POLITECNICO DI MILANO

Scuola di Ingegneria Industriale e dell'Informazione

Course 096125 (095857)

Introduction to Green and Sustainable Chemistry





Chemistry and Organic Chemistry Recall (Symbols and Conventions)

Prof. Attilio Citterio

Dipartimento CMIC "Giulio Natta"

https://iscamapweb.chem.polimi.it/citterio/it/education/course-topics/

Understand Chemistry and its Impact.

Three levels of understanding Chemistry (and, more in general, of other sciences) have been proposed*:

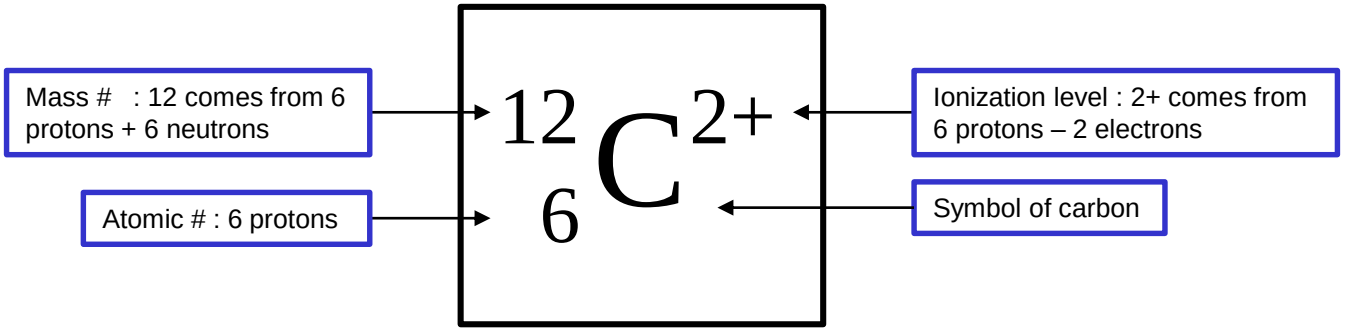
- macroscopic — i.e., the physical and chemical phenomena of chemistry;
- sub microscopic or particle — i.e., the models of chemical behaviour at atomic and molecular scale;
- symbolic — i.e., the symbols, formulae and mathematical relationship used to express the chemical understanding.

The nature of chemistry concepts and the way the concepts are represented (macroscopic, microscopic, or representational) make chemistry difficult to learn, even if the introductive teaching of this matter mainly focus on the more abstract of these levels, the symbolic. An explicit effort of integration of three levels is necessary, but this induces further confusion. Chemistry learning requires much intellectual thought and discernment because the content is replete with many abstract concepts and students found problems in visualizing particulate models and mathematical relationship with sub microscopic level.

* A. H. Johnstone. *J. Computer Assisted Learning* **7**, 75 (1991).

Chemical Notation.

Chemical notation is a specialized system of signs and symbols used in chemistry. It is a short hand system to represent what happens in systems without having to write the details in long hand.



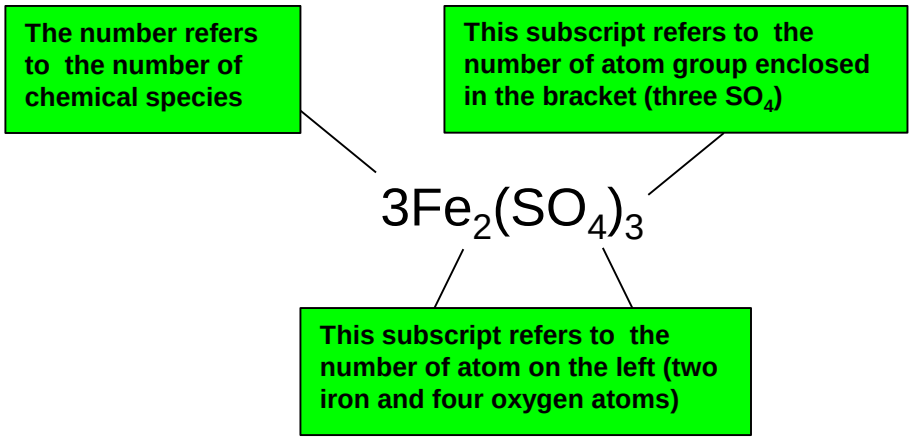
Compounds can be represented in the chemical notation system by the symbols of the elements that compose them. This corresponds to the chemical formula. The name of the compound is part of the chemical nomenclature system . Rules exist for the purpose that have been defined by the **International Union of Pure and Applied Chemistry (IUPAC)** correlating the name to the corresponding chemical formula.

http://www.chem.qmul.ac.uk/iupac/

Chemical Formula.

Formula Unit - is a molecule (for compounds consisting of molecules) or the smallest fragment with the same ratio between atoms, as in compound.

Chemical formula shows how many atoms of each type are present in the formula unit.



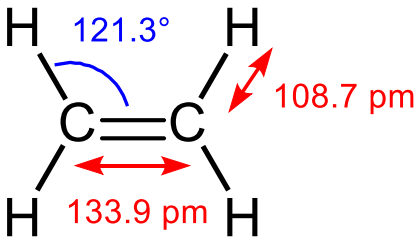
Examples:

1. K_3PO_4 means: "One formula unit of ionic compound K_3PO_4 contains 3 atoms of potassium, 1 atom of phosphorus and 4 atoms of oxygen".
   3, 1, and 4 – **chemical indices**

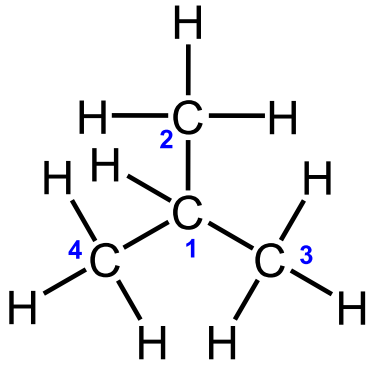
2. $C_3H_8O_3$ means: "One formula unit of compound $C_3H_8O_3$ (molecule, glycerol) contains 3 atoms of carbon, 8 atoms of hydrogen and 3 atoms of oxygen".

Molecular Geometry and Structural Formulas.

The connectivity of a molecule determines its physical and chemical properties and behavior. Isomers (molecules composed of the same numbers of the same types of atoms) are characterized by different structural formula, i.e. which atoms are bonded. From the connectivity, the shape of the molecule can be deduced. A chemical formula supplies information about the types and spatial arrangement of bonds (their distance and angles) and its multiplicity (single, double, triple, aromatic, etc.). A single line (or a pairs of dots) is indicative of the shearing of 2 electrons (sigma bond), two lines indicate that a double bond connects the atoms on either side of them, and three lines indicate a triple bond. Molecules with multiple equal functional groups are expressed by enclosing the repeated group in round brackets.



Ethylene Molecular formula: C_2H_4
Semi-structural formula: $CH_2=CH_2$
Bond distance and bond angles

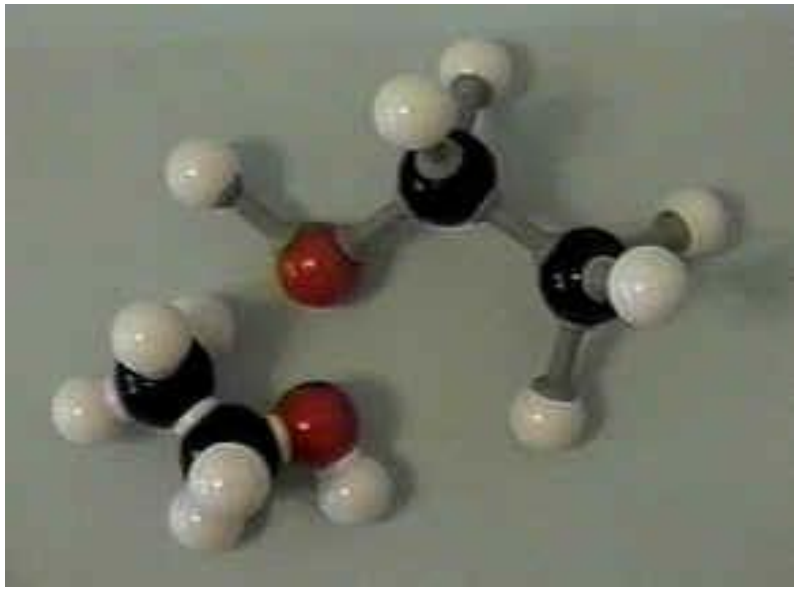


Isobutane (2-methylpropane)
Molecular formula: C_4H_{10}
Semi-structural formula: $(CH_3)_3CH$

Visualization of Molecules.

Heptane

Ball and stick
Vision



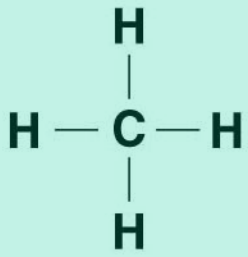
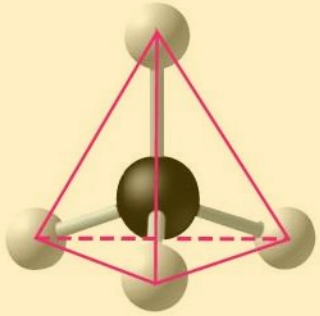
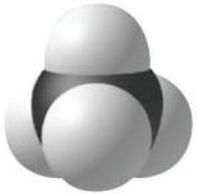
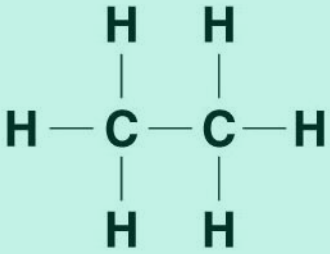
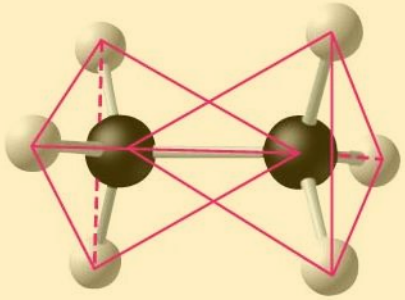
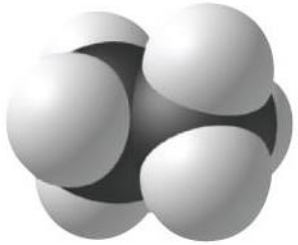
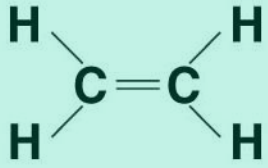
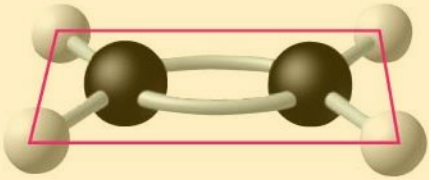
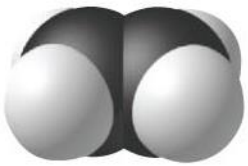
CH_3CH_2OH

Ball and stick vs.
Van der Walls
Vision

Chemical Software:

a) ChemSketch - https://www.acdlabs.com/
b) ChemOffice - http://www.cambridgesoft.com

Writing and Displaying Chemical Formula.

Molecular Formula	Structural Formula	Ball-and-Stick Model	Spatial Model
CH_4 (a) Methane			
C_2H_6 (b) Ethane			
C_2H_4 (c) Ethylene			

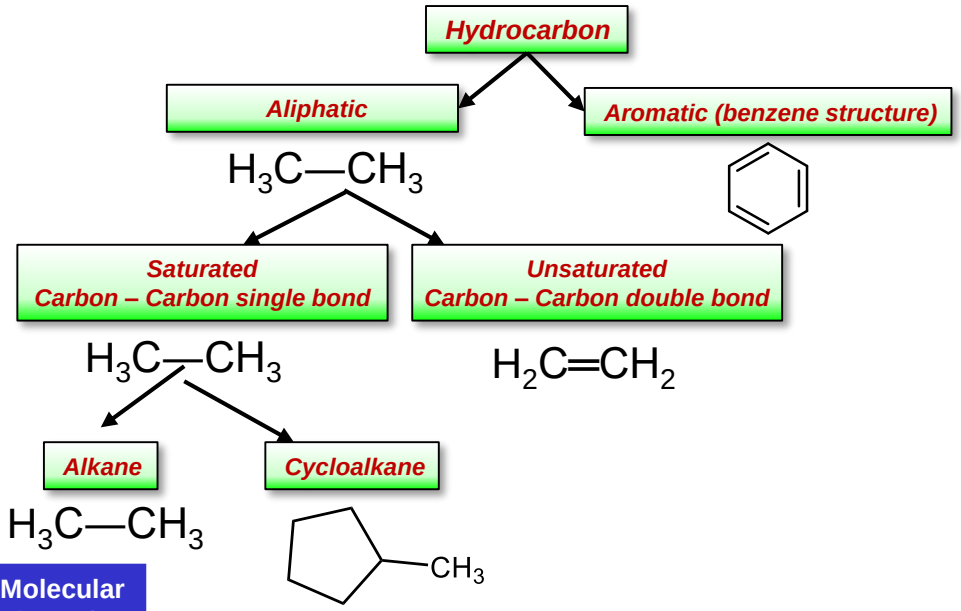
Alkane Hydrocarbon and Nomenclature.

Homologous Series

- same functional group
- each member differ from the next by $-CH_2$ group
- similar chemical properties
- same chemical formula C_nH_{2n+2}
- saturated hydrocarbon
- ending with ane

Physical properties

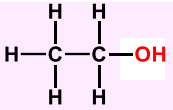
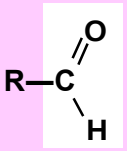
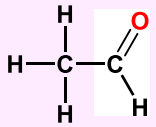
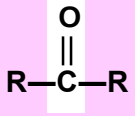
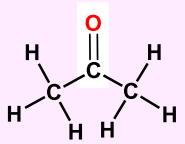
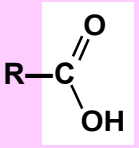
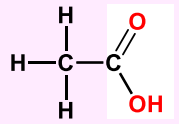
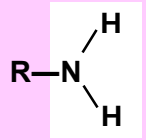
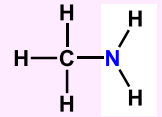
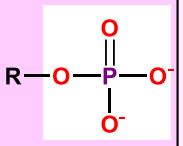
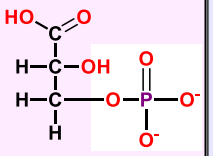
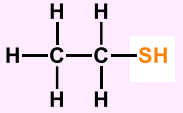
- increase in RMM and molecular size down the series
- van der Waals forces stronger
- melting / boiling point increases
  - $C_1 - C_4$ (gas)
  - $C_5 - C_{18}$ (liquid)
  - C_{19} – (solid)
- insoluble in water, soluble in organic solvent
- non conductors – simple molecules NOT ions



Number of C atoms	Word root	IUPAC name	Structure	Molecular formula
1	Meth	Methane	CH_4	CH_4
2	Eth	Ethane	CH_3-CH_3	C_2H_6
3	Prop	Propane	CH_3-CH_2-CH_3	C_3H_8
4	But	Butane	CH_3-$(CH_2)_2$-CH_3	C_4H_{10}
5	Pent	Pentane	CH_3-$(CH_2)_3$-CH_3	C_5H_{12}
6	Hex	Hexane	CH_3-$(CH_2)_4$-CH_3	C_6H_{14}
7	Hept	Heptane	CH_3-$(CH_2)_5$-CH_3	C_7H_{16}
8	Oct	Octane	CH_3-$(CH_2)_6$-CH_3	C_8H_{18}
9	Non	Nonane	CH_3-$(CH_2)_7$-CH_3	C_9H_{20}
10	Dec	Decane	CH_3-$(CH_2)_8$-CH_3	$C_{10}H_{22}$

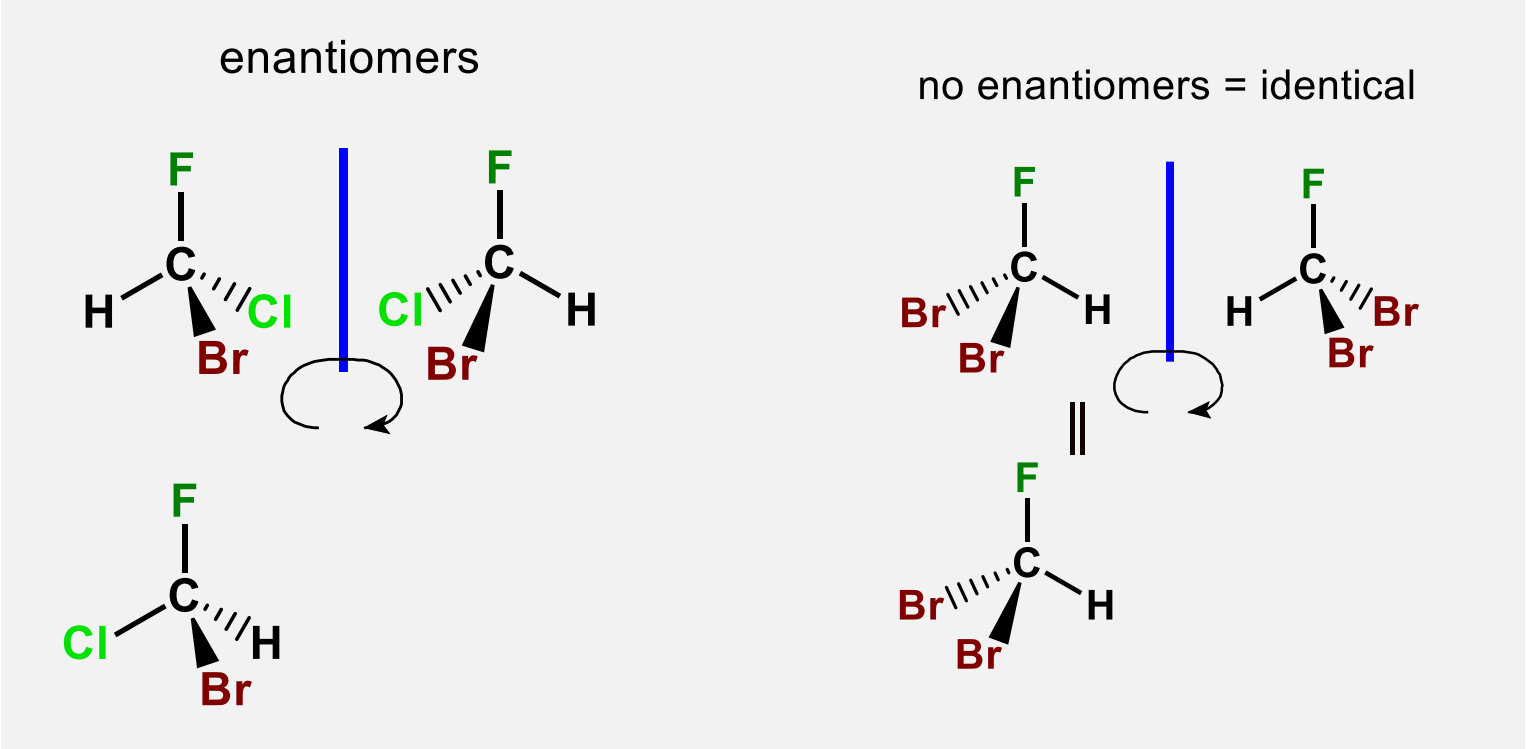
Functional Groups.

