In some cases there may be more than one reference associated with the same structure. If more than one independent investigation of the same structure was performed they are both listed but counted only once in the summary given in Tables 1a, b. In Table 2, the empirical formula is given below the name of the compound. Pertinent remarks regarding the structure are given under the formula (such as coordinate errors in the original paper; the values reported in the "Cambridge Crystallographic Database" are usually correct). Cell constants are given in the next column with only those angles not equal to 90° listed. The next columns contain the space group symbol, space group number used in the "International Tables for X-ray Crystallography" [69H1], the number of molecules per unit cell (Z), and the final R-factor. The reference to the original publication is given to the right of the R-factor.

On the far right an abbreviated type-code is given defining the compound type. The base type (A = adenine, C = cytosine, G = guanine, T = thymine, U = uracil, B = other), sugar type (R = ribose, D = deoxyribose, A = arabinose, S = other), phosphates (3P = 3′-phosphate, 5P = 5′-phosphate, 5PP = 5′-diphosphate, 35P = 3′,5′-cyclic phosphate) are given. An asterisk * indicates that the moiety is substituted or modified. Cyclization is indicated by an at sign (@). For example, U*A*@ refers to a nucleoside with an uracil base, arabinose sugar, containing a cyclic linkage, with modifications on both the base and sugar moieties. In the case of paired bases or nucleosides in the structure, the codes for the individual components are separated by a comma.

About half of the reported structures cover nucleic acid bases while the remaining structures are nucleosides or nucleotides. In addition to the standard bases (adenine, guanine, cytosine, uridine, and thymine), a wide variety of modified bases have been reported, many of which are so heavily modified that they have very little resemblence to the parent compound. Although the standard bases are nearly planar fused ring systems, small but significant distortions from planarity can be observed. Purine bases often display a slight fold ($1 \cdots 2°$) about the C4–C5 bond which fuses the six and five-membered rings.

In the case of nucleosides, a sugar moiety, usually ribose or 2′-deoxyribose (found in the monomeric units of RNA and DNA, respectively), is attached to the N9 position of purine bases and N1 of pyrimidine bases. In some of the structures, ribose may be replaced by arabinose, other furanose sugars, six-membered pyranose ring, or even linear sugars. The ribose atoms are numbered C1′ through C5′ with the ring oxygen atom referred to as O4′, or O1′ in the earlier literature.

The vast majority of the reported nucleotide structures have one phosphate group at the 5′ position and to a lesser extent at the 3′ position while considerably fewer structures have been reported with di- or triphosphate moieties at the 5′ position. Several dinucleotide structures and one trinucleotide have also been reported which provide an insight into the sugar-phosphate backbone chain conformations. Many of the dinucleotide structures are stabilized by planar drug molecules which intercalate between adjacent base pairs. A more detailed discussion of drug-nucleotide interactions is found in chapter 2.6 of this volume.

Many of the reported crystal structures including many nucleotides are associated with metal ions. A summary of the metal ions found in the various type of crystal structures is given in Table 1b. In the case of nucleotides, the positively charged metal ions stabilize the negatively charged phosphate groups either by direct coordination of the phosphate oxygen atoms or via water molecules. The metal ions can also form direct coordination with the exposed nitrogen atoms of the base skeleton such as N1, N3, and N7 of purine nucleosides and N3 for pyrimidine nucleosides as well as bonding to amino or oxo substituents on the base [79S1]. Chapter 2.5 presents the details of the molecular structures of the transition metal-nucleotide and other heavy metal-nucleotide complexes.

2.1.1.4 Nucleoside and nucleotide conformation

The nucleosides and nucleotides have several degrees of freedom which can affect the three-dimensional conformation of the molecule [69S1, 73S1, 75S1, 84S1]. The three major conformational parameters or degrees of freedom for nucleoside structures are (a) the spatial disposition of the base relative to the sugar moiety, (b) the puckering of the ribofuranose sugar ring, and (c) the spatial disposition of the exocyclic O5′ atom relative to the rest of the sugar moiety. These parameters including the endocyclic sugar ring torsion angles are given in Table 6. In the case of the nucleotide structures, tabulated in Table 7, additional conformational parameters associated with the sugar-phosphate chain are also provided. Table 8 lists these values for the di- and trinucleotide structures.

Base disposition: The first of the three basic conformational parameters is measured by the glycosyl torsion angle. In accordance with the recommended IUPAC convention, the glycosyl torsion angle C4–N9–C1′–O4′ for purine bases and the C2–N1–C1′–O4′ torsion angle for pyrimidine bases are used to describe the base disposition. An earlier nomenclature which is still widely used for describing the base disposition is the C8–N9–C1′–O4′ torsion angle for purine bases and the C6–N1–C1′–O4′ torsion angle for pyrimidine bases. Although the torsion angles for the two conventions will differ by approximately 180°, they are both commonly referred to in the literature with the greek symbol chi χ, which is an important point to be observed when interpreting the structural literature. There are two major domains observed for this parameter: the *anti* domain where the six-membered ring of purine bases and the oxygen atoms(s) of pyrimidine bases point away from the ribofuranose ring, and the *syn* where the base is rotated 180° about the glycosyl torsion angle with the six-membered ring of the purine or O2 oxygen atom of the pyrimidine located over the ribofuranose ring.

