Graphical Representation Standards for Organic Chemistry Script: IUPAC Recommendation 2008

Dhatrak N. R
Department of Chemistry
Savitribai Phule Pune University
Pune 411 007, India.
Email:
ajit@chem.unipune.ac.in

In this article, we will acquaint ourselves with IUPAC representation standards in organic chemistry. Most of the chemistry is expressed through graphical representations, and therefore, molecular structures shall be sketched carefully so that the drawing conveys exact and desired meaning. There are specific rules and regulations governing the graphical representations in chemistry. Less literature is available for this purpose, but the importance of the topic can be well understood after knowing that the division of chemical nomenclature and structure representation of IUPAC has published the *Graphical Representation Standards for Chemical Structures* in 2008 comprising 134 pages. The present article focuses on the etiquettes of structure drawings along with a summary in brief about the IUPAC 2008 recommendations in perspectives of organic chemistry.

Introduction

Chemical structures are indeed the language of chemistry. The purpose of the chemical drawing is to convey the exact connectivity of the atoms in two-dimensional forms. The structures drawn must be correct and devoid of any ambiguity. It is anticipated that the structure sketched shall convey a clear and unique meaning. There are certain rules and regulations suggested by IUPAC for the correctness and universal acceptability of structures of molecules. IUPAC has published these rules and norms as IUPAC Recommendations 2008 entitled *Graphical Representation Standards for Chemical Structure Diagrams* [1].

Keywords
IUPAC recommendations 2008, bond-line diagrams, organic chemistry script, molecular diagrams, representation standards.

*Vol.25, No.8, DOI: https://doi.org/10.1007/s12045-020-1029-4

In general, there are two methods in practice for drawing organic molecules.

In general, there are two methods in practice for drawing organic molecules. The very first method involves the depiction of each atom in the molecule and its connectivity in detail or to write condensed molecular formulas. Chemistry students practice this method in their early stages of learning chemistry. This method suffers many problems; like writing every atom and bond makes the molecule challenging to write and understand. Also, the structure becomes more alpha-numeric.

There occurred several modifications in the organic chemistry script with the growing need of writing and representation. An improvised method of sketching organic molecules involves the omission of hydrogen in the aliphatic part and writing the carbon skeleton along with functional groups. This method, in general, referred to as the 'bond-line structure representation'. Drawing structures in bond-line forms give a pictorial touch to organic structures, and the script becomes easy to write and understand. The method soon became famous because of its simplicity and is practiced widely, as it allows fast writing and easy understanding. The method comes handy not only while writing molecules but also facilitates tracing mechanisms of organic reactions. Computer software like ISIS-Draw, ChemWind, ChemDraw, etc., enabled chemists to draw organic molecules with the help of computers and incorporate them in their publications. Below are two structures of butanoic acid and adipic acid drawn using early and advanced methods (*Table* 1).

The present article is a concise and simplified extract of the IUPAC recommendation 2008 from the organic chemistry point of view. Few important recommendations are explained below.

Allowed Font, Colors, Line Thickness

When structures are generated with computer software, IUPAC has mentioned (GR-0.5, [1]) the rules for line width and font. The lines depicting bonds in any structures shall not be thinner than 0.5 points. The sketches generated using software shall use standard fonts and visible alphabet sizes. Roman fonts are al-

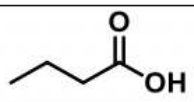
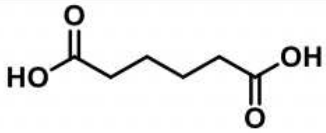
Early method	Advanced method (Bond-Line Structures)
$H_3C\text{-}CH_2\text{-}CH_2\text{-}COOH$	
$HOOC\text{-}CH_2\text{-}CH_2\text{-}CH_2\text{-}CH_2\text{-}COOH$	

Table 1: Structures of butanoic acid and adipic acid drawn using early and advanced methods.

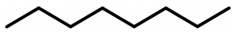


Figure 1. Structure of octane.

lowed, but the font used in sketches shall match the associated text (GR-0.3). Colors in sketches shall be avoided except when an emphasis is desired. A combined use of red and green colors as contrasting colors in a diagram shall be avoided keeping in mind that 10% of readers are color blind (GR-0.5).