The seven basic functional groups present in the biological monomers

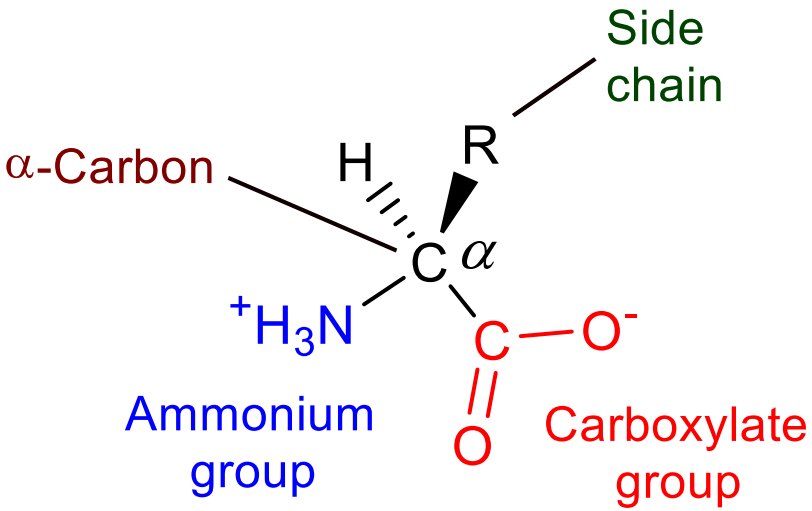
Functional Groups	Class of compounds	Structural formula	Example	Name
Hydroxyl —OH	Alcohols	R—OH		Ethanol
Carbonyl —CHO	Aldehydes			Acetaldehyde
Carbonyl 	Ketones			Acetone
Carboxyl —COOH	Carboxylic acids			Acetic acid
Amino —NH_2	Amines			Methylamine
Phosphate —OPO_3^{2-}	Organic phosphates			3-Phospho-glyceric acid
Sulfhydryl —SH	Thiols	R—SH		Mercapto-ethanol

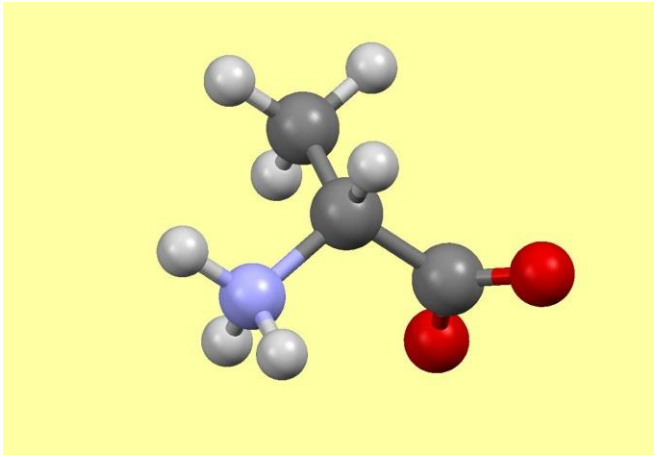
Enantiomers and the Tetrahedral Carbon.

- **Molecules that have one carbon with 4 different substituents have a non-superimposable mirror image – enantiomer**
- **Build molecular models or visualization software to see this:**



Amino Acids: Building Blocks of Proteins.





L-alanine (R = CH_3)

Anatomy of an amino acid. Except for proline and its derivatives, all of the amino acids commonly found in proteins possess this type of structure. The chiral ones are L-stereochemistry at the alpha carbon.

Chemical Notation for Polymers and Ions.

Polymers

For polymers, parentheses are placed around the repeating unit. For example, a $CH_3(CH_2)_{40}CH_3$ *hydrocarbon molecule* contains forty repeating units. If the number of repeating units is unknown or variable, the letter *n* may be used to indicate the formula: $CH_3(CH_2)_nCH_3$.

Polyethylene = $–[CH_2\text{-}CH_2]_n\text{-}$; polyvinyl chloride = $\text{-}[CH_2\text{-}CHCl]_n\text{-}$

Ions

For ions, the charge on a particular atom is denoted with a right-hand superscript, e.g. Na^+ or Cu^{2+}. The total charge on a charged molecule or a polyatomic ion is similarly indicated. For example: H_3O^+ or SO_4^{2-}.

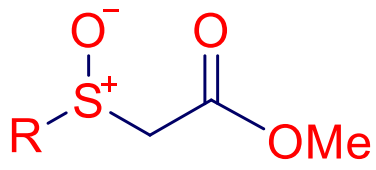
For more complex ions, brackets [ ] are often used to enclose the ionic formula, as in $[B_{12}H_{12}]^{2-}$, which is found in compounds such as $Cs_2[B_{12}H_{12}]$. Parentheses ( ) can be nested inside brackets to indicate a repeating unit, as in $[Co(NH_3)_6]^{3+}$. Here $(NH_3)_6$ indicates that the ion contains six NH_3 groups, and [ ] encloses the entire formula of the complex ion with charge +3.

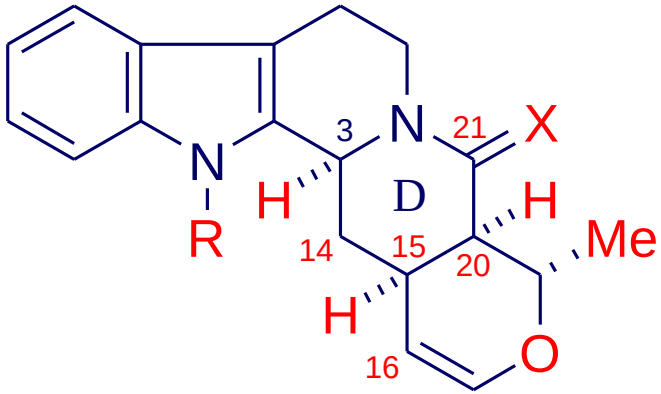
Chemical Structure Diagrams.

- Chemical structure diagrams are a form of representation of chemical compounds.
- Information contained in a structure diagram can be divided into three areas:
  - Atom information

chemical elements,
functional groups,
generic elements,

- vertex label,
- charge,
- atomic weight,
- hybridization,
- etc.

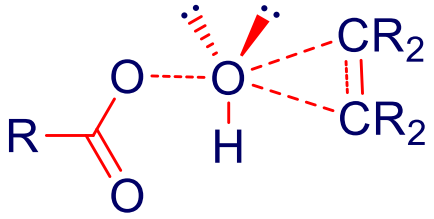


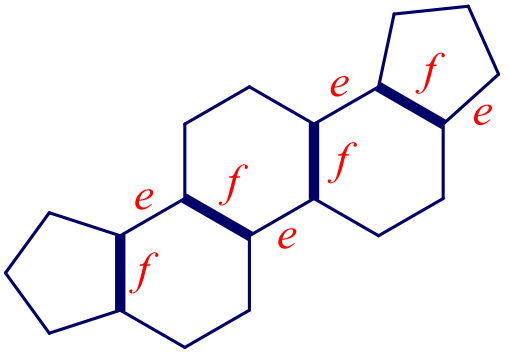


Chemical Structure Diagrams (2).

- Chemical structure diagrams are a form of representation of chemical compounds.
- Information contained in a structure diagram can be divided into three areas:
  - Atom information
  - Bond information

  - Bond orders,
  - Bond styles,
  - Bond labels
  - Transition states



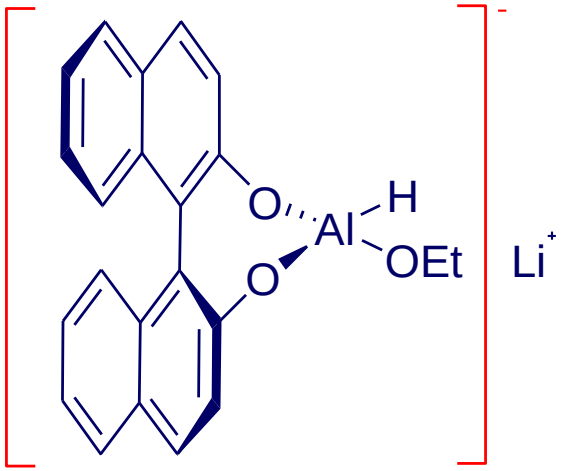


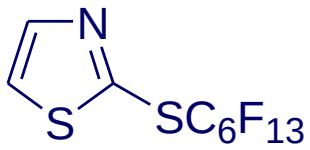
Chemical Structure Diagrams (3).

- Chemical structure diagrams are a form of representation of chemical compounds.
- Information contained in a structure diagram can be divided into three areas:
  - Atom information
  - Bond information
  - Structural information

atom information,
bond information,

- overall charge,
- structure label
- numbering





21

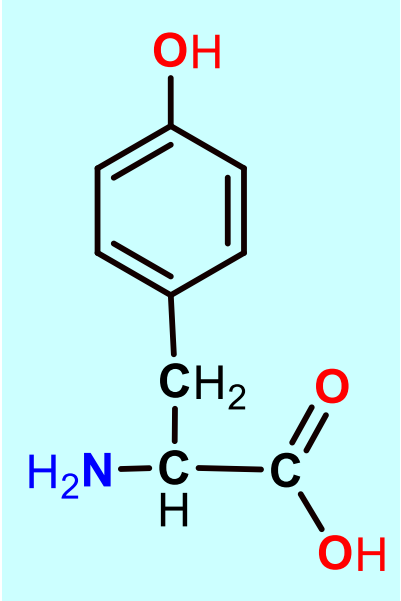
Representing a Chemical Structure.

- How much information do you want to include?
  - atoms present
  - connections between atoms
    - bond types
  - stereochemical configuration
  - charges
  - isotopes
  - 3D-coordinates for atoms

$C_8H_9NO_3$

Representing a Chemical Structure (2).

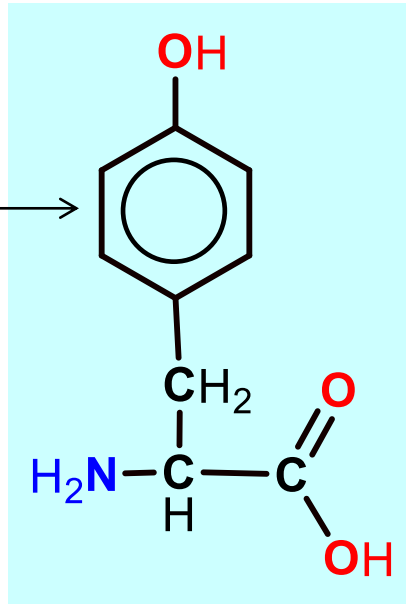
- How much information do you want to include?
  - atoms present
  - connections between atoms
    - bond types
  - stereochemical configuration
  - charges
  - isotopes
  - 3D-coordinates for atoms



Representing a Chemical Structure (3).

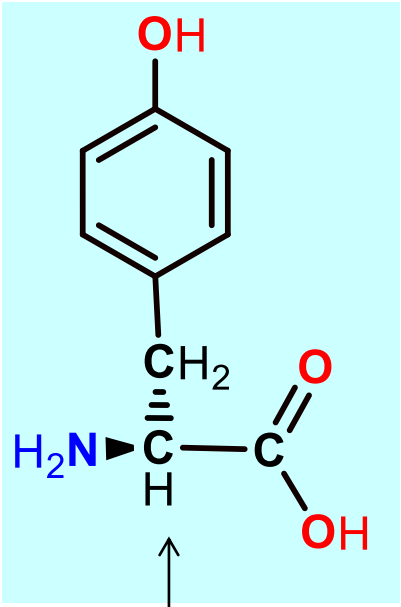
- How much information do you want to include?
  - atoms present (O, N, C, H, etc.)
  - connections between atoms
    - bond types

    (aromatic ring identification)
  - Stereochemical configuration
  - charges
  - isotopes
  - 3D-coordinates for atoms



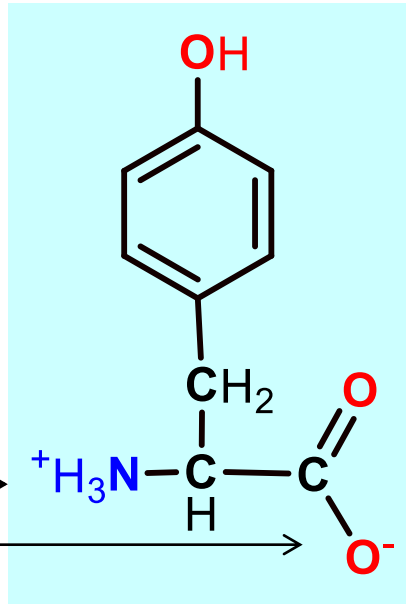
Representing a Chemical Structure (4).

- How much information do you want to include?
  - atoms present
  - connections between atoms
    - Bond types
  - Stereochemical configuration
  - charges
  - isotopes
  - 3D coordinates for atoms



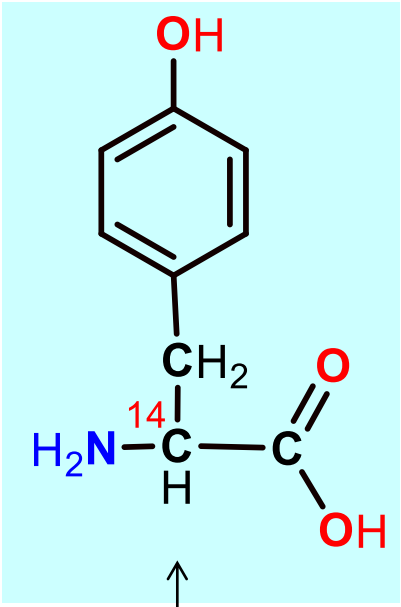
Representing a Chemical Structure (5).

- How much information do you want to include?
  - atoms present
  - connections between atoms
    - Bond types
  - Stereochemical configuration
  - charges
  - isotopes
  - 3D coordinates for atoms



Representing a Chemical Structure (6).

- How much information do you want to include?
  - atoms present
  - connections between atoms
    - bond types
  - stereochemical configuration
  - charges
  - isotopes
  - 3D-coordinates for atoms

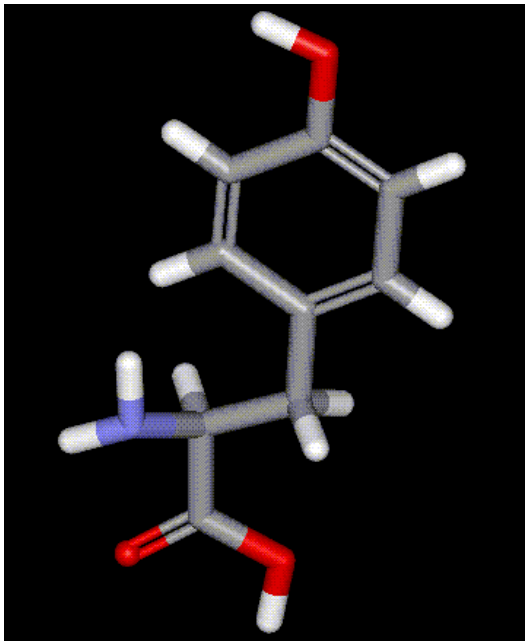


Representing a Chemical Structure (7).

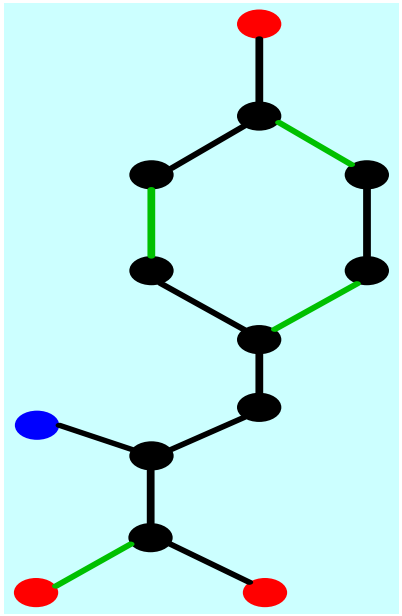
3D structure diagrams.

