

PROVISIONAL

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

and

INTERNATIONAL UNION OF BIOCHEMISTRY JOINT COMMISSION ON BIOCHEMICAL NOMENCLATURE*

ABBREVIATIONS AND SYMBOLS FOR THE DESCRIPTION OF CONFORMATIONS OF POLYNUCLEOTIDE CHAINS

Comments on these recommendations are welcome and should be sent within 8 months from August 1983 to the present Secretary of the Commission

Dr. A. CORNISH-BOWDEN
Department of Biochemistry
University of Birmingham
P.O. Box 363
Birmingham B15 2TT
UK.

Comments from the viewpoint of languages other than English are especially encouraged. These may have special significance regarding the publication in various countries of translations of the nomenclature eventually approved by IUPAC.

*Membership of the Commission for 1981-83 is as follows:

Chairman: P. KARLSON (FRG, 1981-82); H. B. F. DIXON (UK, 1983-); *Secretary:* H. B. F. DIXON (UK, 1981-82); A. CORNISH-BOWDEN (UK, 1983-); *Members:* C. LIÉBECQ (Belgium — as Chairman of IUB Committee of Editors of Biochemical Journals); K. L. LOENING (USA); G. P. MOSS (UK); J. REEDIJK (Netherlands); S. F. VELICK (USA); J. F. G. VLIEGENTHART (Netherlands).

Additional contributors to the formulation of these recommendations:

Expert Panel: D. B. DAVIES (Chairman, UK); C. ALTONA (Netherlands); S. ARNOTT (USA); S. S. DANYLUK (USA); F. E. HRUSKA (Canada); A. KLUG (UK); H.-D. LÜDEMANN (FRG); B. PULLMAN (France); G. N. RAMACHANDRAN (India); A. RICH (USA); W. SAENGER (FRG); R. H. SARMA (USA); M. SUNDARALINGAM (USA); *Associate Members:* P. KARLSON (FRG); O. KENNARD (UK); S.-H. KIM (USA); V. SASISEKHARAN (India); H. R. WILSON (UK). *Nomenclature Committee of IUB* (those additional to JCBN): H. BIELKA (GDR); N. SHARON (Israel).

The document is based on work by a previous subcommission (chairman, F. Cramer) and by an informal committee formed at the Fifth Jerusalem Symposium on Quantum Chemistry and Biochemistry (chairman, M. Sundaralingam).

IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN)

Abbreviations and Symbols for the Description of Conformations of Polynucleotide Chains

Recommendations 1982

CONTENTS

1. General principles of notation
 - 1.1 Chain direction
 - 1.2 Definition of the nucleotide unit
 - 1.3 Atom numbering
 - 1.4 Bonds, bond lengths and interatomic distances
 - 1.5 Bond angles
 - 1.6 Torsion angles
 - 1.7 Conformational regions
 2. The nucleotide unit
 - 2.1 Sugar-phosphate backbone chain
 - 2.2 Sugar ring
 - 2.2.1 Endocyclic torsion angles
 - 2.2.2 Puckered forms
 - 2.2.3 Pseudorotational analysis
 - 2.3 N-Glycosidic bond
 - 2.4 Orientations of side groups
 3. Hydrogen bonds
 - 3.1 Polarity of hydrogen bonds
 - 3.2 Geometry of hydrogen bonds
 4. Helical segments
 - 4.1 Base pairs
 - 4.2 Helix characteristics
- References

INTRODUCTION

Several conventions and notations for polynucleotide conformation have been used by various authors [1–9]. To overcome this confusion, the Joint Commission on Biochemical Nomenclature appointed a panel of experts to review the problem and make recommendations. Their proposals, together with suggestions from the members of

Document of the IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) whose members are P. Karlson (chairman), H. B. F. Dixon, C. Liébecq (as chairman of the IUB Committee of Editors of Biochemical Journals), K. L. Loening, G. P. Moss, J. Reedijk, S. F. Velick and J. F. G. Vliegthart. Comments on and suggestions for future revision of these recommendations may be sent to the secretary of JCBN, H. B. F. Dixon, University Department of Biochemistry, Tennis Court Road, Cambridge, England, CB2 1QW, or to any member. JCBN thanks the expert panel and many others consulted for their work in drafting the document, including other members of the Nomenclature Committee of IUB (H. Bielka and N. Sharon).

The members of the panel were C. Altona, S. Arnott, S. S. Danyluk, D. B. Davies (chairman), F. E. Hruska, A. Klug, H.-D. Lüdemann, B. Pullman, G. N. Ramachandran, A. Rich, W. Saenger, R. H. Sarman, M. Sundaralingam. Associate members were P. Karlson, O. Kennard, S.-H. Kim, V. Sasisekharan and H. R. Wilson. The document is based on work by a previous subcommittee (chairman, F. Cramer) and by an informal committee formed at the Fifth Jerusalem Symposium on Quantum Chemistry and Biochemistry (chairman: M. Sundaralingam).

Reproduced from *Eur. J. Biochem.*, Vol. 131, pp.9–15 (1983)
by courtesy of Springer-Verlag, Heidelberg, FRG.

JCBN and other scientists, are presented here as recommendations that have been approved by IUPAC and IUB.