A combined use of red and green colors as contrasting colors in a diagram shall be avoided keeping in mind that 10% of readers are color blind.

The Principle of Sketching Molecule

In the bond-line method, the molecule is written as a zig-zag line. Each terminus (end) and the bent on the line represents a carbon along with the required number of hydrogen atoms to satisfy valency. Carbons and hydrogens are omitted while writing the structures as it simplifies the structure and improves the readability. In a nutshell, the carbon sigma skeleton is represented in the form of a zig-zag line (*Figure* 1).

Small molecules like methane and ethane, instead of writing as a dot or dash, are written as CH_4 and C_2H_5. Ethylene molecule also, for the same reason, written as $H_2C{=}CH_2$ to avoid a clash with 'equal to sign'. One can practice bond-line structures propane onwards.

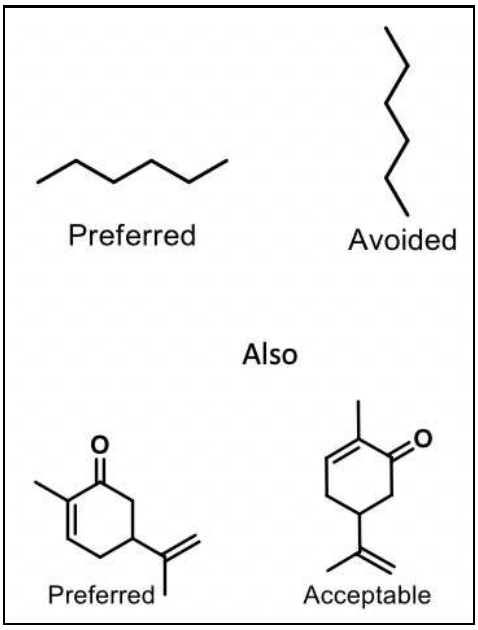


Figure 2. A structure is written horizontally than to write it vertically.

The Rule for Bond Length

All the bonds in a molecule are generally written with identical bond lengths (GR-1.1) although the bond length varies depending upon the bond type. The π-bonds (*double* or *triple*) bonds are written relatively shorter than the single bonds to demark them as π-bonds. The practice facilitates tracing of mechanisms.

The Orientation of the Structure

A horizontal structure should be preferred over vertical as horizontal structures fall in line with the trend of reading; also, they occupy much less space on the paper. For example, horizontal orientation is preferred over the vertical while writing n-hexane molecule, although both convey the same meaning (*Figure* 2).

The Rule for Bond Angle

In reality, the bond angle in the tetrahedral (sp^3) carbon is $109^o.28'$; however, while writing bond-line structures, the bond angle is

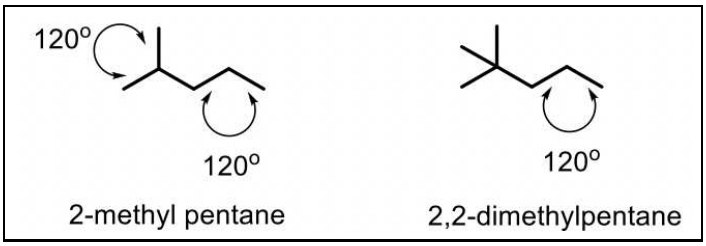


Figure 3. Bond angle in tetrahedral carbon.

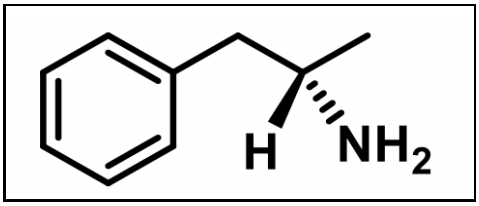


Figure 4. Structure of amphetamine molecule written with the help of stereochemical bonds.

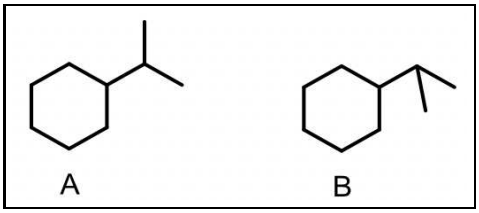


Figure 5. Structure 'A' of isopropyl cyclohexane is preferred over 'B'.

generally maintained at 120°. This furnishes space for the fourth bond, and the carbon appears tetrahedral after writing all the four bonds (*Figure* 3).