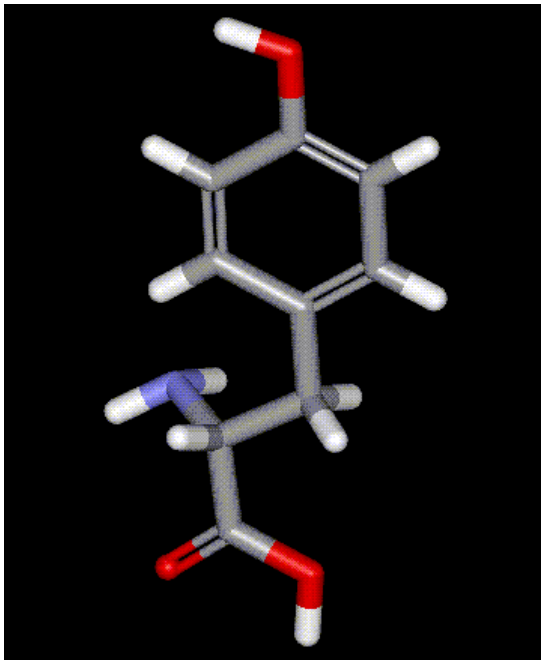
Stereochemistry:

- different compounds with identical connectivity
- same topology, different topography



S-tyrosine

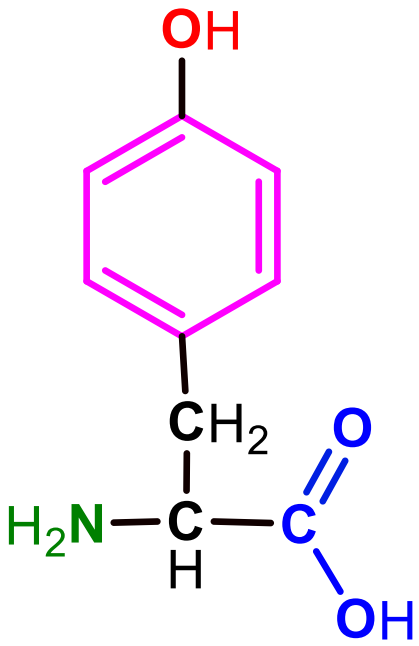




R-tyrosine

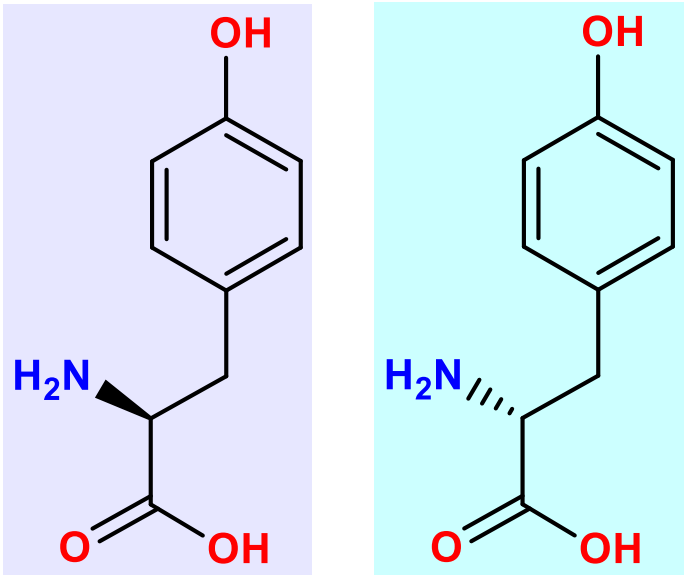
Substructure Fragments.

- Subgraphs can be identified in a structure graph corresponding to functional groups, rings, etc.
  - –OH
  - –NH_2
  - –COOH
  - phenyl
- this can be done by tracing appropriate paths in the graph
- subgraphs may overlap



Stereochemistry: up/down Bonds.

- can be used as additional “colours” for graph edges
  - many connection table formats have special codes for up and down bonds
  - need to know which end of bond is which



- useful for re-generating diagrams for display
- can be used to calculate other stereo descriptors

Registry Numbers of Compound/Substances.

- unique identifiers for compounds or substances
  - catalogue number
- most chemical databases have them
  - Chemical Abstracts ((i.e. 74-85-1 for ethylene)
  - Beilstein
  - private compound registries in pharmaceutical companies
- usually just “idiot numbers”
  - no chemical information
- may have hierarchical structure

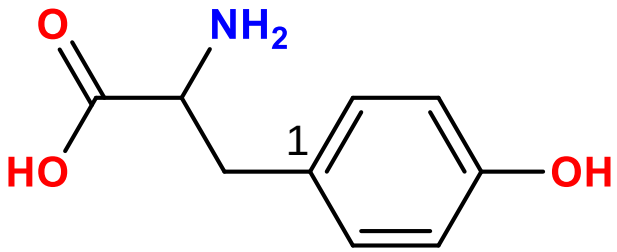
  parent compound → stereoisomer → salt → batch
- need to decide what is a separate compound

Line Notations.

- represent structures as compact linear string of alphanumeric symbols
- easily handled by computer
  - compact storage
  - easily transmitted over a network
- allow rapid manual coding/decoding by trained users
  - much faster for input than using a structure drawing program

Line Notations: SMILES and Beilstein.

- Simplified Molecular Input Line Entry System
  - developed by Dave Weininger (Daylight)

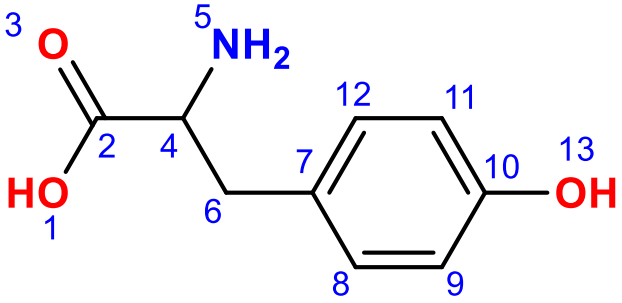


OC(=O)C(N)CC1=CC=C(O)C=C1

- ROSDAL (Beilstein)

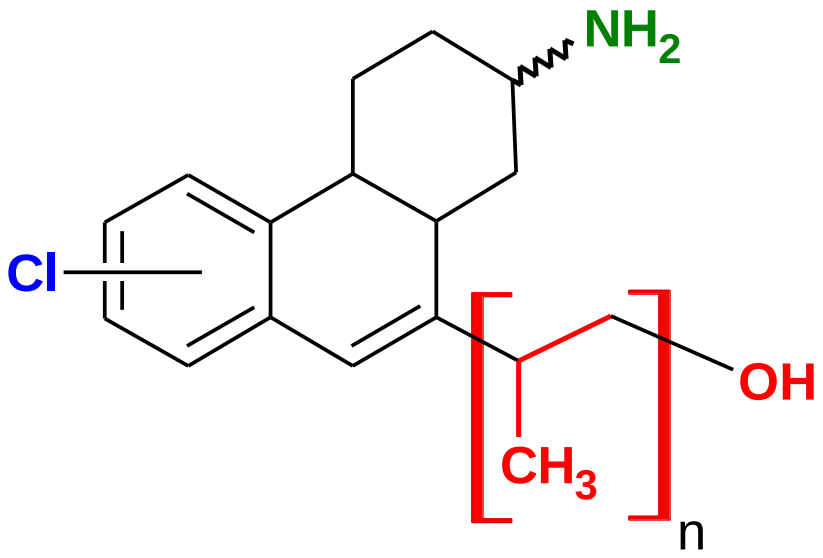
  Representation Of Structure Diagram Arranged Linearly

  1O-2=3O,2-4-5N,4-6-7=-12-7,10-13O



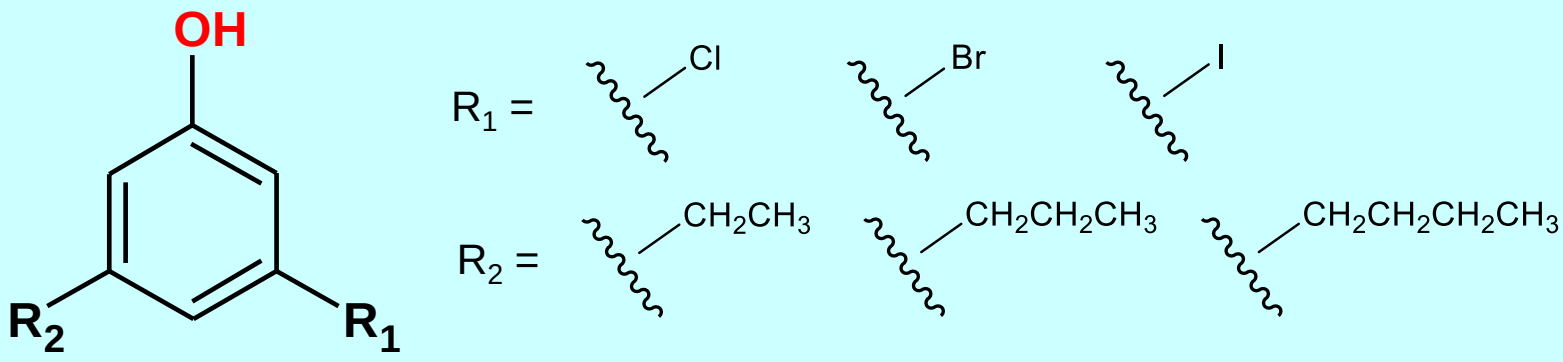
Incompletely-defined Substances.

- unknown stereochemistry
- unknown attachment position
- unknown repetition



Markush ("Generic") Structures.

- structures with R-groups, also called "generic" structures
- shorthand for describing sets of structures with common features



- very important in chemical patents (inventor claims whole class of related compounds)
- can be used to describe combinatorial libraries or as queries in database searches

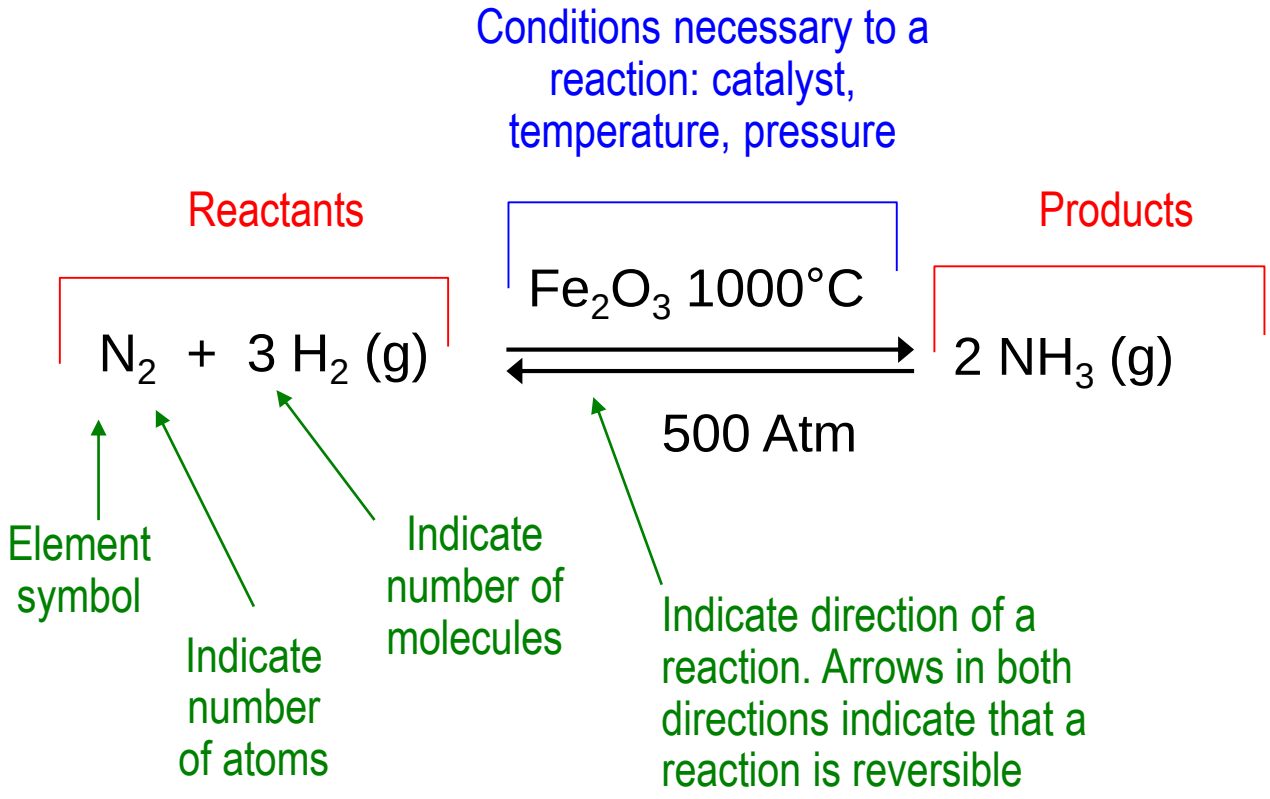
More on Chemical Equations.

- **Chemical equations** show how many formula units react and how many formula units are formed in the reaction.
- An **arrow** shows the direction of the reaction. If the reaction reaches the equilibrium, a **double arrow** with reverse direction is used.
- For the **mass conservation law**, the number of each type of atoms in the left part of the equation must be equal to the number of atoms of the same type in the right part of the equation (**stoichiometry**).
- Upper and below the arrow can be reported reagents and products if appropriate but also solvents and physical conditions (T, P, etc.).

Example:

- $CH_4 + 2\ O_2 \rightarrow CO_2 + 2\ H_2O$ means: “One molecule of compound CH_4 reacts with 2 molecules of the compound O_2 and produces 1 molecule of the compound CO_2 and 2 molecules of the compound H_2O”.
- 1, 2, 1 and 2 - **stoichiometric chemical coefficients** of the reaction.

The Chemical Equation: General Symbols.

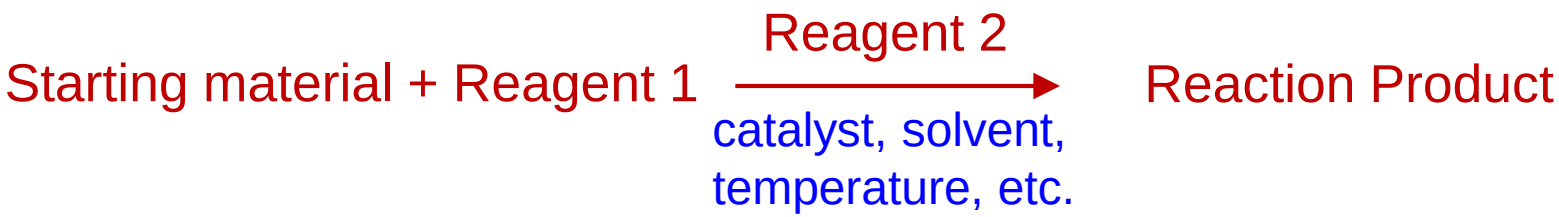


One-step Reaction Schemes.

One-step chemical schemes show, which compounds react, which compounds are formed, which compounds and conditions affect the reaction.

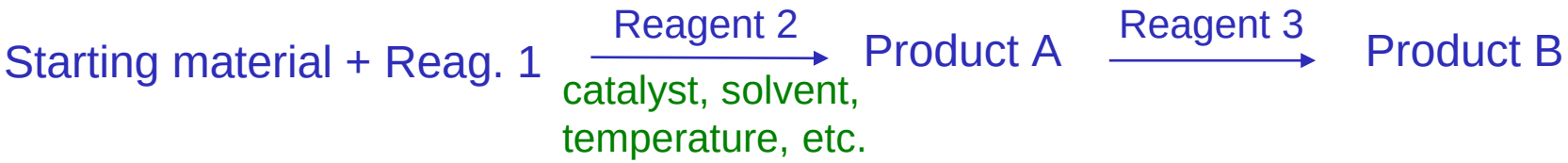
- At least one organic compound must be present in the left part of the scheme
- All other reacting compounds are usually called “reagents” and shown either in the left part of the scheme or over the reaction arrow.
- Chemical coefficients are not shown, except if you need to show some stoichiometric features of the reaction.
- Some reaction products, especially inorganic ones, are often not shown.

Example of generic reaction:



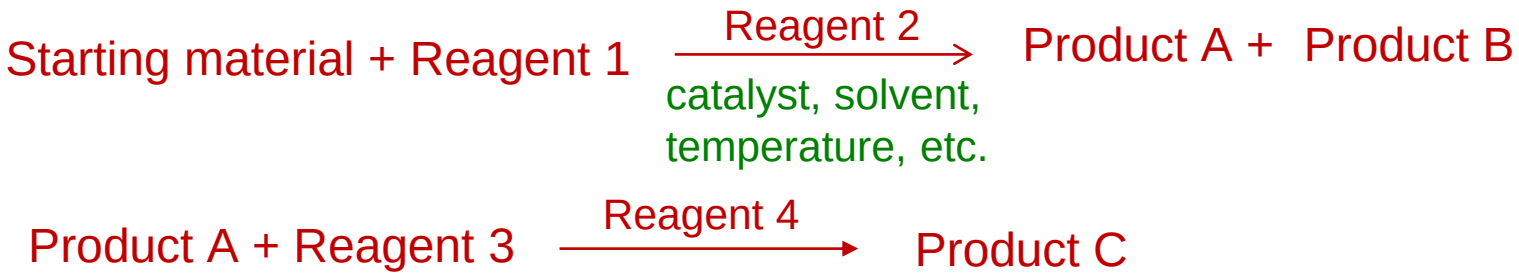
Multi-step Reaction Schemes.

