binding of large protein factors [137], and supercoiled DNA [138] can be visualized. A special application is DNA-hybridization electron microscopy [139, 140]: biotinylated DNA probes are hybridized to selected regions of DNA or RNA and then reacted with avidin and visualized by electron microscopy.

Scanning tunneling microscopy (STM) allows the observation of single-stranded DNA adsorbed on graphite with atomic resolution. It opens up the possibility of structural studies on DNA and of direct sequencing [141–143]. The profiles show excellent correlation with the results of X-ray analysis.

Centrifugation. Ultracentrifugation in cesium chloride gradients can be used to measure the molecular mass of DNA and also for the experimental identification of covalently closed circles. Sedimentation at alkaline pH (> 11.3) causes disruption of the hydrogen bonds and unwinding of the DNA molecules. In the case of linear DNA, two single strands are obtained which have a higher sedimentation coefficient than native DNA at salt concentrations > 0.3 mol/L. Covalently closed circles cannot be denatured into single strands, they collapse and have a sedimentation coefficient that is three times higher than that of native DNA.

In equilibrium centrifugation in cesium chloride, ethidium bromide is firmly bound to DNA in salt solutions by intercalating between base pairs. This lowers the density of the DNA by about 0.15 g/cm^3 and the DNA unwinds with increased binding of ethidium bromide. Unlike linear DNA, covalently closed molecules have a limited absorption capacity for ethidium bromide. Linear DNA is consequently lighter than covalently closed circular DNA. Use of a cesium chloride density gradient with ethidium bromide therefore permits the separation of supercoiled and linear (relaxed) DNA of the same size.

Staining and Quantification. Nucleic acids can be stained with ethidium bromide [144], Stainsall, 1-ethyl-2-[3-(1-ethylnaphtho[1,2 *d*] thiazoline-2-ylidene)-2-methyl-1-propenyl]-naphtho[1,2 *d*] thiazolium bromide [145], or silver [146–148]. DNA stained with ethidium bromide can be quantified in electrophoresis gels after being photographed [144, 149]. Silver staining followed by a densitometer scan is equally sensitive (1 – 2 ng of DNA can be detected) [146]. A quick and sensitive estimation of the amount of nucleic acid can be obtained by

measuring the UV absorption at 260 nm: an absorption of 1.00 corresponds to a concentration of ca. 50 µg/mL of double-stranded DNA.

DNA quantification in the picogram range has been reported with a silicon sensor, that responds to the interaction between DNA-binding proteins and DNA [150]. Quantification in the picogram range is also possible after specific hybridization (see Chap. 8).

Fluorescence imaging of DNA in picogram to nanogram amounts is made possible by laser excitation of a DNA – ethidium homodimer [5,5′-diazadecamethylene) bis(3,8-diamino-6-phenylphenanthridinium) dichloride dihydrochloride] complex at 488 nm [151].

Electrophoresis. Since DNA is negatively charged because of its phosphate groups, it migrates in an electrical field (usually at a pH of ca. 7). The electrophoretic mobility of DNA fragments in solution is independent of the molecular mass (constant linear charge density). However, electrophoresis of DNA in gel matrices (agarose or polyacrylamide) at a constant field strength is an effective method of separating DNA molecules according to molecular mass. The mechanism of separation is not well understood [152].

Important experimental parameters are gel concentration, field strength, temperature, and running time. Gels of 0.5 – 2 % are usually employed for the separation of 50 – 20 000 base pairs at field strengths of 2 – 3 V/cm [153]. The mobilities of linear and supercoiled DNA of equivalent molecular mass are different. Estimation of the molecular mass of supercoiled DNA molecules based on linear standards is therefore not possible.

Polyacrylamide gels (4 – 12 %) give a sharp separation of polynucleotides up to a length of about 600 base pairs. Special gels can separate polynucleotides with up to 2500 base pairs [148]. Mobility varies logarithmically with the size of the DNA at field strengths below 1 V/cm [154]. Separation efficiency is limited in the range of 10 000 – 50 000 base pairs, there is no resolution above 50 000 base pairs.

Denaturing gels (e.g., with urea) are used in the sequencing and analysis of single-stranded nucleic acids. Under special conditions, the behavior of individual DNA molecules in an electric field can be observed with a fluorescence microscope [155].

Pulse-field electrophoresis is used to separate larger linear DNA molecules. Here the direction of the electric field is changed at intervals. The DNA molecules therefore have to change their orientation, the time required for this change depends on the size of the DNA [44, 156]. Unlike conventional electrophoresis (duration 30 – 60 min), a pulse-field run takes 20 – 140 h. Usually, 1.5 % agarose gels are used and 10 – 20 µg of DNA can easily be separated at field strengths of 2.5 – 10 V/cm. Separation can be achieved up to ca. 12×10^6 base pairs [157]. Separation is affected by DNA topology and sequence, pulse time (seconds to minutes), field geometry, field strength, gel composition, sample concentration, temperature, and running time [154, 158].

7. Chemical Synthesis

The chemical synthesis of oligonucleotides began in 1955 when A. M. Michelson and A. R. Todd prepared the first dinucleoside monophosphate [159]. They used the so-called *triester method* to show that oligonucleotides of defined sequence can be prepared by purely chemical methods. In 1956 Khorana et al. successfully synthesized the same nucleotide by the *diester method* [160]. This process dominated oligonucleotide synthesis until the mid-1970s, and was also used by Khorana and coworkers for the first total synthesis of a gene (alanine-specific tRNA from yeast) [161]. The importance of synthetic oligonucleotides was generally underestimated because the synthesis was time-consuming and expensive. It was the progress in genetic engineering that resulted in a dramatic increase in the demand for synthetic oligonucleotides, which, in turn, led to the development of new synthesis concepts and to the automation of DNA and RNA synthesis.

7.1. Synthesis Strategy

The chemical synthesis of DNA consists of three steps:

1) Separate synthesis of the two complementary strands

2) Hybridization of the two strands (formation of hydrogen bonds between A and T, and between C and G)
3) Enzymatic joining of the DNA molecules to give larger DNA units

Three standard methods are used for the synthesis of the 3′,5′-internucleotide phosphodiester bond. They differ in the type of monomer building blocks used and are known as the phosphate triester, phosphoramidite (phosphite), and the *H*-phosphonate methods (Fig. 8, Section 7.4).

An oligomer can be synthesized either by stepwise addition of individual monomer building blocks or by joining oligomer blocks (e.g., dimers or trimers). The former concept is employed in solid-phase synthesis, which was developed almost simultaneously by R. B. MERRIFIELD [162] for peptides and by R. L. LETSINGER [163] for oligodeoxyribonucleotides. The first nucleoside (sugar and base) is covalently bound to an insoluble support at the 3′-OH group of the sugar. Stepwise synthesis of the chain then proceeds by condensation of the appropriate nucleoside monomers. Since the reaction product remains immobilized on the support, it can easily be freed of other reactants by washing; time-consuming chromatographic purification is not required. Synthesis is usually carried out on the 0.2 – 1 μmol scale and all reagents are applied in large excess. Solid-phase

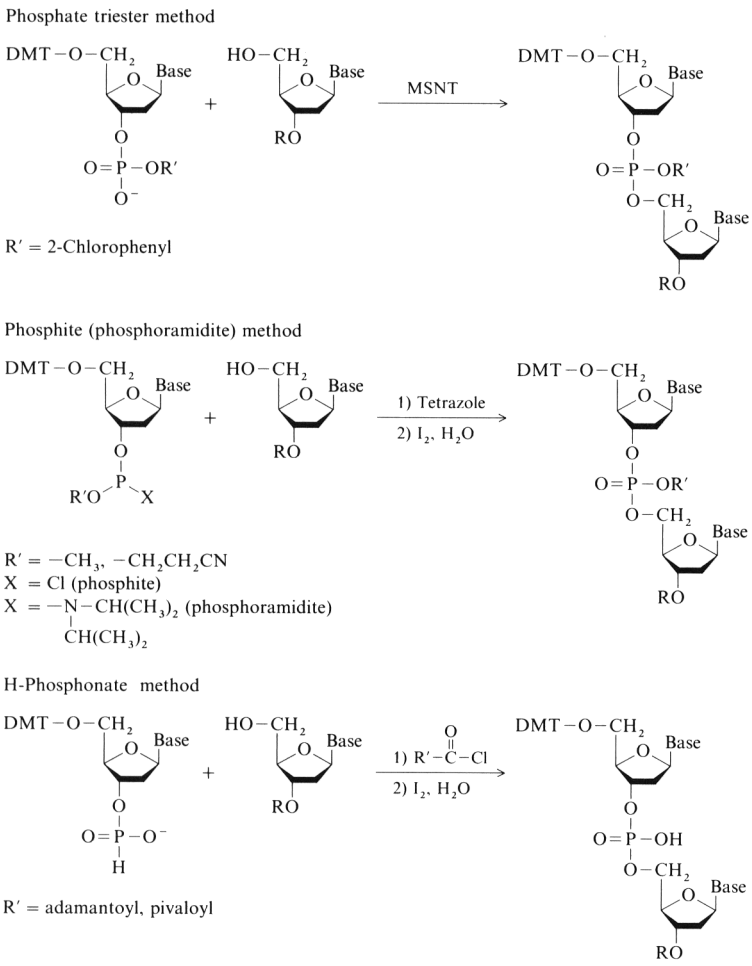


Figure 8. Chemical synthesis of oligodeoxynucleotides
DMT = 4,4′-dimethoxytrityl residue; MSNT = mesitylene sulfonylnitrotriazole; R = polymer support

synthesis of oligonucleotides can be divided into the following steps:

1) Preparation of monomer units that are either protected or can be activated
2) Functionalization of the support
3) Stepwise synthesis of the desired nucleotide sequence by one of the three methods listed above
4) Cleavage of the protecting groups and the support
5) Isolation and purification of the desired sequence

7.2. Protecting Groups

In order to avoid unwanted side reactions during oligonucleotide synthesis, functional groups that do not participate in the linking reaction must be suitably protected. Only a few protective groups have gained acceptance and are based on the work of KHORANA et al. [164].

Two types of groups can be distinguished: permanent and intermediary protecting groups.

Permanent protecting groups are maintained throughout the synthesis and protect the exocyclic amino functions of adenine, cytosine, and guanine; the OH group of phosphorus; and the 2′-OH group of ribose (in RNA synthesis). *Intermediary protecting groups* protect the 5′-OH group of the sugar and are cleaved after each condensation step.

The most important protecting groups are depicted in Figure 9. Adenine and cytosine are usually converted to the corresponding acid amides with excess benzoyl chloride, the amide of guanine is obtained with isobutyric anhydride. Base-protecting groups are introduced with a method developed by TI et al. (transient protection) [165]. After termination of the synthesis, the protecting groups are removed with 32 % aqueous ammonia at 50 – 60 °C in 4 – 5 h. In long, guanine-rich sequences, the O-6 position of guanine can also be protected [166–169], usually with a 4-nitrophenylethyl residue [170]. This is especially important for attaining good yields in the triester process because otherwise sulfonation with the condensing agent easily occurs. More labile amidine or acyl base-protecting groups are of interest for the synthesis of phosphate-modified oligonucleotides because they are more sensitive to hydrolysis than their diester analogues and treatment with ammonia at 60 °C can cause considerable strand breakage [171–173]. Important examples of such groups are the 4-phenoxyacetyl residue (Pac) for adenine and guanine and the isobutyryl residue for cytosine.

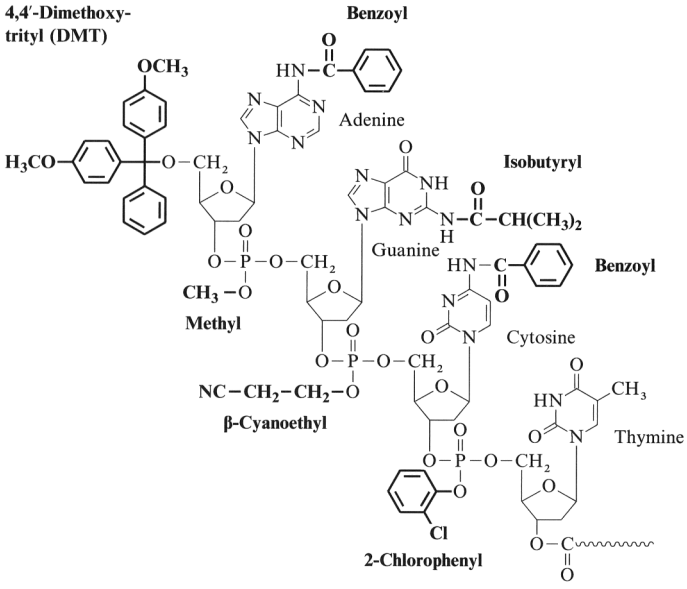


Figure 9. Protecting groups (bold type) used in the synthesis of oligonucleotides

The *tert*-butyldimethylsilyl residue is preferred as the protecting group for the sugar 2′-OH group in the solid-phase synthesis of oligoribonucleotides [174].

The 5′-OH group of deoxyribose or ribose is usually blocked with the 4,4′-dimethoxytrityl group [164], which can easily be removed after each coupling step by mild acid hydrolysis with 3 % di- or trichloroacetic acid in dichloromethane [175].

In the triester method the 2-chlorophenyl group is used to protect the phosphorus [176], while in the phosphoramidite method, the β-cyanoethyl group is preferred to the originally used methyl group [177, 178]. The β-cyanoethyl group can easily be removed at the end of the synthesis by β-elimination with aqueous ammonia [179]. No internucleotide phosphorus-protecting group is required in the *H*-phosphonate process.

