Biochemistry

Volume 9, Number 18 September 1, 1970

IUPAC–IUB Commission on Biochemical Nomenclature. Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains.[*]

Tentative Rules (1969)

Preamble

These rules are based on "A proposal of standard conventions and nomenclature for the description of polypeptide conformation" (Edsall *et al.*, 1966) and have been prepared by a subcommission set up by the IUPAC–IUB Commission on Biochemical Nomenclature in 1966.[1] The original proposals have been modified so as to bring them as far as possible into line with the system of nomenclature current in the fields of organic and polymer chemistry.

Two recommendations are appended to the rules, the first dealing with the terms configuration and conformation and the second with primary, secondary, and tertiary structure. These are formulated as recommendations rather than rules because there is at present no general agreement about their definition.

Note. Two alternative notations are recommended throughout. That with superscripts and subscripts may be used when it is unlikely to cause confusion, *e.g.*, in printed or manuscript material; that without is to be used where superscripts or subscripts may cause confusion or are technically difficult or impossible, *e.g.*, in computer outputs. In the latter connection the following Roman equivalents of Greek letters are recommended: α, A; β, B; γ, G; δ, D; ϵ, E; ζ, Z; η, H; τ T; υ, U; ϕ, F; χ, X; ψ, Q; ω, W.

Rule 1. General Principles of Notation

1.1. *Designation of atoms.* The atoms of the main chain are denoted thus

$$\mathrm{—NH—\underset{|}{C^{\alpha}}H^{\alpha}—CO—}$$

Where confusion might arise the following additional symbolism may be used

$$\mathrm{—N'H'—\underset{|}{C^{\alpha}}H^{\alpha}—C'O'—}$$

1.2. Amino acid residues, -NH-CHR-CO-, are numbered sequentially from the amino-terminal to the carboxyl-terminal end of the chain, the residue number being denoted *i*.

Example

C^{α} of the *i*th residue is written C_i^{α} or Cα(*i*)

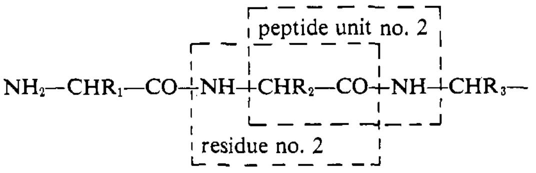
1.3. *Peptide units.* For some purposes it is more convenient to group together the atoms -CHR-CO-NH-. These groups are described as "peptide units," and the peptide unit number, like the residue number, is denoted *i*. It will be noted that the two numbers are identical for all atoms except NH; generally there will be no confusion, because a single document will use either "residues" alone, or "peptide units" alone, but in the latter case explicit reference must be made to this usage at the beginning. If confusion might arise, the symbols N_i^* and H_i^* are to be used for these atoms in the *i*th peptide unit, which are N_i and H_i in the *i*th residue (so that $N_i^* = N_{i+1}$).

---

[*] Document of the IUPAC–IUB Commission on Biochemical Nomenclature (CBN), approved by IUPAC and IUB in 1969 and published by permission of IUPAC and IUB. Comments on and suggestions for future revisions of these tentative rules may be sent to any member of CBN: O. Hoffmann-Ostenhof (Chairman), W. E. Cohn (Secretary), A. E. Braunstein, P. Karlson, B. Keil, W. Klyne, C. Liébecq, E. C. Slater, E. C. Webb, and W. J. Whelan. Reprints of this document may be obtained from Waldo E. Cohn, Director, NAS–NRC Office of Biochemical Nomenclature, Biology Division, Oak Ridge National Laboratory, Box Y, Oak Ridge, Tenn. 37830.

[1] The members of the subcommission were J. C. Kendrew (Chairman), W. Klyne, S. Lifson, T. Miyazawa, G. Némethy, D. C. Phillips, G. N. Ramachandran, and H. A. Scheraga. In addition, the following assisted in the work of the subcommission: R. S. Cahn, R. Diamond, J. T. Edsall, P. J. Flory, C. K. Ingold, A. Liquori, V. Prelog, and J. A. Schellman.

Example



residue notation	N_2	C_2^α	C_2	N_3
peptide unit notation	N_1*	C_2^α	C_2	N_2*

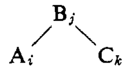
Notes. (i) Residue notation is used throughout these rules.

(ii) Whether "residues" or "peptide units" are being used, ϕ_i and ψ_i always refer to torsion angles about bonds of the same C_i^α.

1.4. *Bond lengths.* If a bond A–B be denoted A_i–B_j or A_i (see Rules 3.1, 4.5), the bond length is written $b(A_i,B_j)$ (or $b(Ai,Bj)$, or b_i^A (or $bA(i)$). An abbreviated notation for use in side chains is indicated in Rule 4.5.

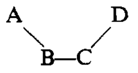
Note. The symbol previously recommended for bond length was *l*. This symbol is no longer recommended, partly because it is easily confused with 1 in many type fonts and partly because it is also used for vibration amplitude in electron diffraction and spectroscopy.

1.5. *Bond angles.* The bond angle included between three atoms



is written $\tau(A_i,B_j,C_k)$, which may be abbreviated, if there is no ambiguity, to $\tau(B_j)$ or τ_j^B or $\tau B(j)$.

1.6. *Torsion angles.* If a system of four atoms



is projected onto a plane normal to bond B–C, the angle between the projection of A–B and the projection of C–D is described as the *torsion angle*[2] of A and D about bond B–C; this angle may also be described as the angle between the plane containing A, B, and C and the plane containing B, C, and D. The torsion angle is written in full as $\theta(A_i,B_j,C_k,D_l)$, which may be abbreviated, if there is no ambiguity, to $\theta(B_j,C_k)$, $\theta(B_j)$, or θ_j^B etc. In the eclipsed conformation in which the projections of A–B and C–D coincide, θ is given the value 0° (synplanar conformation). A torsion angle is considered positive ($+\theta$) or negative ($-\theta$) according as, when the system is viewed along the central bond in the direction B → C (or C → B), the bond to the front atom A (or D) requires rotation to the right or to the left, respectively, in order that it may eclipse the bond to the rear atom D (or A); note that it is immaterial whether the system be viewed from one end or the other. These relationships are illustrated in Figure 1.