The nomenclature proposed here is consistent with that recommended for polypeptide conformation [10] as well as with recommendations for polysaccharide conformation [12] and stereochemistry of synthetic polymers [13]. The recommendations on polypeptide conformation [10] also cover general problems of specifying the conformation of biopolymers. Nomenclature of nucleic acids and symbols for their constituents follow published recommendations [14].

RECOMMENDATIONS

1. GENERAL PRINCIPLES OF NOTATION

1.1. Chain Direction

The atoms of the main chain are denoted in Fig. 1. The direction of progress of a polynucleotide chain is from the 5'-end to the 3'-end of the sugar residue.

Notes

a) The definition of chain direction is in accord with the definition of the nucleotide unit (see 1.2).

b) The definition of chain direction with respect to the *sugar carbon atoms* of the nucleotide unit is consistent with the alternative description of polynucleotide sequences as progressing from the 3'-end of one unit to the 5'-end of the next *through the phosphate group*, i.e. in the chain sequence L–(3'→5')–M–(3'→5')–N, etc. (written as LpMpN or L–M–N for a known sequence, or L, M, N for an unknown sequence).

1.2. Definition of a Nucleotide Unit

A nucleotide unit is the repeating unit of a polynucleotide chain; it comprises three distinct parts: the D-ribose or 2-deoxy-D-ribose (2-deoxy-D-*erythro*-pentose) sugar ring, the

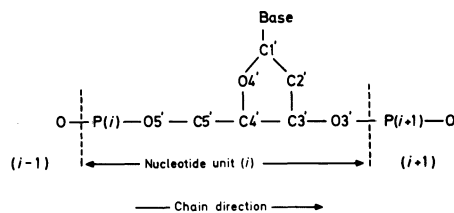


Fig. 1. Designation of chain direction and main chain atoms of *i*th unit in a polynucleotide chain

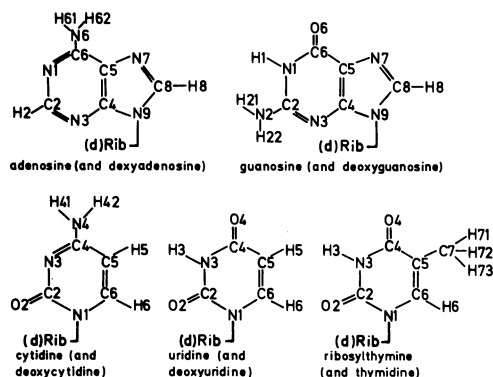


Fig. 2. The atom numbering for the bases of common nucleosides and nucleotides. Hydrogen atoms carry the same numbers as the heavy atoms to which they are attached. The name in parenthesis applies when the 'd' in parenthesis in the formula is present

phosphate group, and the purine or pyrimidine base. The sugar ring and the phosphate group form the backbone of the polynucleotide chain; the base ring linked to the sugar residue constitutes the side chain as shown in Fig. 1.

A nucleotide unit is defined by the sequence of atoms from the phosphorus atom at the 5'-end to the oxygen atom at the 3'-end of the pentose sugar; it includes all atoms of the sugar and base rings.

Specific units ($i, j \dots$ or 3, 4, 5, etc.) are designated by the letter or number in brackets. The units are numbered sequentially in the chain direction, starting at the first nucleotide residue, irrespective of the presence or absence of a phosphate group at the 5'-terminal unit.

The same numbering, A(1), pU(2), pU(3) etc. would apply to the sequence ApUpUp- and pApUpUp-.

1.3. Atom Numbering

The atom numbering of the constituents of the nucleotide unit is shown in Fig. 1–3. The numbering scheme for the bases shown in Fig. 2 is the same as that recommended by IUPAC (Rule B-2.11 on page 58 in [15] and on page 5567 in [16]). The atoms belonging to the sugar moiety are distinguished from those of the base by the superscript prime mark on the atom number. Atoms are specified by the appropriate number after the symbol, e.g. C2, N3 (for base) and C1', C5', O5' (for sugar).

Atoms of a specific unit ($i, j \dots$ or 3, 4, 5) may be designated by the letter or number of the unit in brackets e.g. O3'(i), P(i+1) and N1(3), C2'(4).

Hydrogen atoms carry the same number as the heavy atoms to which they are attached, e.g. base ring H6 (pyrimidine) and H8 (purine) as shown in Fig. 2; sugar ring H1', H2' etc., as shown in Fig. 3. Where there is more than one hydrogen atom (such as at C5' of the sugar ring), the atoms are designated numerically (e.g. H5'1 and H5'2); H5'1 and H5'2 correspond to the *pro-S* and *pro-R* positions [17], respectively (Fig. 3).