7.3. Functionalization of the Support

Polystyrene [180], polyacrylamide [181], cellulose [182], and silica [183] have been used for solid-phase synthesis of oligodeoxyribonucleotides. The preferred support material is controlled pore glass (CPG) [184, 185] functionalized with a long-chain alkylamine spacer. The material is relatively rigid, does not swell, and is inert to all reagents used in the synthesis. In addition, it withstands the high mechanical stress encountered in the large number of reactions performed in automatic synthesizers. The narrow pore-size distribution gives more uniform diffusion parameters and, consequently, more reproducible flow rates and availability of functional support sites.

The support is functionalized with an alkylamine by silanization with 3-aminopropyltriethoxysilane [183, 186]. Heating and subsequent treatment with methanol result in crosslinkage of the silyl groups and hydrolysis of outer, less strongly fixed groups [187]. The nucleoside at the 3′-OH end of the DNA to be synthesized is linked to the support via a succinyl spacer [183, 188, 189]. Amino groups of the support that have not reacted are capped with acetic anhydride [188]. To determine the loading of the support with nucleoside material, an aliquot is treated with strong acid to remove the DMT residue. The resulting orange solution containing the dimethoxytrityl cation [190] can be determined by measuring the absorbance at 495 nm. The average loading is usually 15 – 25 µmol per gram of CPG.

7.4. Methods of Synthesis

The three main methods of oligonucleotide synthesis are the phosphate triester, the phosphoramidate (phosphite), and the *H*-phosphonate methods (see Fig. 8). The phosphoramidite method is the most widely used (Fig. 10, see next page). The reaction cycle involves

1) Cleavage of the 5′-OH DMT group
2) Condensation of the next monomer building block (chain elongation)
3) Masking of 5′-OH groups that have not reacted completely (capping)
4) In the phosphoramidite and *H*-phosphonate processes, an additional oxidation step is required after each condensation step or the end of the synthesis respectively to obtain the desired internucleotide phosphodiester.

Chain elongation in all processes occurs in the $3' \rightarrow 5'$ direction.

Phosphate Triester Method. This method was first used to synthesize oligonucleotides from suitable building blocks in solution [191–194]. It can also be employed for the solid-phase synthesis of nucleotides up to a chain length of about 20 bases. The monomer building block is a protected nucleoside phophodiester derivative (Fig. 8) [193] which reacts with the free 5′-end of a support-bound nucleoside to give a neutral phosphotriester. Various aromatic sulfonic acid derivatives [192, 195, 196], primarily mesitylene sulfonylnitrotriazole (MSNT), serve as condensing agents. The reaction can be accelerated by using catalysts such as *N*-methylimidazole [175]. This method is not widely used any longer due to disadvantages such as long reaction times, a large number of side reactions, and the frequently observed nicks caused by incomplete cleavage of the internucleotide protecting group.

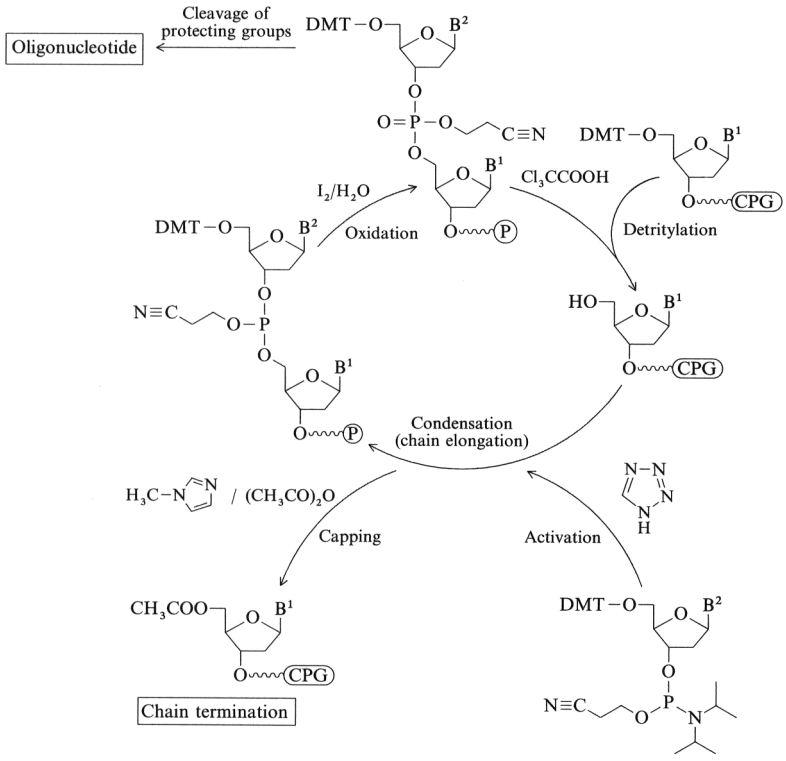


Figure 10. Synthesis cycle according to the phosphoramidite method
B = base; DMTr = 4,4′-dimethoxytrityl residue

Phosphoramidite (Phosphite) Method. This concept is similar to the triester method, but chain synthesis involves trivalent phosphite triester intermediates. The preferred monomer building blocks are the so-called nucleoside phosphoramidites introduced by CARUTHERS et al. [177, 178] that can be activated with tetrazole. After each condensation step, the phosphite triester is converted to the phosphate triester by oxidation with iodine solution. The exceptionally high reactivity of the activated phosphoramidites results in a condensation time of ca. 1 min and $\geq 99\,\%$ yield per chain elongation step.

***H*-Phosphonate Method.** This synthesis method was developed almost simultaneously by GARREGG et al. [197] and by FROEHLER and MATTEUCCI [198]. Condensation is achieved by activation of nucleoside 3′-*H*-phosphonate monomers with pivaloyl or adamantoyl chloride. Since the *H*-phosphonate bond is not cleaved during the synthesis cycle, only one oxidation step is required at the end of the synthesis. This considerably reduces the duration of the cycle. Another advantage over the phosphoramidite method is the higher stability of the monomers. Moreover, excess *H*-phosphonates can be regenerated after the synthesis, thus reducing costs; this makes this method particularly interesting for large-scale synthesis. Since the dinucleoside *H*-phosphonate bond can be easily attacked by nucleophilic reagents [199], oligonucleotide thiophosphates or phosphoamidate analogues can easily be prepared by oxidation with sulfur or amines [199, 200].

Synthesis of Oligoribonucleotides. The solid-phase synthesis of oligoribonucleotides can be conducted by the phosphoramidite [174] or the *H*-phosphonate method [201]. An additional protecting group is required for the 2′-OH group of the ribose [195]. The *tert*-butyldimethylsilyl group has gained acceptance for solid-phase synthesis with ribonucleoside phosphoramidites. It is stable under the acidic conditions required to remove the DMT group and can

be cleaved with tetrabutylammonium fluoride at the end of the synthesis. However, steric hindrance due to the *tert*-butyldimethylsilyl group results in yields of ca. 95 % per chain elongation step. Hence, this method is suitable only for the synthesis of oligomers up to a chain length of about 40 nucleotides.

7.5. Cleavage of Protecting Groups and Purification of Oligonucleotides

Cleavage of Protecting Groups. After synthesis, the oligomer is still completely protected and attached to the support. Cleavage of protecting groups should always begin with the internucleotide protecting group on the phosphorus to prevent strand damage [196, 202].

If the triester method is used, the 2-chlorophenyl group is eliminated by treatment with an "aldoximate" [203]; the base-protecting groups and the support are subsequently removed with 32 % aqueous ammonia at 50 °C. Finally, treatment with 80 % acetic acid cleaves the DMT group, provided it is not required for purification by reversed phase high performance liquid chromatography (HPLC).

If methyl phosphoramidites are used for synthesis, the first deprotection step is treatment with thiophenol to cleave the methyl group [7]. If, however, β-cyanoethyl phosphoramidites are used, then ammonia treatment at 50 °C is sufficient to remove all protecting groups including the support.

Purification of Oligonucleotides. Since condensation does not give 100 % yields per coupling step, short-chain homologues are present in addition to the target sequence. For example, for an oligonucleotide with 70 bases and a 98.5 % yield per condensation step, the target sequence accounts for ≤ 35 % of the total yield. A purification step is therefore required for longer oligomers [204]. High performance liquid chromatography (HPLC) or polyacrylamide gel electrophoresis is commonly used for product isolation. Purification by *reversed phase HPLC* is ideal for the routine separation of oligonucleotides up to ca. 40 – 50 nucleotides [204, 205]. Here, the target sequence that is still tritylated is eluted last. Separation efficiency can be increased with special high-affinity protecting groups allowing purification of chains with 150 bases [206]. If the trityl group is removed before purification, the polyanionic character dominates. It is then more advantageous to separate the desired product by ion-exchange chromatography. *Polyacrylamide gel electrophoresis* under denaturing conditions is the most efficient method for oligonucleotide separation [207, 208]; the homologues are separated according to size. The target sequence, which is usually the longest sequence, has the lowest mobility. The individual fragments become visible on exposure to UV light (λ 254 nm) and are cut out of the gel matrix. Extraction with buffer or electroelution and subsequent desalting (dialysis or gel filtration) yield the desired oligonucleotide.

7.6. Synthesis of Modified Oligonucleotides

Oligonucleotides bind specifically to a defined target sequence under suitable hybridization conditions. Relatively short oligonucleotides (< 20 bases) are used for new drug design strategies involving targeting interference of genetic expression at the level of transcription or translation [209].

For in vivo chemotherapeutic applications based on sequence-specific hybridization (antisense inhibition), nuclease-resistant oligonucleotides are required (the most important are shown in Fig. 11, see next page). They can be obtained by modifying the phosphate backbone of the oligonucleotides, by using α-anomers (**10**) or 2′-O-alkylribosides (**9**). In spite of the changes these oligonucleotides still hybridize with the target sequence and they can also pass through cell membranes more easily due to their lipophilic nature. Applications of these analogues as antisense oligonucleotides are summarized in [209–211].

Phosphothioate analogues (**1**), in which a nonbridging atom of the phosphate group is replaced by a sulfur atom either at a single specific position or at all positions within the chain, can be synthesized by following the phosphoramidite or *H*-phosphonate route [212–214]. The sulfurization reaction is usually carried out with a solution of sulfur in pyridine and carbon disulfide [215]. Other sulfurization reagents that

can also be used in the synthesis are described in [216,217].

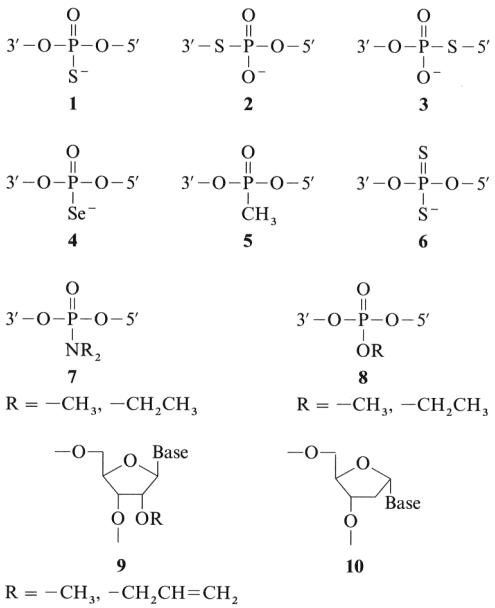


Figure 11. Modified oligonucleotides

Like native DNA, phosphorodithioate analogues (**6**) are stereochemically uniform products and can be prepared by the phosphoramidite method with the corresponding thiophosphoramidite monomers [218].

Methylphosphonates (**5**), phosphoramidates (**7**), and phosphotriesters (**8**) are nonionic and have a more hydrophobic character than naturally occurring nucleic acids. The synthesis of oligonucleoside methylphosphonates is described in [219–222], and is best carried out with nucleoside 3′-methylphosphonous acid imidazolide, which in turn is prepared from methylphosphonous acid bisimidazolide and the fully protected nucleoside. Alternatively, a *H*-phosphonate is reacted with *tert*-butyldimethylsilyl chloride to give silyl phosphite and subsequently oxidized with alkyl or aryl halides [223]. All of the early investigations on the biological activity of these oligonucleotide analogues (which were named "matagen", an acronym for mashing tape for gene expression) were conducted by Ts' O and MILLER [222].

Phosphoramidates (**7**) can be prepared by reacting *H*-phosphonate oligonucleotides with a solution of the appropriate amine in carbon tetrachloride [213].

The 2′-O-methylribonucleotides (**9**) are well established modified nucleosides that are used as monomer building blocks for oligonucleotide synthesis [224]. The synthesis of the corresponding nucleoside phosphoramidites is described in [225]. Oligonucleotides with α-2′-deoxynucleosides (α-DNA) form β-DNA with natural DNA. This form has parallel strands and is also nuclease resistant [226,227].

To increase the binding of an oligonucleotide to the target, an intercalating agent (e.g., 2-methoxy-6-chloro-9-aminoacridine [228,229] or phenazine [230]) can be covalently attached to the oligonucleotide. These agents insert "internally" between the base pairs of the double helix, increasing duplex stability.

A second method involves binding of photoreactive groups (e.g., psoralen [231–233], azidophenacyl or azidoproflavin derivatives, proflavins, and porphyrins [234–238]) that are activated by UV or visible light and can crosslink with the target DNA or RNA. "Sequence-specific artificial endonucleases" [239] are oligonucleotides that bear a group which can induce specific cleavage of the target nucleic acid after binding to the target. Such groups are the above-mentioned photoreactive groups: subsequent treatment with piperidine allows cleavage of the target sequence at the crosslinking site. Direct light-induced cleavage at neutral pH is achieved with ellipticine and diacapyrene derivatives [238]. Metal chelates attached to oligonucleotides (e.g., iron – EDTA [240–242], copper – phenanthroline [243–246], or iron – porphyrin [247]) lead to relatively specific cleavage of the target sequence. A nonspecific DNase or RNase bound to a complementary oligonucleotide can also destroy the target nucleic acid [248,249].