If there is only one shown reaction product, for the next step, write the reagents **over the reaction arrow**. Otherwise the reagents for the next step will be confused with the reaction products for the previous step, which makes the whole reaction scheme wrong.



If there is more than one shown reaction product, there are two ways to write :

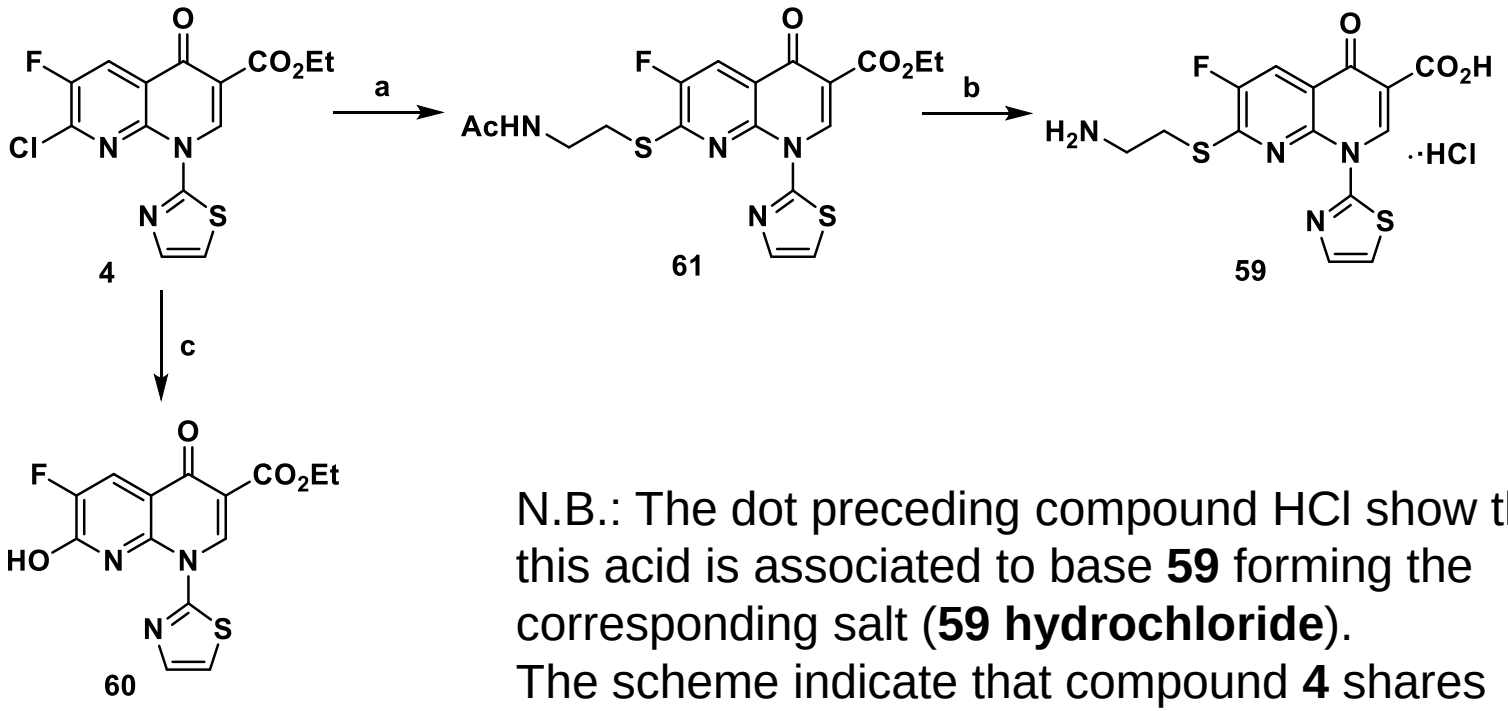
1. *Start each step from a new line, rewriting the product of the preceding step as the starting material form the next step.*



2. *Draw the next reaction arrow from the product of the preceding step, and write the reagent **over the reaction arrow**.*

Some Elements of Chemical Communication.

- Reaction Schemes and numbering:

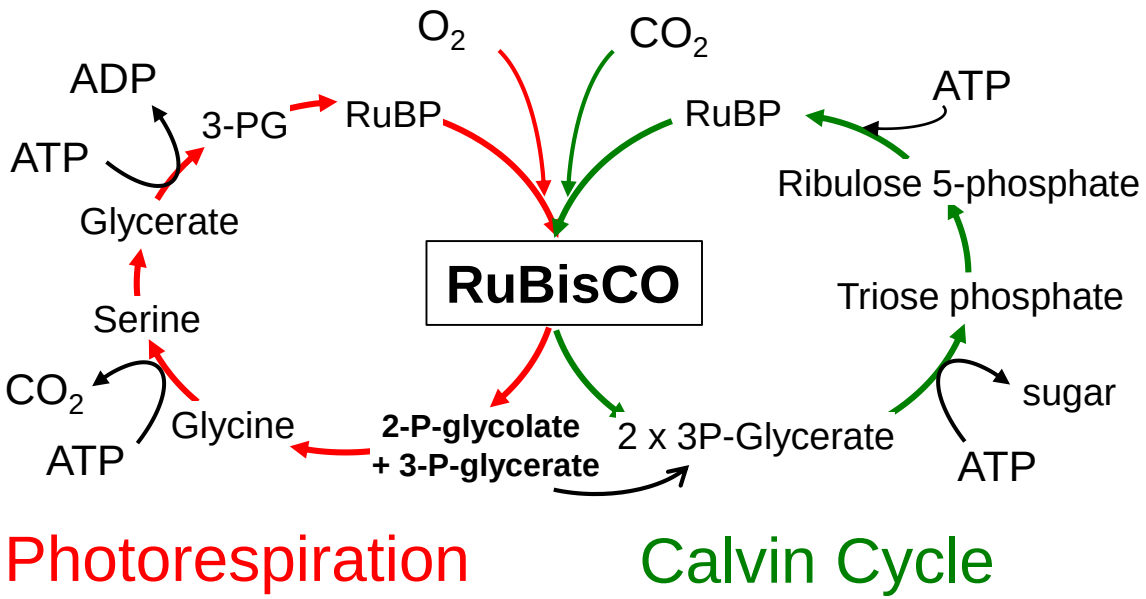


N.B.: The dot preceding compound HCl show that this acid is associated to base **59** forming the corresponding salt (**59 hydrochloride**).
The scheme indicate that compound **4** shares competitively the routes **a** and **c** affording products **61** and **60**, respectively, and that compound **61** is converted to **59** via process **b**.

Chemical Cycles.

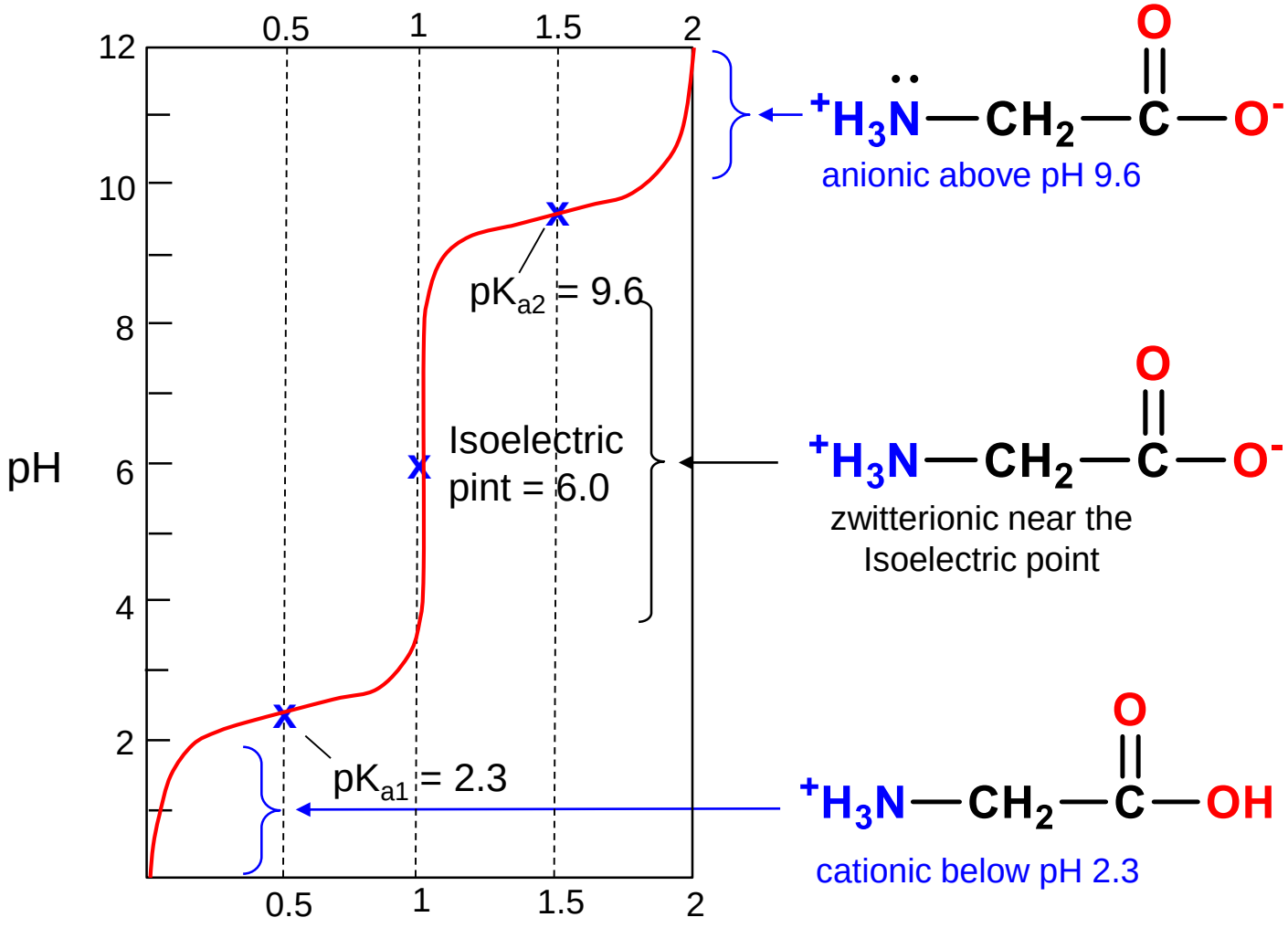
In biochemistry and chemistry is useful to evidence a long sequence of reactions all interconnected in a cyclic way. This is visually indicated as a cyclic sequence of compound interconnected by curved lines (with arrows). Input to the cycle are indicated with arrows from outer side, whereas output of the cycle are indicated with arrows to the outer side.

Example:



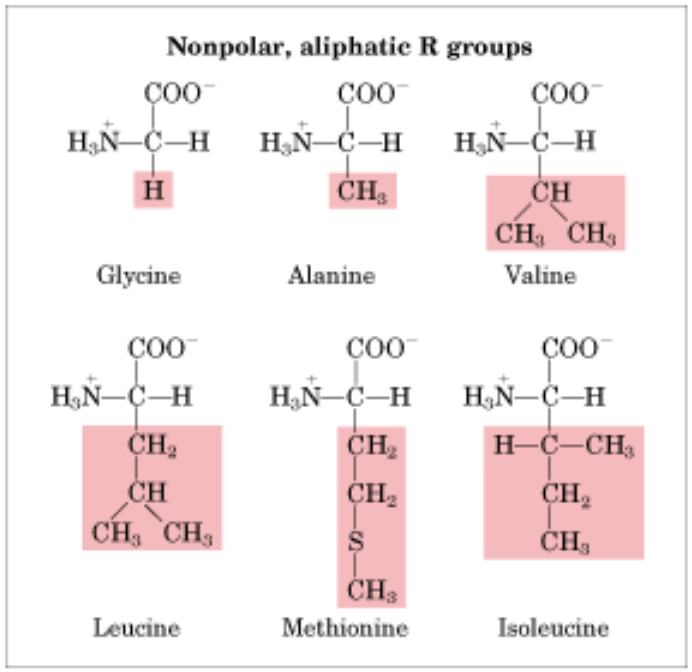
RuBisCO = ribulose-1,5-bisphosphate carboxylase/oxygenase

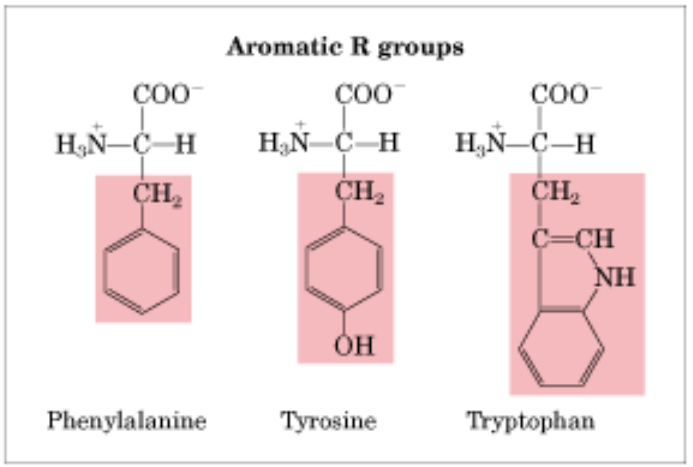
Multiple Equilibria and Product Distribution.

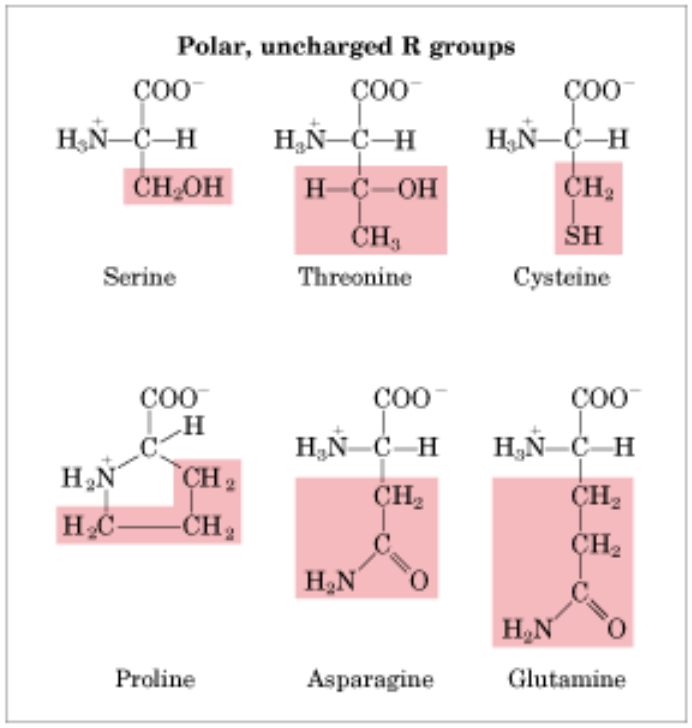


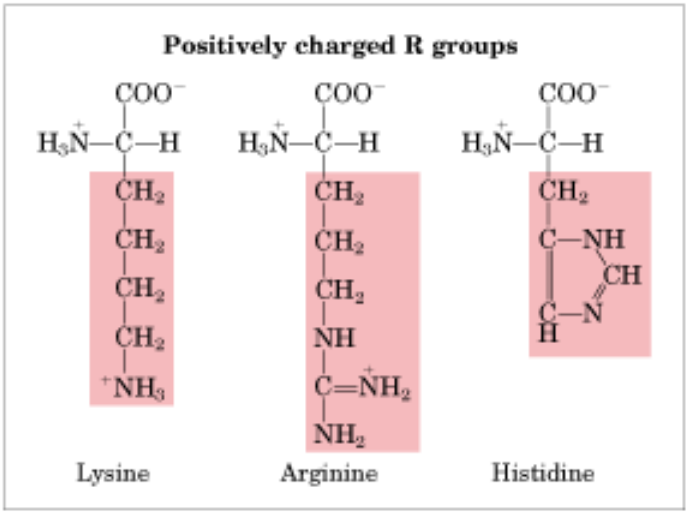
=>

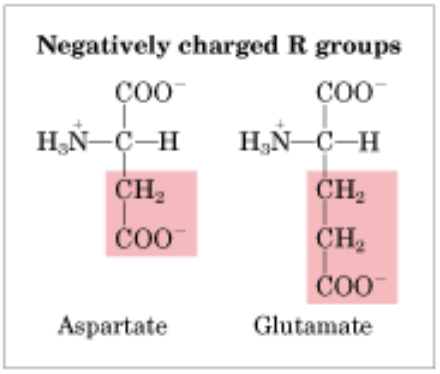
Amino-Acids.


Nonpolar, aliphatic R groups
Glycine
Alanine
Valine
Leucine
Methionine
Isoleucine


Aromatic R groups
Phenylalanine
Tyrosine
Tryptophan


Polar, uncharged R groups
Serine
Threonine
Cysteine
Proline
Asparagine
Glutamine


Positively charged R groups
Lysine
Arginine
Histidine


Negatively charged R groups
Aspartate
Glutamate

POLITECNICO DI MILANO

School of Industrial and Information Engineering

Course 096125 (095857)

Introduction to Green and Sustainable Chemistry





Fundamental Reaction Types

Prof. Attilio Citterio

Dipartimento CMIC “Giulio Natta”

https://iscamapweb.chem.polimi.it/citterio/it/education/course-topics/

Fundamental Reaction Types.