A plot of the observed glycosyl torsion angles for both nucleosides and nucleotides is given in Fig. 1a···j. Several general conclusions can be drawn from the data presented in Fig. 1.

First, the glycosyl torsion angle for pyrimidine bases is found almost entirely within the *anti* domain with very few *syn* structures. Although the majority of the structures with purine bases are also *anti*, there is a significant tendency for these structures to adopt the *syn* conformation. The purine nucleosides display a significantly greater variance for the glycosyl torsion angle than the purine nucleotides while in the case of pyrimidine structures, the increased variability for nucleosides is much less marked. The *syn* conformation for the purine nucleosides is stabilized by an intramolecular hydrogen bond between N3 of the base and the exocyclic O5′ atom of the sugar while the purine nucleotides cannot be similarly stabilized due to the presence of the phosphate group. Even for the case of cyclic nucleotides, the base disposition usually falls either into the *anti* or *syn* domains.

Sugar pucker and pseudorotation: The next major conformational parameter of interest is the pucker of the sugar or ribofuranose ring. This ring puckering could be described by the five endocyclic torsion angles for the five-membered ring, however, ring closure provides a rather severe constraint on the values which these endocyclic torsion angles may adopt. Due to these constraints, the conformational flexibility of any ring can be uniquely described by just two parameters using the pseudorotation concept [72A1, 81R1, 85M1]. This is formally accomplished by a Fourier summation of the endocyclic torsion angles θ_i defined as follows:

$$A = 0.4 \sum_{i=1}^{5} \theta_i \cos[0.8\pi(i-1)],$$

$$B = -0.4 \sum_{i=1}^{5} \theta_i \sin[0.8\pi(i-1)],$$

where A and B may be considered as the real and imaginary components, respectively, of a complex number which uniquely describes the puckering of any five-membered ring (assuming equal bond lengths).

The maximum amplitude t_m of pseudorotation (in degrees) is given by:

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

and the phase angle P of pseudorotation by:

$$P = \tan^{-1}(B/A), \quad \text{if } A < 0 \text{ then } 180° \text{ is added to } P.$$

Note:

This is a formal description of the pucker of a five-membered ring. When the torsion angle θ_i has its maximum value ($\theta_i \equiv t_m$), the phase angle P is zero ($P = 0$).

According to the IUPAC convention [83I1] the endocyclic torsion angles of the furanose ring are designated by v_0, v_1, v_2, v_3 and v_4 and consequently the maximum amplitude should be designated by v_{max} (in contradistinction to the IUPAC which recommends the symbol ψ_m, see section 1.2). Furthermore the phase angle P is defined to be zero when v_2 has its maximum value.

Therefore, for the characterization of the sugar pucker, the following transformation has to be applied:

$$\theta_1 \rightarrow v_2, \quad \theta_2 \rightarrow v_3, \quad \theta_3 \rightarrow v_4, \quad \theta_4 \rightarrow v_0, \quad \theta_5 \rightarrow v_1 \quad \text{and} \quad t_m \rightarrow v_{max}.$$

This nomenclature is used in Tables 6···8 (sections 2.1.2.6···2.1.2.8).

The phase angle of pseudorotation (P) describes the direction of puckering for each of the five ring atoms relative to the C5′ atom. An atom is considered *endo* if on the same side as C5′ and as *exo* if on the opposite side. Figure 2 presents the pseudorotation pathway of the sugar ring illustrating the correlation between the pseudorotation phase angles and the envelope (E) or twist (T) conformations for the sugar. The most severely buckled atom is listed on the left of the letter as a superscript for *endo* and as a subscript for *exo*. In the case of a twist pucker, the less puckered secondary atom is listed to the right of the letter once again either as a superscript or subscript depending on whether the atom is *endo* or *exo*. In the case of a symmetrical twist, both atom numbers are placed to the left of the letter, the *endo* one as a superscript and the *exo* atom as a subscript. The maximum amplitude of pseudorotation is zero for a perfectly planar ring and increase with the degree of ring puckering.