Nonradioactive labeling of synthetic oligonucleotides has become an important technique in molecular biology and several chemical procedures have been developed. Oligonucleotides are covalently attached to "marker groups" such as fluorescent dyes, biotin, and other biologically active groups (see also Section 8.2). These reporter groups can be attached to an oligonucleotide at the 5′-OH terminus, the 3′-OH terminus, the phosphate backbone, or one of the bases. Attachment at the 5′-OH terminus via pro-

tected aminoalkyl or thioalkyl phosphoramidites is most common. These amino- or thiolinkers are added in the last synthesis step and produce a free amino or thiol terminal after cleavage of the protecting groups [250–252]. Subsequent reaction with the activated marker molecule (e.g., biotin-*N*-hydroxysuccinimide) gives the labeled oligonucleotide.

A large number of bifunctional linkers (e.g., maleimidohexanoyl-*N*-hydroxysuccinimide ester) have been described which allow coupling of enzymes [253], peptides, and proteins to linked oligonucleotides via amino groups of the amino acids or thiol groups of cysteine. Attachment at the 3′-OH terminus can be carried out by linking a ribose unit with T 4 RNA ligase, followed by periodate cleavage of the ribose ring, and subsequent reductive amination [254, 255]. Labeling via a 3′-terminal thiol group is described in [256]. Also a modified support is commercially available which produces an aliphatic primary amino group at the 3′-terminus [257]. At the internucleotide bond reporter groups can be attached, e.g., via a phosphoramidate linkage [258–260].

The attachment of reporter groups to modified bases usually occurs at the C-5 position of pyrimidines or the C-8 position of purines [261]. During oligonucleotide synthesis, a fully protected deoxyuridine phosphoramidite [262] modified with a trifluoroacetylaminopropenyl residue is inserted at the desired position in the sequence (usually at C-5); after cleavage of the protecting groups, the group to be coupled is bound via the resulting amino function. Numerous other labeling positions have been described (e.g., the C-4 position [263–265] of 2′-deoxycytidine and the N-6 position of the adenine ring [266] or the exocyclic amino function of the guanine residue [267]).

8. Uses

The direct use of nucleic acids in *nucleic acid probe technology* is becoming increasingly important. Here, labeled nucleic acids are used as probes to detect other nucleic acids by means of hybridization [21, 23]. Identification of the specific nucleotide sequence of a pathogen (virus, bacterium, parasite) can be used to diagnose an infection [268, 269]. More complex analyses are often necessary for the detection of changes in nucleic acids (point mutations, deletions, insertions, translocations, expression level or number of copies of the genetic information) in the diagnosis of hereditary diseases [24], or in the investigation of tumors [25, 26]. This chapter will be limited to describing the uses of nucleic acids in probe technology but other applications will first be briefly summarized.

The second important field of application for nucleic acids is *genetic engineering* (→ Genetic Engineering) which allows the production of proteins in bacteria, fungi, animal or plant cell cultures, or in animals or plants. The nucleic acid that codes for the desired protein is introduced into the organism and expressed.

Antisense technology

is a third field of application of nucleic acids which, however, still only exists in model systems. This technology aims to prevent the expression of nucleic acid sequences that have a negative effect on an organism. This can be achieved by the interception and binding of mRNA with synthetic antisense oligonucleotides (see Section 7.6) or with an antisense RNA from a recombinant expression system. The mRNA forms double strands with the antisense RNA by specific base pairing and cannot be translated [270, 271]. Another method of stopping expression is the specific cleavage of the target nucleic acid by catalytic RNA (ribozymes) [270, 271]. The target nucleic acid is recognized by specific base pairing with the ribozyme. Antisense technology may be used to treat infections caused by viruses, bacteria, or parasites by cleaving or preventing the expression of the genetic information of these pathogens. Another possibility is the treatment of diseases resulting from overexpression of an endogenous gene or its expression in an erroneous target cell.

Finally, nucleic acid technology offers great promise in *research*. As a result of the dramatic improvements in cloning and sequencing large sections of DNA, DNA sequences or genes are increasingly being used as a starting point for the analysis of a biochemical mechanism. Nucleic acid sequences are also used to study evolutionary relationships between different organisms by comparing sequences of conserved genes.

8.1. Hybridization Techniques for Nucleic Acid Detection

Nucleic acid hybridization is based on the fact that two complementary nucleic acid strands (DNA – DNA, RNA – RNA, DNA – RNA) associate to form a double strand according to the rules of base pairing (see also → Genetic Engineering, Chap. 3.1.). The nucleic acid to be detected must be present as a single strand, i.e., denaturation of double strands is required before hybridization. The nucleic acid probe must contain a sequence complementary to that of the nucleic acid under investigation and have a label to permitidentification of the double strands formed by hybridization.

Solid-Phase Hybridization. The labeled hybrid strands must also be distinguishable from unhybridized probes. The simplest way of doing this is to use methods in which the nucleic acid of interest is bound to a solid phase (usually a nitrocellulose or nylon membrane). The potential nonspecific binding sites for the probe on the solid phase are blocked prior to the hybridization reaction with unlabeled nucleic acids of different origins, proteins, and polymers. Consequently, the labeled probe only binds to the solid phase via specific base pairing to the target nucleic acid. The unbound probe is subsequently washed away.

In the *dot blot process* nucleic acid samples are spotted on a membrane [272]. After hybridization the detection of bound label reveals which spots contain the nucleic acid of interest. Quantification is possible by comparison with the signal from a standard [273].

In *Southern blotting* (→ Genetic Engineering) the DNA to be analyzed is separated according to size by gel electrophoresis and subsequently transferred to a membrane without changing the spatial distribution of the DNA fragments [21, 274]. After hybridization the size of the fragment containing the sequences of interest can be analyzed. Analogous separation and hybridization of RNA is referred to as *Northern blotting*.

In *in situ hybridization* the nucleic acid of interest is fixed in such a manner that its natural cellular compartmentalization is maintained [21, 23, 275]. As in histological studies, entire cells are applied and fixed to microscope slides or tissue is cut into ultrathin layers and fixed. Hybridization takes place on the slide between the labeled probes and the nucleic acids in the sample material. Examination under the microscope reveals which and how many cells contain the nucleic acid of interest, as well as its subcellular location. In situ hybridization permits differentiation between individual cells, whereas in solid-phase blotting methods the nucleic acids from many cells are mixed together. In situ hybridization also permits the assignment of a nucleic acid sequence to a chromosome by analyzing cells that are in the metaphase of cell division [276, 277]. The areas of the chromosomes containing the complementary sequences are specifically labeled by the nucleic acid probe.

Hybridization in Solution. Hybridization reactions in which both reaction partners are in solution are more efficient. The hybridized probe, as part of a double strand, can be separated from the unreacted single-stranded probe by *electrophoresis* or *chromatography*.

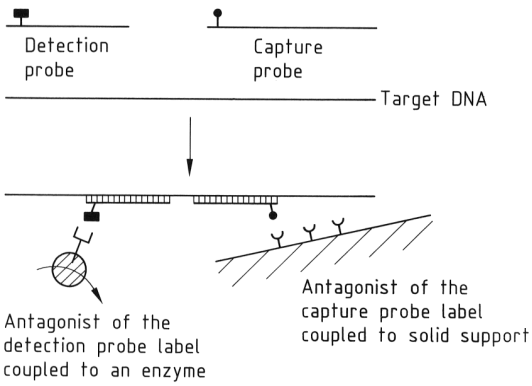


Figure 12. Sandwich hybridization
The oligonucleotide detection and capture probes are labeled with two different labels.

Sandwich hybridization offers an alternative way of separating hybrids formed in solution (Fig. 12) [278, 279]. The nucleic acid of interest is hybridized with two different probes that recognize different sections of the nucleic acid. One of the probes (detection probe) is labeled to allow detection of the hybrid. The second probe acts as a capture probe to bind the hybrid consisting of the nucleic acid of interest and the detection probe to a solid phase. The detection probe only binds to the solid phase if the

nucleic acid of interest is present in the solution and forms a bridge between the two probes. The capture probe can be covalently coupled to the solid phase [280], or can be labeled with a molecule that is distinct from the detection label and is bound to the solid phase via a fixed antagonist. Examples are biotin-labeled probes with a streptavidin- or avidin-loaded solid phase [281] or antigen-loaded probes with an appropriate antibody on the solid phase.

Strand displacement is another method for detecting hybrids formed in solution (Fig. 13) [282, 283]. Here, the probe is a hybrid of a short, labeled oligonucleotide and a longer nucleic acid strand forming a partially single-stranded, partially double-stranded molecule. The longer strand is complementary to the nucleic acid of interest. If the nucleic acid of interest is present in the assay, double strands are formed with the single-stranded region of the probe. This double-stranded region of the unlabeled probe strand and the nucleic acid of interest is longer than the double-stranded region formed between the two probe strands. Therefore complete hybridization occurs between the longer probe strand and the nucleic acid of interest. As a result the short, labeled probe strand is released and can be detected.

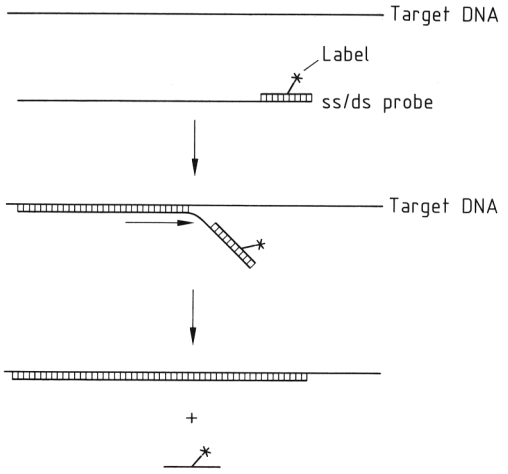


Figure 13. Strand displacement
ss = single-stranded; ds = double stranded

Finally, there are a series of modifications of a hybridization principle which does not require the separation of bound from unbound probe. Here, two labeled probes hybridize with the nucleic acid of interest at adjacent positions. Spatial proximity of the two labels is only guaranteed if the nucleic acid of interest is present and produces a signal that can be measured (e.g., enzymatic activity, light of a certain wavelength) [284].

8.2. Labeling and Detection Systems

Radioactive labeling is losing importance due to the development of sensitive, nonradioactive labels. Nucleic acids are usually radioactively labeled by enzymatic incorporation of nucleotides containing ^{3}H, ^{35}S, or ^{32}P isotopes with polymerases. Alternatively, the 5′-OH terminus of DNA can be labeled with ^{32}P by polynucleotide kinase [21]. Radioactive nucleic acids are detected by exposure to X-ray film or a photoemulsion (in situ hybridization) or by measurement in a scintillation counter.

In nonradioactive methods, the nucleic acid can be directly labeled with a fluorescent or luminescent dye [285], bound to an enzyme (e.g., β-galactosidase, horseradish peroxidase, alkaline phosphatase) or coupled with biotin or a hapten (antigenic determinant) [284]. The hapten (or biotin) is detected by the binding of an appropriate antibody (or streptavidin/avidin). The antibody and streptavidin/avidin are, in turn, directly coupled to a dye molecule or to an enzyme [21]. Additional chemical methods for labeling nucleic acids are described in [21, 286–288] and in Section 7.6.

Photolabeling is a simple and fast labeling method [21, 289]. Biotin or a hapten is coupled to a photoreactive group (photobiotin [289], photodinitrophenyl phosphate (DNP) [290], or photodigoxigenin [291]) which reacts with the amino groups of the nucleotides. On irradiation with light the photoreactive group is activated resulting in the formation of a covalent bond between biotin or the hapten (DNP, digoxigenin) and the nucleic acid.

Nucleic acids from biological sources (e.g., plasmids) are usually labeled with the aid of enzymes. Modified nucleotides are inserted with polymerases [21]; DNA polymerases are used to insert the deoxyribonucleotide triphosphate derivatives (e.g., biotin – dUTP, biotin-dATP, digoxigenin – dUTP, bromo – dUTP, and fluorescein – dUTP) in nick

translation (→ Enzymes, Chap. 6.2.1.) or by random primed labeling. The 3′-terminus of nucleic acids can be labeled with these nucleotides by terminal transferase. Modified ribonucleotides (e.g., biotin – UTP and digoxigenin – UTP) are inserted into RNA probes by in vitro transcription with RNA polymerases, particularly with phage-coded RNA polymerases [21,22] (→ Enzymes, Chap. 6.3.).

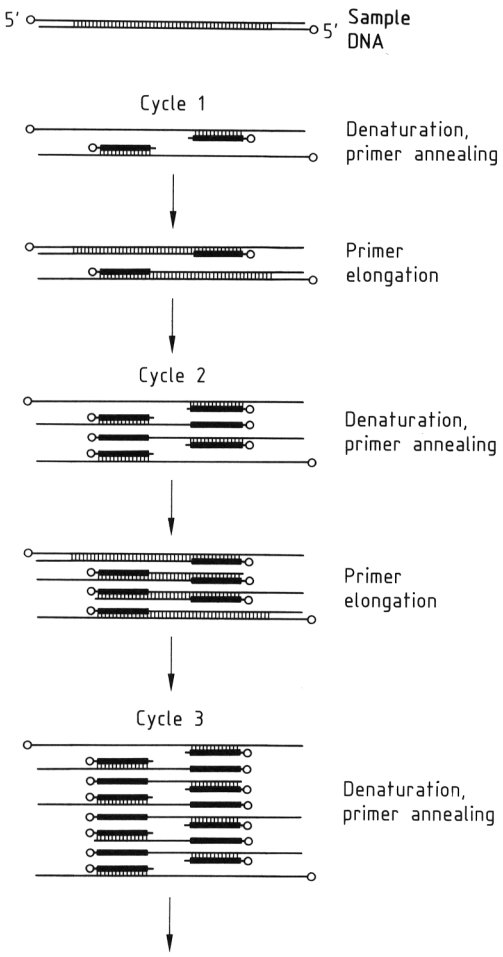


Figure 14. Polymerase chain reaction (PCR)

8.3. Amplification Systems

The sensitivity limit of nucleic acid probe technology can be increased by specifically amplifying (i.e., increasing the amount of) the nucleic acid of interest [21].