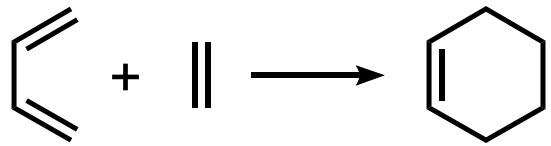
Elementary Mechanisms:

- Decomposition
- Pericyclic (Cycloaddition)*
- Recombination
- Rearrangement

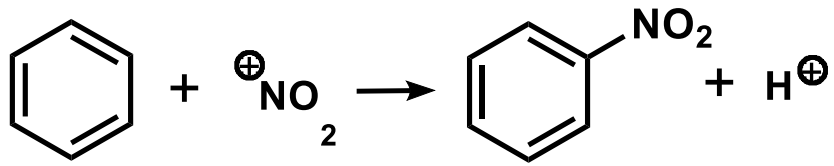
Complex Mechanisms:

- Insertion
- Addition
- Elimination
- Substitution**

*



**



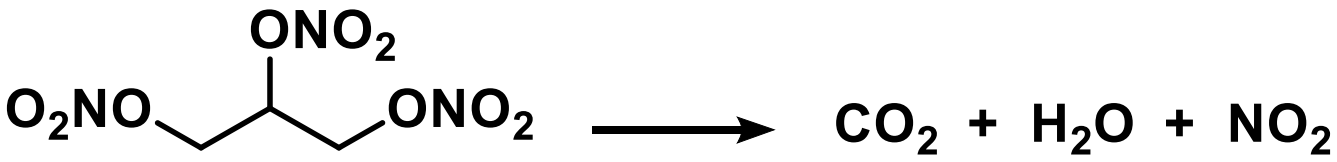
- **Molecular weight increasing in each step (minimum protections, co-products with low MW)**
- **Limited or absent reaction media (high productivity)**

Intrinsic Atom Economy in Reactions.

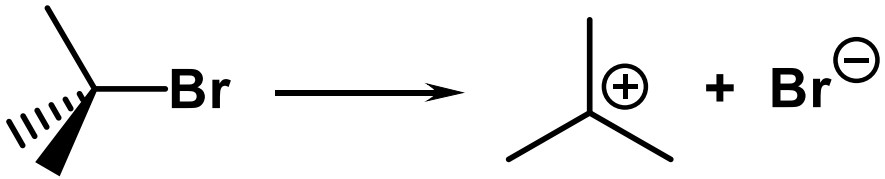
Reaction Type	*Waste Generation Potential*
Addition Reaction Isobutylene + methanol → methyl tert-butyl ether $C_4H_8 + CH_3OH \rightarrow (C_4H_9)\text{-}O\text{-}CH_3$	• completely incorporate starting material into product
Substitution Reaction Phenol + ammonia → analine + water $C_6H_5\text{-}OH + NH_3 \rightarrow C_6H_5\text{-}NH_2 + H_2O$	• stoichiometric amounts of waste are generated
Elimination Reaction Ethylbenzene → styrene + hydrogen $C_6H_5\text{-}C_2H_5 \rightarrow C_6H_5\text{-}C_2H_3 + H_2$	• stoichiometric amounts of waste are generated

Examples of Atom Economic Reactions	Examples of not-Economic Reactions
Rearrangements	Substitutions
Addition	Eliminations
Diels-Alder	Wittig
Other concerted reactions	Grignard

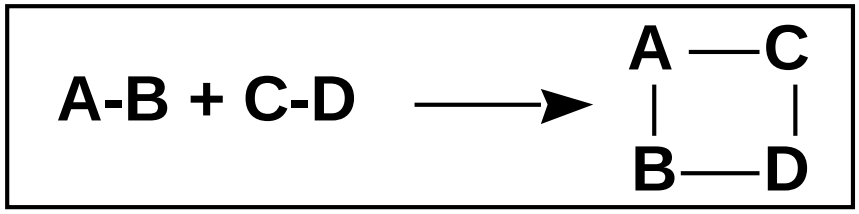
Decomposition Reaction (less than 100% AE).

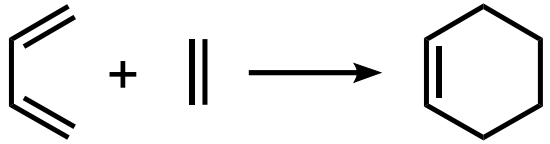


trinitroglicerine

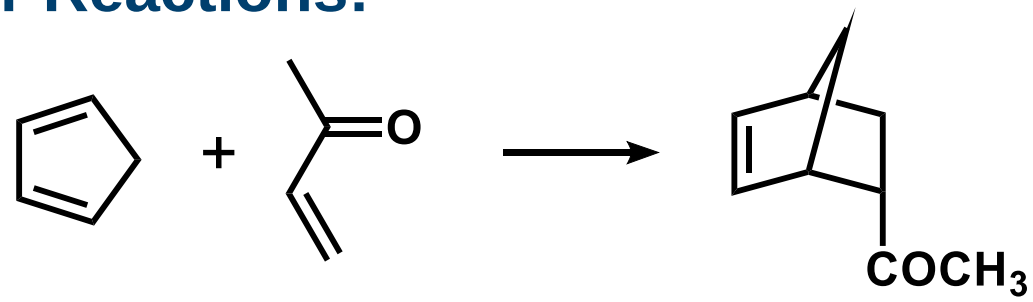


Pericyclic Reaction (Cycloaddition) (100% AE).



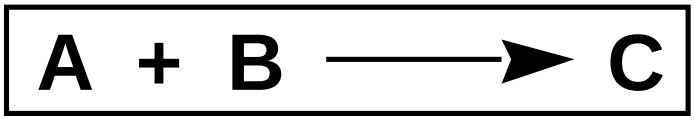


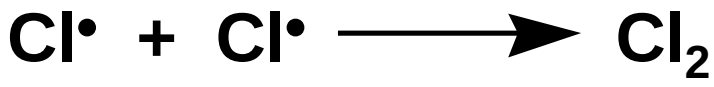
Diels-Alder Reactions:

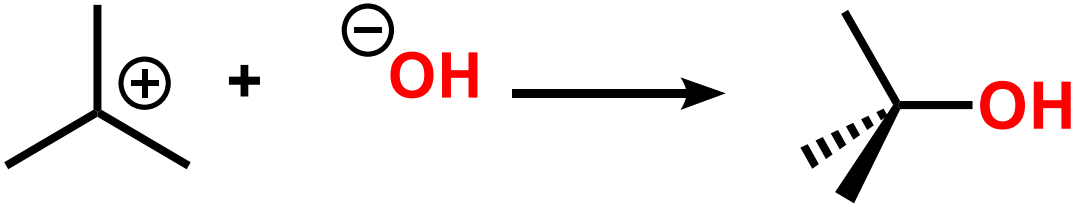


- Excellent method to form simultaneously two C-C bonds
- Concerted mechanism, therefore high regio- and stereo-selectivity
- Some D-A reactions can be carried out in water or ionic liquids, which can also act as catalyst.

Recombination Reaction (100 % AE).

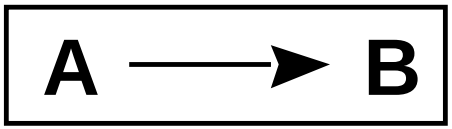


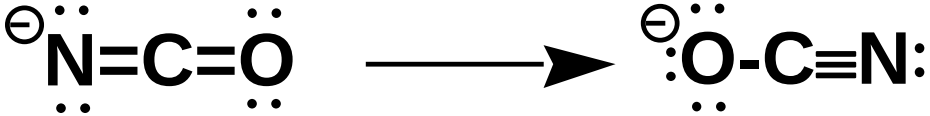




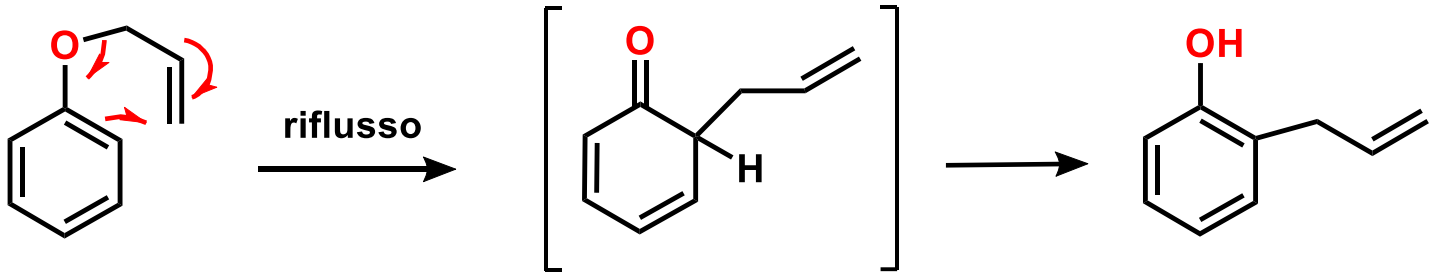
- Excellent method to form a single bond
- Bimolecular reaction, with low to high regio- and stereo-selectivity depending on the mechanism (radical low – ionic high)

Rearrangement Reactions (100% AE).





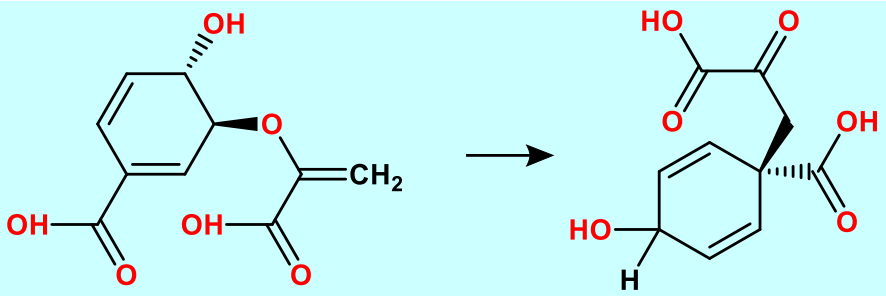
Several rearrangements (i.e. Claisen) occur by heating

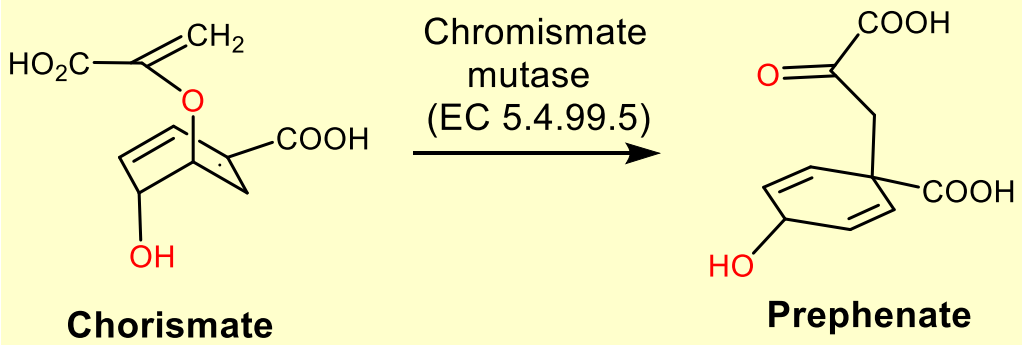


Excellent type of reaction : 100% AE

Which factors can influence the rate of these reactions?

Claisen Rearrangement and Solvent Polarity.



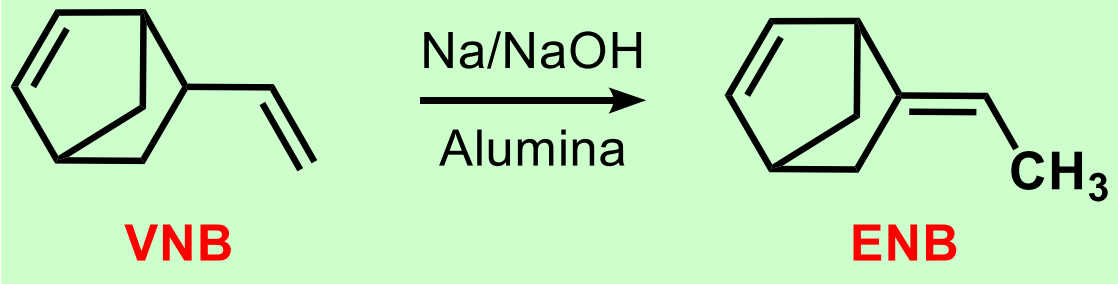


Reaction 100 time faster in water than in methanol - more polar solvent.

Copley, *J. Am. Chem. Soc*. 1987, 109, 2628

Catalyzed Isomerization (100% AE).

Some isomerizations are catalyzed by bases:



Industrial processes catalyzed by solid bases are relatively rare.

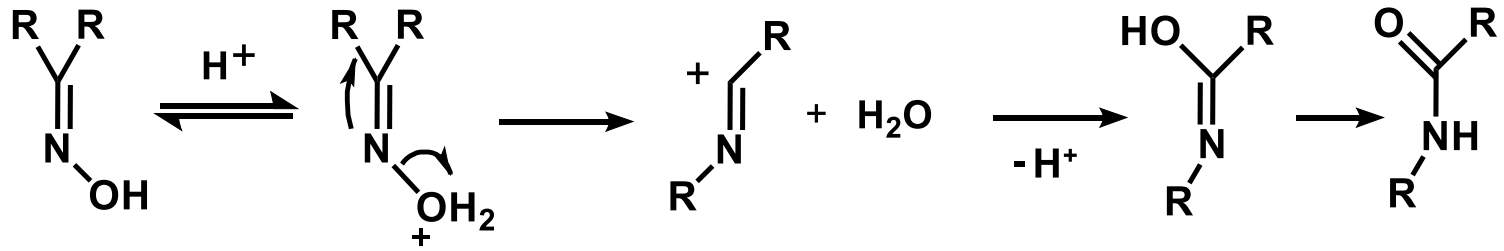
Sumitomo has developed a process to convert VNB to ENB.

This process eliminate the use of Na/K in liquid ammonia.

ENB is used as key component in EPDM rubber.

Rearrangement – Caution!

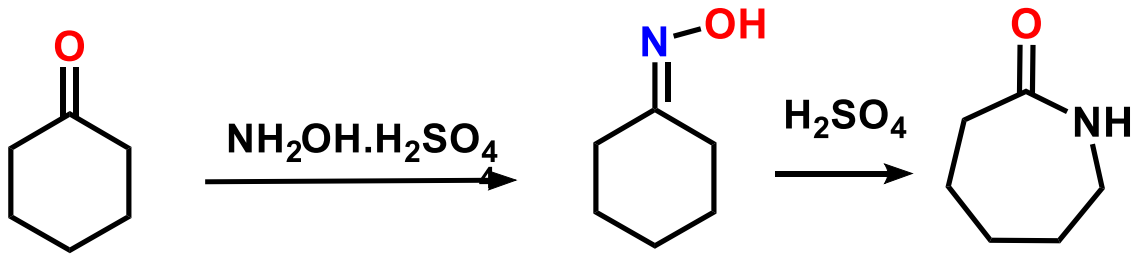
Beckmann Rearrangement



Several rearrangements, i.e. the Beckmann one, need acids to ‘catalyze’ the reactions - in some cases these compounds are used in very high volumes - ancillary reagents instead of catalysts.

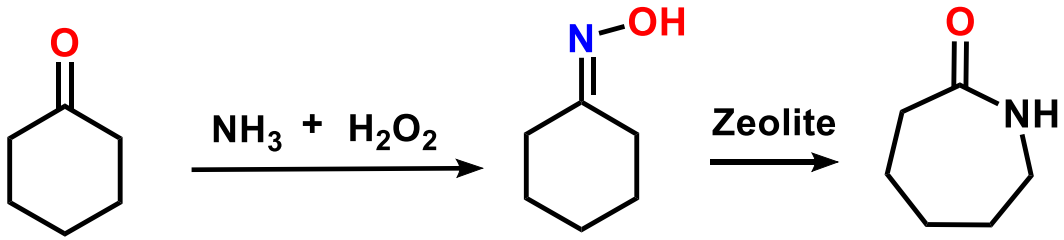
Caprolactam Route (Nylon© Intermediate).

Traditional:



- $(NH_4)_2SO_4$ in high amount
- E = 8 (kg waste per kg of product)

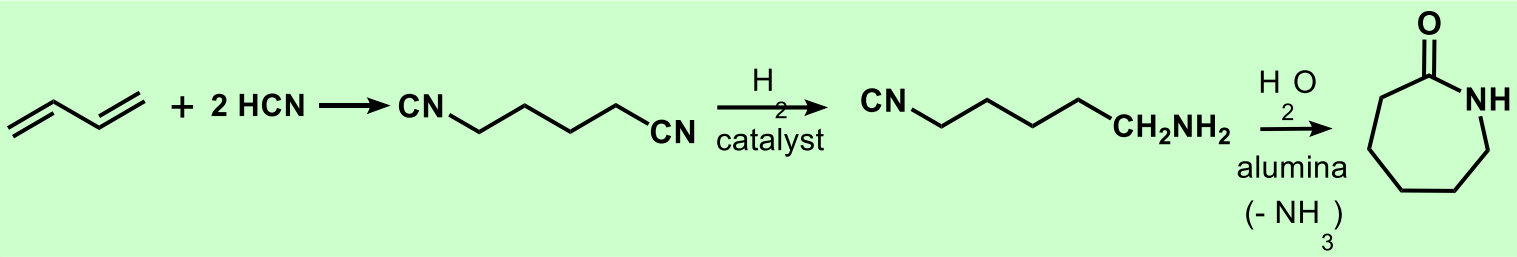
Improvement (use of heterogeneous nanostructured catalyst - zeolites):



- minimum waste (water) and without salts
- E = 0.32

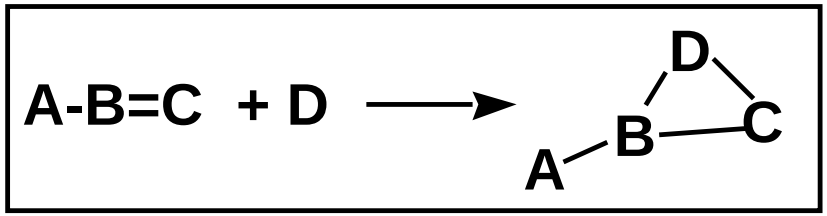
Alternative Route to Caprolactam.

