Nucleotides: Structure and Properties

Richard Peter Bowater, *University of East Anglia, Norwich, UK*

Nucleotides consist of a nitrogen-containing base, a five-carbon sugar, and one or more phosphate groups. Cells contain many types of nucleotides, which play a central role in a wide variety of cellular processes, including metabolic regulation and the storage and utilization of genetic information.

Nomenclature

A complex and conflicting terminology for nucleotides and related compounds has been used, but standard definitions have been suggested to aid understanding. Nomenclature in general use is listed in **Table 1**.

All nucleotides have three fundamental components (**Figure 1**).

1. Base: Also referred to as heterocycles, these nitrogen-containing ring compounds are derivatives of purine or pyrimidine. The atoms within the purine/pyrimidine rings do not vary and they are given the same number in different nucleotides. A variety of chemical groups can be bonded at different positions to the ring constituents. Although one structure predominates for each base, tautomeric forms occur in minor amounts (**Figure 2**).
2. Sugar: A five-carbon sugar, usually ribose, is linked to the base. Each carbon atom is numbered and, to allow their distinction from atoms in the base, the number is followed by a prime mark: thus ribose has five carbon atoms, numbered 1′ to 5′. The sugars are locked into a five-membered furanose ring by the bond from C1′ of the sugar to the base.
3. Phosphate ester: Phosphate groups are attached to the sugar as esters, with the most common site of ester-ification in natural compounds being via the hydroxyl at the C5′. Typically, one, two or three phosphates are joined, producing mono-, di- and triphosphates, respectively. (Note that the prefixes bis and tris are used in reference to molecules with two or three phosphates attached to different hydroxyls on the sugar.)

Nucleosides have very similar structures to nucleotides, but consist only of the base bonded to a sugar. Thus, a nucleotide can also be referred to as a nucleoside phosphate. It should be emphasized that, in the majority of situations, the structures of nucleosides and related nucleotides are very similar.

Table 1 Nomenclature of commonly occurring bases, ribonucleosides and ribonucleotides

Base	Ribonucleoside	Ribonucleotide (5′-monophosphate)
Purines (*Pur*)	Purine nucleoside (R or Puo)	
Adenine (Ade)	Adenosine (A or Ado)	Adenylate (5′-AMP or Ado-5′ or pA)
Guanine (Gua)	Guanosine (G or Guo)	Guanylate (5′-GMP or Guo-5′-P or pG)
Hypoxanthine (Hyp)	Inosine (I or Ino)	Inosinate (5′-IMP or Ino-5′-P or pI)
Xanthine (Xan)	Xanthosine (X or Xao)	Xanthosinate (5′-XMP or Xao-5′-P or pX)
Pyrimidines (*Pyr*)	Pyrimidine nucleoside (Y or Pyd)	
Cytosine (Cyt)	Cytidine (C or Cyd)	Cytidylate (5′-CMP or Cyd-5′-P or pC)
Orotate (Oro)	Orotidine (O or Ord)	Orotidylate (5′-OMP or Ord-5′-P or pO)
Thymine (Thy)	Thymidine (T or Thd)	Thymidylate (5′-TMP or Thd-5′-P or pT)
Uracil (Ura)	Uridine (U or Urd)	Uridylate (5′-UMP or Urd-5′-P or pU)

Frequently used abbreviations are given in parentheses. Designation of the 2′-deoxyribo form of these nucleosides and nucleotides is done by prefixing with deoxy (or d for abbreviated forms). Multiply phosphorylated and nucleotides occur and abbreviated notations use D for diphosphate and T for triphosphate for substitutions on the same sugar hydroxyl. Bis and tris are used to refer to substitutions of the same group on different sugar hydroxyls, e.g. 3′,5′-pAp is a biphosphate.
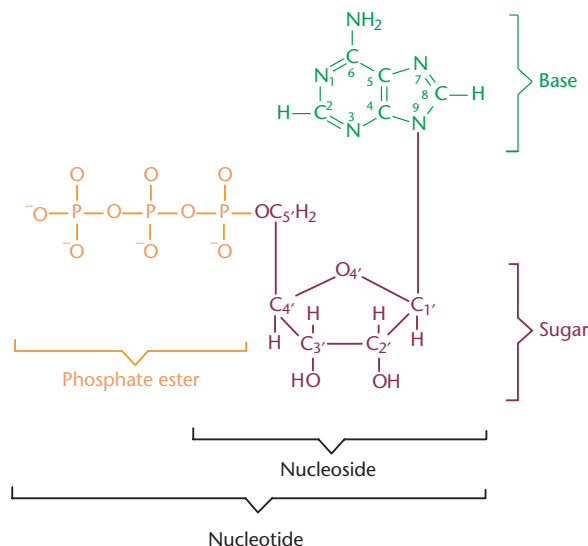
Figure 1 Chemical structure of adenosine triphosphate (ATP), a nucleotide. All nucleotides consist of a base, a sugar and a phosphate ester. These constituent parts are shown for ATP, where the base is adenine (shown in green), the sugar is β-D-ribose (shown in purple), and the phosphate is triphosphate (shown in orange). The base-and-sugar moiety is referred to as a nucleoside (termed adenosine for the nucleoside shown). The positions of atoms in the base and sugar are numbered. Purine bases contain atoms numbered 1–9 as shown, and pyrimidine bases are numbered 1–6 (see **Figure 2**). Atoms within the sugar (identified by the prime mark) are numbered similarly for all nucleotides containing β-D-ribose. Bond lengths are not drawn to scale.