The atom-numbering scheme for the 2'-deoxyribonucleotide chain is the same as that for the ribonucleotide chain shown in Fig. 1. The numbering for the sugar ring atoms of both D-ribose and 2-deoxy-D-ribose rings is shown in Fig. 3. (Note the absence of a prime in '2-deoxy-D-ribose' in the

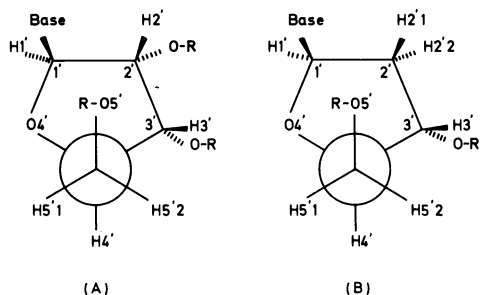


Fig. 3. Designation of sugar ring atoms and hydrogen atoms. (A) In β -D-nucleosides and nucleotides; (B) in their 2'-deoxy derivatives

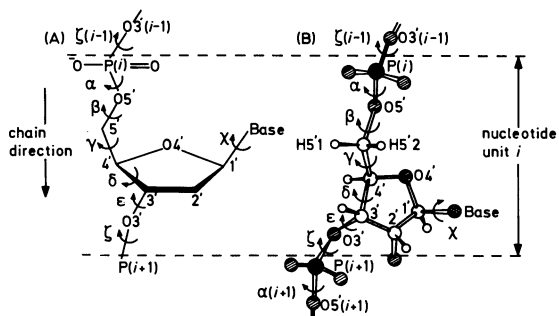


Fig. 4. Section of a polynucleotide backbone showing the atom numbering and the notation for torsion angles. (A) Conventional representation; (B) absolute stereochemistry

previous sentence; C2' of a nucleotide is C2 of its ribose residue.) The two hydrogen atoms attached to the C2' atom of a nucleoside are denoted by H2'1 and H2'2, corresponding to the *pro-S* and *pro-R* positions respectively (Fig. 3).

The hydrogen atoms of hydroxyl groups are specified as in O5'H, O3'H and O2'H, where appropriate, whereas the hydroxyl groups are specified as OH5', etc.

Notes

- Designation of the sugar-ring oxygen atom by O4' conforms with chemical nomenclature; it has been widely but inaccurately denoted by O1' in the past.
- Detailed atom numbering for the modified nucleotides is not considered here.
- The numerical designation of C5' and C2' methylene hydrogen atoms supersedes that introduced by Davies [9].

1.4. Bonds, Bond Lengths and Interatomic Distances

Covalent bonds are denoted by a hyphen between atoms, e.g. O5'–C5', C5'–H5'1 and C2–N3. Atoms in specified nucleotide units are indicated by putting the number of the unit in parentheses, e.g. O5'(i)–C5'(i), O3'(i)–P(i+1). Bond lengths are denoted by $b(\text{O5}', \text{C5}')$ or $b[\text{O3}'(i), \text{P}(i+1)]$. Use of the symbol l for bond length is avoided because it can be confused with the numeral 1 and because l is used for vibration amplitude in electron diffraction (section 1.4 of [10]).

Hydrogen bonds are denoted by a dotted line, with the donor atom being written first, if it can be specified, e.g.

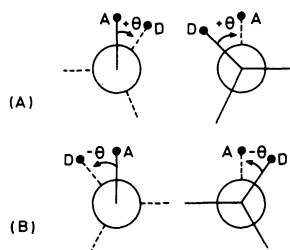


Fig. 5. Newman projections illustrating (A) positive and (B) negative torsion angles. (A) A clockwise turn of the bond containing the front atom about the central bond is needed for it to eclipse the bond to the back regardless of the end from which the system is viewed; hence the value of θ is positive ($+\theta$). (B) A counterclockwise turn of the bond containing the front atom is needed for it to eclipse the bond to the back atom regardless of the end from which the system is viewed; hence the value of θ is negative ($-\theta$).

intramolecular $O5' \cdots N3$ hydrogen bonding in some purine derivatives or $O5'(i) \cdots N3(j)$ for intermolecular hydrogen bonding. The position of the hydrogen may also be indicated as in $O5'-H \cdots N3$. Hydrogen-bonded base pairs are considered separately (see section 4.1).

Distances between non-bonded atoms are denoted by a dot, e.g. $O5'(i) \cdots O3'(j)$.

1.5. Bond Angles

The bond angle included between three atoms $A-B-C$ is written $\tau(A, B, C)$, which may be abbreviated to $\tau(B)$ if there is no ambiguity.

1.6. Torsion Angles [17]

If a system of four atoms $A-B-C-D$ is projected onto a plane normal to $B-C$, the angle between the projection of $A-B$ and the projection of $C-D$ is described as the *torsion angle* about bond $B-C$; this angle may also be described as the angle between the plane containing atoms A, B and C , and the plane containing atoms B, C and D .

The torsion angle is written in full as $\theta(A, B, C, D)$, which may be abbreviated, if there is no ambiguity, to $\theta(B, C)$. In the statement of this rule, the angle θ is used as a general angle rather than as referring to any particular bond (see Rule 2 for designation of main-chain torsion angles).

The zero-degree torsion angle ($\theta = 0^\circ$) is given by the conformation in which the projections of $A-B$ and $C-D$ coincide (this is also known as the eclipsed or *cis* conformation). When the sequence of atoms $A-B-C-D$ is viewed along the central bond $B-C$, a torsion angle is considered positive when the bond to the front must be rotated clockwise in order that it may eclipse the bond to the rear as shown in Fig. 5A. When the bond to the front must be rotated counterclockwise in order to eclipse the bond at the rear, the angle is considered negative as shown in Fig. 5B.

Angles are usually measured from 0° to 360° , but they may be expressed as -180° to $+180^\circ$ when special relationships between conformers need to be emphasized.