The pseudorotation parameters for both nucleosides and nucleotides are graphically depicted in Fig. 3. As seen in the figure, the phase angles are not uniformly distributed but instead tend to adopt one of two preferred puckering conformations. The C3′-*endo* domain in the region $P = 340° \cdots 40°$ and the C2′-*endo* domain from $P = 140° \cdots 200°$ are by far the most preferred. While the ribonucleoside structures appear to be equally distributed between these two preferred domains, the deoxyribonucleosides appear to favor the C2′-*endo* domain as opposed to the C3′-*endo* domain. In the case of the cyclized nucleosides, where the sugar conformation is often restricted, the region from $P = 210° \cdots 270°$ is prevalent, in contrast to the standard nucleosides and nucleotides where virtually no structures are found in this 'forbidden' region. The puckering amplitude is also constrained in the cyclic structures with a number of relatively flat rings (ψ_m less than 30°) in contrast to the standard sugars (where ψ_m is almost always greater than 30°).

Disposition of the exocyclic O5′ atom: The exocyclic C3′–C4′–C5′–O5′ torsion angle describes the disposition of the O5′ atom and the attached phosphorus atom, if any, relative to the other parts of the structure. For this parameter, there are three favored domains: the C3′–C4′–C5′–O5′ torsion angle can be in the vicinity of $-60°$(*−gauche*), 180°(*trans*), or $+60°$(*+gauche*). A torsion angle plot is given in Fig. 4. For both nucleosides and nucleotides, the *+gauche* region is the most preferred. While there are also a significant number of *trans* structures in the nucleoside series, very few *−gauche* structures are observed. The *+gauche* conformation brings the nucleoside O5′ atom in closest proximity to the base to facilitate hydrogen bonding interactions with the base, however, in rare cases the O5′ can still participate in a hydrogen bond even in the *trans* conformation. In contrast to the nucleosides, the nucleotides prefer the *−gauche* region over the *trans* region for both the purine and the pyrimidine bases.

Disposition of phosphate group: In the case of 5′-mononucleotides, the C4′–C5′–O5′–P torsion angle is also an important conformational feature (Fig. 5). This torsion angle is *trans* for the vast majority of the mononucleotides. In the case of di- or trinucleotides, there are additional backbone torsion angles at the phosphorus atoms which are included in Table 8: C5′–O5′–P–O3′ and O5′–P–O3′–C3′. The conformations around these internucleotide P–O bonds provide a measure of the relative disposition of the linked nucleotides. These torsion angles are highly sensitive to the helical sense of the DNA or RNA polymers (such as left or right handed helices, and number of residues per turn). The two aforementioned torsion angles are actually highly correlated since only certain combinations will allow the bases of adjacent nucleotide residues to properly stack over each other. A plot of the C5′–O5′–P–O3′ versus O5′–P–O3′–C3′ torsion angles for dinucleotides is given in Fig. 6a which shows one highly preferred domain although a variety of other combinations are also allowed. In contrast, the corresponding torsion angles for the 3′–5′-cyclic mononucleotides are extremely constrained resulting in only one allowed domain as can be seen from Fig. 6b.

In the case of 5′-mononucleotides, the presence of the phosphate group at the O5′ position limits the conformational flexibility of the remainder of the molecule. For example, in 5′-nucleotides the *anti* conformation is favored since the *syn* conformation results in unfavorable steric interactions between the base and phosphate groups. The *syn* conformation will necessarily involve rotation around the C4′–C5′ bond resulting in the *trans* or *−gauche* conformation.

