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. 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: \( \alpha, A; \beta, B; \gamma, G; \delta, D; \epsilon, E; \zeta, Z; \eta, H; \tau \)
\( T; \upsilon, U; \phi, F; \chi, X; \psi, Q; \omega, W. \)

Rule 1. General Principles of Notation

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

\[ \text{C}^\alpha, \text{N}^\beta, \text{C}^\gamma, \text{O} \]

where confusion might arise the following additional symbolism may be used

\[ \text{C}^\alpha, \text{N}^\beta, \text{C}^\gamma, \text{O} \]

1.2. Amino acid residues, \(-\text{NH}–\text{CHR}–\text{CO}–\), are numbered sequentially from the amino-terminal to the carboxyl-terminal end of the chain, the residue number being denoted \( i \).

Example

\( C^\alpha \) of the \( i \)th residue is written \( C^i_{\alpha} \) or \( C_{\alpha(i)} \)

1.3. Peptide units. For some purposes it is more convenient to group together the atoms \(-\text{CHR}–\text{CO}–\text{NH}–\). 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 \( N^\alpha_i \) and \( H^\alpha_i \) 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 \( N^\alpha_{i+1} = N_{i+1} \)).
Example

\[ \text{residue notation} \quad N_1^* \quad C_3^* \quad C_2 \quad N_2^* \]

\[ \text{peptide unit notation} \quad N_1 \quad C_3 \quad C_2 \quad N_2 \]

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

(ii) Whether "residues" or "peptide units" are being used, \( \phi \) and \( \psi \) always refer to torsion angles about bonds of the same \( C \).

1.4. Bond lengths. If a bond \( A-B \) be denoted \( A_1-B_1 \) or \( A_i \)
(see Rules 3.1, 4.5), the bond length is written \( b(A_i,B_j) \)
(or \( b(A_i,B_j) \) or \( b_iA \) (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 \( l \) 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

\[ \begin{array}{c}
A_i \\
B_j \\
C_k
\end{array} \]

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

1.6. Torsion angles. If a system of four atoms

\[ \begin{array}{c}
A \\
B \\
C \\
D
\end{array} \]

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 \( \theta(B_j) \). In the eclipsed conformation
in which the projections of \( A-B \) and \( C-D \) coincide, \( \theta \) is
given the value 0° (symplanar 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 \( B \rightarrow C \) (or \( C \rightarrow B \)), the bond to the front atom
\( (D) \) requires rotation to the right or to the left, respectively,
in order that it may eclipse the bond to the rear atom \( D \);
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 \(-180 < \theta \leq
+180^\circ\), 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 \( \phi \), \( \psi \), \( \omega \), \( \psi \), and \( \chi \)
(see Rules 3.2, 4.5.2). In the above \( \theta \) 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 preced-
ce 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 applica-
tion 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 \( C \) over the side chain
and at \( C \) over \( C \).

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

\[ \text{Cl} \]
\[ \text{In} \]
\[ \text{H} \]

In \( \text{Br-C-CH}_3 \) the order of priority is \( \text{Br, Cl, CH}_3, \text{H.} \)

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) \( \text{CH}_3\text{CH}_2-\text{C-CH}_3 \)

\[ \begin{array}{c}
\text{Cl} \\
\text{H}
\end{array} \]

\(^{2}\) 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 Elie1
(1962).

\(^{3}\) Other local systems are available analogously for steroids, carbo-
hydrates, and cyclitols, where the sequence rule is applied when the
local system does not suffice.
1. In general, the order is Cl, (CH₂CH₃), CH₃, H. (CH₃CH₂) takes precedence over C₃H₂ because C₃ is bonded to C, H, H and C₂ to H, H, H).

(ii) In

\[
\begin{align*}
\text{CH₂Cl} \\
\text{HO-} & \text{C-CH₂OH} \\
\text{H}
\end{align*}
\]

the order is OH, CH₂Cl, CH₂OH, H.

(iii) In

\[
\begin{align*}
\text{OH} \\
\text{CH₃CH₂-} & \text{C-CH(CH₂)₂} \\
\text{H}
\end{align*}
\]

the order is OH, CH(CH₃)₂, CH₂CH₃, H.

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

\[
\begin{align*}
\text{(O)} & \text{(C)}
\end{align*}
\]

Example

In CH₂CO-OH the order is =O, -OH, CH₃.