Illustrations of the definition of torsion angles are shown for positive and negative values of θ in Fig. 5A and B respectively. It should be noted that a clockwise turn of the bond containing the front atom about the central bond gives a

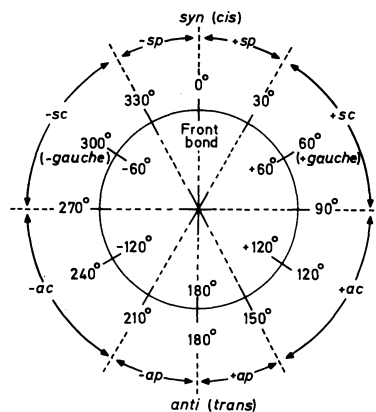


Fig. 6. Relationship between the syn-anti terminology for describing conformational regions [17, 18] and the magnitude of the torsion angle $0-360^\circ$ (or $0 \pm 180^\circ$) with the front bond specifying the zero position. *sp*, synperiplanar; *sc*, synclinal; *ap*, antiperiplanar; *ac*, anticlinal. Other descriptions of particular torsion angles are also given for comparison: 0° , *cis* (*c*); 60° , *+gauche* (*g*⁺); 180° , *trans* (*t*); 300° , *-gauche* (*g*⁻).

positive value of θ from whichever end the system $A-B-C-D$ is viewed in Fig. 5A; similar considerations apply to the conformation with a negative torsion angle (Fig. 5B).

1.7. Conformational Regions

If the precise torsion angle for a conformation is not known, it may be convenient to specify it roughly by naming a conformational region, i.e. a range in which the torsion angle lies. For this the Klyne-Prelog nomenclature [17, 18], accepted in organic chemistry, is recommended. The relationship between the terms used, \pm synperiplanar ($\pm sp$), \pm synclinal ($\pm sc$), \pm anticlinal ($\pm ac$) and \pm antiperiplanar ($\pm ap$), and the magnitudes of the torsion angles are shown in Fig. 6.

The range $0 \pm 90^\circ$ is denoted as *syn* and the range $180 \pm 90^\circ$ is denoted as *anti*.

Note

In order that conformations described by the torsion angle defined in rule 1.6 be consistent in sign and magnitude with the conformational regions ($\pm syn$, $\pm anti$) shown in Fig. 6, it is necessary, when looking down the $B \rightarrow C$ (or $C \rightarrow B$) bond, that the front bond $A-B$ (or $D-C$) should define the zero (0°) position, and that the back bond should define the conformational region.

Examples

a) In Fig. 3 the $O5'-C5'$ bond makes an angle of $+60^\circ$ to $C4'-C3'$ and 300° (-60°) to the $C4'-O4'$ bond. With the $O5'-C5'$ bond defining the zero position these conformations correspond to the $+sc$ and $-sc$ regions, respectively.

b) For the sequence of atoms $A-B-C-D$ as shown in Fig. 5A the same conformation with the torsion angle θ being positive ($+\theta$) is found when looking along either the $B-C$ bond (A to the front) or the $C-B$ bond (D to the front);

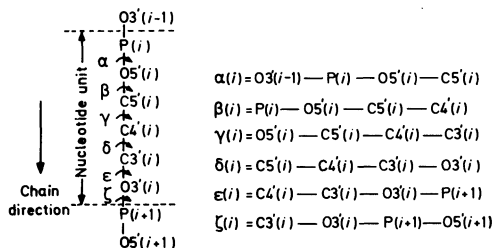


Fig. 7. Torsion angles for backbone conformations of the *i*th nucleotide in polynucleotide chains

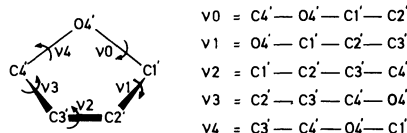


Fig. 8. Torsion angles in sugar rings of β -D-nucleosides and nucleotides

The sequence of atoms used to define each backbone torsion angle is shown in Fig. 8; e.g. ν_0 refers to the torsion angle of the sequence of atoms $C4' - O4' - C1' - C2'$, etc.

this conformation is described as *+sc*. Similarly the conformation designated by a negative value of θ ($-\theta$), shown in Fig. 5B, corresponds to the $-sc$ region.

c) See section 2.3, notes (a) and (c), for the example in Fig. 11.

2. THE NUCLEOTIDE UNIT

The notations used to designate the various torsion angles in the nucleotide unit are indicated in three sections: sugar-phosphate backbone chain, sugar ring and sugar-base side chain.

2.1. Sugar-Phosphate Backbone Chain (main chain)

The backbone of a polynucleotide chain consists of a repeating unit of six single bonds as shown in Fig. 1, viz. $P - O5', O5' - C5', C5' - C4', C4' - C3', C3' - O3'$ and $O3' - P$. The torsion angles about these bonds are denoted, respectively, by the symbols $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$. The symbols $\alpha(i) - \zeta(i)$ are used to denote torsion angles of bonds within the *i*th nucleotide unit as shown in Fig. 4 and 7.

The sequence of main-chain atoms used to define each backbone torsion angle is shown in Fig. 7.