Notes. (i) Angles are measured in the range $-180 < \theta \leq +180°$, rather than from 0 to 360°, so that the relationship between enantiomeric configurations or conformations can be readily appreciated.

(ii) The symbols actually used to describe the various torsion angles important in polypeptides are ϕ, ψ, ω, υ, and χ (see Rules 3.2, 4.5.2). In the above θ is used simply as an illustrative generic symbol covering all these.

Rule 2. The Sequence Rule, and Choice of Torsion Angle

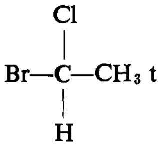
2.1. The rules here enunciated for use in the field of synthetic polypeptides and proteins are in general harmony with the sequence rule of Cahn *et al.*,[3] with the exceptions of Rules 2.1.1 and 2.2.2 (cases II and III), and later rules dependent upon these. The sequence rule was formulated as a universal and unambiguous means of designating the "handedness" or chirality of an element of asymmetry. It includes subrules for the purpose of arranging atoms or groups in an order of precedence or preference, and this system may conveniently be used in the description of steric relationships across single bonds (see Klyne and Prelog, 1960). Here its function is to determine the priority or precedence of different atoms or groups attached to the same atom. However, Rule 2.1.1 below overrides the precedences of the sequence subrules, providing a new "local" (specialist) system for use with the general sequence rule.[4] After application of Rule 2.1.1, the normal procedure of the sequence rule is applied, but modified by Rule 2.2.2; in this connection the only parts of the sequence rule required are given in Rules 2.1.2–2.1.5.

2.1.1. The main chain is given formal priority over branches, notwithstanding any conflict with the following rules. Thus the main chain has precedence at C^α over the side chain and at C′ over O′.

Note. This rule has not yet been formally accepted except in the present context.

2.1.2. The order of (decreasing) priority is the order of (decreasing) atomic number.

Example

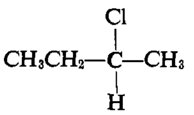


In $Br—\overset{Cl}{\underset{H}{C}}—CH_3$ the order of priority is Br, Cl, CH_3, H.

2.1.3. If two atoms attached to the central atom are the same, the ligands attached to these two atoms are used to determine the priority.

Examples

(i) In



$$CH_3CH_2—\overset{Cl}{\underset{H}{C}}—CH_3$$

---

[2] The terms dihedral angle and internal rotation angle are also used to describe this angle, and may be regarded as alternatives to torsion angle though the latter has been used throughout these rules.

[3] See Cahn *et al.* (1966), and IUPAC Tentative Rules for the Nomenclature of Organic Chemistry, Section E, IUPAC Information Bulletin (1969). Earlier papers: Cahn and Ingold (1951); Cahn *et al.* (1956). For a partial, simplified account see Cahn (1964) and Eliel (1962).

[4] Other local systems are available analogously for steroids, carbohydrates, and cyclitols, where the sequence rule is applied when the local system does not suffice.

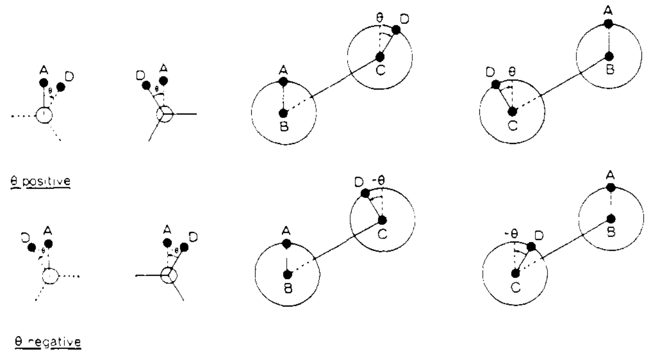
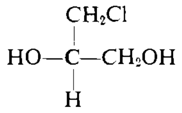


FIGURE 1: Newman and perspective projections illustrating positive and negative torsion angles. Note that a right-handed turn of the bond to the front atom about the central bond gives a positive value of θ from whichever end the system is viewed.

the order is Cl, (CH_2CH_3), CH_3, H. ($C^xH_2CH_3$ takes precedence over C^yH_3 because C^x is bonded to C, H, H and C^y to H, H, H).

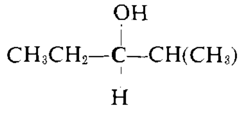
(ii) In



$$\begin{array}{c}\mathrm{CH_2Cl}\\ |\\ \mathrm{HO{-}C{-}CH_2OH}\\ |\\ \mathrm{H}\end{array}$$

the order is OH, CH_2Cl, CH_2OH, H.

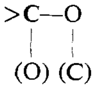
(iii) In



$$\begin{array}{c}\mathrm{OH}\\ |\\ \mathrm{CH_3CH_2{-}C{-}CH(CH_3)_2}\\ |\\ \mathrm{H}\end{array}$$

the order is OH, $CH(CH_3)_2$, CH_2CH_3, H.

2.1.4. A double bond is formally treated as though it were split. Thus >C=O is treated as



$$>\underset{\mathrm{(O)}}{\overset{}{\mathrm{C}}}\!-\!\underset{\mathrm{(C)}}{\mathrm{O}}$$

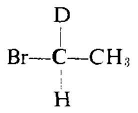
Example

In $CH_3CO{-}OH$ the order is =O, —OH, CH_3.

2.1.5. If two ligands are distinguished only by having different masses (*e.g.*, deuterium and hydrogen), the heavier takes precedence.