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

Example

In

\[
\begin{align*}
\text{D} \\
\text{Br-} & \text{C-CH₃} \\
\text{H}
\end{align*}
\]

the order is Br, CH₃, 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).

2.2.1. If, in a compound

\[
\begin{align*}
\text{A} & \text{D} \\
\text{P-B-C} & \text{E} \\
\text{Q} & \text{F}
\end{align*}
\]

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

\[
\begin{align*}
\text{C-D}, & \text{ C-E}, \text{ and C-F}
\end{align*}
\]

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 = C. 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 CD¹ to CD² 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 assigned the highest priority and the principal torsion angle is calculated.

\[ \text{Note. This rule is to be used only if the two previous rules do not give a decision.} \]

\[ \text{2.2. Choice of torsion angle and numbering of branches (tetrahedral configurations).} \]

\[ \text{2.2.1. If, in a compound} \]

\[ \text{2.2.2. If two branches are identical, and the third is different} \]

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

\[ \text{Note. This rule is to be used only if the two previous rules do not give a decision.} \]
lowest number (I) (see Figure 3, IV, V); if two branches have torsion angles respectively $+60$ and $-60^\circ$, the former is chosen (see Figure 3, VI). The others are numbered in a clockwise sense when viewed in the direction $B \rightarrow C$.

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

```
A
P\rightarrow B\rightarrow C\rightarrow D
Q
```

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

$C$-$D$, $C$-$E$

2.3.2. If the branches are identical, that giving the smallest positive or negative value of the principal torsion angle is normally assigned the highest priority and the lowest number (I); if the two branches have torsion angles respectively $+90$ and $-90^\circ$, the former is chosen (see Figure 4).

Note. (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 $180^\circ$ to, or subtracting $180^\circ$ 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-NH, it is normally possible for $\omega$ to assume values only in the neighborhood of $0$ or $180^\circ$. $\omega \sim 180^\circ$ is the value which is generally found (i.e., the trans conformation).

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

(iv) Table II gives values of $\phi$ and $\psi$ for various well-known regular structures. It is noteworthy that a right-handed $\alpha$ helix has negative torsion angles.

(v) Figure 6 is a typical conformational map ($\phi$-$\psi$ 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

$$C_{\alpha}N^{i+1}C^{i+1}C_{\alpha}C_{\alpha}$$
TABLE I: Main-Chain Torsion Angles for Various Conformations in Peptides of L-Amino Acids.\textsuperscript{a,b}

| \( \phi \) (deg) | Rotation about N–C\(^\alpha\) | \( \psi \) (deg) | Rotation about C\(^\alpha\)–C | \\
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>C(^\alpha)–C trans</td>
<td>0</td>
<td>C(^\alpha)–N trans</td>
</tr>
<tr>
<td>+60</td>
<td>C(^\alpha)–H cis</td>
<td>+60</td>
<td>C(^\alpha)–K cis</td>
</tr>
<tr>
<td>+120</td>
<td>C(^\alpha)–R trans</td>
<td>+120</td>
<td>C(^\alpha)–H trans</td>
</tr>
<tr>
<td>180</td>
<td>C(^\alpha)–C cis</td>
<td>C–H cis</td>
<td>C(^\alpha)–N cis</td>
</tr>
<tr>
<td>-120</td>
<td>C(^\alpha)–H trans</td>
<td>-120</td>
<td>C(^\alpha)–R trans</td>
</tr>
<tr>
<td>-60</td>
<td>C(^\alpha)–R cis</td>
<td>-60</td>
<td>C–H cis</td>
</tr>
</tbody>
</table>

\( \text{a} \) \( \text{trans to N}_i\)\(–H_i \) is the same as \( \text{cis to N}_i\)\(–C_{i-1} \); \( \text{trans to C}_{i}–O_{i} \) is the same as \( \text{cis to C}_{i}\)\(–N_{i+1} \) (see Figure 5). \( \text{b} \) For the description of D-amino acids, interchange \( \text{C}^\alpha–\text{H} \) and \( \text{C}^\alpha–\text{R} \) in the table.