Note

The recommended $\alpha - \zeta$ notation differs from the ϕ, ψ, ω notation [4, 7] adopted by many workers and from an earlier $\alpha - \zeta$ notation [8].

A substantial majority of the subcommittee favoured the $\alpha - \zeta$ notation because it is convenient to remember for a backbone repeat of six bonds. The recommended $\alpha - \zeta$ notation is the second of the systems proposed by Seeman et al. [8] and was chosen because it starts at the phosphorus atom which is the first atom of the nucleotide unit, has the highest atomic number, and is the only atom of its kind in the backbone.

2.2. Sugar Ring

2.2.1. Endocyclic Torsion Angles

The sugar ring occupies a pivotal position in the nucleotide unit because it is part of both the backbone and the side chain. In order to provide a complete description of the ring conformation, it is necessary to specify the endocyclic torsion angles for the ring as well as the bond lengths and bond angles. The five endocyclic torsion angles for the bonds $O4' - C1', C1' - C2', C2' - C3', C3' - C4'$ and $C4' - O4'$ are denoted by the symbols, $\nu_0, \nu_1, \nu_2, \nu_3$ and ν_4 , respectively.

Notes

a) The backbone torsion angle δ and the endocyclic torsion angle ν_3 both refer to rotation about the same bond, $C4' - C3'$. Both angles are needed for complete description of the main-chain and sugar-ring conformations in some studies.

b) The notation $\tau_0 - \tau_4$ previously used [4, 19-21] to represent torsion angles about the bonds in the sugar ring is superseded by the present notation (ν), which is consistent with polysaccharide nomenclature [12]. The symbol τ is now used to denote a bond angle, which is consistent with polypeptide nomenclature [10].

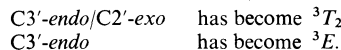
2.2.2. Puckered Forms

Since the sugar ring is generally non-planar, its conformation may need designation. If four of its atoms lie in a plane, this plane is chosen as a reference plane, and the conformation is described as envelope (*E*); if they do not, the reference plane is that of the three atoms that are closest to the five-atom, least-squares plane, and the conformation is described as twist (*T*) [22, 23]. Atoms that lie on the side of the reference plane from which the numbering of the ring appears clockwise are written as superscripts and precede the letter (*E* or *T*); those on the other side are written as subscripts and follow the letter (Fig. 9). These definitions [23] mean that atoms on the same side of the plane as C5 in D-ribofuranose derivatives are written as preceding superscripts.

Notes

a) The present *E* and *T* notations for puckered forms of the sugar ring conform to those recommended for the conformational nomenclature of five and six-membered rings of monosaccharidés and their derivatives [23].

b) The *E/T* notation has superseded the *endo/exo* description [24], in which atoms now designated by superscripts were called *endo*, and those now designated by subscripts were called *exo*. Fig. 10 shows both systems of designation. Examples:



c) Symmetrical twist conformations, in which both atoms exhibit equal displacements with respect to the five-atom plane, are denoted by placing the superscript and subscript on the same side of the letter *T*, e.g. ${}^3T, {}^4T$, etc.

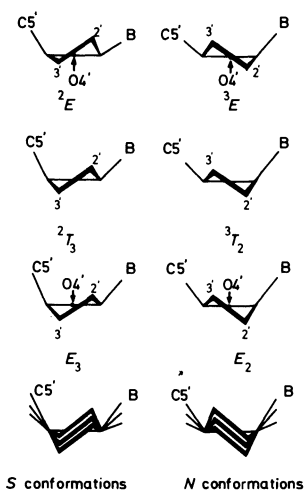


Fig. 9. Diagrammatic representation of sugar-ring conformations of β -D-nucleosides and their relation to the pseudorotational N-type and S-type conformers (Section 2.2.3.). The purine or pyrimidine base is represented by B. (Figure adapted from that of Saenger [25])

2.2.3. Pseudorotational Analysis

The sugar ring conformation has also been described by Altona and Sundaralingam [26] using the concept of pseudorotation, which has been found advantageous in describing the conformational dynamics of the sugar ring [27].

Each conformation of the furanose ring can be unequivocally described by two pseudorotational parameters: the phase angle of pseudorotation, P , and the degree of pucker, ψ_m . A standard conformation ($P = 0^\circ$) is defined with a maximally positive $C1'-C2'-C3'-C4'$ torsion angle [i.e. the symmetrical 3T form], and P has value $0-360^\circ$. Conformations in the upper or northern half of the circle ($P = 0 \pm 90^\circ$) are denoted *N* and those in the southern half of the circle ($P = 180 \pm 90^\circ$) are denoted *S* conformation. The relationship between P and the *endo/exo* and *T/E* notations is illustrated in Fig. 10. It may be seen that the symmetrical twist (*T*) conformations arise at even multiples of 18° of P and the symmetrical envelope (*E*) conformations arise at odd multiples of 18° of P .

Note. The present designation of the degree of pucker (ψ_m) differs from the original notation (τ_m) of Altona and Sundaralingam [26] in order to avoid confusion with the notation for bond angles.

2.3. N-Glycosidic Bond

The torsion angle about the N-glycosidic bond ($N-C1'$) that links the base to the sugar is denoted by the symbol χ which is the same as the notation used to denote side-chain torsion angles in polypeptides [10]. $\chi(i)$ denotes the torsion angle in the i th nucleotide unit.