Example

In



$$\begin{array}{c}\mathrm{D}\\ |\\ \mathrm{Br{-}C{-}CH_3}\\ |\\ \mathrm{H}\end{array}$$

the order is Br, CH_3, D, H.

Note. This rule is to be used only if the two previous rules do not give a decision.

2.2. *Choice of torsion angle and numbering of branches (tetrahedral configurations).*

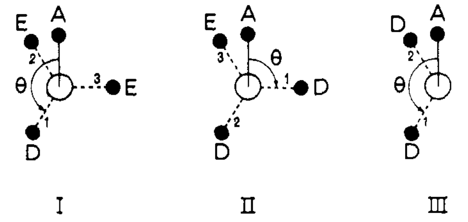
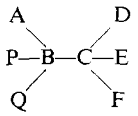


FIGURE 2: Tetrahedral configurations: case I, D > E = E; case II, D = D > E; case III, D = D (E nonexistent).

2.2.1. If, in a compound



the sequence rule gives the priorities A > P, Q and D > E > F, then the *principal torsion angle θ* is that measured by reference to the atoms A–B–C–D as in Rule 1.6. The branches beginning at C are numbered

$$\underset{1}{\mathrm{C{-}D}},\ \underset{2}{\mathrm{C{-}E}},\ \text{and}\ \underset{3}{\mathrm{C{-}F}}$$

2.2.2. If two branches are identical, and the third is different (or nonexistent), they are numbered in a clockwise sense when viewed in the direction B → C, as follows (see Figure 2).

Case I: D > E = E. D has the highest priority and is given the smallest number (1).

Case II: D = D > E. E has the lowest priority and is given the largest number (3).

Case III: D = D, numbered 1 and 2 (E nonexistent). In each case the principal torsion angle is measured between A–B and branch 1.

Notes. (i) The rule given in case II differs from conformational selection rule b of the sequence rule (see Cahn *et al.*, 1966, p 406), according to which if an identity among the groups of a set leaves one group unique, the unique group is fiducial. The reason for the difference is that the sequence rule would define principal torsion angle in terms of a hydrogen atom whenever a single such atom formed part of the set; in the X-ray technique, nearly always used to establish structures of the type under discussion, hydrogen atoms are usually unobservable, and even at best not accurately locatable, so that the position of one used to define a principal torsion angle could only be established by calculation based on (perhaps unjustified) assumptions about the bond angles concerned. These considerations apply with even more force to case III, where one branch is nonexistent; the "phantom atom" of zero atomic number would be given highest priority because it is unique.

(ii) In case III the clockwise passage from CD^1 to CD^2 shall be by the shorter of the two possible routes.

2.2.3. If all three branches are identical, that giving the smallest positive or negative value of the principal torsion angle is normally[5] assigned the highest priority and the

---

[5] The qualification "normally" is added to avoid the need to renumber the branches if by chance the rule would demand this in consequence of a movement during refinement of a structure. In this or similar cases the symbolism should remain unchanged.

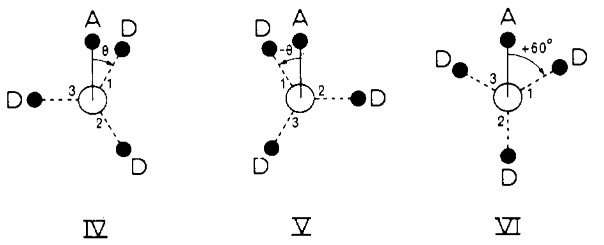


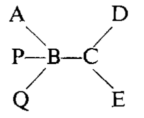
FIGURE 3: Tetrahedral configurations. Three identical branches: IV, general case, θ positive; V, general case, θ negative; VI, θ = 60°.

lowest number (1) (see Figure 3, IV, V); if two branches have torsion angles respectively +60 and −60°, the former is chosen (see Figure 3, VI). The others are numbered in a clockwise sense when viewed in the direction B → C.

Note. Rule 2.2.3 introduces a new principle, not invoked in 2.2.1 or 2.2.2, that the precedence depends on the conformation. This must necessarily be done since in this case the branches are distinguishable only in this respect. (The same applies to Rule 2.3.2.)

2.3. *Choice of torsion angle and numbering of branches (planar trigonal configurations).*

2.3.1. If, in a compound



such that B, C, D, and E are coplanar, or nearly so, the sequence rule gives the priorities A > P, Q and D > E, then the principal torsion angle is that measured by reference to atoms A–B–C–D as in Rule 1.6 above. The branches beginning at C are numbered

$$\underset{1}{\text{C—D}},\ \underset{2}{\text{C—E}}$$

2.3.2. If the branches are identical, that giving the smallest positive or negative value of the principal torsion angle is normally[5] assigned the highest priority and the lowest number (1); if the two branches have torsion angles respectively +90 and −90°, the former is chosen (see Figure 4).

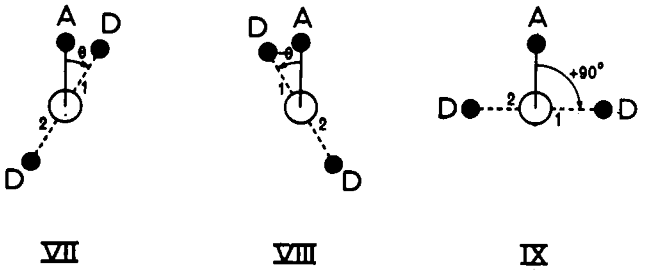


FIGURE 4: Planar trigonal configurations. Identical branches: VII, θ positive; VIII, θ negative; IX, θ = 90°.