This rule (120° bond angle) also facilitates the accommodation of stereochemical bonds, and it becomes possible to write the stereochemical structures. E.g., the amphetamine molecule is written with the help of stereochemical bonds (*Figure* 4).

As long as possible, the rule of bond angle (120°) is obeyed to impart symmetric attributes to the molecule. While writing isopropyl cyclohexane, structure 'A' is preferred over 'B' as it is symmetrically written (*Figure* 5).

Customarily, the bond angle of 180° is maintained while sketch-

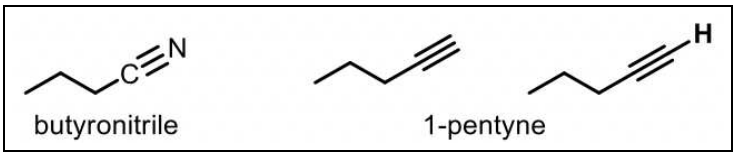


Figure 6. Two sp-hybrid atoms, along with two adjacent atoms, are written in a straight line.

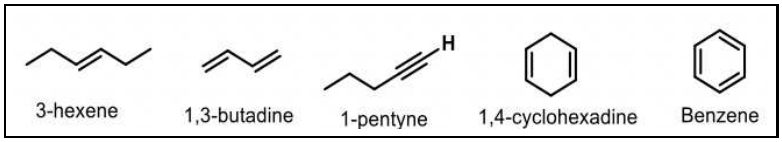


Figure 7. The method to write the 'σ' and the 'π' frameworks.

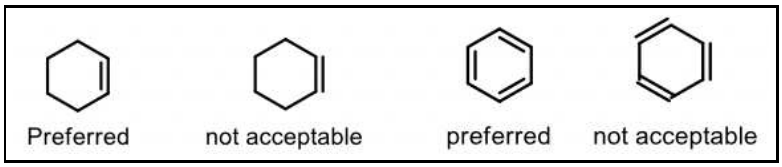


Figure 8. Double bonds should have an offset towards the center of the ring.

ing alkynes, and nitriles (cyanides). The two sp-hybrid atoms, along with two adjacent atoms, are written in a straight line (*Figure* 6).

Sigma and π-frameworks

Molecules may have different functional groups and thus have the sigma and π-bonds. All the sigma bonds of a molecule are written as a continuous framework. The π-bonds, however, are written as isolated dashes and relatively shorter in length as compared to sigma bonds (*Figure* 7).

All the double bonds associated with a ring shall be written inside the ring and slightly offset towards the center of the ring.

This helps in distinguishing the π-bonds from the sigma bonds especially while writing mechanisms where it is essential to show the delocalization of π-bonds. Double bonds in dienes and polyenes shall be written on the same side. GR-1.10 has given clear convention about the sidedness of the double bonds. The double bonds associated with rings shall offset towards the center of the ring (*Figure* 8).

Both the sketches in *Figure* 9 are strongly discouraged (GR-1.10)

Figure 9. Unusual sidedness of the π-bonds.

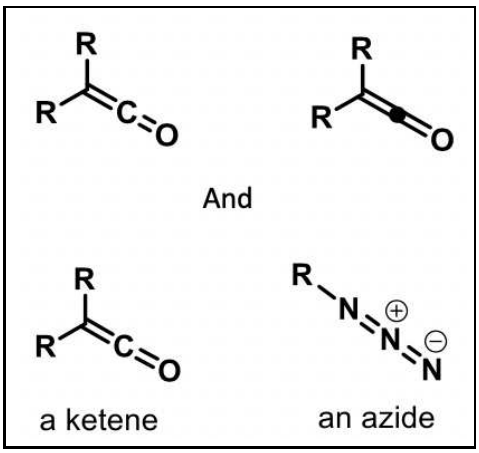


Figure 10. Depiction of the allene type carbon (or hetero-atom).

because of the unusual sidedness of the π-bonds (*Figure* 9).

The central carbon (or hetero-atom) is essentially depicted while writing cumulated double bonds (GR-2.1.2) (*Figure* 10).

Implicit and Explicit Hydrogen

Two types of hydrogen may be associated with a molecular sketch; hydrogen those accompany the carbons in aliphatic or aromatic skeletons and those which accompany the heteroatoms. The hydrogens which are not written in general but are supposed to be there in an appropriate number are called the implicit hydrogen (*hidden*) and those must be written are called the explicit hydrogen (*open*) (*Figure* 11).