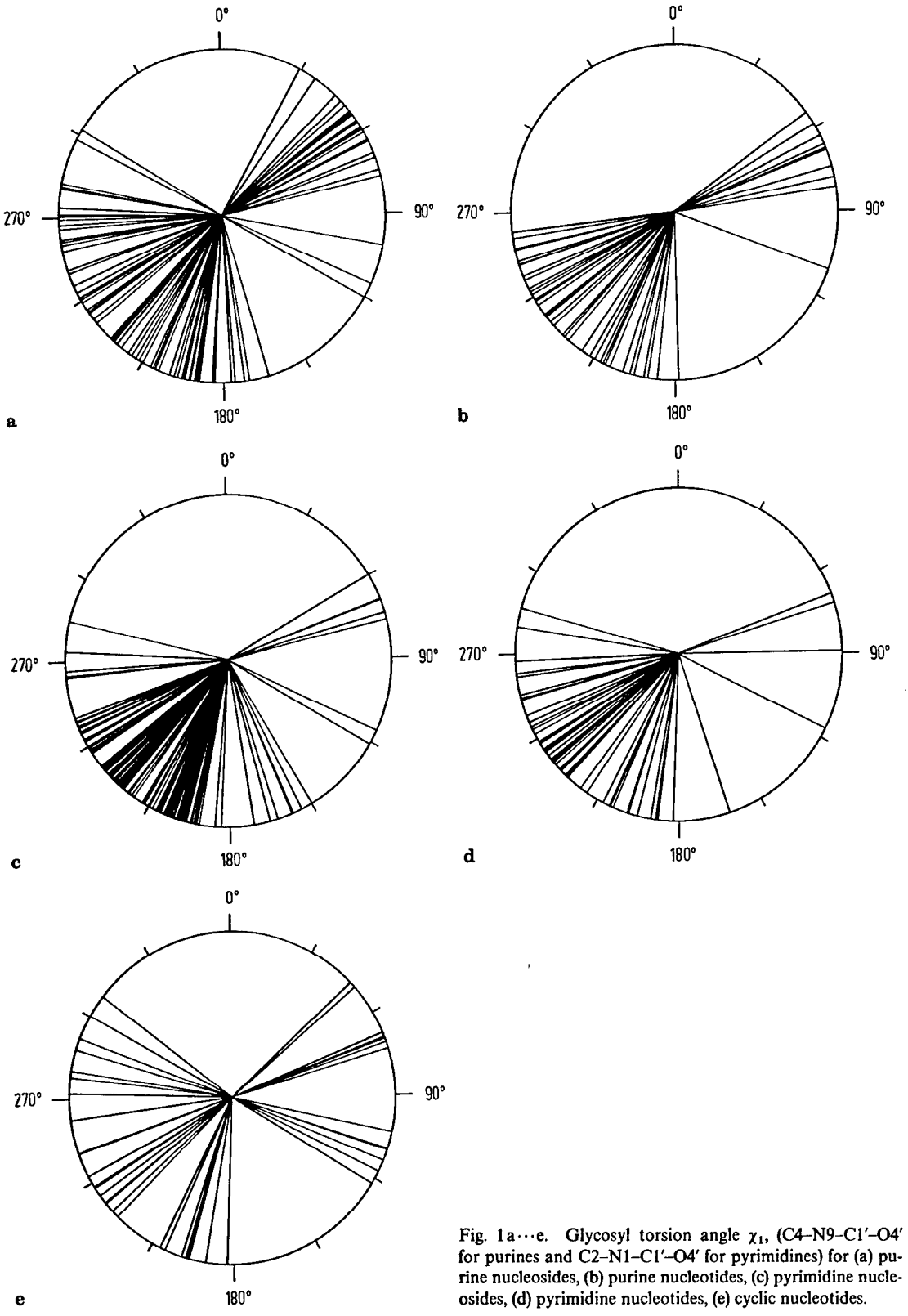
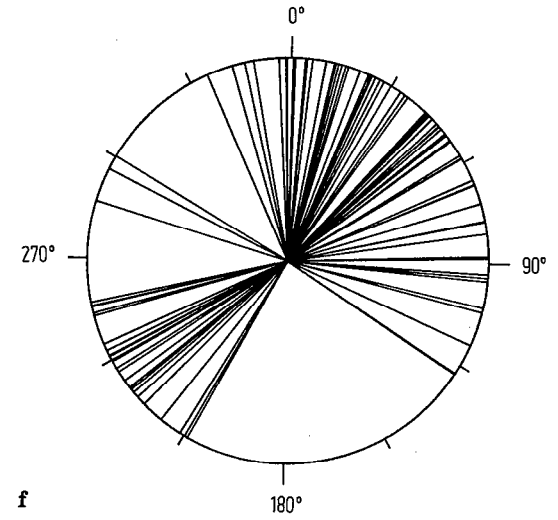
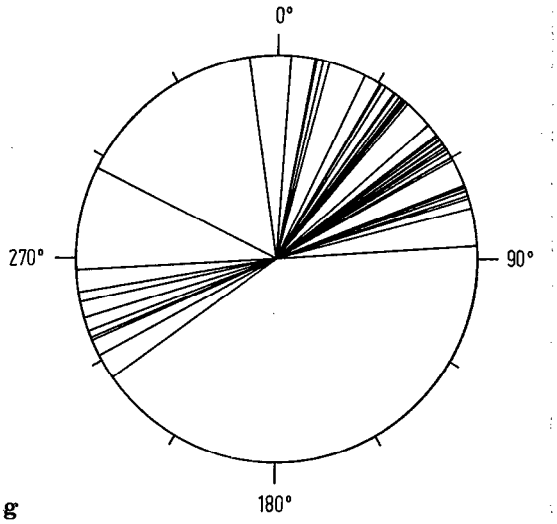