In addition to their monomeric state, nucleotides exist in polymeric forms, called nucleic acids, and there are two closely related types: ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). These polymeric forms have a directional polarity within their sequence, which occurs because the 3′-OH of one nucleotide binds to the 5′-phosphate of the next by a phosphodiester linkage. Thus, one end of the molecule has a 5′-phosphate and the other end has a 3′-OH. It is common practice that sequences are written starting with the nucleotide containing the 5′-phosphate at the left.

Generally, the deoxy forms of nucleotides and nucleosides are specified. Thymine bases were originally thought to occur only in DNA and the terms thymidine, thymidine monophosphate and thymidylate were accepted to mean the deoxy form. However, ribonucleotides with thymine as the base do occur in transfer RNA and have been synthesized *in vitro*. In current practice, the terms thymidine and deoxythymidine are used interchangeably, and the prefix ribo- should be used for ribonucleotides of thymine. To avoid ambiguity, ribothymidines should always be designated as such.

Nucleotides are usually referred to in an abbreviated form (**Table 1**). For example, molecules with respectively one, two and three phosphates attached to adenosine are AMP (adenosine monophosphate), ADP (adenosine

diphosphate) and ATP (adenosine triphosphate); the respective deoxy variants are dAMP, dADP and dATP. Nucleotides without specified bases are written as NMP, NDP and NTP. Polynucleotides are referred to by a number of different shortened nomenclatures, usually incorporating single capital letters to distinguish the different bases of the nucleotides. The distinction between deoxyribonucleotides and ribonucleotides is frequently not specified other than saying that the sequence relates to RNA or DNA. In some abbreviations it is desirable to indicate the phosphates within the polynucleotide thus: pApGpTpC or $^{5'}$pApGpTpC$^{3'}$-OH. Specific nomenclature is used for nucleic acids containing repetitive sequences. For example, a homopolymer made from deoxyadenylate is called poly(deoxyadenylate) or poly(dA), while heteropolymers with alternating sequences include poly(deoxyadenylate-deoxythymidylate) or poly(dA-dT). Note that a related molecule with dA and dT distributed randomly over the polymeric chain is named similarly, but a comma replaces the hyphen to give poly(dA,dT).

The bases of certain nucleotides can bind to others via hydrogen bonds. These noncovalent interactions (called base pairing) allow different nucleic acid molecules to recognize each other. Generally, G pairs with C and A with T (in DNA) or U (in RNA). Thus, two nucleic acid polymers containing complementary sequences can form very stable interactions, and such pairing forms the basis of the double-stranded nature of the genetic material. Apart from a few specific exceptions, base pairing between two different strands occurs in such a way that the strands have opposite polarity. References to double-stranded nucleic acids frequently give the sequence of only one strand because that of the complementary strand is fixed by the base-pairing rules.

Under certain conditions, unusual base pairs can be formed and these allow molecules to be formed from the interaction of higher numbers of strands (such as triplexes and quadruplexes).

Nucleosides and Nucleotides

Within cells, ribonucleotides are synthesized *de novo* from simple building blocks or they are obtained from the recycling of preformed bases. Deoxyribonucleotides are synthesized by the reduction of ribonucleotides by the reduced form of nicotinamide–adenine dinucleotide phosphate (NADPH). Purine and pyrimidine bases are built *de novo* from amino acids, tetrahydrofolate derivatives, NH_4^+ and CO_2. The sugar phosphate moiety of ribonucleotides is donated by 5-phosphoribosyl 1-pyrophosphate. Nucleotides can also be synthesized via a salvage reaction using preformed bases, and this is simpler and requires much less energy than the reactions of the *de novo* pathway.