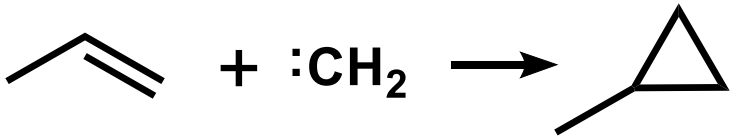
From Butadiene (alternative raw material) and Hydrogen Cyanide

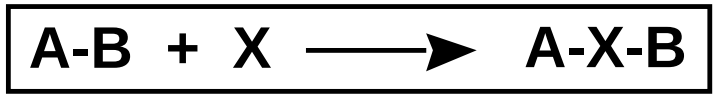


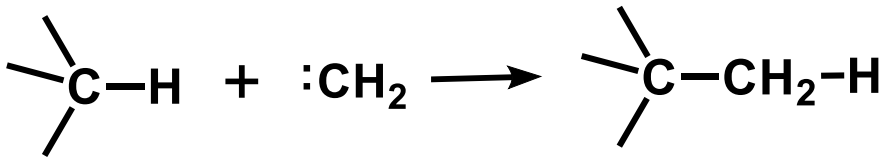
- minimal waste
- avoid the use of benzene raw material (via cyclohexane/cyclohexanone)
- without salts
- E = 0.13
- However HCN is very toxic!

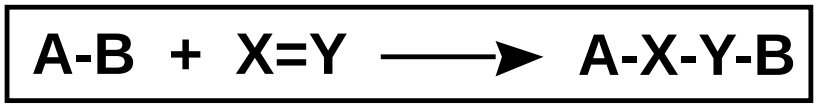
Insertion Reaction (100% atom economy).

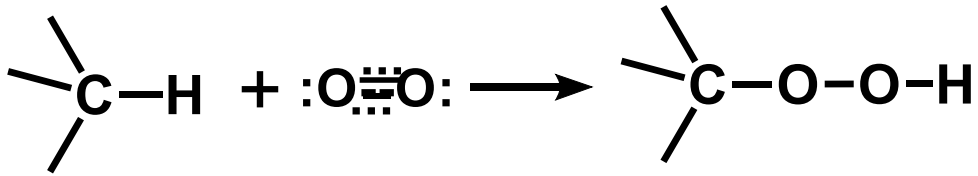




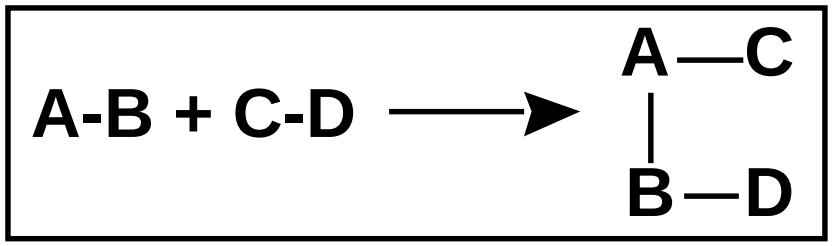


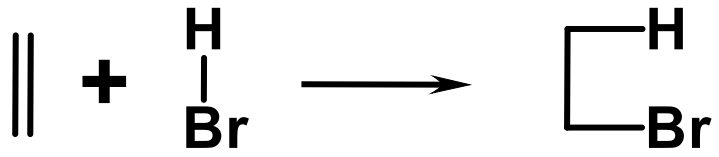




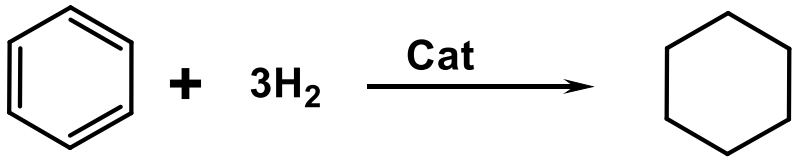


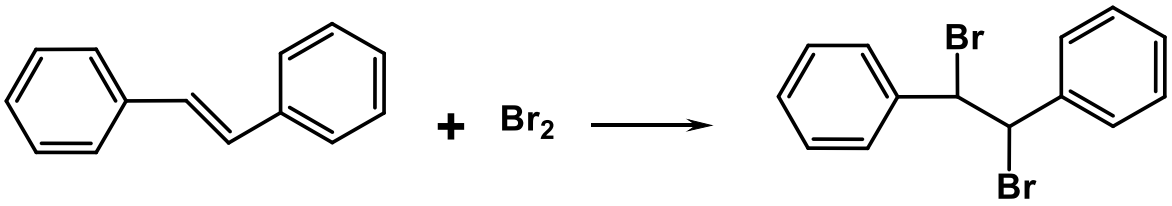
Addition Reaction (100% atom economy).



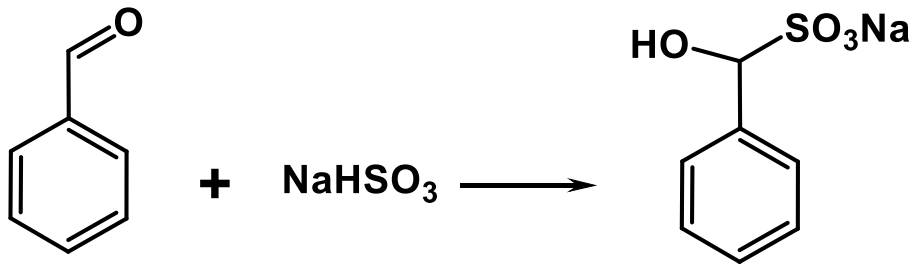


Examples:
Hydrogenation



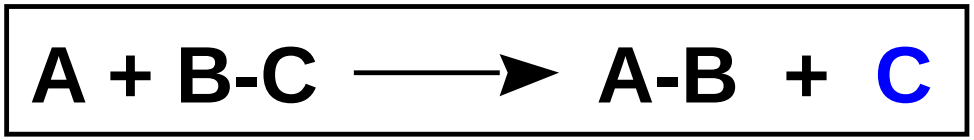


Electrophilic additions to alkenes

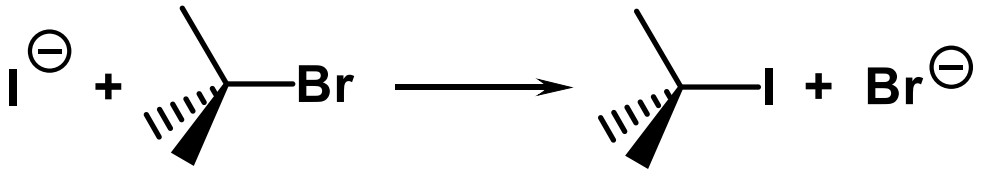


Nucleophilic addition to carbonyls

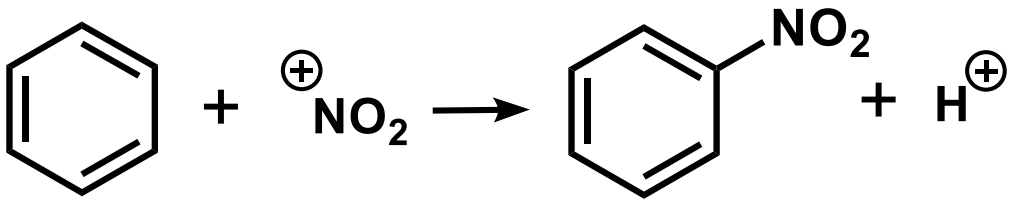
Substitution Reaction (lower than 100% AE).



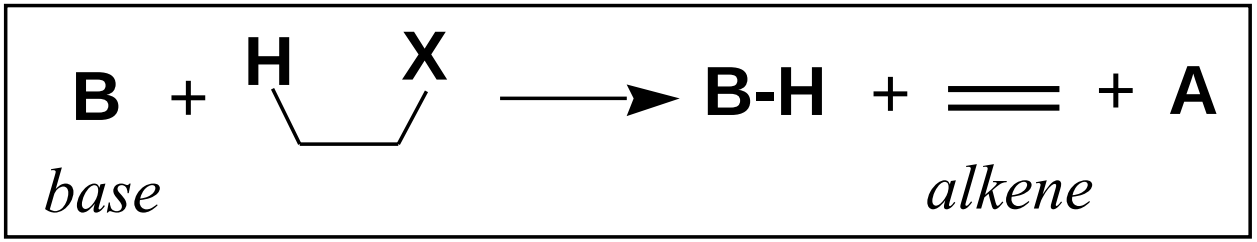
Aliphatic Nucleophilic Substitution

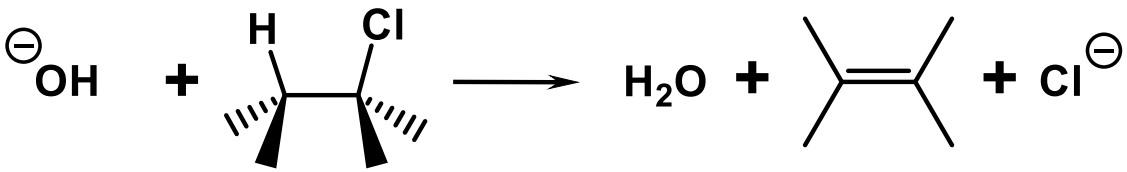


Aromatic Electrophilic Substitution



Elimination Reaction.





- Method with low AE but normally easy to carried out
- Several type of mechanism, variable selectivity, sometime very high.

POLITECNICO DI MILANO

School of Industrial and Information Engineering

Course 096125 (095857)

Introduction to Green and Sustainable Chemistry





Reaction Mechanism and Chemical Communication.

Prof. Attilio Citterio

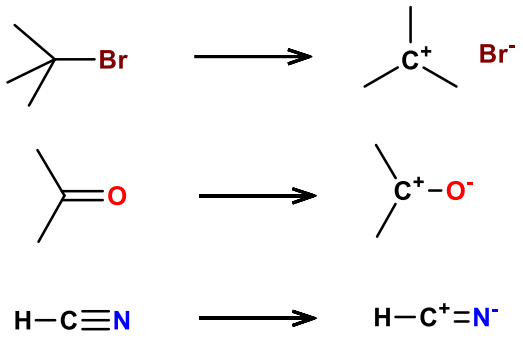
Dipartimento CMIC “Giulio Natta”

https://iscamapweb.chem.polimi.it/citterio/it/education/course-topics/

Electron Shift in Organic Chemistry.

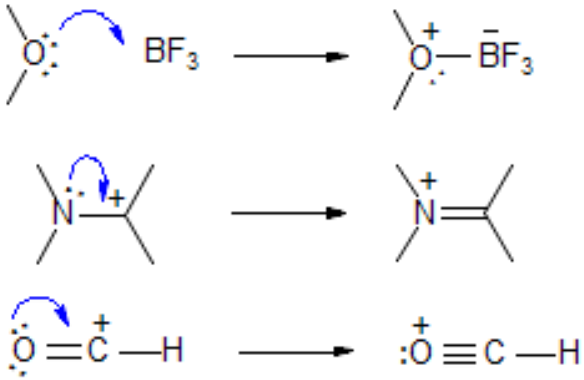
The movement of a pair of electrons from an electron-rich site (a lone pair of electrons or a bond) to a electron-poor substrate. Arrows of electron transfer are used as device "accounting« to trace easily the bonds and the formal charges in the interconversion of resonance structures or to represent reactions. Used correctly, they are very powerful not only to understand reactions, but also to foreseen the chemical reactivity. There are two fundamental types of electron shift in chemistry:

Bond breaking



The arrow starts from a bond electron pair and end on an atom, producing an atom with six electrons (or H^+) on the origin atom and a new electron pair on the terminal atom.

Bond Formation

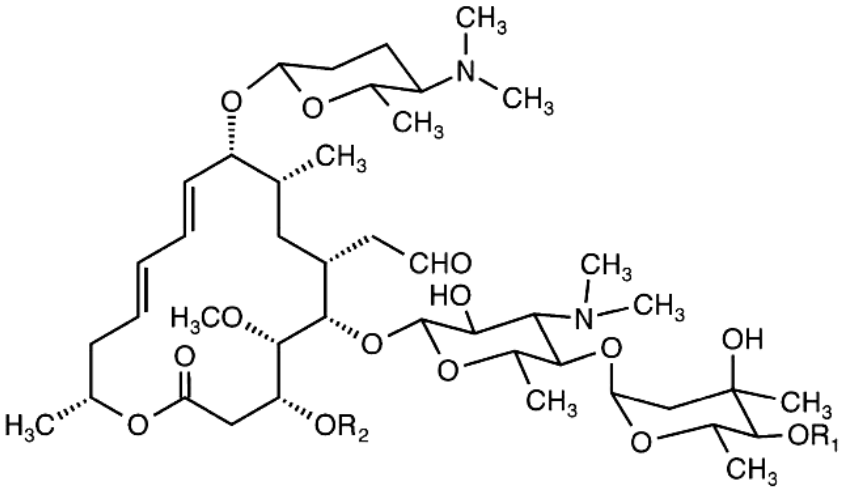


The arrow starts from an electron pair and must finish on a six electron atom (Lewis acid or H^+), forming a new single bond or converting a single bond into a double bond.

Some Elements of Chemical Communication.

- Reaction schemes
- Interpretation of a generic text (substituents R_1, R_2, X, etc.)

N atoms are stronger basic centers than all O atoms, which in turn are weakly basic (ethereal O are more basic than estereal O and O in carbonyl C=O.

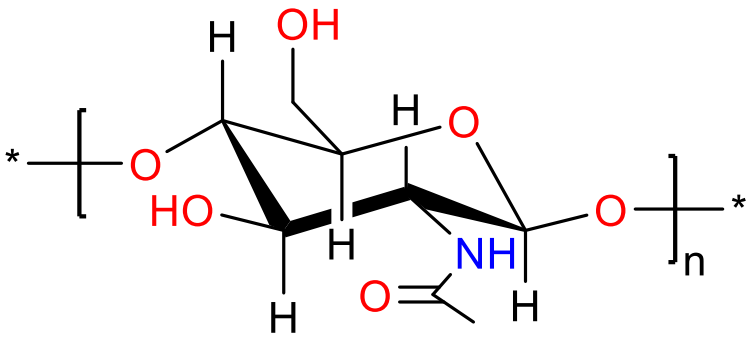


	R_1	R_2
ISV-SPM I	$COCH_2CH(CH_3)_2$	H
ISV-SPM II	$COCH_2CH(CH_3)_2$	$COCH_3$
ISV-SPM III	$COCH_2CH(CH_3)_2$	COC_2H_5
SPM I	H	H
SPM II	H	$COCH_3$
SPM III	H	COC_2H_5

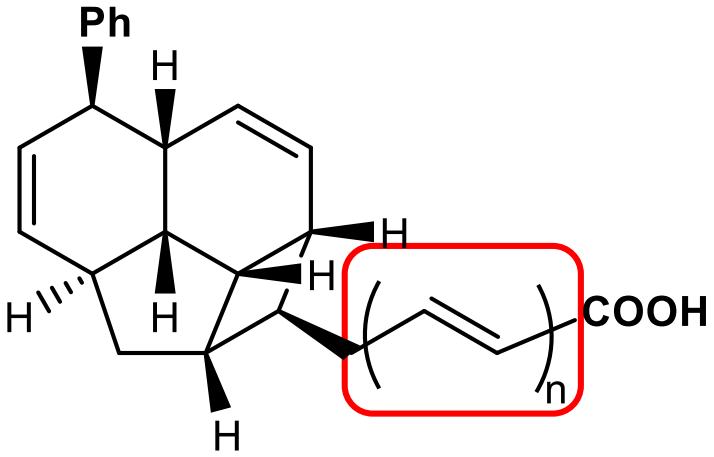
Natural Macrolide Family

Some Elements of Chemical Communication (2).

- Reaction schemes
- Interpretation a generic text
- Frequency change in Markush structures



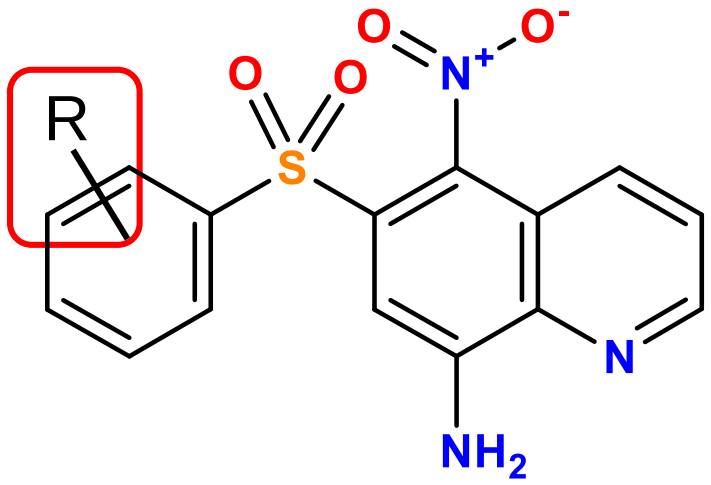
Numero di unità ripetitive (monomero) in una molecola ad alto PM (oligomero, polimero)
I numeri piccoli indicano la numerazione adottata per gli atomi



Numero di unità ripetitive come parte di una molecola a basso PM.