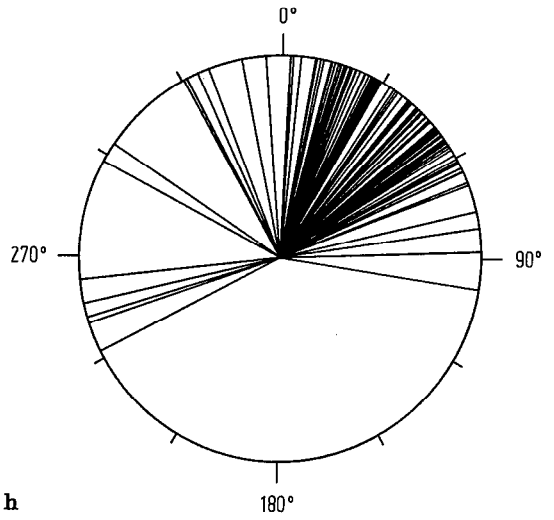
Fig. 1a···e. Glycosyl torsion angle χ_1, (C4–N9–C1′–O4′ for purines and C2–N1–C1′–O4′ for pyrimidines) for (a) purine nucleosides, (b) purine nucleotides, (c) pyrimidine nucleosides, (d) pyrimidine nucleotides, (e) cyclic nucleotides.



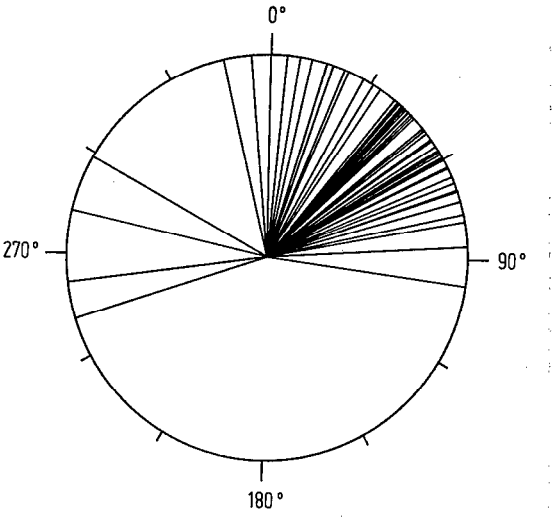
f



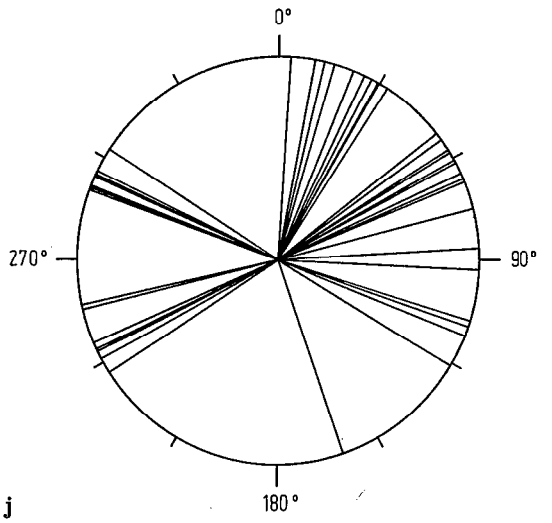
g



h



i



j

Fig. 1f···j. Glycosyl torsion angle χ_2 (C8–N9–C1′–O4′ for purines and C6–N1–C1′–O4′ for pyrimidines, according to an older convention) for (f) purine nucleosides, (g) purine nucleotides, (h) pyrimidine nucleosides, (i) pyrimidine nucleotides, (j) cyclic nucleotides.

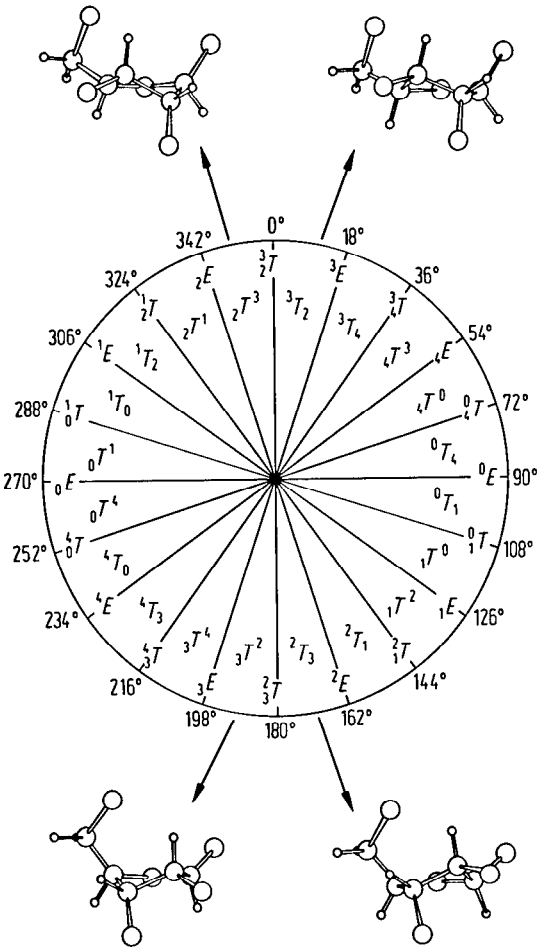


Fig. 2. Pseudorotation pathway of the sugar ring illustrating the correlation between the pseudorotation phase angle P and the envelope E and twist T conformations of the sugar. Molecular diagrams for four types of puckers are also shown. The pucker on the top (right), the puckers on the bottom, and slight variants of these are the most familiar puckers. In addition to them, the intermediate O4'-endo pucker also appears to be important. The symmetrical twist conformations ${}^{3}_{2}T$, ${}^{3}_{4}T$, etc. ($P = 0°$, 36°, etc.) have a diad symmetry in the ring, while the envelope conformations ${}^{3}E$, ${}_{4}E$, etc. ($P = 18°$, 54°, etc.) have a mirror plane symmetry.