For example, in *Figure* 11, hydrogen those accompany oxygen and nitrogen, are essentially written and are referred to as explicit hydrogen, while those accompanying the carbon in aliphatic chains are not written and presumed to be there in an appropriate number to satisfy the carbon valency, and are called as implicit

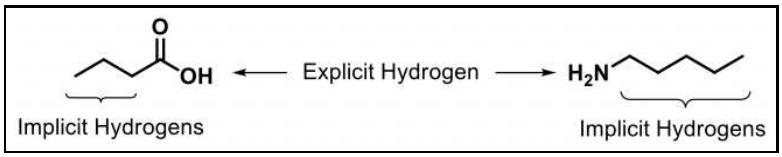


Figure 11. Implicit and explicit hydrogen.

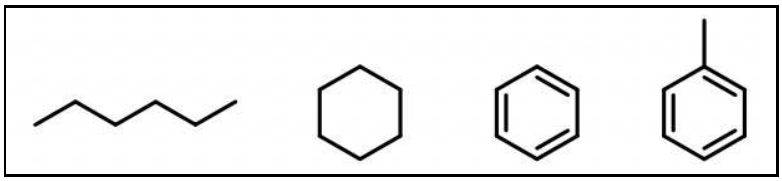


Figure 12. Implicit hydrogen.

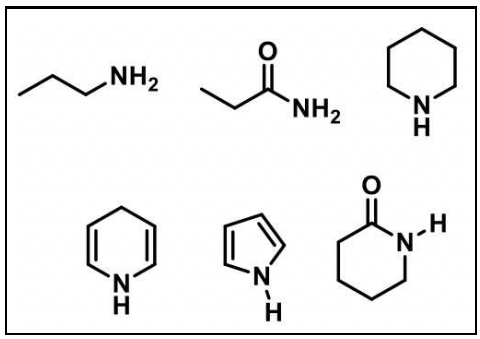


Figure 13. Explicit hydrogen in nitrogen compounds.

hydrogen. The omission of implicit hydrogen improves the simplicity and readability of the structure, and the functional groups become prominently visible.

Optional hydrogens are omitted in the molecular sketches, but they are supposed to be there in the required number and are called implicit hydrogen (*Figure* 12).

Hydrogens accompanying the heteroatoms are treated as explicit and are always written. For example, the hydrogen shown with nitrogen in the following compounds are explicit and are written without fail (*Figure* 13).

Hydrogens accompanying the oxygen atom in alcohols, acids, peroxides, and even in the water molecule are explicit and must be written (*Figure* 14).

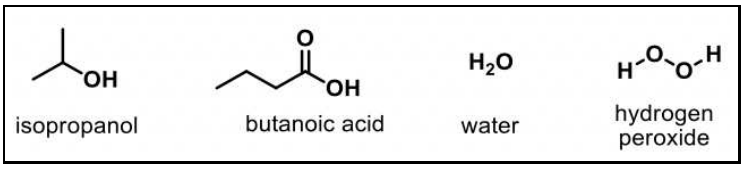


Figure 14. Explicit hydrogen accompanying the oxygen atom in alcohols, acids, peroxides.

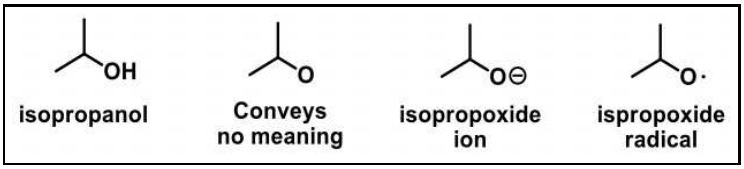


Figure 15. Explicit hydrogen with the hetero atoms must be written.

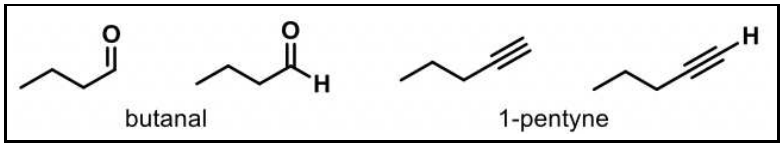


Figure 16. Representation of functional hydrogen in an aldehyde and an alkyne.

