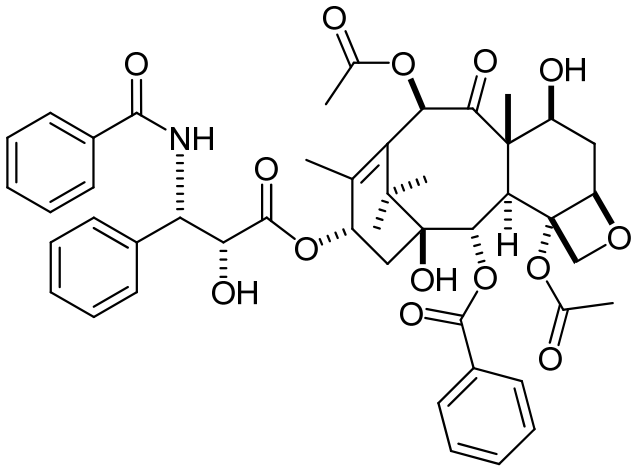
CHM508

Chimie des Substances Naturelles



Notes de cours du Professeur Claude Spino

avec la participation du Professeur Jean Lessard

Département de Chimie

Université de Sherbrooke

1995-2017

Table des matières

Introduction.

The term "natural product" is one that is often misused and misunderstood. A natural product is a compound found in nature regardless of its natural source and of its structure. Thus, sodium chloride is as much a natural product as is cholesterol, and the contraceptive pill is as "synthetic" as is nylon 66. In addition, a man made natural product is still a natural product and should not be called a "synthetic product". The term man-made natural product applies in this case. The term "natural product" should never be used to mean "healthy" or "good", nor should the term "synthetic" be associated with "toxic" or "unhealthy", etc.

In the chemical world, the term "natural product" is often associated with organic chemistry because that is where the term was first proposed and is most often used. In addition, within the context of organic chemistry, the term is loosely used to refer to life metabolites (vide infra). Because of its wide acceptance, that meaning will be retained throughout this course.

Natural products are broken down into two main categories:

PRIMARY METABOLITES

(AMINO ACIDS, PROTEINS, CARBOHYDRATES, DNA, ETC.)

Essential and similar for all forms of life. It contains the blue print of life.

↓

CELLULAR LIFE

↓

SECONDARY METABOLITES

(CHEMICALS WITH PARTICULAR BUT SECONDARY FUNCTIONS: HORMONES, NEUROTRANSMITORS, VITAMINS, ETC.)

Wide variety differing according to species. **Made from the primary metabolites** via only a handful of different pathways.

It is the second category, the secondary metabolites, that will be the subject of our study. Life fabricates these secondary metabolites for specific purposes: defense, communication, sexual functions etc. Natural products are classified according to four main areas:

-Chemical Structure; classification according to broad structural features, e.g., aliphatic, aromatic, heterocyclic, etc.

-Physiological Activity; classification according to their role in biological systems, e.g., antibiotics, vitamins, hormones, antineoplastic, etc.

-Taxonomy; classification according to the source of the compound. Based on plant taxonomy, e.g., alkaloids, terpenoids, etc.

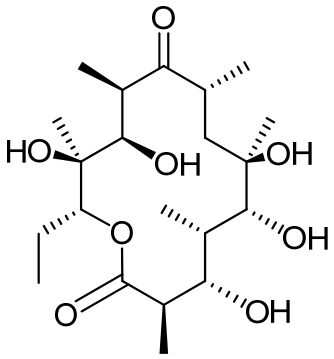
-Biogenesis; classification according to their chemical origin or biosynthesis. Usually the compounds are named after the mechanisms by which they are produced, e.g., isoprenoids, acetogenins (polyketides), shikimic acid pathway, etc.

Table 1 gives examples for each class with their subdivisions. Figure 1 illustrates some actual natural products. It is not rare for a compound to be classified in more than one category or subdivision because the different categories are independent from each other. For example geraniol is clearly an open chain aliphatic compound (class I.a.) but it is often included with other terpenoids and steroids under class III.b. Not all terpenoids (class III.b.) are hormones (class II.a.) and not all hormones are terpenoids. So one cannot simply treat hormones as a subdivision of terpenoids or terpenoids as a subdivision of hormones. Yet, **testosterone**, is a steroid (terpenoid) and a hormone, but other steroids are not hormones, like cholic acid (a bile acid). All terpenoids are isoprenoids, and all isoprenoids are terpenoids, however both words have a different meaning in terms of classification. Whereas the word terpenoids comes from plant taxonomy, the word isoprenoids refers to the "isoprene" biosynthetic pathway. It is necessary to keep an open mind about the classification of natural products. Usually, the context will obviate the classification area that is used to describe the natural products in question.