MAJOR FORMS

MINOR FORMS
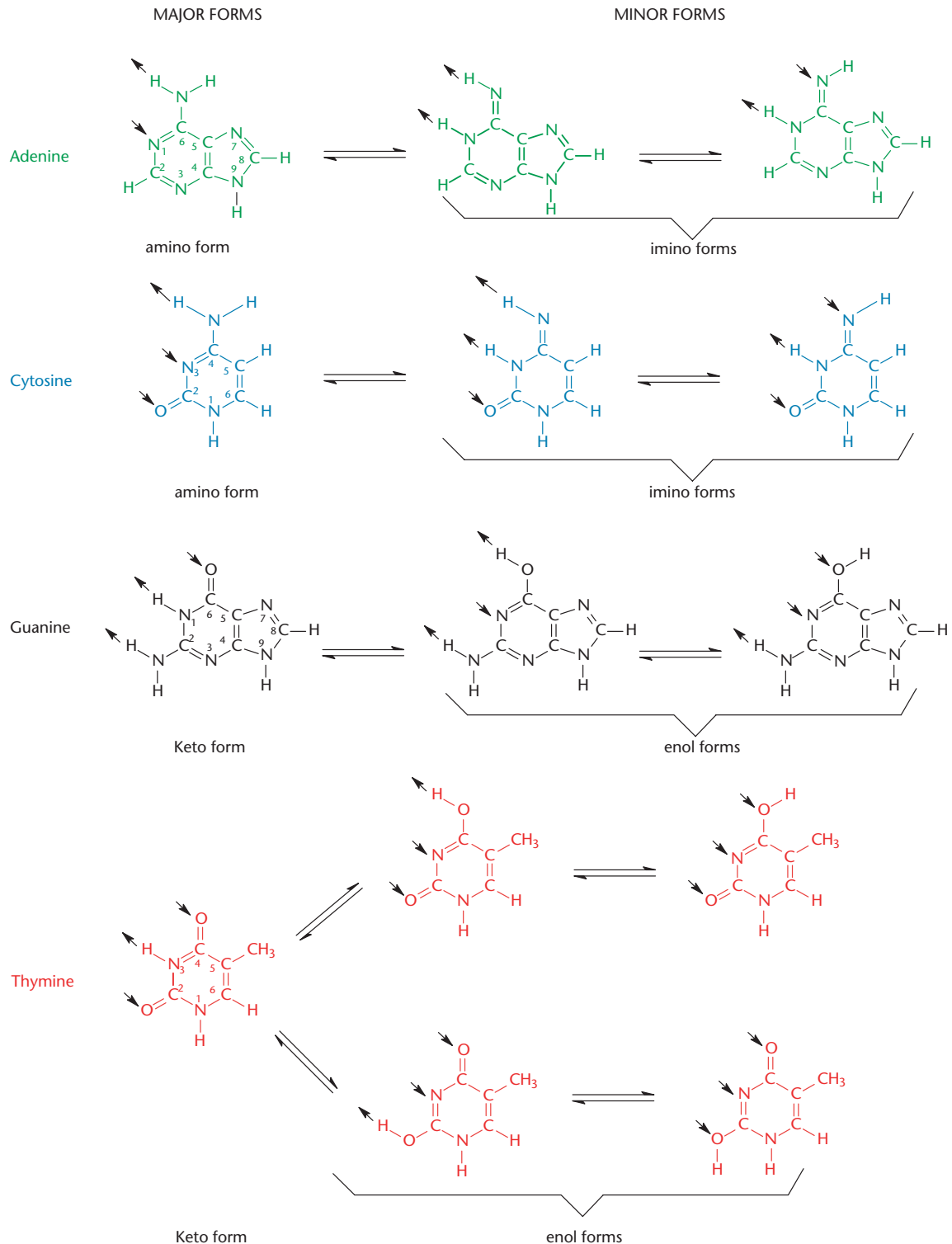
Figure 2 Structures and tautomeric equilibria of the DNA bases. Atoms within bases are numbered, with N1 of pyrimidines and N9 of purines being bonded to C1′ of the sugar in nucleosides and nucleotides. Tautomeric forms shown to the left are the major ones and the imino and enol forms shown to the right are present in very small amounts. Participation in hydrogen bonds is shown by the arrows: hydrogen bond donors are shown by arrows directed away from the atom, acceptors are shown by arrows directed towards the atom.

Nucleotides play key roles in nearly all biochemical processes, including the following:

- Nucleic acid synthesis: Nucleotides are the activated precursors of DNA and RNA.
- Energy transfer: ATP is used as a universal currency of energy in biological systems and GTP is utilized in many systems involved in movements of macromolecules.
- Coenzymes: Adenine nucleotides are components of some major coenzymes.
- Metabolism: Nucleotide derivatives are activated intermediates in many biosynthetic reactions, and they also serve as mediators to regulate metabolism.

Variations in the nucleotide composition of different cells are striking because their amounts are fundamentally linked to the rate of cell growth. Quantification is difficult in organisms in which growth rates may vary dramatically, but nucleotides and nucleic acids are estimated to comprise 5–10% of the total weight of rapidly dividing bacteria. In mammals the amounts vary from approximately 1% in muscle to 15–40% in thymus gland and sperm cells. The total concentration of nucleotides is approximately 300 mmol L^{-1} in bacteria in exponential growth but is usually a factor of 10 lower in mammalian tissue (**Table 2**).

Nucleotides are in constant flux between their free and polymeric states, and their levels within cells are regulated very efficiently. This is illustrated most vividly with the precursors for DNA, since an inadequate supply of any of the dNTPs is lethal, but an overabundance is mutagenic.

Table 2 Amounts of nucleotides in bacterial and mammalian cells

	Amount in bacterial cells (mmol L^{-1})	Amount in mammalian cells (mmol L^{-1})
DNA precursors		
dNMPs	43	Nd
DATP	0.18	0.013
DGTP	0.12	0.005
DCTP	0.07	0.022
DTTP	0.08	0.023
RNA precursors		
rNMPs	270	Nd
ATP	3.0	2.8
GTP	0.92	0.48
CTP	0.52	0.21
TTP	0.89	0.48

Values shown for bacterial cells are for *Escherichia coli* and *Salmonella typhimurium* in exponential growth, and those for mammalian cells are determined from diverse mammalian tissues. nd = not determined. Data are from Kornberg and Baker (1992), p. 54.