The omission of explicit hydrogen leads to confusion (GR-2.1.1). For example, the omission of hydrogen accompanying the oxygen atom in isopropanol leads to a deviated meaning or no meaning. Explicit hydrogens are an integral part of the structure. These can only be replaced with cationic, anionic charges, or free radical notations (*Figure* 15).

Hydrogen in few molecules may be implicit or explicit depending on its utility as a functional group. For example, omission of the aldehyde hydrogen in aldehydes and terminal alkyne proton from alkynes leads to no alteration of meaning. These hydrogens shall not be excluded if these are the part of some chemical reaction under discussion. For instance, aldehyde hydrogen can be omitted while writing an aldol condensation but not while writing Cannizzaro reaction (*as it is functional hydrogen*). Terminal alkyne proton must be shown while sketching anion formation from terminal alkynes (*as the proton constitutes a function group*) (*Figure* 16).

Figure 17. Stereo-descriptors to show geometrical isomerism.

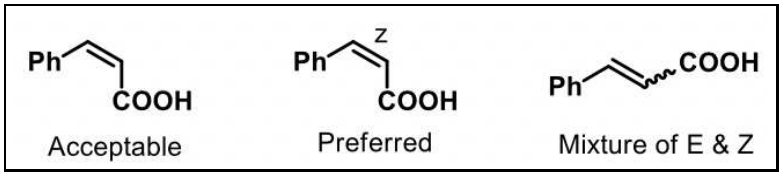


Figure 18. Wavy bond to depict a mixture of *cis* and *trans* isomers.

The Depiction of Geometrical Isomerism

All intermittent double bonds (except terminal) spontaneously appear as trans only as zig-zag writing is the essence of the sketch. Hence its good practice to write *E* or *Z* along with double bond. Mixtures can be depicted using a wavy bond.

All intermittent double bonds (*except terminal*) spontaneously appear as trans only as zig-zag writing is the essence of the sketch. Hence, a double bond shall not be taken as trans unless it is specified as trans. Therefore, it is suggested to write a stereo-descriptor (*cis*, *trans* or *E*, *Z*) with the double bond (*Figure* 17).

On the contrary, writing a cis bond is a deliberate action, and a stereo-descriptor may not be required. However, stereo-descriptor shall be written to overcome any ambiguity (GR-11.2). A wavy bond may be used to depict a mixture of *cis* and *trans* isomers (*Figure* 18).

While writing stereoisomers, it is good practice to keep the orientation of isomers same and to depict the stereocenters by altering the stereochemical bond.

Stereochemical Bonds

Generally, two wedge-shaped stereochemical bonds are used while sketching molecules with stereochemistry. Bonds to the atom above the plane of a drawing are depicted by a bold wedge ( ) bond while, bonds to the atoms below the plane are shown using hash wedge ( ) bonds. The narrow end of these bonds should point towards the stereogenic center. Two enantiomers of 2-chlorobutate are shown below using stereochemical bonds (*Figure* 19).

Figure 19. Two enantiomers of 2-chlorobutate shown using stereochemical bonds.

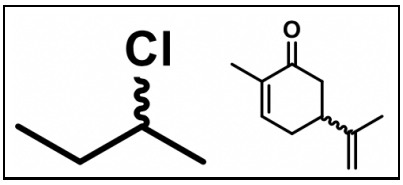


Figure 20. Racemic forms of 2-chlorobutane and carvone.

One more stereochemical bond (~~~~~) is practiced to denote racemic mixtures. This bond is called a 'wavy bond' or 'either bond.' The molecules shown below depict racemic forms of 2-chlorobutane and carvone respectively (*Figure* 20).

IUPAC has also published one more publication focusing on graphical representation of stereochemical structure in 2006 [2].

Depicting Cycloalkanes

Cycloalkanes are generally depicted by wiring equilateral polygons when stereochemical aspects are not to be emphasized (GR-3.4.2). E.g., cyclohexane can be drawn as a regular hexagon (*Figure* 21), but the chair or boat forms are preferred when an emphasis is desired on a specific conformer.

While writing bigger ring sizes or macrocyclic alkanes, a specified pattern may be followed. The number of carbons may be mentioned in the center of such rings so that the number of carbons in the ring could quickly be noted (*Figure* 22).