Polymerase Chain Reaction (PCR). The polymerase chain reaction is an established technique [15,27,292,293] which is used to amplify a particular nucleic acid from a complex mixture of sequences. The amplified DNA can be subsequently sequenced and/or cloned using genetic engineering techniques, or can be used to increase the sensitivity of pathogen detection. The sensitivity can be increased to such an extent that, for instance, one genome from the human immunodeficiency virus (HIV) integrated in the DNA of a single cell can be detected in the presence of many uninfected cells.

The PCR is based on the repetition of three partial reactions in 20 – 50 cycles (Fig. 14):

1) An oligonucleotide primer is annealed to each of the two denatured complementary strands of the DNA sequence to be amplified.
2) Primer extension catalyzed by a DNA polymerase converts the region to be amplified into two double strands.
3) The double strands are denatured by heating.

Theoretically the number of templates can be doubled in each cycle, so that 2^n templates would be created in n cycles; 25 cycles would result in an amplification rate of 3.4×10^7. However, the partial reactions are not perfectly efficient, and amplification rates of ca. $(1-3)\times10^6$ are obtained in 25 cycles. Implementation of the PCR is facilitated by the use of thermostable DNA polymerases because otherwise new polymerase would have to be added at the start of each cycle. Typical PCR cycles consist of a few minutes of primer annealing at ca. 55 °C, a few minutes of elongation at ca. 70 °C, and 1 – 2 min of denaturation at 95 °C. These steps can be carried out automatically in programmable "thermocyclers".

RNA can also be amplified by PCR if a DNA template is created in the first step with the help of reverse transcriptase. The PCR product can be detected by electrophoresis with or without previous cleavage with restriction endonucleases, and with or without hybridization with a probe. The PCR product can be labeled directly by using labeled primers [294] or by incorporating labeled nucleotides (Section 8.2) [295,296]. Con-

tamination is a serious problem in PCR because of the enormous sensitivity of the method. For example, contamination of a sample with a single molecule of the DNA of interest can lead to an incorrect positive result.

Other Amplification Systems. In *transcription-based amplification* oligonucleotide primers are annealed to the nucleic acid of interest which, in addition to complementary sequences to the target, also contain a promoter (binding site) for RNA polymerase [297]. After primer elongation RNA polymerases are used to produce RNA transcripts starting from the promoters. The transcripts can be converted to DNA with reverse transcriptase and used again as targets for the oligonucleotide primers.

The *ligase chain reaction* uses a ligase instead of a polymerase in temperature cycles analogous to the PCR [298, 299]. Two oligonucleotides are annealed adjacent to one another on the template strand and joined by a ligase. The ligated strand is then removed by denaturation and two oligonucleotides can anneal again in the next cycle.

Other amplification systems do not require cycles with different temperatures but proceed isothermally. In *self-sustained sequence replication* (3 SR) [300] and *nucleic acid sequence based amplification* (NASBA) [301], three enzymes act simultaneously at a constant temperature to amplify the target sequence. Reverse transcriptase synthesizes DNA on an RNA template by extending oligonucleotide primers that contain promoter sequences for RNA polymerases. The RNA strand of the resulting DNA–RNA hybrid is digested with RNase H allowing double-stranded DNA to be produced by reverse transcriptase with a second primer. RNA polymerase then transcribes the double-stranded DNA starting at the promoters, and the resulting RNA is again converted to DNA with reverse transcriptase.

Another isothermal amplification technique is based on the replication of RNA genomes with the help of *Qβ replicase*–a RNA-dependent RNA polymerase from phage Qβ [302, 58]. Natural substrates of Qβ replicase are characterized by RNA sequences forming a special secondary structure. For nucleic acid detection these sequences are modified by the insertion of short probe sequences. Unlike the amplification methods described above, the target sequence itself is not multiplied, instead the probe is amplified after it has located the target sequence (signal amplification). Nonspecifically bound and contaminating probe molecules cause background problems that can be alleviated by repeating the hybridization in cycles (reversible target capture) [303].

Another method of signal amplification involves the formation of probe networks in which the target-specific probe has target sequences for secondary probes. A "Christmas tree" of labeled probe sequences is produced for each original target molecule [304–306].

8.4. Applications of Probe Technology

Detection of Infectious Microorganisms. Detection of nucleic acids can be used to identify pathogenic viruses, bacteria, or parasites. In comparison with the detection of proteins or antigens, nucleic acid detection has the advantage that it allows identification of infectious organisms that frequently change their antigens (e.g., HIV) or that exist in a variety of serotypes. Furthermore, viral infections can be detected at the latent stage when no detectable synthesis of viral proteins occurs (e.g., HIV or herpes viruses). DNA probe technology can be more sensitive than immunological techniques due to the possibility of nucleic acid amplification. Although immunological tests are still easier, faster, and cheaper, competitive nucleic acid tests are to be expected in the future.

Detection of the nucleic acid of a pathogen can also serve as a direct measure of the infectiousness of a patient or test material [307]. Quantitative determination of the pathogenic nucleic acid can be used to follow the progress of treatment.

Detection of viruses is possible with a broad spectrum of nucleic acid probes, ranging from synthetic oligonucleotides to the complete cloned viral genome. Infections caused by some virus groups (e.g., papilloma viruses or enteroviruses [308]) can be diagnosed with a probe containing group-specific sequences identical in all members of the group and by using probes with nucleic acid sequences specific for each member.

For the detection of bacteria, the fact that each bacterium contains approximately 10 000 ribosomes and thus 10 000 copies of rRNAs can be used to increase the sensitivity level. The rRNA sequences are relatively well conserved and can used to establish relationships between different organisms in phylogenetic trees. Tests for individual bacterial species, for groups, or even for a wide range of bacteria can be carried out with appropriate oligonucleotide probes [309]. Another group of nucleic acid probes contains sequences of bacterial pathogenicity factors and allows differentiation between pathogenic and apathogenic types of a species (e.g., *Escherichia coli*) [268, 269]. Analogously, detection of resistance genes permits detection of antibiotic resistance and simultaneous identification of the pathogen [310]. Bacterial identification is also applied in the detection of pathogens in food (e.g., *Salmonella*).

Detection of Human DNA Sequences. Nucleic acid probe technology can be used for the *diagnosis of hereditary diseases* and *genetic counseling* [24]. If the genetic basis of the disease is known, the alteration in the nucleic acid can be detected by hybridization with the appropriate region. Such alterations include point mutations (sickle-cell anemia), deletions (Duchenne's muscular dystrophy), or variations in expression (forms of thalassemia). If the genetic cause of the disease is unknown, genetic markers that are inherited together with the defect gene are analyzed, e.g., recognition sites for restriction enzymes (restriction fragment length polymorphism, RFLP) [311].

The analysis of human sequences can also be applied in *tumor diagnosis* [25, 26]. Some human tumors are associated with characteristic chromosomal changes or alterations in genes (oncogenes) [312]. Suitable nucleic acid probes can be used for the highly sensitive detection of DNA sequences of different chromosomes connected by chromosomal translocations or for the detection of mutations in oncogenes that are associated with tumors. For instance, the Philadelphia chromosome (a translocation between chromosome 9 and 22) is characteristic for chronic myelogenous leukemia [313]. Point mutations in the *ras* oncogene are associated with several tumors, and amplification of the *N-myc* gene with neuroblastomas. In the case of recessive oncogenes (tumor suppressor genes) [314, 315], the loss of gene expression results in a tumor; for example, the loss of the retinoblastoma gene results in an eye tumor, retinoblastoma. Another area of tumor diagnosis is the detection of clonality of a leucocyte population caused by the transformation of a white blood cell [316]. Since all lymphocytes have their own characteristic rearrangement of immunoglobulin genes or T-cell receptor genes, the appearance of a defined pattern in many lymphocytes indicates proliferation of a single cell.

Methods that use the genetic diversity of humans to study individuals have wide applicability. Determination of the genes of the major histocompatibility locus (HLA) can be used both for diagnosis (transplantation medicine, prediction of risk of disease, autoimmune diseases) and for differentiation between individuals [317]. Probes which hybridize with repetitive sequences that occur in each individual in a slightly different, genetically determined arrangement can be used for finger printing [28]. Here nucleic acids cleaved by restriction enzymes are separated electrophoretically and hybridized with these probes in the Southern blot technique (Section 8.1). Several fragments of different sizes, each containing these repetitive sequences, produce a signal. Each individual inherits half the sequences from the father and half from the mother, and therefore has a unique fragment pattern [318–320]. This technique is used in forensic medicine to identify criminals by DNA isolated from cells (hair, blood etc.). Other fields of application are in the determination of paternity and family relationships.

9. Nucleosides and Nucleotides

9.1. Nucleosides

The term nucleosides was originally applied to the ribose derivatives of purines that can be isolated from an alkaline RNA hydrolysate. Later, this name was given to all purine and pyrimidine *N*-glycosides of D-ribose and 2-deoxy-D-ribose. Today, this term refers to all natural or synthetic compounds which consist of a heterocyclic nitrogen-containing base (aglycon) and a carbohydrate residue (glycon). The nitrogen atom (*N*-nucleoside) or carbon atom (*C*-

nucleoside) of the heterocycle is linked to the anomeric carbon atom of the sugar residue.

Ribonucleosides and deoxyribonucleosides are obtained from naturally occurring RNA and DNA by enzymatic or chemical hydrolysis. Most nucleosides are stable within a wide pH range. However, under strongly acidic or alkaline conditions at elevated temperature cleavage of the *N*-glycosidic bond may occur. The purine nucleosides are adenosine [*58-61-7*], deoxyadenosine [*958-09-8*], guanosine [*118-00-3*], and deoxyguanosine [*961-07-9*]. The pyrimidine nucleosides are cytidine [*65-46-3*], deoxycytidine [*951-77-9*], thymidine [*50-89-5*], and uridine [*58-96-8*].

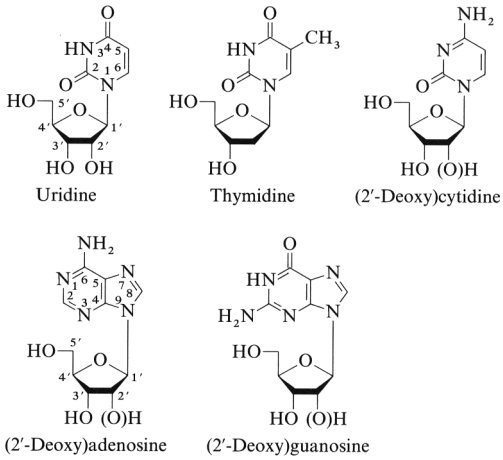
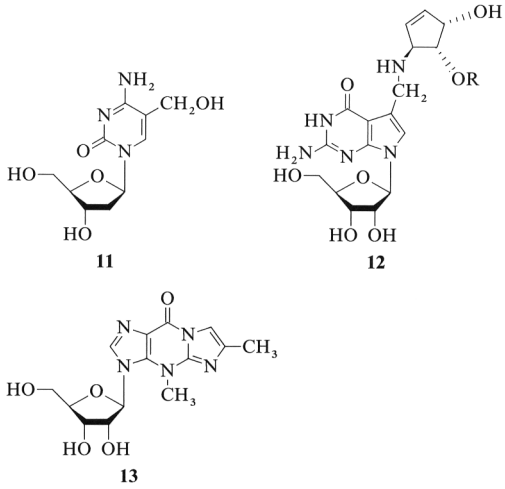


Figure 15. Common nucleosides

The furanoses D-ribose and 2-deoxy-D-ribose are linked to the bases by a β-glycosidic bond at C-1 (Fig. 15). A limited number of nucleosides in the α-D configuration (e.g., 5,6-dimethylbenzimidazole-α-D-ribofuranoside of vitamin B 12) can be isolated from natural substances.

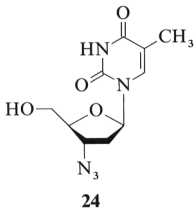
Rare nucleosides and nucleoside antibiotics are found in biopolymers as fermentation products of microorganisms or as building blocks of nucleic acids [29]. They exhibit structural modifications of the nucleobase or the glycon residue.

An example of a rare building block of viral DNA is 2′-deoxy-5-hydroxymethylcytidine (**11**) and those of tRNA are the hypermodified nucleosides queuosine [*57072-36-3*] (**12**) and wyosine [*52662-10-9*] (**13**).



Examples of nucleosides formed by microorganisms are tubercidin [*69-33-0*] (**14**), formycin [*6742-12-7*] (**15**), cordycepin [*73-03-0*] (**16**), blasticidin S [*2079-00-7*] (**17**), vidarabin [*24356-66-9*] (**18**), oxetanocin [*103913-16-2*] (**19**), aristeromycin [*19186-33-5*] (**20**), neplanocin A [*72877-50-0*] (**21**), sinefungin [*58944-73-3*] (**22**), and tunicamycin A [*66081-37-6*] (**23**, $n = 9$), tunicamycin B [*66081-36-5*] (**23**, $n = 10$), tunicamycin C [*66081-37-6*] (**23**, $n = 8$), and tunicamycin D [*66081-38-7*] (**23**, $n = 11$) (Fig. 16).