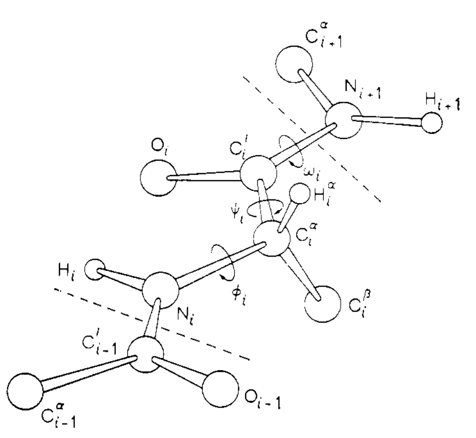


FIGURE 5: Perspective drawing of a section of polypeptide chain representing two peptide units. The limits of a *residue* are indicated by dashed lines, and recommended notations for atoms and torsion angles are indicated. The chain is shown in a fully extended conformation ($\phi_i = \psi_i = \omega_i = 180°$), and the residue illustrated is L.

Rule 3. The Main Chain (or Polypeptide Backbone)

3.1. *Designation of bonds.* Bonds between main-chain atoms are denoted by the symbols of the two atoms terminating them, *e.g.*, N_i–C_i^{α}, C_i^{α}–C_i, C_i–N_{i+1}, C_i–O_i, N_i–H_i. Abbreviated symbols should not be used. Bond lengths are written $b(C_i,N_{i+1})$, etc.

3.2. *Torsion angles.*

3.2.1. The principal torsion angle describing rotation about N–C^{α} is denoted by ϕ, that describing rotation about C^{α}–C is denoted by ψ, and that describing rotation about C–N is denoted by ω. The symbols ϕ_i, ψ_i, and ω_i are used to denote torsion angles of bonds within the *i*th residue in the case of ϕ and ψ and between the *i*th and $(i + 1)$th residues in the case of ω; specifically, ϕ_i refers to the torsion angle of the sequence of atoms $C_{i-1},N_i,C_i^{\alpha},C_i$; ψ_i to the sequence $N_i,C_i^{\alpha},C_i,N_{i+1}$; and ω_i to the sequence C_i^{α}, $C_i,N_{i+1},C_{i+1}^{\alpha}$ (see Figure 5). In accordance with Rules 1.6 and 2.1.1, these torsion angles are ascribed zero values for eclipsed conformations of the main-chain atoms N, C^{α}, and C, that is, for the so-called *cis* conformations (see Table I).

Notes. (i) This convention differs from that proposed by Edsall *et al.* (1966). The new designation of angles may be derived from the old by adding 180° to, or subtracting 180° from, the latter. (This statement is precisely correct only if the peptide bond is exactly planar, which is not generally the case in experimentally determined structures).

(ii) Owing to the partial double-bond character of CO$\overset{\cdots}{-}$NH, it is normally possible for ω to assume values only in the neighborhood of 0 or 180°. $\omega \sim 180°$ is the value which is generally found (*i.e.*, the *trans* conformation),

(iii) A "fully extended" polypeptide chain is characterized by $\phi = \psi = \omega = +180°$. The case of $\phi = \psi = 0°$ would involve the relations indicated in Table I.

(iv) Table II gives values of ϕ and ψ for various well-known regular structures. It is noteworthy that a right-handed α helix has negative torsion angles.

(v) Figure 6 is a typical conformational map (ϕ–ψ plot) using the rules enunciated above.

3.2.2. There may occasionally be a need to consider torsion angles differing from zero for the sequences of atoms

TABLE I: Main-Chain Torsion Angles for Various Conformations in Peptides of L-Amino Acids.[a,b]

ϕ (deg)	Rotation about N-C^α		ψ (deg)	Rotation about C^α-C	
0	C^α–C *trans*	to N-H	0	C^α–N *trans*	to C-O
+60	C^α–H *cis*		+60	C^α–R *cis*	
+120	C^α–R *trans*		+120	C^α–H *trans*	
+180	C^α–C *cis*		+180	C^α–N *cis*	
−120	C^α–H *trans*		−120	C^α–R *trans*	
−60	C^α–R *cis*		−60	C^α–H *cis*	

[a] *trans* to N_i–H_i is the same as *cis* to N_i–C_{i-1}; *trans* to C_i–O_i is the same as *cis* to C_i–N_{i+1} (see Figure 5). [b] For the description of D-amino acids, interchange C^α–H and C^α–R in the table.

TABLE II: Approximate Torsion Angles for Some Regular Structures.[a]

	ϕ (deg)	ψ (deg)	ω (deg)	Reference
Right-handed α helix (α-poly(L-alanine))	−57	−47	+180	Arnott and Dover (1967)
Left-handed α helix	+57	+47	+180	Arnott and Dover (1967)
Parallel-chain pleated sheet	−119	+113	+180	Schellman and Schellman (1964)
Antiparallel-chain pleated sheet (β-poly(L-alanine))	−139	+135	−178	Arnott *et al.* (1967)
Polyglycine II	−80	+150	+180	Ramachandran *et al.* (1966)
Collagen	−51, −76, −45	+ 153, +127, +148,	+180	Yonath and Traub (1969)
Poly(L-proline) I	−83	+158	0	Ramachandran and Sasisekharan (1968), calculated from Traub and Schmueli (1963)
Poly(L-proline) II	−78	+149	+180	Arnott and Dover (1968)

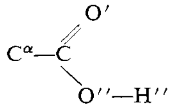
[a] For a fully extended chain $\phi = \psi = \omega = +180°$.

O=C—N—C^α and C^α—C—N—H, in cases where C=O or N—H lies out of the peptide plane. These angles may be represented υ^{O} and υ^{H} (Greek upsilon).

3.3. *Chain terminations.*

3.3.1. If the terminal amino group of the chain is protonated the three hydrogen atoms are denoted $H_1{}^1$, $H_1{}^2$, and $H_1{}^3$; the hydrogen atom giving the smallest (positive or negative) value of the principal torsion angle H–N–C^α–C is denoted $H_1{}^1$ and the others are numbered in a clockwise sense when viewed in the direction $C^\alpha \rightarrow$ N. The corresponding torsion angles are denoted $\phi_1{}^1$, $\phi_1{}^2$, and $\phi_1{}^3$. If the terminal amino group is not protonated the hydrogen atoms are denoted $H_1{}^1$ and $H_1{}^2$ in accordance with Rule 2.2.2, and the corresponding torsion angles, $\phi_1{}^1$ and $\phi_1{}^2$.