Bond Overlap

Overlap of bonds is inevitable while drawing a cage and three-dimensional structures. The rear bond must be shown discontinuous at the point of apparent intersection. It will help in un-

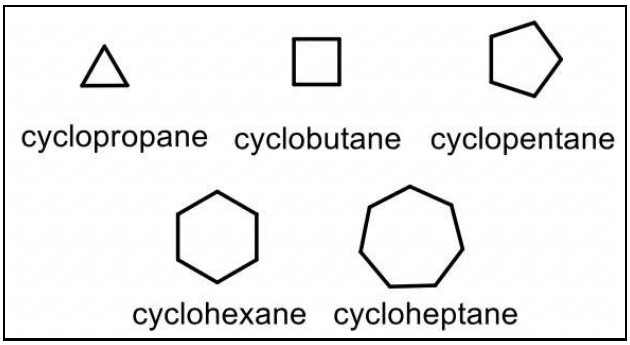


Figure 21. Cycloalkanes are written as regular polygons.

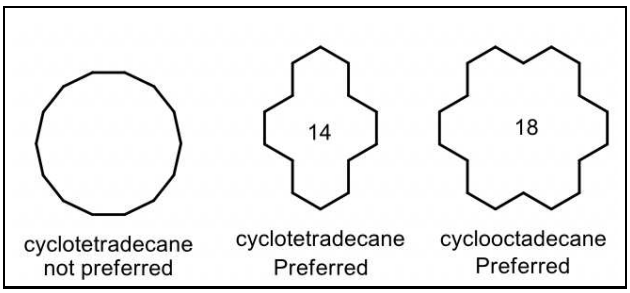


Figure 22. Number of carbons mentioned in the center of macrocyclic alkane rings.

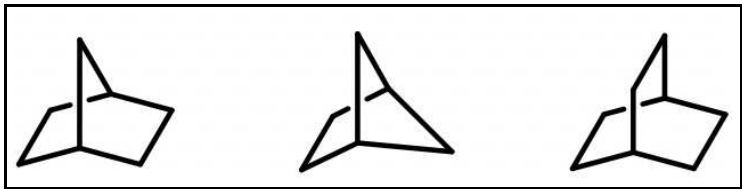


Figure 23. Intersections of bonds.

derstanding the pattern of overlapping. The intersections of the bonds sometimes may be taken as an additional carbon if the rule is neglected (*Figure* 23).

Molecular drawing further can be simplified with accepted abbreviations. The passive part of the molecule can be abbreviated.

Additional thickness of some bonds is allowed when emphasis is desired for front facing bonds (*Figure* 24).

Use of Abbreviations

Molecular drawing further can be simplified with accepted abbreviations. The passive part of the molecule can be abbreviated.

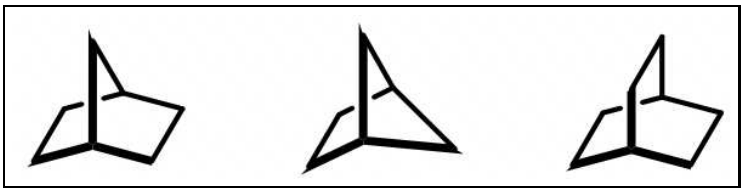


Figure 24. Additional thickness of bonds to emphasize front facing bonds.

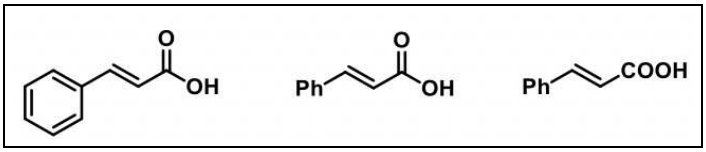


Figure 25. Standard abbreviations incorporated to structures.

For example, cinnamic acid may be written in three ways. Few standard abbreviations can intelligently be incorporated to sketch a molecule as simply as possible without having adjusted with the essence of drawing (*Figure* 25).

As a rule, the passive part of the molecule shall only be abbreviated. Few most commonly used abbreviations are: Me = methyl, Et = ethyl, *n*-Pr = *n*-propyl, *i*-Pr = isopropyl, *n*-Bu = *n*-butyl, *t*-Bu =, tert-butyl, Ph = phenyl, Bz = benzyl, Ac = acetyl. More abbreviations may be noted in [4].