Naturally occurring nucleosides show a wide variety of biological effects. Accordingly a large number of derivatives have been synthesized in order to find useful chemotherapeutics [30]. The 2′,3′-dideoxyribonucleosides are important synthetic nucleoside derivatives with modified sugar moieties. In the cell the 5′-triphosphate derivatives of these compounds inhibit the reverse transcriptase of retroviruses and thus are potentially important therapeutic drugs (e.g., 3′-azidothymidine [*30516-87-1*] (**24**) (AZT), the first approved drug against HIV infection.



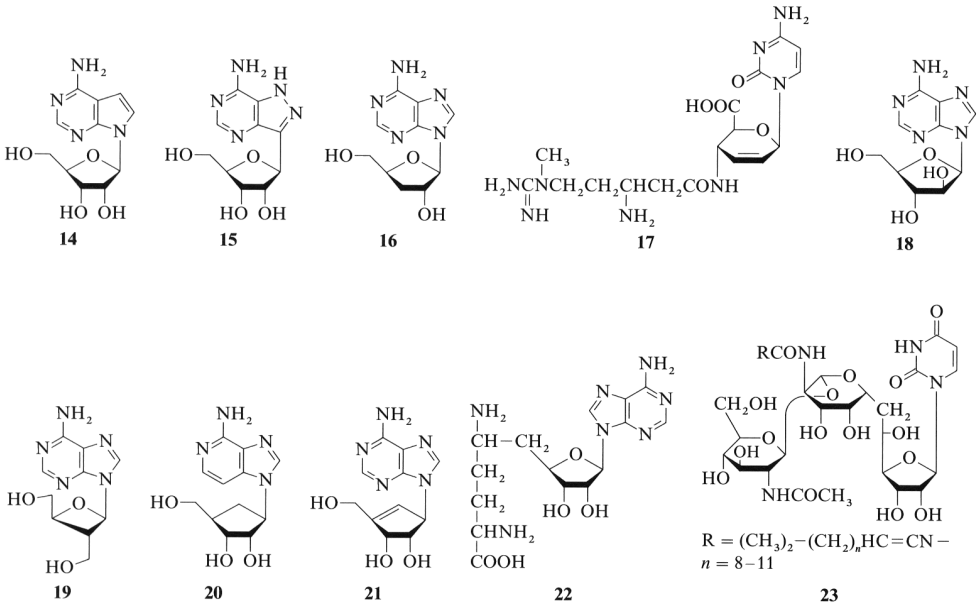
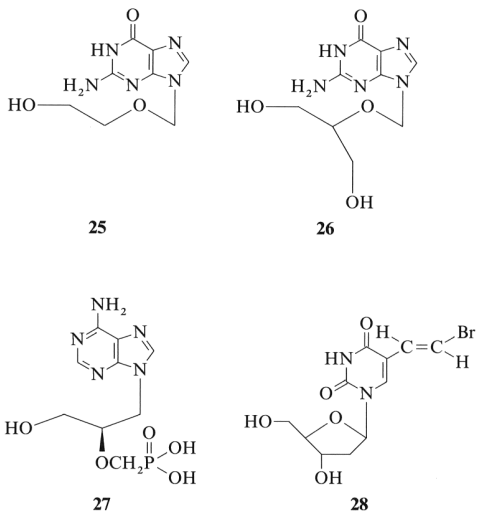


Figure 16. Nucleosides formed by microorganisms

Other derivatives which contain only a part of the ribofuranosyl residue (e.g., acyclonucleosides) also show antiviral activity. This class of substances includes acyclovir (**25**), ganciclovir (**26**) and phosphonylmethyl analogues [e.g., (*S*)-9-(3-hydroxy-2-phosphonylmethoxypropyl) adenine (**27**)]. Nucleosides with a halogen substituent in the heterocyclic moiety are important antiviral drugs (e.g., 5-iodo-2′-deoxyuridine, 5-bromovinyl-2′-deoxyuridine (**28**), or 2-chloro-2′-deoxyadenosine).



The nucleosides are obtained by modification of existing nucleosides or total synthesis. Total synthesis can be carried out in a linear or convergent manner. Due to the chiral centers on the glycon residue, preparation of optically pure nucleosides requires stereoselective methods of synthesis, especially for the glycosidation of the aglycons. The most important methods of total synthesis will be described briefly. Detailed surveys of the chemical synthesis of and on nucleoside derivatives are given in [31–35].

Synthesis of *N*-Nucleosides. The first syntheses date back to the work of Fischer and Helferich who condensed acetyl-protected α-D-glucopyranosyl halides with the silver salts of heterocycles. The analogous chemical synthesis of the natural nucleosides adenosine and guanosine was described by Todd and coworkers [321]. This method was improved after the introduction of the chloromercury purine derivatives [322] and the use of mercury cyanide (which avoids the tedious preparation of the heavy metal salts of the nucleobases) [323]. The Hilbert–Johnson method was developed for the preparation of pyrimidine nucleosides [324]. This method starts with 2,4-dialkoxypyrimidines, which yield the corresponding nucleosides on reaction with O-acetyl-protected

1-halogenosugars (glycosyl halides) and subsequent ammonolysis or hydrolysis of the 4-alkoxy intermediates [325].

The introduction of the melt process according to SATO [326] represented another advance in nucleoside synthesis. Here, a polyacylated sugar is heated with the purine or pyrimidine base and a catalyst (e.g., 4-toluenesulfonic acid or zinc chloride) in vacuum. This method has the advantage that the free heterocyclic bases (instead of their heavy metal salts) and the 1-O-acetyl sugars (instead of the corresponding glycosyl halides) can be used.

Another synthesis, developed simultaneously by BIRKHOFER and coworkers [327] and NISHIMURA and coworkers [328], involves reaction of trialkylsilylated heterocycles with peracylated 1-halogenosugars in the presence of silver perchlorate as catalyst.

This method was improved by VORBRÜGGEN and is a silyl version of the Hilbert–Johnson method [329]. It is the method of choice for the synthesis of pyrimidine nucleosides. A similar process is based on transglycosylation, i.e., the transfer of the sugar residue of a nucleoside (donor) to a heterocyclic base (acceptor). Apart from chemical transglycosidation, enzymatic sugar transfer is also possible [330].

An elegant method for the synthesis of the nucleosides of purines or purine analogues, developed by SEELA and coworkers, is based on phase transfer catalysis [331]. It gives high yields of the desired β-anomers and the anomeric ratio can be controlled by the choice of catalyst. Finally, reference should be made to the synthesis of the heterocyclic moiety from the C-1 functionalized sugar molecule [332]. This is the method of choice for the synthesis of carbocyclic nucleosides [333].

Synthesis of *C*-Nucleosides. Two methods are available for synthesizing the C–C linkage between the glycosidic and the aglyconic parts of the molecule [334]:

1) Synthesis according to the Fischer–Helferich principle
2) Synthesis via a preformed function on C-1 of the sugar

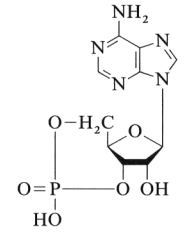
Unlike *N*-glycosides, the glycosidic bond of *C*-nucleosides is stable to acids and these compounds are not attacked by nucleoside phosphorylases. Consequently *C*-nucleosides are more stable than *N*-nucleosides in biological systems.

9.2. Nucleotides

The term nucleotide generally refers to the phosphate esters of nucleosides. Only mononucleotides (i.e., nucleotides consisting of only one nucleotide unit) are considered here. Mononucleotides can be joined together to form an oligonucleotide (up to 200 nucleotide units) or a polynucleotide (> 200 nucleotide units).

Nucleotides are obtained by the chemical or enzymatic hydrolysis of nucleic acids and by the chemical or enzymatic phosphorylation of nucleosides [335–339]. They may be ribonucleotides (from RNA) or 2′-deoxyribonucleotides (from DNA). Depending on the position at which the sugar moiety is phosphorylated nucleotides can exist as 2′-, 3′-, or 5′-phosphoric acid monoesters. Cyclic 2′,3′- and 3′,5′-phosphoric acid diesters can be formed by intramolecular dehydration.

Cyclic adenosine monophosphate [*60-92-4*] (cAMP) is especially important as a mediator of the action of a large number of hormones (second messenger) [340].

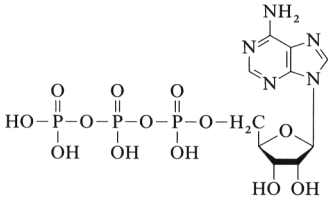


Adenosine
3′,5′-monophosphoric acid
(cyclic AMP)

Cyclic AMP is formed in the cell from ATP by the enzyme adenylate cyclase. It was first synthesized in the laboratory by COOK and coworkers [341]. Numerous attempts have been made to specifically alter the pharmacological effects of cAMP by chemical modification [342].

The nucleotides can contain more than one phosphate group. The 5′-triphosphates of the ribonucleosides and 2′-deoxyribonucleosides of

the purines adenine (ATP or dATP) and guanine (GTP or dGTP) and of the pyrimidines cytosine (CTP or dCTP) and thymine (dTTP) or uracil (UTP) are the most common natural nucleotides. They serve as building blocks for the polymerase-catalyzed synthesis of nucleic acids. Adenosine 5′-triphosphate [*56-65-5*] is also the main source of chemical energy in the living world.

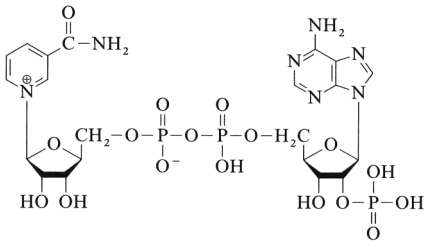


Adenosine
5′-triphosphoric acid
(ATP)

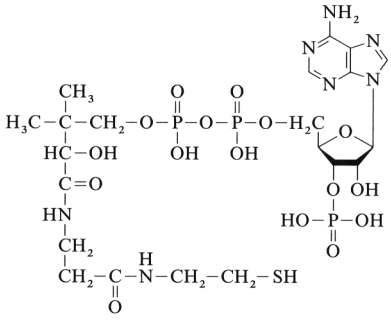
In contrast to many nucleosides, the corresponding phosphate esters are readily soluble in water as their alkali salts. Although the nucleoside monophosphates are stable compounds, the corresponding nucleoside di- and triphosphates undergo slow hydrolytic cleavage of the pyrophosphate bond, even at ambient temperature and neutral pH. Nucleoside mono- and polyphosphates are dephosphorylated by phosphatases (phosphomonoesterases) to yield the corresponding nucleosides.

Analogues containing modified phosphate groups (e.g., thiophosphate [343], imidophosphate [344], or methylene phosphonate functions [345]) are not or only very slowly attacked by phosphomonoesterases. The triphosphate derivatives of those analogues which are only slightly modified compared to the parent nucleotides are valuable tools for studying the reaction mechanisms of ATP-dependent enzyme systems.

The naturally occurring cyclic nucleoside 3′,5′-monophosphates are also very stable. They are converted to acyclic 5′-monophosphates by specific phosphodiesterases. Other important nucleotide building blocks occurring in nature are the adenosine phosphate residues in nicotin-amide adenine dinucleotide phosphate [*53-59-8*] (NADP), coenzyme A [*85-61-0*], and flavine adenine dinucleotide (FAD) [*146-14-5*].



β-Nicotinamide adenine dinucleotide phosphate
(NADP)



Coenzyme A

Synthesis of Nucleotides. For a review of chemical nucleotide syntheses, see [346]. Nucleoside monophosphates are usually prepared by chemical phosphorylation of the nucleosides with phosphorus oxytrichloride. If the reaction is carried out with a trialkyl phosphate as solvent the 5′-monophosphates are obtained almost exclusively [347]. Protection of the 2′- and/or 3′-OH functions of the sugar is not required. Various methods are available for the chemical synthesis of nucleoside di- and triphosphates from nucleoside monophosphates.

In the *morpholidate method* the nucleoside monophosphate is reacted with morpholine and dicyclohexylcarbodiimide to give the corresponding phosphomorpholidate [348]. The latter is subsequently reacted with orthophosphate to give the nucleoside diphosphate or with pyrophosphate to yield the nucleoside triphosphate.

In the *imidazolidate method* the nucleoside monophosphates are reacted with 1,1′-carbonyldiimidazole to give the corresponding nucleoside monophosphate imidazolidates [349]. These compounds are in turn treated with inorganic ortho- or pyrophosphate to yield the nucleoside di- or triphosphates. In comparison with

the morpholidate method, this process has the advantage of being a "one-pot" reaction.

In the *anhydride method* the nucleoside monophosphates are reacted with diphenyl chlorophosphate [350]. The resulting nucleoside diphenyl pyrophosphates can be subjected to nucleophilic exchange by treatment with anions such as ortho- or pyrophosphate.

These methods can be used analogously to prepare the nucleotides of the coenzyme type by replacing the orthophosphate with, for example, sugar phosphates. The possibility of reacting the nucleoside phosphodichloridate, obtained after treatment with phosphorus oxytrichloride, directly with pyrophosphate in a "one-pot" reaction has been described [351].

Enzymatic syntheses of nucleotides are described in [335–338].

