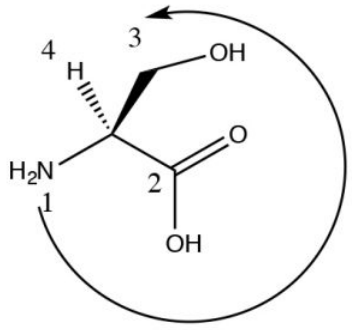
3.3: Configurations of Chiral Molecules- the Cahn-Ingold-Prelog Convention

Since we cannot distinguish between enantiomers using the naming conventions that we have considered so far, we have to invoke a new convention to unambiguously specify the arrangement of bonds around a chiral center which is known as the **configuration** (not to be confused with conformation). The most common naming strategy used is referred to as the Cahn-Ingold-Prelog Convention: a set of rules that assigns a configuration (known as R or S) to a specific chiral (or stereogenic) center. An important fact to remember is that the configuration **assigned** to a molecule in this way has nothing to do with the molecule's **observed** optical rotation.

Cahn-Ingold-Prelog Convention

1. Look at each atom **directly** connected to a chiral carbon and rank by **atomic number** (Z); **highest first** (e.g. $\mathrm{O} > \mathrm{N} > \mathrm{C}$). For isotopes the higher atomic mass receives a higher priority (e.g. $\mathrm{D} > \mathrm{H}$).
2. If you can't make a decision here, go out to the next atom in each chain—and so on until you get to the **first point of difference**. So for example, $-\mathrm{CH_2CH_2OH}$ takes priority over $-\mathrm{CH_2CH_2CH_2OH}$.
3. Multiple bonds are equivalent to the same number of single-bonded atoms. For example, $\mathrm{CH_2{=}\,CH_2}$ takes priority over $-\mathrm{CH_2CH_3}$ because it is counted as if it were $\mathrm{CH_2(CH_3)_2}$
4. Now place your eye so that you are looking down the bond from carbon to the lowest ranking group.
   - If the sequence, high→low, is clockwise: **R (Rectus).**
   - If the sequence, high→low, is counterclockwise: **S (Sinister**).

The skill of assigning configurations is one that takes practice and it is necessary to understand what the rules mean. Some **common errors** that people make are:



a. Looking at the substituent as a whole rather than looking only at the atom (or the first atom that is different). So for example, $\mathrm{O} > \mathrm{N} > \mathrm{C}$, but also $-\mathrm{OH} > -\mathrm{CO_2H}$ even though $\mathrm{CO_2H}$ is overall larger, it is attached by a C which is a lower priority than an O. Consider the example of this stereoisomer of the amino acid serine (It's the L isomer but we will get to that in a while). The highest ranking group is the N (since its atomic number is higher than C), next is the $\mathrm{CO_2H}$ because there are an equivalent of three $\mathrm{C-O}$ bonds on that carbon and only one $\mathrm{C-O}$ bond on the $\mathrm{CH_2OH}$ group. Since the H is pointing back, we can now directly see that $1 \rightarrow 2 \rightarrow 3$ is counterclockwise and the configuration is S.
b. Forgetting to look down the bond from the C to the lowest ranking group (often, but not always, an H). If you look down any other bond (or from the low ranking atom to the carbon) then you will almost certainly get the wrong assignment.

If the molecule is not drawn with the lowest ranking group pointing away, there are a number of options:

1. Particularly when you are first getting started, you should MAKE A MODEL. There is research evidence that shows using an actual, physical model is the best way to learn this particular skill. Be careful to construct the model so that it looks exactly like the drawing. Then you can physically rotate it and identify the configuration.
   or
2. You can switch two groups—either redrawing or in your imagination (we strongly suggest you redraw it), so that the lowest group is now pointing back out of the plane. Then the configuration you get will be the opposite of the actual configuration.
   or
3. You could imagine yourself moving in space to look down the bond from C to the lowest ranking group. This is quite difficult for some people.

or

4. You could rotate the molecule in your head so that you are looking down the bond from C to the lowest ranking group. This is also quite difficult for some people.

