1.2 Description of conformation

a) Bond distances, bond angles, torsion angles

The three-dimensional structure of any molecule can be described by:

bond distances A–B between two covalently bound atoms A,B
bond angles A–B–C between three covalently bound atoms A,B,C
torsion angles, which give the relative orientation of four covalently bound atoms A–B–C–D.

For the torsion angles, the IUPAC-IUB commission has recommended the following definition [83I1]:
The torsion angle θ in Fig. 3 is described as the angle between projected bonds A–B and C–D when looking along the central bond either in direction B→C or in the opposite sense C→B. It is defined as 0° if A–B and C–D are eclipsed (*cis* and *coplanar*), and the sign of θ is positive if the front bond A–B (if looking in direction B→C) has to be rotated clockwise to eclipse it with the rear bond C–D. If it has to be rotated anti-clockwise, θ is negative. The torsion angle θ is reported either in the range 0° to 360° or −180° to +180°.

Rather than describing the torsion angle θ in terms of an angle between projected bonds, it can also be formulated as an angle between the two planes containing atoms, A, B, C, and B, C, D. Another definition uses the angle between the normal to these planes. This *dihedral angle* (Fig. 3) is in fact the complement of the torsion angle θ. In the literature, the nomenclature *torsion* and *dihedral* angles are often confused, as are the definitions, and the term "dihedral angle" is used to describe, in fact, a torsion angle. Therefore, one has to be careful if reading the literature.

Because rotations about bonds are usually restricted by steric requirements, it is often sufficient to describe a molecular conformation by a *torsion angle range* rather than by the proper torsion angle. The ranges commonly used in organic chemistry are those proposed by Klyne and Prelog, *syn* ($\approx 0°$), *anti* ($\approx 180°$), $\pm$*synclinal* ($\approx \pm 60°$), and $\pm$*anticlinal* ($\approx \pm 120°$) [60K1]. In spectroscopic and crystallographic publications, the notation *cis* ($\approx 0°$), *trans* ($\approx 180°$), $\pm$*gauche* ($\approx \pm 60°$) is most frequently employed (Fig. 4).
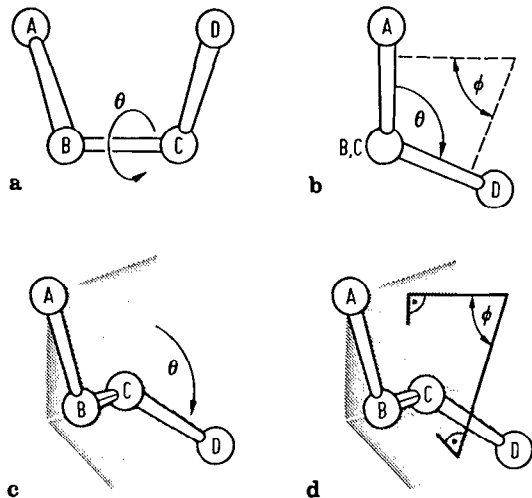