9.3. Therapeutically Important Nucleoside and Nucleotide Derivatives

Antiviral Derivatives. The strategies involved in the development of antiviral nucleosides are reviewed in [352]. For the development and applications of antiviral active nucleotide analogues, see [353]. Of the large number of compounds that have been synthesized since the early 1970s, only a few derivatives have so far been approved. Among them are:

1) 5-iodo-2′-deoxyuridine [*54-42-2*] (IDU)
2) 5-trifluoromethyl-2′-deoxyuridine [*70-00-8*] (trifluorothymidine, TFT)
3) 9-β-D-arabinofuranosyladenine [*24356-66-9*] (Ara-A, vidarabin) (**18**)
4) 3′-azido-3′-deoxythymidine [*30516-87-1*] (AZT, zidovudine) (**24**)
5) 1-β-D-ribofuranosyl-1,2,4-triazole-3-carboxamide [*36791-04-5*] (ribavirin, virazol)
6) 9-[(2-hydroxyethoxy)methyl]guanine [*59277-89-3*] (acyclovir) (**25**)

The following nucleoside analogues are active against HIV infections. However, clinical trials of these compounds are still incomplete or they have not yet been approved partly because of serious side effects:

1) 2′,3′-dideoxycytidine [*7481-89-2*] (ddC), 2′,3′-dideoxyadenosine [*4097-22-7*] (ddA), and 2′,3′-dideoxyinosine [*69655-05-6*] (ddI)
2) 9-(1,3-dihydroxy-2-propoxymethyl)guanine [*82410-32-0*] (ganciclovir, DHPG) (**26**)
3) (*E*)-5-(2-bromovinyl)-2′-deoxyuridine [*73110-56-2*] (BVDU) (**28**)

Antitumor Derivatives. For details, see → Cancer Chemotherapy, Chap. 2.4.

10. References

General References

1. L. Stryer: *Biochemistry,* W. H. Freeman, New York 1988.
2. D. Freifelder: *Molecular Biology,* Jones and Bartlett Publishers, Boston 1987.
3. J. Darnell, H. Lodish, D. Baltimore: *Molecular Cell Biology,* Scientific American Books, W. H. Freeman, New York 1990.
4. B. Alberts: *Molecular Biology of the Cell,* Garland Publishing, New York-London 1983.
5. J. D. Watson et al.: *Molecular Biology of the Gene,* The Benjamin/Cummings Publishing Comp., Menlo Park, California 1987.
6. B. Lewin: *Gene,* VCH Verlagsgesellschaft, Weinheim 1988.
7. J. Sambrook, E. F. Fritsch, T. Maniatis: *Molecular Cloning. A Laboratory Manual,* 2nd ed., Cold Spring Harbor Laboratory Press 1989.
8. E.-L. Winnacker: *From Genes to Clones. Introduction to Gene Technology,* VCH Verlagsgesellschaft, Weinheim 1987.
9. A. Kornberg: *DNA Replication,* W. H. Freeman, San Francisco 1980.
10. E. C. Friedberg: *DNA Repair,* W. H. Freeman, New York 1985.
11. W. Saenger: *Principles of Nucleic Acid Structure,* Springer Verlag, New York 1984.
12. J. E. Dahlberg, J. N. Abelson (eds.): *Methods in Enzymology,* vol. 180, Part A, RNA processing, Academic Press, New York 1989.
13. P. M. Wassarman, R. D. Kornberg (eds.): *Methods in Enzymology,* vol. 170, Nucleosomes, Academic Press, New York 1989.
14. H. F. Noller, Jr., K. Moldave (eds.): *Methods in Enzymology,* vol. 164, Ribosomes, Academic Press, New York 1988.
15. M. A. Innis, D. H. Gelfand, J. J. Sninsky, T. J. White (eds.): *PCR Protocols – A Guide to Methods and Applications,* Academic Press, New York 1990.

16. B. S. Sproat, M. J. Gait in M. J. Gait (ed.): *Oligonucleotide Synthesis, A Practical Approach,* IRL Press, Oxford-Washington 1984.
17. H. G. Gassen, A. Lang (eds.): *Chemical and Enzymatic Synthesis of Gene Fragments: A Laboratory Manual,* Verlag Chemie, Weinheim 1982.
18. E. Sonveaux: "The Organic Chemistry Underlying DNA Synthesis," *Bioorg. Chem.* **14** (1986) 274 – 325.
19. J. W. Engels, E. Uhlmann in A. Fichter (ed.): *Gene Synthesis, Advances in Biochemical Engineering/Biotechnology,* vol. 37, Springer Verlag, Berlin – Heidelberg 1988.
20. G. Zon: "Oligonucleotide Analogues as Potential Chemotherapeutic Agents," *Pharm. Res.* 1988, no. 5, 539 – 549.
21. G. H. Keller, M. M. Manak: *DNA Probes,* Macmillan Publishers, New York 1989.
22. H.-J. Höltke, C. Kessler, *Nucleic Acid Res.* **18** (1990) 5843 – 5851.
23. B. D. Hames, S. J. Higgins (eds.): *Nucleic Acid Hybridization: A Practical Approach,* IRL Press, Oxford 1985.
24. K. E. Davies (ed.): *Human Genetic Diseases: A Practical Approach,* IRL Press, Oxford 1986.
25. J. Cossman (ed.): *The Molecular Genetics and the Diagnosis of Cancer,* Elsevier Science Publishing, New York 1990.
26. M. Furth, M. Greaves (eds.): *Molecular Diagnostics of Human Cancer, Cancer Cells 7,* Cold Spring Harbor Laboratory, New York 1989.
27. H. A. Erlich (ed.): *PCR Technology. Principles and Applications for DNA Amplification,* Stockton Press, New York-London-Tokyo-Melbourne-Hong Kong 1989.
28. L. T. Kirby: *DNA Fingerprinting, An Introduction,* Stockton Press, New York – London – Tokyo –Melborne – Hong Kong 1990.
29. R. J. Suhadolnik: *Nucleoside Antibiotics,* Wiley-Interscience, New York 1970.
30. R. J. Suhadolnik: *Nucleosides as Biological Probes,* Wiley-Interscience, New York 1979.
31. C. A. Decker, L. Goodman in W. Pigman, D. Horton (eds.): *The Carbohydrates,* vol. IIA, Academic Press, New York-London 1970.
32. W. Zorbach, R. S. Tipson: *Synthetic Procedures in Nucleic Acid Chemistry,* vol. I, Interscience Publishers, New York 1968.
33. L. Goodman in P. O. P. Ts'o (ed.): *Basic Principles in Nucleic Acid Chemistry,* **vol. I,** Academic Press, New York-London 1974, p. 93.
34. L. B. Townsend, R. S. Tipson: *Nucleic Acid Chemistry,* Wiley-Interscience, New York, Parts 1 and 2, 1978; Part 3, 1986.
35. Y. Mizuno; *The Organic Chemistry of Nucleic Acids,* Elsevier, Amsterdam 1986.

Specific References

36. G. F. Joyce, *Nature (London)* **338** (1989) 217.
37. A. I. Lamond, T. J. Gibson, *Trends Genet.* **6** (1990) 145.
38. J. D. Watson, F. H. C. Crick, *Nature (London)* **171** (1953) 737.
39. S. B. Zimmerman, *Annu. Rev. Biochem.* **51** (1982) 395.
40. R. Schleif, *Science (Washington, D. C.)* **241** (1988) 1182.
41. L. Joshua-Toret et al., *Nature (London)* **334** (1988) 82.
42. M. Miller et al., *Nature (London)* **334** (1988) 85.
43. A. Nordheim: *ISI Atlas of Science: Biochemistry* **1** (1988) 279.
44. H. R. Drew et al., *Annu. Rev. Cell Biol.* **4** (1988) 1.
45. J. Lebowitz, *TIBS* **15** (1990) 202.
46. J. Grinsted, P. M. Bennett, *Methods Microbiol.* **21** (1988) 129.
47. G. J. Pruss, K. Drlica, *Cell* **56** (1989) 521.
48. A. A. Travers, *Cell* **60** (1990) 177.
49. C. F. McAllister, E. C. Achberger, *J. Biol. Chem.* **264** (1989) 10 451.
50. A. A. Travers, *Annu. Rev. Biochem.* **58** (1989) 427.
51. D. M. Crothers et al., *J. Biol. Chem.* **265** (1990) 7093.
52. T. T. Eckdahl, J. N. Anderson, *Nucleic Acids Res.* **18** (1990) 1609.
53. A. Travers, A. Klug, *Nature (London)* **327** (1987) 280.
54. L. Bracco et al., *EMBO J.* **8** (1989) 4289.
55. M. S. Z. Horwitz, L. A. Loeb, *Science (Washington, D. C.)* **241** (1988) 703.
56. U. R. Müller, W. M. Fitch, *Nature (London)* **298** (1982) 582.
57. R. D. Wells, *J. Biol. Chem.* **263** (1988) 1095.
58. Y. Kohwi, T. Kohwi-Shigematsu, *Proc. Natl. Acad. Sci. USA* **85** (1988) 3781.
59. J. H. van de Sande et al., *Sience (Washington, D. C.)* **241** (1988) 551.
60. L. Perrouault et al., *Nature (London)* **344** (1990) 358.
61. C. Hélène, J.-J. Toulmé, *Biochim. Biophys. Acta* **1049** (1990) 99.

62. C. B. Harley et al., *Nature (London)* **345** (1990) 458.
63. G.-L. Yu et al., *Nature (London)* **344** (1990) 126.
64. V. A. Zakian et al., *Trends Genet.* **6** (1990) 12.
65. W. I. Sundquist, A. Klug, *Nature (London)* **342** (1990) 825.
66. J. D. Boeke, *Cell* **61** (1990) 193.
67. D. Sen, W. Gilbert, *Nature (London)* **344** (1990) 410.
68. F. Baudin, P. J. Romaniuk, *Nucleic Acids Res.* **17** (1989) 2043.
69. J. Normanly, J. Abelson, *Annu. Rev. Biochem.* **58** (1989) 1029.
70. A. Bhattacharyya et al., *Nature (London)* **343** (1990) 484.
71. R. M. W. Mans et al., *Nucleic Acids Res.* **18** (1990) 3479.
72. J. R. Wyatt et al., *BioEssays* **11** (1989) 100.
73. S. Arnott et al., *J. Mol. Biol.* **188** (1986) 631.
74. H. I. Elsner, E. B. Lindblad, *DNA* **8** (1989) 697.
75. H. Donis-Keller et al., *Nucleic Acids Res.* **4** (1977) 2527.
76. G. Krupp, H. J. Gross, *Nucleic Acids Res.* **6** (1979) 3481.
77. J. Parker, *Microbiol. Rev.* **53** (1989) 273.
78. M. Meselson, F. W. Stahl, *Proc. Natl. Acad. Sci. USA* **44** (1958) 671.
79. A. Kornberg, *Biochim. Biophys. Acta* **951** (1988) 235.
80. J. P. Vaughn et al., *Nucleic Acids Res.* **18** (1990) 1965.
81. D. Riesner, H. J. Gross, *Annu. Rev. Biochem.* **54** (1985) 531.
82. T. D. Fox, *Annu. Rev. Genet.* **21** (1987) 67.
83. H. M. Weintraub, *Sci. Am.* **262** (1990) 34.
84. J. Shine, L. Dalgarno, *Proc. Natl. Acad. Sci. USA* **71** (1974) 1342.
85. J. G. Belasco, C. F. Higgins, *Gene* **72** (1988) 15.
86. P. F. Johnson, S. L. McKnight, *Annu. Rev. Biochem.* **58** (1989) 799.
87. A. Baniahmad et al., *Cell* **61** (1990) 505.
88. D. R. Herendeen et al., *Science (Washington, D. C.)* **248** (1990) 573.
89. L. Gold, *Annu. Rev. Biochem.* **57** (1988) 199.
90. J. A. Steitz, *Sci. Am.* **258/6** (1988) 56.
91. M. R. Green, *Annu. Rev. Genet.* **20** (1986) 671.
92. S. M. Mount, *Nucleic Acids Res.* **10** (1982) 459.
93. A. M. Weiner, N. Maizels, *Cell* **61** (1990) 917.
94. V. Volloch et al., *Nature (London)* **343** (1990) 482.
95. D. Shippen-Lentz, E. H. Blackburn, *Science (Washington, D. C.)* **247** (1990) 546.
96. E. H. Blackburn, *J. Biol. Chem.* **265** (1990) 5919.
97. A. E. Dahlberg, *Cell* **57** (1989) 525.
98. L. Gold et al., *Annu. Rev. Microbiol.* **35** (1981) 365.
99. C. O. Gualerzi, C. L. Pon, *Biochemistry* **29** (1990) 5881.
100. W. Doerfler, *Annu. Rev. Biochem.* **52** (1983) 93.
101. E. U. Selker, *TIBS* **15** (1990) 103.
102. W. Zacharias et al., *J. Bacteriol.* **172** (1990) 3278.
103. N. D. F. Grindley, R. R. Reed, *Annu. Rev. Biochem.* **54** (1985) 863.
104. G. C. Walker, *Annu. Rev. Biochem.* **54** (1985) 425.
105. B. van Houten, *Microbiol. Rev.* **54** (1990) 18.
106. G. M. Myles, A. Sancar, *Chem. Res. Toxicol.* **2** (1989) 197.
107. L. A. Grivell, *Nature (London)* **344** (1990) 110.
108. D. L. Robertson, G. F. Joyce, *Nature (London)* **344** (1990) 467.
109. K. Kruger et al., *Cell* **31** (1982) 147.
110. C. Guerrier-Takada et al., *Cell* **35** (1983) 849.
111. O. C. Uhlenbeck, *Nature (London)* **328** (1987) 596.
112. T. R. Cech, *Science (Washington, D. C.)* **236** (1987) 1532.
113. J.-P. Perreault et al., *Nature (London)* **344** (1990) 565.
114. D. Herschlag, T. R. Cech, *Nature (London)* **344** (1990) 405.
115. J. J. Rossi, N. Sarver, *TIBTECH* **8** (1990) 179.
116. G. J. Raymond et al., *Anal. Biochem.* **173** (1988) 125.
117. J. Marmur, *J. Mol. Biol.* **3** (1961) 208.
118. U. Gubler, B. J. Hoffman, *Gene* **25** (1983) 263.
119. J. A. Thompson, R. D. Wells, *Nature (London)* **334** (1988) 87.
120. G. H. Hamm, G. N. Cameron, *Nucleic Acids Res.* **14** (1986) 5 – 9.
121. H. S. Bilofsky et al., *Nucleic Acids Res.* **14** (1986) 1 – 4.
122. E. Y. Chen, P. H. Seeburg, *DNA* **4** (1985) 165.
123. F. Sanger et al., *Proc. Natl. Acad. Sci. USA* **74** (1977) 5463.
124. A. M. Maxam, W. Gilbert, *Proc. Natl. Acad. Sci. USA* **74** (1977) 560.
125. M. M. Yang, D. C. Youvan, *Bio/Technology* **7** (1989) 576.
126. J. A. Brumbaugh et al., *Proc. Natl. Acad. Sci. USA* **85** (1988) 5610.
127. L. M. Smith, *Genet. Eng.* **10** (1988) 91.