TABLE II: Approximate Torsion Angles for Some Regular Structures.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Structure</th>
<th>( \phi ) (deg)</th>
<th>( \psi ) (deg)</th>
<th>( \omega ) (deg)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right-handed ( \alpha ) helix (( \alpha )-poly(L-alanine))</td>
<td>-57</td>
<td>-47</td>
<td>+180</td>
<td>Arnott and Dover (1967)</td>
</tr>
<tr>
<td>Left-handed ( \alpha ) helix</td>
<td>+57</td>
<td>+47</td>
<td>+180</td>
<td>Arnott and Dover (1967)</td>
</tr>
<tr>
<td>Parallel-chain pleated sheet</td>
<td>-119</td>
<td>+113</td>
<td>+180</td>
<td>Schellman and Schellman (1964)</td>
</tr>
<tr>
<td>Antiparallel-chain pleated sheet (( \beta )-poly(L-alanine))</td>
<td>-139</td>
<td>+135</td>
<td>-178</td>
<td>Arnott et al. (1967)</td>
</tr>
<tr>
<td>Polyglycine II</td>
<td>-80</td>
<td>+150</td>
<td>+180</td>
<td>Ramachandran et al. (1966)</td>
</tr>
<tr>
<td>Poly(( l )-proline) I</td>
<td>-83</td>
<td>+158</td>
<td>0</td>
<td>Ramachandran and Sasishekar (1968), calculated from Traub and Schmüll (1963)</td>
</tr>
<tr>
<td>Poly(( l )-proline) II</td>
<td>-78</td>
<td>+149</td>
<td>+180</td>
<td>Arnott and Dover (1968)</td>
</tr>
</tbody>
</table>

\( \text{a} \) For a fully extended chain \( \phi = \psi = \omega = +180^\circ \).

O–C–N–C\(^\alpha\) and C\(^\alpha\)–C–N–H, in cases where C=O or N–H lies out of the peptide plane. These angles may be represented \( \psi^0 \) and \( \psi^\beta \) (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\(_1\), H\(_2\), and H\(_3\); the hydrogen atom giving the smallest (positive or negative) value of the principal torsion angle H–N–C\(^\alpha\)–C is denoted H\(_1\); and the others are numbered in a clockwise sense when viewed in the direction C\(^\alpha\)→N. The corresponding torsion angles are denoted \( \phi_1, \phi_2, \text{ and } \phi_3 \). If the terminal amino group is not protonated the hydrogen atoms are denoted H\(_1\) and H\(_2\) in accordance with Rule 2.2.2, and the corresponding torsion angles, \( \phi_1, \phi_2, \text{ and } \phi_1, \phi_2 \).

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

\[ \text{C}^\alpha–\text{H} \quad O' \quad O''–\text{H}'' \]

The torsion angles about the C\(^\alpha\)–C bond are written \( \psi_1 \) and \( \psi_2 \) (or \( \psi_1(T), \psi_2(T) \)); the torsion angle about the C–O'' bond, defining the orientation of the hydrogen atom of the hydroxyl group relative to C\(^\alpha\), is written \( \phi_0^e \) (or \( \theta(T) \)). 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. \( \phi_1 \) may be used instead of \( \psi_4 \), 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) \textit{N-formyl group}

\[ \text{H}_6–\text{C} = \text{O}_6–\text{N}_6\text{H}_5–\text{C}^\circ\text{H}_6^\circ– \ldots \]

(ii) \textit{N-acetyl group}

\[ \text{C}_6(\text{H}_6^\circ,\text{H}_6^1,\text{H}_6^2)–\text{C} = \text{O}_6–\text{N}_6\text{H}_5–\text{C}^\circ\text{H}_6^\circ– \ldots \]

(iii) \textit{C-amido group}

\[ \text{C}^\circ\text{H}_7^\circ–\text{C}_7 \quad \text{H}_{7+1} \quad \text{H}_{7+2} \]
Rule 4. Side Chains

4.1. Atoms are lettered, or lettered and numbered, from C', and bonds are numbered from C', 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., Cγ (or Cβ(i)), Nβ (or Nβ(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' and by the number of their branch (see Rules 2.2 and 2.3); e.g., in valine Cγ1 and Cγ2 (or Cβ1(i), Cβ2(i)). The branch number need not be indicated where no ambiguity results, e.g., in threonine Oγ and Cγ instead of Oγ1 and Cγ2; in hydroxyproline Oγ, Cγ instead of Oγ1, Cγ2; and in histidine Cγ, Nγ etc., instead of Cγ2, Nγ. 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)γ1, (NO)γ2, or (NO)γ1, (NO)γ2, 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 Hβ (or Hβ(i)). Where three hydrogen atoms are attached to a single non-hydrogen 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 θ = ±60°), e.g., in valine Hγ1, Hγ1, Hγ1, Hγ1 (or Hγ1(i), etc.). Where only two hydrogen atoms are present, they are designated in accordance with Rule 2.2.2, case I, for –CH2-R and case III for –NH2.