3.3.2. At the carboxyl terminus of the chain (i = T) the double-bonded oxygen is written as O′ and the other oxygen as O′′, thus



The torsion angles about the C^α–C bond are written $\psi_T{}^1$ and $\psi_T{}^2$ (or ψ1(T), ψ2(T)); the torsion angle about the C–O′′ bond, defining the orientation of the hydrogen atom of the hydroxyl group relative to C^α, is written $\theta_T{}^C$ (or θC(T)). If the terminal carboxyl group is ionized the oxygen atoms are denoted O′ and O′′, the precedence being determined by Rule 2.3.2, and the torsion angles are written as before.

Note. Instead of O′ and O′′ the alternative notations O^1 and O^2 may be used. ψ_T may be used instead of $\psi_T{}^1$, in conformity with the convention for the middle of the chain, so long as confusion does not arise.

3.3.3. *Substituted terminal groups.* Natural extensions of the above rules may be devised, *e.g.*

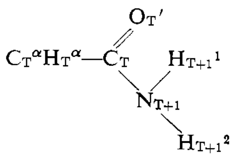
(i) *N-formyl group*

$$H_0\text{—}C_0O_0\text{—}N_1H_1\text{—}C_1{}^\alpha H_1{}^\alpha\text{—}\ \ldots\ldots$$

(ii) *N-acetyl group*

$$C_0(H_0{}^1,H_0{}^2,H_0{}^3)\text{—}C_0O_0\text{—}N_1H_1\text{—}C_1{}^\alpha H_1{}^\alpha\text{—}\ \ldots\ldots$$

(iii) *C-amido group*



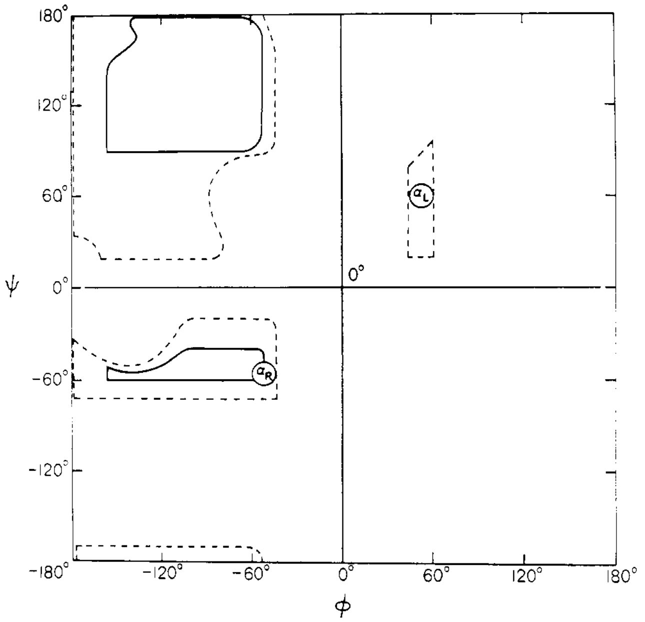


FIGURE 6: Typical conformational map (Ramachandran *et al.*, 1963) transposed into the standard conventions. Note: this diagram is identical with that of Edsall *et al.* (1966) except that the origin is now at the center, instead of at the lower left-hand corner. The solid lines enclose the freely allowed values of φ and ψ for an alanyl residue in a polypeptide; the dotted lines enclose "outer limit" values based on the shortest known van der Waals radii in related structures. Analogous diagrams for other residues, and for slightly different assumptions, are given by Ramachandran and Sasisekharan (1968; note that these authors used the earlier convention with the origin at the corner).

Rule 4. Side Chains

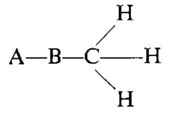
4.1. Atoms are lettered, or lettered and numbered, from C^{α}, and bonds are numbered from C^{α}, working outwards away from the main chain.

4.2. *Designation of atoms other than hydrogen.* Atoms other than hydrogen are designated in the usual way by Greek letters, β, γ, δ etc., *e.g.*, C_i^{β} (or Cβ(*i*)), N_i^{ζ} (or Nζ(*i*)).

Note. The notations for the amino acids normally occurring in proteins are given in Table III.

4.3. *Designation of branches.* If a side chain is branched, the branches are numbered 1 and 2, the order being determined (i) in cases where the branches are different, by application of Rule 2.2.1 or 2.3.1, (ii) in cases where two branches are identical (*e.g.*, in valine, phenylalanine), by the application of Rule 2.2.2 (valine) or 2.3.2 (phenylalanine). Nonhydrogen atoms in different branches are designated by the Greek letter indicating their degree of remoteness from C^{α} and by the number of their branch (see Rules 2.2 and 2.3); *e.g.*, in valine $C_i^{\gamma 1}$ and $C_i^{\gamma 2}$ (or Cγ1(*i*), Cγ2(*i*)). The branch number need not be indicated where no ambiguity results, *e.g.*, in threonine O^{γ} and C^{γ} instead of $O^{\gamma 1}$ and $C^{\gamma 2}$, in hydroxyproline O^{δ}, C^{δ} instead of $O^{\delta 1}$, $C^{\delta 2}$, and in histidine C^{δ}, N^{ϵ}, etc., instead of $C^{\delta 2}$, $N^{\epsilon 2}$. For asparagine or glutamine, in cases where nitrogen and oxygen in the amide group have not yet been distinguished, these atoms may be written $(NO)^{\delta 1}$, $(NO)^{\delta 2}$, or $(NO)^{\epsilon 1}$, $(NO)^{\epsilon 2}$, the indices 1 and 2 being determined by Rule 2.3.2.