The sequence of atoms chosen to define this angle is $O4'-C1'-N9-C4$ for purine and $O4'-C1'-N1-C2$ for pyrimidine derivatives. Thus when $\chi = 0^\circ$ the $O4'-C1'$ bond is eclipsed with the $N9-C4$ bond for purine and the $N1-C2$ bond for pyrimidine derivatives. The definitions of torsion angles (section 1.6) of the N-glycosidic bond are illustrated, looking along the bond, in Fig. 11.

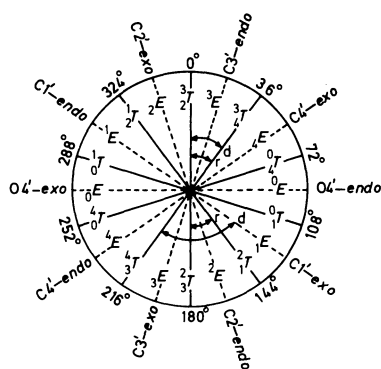


Fig. 10. The pseudorotational pathway of the β -D-aldofuranose ring, showing the relation between phase angle of pseudorotation P ($0-360^\circ$), the envelope (*E*) and twist (*T*) notations and the *endo* and *exo* notations. *N*-Type conformations correspond to the northern half ($P = 0 \pm 90^\circ$) and *S*-type correspond to the southern half of the pseudorotational cycle. The symbols 'r' and 'd' represent the usual range of P values for *N* and *S* conformations of ribo- (r) and 2'-deoxyribo- (d) furanose rings of β -D-nucleosides and nucleotides. (Diagram adapted from the work of Altona and Sundaralingam [26])

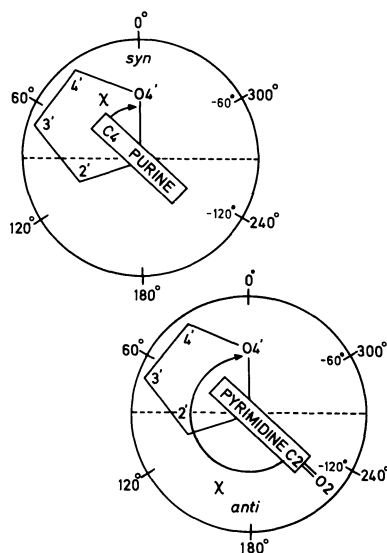


Fig. 11. Diagrammatic representation of the N-glycosidic bond torsion angle χ and the *syn* and *anti* regions for purine and pyrimidine derivatives. The purine derivative (upper left) is viewed down the $N9-C1'$ bond and is shown in the $+sc$ conformation. The pyrimidine derivative (lower right) is viewed down the $N1-C1'$ bond and is shown in the $-ac$ conformation. The sugar ring is shown as a regular pentagon

Notes

a) The choice of bond sequence to define χ is based on accepted chemical nomenclature [13] and, at the same time, the use of the terms *syn* and *anti* to describe different conformational regions of χ for purine and pyrimidine derivatives

is now consistent with accepted chemical nomenclature (Rule 1.7), viz.

$$\begin{aligned} \text{syn, } \chi &= 0 \pm 90^\circ \\ \text{anti, } \chi &= 180 \pm 90^\circ. \end{aligned}$$

Examples of *syn* and *anti* conformations are shown in Fig. 11. In the new convention most conformations formerly described as *syn* and *anti* remain *syn* and *anti* respectively, except the *high-anti* region which may be described as *-synclinal (-sc)*, Rule 1.7.

b) Many of the conventions for defining the torsion angle of a bond in nucleic acids [1-7, 28, 29] have been based on the sugar ring O4' atom [4, 5, 25] or C2' atom [3, 29] in conjunction with the base ring C8(Pur)/C6(Pyr) atoms [4, 5, 29] or base ring C4(Pur)/C2(Pyr) atoms [3, 28]. Approximate relationships between the definitions of torsion angles have been summarised by Sundaralingam [30], and these aid the comparison of conformations described in the older literature.

A substantial portion of the literature has used the nomenclature based on sugar ring O4'—C1' and base ring N9—C8 (purine) and N1—C6 (pyrimidine). χ_{old} is related to the present definition χ_{new} by the relation,

$$\chi_{new} \approx \chi_{old} \pm 180^\circ.$$

c) Following the definitions of conformational regions in Rule 1.7 and Fig. 6 the *anti* conformation of the pyrimidine example in Fig. 11 corresponds to the *-anticalinal (-ac)* conformation and the *syn* conformation of the purine derivative corresponds to the *+synclinal (+sc)* conformation.

2.4. Orientation of Side Groups

For the precise definition of the orientation of any pendant groups specification of the torsion angle about the exocyclic bond is necessary. The exocyclic torsion angle may be denoted by the symbol η with a locant to indicate the atom to which it refers. Examples:

Ribose Rings. The symbol $\eta_{2'}$ may be used to denote the torsion angle about the C2'—O2' bond for the sequence of atoms C1'—C2'—O2'—X where X = H, CH₃, PO₃²⁻, etc. If no confusion is possible, the symbol η (without any additional index) may be used for the C2'—O2' bond.