To aid regulation of cellular metabolism, cells have small pools of free nucleotides (with the exception of ATP).

Each somatic cell in an organism has the same amount (and sequence) of DNA, but these vary dramatically between species. The genome of each organism or virus is organized into a unique number of DNA–protein complexes, called chromosomes. The smallest genomes are carried by viruses and may be fewer than 10^4 base pairs. *Escherichia coli* has one chromosome of about 4.5×10^6 base pairs and *Saccharomyces cerevisiae* has 16 chromosomes for a total of approximately 2×10^7 base pairs. The DNA content of humans is approximately 3×10^9 base pairs in total, carried on 23 different pairs of chromosomes (46 chromosomes in each somatic cell).

In addition to the major nucleotides listed in **Table 1**, many other types are present as minor components within cells. Some of the minor nucleotides are required for specific cellular functions and are particularly prevalent within stable RNAs (rRNA and tRNA). In some cases the minor forms are unwanted byproducts of reactions occurring within the cell and these can be deleterious to cell survival. For example, endogenous and exogenous damage to DNA generates a variety of unusual nucleotides, and the cell has evolved a wide range of metabolic pathways to allow their removal. Molecules with more complicated structures than those shown in **Figure 1** and **Figure 2** are also utilized. For example, variants of purine nucleotides exist with polyphosphates (e.g. ppGpp and pppGpp) and with a monophosphate bound to the base in a cyclical fashion (e.g. cyclic AMP and cyclic GMP). Other nucleotides with diverse structures have been synthesized chemically; many have been tested as potential pharmaceuticals and they have also been particularly valuable in scientific analysis of nucleotide metabolism. Fluorescent derivatives of nucleotides are of exceptional importance. Such molecules are used during high-throughput automated sequencing of DNA, which is generating vast amounts of genome sequence data.

Physical properties

A variety of atoms within nucleotides and nucleosides are available to interact with other molecules, particularly via hydrogen-bond formation (**Figure 2**). The capacity of chemical groups (such as amino and keto groups) to act as hydrogen-bond acceptors or donors is linked to the propensity of their constituent atoms to dissociate and create a charge upon that group. Such dissociation relates to the acid–base behaviour of the chemical group and is usually quantified in terms of its pK_a, with lower numbers indicating greater ability to dissociate. The pK_a values for the five major bases found in these compounds are listed in **Table 3**. In reality, the bases are only weakly basic and A, C and G become protonated only below pH 4. The amide NH groups in G, T and U are deprotonated at pH above 9.

Table 3 pK_a values for the bases found in the major nucleosides and nucleotides

Base (site of protonation)	Nucleoside	3′-Nucleotide	5′-Nucleotide
Adenine (N1)	3.52	3.70	3.88
Cytosine (N3)	4.17	4.43	4.56
Guanine (N7)	3.30	3.50	3.60
Guanine (N1)	9.42	9.84	10.00
Thymine (N3)	9.93	–	10.47
Uracil (N3)	9.38	9.96	10.06

Data are from Blackburn and Gait (1996), p.18, and relate to 20°C and zero salt concentration. Since pK_a values greater than 9 correspond to loss of a proton and pK_a values less than 5 to capture of a proton, all bases are uncharged at physiological pH.

Thus, all of the bases are uncharged in the physiological pH range, and the same is true for the sugar. The phosphate groups in nucleotides display two main pK_a values: at pH values of 1–2, deprotonation occurs for each phosphate; at approximately pH 7, deprotonation occurs for the terminal phosphate of the nucleotide.

As is evident from this discussion, the manner in which nucleotides interact with other molecules is determined predominantly by their charge density. Charge density varies throughout the different nucleotides and is intimately linked to physical characteristics of these molecules, such as their tautomeric structure (**Figure 2**). Tautomers occur frequently in heterocyclic molecules owing to the possibility of hydrogen atoms attached to nitrogens being able to migrate to other, free nitrogens or to keto oxygens within the same molecule. An important consequence of this type of isomerization is that the hydrogen-bonding possibilities differ for each tautomer. Thus, base amino groups bear hydrogen atoms with partial positive charge and are good hydrogen-bond donors, whereas base keto oxygens and nitrogens are good acceptors of hydrogen bonds. In general, the oxygens of the sugar ring and phosphate ester are less important as hydrogen-bond acceptors, although these can make important contacts in interactions within nucleic acids and between nucleic acids and proteins.

Conformational features

The molecular geometry and conformational properties of the various components of a nucleotide are intrinsically related. Therefore, the conformation of specific atoms must be considered in relation to the whole molecule. Overall, nucleotides have rather compact shapes owing to interactions between several of their nonbonded atoms.