4.4. *Designation of hydrogen atoms.* Hydrogen atoms are designated by the Greek letter and/or number of the atom to which they are attached, *e.g.*, in valine H_i^{β} (or Hβ(*i*)). Where three hydrogen atoms are attached to a single nonhydrogen atom, they are designated 1, 2, and 3; in the situation

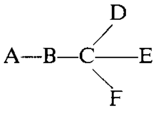


the hydrogen atom giving the smallest (positive or negative) value of the principal torsion angle is designated 1, and the others are numbered in a clockwise sense when viewed in the direction B → C (see Rule 2.2.3, which also covers the case where $\theta = \pm 60°$), *e.g.*, in valine $H_i^{\gamma 11}$, $H_i^{\gamma 12}$, $H_i^{\gamma 13}$ and $H_i^{\gamma 21}$, $H_i^{\gamma 22}$, $H_i^{\gamma 23}$ (or Hγ11(*i*), etc.). Where only two hydrogen atoms are present, they are designated in accordance with Rule 2.2.2, case I, for $-CH_2-R$ and case III for $-NH_2$.

4.5. *Designation of bonds and torsion angles* (see Table III).

4.5.1. Bonds are designated by means of the two atoms terminating them, *e.g.*, C_i^{α}–C_i^{β}, N_i^{ζ}–$H_i^{\zeta 2}$, or, if no ambiguity results, by the symbol of the first atom of the bond, *e.g.*, C_i^{α}, $C_i^{\gamma 1}$. In superscripts the bond may be denoted either by α; β; γ1; γ2, etc., or by 1; 2; 3,1; 3,2; etc. Bond lengths are denoted $b(C_i^{\alpha}, C_i^{\beta})$, bC_i^{α}, b_1^{1}, $b_i^{3,1}$, etc.

4.5.2. Torsion angles are denoted by χ and are specified by two (or three) superscripts, the first one (or two), in the situation



indicating the bond B–C about which the angle is measured, and the last indicating whether the angle is measured relative to D, E, or F. The principal torsion angle is defined by Rule 2.2.1, and if there is no ambiguity the last superscript may be omitted in referring to it.

Thus in valine $\chi_i^{2,1}$ and $\chi_i^{2,2}$ refer to the torsion angles specifying atoms $C_i^{\gamma 1}$ and $C_i^{\gamma 2}$; in leucine $\chi_i^{3,1,1}$, $\chi_i^{3,1,2}$, and $\chi_i^{3,1,3}$ refer to the torsion angles specifying the three hydrogen atoms attached to $C^{\delta 1}$. If there is no ambiguity the principal torsion angles may be referred to, in valine and leucine, as χ_i^{2} and $\chi_i^{3,1}$, respectively. Corresponding notations without subscripts are χ2,1(*i*), χ2(*i*); χ3,1,1(*i*), χ3,1(*i*).

Note. By the sequence rule, when $\chi_1 = 0$, C^{γ} (or $C^{\gamma 1}$) is in the eclipsed position relative to N.

Rule 5. Hydrogen Bonds

5.1. *Polarity of hydrogen bonds.* In specifying a hydrogen bond as directed from residue *i* to residue *k* (or from atom X_i to atom Y_k), the direction X–H to :Y is implied; *i.e.*, the atom covalently linked to the hydrogen atom is mentioned first.

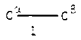
Example

In the α helix the N—H of residue *i* is hydrogen bonded

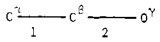
TABLE III: Symbols for Atoms and Bonds in the Side Chains of the Commonly Occurring L-Amino Acids.