Base. Torsion angles for bonds in base rings e.g. C6—N6 in adenine, C2—N2 in guanine and C4—N4 in cytosine may be specified by η_6 , η_2 and η_4 , respectively.

When the groups are substituted by hydrogen atoms only (see Fig. 2), the relevant dihedral angles defined by the sequence rules are:

$$\begin{aligned} \eta_{61} &= \text{N1—C6—N6—H61} \\ \eta_{62} &= \text{N1—C6—N6—H62} \\ \eta_{21} &= \text{N1—C2—N2—H21} \\ \eta_{22} &= \text{N1—C2—N2—H22} \\ \eta_{41} &= \text{N3—C4—N4—H41} \\ \eta_{42} &= \text{N3—C4—N4—H42}. \end{aligned}$$

The rules may be adapted for substituted base and sugar rings, such as those of minor components of tRNA. Examples:
1-Methyladenosine: use η_{11} , η_{12} , η_{13} for the C—CH₃ group conformation.

2'-O-Methyladenosine: use $\eta_{2'}$ for rotation about the C2'—O2' bond and $\eta_{2'1}$, $\eta_{2'2}$, $\eta_{2'3}$ for the OCH₃ group.

Note. Recommendations governing the description of conformations of side chains and derivatives follow the

sequence rules (Rule 2) and the side-chain rules (Rule 4) of the recommendations for polypeptides [10].

3. HYDROGEN BONDS

3.1. Polarity of Hydrogen Bonds

In specifying a hydrogen bond the atom covalently linked to the hydrogen atom is mentioned first, as in X—H...Y. The polarity of a hydrogen bond is from the hydrogen-atom donor to the acceptor.

3.2. Geometry of Hydrogen Bonds

The hydrogen bond may be described by extension of the nomenclature of sections 3.1, 1.4, 1.5 and 1.6, so that for the hydrogen bond in the system C(i)—X(i)—H(i)...Y(k)—C(k) the following symbols may be used:

$$\begin{aligned} b[\text{H}(i) \cdots \text{Y}(k)] & \quad \text{or} \quad b[\text{H}(i), \text{Y}(k)] \\ \tau[\text{X}(i) - \text{H}(i) \cdots \text{Y}(k)] & \quad \text{or} \quad \tau[\text{X}(i), \text{H}(i), \text{Y}(k)] \\ \tau[\text{H}(i) \cdots \text{Y}(k) - \text{C}(k)] & \quad \text{or} \quad \tau[\text{H}(i), \text{Y}(k), \text{C}(k)]. \end{aligned}$$

Where the positions of hydrogen atoms are not available the following may be used:

$$b[\text{X}(i), \text{Y}(k)] \quad \text{and} \quad \tau[\text{C}(i), \text{X}(i), \text{Y}(k)].$$

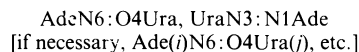
4. HELICAL SEGMENTS

A regular helix is strictly of infinite length, with the torsion angles in a nucleotide unit the same for all units. Two or more polynucleotide chains may associate in a helical complex by hydrogen bonding between base pairs. Torsion angles for each residue may differ for different chains in the same double, triple, etc. complex.

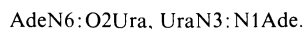
A helical *segment* of a polynucleotide chain may be described in terms of torsion angles of the nucleotide units or in terms of the helix characteristics summarized in section 4.2.

4.1. Base Pairs

Base pairs with different geometries have been observed. These geometries should be denoted by the appropriate hydrogen-bonding scheme specifying both the heterocyclic base (e.g. Ade, Ura, Gua, Cyt) and the heteroatoms involved in the hydrogen bonding. In some cases it is also desirable to specify the nucleotide unit (*i, j*, etc.). Typical examples are: for a Watson-Crick A:U base pair:



for a reversed Watson-Crick A:U base pair:



4.2. Helix Characteristics

In the description of helices or helical segments the following symbols should be used:

$$\begin{aligned} n &= \text{number of residues per turn} \\ h &= \text{unit height (translation per residue along the helix axis)} \\ t &= 360^\circ/n = \text{unit twist (angle of rotation per residue about the helix axis)} \\ p &= \text{pitch height of helix} = n.h. \end{aligned}$$

A polynucleotide may be accurately described in terms of the polar atomic co-ordinates r_i , ϕ_i , z_i where for each atom i , r_i is the radial distance from the helix axis and ϕ_i and z_i are the angular and height differences respectively, relative to a reference point. The reference point should be either a symmetry element, as in RNA and DNA, or the C1' atom of a nucleotide if no symmetry element between polynucleotide chains is present.