Pyrimidine and purine bases have a similar range of conformations whether they are part of a nucleotide or as individual units. In general, the heterocycles of the bases are planar. Although several tautomeric forms exist in principle, the keto and amino forms predominate (**Figure 2**). Typically, the enol and imino forms are no more than 0.1% of the total, although this amount can change dramatically if modified bases are considered or if base pairs are formed.

Details of the conformation of a nucleotide are accurately defined by a series of torsion angles specific for each bond. Each torsion angle accurately depicts the geometric relationship of four atoms in a sequence, e.g. A–B–C–D. The angle is defined as that between projected bonds A–B and C–D when viewed along the central bond B–C. Definitions of the torsion angles for nucleotides are shown in **Figure 3a** and **Table 4**. Torsion angles are given positive and negative values by the position of the first and last atoms (A and D) relative to the bond angle: if the smallest movement for the front atom to eclipse the back atom is in a clockwise direction it is given a positive value; if it is anticlockwise, it is given a negative value.

As with most molecules with rotational freedom about single bonds, certain conformations of nucleotides are usually preferred as a result of steric constraints. Ranges of torsion angle are commonly defined by the Klyne–Prelog nomenclature developed for spectroscopists, in which those of approximately 0° are referred to as *syn* and those of approximately 180° are referred to as *anti*. (Note, however, that the terms *syn* and *anti* have special meanings in regard to the glycosolic bonds of nucleosides (**Figure 3c**).) Angles subtending between 30° and 150° and between 210° and 330° are defined as clinal, with the remaining angles around 0° and 180° defined as periplanar. Since many of the torsion angles of nucleotides are interdependent, the approximate shapes of nucleotides can be described by considering only four parameters: the sugar pucker, the conformation of the bond between the base and sugar (β-glycosolic linkage), the orientation of C4′–C5′, and the shape of the phosphate ester bonds.

The five-membered furanose ring systems in nucleosides are twisted out of plane in order to minimize unfavourable nonbonded interactions between their substituents. This 'puckering' is in an envelope (E) form when four atoms are in a plane or in a twist (T, also termed 'half-chair') form when two adjacent atoms are displaced on opposite sides of a plane through the other three atoms (**Figure 3b**). Atoms most usually involved in deviations from the plane of the other sugar ring atoms are C2′ and C3′. If the carbon atom lies above the sugar ring, i.e. towards the base and C5′, its conformation is *endo*; when it is below the sugar ring it is termed *exo*. This method provides only an approximate description of sugar puckering; if a more exact description is required (e.g. for intermediate values), the sugar ring conformation can also be defined using the concept of pseudorotation around the torsion angle ν(C1′–C2′–C3′–C4′). A standard conformation ($P = 0°$) is defined when this angle is maximally positive (equivalent to symmetrical C2′-*exo*–C3′-*endo*). P has values of 0–360° and is usually visualized as a circle with 0/360° at the top. Conformations
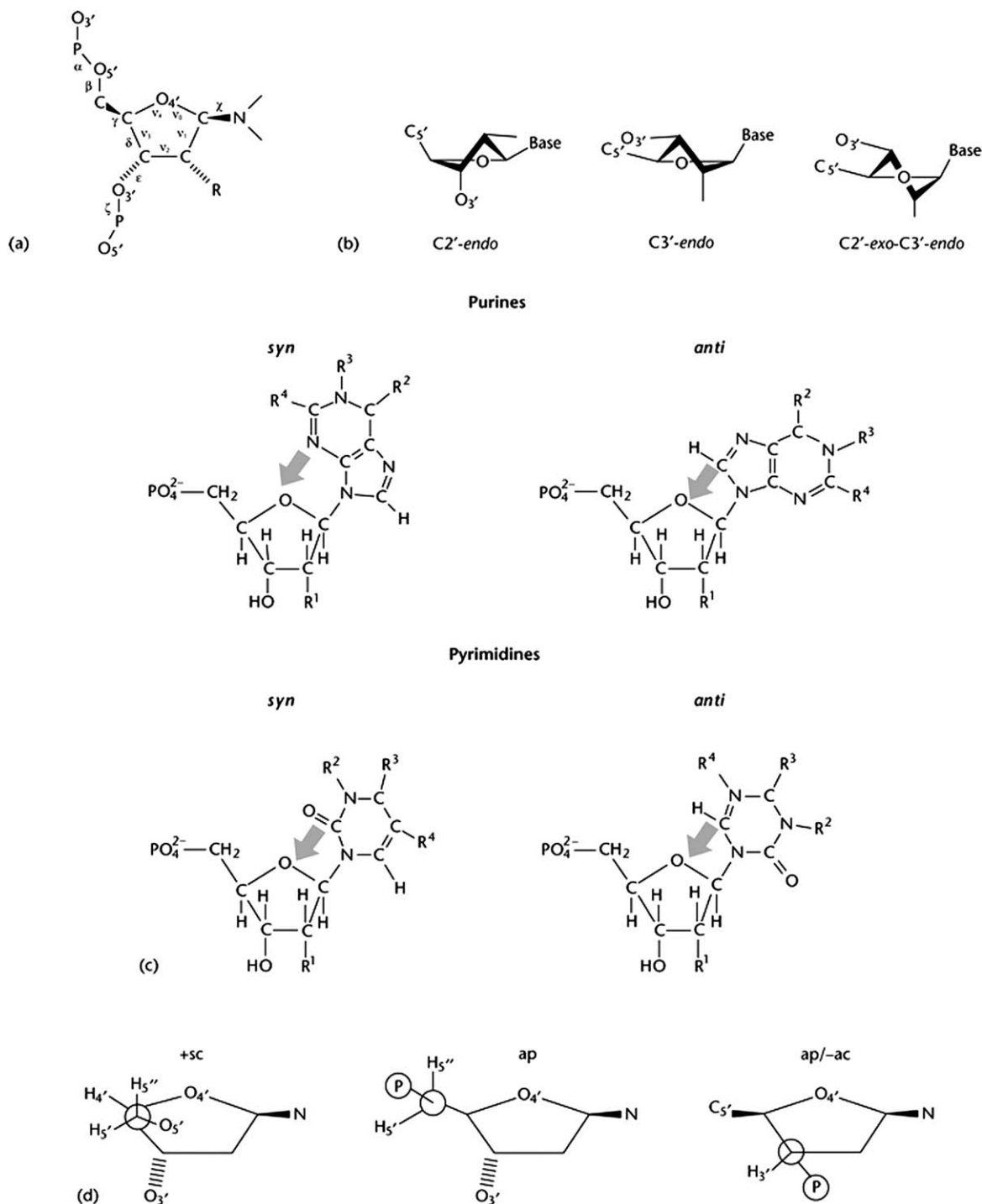
Figure 3 Conformations of nucleotides. (a) Torsion angle notation (IUPAC definitions) for a polynucleotide. A conventional representation showing the torsion angles relative to atom number of the nucleotide. See **Table 4** for a complete description of the atoms defining each angle. (b) Representation of puckering modes of furanose sugar. Three twist (T) forms are shown in which three ring atoms (C4', O4', C1') are planar and the other two (C2', C3') lie on opposite sides of this plane. 'Base' refers to any purine or pyrimidine base. (c) *Anti* and *syn* conformations of the glycosolic bonds for a purine or pyrimidine base. The arrow indicates which base atoms lie above the furanose sugar ring. For purine nucleotides: R^1 = H or OH; R^2 = NH_2 or O; R^3 = H or double bond to C6'; R^4 = H or NH_2. For pyrimidine nucleotides: R^1 = H or OH; R^2 = H or double bond to C4; R^3 = NH_2 or O; R^4 = H or CH_3. (d) Preferred nucleotide conformations at C4'–C5' and C3'–O3'. The structures show synclinal (sc) and antiperiplanar (ap) rotamers of the C4'–C5' bond and the typical antiperiplanar/anticlinal (ap/-ac) conformation of the C3'–O3' bond.