You should always use at least two methods – and make sure you get the same answer both times!

After a while, this skill will come more easily to you.

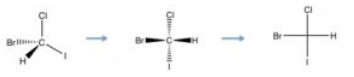
R and S and D and L isomers:

As you probably already know, most biological molecules are chiral. For example, all of the naturally-occurring amino acids (aside from glycine) have a chiral carbon (see above), sugars have several chiral centers, and so the large molecules made up from these smaller monomers (such as nucleic acids, carbohydrates, lipids, polypeptides, and a range of smaller molecules) contain one or more (sometimes many) chiral carbons. For historical reasons, however, most simple biological molecules (that is the monomers from which polymers are constructed) are referred to as either D or L isomers, rather than R or S. The convention for naming substances as D or L dates back to the early 1900s and has to do with a compound's similarity to glyceraldehyde which exists as a pair of enantiomers. The enantiomer that rotated light to the right (+) was identified as D, and the other one (−) as L. For example, the naturally-occurring amino acids are the L-isomer, and many of the simple sugars are the D isomer. It should be noted that a compound such as glucose can be identified as D-glucose, even though it contains 5 chiral centers. The D/L nomenclature is applied to the molecule as a whole, whereas the R/S nomenclature applies to a specific chiral carbon. The D/L nomenclature does not, in fact, predict the direction of rotation of polarized light in most larger molecules—but it is a much simpler way of naming compounds with more than one chiral center.

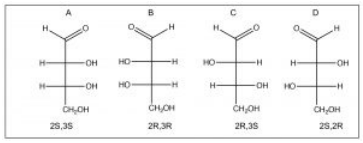
Speaking of which…

Molecules with two or more chiral (stereogenic) centers

Molecules that have more than one chiral center bring in another level of complexity. It can be quite difficult to draw molecules accurately that have more than one chiral center, and therefore we often turn to a representation that is more stylized than the wedge-dash; this is known as the Fischer projection, named after Emil Fischer (1852-1919) who elucidated the structures of sugars (and who also introduced the D/L nomenclature). Fischer projections are written as vertical and horizontal lines. The carbon backbone is the vertical line, and the other substituents are horizontal. We assume that the horizontal bonds are coming out towards you (↓). This convention makes it easier to draw and to assign configuration to the chiral centers (this structure is S).



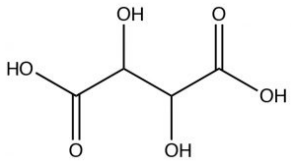
Now if we look at Fischer projections for a four carbon sugar that has two chiral centers, we see that there can be several possibilities.



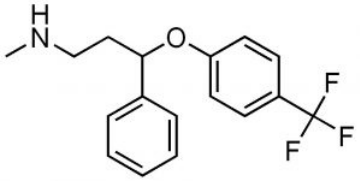
Numbering from the aldehyde carbon as carbon 1, both C_2 and C_3 are chiral. We also see that there are two pairs of enantiomers: A and B are mirror images of each other and so are C and D. In general, for a molecule with n chiral centers, there are a possibility of 2^n stereoisomers. When we look at the relationship between isomer A and isomer C (or D) we note that they are not mirror images of each other (the chiral centers are the same at one carbon and different at the other). These compounds are known as **diastereomers**, that is, stereoisomers that are not mirror images of one another; they do not have identical distances between all of their atoms. Because they have different arrangements in space, diastereomers have different properties, both physical and chemical, and can be separated. In fact, the two compounds have different names: A and B are called erythrose while C and D are called threose. Diastereomers actually belong to the same class of stereoisomers as cis/trans ring compounds (and alkenes), they

have the same molecular formula and connectivity, but have different arrangements in space and cannot be interconverted by bond rotations. In contrast, enantiomers have the same atomic arrangements in space, but cannot be superimposed on one another.