128. W. Ansorge et al., *Nucleic Acids Res.* **15** (1987) 4593.
129. E. R. Mardis, B. A. Roe, *BioTechniques* **7** (1989) 840.
130. J. H. Jett et al., *J. Biomol. Struct. Dyn.* **7** (1989) 301.
131. J. A. Sorge, L. A. Blinderman, *Proc. Natl. Acad. Sci. USA* **86** (1989) 9208.
132. C. D. Carpenter, A. E. Simon, *BioTechniques* **8** (1990) 26.
133. E. P. Nikonowicz et al., *Biochemistry* **29** (1990) 4193.
134. F. J. M. van de Ven, C. W. Hilbers, *Eur. J. Biochem.* **178** (1988) 1.
135. C. de los Santos et al., *Biochemistry* **28** (1989) 7282.
136. S.-H. Chou et al., *Biochemistry* **28** (1989) 2435.
137. E. P. Geiduschek, G. P. Tocchini-Valentini, *Annu. Rev. Biochem.* **57** (1988) 873.
138. J. M. Sperrazza et al., *Gene* **31** (1984) 17.
139. M. I. Oakes, J. A. Lake, *J. Mol. Biol.* **211** (1990) 897.
140. M. I. Oakes et al., *J. Mol. Biol.* **211** (1990) 907.
141. D. D. Dunlap, C. Bustamante, *Nature (London)* **342** (1989) 204.
142. P. K. Hansma et al., *Science (Washington, D. C.)* **242** (1988) 209.
143. R. J. Driscoll et al., *Nature (London)* **346** (1990) 294.
144. C. Malvy, *Anal. Biochem.* **143** (1984) 158.
145. A. E. Dahlberg et al., *J. Mol. Biol.* **41** (1969) 139.
146. S. Peats, *Anal. Biochem.* **140** (1984) 178.
147. C. R. Merril, *Nature (London)* **343** (1990) 779.
148. N. C. Mills, J. Ilan, *Electrophoresis (Weinheim, Fed. Republ. Ger.)* **6** (1985) 531.
149. C. E. Willis, G. P. Holmquist, *Electrophoresis (Weinheim, Fed. Republ. Ger.)* **6** (1985) 259.
150. V. T. Kung et al., *Anal. Biochem.* **187** (1990) 220.
151. A. Glazer et al., *Proc. Natl. Acad. Sci. USA* **87** (1990) 3851.
152. J. Maddox, *Nature (London)* **345** (1990) 381.
153. N. C. Stellwagen, *Adv. Electrophor.* **1** (1987) 177.
154. C. R. Cantor et al., *Annu. Rev. Biophys. Chem.* **17** (1988) 287.
155. S. B. Smith et al., *Science (Washington, D. C.)* **243** (1989) 203.
156. S. M. Clark et al., *Sience (Washington, D. C.)* **241** (1988) 1203.
157. E. Lai et al., *BioTechniques* **7** (1989) 34.
158. M. V. Olson, *J. Chromatogr.* **470** (1989) 377.
159. A. M. Michelson, A. R. Todd, *J. Chem. Soc.* 1955, 2632 – 2638.
160. H. G. Khorana, G. M. Tener, J. G. Moffatt, E. H. Pol, *Chem. Ind. (London)* **34** (1956) 1523 – 1531.
161. H. G. Khorana, *Science (Washington, D. C.)* **203** (1979) 614.
162. R. B. Merrifield, *J. Am. Chem. Soc.* **85** (1962) 3821 –3827.
163. R. L. Letsinger, W. B. Lunsford, *J. Am. Chem. Soc.* **85** (1963) 3045 – 3046.
164. H. Schaller, G. Weimann, B. Lerch, H. G. Khorana, *J. Am. Chem. Soc.* **85** (1963) 3821 – 3827.
165. G. S. Ti, B. L. Gaffney, R. A. Jones, *J. Am. Chem. Soc.* **104** (1982) 1316 – 1319.
166. C. B. Reese, A. Ubasawa, *Tetrahedron Lett.* **21** (1980) 2265 – 2268.
167. A. Krazewski, J. Stawinski, M. Wiewiorowski, *Nucleic Acids Res.* **8** (1980) 2301 – 2305.
168. R. T. Pon, N. Usman, M. J. Damha, K. K. Ogilvie, *Nucleic Acids Res.* **14** (1985) 6453 – 6469.
169. K. K. Ogilvie, R. T. Pon, M. J. Damha, *Nucleic Acids Res.* **13** (1985) 6447 – 6465.
170. F. Himmelsbach et al., *Tetrahedron Lett.* **11** (1981) 59 – 72.
171. B. C. Froehler, M. D. Matteucci, *Nucleic Acids Res.* **22** (1983) 8031 – 8036.
172. J. C. Schulhof, D. Molko, R. Teoule, *Nucleic Acids Res.* **15** (1987) 397 – 416.
173. J. C. Schulhof, D. Molko, R. Teoule, *Nucleic Acids Res.* **16** (1988) 319 – 326.
174. N. Usman, K. K. Ogilvie, M.-Y. Jiang, R. J. Cedergren, *J. Am. Chem. Soc.* **109** (1987) 7845 – 7854.
175. B. S. Sproat, M. J. Gait in M. J. Gait (ed.): *Oligonucleotide Synthesis, A Practical Approach,* IRL Press, Oxford – Washington 1984, pp. 83 – 115.
176. C. B. Reese, R. C. Titmas, L. Yau, *Tetrahedron Lett.* **19** (1978) 2727 – 2730.
177. S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **22** (1981) 1859 – 1862.
178. L. J. McBride, M. H. Caruthers, *Tetrahedron Lett.* **24** (1983) 245.
179. N. D. Sinha, J. Biernat, J. McMagnus, H. Köster, *Nucleic Acids Res.* **12** (1984) 4539 – 4557.
180. H. Ito, Y. Ike, S. Ikuta, K. Itakura, *Nucleic Acids Res.* **10** (1982) 1755.
181. K. Miyoshi, R. Arentzen, T. Huang, K. Itakura, *Nucleic Acids Res.* **8** (1980) 5507.
182. R. Crea, T. Horn, *Nucleic Acids Res.* **8** (1980) 2331.

183. M. D. Matteucci, M. H. Caruthers, *J. Am. Chem. Soc.* **103** (1981) 3171 – 3174.
184. V. A. Efimov, S. V. Reverdatto, O. B. Chakhmakhcheva, *Nucleic Acids Res.* **9** (1982) 6675 – 6694.
185. S. P. Adams et al., *J. Am. Chem. Soc.* **105** (1983) 661 – 663.
186. M. J. Nemer, K. K. Ogilvie, *Tetrahedron Lett.* **21** (1980) 4159 – 4162.
187. T. G. Wadell, D. E. Leyden, M. T. de Bello, *J. Am. Chem. Soc.* **103** (1981) 5303 – 5307.
188. M. H. Caruthers in H. G. Gassen, A. Lang (eds.): *Chemical and Enzymatic Synthesis of Gene Fragments: A Laboratory Manual,* Verlag Chemie, Weinheim 1982, pp. 71 – 79.
189. W. Bannwarth, P. Iaiza, *DNA* **5** (1986) 413 – 419.
190. F. Chow, T. Kempe, G. Palm, *Nucleic Acids Res.* **9** (1981) 2807.
191. J. C. Catlin, F. Cramer, *J. Org. Chem.* **38** (1973) 245 –250.
192. S. A. Narang, H. M. Hsiung, R. Brousseau, *Methods Enzymol.,* **68** (1979) 90 – 98.
193. S. S. Jones et al., *Tetrahedron* **36** (1980) 3075 – 3085.
194. J. H. van Boom et al., *J. Chem. Soc. Chem. Commun.* 1971, 869 – 871.
195. E. Ohtsuka, S. Iwai in S. Narang (ed.): *Synthesis and Application of DNA and RNA,* Academic Press, San Diego – New York – Berkley – Boston 1987.
196. J. H. van Boom, *J. Chem. Soc. Chem. Commun.* 1976, 167.
197. P. J. Garegg, T. Regberg, J. Stawinski, R. Strömberg, *Chem. Scr.* **25** (1985) 280 – 282.
198. B. C. Froehler, M. D. Matteucci *Tetrahedron Lett.* **27** (1986)469 – 472.
199. B. C. Froehler, *Tetrahedron Lett.* **27** (1986) 5575 –5578.
200. B. C. Froehler, P. Ng, M. D. Matteucci, *Nucleic Acids Res.* **16** (1988) 4831 – 4839.
201. J. Stawinski, R. Strömberg, M. Thelin, E. Westman, *Nucleic Acids Res.* **16** (1988) 9285 – 9298.
202. J. H. van Boom et al., *Nucleic Acids Res.* **4** (1977) 1047.
203. E. E. van Tamelen, S. V. Daub, *J. Am. Chem. Soc.* **99** (1977) 3526.
204. L. W. McLaughlin, J. U. Krusche in H. G. Gassen, A. Lang (eds.): *Chemical and Enzymatic Synthesis of Gene Fragments A Laboratory Manual,* : Verlag Chemie, Weinheim 1982.
205. L. W. McLaughlin, N. Piel in M. J. Gait (ed.): *Oligonucleotide Synthesis, A Practical Approach,* IRL Press, Oxford – Washington 1984, pp. 117 –132.
206. G. Schmidt, H. Seliger, *J. Chromatogr.* **397** (1987) 141 – 151.
207. R. Wu in M. J. Gait (ed.): *Oligonucleotide Synthesis, A Practical Approach,* IRL Press, Oxford – Washington 1984, pp. 135 – 151.
208. A. M. Maxam, W. Gilbert, *Methods Enzymol.* **65** (1980) 499 – 503.
209. G. Zon, *Pharmaceutical Research* **5** (1988) no. 9, 539 – 549.
210. C. Hélène, J.-J. Toulmé, *Biochim. Biophys. Acta* **1049** (1990) 99 – 125.
211. C. A. Stein, J. S. Cohen, *Cancer Research* **48** (1988) 2659 – 2668.
212. W. J. Stec, G. Zon, W. Egan, B. Stec, *J. Am. Chem. Soc.* **106** (1984) 6077 – 6079.
213. B. C. Froehler, *Tetrahedron Lett.* **27** (1986) 5565 – 5568.
214. A. Andrus, J. W. Efcavitch, L. J. McBride, B. Giusti, *Tetrahedron Lett.* **29** (1988) 861 – 864.
215. B. H. Dahl, K. Bjergarde, V. B. Sommer, O. Dahl, *Nucleosides & Nucleotides,* **8** (1989) 1023 – 1027.
216. S. L. Beaucage et al., *J. Am. Chem. Soc.* **112** (1990) 1254 – 1255.
217. P. C. J. Kamer et al., *Tetrahedron Lett.* **30** (1989) 6757 – 6760.
218. W. K.-D. Brill, J.-Y. Tang, Y.-X. Ma, M. H. Caruthers, *J. Am. Chem. Soc.* **111** (1989) 2321 – 2322.
219. P. S. Miller, K. N. Fang, N. S. Kondo, P. O. P. Ts'o, *J. Am. Chem. Soc.* **93** (1971) 6657 – 6665.
220. P. S. Miller et al., *Biochimie* **67** (1985) 769 – 776.
221. P. S. Miller et al., *Nucleic Acids Res.* **11** (1983) 6225 – 6242.
222. P. S. Miller et al., *Biochemistry* **25** (1986) 5092 – 5097.
223. E. de Vroom et al., *Recl. Trav. Chim. Pays-Bas* **106** (1987) 65 – 66.
224. H. Inoue et al., *Nucleic Acids Res.* **15** (1987) 6131 – 6148.
225. B. S. Sproat, B. Beijer, A. Iribarren, *Nucleic Acids Res.* **18** (1990) 41 – 49.
226. U. Sequin, *Experientia* **29** (1973) 1059 – 1062.
227. J. L. Imbach, B. Rayner, F. Morvan, *Nucleosides & Nucleotides* **8** (1989) 627 – 648.
228. C. Hélène et al., *Biochimie* **66** (1985) 777 – 783.
229. A. Zerial, N. T. Thuong, C. Hélène, *Nucleic Acids Res.* **15** (1987) 9909 – 9919.