Table 4 Definition of torsion angles in nucleotides. Atoms designated $(n-1)$ and $(n+1)$ belong to adjacent nucleotide units

Torsion angle	Atoms defining angle
α	$(n-1)$O3′–P–O5′–C5′
β	P–O5′–C5′–C4′
γ	O5′–C5′–C4′–C3′
δ	C5′–C4′–C3′–O3′
ε	C4′–C3′–O3′–P
ξ	C3′–O3′–P–O5′$(n+1)$
χ	O4′–C1′–N1–C2 (pyrimidines)
χ	O4′–C1′–N9–C4 (purines)
v_0	C4′–O4′–C1′–C2′
v_1	O4′–C1′–C2′–C3′
v_2	C1′–C2′–C3′–C4′
v_3	C2′–C3′–C4′–O4′
v_4	C3′–C4′–O4′–C1′

in the upper, or northern, half of the circle $(P = 0 \pm 90°)$ are denoted N and those in the southern half $(P = 180 \pm 90°)$ are denoted S. By this method, sugars with conformation can also be defined using the concept of pseudorotation, since sugars with C2′-*endo* and C3′-*endo* are located in the south (S) and north (N) cycles, respectively.

Observed conformations rarely adhere to these strict definitions since transitions between the possible conformations occur easily and are in rapid equilibrium. Also, atoms defining four-atom planes are usually not perfectly coplanar and perfect envelope forms are rare. Thus, most conformations are twist forms and the conformation defined for the furanose ring is determined by which atom has the greatest displacement (**Figure 3b**): if the *endo* displacement of C2′ is greater (major puckering) than the *exo* displacement of C3′ (minor puckering), the conformation is called C2′-*endo*. An abbreviated form of notation describes all aspects of the sugar conformation by using superscripts for *endo* atoms and subscripts for *exo* atoms; these notations precede or follow the letter E or T for major or minor puckering, respectively. For example, the abbreviated forms for the conformations shown in **Figure 3b** are C2′-*endo* (with minor C3′-*exo*) = ${}^{2}T_{3}$; C3′-*endo* (with minor C2′-*exo*) = ${}^{3}T_{2}$; C2′-*exo*–C3′-*endo* (i.e. symmetrical puckering) = . Similarly, a C2′-*exo* envelope is given as ${}_{2}E$.