Fig. 3a⋯d. Definition of torsion and dihedral angles. (a) Torsion angle θ(A–B–C–D) describing orientations of bonds A–B and C–D with respect to the central bond B–C. (b) View along B→C. θ is the torsion angle between the projected bonds A–B and C–D; the complement ϕ is called the dihedral angle. If A–B and C–D are *cis*-planar (coinciding in projection), angles $\theta = 0°$ and $\phi = 180°$ ($\theta = 180°-\phi$); they are counted positive if the near bond A–B has to be rotated clockwise to bring it into *cis*-planar position with the far bond C–D. (c) θ is defined as the angle between planes A–B–C and B–C–D. (d) The dihedral angle ϕ represents the angle between normals to these planes [84S1]. Note: in the literature, the terms "torsion" and "dihedral" are often confused and used synonymously. Most frequently "dihedral" means angles defined as θ in Fig. 3, which are, in fact, "torsion" angles.
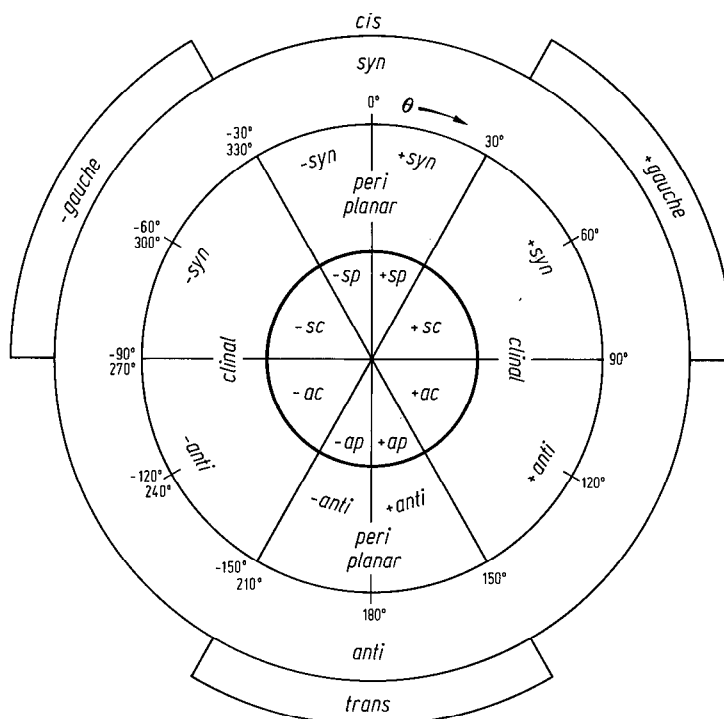
Fig. 4. Correlation of torsion angle ranges (*cis, trans, +gauche, −gauche*) with ranges defined by Klyne and Prelog [60K1] (*syn* or *synperiplanar, anti* or *antiperiplanar, +synclinal, −synclinal, +anticlinal, −anticlinal*). The terms *syn, anti* have a special meaning in nucleotide stereochemistry (Fig. 8) [84S1].

b) Definition of the nucleotide unit

A nucleotide unit is the repeating unit of a polynucleotide chain, and defined by the sequence of atoms from the phosphorous atom at the 5′-end to the oxygen atom at the 3′-end of the pentose sugar (Fig. 5).

c) Definition of backbone torsion angles in the nucleotide unit

The repeating unit of the backbone of a polynucleotide chain consists of six bonds as shown in Fig. 5: P–O5′, O5′–C5′, C5′–C4′, C4′–C3′, C3′–O3′, O3′–P. The torsion angles about these bonds are denoted by greek symbols in sequential and in alphabetical order: α, β, γ, δ, ε, ξ. In another notation, the nomenclature used in polypeptides was adapted, as: ω, ϕ, ψ, ψ', ϕ', ω', but was not recommended by the IUPAC-IUB commission [83I1]. It is, however, still in use in some laboratories.

d) Endocyclic and exocyclic sugar torsion angles

In nucleic acids, the furanose sugar ring is part of both the backbone and the sugar-base side-chain. The conformation of the sugar is described by the endocyclic torsion angles for the bonds O4′–C1′, C1′–C2′, C2′–C3′, C3′–C4′, C4′–O4′, which are denoted by the symbols ν_0, ν_1, ν_2, ν_3, ν_4, respectively (Fig. 5).

The backbone torsion angle δ and the endocyclic sugar torsion angle ν_3 refer to the same bond C3′–C4′, but one is exocyclic, the other endocyclic. They are both needed to properly describe the nucleotide conformation.