Rules for Side Chain Representation

Side alkyl chains shall not be abbreviated if there is a possibility of isomerism (GR-2.3.9). It is recommended to write desired alkyl chains to avoid any ambiguity. For example, abbreviating the side chain as C_4H_9 can imply five possible isomers (*Figure* 26).

The Depiction of Ionic Bonds

Ionic bonds shall never be written with a 'dash' (GR-7.1), as the depiction may contradict with a covalent bond. The atoms should be written with explicit positive and negative charges to mention the ionic bonding (*Figure* 27).

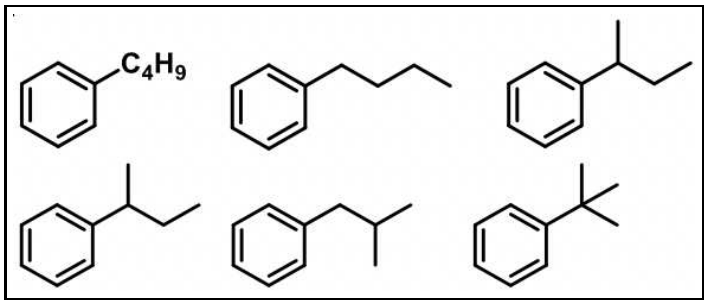


Figure 26. Abbreviating the side chain as C_4H_9 can imply five possible isomers.

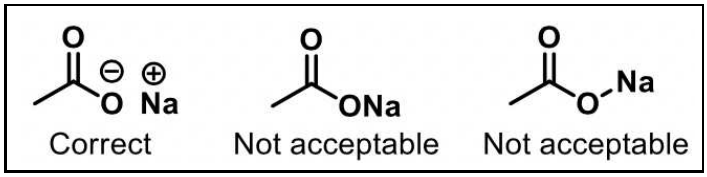


Figure 27. Atoms written with explicit positive and negative charges to mention the ionic bonding .

The Orientation of Atom Label

Atom labels should be connected properly with the primary structures (GR-2.1.6). The atom connectivity must be obeyed correctly (*Figures* 28(a) and (b)).

Similarly, the structures of esters and carboxylic acids must be written with care as often these are miswritten.

Care must be taken while writing functional groups appearing at the left-hand side to avoid errors, or the functional groups in the ambit of reaction shall be placed at the right of the primary structure (*Figure* 28(d)).

In the same way, abbreviations must fall in line with the usual way of reading (*left to right*). For example, in *Figure* 28(e), EWG (*electron-withdrawing group*) is incorrectly written as GWE. TMS (*trimethylsilyl*) sometimes is wrongly written as SMT (*Figure* 28(e)).

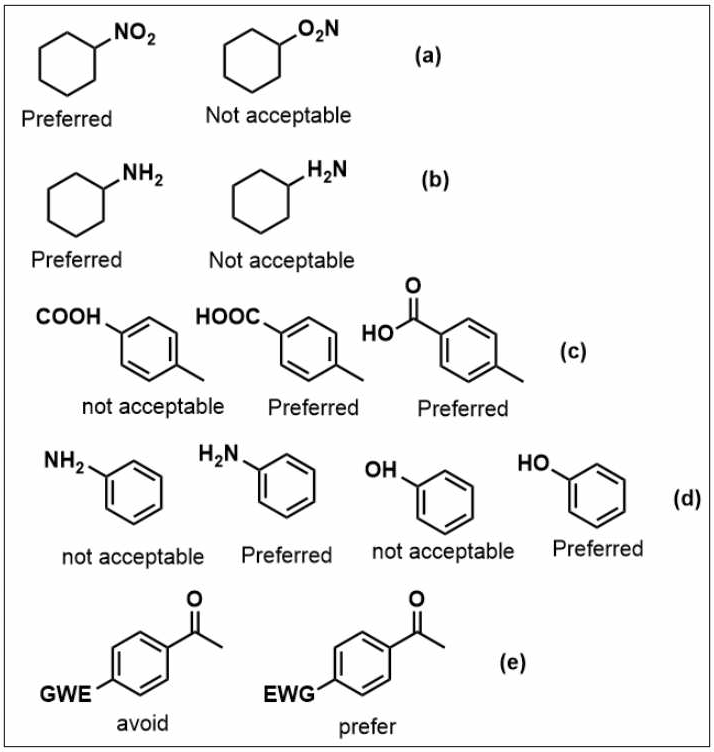


Figure 28. Atom connectivity.