The plane of the base is almost perpendicular to that of the sugar and approximately bisects the angle subtended by atoms O4′–C1′–C2′. This allows the bases to occupy either of two principal orientations relating to the angle χ of the glycosolic linkage: the *syn* conformer is defined as that with the larger O2 (pyrimidine) or N3 (purine) atom above the sugar ring, while the *anti* conformer has the smaller H6 (pyrimidine) or H8 (purine) in that position (**Figure 3c**).

The exocyclic C4′–C5′ bond can rotate to allow O5′, and the 5′-phosphate, to assume different positions relative to the sugar. Three conformations with all substituents in staggered positions are the favoured rotamers of this bond (γ) (see **Figure 3d**). Phosphate diesters are tetrahedral at phosphorus and typically show antiperiplanar (ap) conformations for the C5′–O5′ bond. Similarly, the C3′–O3′ bond lies in the antiperiplanar to anticlinal (ac) sector (**Figure 3d**).

These principal conformational features of nucleotides are common to the monomeric or polymeric states. However, polymers involve additional parameters relating to their helical symmetry and base pairing (see below).

Pyrimidines and Pyrimidine Nucleotides: Structure, Occurrence and Properties

Structure

The major pyrimidines found in nucleotides (**Table 1**) are cytosine, thymine, uracil and orotate. Relative to the archetypal pyrimidine ring structure, cytosine contains oxo (keto) and amino groups at positions 2 and 4, respectively; orotate has oxo groups at positions 2 and 4 and a carboxyl group at position 6; and uracil contains oxo groups at positions 2 and 4. Thymine is equivalent to a methyl group being added at position 5 of uracil.

The β-glycosolic linkage of all of the pyrimidine nucleotides is between N1 of the base and C1 of the sugar and occupies a narrow range of *anti* conformations. For pyrimidine nucleosides and nucleotides, $+sc$ is the preferred conformation of the phosphate ester bond.

Occurrence

Within cells, the majority of pyrimidine nucleotides exist in their polymeric form since they are present within DNA and RNA molecules. Cytosine is found in both DNA and RNA, but uracil is only present within RNA and thymine is, in general, only found in DNA. (Note, however, that thymine does occur occasionally in specific RNA molecules.) Orotate is synthesized as the precursor for the other pyrimidine nucleotides.

Because of the pairing of specific bases (Thy with Ade, Cyt with Gua), pyrimidine nucleotides constitute 50% of double-stranded DNA. The ratio of cytosines to thymines is constant within a single species, but this ratio varies between species that have different coding genes. This situation is particularly noticeable in species that live in extreme environments, such as high temperatures, where C•G base pairs are utilized at a high frequency to increase the stability of the DNA molecule.

Properties

The primary role of pyrimidine nucleotides in all cells is their involvement in DNA and RNA, the molecules that store and process genetic information. However, pyrimidine nucleotides are also utilized as activated intermediates in some biosynthetic reactions. For example, UDP-glucose and CDP-diacylglycerol are precursors of glycogen and phosphoglycerides, respectively.

Purines and Purine Nucleotides: Structure, Occurrence and Properties

Structure

Table 1 contains nomenclature for four of the purine nucleotides common to cellular metabolism. Relative to the archetypal purine ring structure, the bases contain the following additional groups: adenine has an amino group at position 6; guanine has oxo and amino groups at positions 6 and 2, respectively; hypoxanthine has an oxo group at position 6; and xanthine has oxo groups at positions 2 and 6.

In contrast to the pyrimidines, purines display a wide range of conformations at their β-glycosolic linkage, with most nucleotides being in the *anti* conformation. Notably, purines can also adopt a *syn* β-glycosolic linkage, as is observed in some nucleic acid sequences that take up a specific DNA secondary structure known as Z-DNA. For purine nucleosides, +sc and ap are equally preferred conformations of the phosphate ester bond. However, in the nucleotides, the 5′-phosphate reduces the conformational freedom and the dominant conformer is +sc.

Specific forms of purine nucleotides are synthesized for various aspects of cellular metabolism. For example, 3′,5′-cyclic AMP (cAMP) consists of a displacement on the most internal phosphorus atom of ATP by the 3′-hydroxyl group of the ribose ring. Adenine nucleotides exist in the form of a number of coenzymes: nicotinamide–adenine dinucleotides (NAD^+, NADH, $NADP^+$, NADPH) contain a modified version of the vitamin nicotinate attached to a ribose which is bound to AMP; flavin–adenine dinucleotide (FAD) contains AMP attached to the coenzyme riboflavin 5′-phosphate; coenzyme A (CoA) consists of AMP, with an extra phosphoryl group on the 3′-hydroxyl, linked at the 5′-end to another phosphate group which is esterified to a combination of pantothenate and cysteine.

Occurrence

Nucleotides containing adenine and guanine are utilized in many aspects of cellular biochemistry, while those containing hypoxanthine and xanthine are formed during the biosynthesis of adenylate and guanylate. As with the pyrimidine nucleotides, a large proportion of the major purine nucleotides are found in polymers of DNA and RNA. Unlike the case for pyrimidine nucleotides, the same nucleotides are utilized by both the ribo and deoxyribo forms.