Some Elements of Chemical Communication (3).

- Reaction schemes
- Interpretation of a generic text
- Frequency variation in Markush structures
- Positional variation in Markush structures (i.e. generic substituents on aromatic ring)

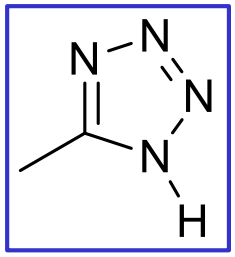


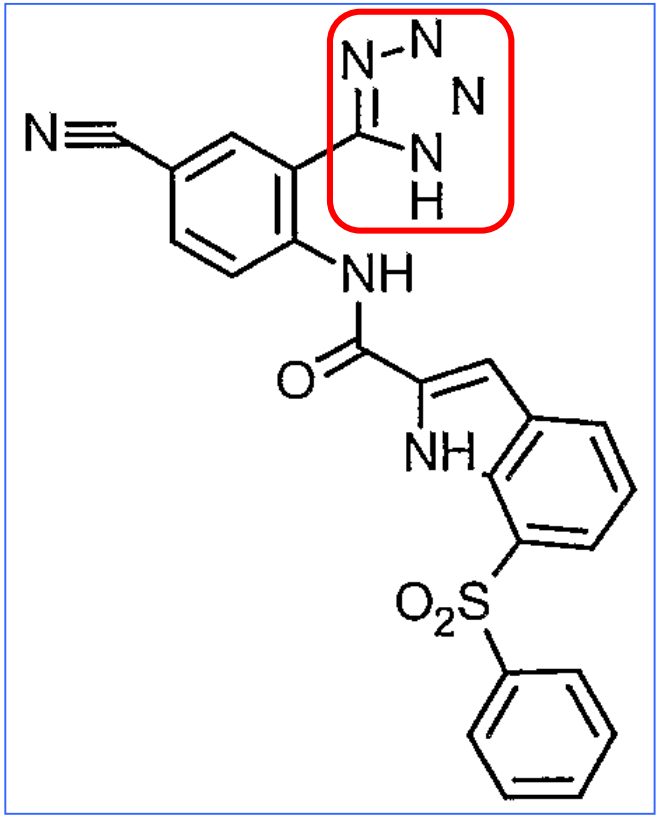
432; R = 4-NH_2
433; R = 2-F

Some Elements of Chemical Communication (4).

- Reaction schemes
- Interpretation of a generic text
- Frequency variation in Markush structures
- Positional variation in Markush structures (i.e. substituents)
- Other difficult situations (e.g. missing bonds between ring atoms)

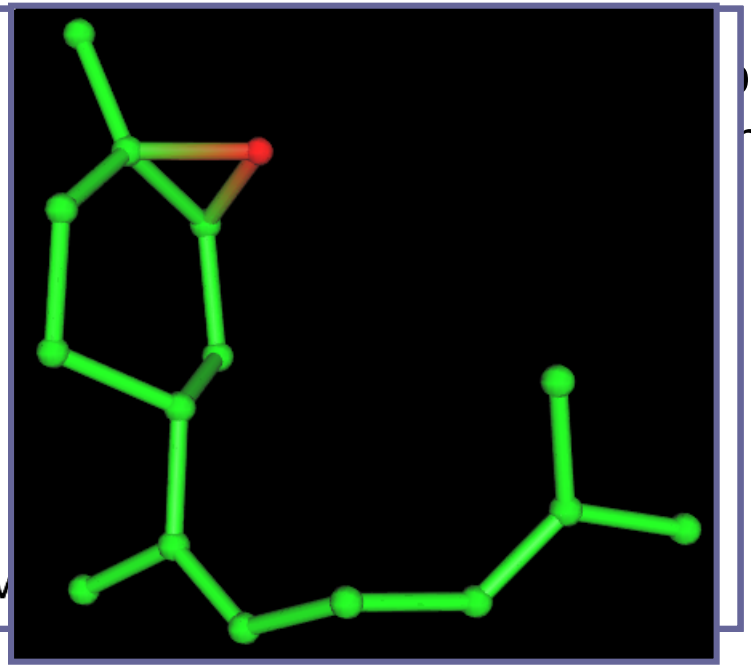
This is a notation problem easily solved with an appropriate character selection.





3-D Chemical Structures.

The tri-dimensio…om Ray-X or Neutron Diffraction or NM…hanical calculations.



1 Chemical image

2 Classification of connected components

3 Construction of dashed bonds

4 Vectorization

5 Construction of atom labels

6 Construction of connection table

2 Class… 6 … 4 … 3 …

…on the Hough transform method [4]

…a polygonal approximation method [5]

…exploiting the structural information 2D and converting the structure from 2D to 3D

Connecting lines to atoms

Grouping characters into atom labels

Recognition of superatoms

Classification by Reaction Names.

- **Chemists are familiar with Name Reactions (Diels-Alder, Michael, etc.)**
  - **Papers in a one issue of JOC (*22*, 2004) mentioned 20 name reactions, known and lesser known, some multiple times**
    - **e.g., Mitsunobu reaction, Nazarov reaction, Wolff rearrangement, etc.**
  - **Several books dealing exclusively with Name Reactions* (~ 700 reactions)**

- **Use of Name Reactions facilitates reaction retrieval**
  - **Complementary to other searches**
  - **Used in combination with other data**
  - **Easier alternative to formulating complex RSS queries**

- **Excellent browsing tool**
  - **Overview of scope and limitations of a given reaction, e.g. Aldol reaction**
  - **Combining different reaction types leading to same compound class**
    - **Hantzsch pyridine synthesis from dihydropyridines or ß-keto esters**
    - **Fischer Indole synthesis from hydrazines or hydrazones**
    - **Darzens reaction of epoxides from esters, amides, sulfones, or nitriles**

*References:

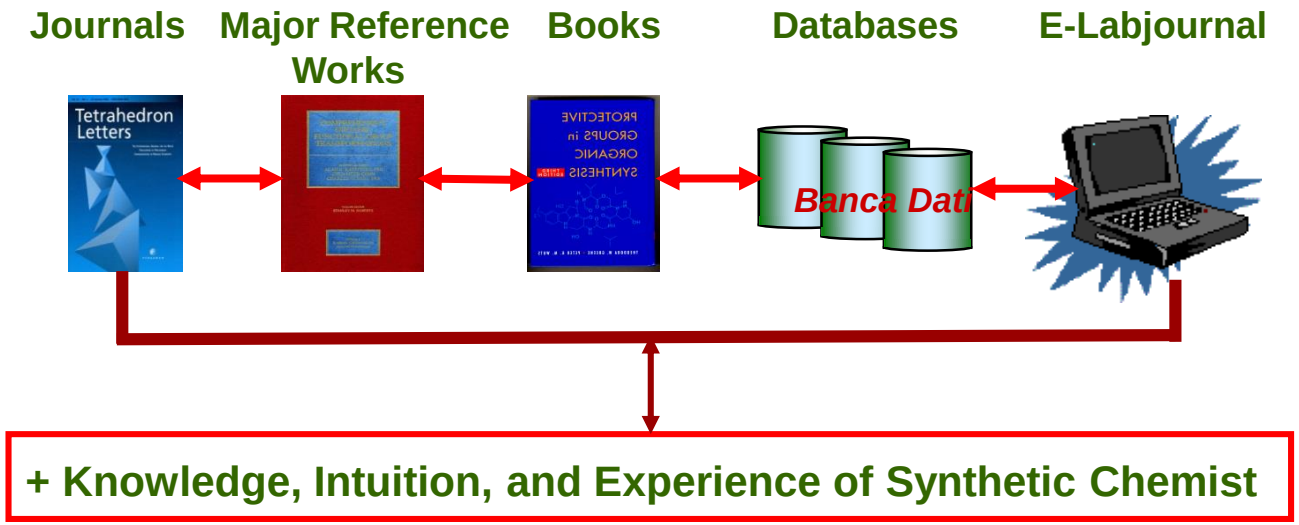
Named Organic Reactions, Laue, T. and Plagens, A., Eds., John Wiley &Sons, 1st Edition **1999,** 2nd Edition **2005**
Organic Syntheses Based on Name Reactions, Hassner, A. and Stumer,C., Eds., Elsevier Science,1st Edition **1994;** 2nd Edition **2002**
Name Reactions, Li, J. J., Ed., Springer, **2002**
Strategic Applications of Named Reactions, Kürti, L. and Czakó, B., Eds., Elsevier, **2005**
Name Reactions and Reagents in Organic Synthesis, Mundy, B.P; Ellerd, M.G. and Favaloro, F.G., Jr. Wiley Interscience **2005**

Use of Chemical Information in Organic Synthesis.

- **"The design of organic syntheses by chemists without the help of computers proceeds in anything but a systematic stepwise manner from the target molecule to available starting materials. A systematic stepwise approach is more the exception than the rule".**
- **"The human mind solves problems by lateral thinking, jumping from one idea to the next, from one question to a different one, from retrosynthetic thinking to considering the course and outcome of a reaction, etc."**

Gasteiger, J.; Ihlenfeldt, W.D.; Roese, P. *Recl. Trav. Chim. Pays-Bas* 1992, 111, 270.

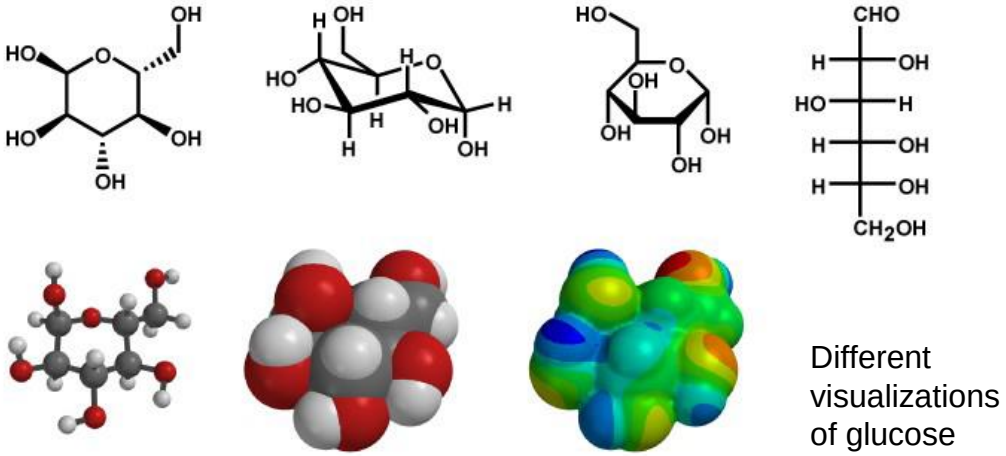
The paradigm in an ideal electronic world:



Molecular Visualization.

Visualizations are symbolic constructions used to codify information in order to make it meaningful to learners. Visualizations (graphs, tables, illustrations, animations) are valuable in educational settings because they “help make complex information accessible and cognitively tractable”, and “help us think in visual rather than abstract, symbolic terms”. In chemistry education, where the principle actors cannot be seen by eye, visualizations are of even greater significance.

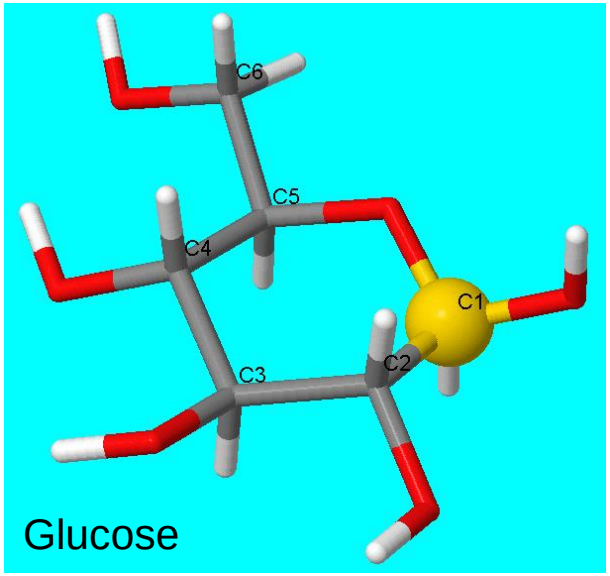
The ability to switch rapidly between the representational levels and relate the complementary information that they offer can provide a Deeper understanding of chemical reactivity.

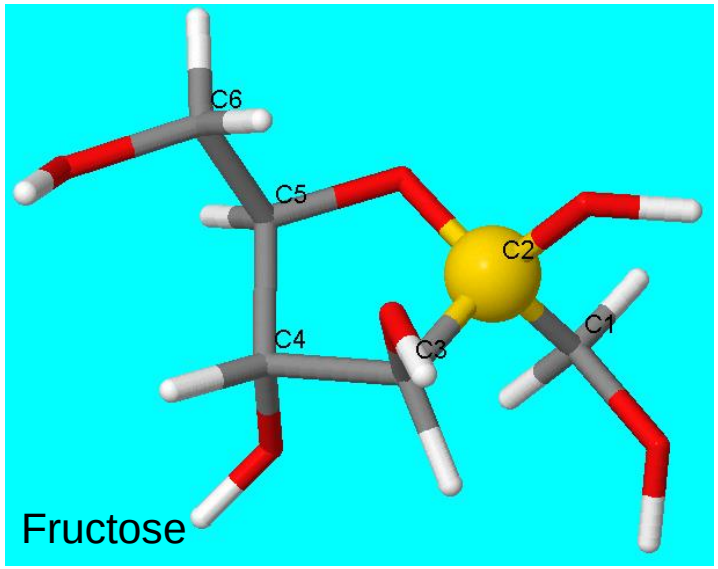


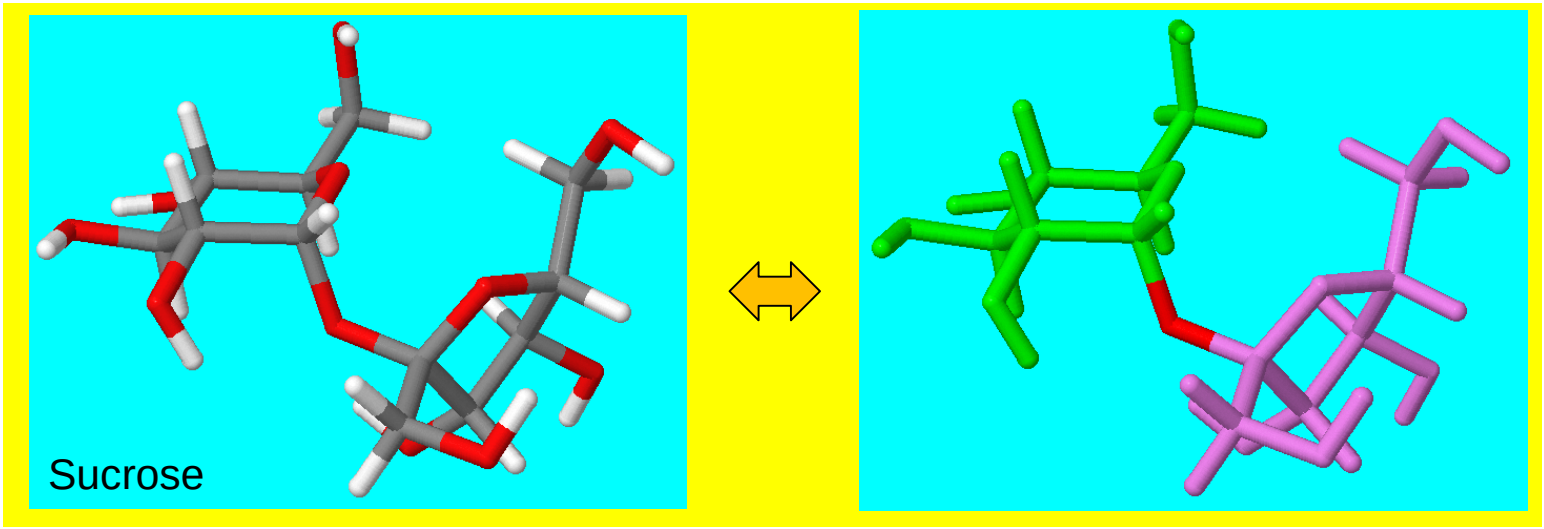
Different visualizations of glucose molecule

For a map of visualization methods see
http://www.visual-literacy.org/periodic_table/periodic_table.html

Highlighting of Portion of a Molecule. - Ex. Sugars: Glucose, Fructose and Sucrose.