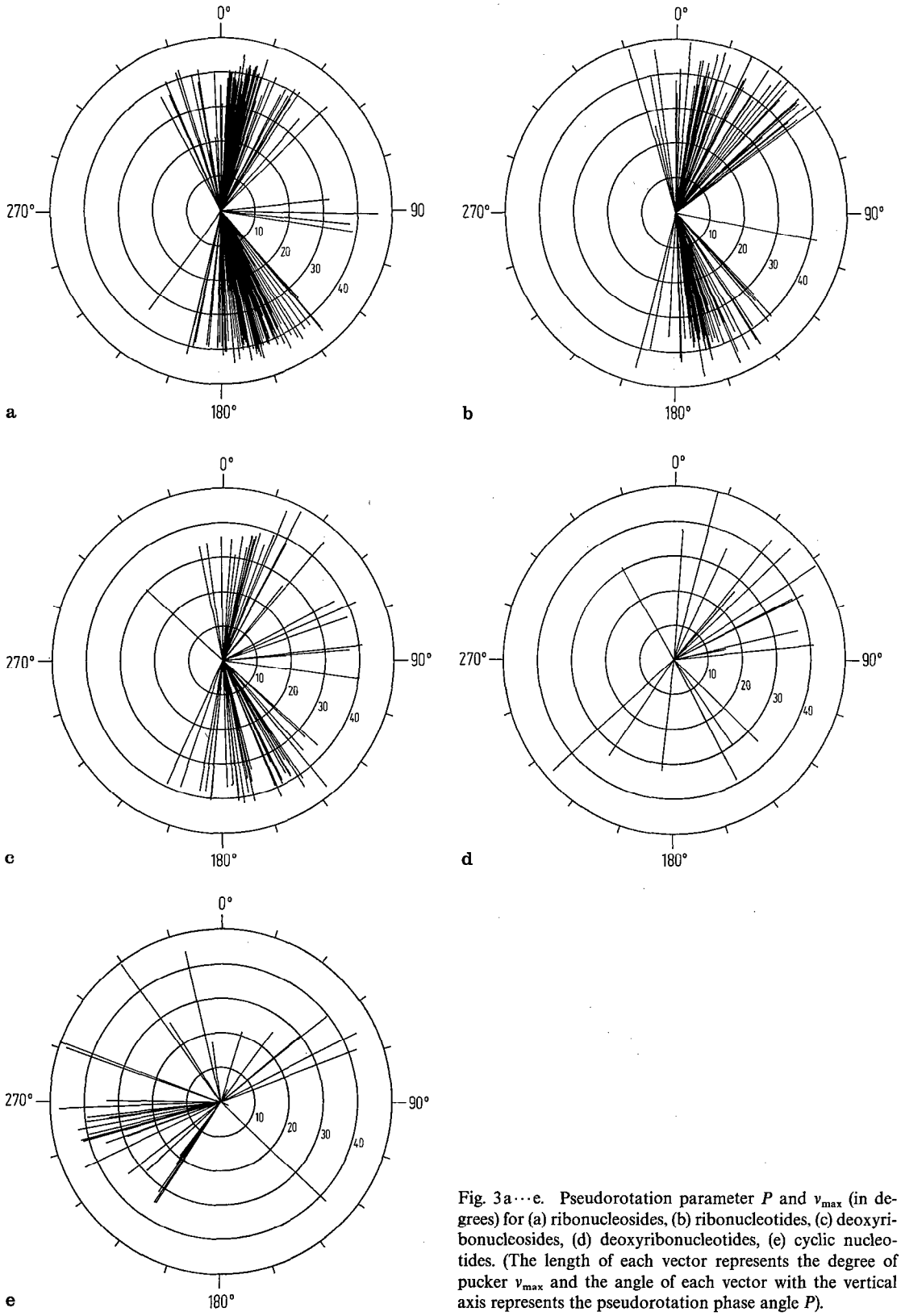


Fig. 3a···e. Pseudorotation parameter P and ν_{max} (in degrees) for (a) ribonucleosides, (b) ribonucleotides, (c) deoxyribonucleosides, (d) deoxyribonucleotides, (e) cyclic nucleotides. (The length of each vector represents the degree of pucker ν_{max} and the angle of each vector with the vertical axis represents the pseudorotation phase angle P).