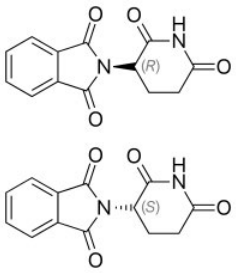
So now you might well ask yourself, given that stereoisomers have the same physical properties: how could you possibility separate enantiomers? The existence of enantiomers was first observed by Louis Pasteur (1882-1895), who was actually able to identify different crystalline forms of tartaric acid[3] (actually, the potassium salt), which turned out to be the two enantiomers (but not the meso isomer), which appeared as tiny crystals that were mirror images of each other. In fact, this selective recrystallization is highly unusual, and it is normally not possible to selectively recrystallize isomers.



So why does the stereo-asymmetry of a molecule matter? The answer is that organisms (living systems) are—apparently due to an accident of their origin and subsequent evolution—asymmetric at the molecular level. For example, polypeptides and proteins are polymers of amino acids. While organisms can make and use both D- and L-form amino acids, all of the amino acids used in polypeptide/protein synthesis (with the exception of glycine, which is achiral) exist in L- and D- forms. Only L-form amino acids are used in polypeptides and proteins. The use of only L-form amino acids means that each polypeptide/protein has a distinct three-dimensional shape that influence how other molecules bind to it. For example, a particular drug can be designed to inhibit a particular enzyme (protein-based catalyst).



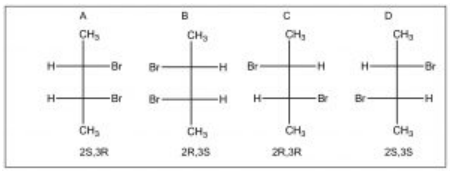
If the drug is itself chiral, then it is extremely likely that one or the other of its chiral forms will be an effective inhibitor while the others will not. In fact, it is common for drug companies to first patent the racemic mixture of a drug, and then later the purified enantiomers. For example, Fluoxetine (Prozac)(→) is a racemic mixture of the two enantiomers (can you find the chiral center?). In the human body, the two forms of fluoxetine are metabolized into corresponding forms of norfluoxetine, one of which is significantly more active than the other.[4] Another example of the use of a racemic drug is the terrible story of thalidomide, which was marketed in the 1960's in Europe and Canada as a drug for morning sickness.



The original drug was administered as the racemic mixture but we know that it is the R isomer that is the sedative. The S isomer is teratogenic. That is, it causes birth defects in this case associated with the development of limbs. Many children were born without limbs before Thalidomide was removed from the market. Now we know that even the pure R enantiomer racemizes at physiological pH. In fact, Thalidomide was re-introduced to the market because it is one of the few drugs that can be used to treat leprosy.

Non-chiral species with chiral centers

Just because a compound has a chiral (stereogenic) center, doesn't mean the actual compound is optically active. Remember that the requirement for a chiral substance is that there is no symmetry element. Consider the set of compounds below.



There appear to be four stereoisomers here, but, in fact, there are only three. Isomers A and B are identical due to the a mirror plane through the center of the molecule (bisecting the $C_2 - C_3$ bond), the top half of the molecule is identical to the bottom half. Compounds like this are called **meso** isomers. The meso isomer is a diastereomer of the pair of enantiomers C and D.

Questions to Answer

- Without looking at the diagram above, construct your own representation and explanation for the various types of isomers that we have encountered. Give an example of each type of isomerism (using two such isomers) and explain how and why they differ from each other and from other types of isomers.
- How is it possible that compounds with chiral carbons are not themselves chiral?
- Do you think it is possible to have a chiral compound that does not have chiral carbons in it? What structural features would you look for? Why?

Questions to Ponder

- Most biomolecules are chiral, how do you think they got that way?
- Most biomolecules are chiral, what are the implications for the synthesis of pharmacologically active compounds?