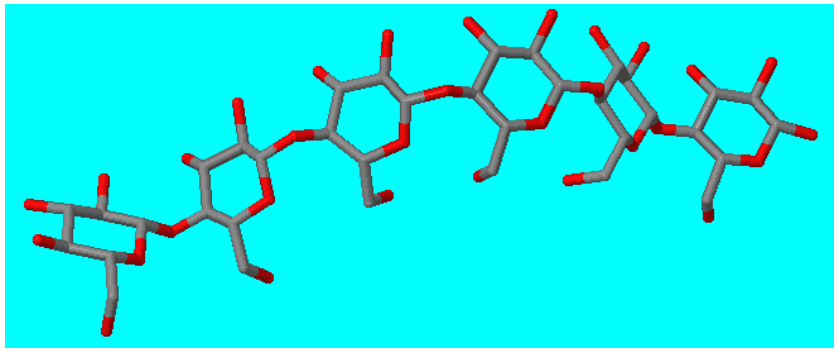


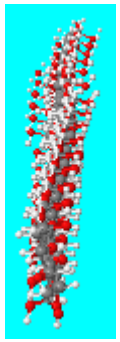


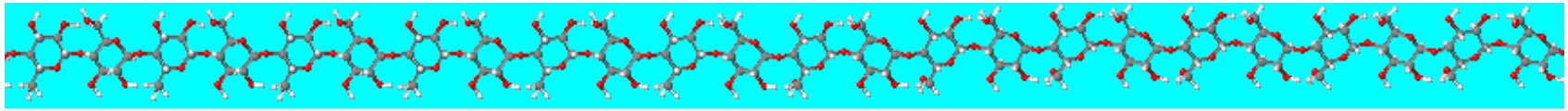


Spatial Visualization: Cellulose.

- It's the major polysaccharide in wood and in structural parts of plants, such as stems and leaves.
- Its structure is formed by the linkage of D-glucose molecules through glycosidic beta(1-4) bonds. Observe its linear structure and the flat non symmetric lateral view, which explain its tendency to organize in fibrils.





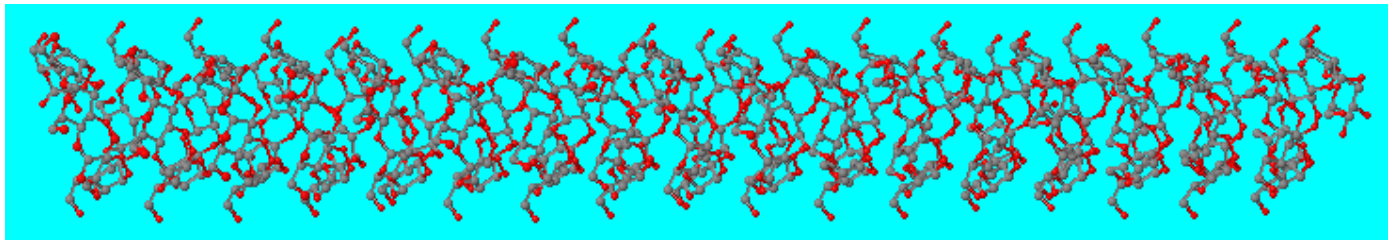


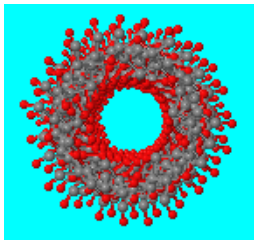
Spatial Visualization: Starch.

This polysaccharide, with reserve function in plants, is abundant in all starches, such as potato, rice, or wheat.

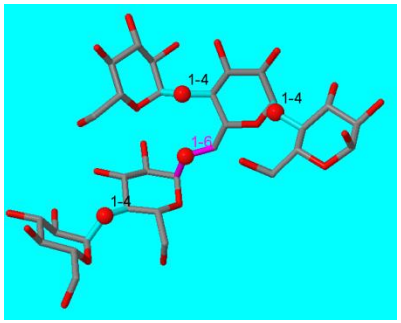
On its structure we can distinguish two types of components:

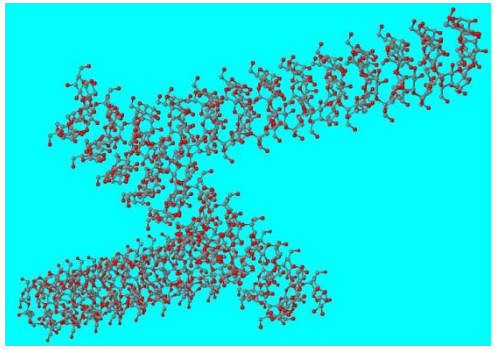
- **Amylose**, a linear polysaccharide formed also by **D-glucose** molecules, but in this case linked by glycosidic **alpha**(1-4) bonds. It form a coil, hence the molecule is very different from cellulose:





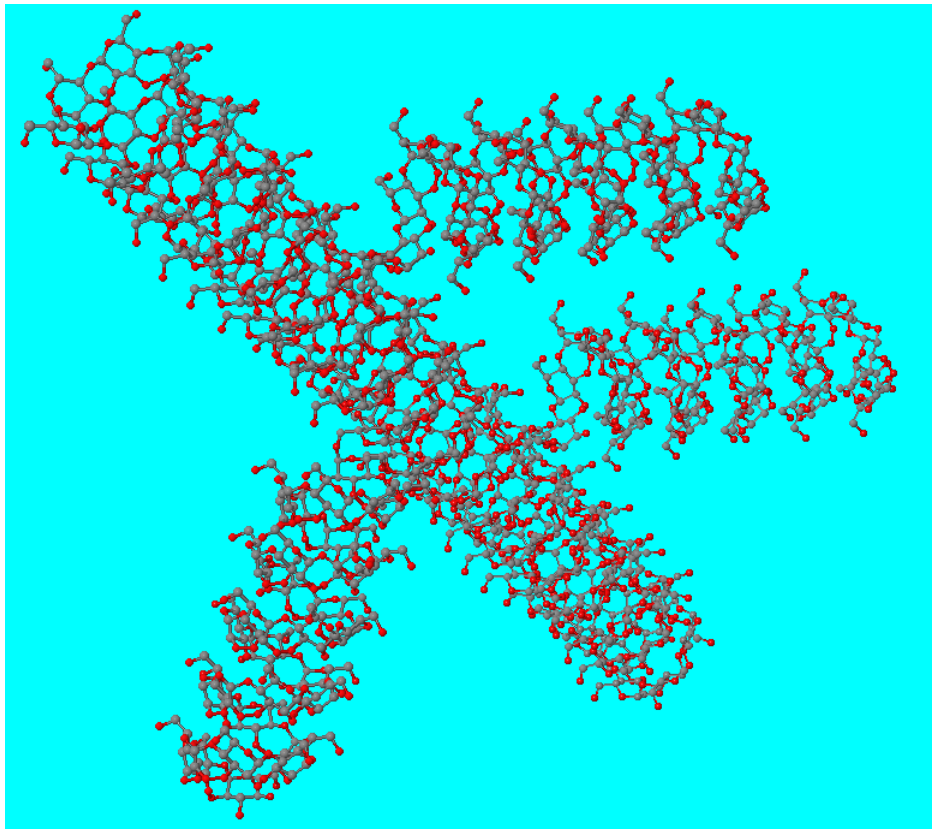
- **Amylopectin**, whose chain is similar but branched, due to **alpha(1-6)** bonds in addition to alpha(1-4) bonds.





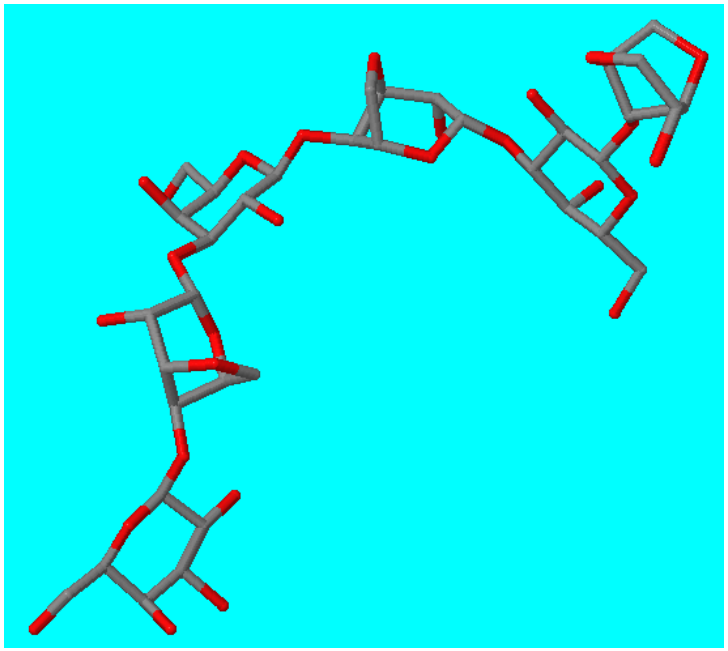
Glycogen.

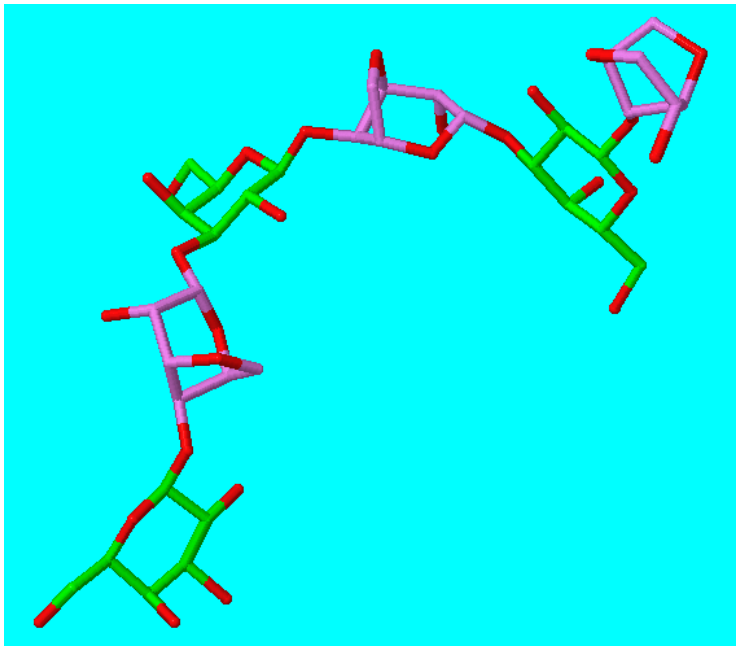
- Is the reserve polysaccharide in animals; in humans it accumulates specially in the liver.
- Its structure is identical to that of amylopectin, except that the branches are more frequent



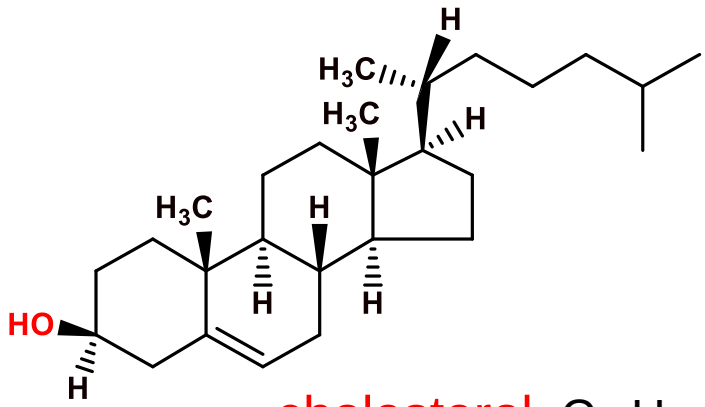
Agarose.

- It's a polysaccharide synthesized by some algae, that forms part of agar-agar (an ingredient for culture media and gel).
- Its structure is formed by the alternating linkage of beta-D-galactose and 3,6-anhydro-alpha-L-galactose, by glycosidic beta(1-4) and alpha(1-3) bonds.

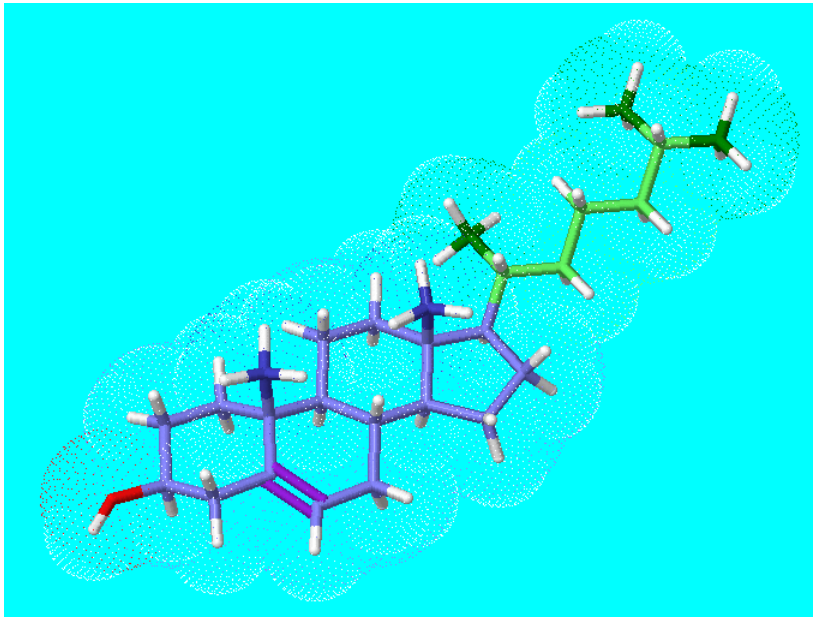




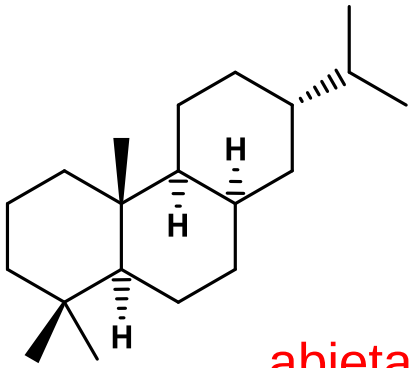
Steroids and Terpenes: Cholesterol and Abietane.



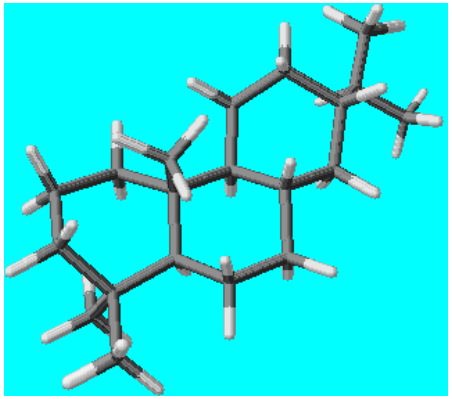
cholesterol $C_{27}H_{46}O$



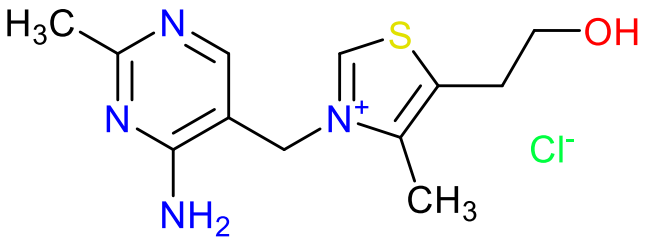
Telomers of C_5 units



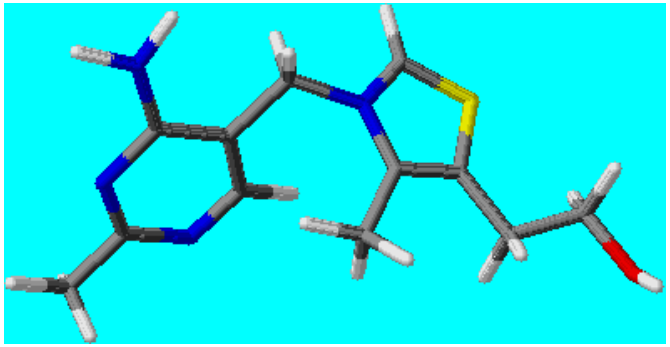
abietane $C_{20}H_{36}$

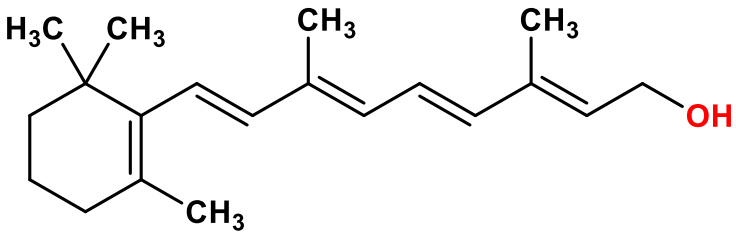


Vitamins.

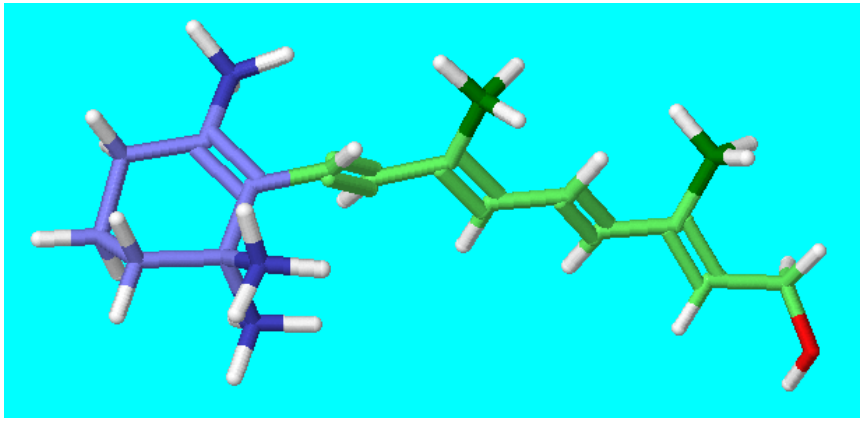


Vitamin B_1
Thiamine chloride





Vitamin A



RNA.

In contrast to DNA, RNA is formed by a single *strand*. This adopts different structures depending on its nucleoside sequence (or base sequence). Here we will see just an example, the transfer RNA, or tRNA. The fictitious line joining phosphorus atoms (*backbone*) allows us to observe how the chain folds.