(a) Unbranched side chains

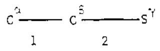
Alanine



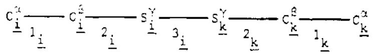
Serine



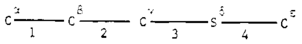
Cysteine



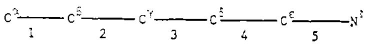
Cystine



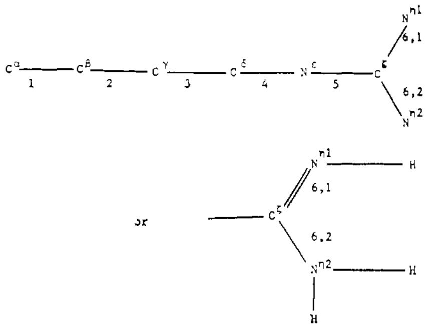
Methionine



Lysine

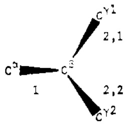


Arginine

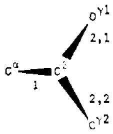


(b) Branched side chains

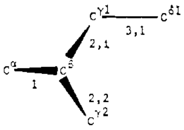
Valine



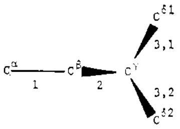
Threonine



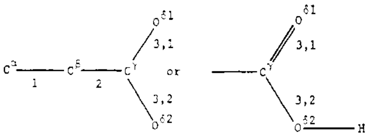
Isoleucine



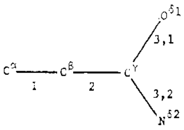
Leucine



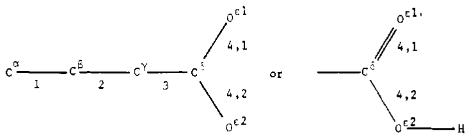
Aspartic acid



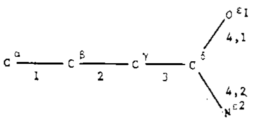
Asparagine



Glutamic acid

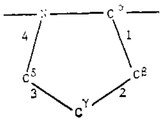


Glutamine

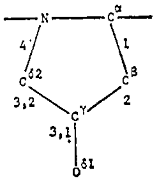


(c) Cyclic side chains

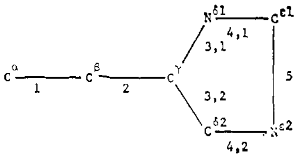
Proline



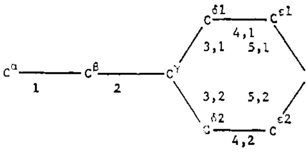
Hydroxyproline



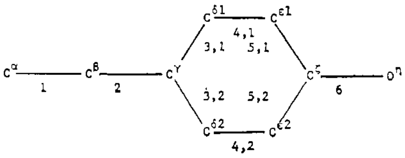
Histidine



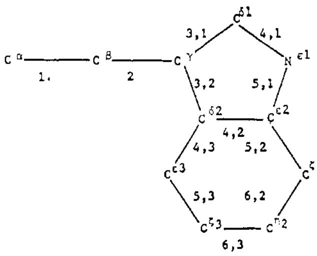
Phenylalanine



Tyrosine

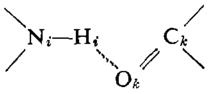


Tryptophan



to the O=C of residue ($i - 4$). Therefore, the α helix is described as having i to ($i - 4$), or ($5 - 1$), hydrogen bonding.

5.2. *Dimensions of hydrogen bonds.* Dimensions may be denoted by natural extensions of the nomenclature given above. For example, in



the following symbols might be used: $b(H_i, O_k)$, $\tau(N_i,H_i,O_k)$, $\tau(H_i,O_k,C_k)$, $\theta(H_i,O_k)$, $\theta_i(N,H)$, $\theta_k(C,O)$.

Rule 6. Helical Segments

A regular helix is strictly of infinite length, with all ϕ's identical and all ψ's identical. A helical *segment* of polypeptide chain may be defined either in terms of ϕ and ψ or in terms of symmetry and hydrogen-bond arrangement.

6.1. In the description of helices or helical segments the following symbols should be used: n = number of residues per 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).

6.2. *Definition in terms of ϕ and ψ.* Under this definition a helical segment is referred to as a (ϕ,ψ) helix; thus a right-handed α helix would be a ($-57°,-47°$) helix. The *first* and *last* residues of the helical segment are taken to be the first and last residues which have ϕ and ψ values equal to those defining the helix, within limits which should be defined in the context. No account is taken of hydrogen-bonding arrangements.

6.3. *Definition in terms of symmetry and hydrogen-bond arrangement.* A helix is referred to as an n_r helix, where n = number of residues per turn and r = number of atoms in ring formed by a hydrogen bond and the segment of main chain connecting its extremities. Thus an α helix would be 3.6_{13}. The *first* helical residue is taken as the first whose CO group is *regularly* bonded to NH along the helix (in the case of an α helix, to the NH of the fifth residue); the *last* helical residue is the last whose NH is *regularly* hydrogen bonded to CO along the helix (in the case of an α helix, to the CO of the residue last but four). Irregular hydrogen-bonding arrangements are not considered to form part of the helix.

Notes. (i) A helical segment defined by Rule 6.2 may, but need not necessarily, be two residues shorter than the same segment defined by Rule 6.3.

(ii) These rules prescribe no definitions for irregular helical segments.

Appendix

Recommendation A. Conformation and Configuration

There is at present no agreed definition of these two terms for general sterochemical usage.

In polypeptide chemistry the term "conformation" should be used, in conformity with current usage, to describe different spatial arrangements of atoms produced by rotation about covalent bonds; a change in conformation does not involve the breaking of chemical bonds (except hydrogen bonds) or changes in *chirality* (see Cahn *et al.*, 1966).

On the other hand in polypeptide chemistry the term "configuration" is currently used to describe spatial arrangements of atoms whose interconversion requires the formal breaking and making of covalent bonds (note: this usage takes no account of the breaking or making of hydrogen bonds). For a more extensive discussion see IUPAC Tentative Rules for the Nomenclature of Organic Chemistry, Section E, Fundamental Stereochemistry, *IUPAC Information Bulletin* **No. 35,** 71-80 (1969).

Recommendation B. Definitions of Primary, Secondary, Tertiary, and Quaternary Structure

These concepts, originally introduced by Linderstrøm-Lang (1952),[6] cannot be defined with precision, but the definitions given below may be helpful.

B.1. The *primary structure* of a segment of polypeptide chain or of a protein is the amino acid sequence of the polypeptide chain(s), without regard to spatial arrangement (apart from configuration at the α-carbon atom).

Note. This definition does not include the positions of disulfide bonds and is therefore not identical with "covalent structure."

B.2. The *secondary structure* of a segment of polypeptide chain is the local spatial arrangement of its main-chain atoms without regard to the conformation of its side chains or to its relationship with other segments.

B.3. The *tertiary structure* of a protein molecule, or of a subunit of a protein molecule, is the arrangement of all its atoms in space, without regard to its relationship with neighboring molecules or subunits.

B.4. The *quaternary structure* of a protein molecule is the arrangement of its subunits in space and the ensemble of its intersubunit contacts and interactions, without regard to the internal geometry of the subunits.

Note. A protein molecule not made up of at least potentially separable subunits (not connected by covalent bonds) possesses no quaternary structure. Examples of proteins without quaternary structure are ribonuclease (one chain) and chymotrypsin (three chains).

References

Arnott, S., and Dover, S. D. (1967), *J. Mol. Biol. 30,* 209.

Arnott, S., and Dover, S. D. (1968), *Acta Cryst. B24,* 599.

Arnott, S., Dover, S. D., and Elliott, A. (1967), *J. Mol. Biol. 30,* 201.

Cahn, R. S. (1964), *J. Chem. Educ. 41,* 116.

Cahn, R. S., and Ingold, C. K. (1951), *J. Chem. Soc.* 612.

Cahn, R. S., Ingold, C. K., and Prelog, V. (1956), *Experientia 12,* 81.

Cahn, R. S., Ingold, C. K., and Prelog, V. (1966), *Angew. Chem., Int. Ed. Engl. 5,* 385, 511; *Angew. Chem. 78,* 413.