230. V. V. Vlassov et al., *Nucleic Acids Res.* **14** (1986) 4065 – 4076.
231. H. B. Gamper, G. D. Cimino, J. E. Hearst, *J. Mol. Biol.* **197** (1987) 349 – 362.
232. Y. B. Shi, H. B. Gamper, J. E. Hearst, *J. Mol. Biol.* **263** (1988) 527 – 534.
233. P. S. Miller, P. O. P. Ts'o, *Anti-Cancer Drug Des.* **2** (1987) 117 – 128.
234. U. Asseline et al., *Proc. Natl. Acad. Sci USA* **81** (1984) 3297 – 3301.
235. Le Doan et al., *Nucleic Acids Res.* **15** (1987) 7749 – 7760.
236. D. Praseuth, L. Perrouault, T. Le Doan, *Proc. Natl. Acad. Sci USA* **85** (1988) 1349 – 1353.
237. D. Praseuth et al., *Biochemistry* **27** (1988) 3031 – 3038.
238. C. Hélène, T. Le Doan, N. T. Thuong in P. E. Nielsen (ed.): *Photochemical Probes in Biochemistry,* Kluwer Publ., Norwell MA, 1989, pp. 219 – 229.
239. C. Hélène, N. T. Thuong, T. Saison-Behmoaras, J. C. Francois, *Tibtech* **7** (1989) 310 – 315.
240. A. S. Boutorin et al., *FEBS Lett.* **172** (1984) 43 –46.
241. B. C. F. Chu, L. E. Orgel, *Proc. Natl. Acad. Sci USA* **82** (1985) 963 – 967.
242. G. B. Dreyer, P. E. Dervan, *Proc. Natl. Acad. Sci. USA* **82** (1985) 968 – 972.
243. C. H. B. Chen, D. S. Sigman, *Proc. Natl. Acad. Sci. USA* **83** (1986) 7147 – 7151.
244. C. H. B. Chen, D. S. Sigman, *J. Am. Chem. Soc.* **110** (1988) 6570 – 6572.
245. J.-C. Francois et al., *Biochemistry* **27** (1988) 2272 – 2276.
246. J.-C. Francois et al., *J. Biol. Chem.* **264** (1989) 5891 – 5898.
247. T. Le Doan et al., *Nucleic Acids Res.* **15** (1987) 8643 – 8659.
248. D. Corey, P. Schultz, *Science (Washington, D. C.)* **238** (1987) 1401 – 1403.
249. R. N. Zuckermann, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **86** (1989) 1766 – 1770.
250. Applied Biosystems: DNA Synthesizer, User Bulletin, No. 49, Foster City 1988.
251. Beckman Instruments, Aminomodifiers and Easy Label Kits, User Manual, Palo Alto 1990.
252. B. C. F. Chu, G. M. Wahl, L. E. Orgel, *Nucleic Acids Res.* **11** (1985) 6513 – 6529.
253. E. Jablonski, E. W. Moomaw, R. H. Tullis, J. L. Ruth, *Nucleic Acids Res.* **14** (1986) 6115 – 6128.
254. M. Lemaitre, B. Bayard, B. Lebleu, *Proc. Natl. Acad. Sci. USA* **84** (1987) 648 – 652.
255. M. Lemaitre, C. Bisbal, B. Bayard, B. Lebleu, *Nucleosides & Nucleotides* **6** (1987) 311 – 315.
256. R. Zuckermann, D. Corey, P. Schultz, *Nucleic Acids Res.* **15** (1987) 5305 – 5321.
257. P. S. Nelson, R. A. Frye, E. Liu, *Nucleic Acids Res.* **17** (1989) 7187 – 7194.
258. S. Agrawal, J.-Y. Tang, *Tetrahedron Lett.* **31** (1990) 1543 – 1546.
259. R. L. Letsinger, M. E. Schott, US 4 547 569, 1985.
260. A. Jäger, M. J. Levy, S. M. Hecht, *Biochemistry* **27** (1988) 7237 – 7246.
261. R. P. Langer, A. A. Waldrop, D. C. Ward, *Proc. Natl. Acad. Sci. USA* **78** (1981) 6635 – 6637.
262. A. F. Cook, E. Vuocolo, C. L. Brakel, *Nucleic Acids Res.* **16** (1988) 4077 – 4095.
263. M. S. Urdea et al., *Gene* **61** (1987) 253 – 264.
264. I. C. Gillam, G. M. Tener, *Anal. Biochem.* **157** (1986) 199 – 207.
265. U. Pieles, B. S. Sproat, P. Neuner, F. Cramer, *Nucleic Acids Res.* **17** (1989) 8967 – 8978.
266. G. Gebeyehu et al., *Nucleic Acids Res.* **16** (1987) 4937 – 4534.
267. P. Tchen, R. P. P. Fuchs, E. Sage, M. Leng, *Proc. Natl. Acad. Sci. USA* **81** (1984) 3466 – 3470.
268. J. A. Washington, G. L. Woods in B. Swaminathan, G. Prakash (eds.): *Nucleic Acid and Monoclonal Antibody Probes,* Marcel Dekker, New York 1989, p. 319.
269. F. C. Tenover, *Clin. Microbiol. Rev.* **1** (1988) 82.
270. E. Uhlmann, A. Peyman, *Chemical Rev.* **90** (1990) 543 – 584.
271. C. Hélène, J.-J. Toulmé, *Biochim. Biophys. Acta* **1049** (1990) 99 – 125.
272. J. Brandsma, G. Miller, *Proc. Natl. Acad. Sci. USA* **77** (1980) 6851.
273. P. McIntyre, G. R. Stark, *Anal. Biochem.* **174** (1988) 209.
274. E. M. Southern, *J. Mol. Biol.* **98** (1975) 503.
275. T. R. Moench, *Mol. Cell. Probes* **1** (1987) 195.
276. J. E. Landegent et al., *Nature (London)* **317** (1985) 175.
277. P. Lichter et al., *Science (Washington, D. C.)* **247** (1990) 64.
278. M. Ranki et al., *Gene* **21** (1983) 77.
279. P. J. Nicholls, A. D. B. Malcolm, *J. Clin. Lab. Anal.* **3** (1989) 122.
280. M. Virtanen et al., *J. Clin. Microbiol.* **20** (1984) 1083.
281. A. C. Syvaenen et al., *Nucleic Acids Res.* **14** (1986) 5037.
282. M. S. Ellwood et al., *Clin. Chem.* **32** (1986) 1631.

283. C. P. H. Vary, *Nucleic Acids Res.* **15** (1987) 6883.
284. J. A. Matthews, L. J. Kricka, *Anal. Biochem.* **169** (1988) 1.
285. L. J. Arnold, Jr., et al., *Clin. Chem.* **35** (1989) 1588.
286. R. P. Viscidi et al., *J. Clin. Microbiol.* **23** (1986) 311.
287. A. H. N. Hopman et al., *Exp. Cell. Res.* **169** (1987) 357.
288. G. H. Keller et al., *Anal. Biochem.* **170** (1988) 441.
289. A. C. Forster et al., *Nucleic Acids Res.* **13** (1985) 745.
290. G. H. Keller et al., *Anal. Biochem.* **177** (1989) 392.
291. K. Mühlegger et al., *Biol. Chem. Hoppe-Seyler* **371** (1990) 953.
292. K. B. Mullis, F. A. Faloona, *Methods Enzymol.* **155** (1987) 335.
293. R. K. Saiki et al., *Science (Washington, D. C.)* **230** (1985) 1350.
294. H. Lee et al., *Science (Washington, D. C.)* **244** (1989) 471.
295. D. B. Schowalter, S. S. Sommer, *Analyt. Biochemistry* **177** (1988) 90 – 94.
296. T. Lion, O. A. Haas, *Analyt. Biochemistry* **188** (1990) 335 – 337.
297. D. Y. Kwoh et al., *Proc. Natl. Acad. Sci. USA* **86** (1989) 1173.
298. D. Y. Wu, R. B. Wallace, *Genomics* **4** (1989) 560.
299. K. J. Barringer et al., *Gene* **89** (1990) 117.
300. J. C. Guatelli et al., *Proc. Natl. Acad. Sci. USA* **87** (1990) 1874.
301. J. van Brunt, *Bio/Technology* **8** (1990) 291.
302. P. M. Lizardi et al., *Bio/Technology* **6** (1988) 1197.
303. H. Lomeli et al., *Clin. Chem. (London)* **35** (1989) 1826.
304. H. Wolf et al., *J. Virol. Methods* **13** (1986) 1.
305. M. S. Urdea et al., *Gene* **61** (1987) 253.
306. P. D. Fahrlander, *Bio/Technology* **6** (1988) 1165.
307. M. Berninger et al., *J. Med. Virol.* **9** (1982) 57.
308. H. A. Rotbart et al., *J. Clin. Microbiol.* **26** (1988) 2669.
309. J. J. Hogan in B. Swaminathan, G. Prakash (eds.): *Nucleic Acid and Monoclonal Antibody Probes,* Marcel Dekker, New York 1989.
310. S. Huovinen et al., *Antimicrob. Agents Chemother.* **32** (1988) 175.
311. D. N. Cooper, J. Schmidtke, *Hum. Genet.* **73** (1986) 1.
312. S. Nishimura, T. Sekiya, *Biochem. J.* **243** (1987) 313.
313. P. Benn et al., *Cancer Genet. Cytogenet.* **29** (1987) 1.
314. A. J. Levine, *BioEssays* **12** (1990) 60.
315. R. A. Weinberg, *Cancer Res.* **49** (1989) 3713.
316. M. A. Lovell, *Clin. Chem. (Winston-Salem, N.C.),* Suppl. 35 (1989) 1343.
317. H. A. Erlich et al., *Bio/Technology* **4** (1986) 975.
318. A. J. Jeffreys, *Biochem. Soc. Trans.* **15** (1987) 309.
319. H. Zischler et al., *Hum. Genet.* **82** (1989) 227.
320. A. H. Cawood, *Clin. Chem. (Winston-Salem, N.C.)* **35** (1989) 1832.
321. J. Davoll, B. Lythgoe, A. R. Todd, *J. Chem. Soc.* 1948, 1685.
322. J. Davoll, B. A. Lowy, *J. Am. Chem. Soc.* **73** (1951) 1650.
323. N. Yamaoka, K. Aso, H. Matsuda, *J. Org. Chem.* **30** (1965) 149.
324. G. E. Hilbert, T. B. Johnson, *J. Am. Chem. Soc.* **52** (1930) 4489.
325. J. Pliml, M. Prystas, *Adv. Heterocycl. Chem.* **8** (1967) 115.
326. T. Sato, T. Shimadate, Y. Ishido, *Chem. Abstr.* **56** (1962) 11 692 g.
327. L. Birkhofer, A. Ritter, H. P. Küelthau, *Angew. Chem.* **75** (1963) 209.
328. T. Nishimura, B. Shimizu, I. Iwai, *Chem. Pharm. Bull.* **11** (1963) 1470.
329. U. Niedballa, H. Vorbrüggen, *Angew. Chem., Int. Ed. Engl.* **9** (1970) 461.
330. A. Holy, I. Votruba, *Nucleic Acid Symp. Ser.* **18** (1987) 69.
331. F. Seela, H. D. Winkeler, *J. Org. Chem.* **47** (1982) 226.
332. D. H. Shannahoff, R. A. Sanchez, *J. Org. Chem.* **38** (1973) 593.
333. V. E. Marquez, M. I. Lim, *Med. Res. Rev.* **6** (1986) 1.
334. S. Hanessian, A. C. Parnet, *Adv. Carbohydr. Chem. Biochem.* **33** (1976) 111.
335. K. Ogata, *Adv. Appl. Microbiol.* **19** (1975) 209.
336. K. Ogata, S. Kinoshita, T. Tsunoda, K. Aida: *Microbial Production of Nucleic Acid-Related Substances,* Halsted Press, New York 1976.
337. S. L. Haynie, D. N. Whitesides, *Appl. Biochem. Biotech.* **23** (1990) 205.
338. D. J. Merkler, V. L. Schramm, *Anal. Biochem.* **167** (1987) 148.
339. L. A. Slotin, *Synthesis* 1977, 737.
340. G. A. Robinson, R. W. Butcher, E. W. Sutherland: *Cyclic AMP,* Academic Press, New York – London 1971.

341. W. H. Cook, D. Lipkin, R. Markham, *J. Am. Chem. Soc.* **79** (1957) 3607.
342. M. S. Amer in N. J. Harper, A. B. Simmonds (eds.): *Advances in Drug Research,* **vol. 12,** Academic Press, New York – London 1977, p. 1.
343. F. Eckstein, *Ann. Rev. Biochem.* **54** (1985) 367.
344. R. G. Yount, D. Babcock, D. Ojala, W. Ballantyne, *Biochemistry* **10** (1971) 2484.
345. T. C. Myers, K. Nakamura, J. W. Flesher, *J. Am. Chem. Soc.* **85** (1963) 3292.
346. G. R. Pettit: *Synthetic Nucleotides,* vol. I, Van Nostrand Reinhold, New York 1972.
347. M. Yoshikawa, T. Kato, T. Takenishi, *Tetrahedron Lett.* **50** (1967) 5065.
348. J. G. Moffatt, H. G. Khorana, *J. Am. Chem. Soc.* **83** (1961) 649.
349. D. E. Hoard, D. G. Ott, *J. Am. Chem. Soc.* **87** (1965) 1785.
350. A. M. Michelson, *Biochim. Biophys. Acta* **91** (1964) 1.
351. J. Ludwig, *Acta Biochim. Biophys. Acad. Sci. Hung.* **16** (1981) 131.
352. E. De Clercq, *Int. Congr. Ser. Excerpta Med.* **750** (1987) 631.
353. J. C. Martin: "Nucleotide Analogues as Antiviral Agents," *ACS Symp. Ser.* **401** (1989).