REFERENCES

1. Donohue, J. & Trueblood, K. N. (1960) *J. Mol. Biol.* **2**, 363–371.
2. Sasisekharan, V., Lakshminarayanan, A. V. & Ramachandran, G. N. (1967) in *Conformation of Biopolymers II* (Ramachandran, G. N., ed.) pp. 641–654, Academic Press, New York.
3. Arnott, S. & Hukins, D. W. L. (1969) *Nature (Lond.)*, **224**, 886–888.
4. Sundaralingam, M. (1969) *Biopolymers*, **7**, 821–860.
5. Lakshminarayanan, A. V. & Sasisekharan, V. (1970) *Biochim. Biophys. Acta*, **204**, 49–59.
6. Olson, W. K. & Flory, P. J. (1972) *Biopolymers*, **11**, 1–23.
7. Sundaralingam, M., Pullman, B., Saenger, W., Sasisekharan, V. & Wilson, H. R. (1973) in *Conformations of Biological Molecules and Polymers* (Bergman, E. D. & Pullman, B., eds) pp. 815–820, Academic Press, New York.
8. Seeman, N. C., Rosenburg, J. M., Suddath, F. L., Kim, J. J. P. & Rich, A. (1976) *J. Mol. Biol.* **104**, 142–143.
9. Davies, D. B. (1978) in *NMR in Molecular Biology* (Pullman, B., ed.) pp. 509–516, Reidel, Dordrecht.
10. IUPAC-IUB Commission on Biochemical Nomenclature (CBN) Abbreviations and symbols for the description of the conformation of polypeptide chains, Tentative rules 1969 (approved 1974), *Arch. Biochem. Biophys.* **145**, 405–421 (1971); *Biochem. J.* **121**, 577–585 (1971); *Biochemistry*, **9**, 3471–3479 (1970); *Biochim. Biophys. Acta*, **229**, 1–17 (1971); *Eur. J. Biochem.* **17**, 193–201 (1970); *J. Biol. Chem.* **245**, 6489–6497 (1970); *Mol. Biol.* (in Russian) **7**, 289–303 (1973); *Pure Appl. Chem.* **40**, 291–308 (1974); also on pp. 94–102 in [11].
11. International Union of Biochemistry (1978) *Biochemical Nomenclature and Related Documents*, The Biochemical Society, London.
12. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) Symbols for specifying the conformation of polysaccharide chains, Recommendations 1981, *Eur. J. Biochem.* **131**, 5–7 (1983).
13. IUPAC Commission on Macromolecular Nomenclature (CMN), Stereochemical definitions and notations relating to polymers, *Pure Appl. Chem.* **53**, 733–752 (1981).
14. IUPAC-IUB Commission on Biochemical Nomenclature (CBN) Abbreviations and symbols for nucleic acids, polynucleotides and their constituents, Recommendations 1970, *Arch. Biochem. Biophys.* **145**, 425–436 (1971); *Biochem. J.* **120**, 449–454 (1970); *Biochemistry*, **9**, 4022–4027 (1970); *Biochim. Biophys. Acta*, **247**, 1–12 (1971); *Eur. J. Biochem.* **15**, 203–208 (1970) corrected 25, 1 (1972); *Hoppe-Seyler's Z. Physiol. Chem.* (in German) **351**, 1055–1063 (1970); *J. Biol. Chem.* **245**, 5171–5176 (1970); *Mol. Biol.* (in Russian) **6**, 167–174 (1972); *Pure Appl. Chem.* **40**, 277–290 (1974); also on pp. 116–121 in [11].
15. International Union of Pure and Applied Chemistry (1979) *Nomenclature of Organic Chemistry, Sections A, B, C, D, E, F and H*, (Rigandy, J. & Klesney, S. P., eds) Pergamon Press, Oxford.
16. IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC) Definitive rules for nomenclature of organic chemistry, *J. Am. Chem. Soc.* **82**, 5545–5574 (1960).
17. IUPAC Commission on Nomenclature of Organic Chemistry (CNOC) Rules for the nomenclature of organic chemistry, Section E: Stereochemistry, Recommendations 1974, *Pure Appl. Chem.* **45**, 11–30 (1976); also on pp. 473–490 in [15] and on pp. 1–18 in [11].
18. Klyne, W. & Prelog, V. (1960) *Experientia*, **16**, 521–523.
19. Arnott, S. & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun.* **47**, 1504–1509.
20. Pullman, B., Perahia, D. & Saran, A. (1972) *Biochim. Biophys. Acta*, **269**, 1–14.
21. Saran, A. & Govil, G. (1971) *J. Theor. Biol.* **33**, 407–418.
22. Hall, L. D. (1963) *Chem. Ind. (Lond.)* 950–951.
23. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) Conformational nomenclature for five and six-membered ring forms of monosaccharides and their derivatives, Recommendations 1980, *Arch. Biochem. Biophys.* **207**, 469–472 (1981); *Eur. J. Biochem.* **111**, 295–298 (1980); *Pure Appl. Chem.* **53**, 1901–1905 (1981).
24. Jardetzky, C. D. (1960) *J. Am. Chem. Soc.* **82**, 229–233.
25. Saenger, W. (1973) *Angew. Chem. Int. Ed. Engl.* **12**, 591–601.
26. Altona, C. & Sundaralingam, M. (1972) *J. Am. Chem. Soc.* **94**, 8205–8212.
27. Altona, C. & Sundaralingam, M. (1973) *J. Am. Chem. Soc.* **95**, 2333–2334.
28. Kang, S. (1971) *J. Mol. Biol.* **58**, 297–315.
29. Saenger, W. & Scheit, K. H. (1970) *J. Mol. Biol.* **50**, 153–169.
30. Sundaralingam, M. (1973) in *Conformations of Biological Molecules and Polymers* (Bergman, E. D. & Pullman, B., eds) pp. 417–455, Academic Press, New York.