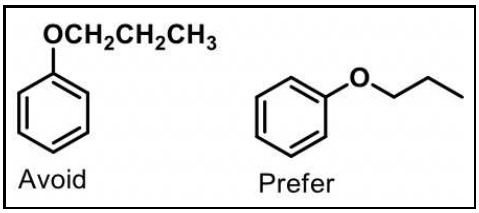


Figure 29. Structures should be written in a consistent bond line form.

Some Simple Expectations Worth Mentioning

While writing the bond-line structures, the entire molecule shall be written consistently in bond-line form only. The first structure in (*Figure* 29) may be avoided and should be written in a consistent bond line form.

The overlap between the substituents or too close placement of functional groups can be skillfully avoided without adjusting the

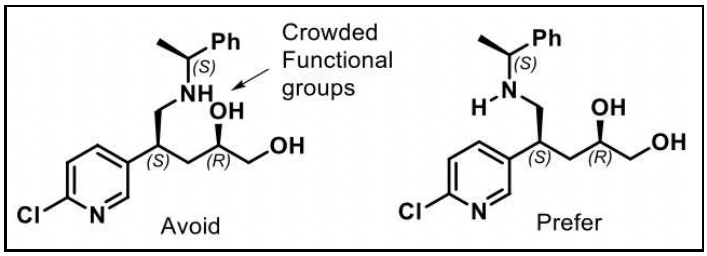


Figure 30. The overlap between the substituents or too close placement of functional groups can be skillfully avoided without adjusting the essence isomerism.

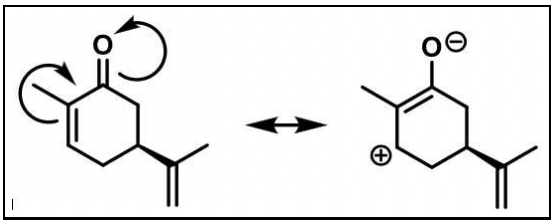


Figure 31. The resonance structures of carvone.

essence of isomerism (*Figure* 30).

Besides, the above-mentioned concise rules and regulations, a few good practices make chemical sketch simple. The orientation of the molecule may be kept unchanged while writing the mechanisms to convey the net change in the structure. For example, the resonance structure of carvone is shown below, and the direction of the isopropenyl group is kept unaltered while writing the next step (*Figure* 31).

Understanding stereochemical aspects of a molecule indeed become more comfortable if the orientation of the molecule is kept constant, and the change in stereochemistry is depicted using stereochemical bonds (*Figure* 32).

Conclusion

There is an unambiguous grammar for organic chemistry script. Violation of this grammar can lead to severe confusion and deviate meanings, and hence it is essential to follow the standards suggested by IUPAC.

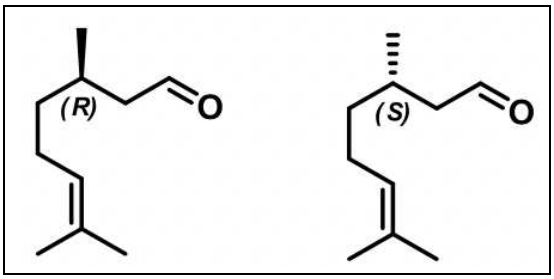


Figure 32. Face of the molecule is retained as it is while writing a pair of enantiomers.

Suggested Reading

[1] **J. Brecher, Graphical Representation Standards for Chemical Structure Diagrams, IUPAC Recommendation-2008, *Pure Appl. Chem*., Vol.80, No.2, pp.277–410, 2008.**

[2] **J. Brecher, Graphical Representation of Stereochemical Configuration, IUPAC Recommendation-2008, *Pure Appl. Chem*., Vol.80, No.2, pp.277–410, 2008.**

[3] **J. Clayden, N. Greeves, S. Warren, P. Wothers, *Organic Chemistry*, 1st edn, Oxford University Press, New York, pp.19–46, 2001.**

[4] **Standard Abbreviations, http://pubs.acs.org/page/joceah/submission/authors.html**