Edsall, J. T., Flory, P. J., Kendrew, J. C., Liquori, A. M., Némethy, G., Ramachandran, G. N., and Scheraga, H. A. (1966), *J. Biol. Chem. 241,* 1004; *Biopolymers 4,* 130; *J. Mol. Biol. 15,* 339.

Eliel, E. L. (1962), Stereochemistry of Carbon Compounds,

---

[6] The use of the terms "primary, secondary, tertiary, and quaternary structure" has been criticized as being imprecise by Weltaufer (1961). He has proposed an alternative terminology.

McGraw-Hill, New York, N. Y., pp 92–94.
IUPAC Information Bulletin (1969), No. 35, 36–80.
Klyne, W., and Prelog, V. (1960), *Experientia 16*, 521.
Linderstrøm-Lang, K. (1952), Proteins and Enzymes, Lane Memorial Lectures, Stanford University Press, Stanford, Calif., p 54.
Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963), *J. Mol. Biol. 7*, 95.
Ramachandran, G. N., and Sasisekharan, V. (1968), *Advan. Protein Chem. 23*, 283.
Ramachandran, G. N., Sasisekharan, V., and Ramakrishnan, C. (1966), *Biochim. Biophys. Acta 112*, 168.
Schellman, J. A., and Schellman, C. (1964), *Proteins 2*, 1.
Traub, W., and Schmueli, U. (1963), *in* Aspects of Protein Structure, Ramachandran, G. N., Ed., Academic Press, London, p 81.
Wetlaufer, D. B. (1961), *Nature 190*, 1113.
Yonath, A., and Traub, W. (1969), *J. Mol. Biol. 43*, 461.

Influence of the Phosphodiester Linkage (3′–5′, 2′–5′, and 5′–5′) on the Conformation of Dinucleoside Monophosphate*

Norman S. Kondo,† Helen M. Holmes, Lloyd M. Stempel, and Paul O. P. Ts'o

ABSTRACT: The interaction of two adenines in dinucleoside monophosphates joined at different positions by a phosphodiester linkage (2′–5′, 3′–5′, and 5′–5′) have been studied by proton magnetic resonance, circular dichroism, and ultraviolet absorbance over the temperature range of 4–60°. The proton magnetic resonance spectra have been measured in dimethyl sulfoxide-d_6 and in D_2O. Similarly, the interactions of adenine and cytosine in $A_{2'}p_{5'}C$ and $A_{3'}p_{5'}C$ were also investigated. While these five dimers possess certain basic features in common in their conformations, the geometrical relationship between the two constitutive units as well as the temperature effect on this geometrical relationship is quite different for each dimer when examined in sufficient detail. These results indicate that the conformation of the dimers is strongly influenced by the position of the phosphodiester linkage. Studies on the interaction of adenine–adenine (or adenine–cytosine) having three types of geometrical relationships by three physicochemical methods offer a unique comparison among different types of dependence of these three methods (proton magnetic resonance, circular dichroism, and ultraviolet absorbance) on the geometrical relationships between two bases. From the analysis of the proton magnetic resonance data at 4° based on a most recent calculation on diamagnetic effects of ring-current magnetic anisotropy (Giessner-Prettre, C., and Pullman, B. (1970), *J. Theor. Biol. 27*, 87, 341), conformational models for these five dimers have been proposed. From these models, both the magnetic resonance data and the optical data can be correlated and unified qualitatively in accordance with our current understanding of the effect of base–base interaction on the optical properties in a homodimer (Glaubiger *et al.* (1968), *Biopolymers 6*, 409). The base planes in these five dimer models are parallel to each other with a vertical distance of about 3–4 Å. In this model the nucleosidyl units all have the *anti* conformation with respect to the sugar–base torsion angle, and the turn of the screw axis of the stack is right handed. The models for $A_{3'}p_{5'}A$, $A_{2'}p_{5'}A$, $A_{3'}p_{5'}C$, and $A_{2'}p_{5'}C$ are asymmetric, but the model of $A_{5'}p_{5'}A$ is constructed with a twofold axis of symmetry (C_{2v}) located at the phosphate group.

In our previous paper of this series, we reported the studies by proton magnetic resonance on 15 nucleosides and nucleotides and on 25 dinucleoside mono- and diphosphates (Ts'o *et al.*, 1969a). From these studies, a general conformational model for all the dimers is constructed. In this model the nucleosidyl units all have the *anti* conformation with respect to the sugar–base torsion angle, and the turn of the 3′–5′ screw axis of the stack is right handed. In this paper, we wish to study the influence of the phosphodiester linkage on the conformation of the dinucleoside monophosphates. For this reason, a comparative study was made on $A_{3'}p_{5'}A$, $A_{2'}p_{5'}A$, and $A_{5'}p_{5'}A$ as well as on $A_{3'}p_{5'}C$ and $A_{2'}p_{5'}C$. The base–base interactions in these five dimers have been investigated concurrently by proton magnetic resonance, circular dichroism, and ultraviolet absorbance over the temperature range of 4–60°. The proton magnetic resonance spectra have been measured in dimethyl sulfoxide-d_6 and in D_2O. The results indicate that the geometric relationships between the two bases in these isomeric dimers are strongly influenced by the position of the phosphodiester linkage they possess. In applying three physicochemical methods to examine the adenine–adenine (or adenine–cytosine) interaction in dimers having different geometrical relationships, we are able to compare different types of depen-

---

* From the Department of Radiological Sciences, The Johns Hopkins University, Baltimore, Maryland. *Received March 12, 1970*. This work was supported in part by a grant from the National Science Foundation (GB-8500) and a grant from the National Institutes of Health (GM-16066-02). This is paper no. 2 in a series entitled: Studies of the Conformation and Interaction of Dinucleoside Mono- and Diphosphates.

† U. S. Public Health Service Postdoctoral Fellow 1967–1969.