In oligonucleotide crystal structure analyses, it is common usage to describe the sugar pucker only with the exocyclic torsion angle δ, defined by C5′–C4′–C3′–O3′. Although this might be sufficient in an *ideal* system with undistorted bond geometry, its use should be discouraged in favor of the better and more reliably defined pseudorotation parameters given in subsections (e) and (f). They can be easily derived once the atomic coordinates of an oligonucleotide are known.
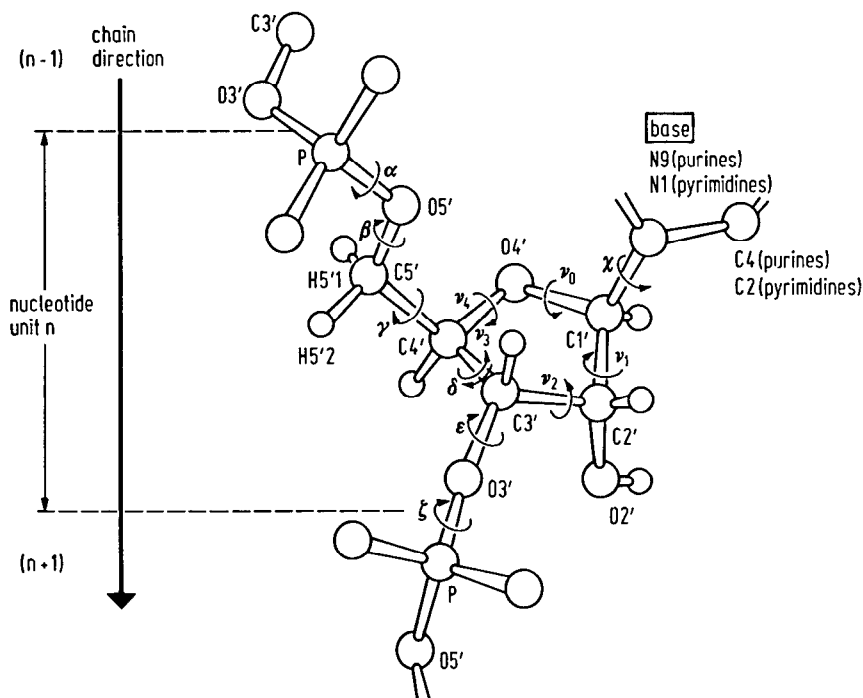
Fig. 5. Atomic numbering scheme and definition of torsion angles for a polyribonucleotide chain (arrows indicate positive rotation of angles A–B–C–D when looking along the central bond B→C given in the table). Counting of nucleotides is from top to bottom, i.e., in the direction O5'→O3'. Hydrogens at C5' are differentiated by 1 and 2. In deoxyribose, the hydrogen replacing O2' is labelled 2, the other one, 1. The full description of torsion angles is given in the following table [84S1].

Torsion angle	Atoms involved[1])
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
ζ	C3'–O3'–P–O5'$_{(n+1)}$
χ	O4'–C1'–N1–C2 (pyrimidines)
	O4'–C1'–N9–C4 (purines)
ν_0	C4'–O4'–C1'–C2'
ν_1	O4'–C1'–C2'–C3'
ν_2	C1'–C2'–C3'–C4'
ν_3	C2'–C3'–C4'–O4'
ν_4	C3'–C4'–O4'–C1'

[1]) Atoms designated (n−1) and (n+1) belong to adjacent units.

e) Description of sugar pucker

The sugar ring is generally puckered such that in the *envelope* (E) conformation one ring atom deviates from the plane defined by the other four atoms, or in the *twist* (T) conformation two atoms deviate from the plane defined by the other three atoms (Fig. 6a, b). The pucker is described relative to the exocyclic atom C5′, and called *endo* if the puckered atom is on the same side of the plane as C5′, otherwise *exo*.

Thus, if atom C3′ is on the same side as C5′ in an envolope form, the pucker is described as

C3′-*endo* or ^{3}E (with a 3 preceding E as a superscript) and if it is on the opposite side, we have

C3′-*exo* or $_3$E (with a 3 preceding E as a subscript).

In the twist form, the twist can be symmetrical, e.g.:

C3′-*endo*, C2′-*exo* or $_2^3$T,

but it can also be unsymmetrical, with more pronounced C2′-*exo* than C3′-*endo* pucker. Then the major pucker atom precedes the letter T (as sub- or superscript) and the minor pucker atom follows, e.g.:

C2′-*exo*, C3′-*endo*, or $_2$T^3 (major pucker is C2′-*exo*, minor pucker is C3′-*endo*).
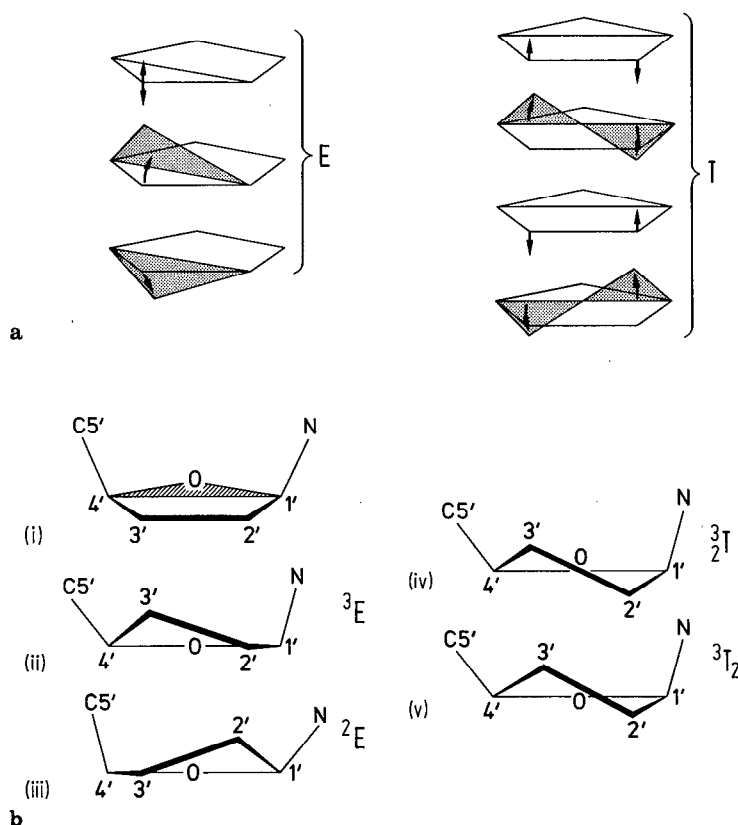

Fig. 6a, b. (a) Puckering of five-membered ring into envelope (E) and twist (T) forms. In E, four of the five atoms are coplanar and one deviates from this plane; in T, three atoms are coplanar and the other two lie on opposite sides of this plane. (b) Definition of sugar puckering modes. (i) Starting position with flat five-membered sugar, a situation never observed. Plane C1′–O4′–C4′ is shown hatched. (ii⋯v) View with this plane perpendicular to the paper. (ii) Envelope C3′-*endo*, ^{3}E. (iii) Envelope C2′-*endo*, ^{2}E. (iv) Symmetrical twist or half-chair C2′-*exo*, C3′-*endo*, $_2^3$T. (v) Unsymmetrical twist with major C3′-*endo* and minor C2′-*exo* pucker, ^{3}T$_2$ [84S1].

f) Pseudorotational analysis (Fig. 7a, b)

Each sugar ring conformation can be neatly described by two pseudorotational parameters which can be derived by mathematical formulae from the five endocyclic torsion angles [72A2, 73A1, 81R1, 85M1]: the phase angle of pseudorotation, P, and the degree of pucker, v_{max} (called ψ_m in [72A2, 73A1, 83I1] and t_m in [81R1, 85M1]), which gives a measure of the maximum out-of-plane pucker of the furanose ring atoms. The pseudorotation phase angle is defined as $0°$ if torsion angle C1′–C2′–C3′–C4′ is maximally positive corresponding to the symmetrical form C3′-*endo*, C2′-*exo*, or ${}^3_2T$, and P adopts values $0°$ to $360°$.

Conformations in the upper ("northern") half of the circle (Fig. 7a) ($P = 0° \pm 90°$) are denoted N, those in the lower ("southern") half of the circle ($P = 180° \pm 90°$) are denoted S. As illustrated in Fig. 7, envelope and symmetrical twist pucker modes alternate every $18°$, with E at even and T at odd multiples of $18°$ of the pseudorotation phase cycle.

For the calculation of P and v_{max}, two different mathematical equations have been derived. Since different endocyclic torsion angles θ_j and θ'_i are used which are also different from those described in Fig. 5, one has to be careful in their application. These equations are:

I. The original formulation [72A2, 73A1]:[1]

$$\tan P = \frac{(\theta_2 + \theta_4) - (\theta_1 + \theta_3)}{2\theta_0 (\sin 36° + \sin 72°)}$$

$$\theta_j = \theta_m \cos(P + j\delta) \quad (\delta = 144°)$$

$$\theta_m = \theta_0 / \cos P$$

II. In a new formulation, a Fourier-type equation is used [81R1, 85M1]:[2]

$$A = 0.4 \sum_{i=1}^{5} \theta'_i \cos[144°(i-1)]$$

$$B = -0.4 \sum_{i=1}^{5} \theta'_i \sin[144°(i-1)]$$

$$\tan P = B/A \quad \text{(if } A < 0, \text{ then } 180° \text{ is added to } P\text{)}$$

$$t_m = \sqrt{A^2 + B^2}$$

In these equations, θ_j and θ'_i are different and they deviate from the IUPAC-IUB recommendation for the v_j (Fig. 5). The necessary translation table is:

Equation I	θ_3	θ_4	θ_0	θ_1	θ_2	θ_m
Equation II	θ'_4	θ'_5	θ'_1	θ'_2	θ'_3	t_m
IUPAC-IUB	v_0	v_1	v_2	v_3	v_4	ψ_m
This volume	v_0	v_1	v_2	v_3	v_4	v_{max}

[1] There is an error in the monograph "Principles of Nucleic Acid Structure" [84S1] concerning the formulae for v_j and v_{max}. This will be corrected in the second print (1988).
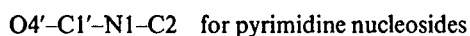[2] See also section 2.1.1.4.

Fig. 7a, b. (a) Pseudorotation cycle of the furanose ring in nucleosides. Values of phase angles P given in multiples of $36°$. Envelope E and twist T forms alternate every $18°$. After rotation by $180°$ the mirror image of the starting position is found. On the periphery of the cycle, riboses with signs of endocyclic torsion angles are indicated. ($+$)Positive, ($-$)negative, (0)angle is $0°$ [72A2]. (b) Schematic representation of the most frequently observed puckering modes, corresponding to the pseudorotation. Horizontal transitions are continuous and at the same energy level whereas vertical transitions are separated by a (shallow) energy barrier and describe an N⇌S interchange. Note that directions of exocyclic C1′–N and C4′–C5′ bonds are intrinsically related to sugar conformation [84S1].

Fig. 7a, b
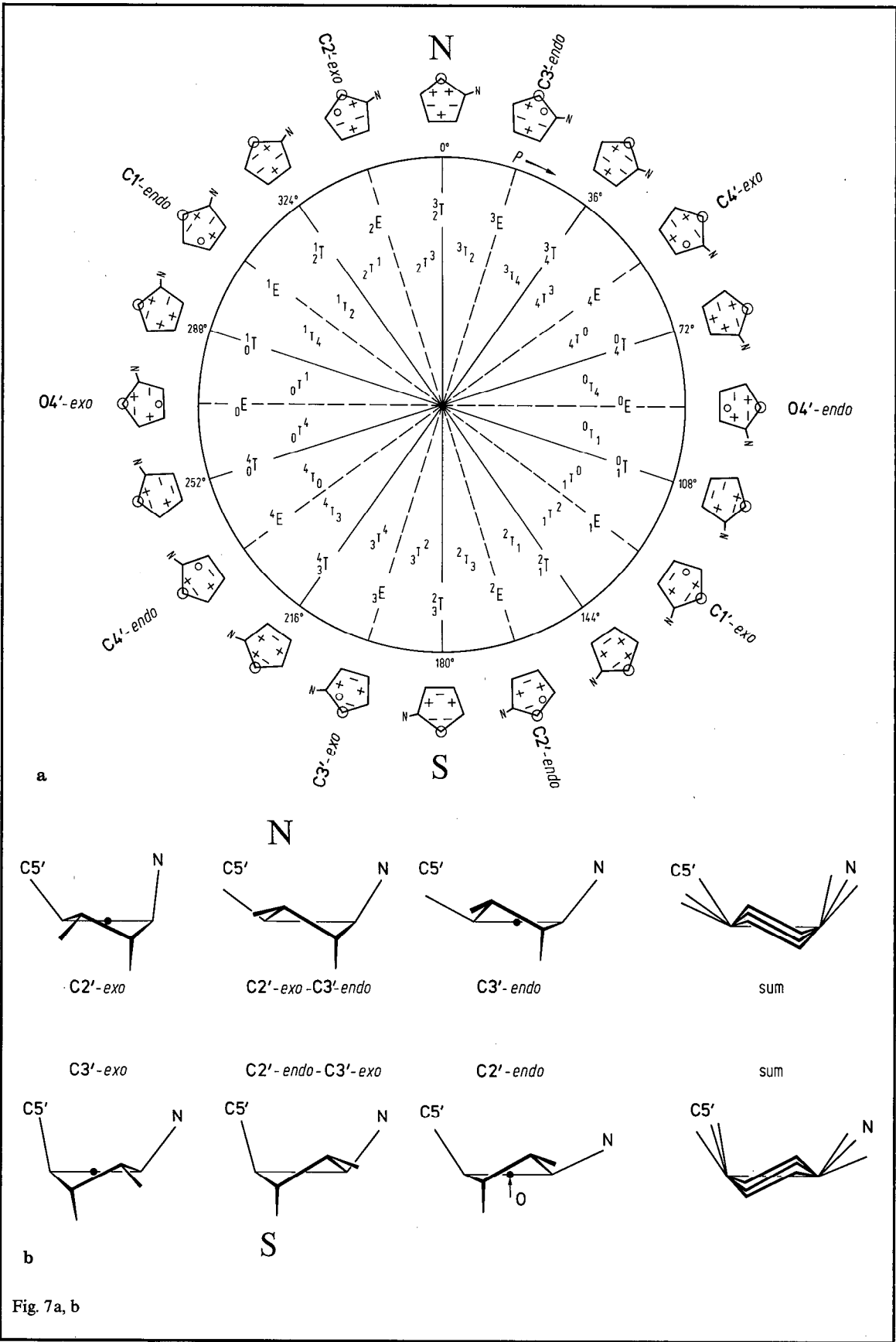
g) Orientation χ about the glycosyl bond: *syn* **and** *anti*