Another difference compared to pyrimidine nucleotides is that significant amounts of purine ribonucleotides exist within cells owing to their utilization in a wide variety of energy-requiring processes (**Table 2**). Also ubiquitous to all life forms are modified versions of purine nucleotides, such as cAMP, variants of NAD^+ and FAD and coenzyme A. The concentrations of these nucleotides can vary significantly; notably, although cyclic nucleotides are relatively stable, cAMP has a short half-life within cells because it is degraded rapidly by enzymes.

Properties

In addition to their distinct roles in DNA and RNA, perhaps the most important property of purine-based nucleotides is their utilization as an energy pool for many enzymatic processes. GTP powers the movements of macromolecules, including the translocation of nascent peptide chains on ribosomes and the activation of signal-coupling proteins. ATP is a universal currency of energy in biological systems: this arises from its high potential for phosphate transfer because of instabilities arising from the electrostatic repulsion of phosphate moieties in the pyrophosphate. The hydrolysed products are more stable relative to ATP, although this is highly dependent on pH and Mg^{2+} concentration.

Further to their major involvement in cellular energetics, several chemical variants of nucleotides containing adenine adapt to a number of dramatically different biochemical roles: cAMP is a 'second messenger' that controls and mediates the activities of peptide hormones; bacterial cells use guanosine polyphosphates (such as ppGpp and pppGpp) to link growth rate to nutrient availability; puromycin is a potent inhibitor of protein biosynthesis; arabino or 8-aza derivatives of adenosine display antibiotic activities; distinct forms of NAD^+ and FAD are also utilized in a variety of oxidation–reduction reactions; acetyl-CoA acts as a carrier of activated acetyl or other acyl groups; *S*-adenosylmethionine carries an activated methyl group that can be transferred to other molecules as required.

Dinucleotides and Oligonucleotides

In their polymeric forms, individual nucleotides are usually linked via 3′,5′-phosphodiester bonds. However, RNA molecules with 2′,5′-phosphodiester bonds have been

synthesized chemically and oligoadenylates with a 2′,5′ linkage are generated in mammalian cells in response to some viral infections.

The covalent structure of nucleic acids is relatively stable, but it is involved in a variety of chemical reactions within the cell. Both DNA and RNA are easily broken down by acid-catalysed hydrolysis, with the glycosolic linkages of DNA being more labile than those of RNA, and purine nucleotides being more easily hydrolysed than pyrimidine nucleotides. The presence of the 2′-OH group in RNA makes these molecules susceptible to hydrolysis in alkaline solutions, producing a mixture of 2′- and 3′-nucleotides. In addition, a number of enzymes, termed nucleases, have been identified that degrade nucleic acids (either RNA or DNA) by specific mechanisms.

DNA and RNA polymers have a unique primary structure that is specified by the sequence of bases in each chain. However, in polynucleotides, individual nucleotides are not independent of each other. Rather, they interact by stacking of the bases and they are organized into a helical form. Generally, these helices have a right-handed sense, but specific sequences are able to adopt left-handed helices. Nucleic acids have a continuum of different structures, but a variety of specific DNA secondary structures have been characterized. The conformation thought to be the most common in cells is the right-handed B-form; other common conformations are designated A (a right-handed helix) and Z (a left-handed helix). Furthermore, sequence-dependent modulation of DNA structure occurs because the two bases in a base pair can undergo a wide variety of potential rotations relative to each other and to neighbouring base pairs. The variety of intricate nucleic acid conformations that exist are stabilized by and are only possible because of reversible interactions with water and its solutes.

RNA has a much higher rate of metabolic turnover than DNA, meaning that RNA is more suited to coordinated expression and regulation of genes. Although RNAs are generally synthesized as single-stranded molecules, significant intra-strand basepairing occurs within molecules and they have defined secondary and tertiary structures. RNA species vary in size from 65 nucleotides upwards, but all double-stranded regions of RNA have a helical structure that is, typically, of a right-handed sense and is closely related to the A conformation of DNA.

Further Reading

Blackburn GM and Gait MJ (1996) *Nucleic Acids in Chemistry and Biology*, 2nd edn. Oxford: Oxford University Press.

Diekmann S (1989) Definitions and nomenclature of nucleic acid structure parameters. *EMBO Journal* **8**: 1–4.

IUPAC-IUB Commission on Biochemical Nomenclature (1970) Abbreviations and symbols for nucleic acids, polynucleotides and their constituents. *European Journal of Biochemistry* **15**: 203 –208.

IUPAC-IUB Joint Commission on Biochemical Nomenclature (1983) Abbreviations and symbols for the description of conformations of polynucleotide chains. *European Journal of Biochemistry* **131**: 9–15.

Kornberg A and Baker TA (1992) *DNA Replication*, 2nd edn. New York: Freeman.

Neidle S (1999) *Oxford Handbook of Nucleic Acid Structure*. Oxford: Oxford University Press.

Neidle S (2002) *Nucleic Acid Structure and Recognition*. Oxford: Oxford University Press.

Saenger W (1984) *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.