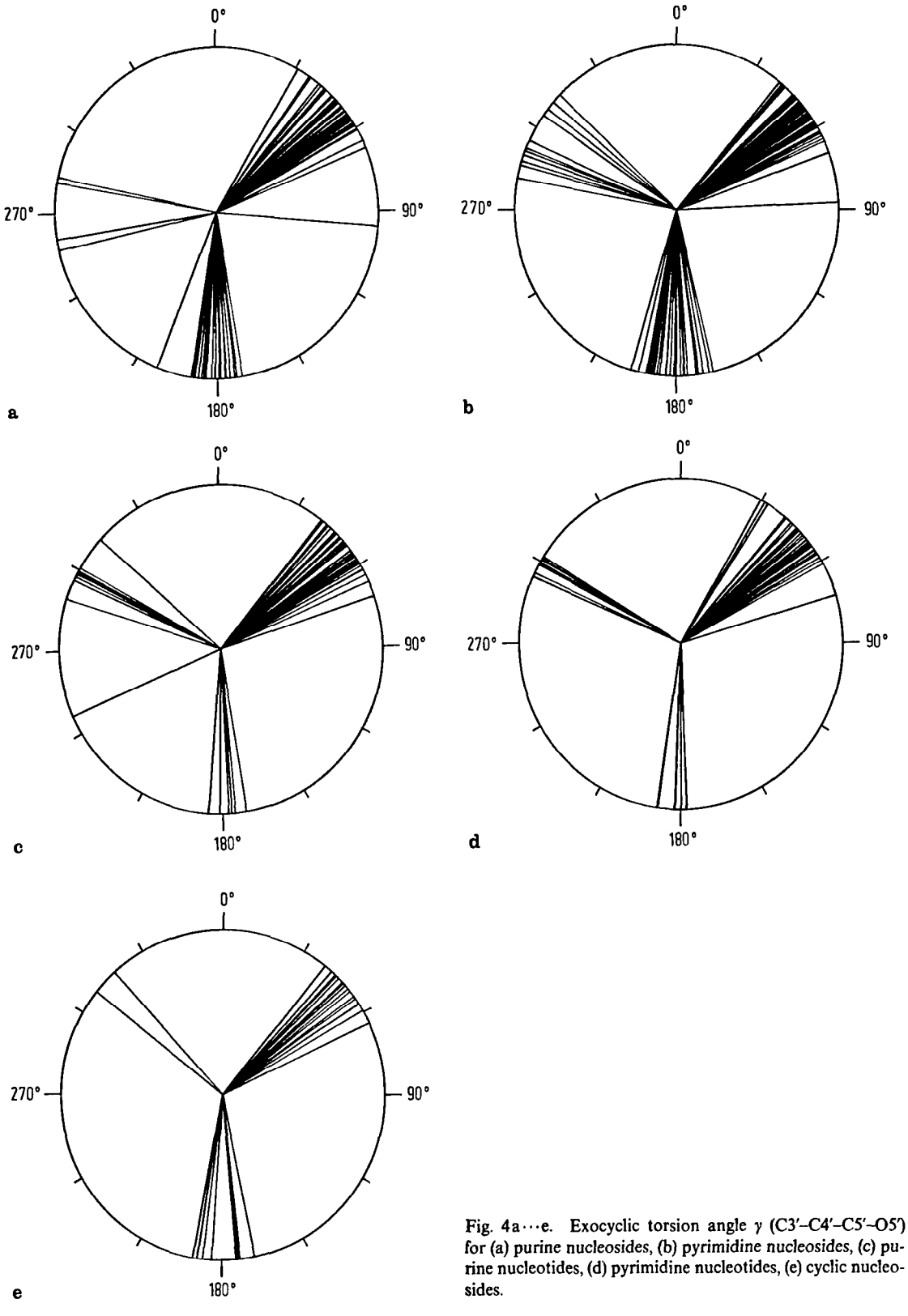


Fig. 4a···e. Exocyclic torsion angle γ (C3′–C4′–C5′–O5′) for (a) purine nucleosides, (b) pyrimidine nucleosides, (c) purine nucleotides, (d) pyrimidine nucleotides, (e) cyclic nucleosides.