The base can rotate about the glycosyl C1'–N bond that links the furanose C1' atom with N1 of pyrimidine bases (Ura, Cyt) and with N9 of purine bases (Ade, Gua). There are, however, steric limitations which restrict the rotation.

The orientation of the base relative to the sugar is denoted by the torsion angle χ which is defined as

$$O4'-C1'-N1-C2 \quad \text{for pyrimidine nucleosides}$$
$$O4'-C1'-N9-C4 \quad \text{for purine nucleosides.}$$

This convention follows IUPAC recommendations and is consistent with accepted chemical nomenclature.

There are a number of earlier definitions of torsion angle χ which used different atoms on both sides of the C1'–N link: sugar O4' or C2', and pyrimidine C2 or C6 and purine C4 or C8.

The conversions from these χ angles to the now accepted definition are summarized in Table 3.

Table 3. Conversion of different definitions for torsion angle about the glycosyl C1'–N linkage [$\chi_{present} = \chi_{other} + $ difference]. Differences according to [73S1].

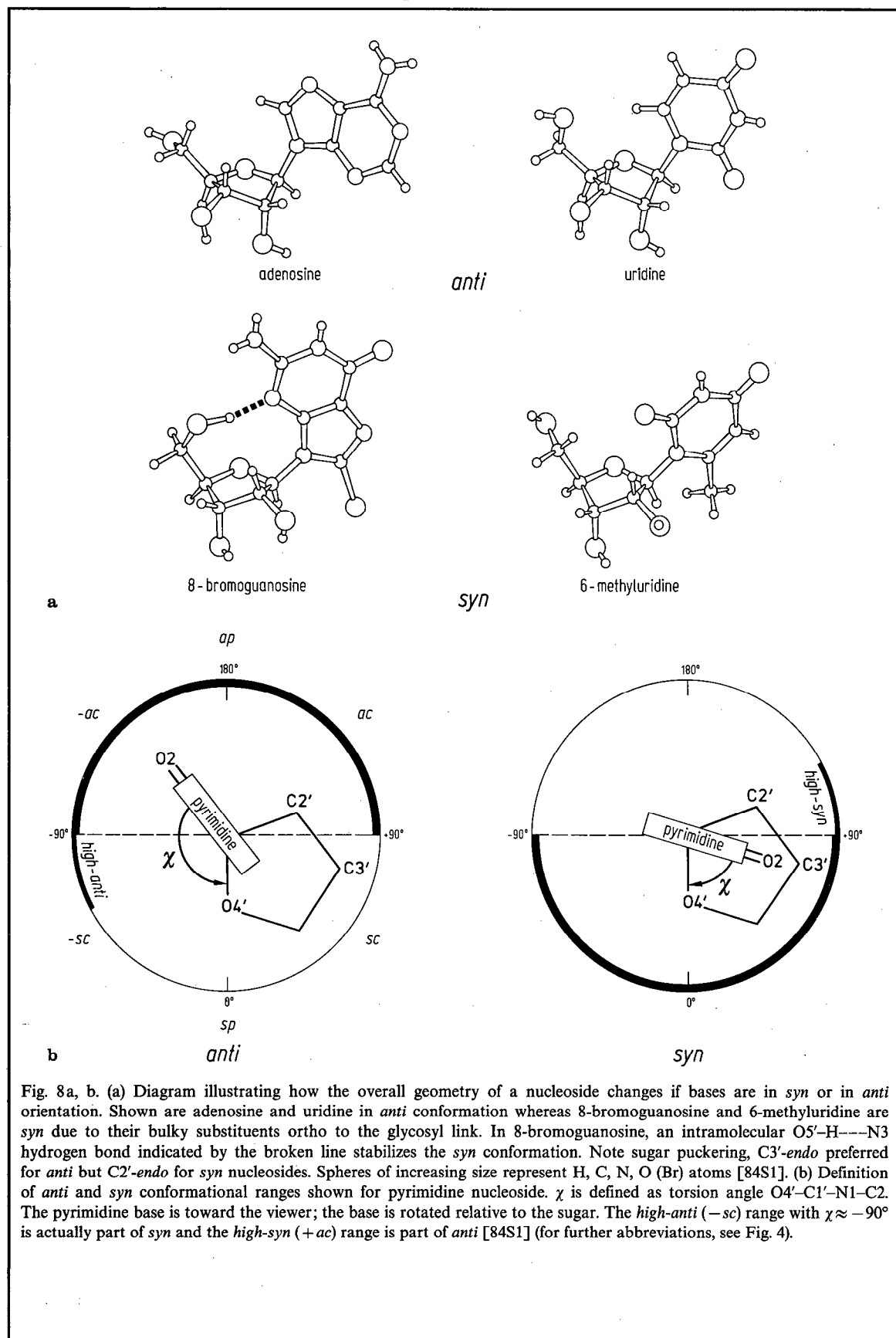|	Present χ	Other definition of χ	Difference between present and other[1]
Purine		O4'–C1'–N9–C8	$+180°$
Pyrimidine		O4'–C1'–N1–C6	
Purine	O4'–C1'–N9–C4	C2'–C1'–N9–C8	$-62.5°$
Pyrimidine	O4'–C1'–N1–C2	C2'–C1'–N1–C6	
Purine		C2'–C1'–N9–C4	$+116.5°$
Pyrimidine		C2'–C1'–N1–C2	

[1]) Approximate values.

There are two main orientations for the relative positions of sugar and base, *anti* and *syn* (Fig. 8a). In *anti*, the bulk of the base is rotated away from the five-membered sugar ring whereas in *syn*, it is located over the sugar ring. Since in the latter orientation steric interference might or does occur, the former is, in general, preferred. The *anti* and *syn* ranges are:

$$anti: \quad \chi = 180° + 90°$$
$$syn: \quad \chi = \quad 0° \pm 90°.$$

There is also a *high-anti* region corresponding to χ in the $-90°$ region, or *$-synclinal$* $(-sc)$ (Fig. 8b).

Fig. 8a, b. (a) Diagram illustrating how the overall geometry of a nucleoside changes if bases are in *syn* or in *anti* orientation. Shown are adenosine and uridine in *anti* conformation whereas 8-bromoguanosine and 6-methyluridine are *syn* due to their bulky substituents ortho to the glycosyl link. In 8-bromoguanosine, an intramolecular O5'-H---N3 hydrogen bond indicated by the broken line stabilizes the *syn* conformation. Note sugar puckering, C3'-*endo* preferred for *anti* but C2'-*endo* for *syn* nucleosides. Spheres of increasing size represent H, C, N, O (Br) atoms [84S1]. (b) Definition of *anti* and *syn* conformational ranges shown for pyrimidine nucleoside. χ is defined as torsion angle O4'-C1'-N1-C2. The pyrimidine base is toward the viewer; the base is rotated relative to the sugar. The *high-anti* ($-sc$) range with $\chi \approx -90°$ is actually part of *syn* and the *high-syn* ($+ac$) range is part of *anti* [84S1] (for further abbreviations, see Fig. 4).

b) Orientation γ about the C4′–C5′ bond

The rotation about the C4′–C5′ bond allows atom O5′ to assume different positions relative to the furanose. Three main conformations with all substituents staggered are possible (Fig. 9).