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., Cγ–Cβ, Nβ–Hγ3, or, if no ambiguity results, by the symbol of the first atom of the bond, e.g., Cγ, Cβ. In superscripts the bond may be denoted either by α; β; γ1; γ2, etc., or by 1; 2; 3, 1; 2; etc. Bond lengths are denoted b(Cγ, Cβ), b(Cβ, b1, b2, b1, b2, 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 χγ1 and χγ2 refer to the torsion angles specifying atoms Cγ1 and Cγ2; in leucine χγ1, χγ2, and χγ3, χγ3, refer to the torsion angles specifying the three hydrogen atoms attached to Cγ. If there is no ambiguity the principal torsion angles may be referred to, in valine and leucine, as χγ1 and χγ3, respectively. Corresponding notations without subscripts are χ1(0), χ2(0); χ3,1,1(0), χ3,1(0).

Note. By the sequence rule, when χγ1 = 0, Cγ1 (or Cγ) 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 Xi to atom Yγ), the direction X–H to Yγ 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 with the oxygen of residue i+4.
TABLE III: Symbols for Atoms and Bonds in the Side Chains of the Commonly Occurring L-Amino Acids.

(a) Unbranched side chains

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2$</td>
</tr>
<tr>
<td>Serine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2$</td>
</tr>
<tr>
<td>Cysteine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Cystine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Methionine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Lysine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}_1$</td>
</tr>
</tbody>
</table>

(b) Branched side chains

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Threonine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Leucine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Asparagine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
</tbody>
</table>

(c) Cyclic side chains

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenylalanine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Histidine</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
<tr>
<td>Hydroxyproline</td>
<td>$\mathrm{C}_3\mathrm{H}_7\mathrm{O}_2\mathrm{N}_2\mathrm{S}$</td>
</tr>
</tbody>
</table>
to the O=C of residue \((i - 4)\). Therefore, the \(\alpha\) 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

\[ \begin{align*}
\text{N} & \quad \text{H} \\
\text{O} & \quad \text{C} \\
\end{align*} \]

the following symbols might be used: \(b(H_{i}, O_{i})\), \(\tau(N_{i}, H_{i}, O_{i})\), \(\pi(H_{i}, O_{j}, C_{j})\), \(\theta(H_{i}, O_{i})\), \(\theta(N, H)\), \(\theta(C, O)\).

**Rule 6. Helical Segments**

A regular helix is strictly of infinite length, with all \(\phi\)'s identical and all \(\psi\)'s identical. A helical segment of polypeptide chain may be defined either in terms of \(\phi\) and \(\psi\) 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); \(t\) = \(360\degree / n\) = unit twist (angle of rotation per residue about the helix axis).

6.2. **Definition in terms of \(\phi\) and \(\psi\).** Under this definition a helical segment is referred to as a \((\phi, \psi)\) helix; thus a right-handed \(\alpha\) helix would be \((-57\degree, -47\degree)\) helix. The first and last residues of the helical segment are taken to be the first and last residues which have \(\phi\) and \(\psi\) 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 \(n,\) 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 \(\alpha\) helix would be \(3,6\). 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 \(\alpha\) helix, to the NH of the fifth residue); the last helical residue is the last whose NH is regularly hydrobnded to CO along the helix (in the case of an \(\alpha\) 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 stereochemical 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 Linderstrom-Lang (1952), 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 \(\alpha\)-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**

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

*Acknowledgment*—We are indebted to the Editors of *Biochemistry* for permission to reproduce this document that appeared earlier (*Biochemistry*, 9, 3471 (1970)).
Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains: Tentative Rules (1969)
IUPAC-IUB Commission on Biochemical Nomenclature


Access the most updated version of this article at http://www.jbc.org/content/245/24/6489.citation

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/245/24/6489.citation.full.html#ref-list-1