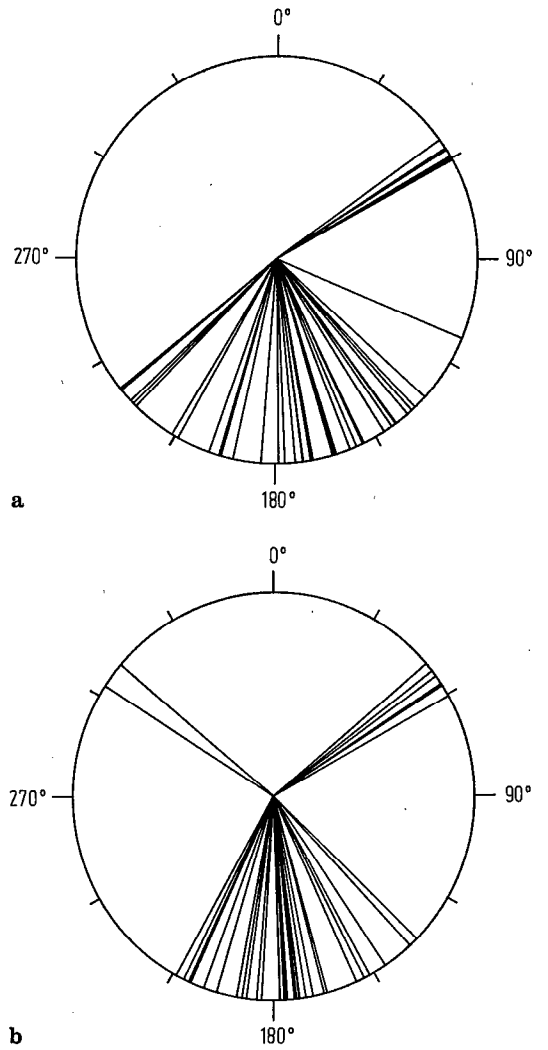


Fig. 5a, b. Torsion angle β (C4′–C5′–O5′–P5′) for (a) purine nucleotides, (b) pyrimidine nucleotides.

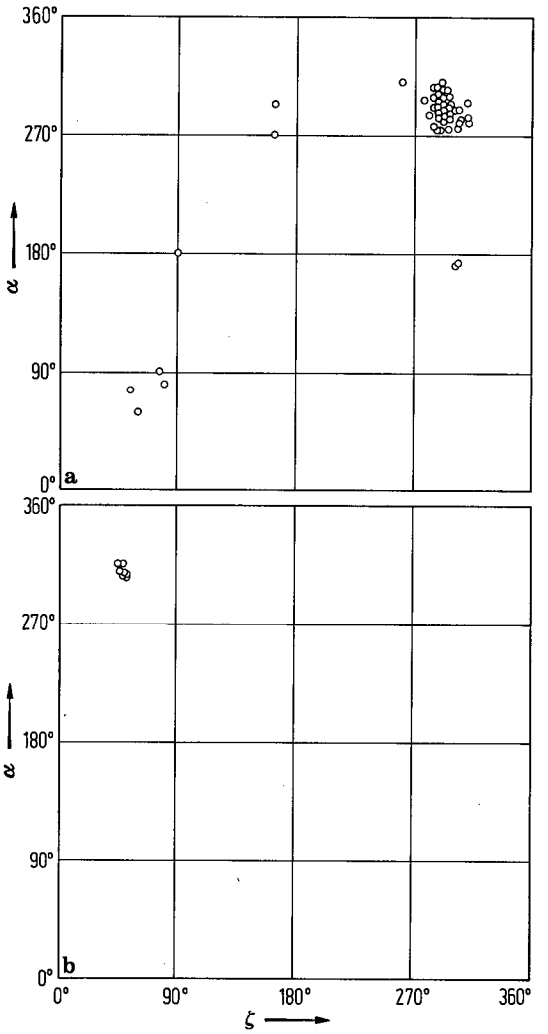


Fig. 6a, b. Torsion angle α (O3′–P–O5′–C5′) (vertical axis) vs. torsion angle ζ, (C3′–O3′–P–O5′) (horizontal axis) for (a) di(tri)nucleotides and (b) 3′,5′-cyclic nucleotides.

2.1.1.5 Analog structures

Most of crystal structures which have been reported are analogs of the naturally occurring bases, nucleosides, and nucleotides. Some of these analogs are found in nature, however, most of them have been synthesized to further our understanding of various enzyme mechanisms and structure-function relationships. These analogs have been used to study DNA, RNA, and protein synthesis, as well as to study the precursors in nucleic acid biosynthesis. They have also been used as substrate probes in numerous enzyme systems, many of which display some degree of biological activity, with a few analogs reacting even faster than the standard substrate. They have also furthered our understanding of various aspects of intermediary metabolism, subcellular organization, and certain hormone systems. Usually, the biological activities of the analogs result from their structural similarity with the corresponding natural enzyme substrates. Many analogs function as enzyme inhibitors, thereby interfering with nucleic acid or protein synthesis, a function which proves useful for the design of various antibiotic drugs. The analogs can be divided into the following general categories: minor substitutions, ring skeleton substitutions, cyclization, and major substitution or rearrangement.