These conformations can be denoted either by the torsion angle γ, or, still frequently in use, by the two torsion angles $\phi_{OO} = O5'–C5'–C4'–O4'$ and $\phi_{OC} = O5'–C5'–C4'–C3' = \gamma$. The ±gauche, trans notation is used most commonly.

Table 4. Definitions for orientation about the C4′–C5′ bond.

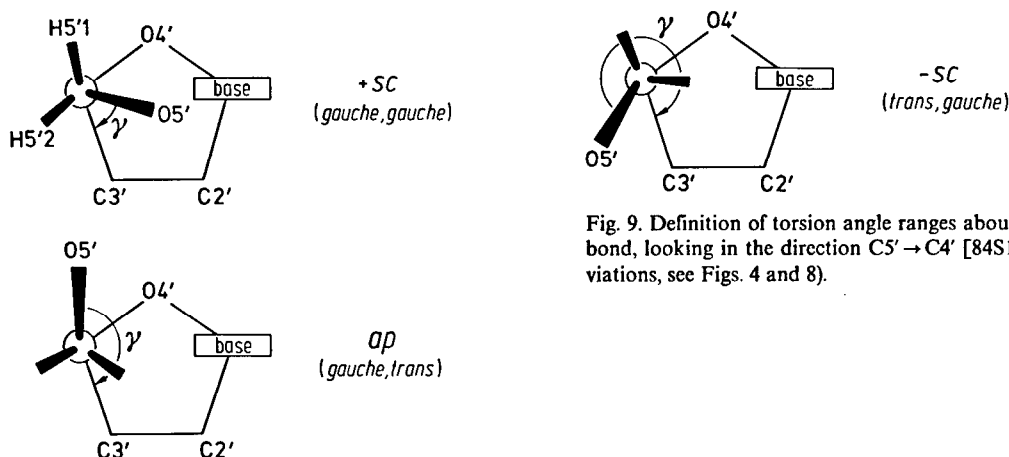| γ	ϕ_{OO}, ϕ_{OC}
+sc	gauche, gauche; gg; (+)gauche; +g
−sc	trans, gauche; tg; trans; t
ap	gauche, trans; gt; (−)gauche; −g

+SC
(gauche, gauche)

−SC
(trans, gauche)

Fig. 9. Definition of torsion angle ranges about the C4′–C5′ bond, looking in the direction C5′→C4′ [84S1] (for abbreviations, see Figs. 4 and 8).

ap
(gauche, trans)

1.3 Geometry of helices

The nucleic acids preferentially form the well-known Watson-Crick double helix but there are also single, triple and quadruple helices which are found with e.g. poly(A), poly(A)·2poly(U), [poly(G)]$_4$.

a) Helical parameters

The parameters used to define the geometry of a helix are (Fig. 10):

n = number of residues (nucleotides) per helix turn
h = unit height (translation per residue along the helix axis)
t = $360°/n$ = unit twist (angle of rotation per residue about the helix axis)
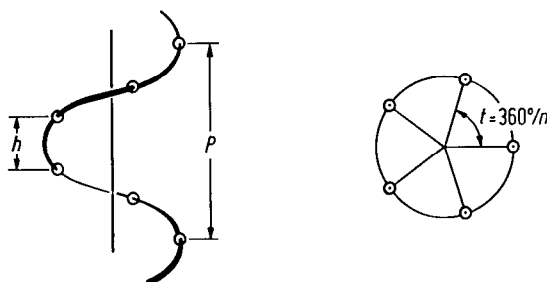P = pitch height of helix (the identical repeat) = $n \cdot h$.

Fig. 10. Definition of helical parameters: pitch P, axial rise per residue h, and unit twist t, shown for a right-handed helix with $n = 5$ residues per turn [84S1].