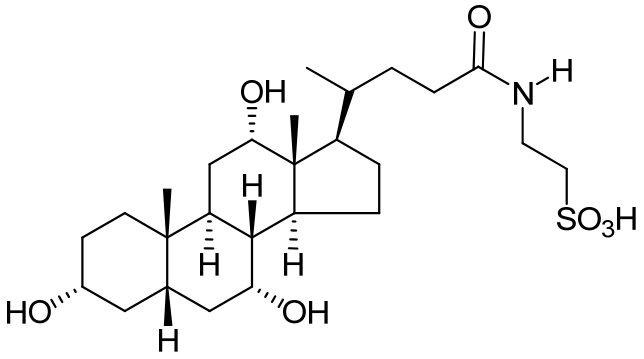
Table 1 The four main areas of classification of natural products.

I. Structure	**II. Physiology**	**III. Taxonomy**	**IV. Biogenesis**
I.a. Aliphatics Fatty acids, sugars, amino acids, some monoterpenes...	II.a. Hormones Some steroids, prostaglandins, neurotransmitters	III.a. Alkaloids Opium alkaloids, ergot alkaloids...	IV.a. Isoprenoids Terpenoids, steroids, carotenoids...
I.b. Cyclic Terpenoids, steroids, some indole alkaloids...	II.b. Vitamins Some steroids, aromatics, carotenoids...	III.b. Terpenoids Steroids, mono, di, and triterpenes...	IV.b. Polyketides Phenolics, fatty acids, polyethers...
I.c. Aromatics Phenolics, quinones...	II.c. Antibiotics Macrolides, polyethers, penicillins...		IV.c. Shikimic acid Some indole alkaloids, some aromatic amino acids...

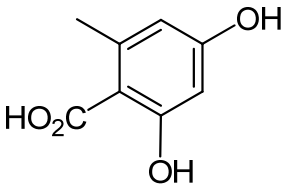
Natural products, whether collected from nature or synthesized in the laboratory, play a large and increasing role in our everyday lives. Besides the obvious role of keeping us alive (as do vitamins, hormones, etc.) natural products are used as therapeutic agents, as flavoring additives, as insecticides, for skin protection, and in many other applications. They were used ever since man learned to interact with its environment. Centuries ago, they were administered as plant or animal extracts, or in other forms, and they were used to treat diseases or as hallucinogens, etc. The effects of the extracts were known but the agent(s) responsible for the effect could not then be identified. Today, with the help of modern science and technology, chemists are able to identify the components of an extract and to assign a biological activity or other function to each component. This branch of chemistry is called "natural product isolation and structure elucidation". But the science doesn't stop there. Chemists are looking for ways to synthesize natural products in the laboratory. That way, they can obtain large supplies of rare, scarce and difficultly accessible natural products. They also can modify the structure of natural products and test the analogues for increased biological activity, decreased side effects, or metabolic stability. New chemicals can be obtained that are more powerful, more selective, or cheaper than the original natural products.



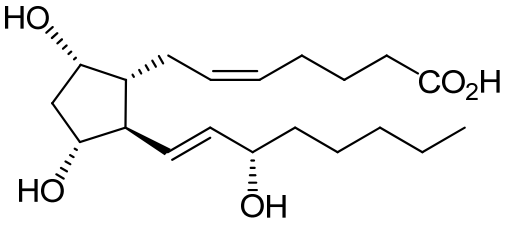
Erythronolide A
(macrolide antibiotic)
(Polyketide)



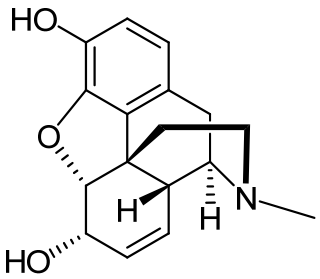
Tauro-cholic acid
(Bile acid)



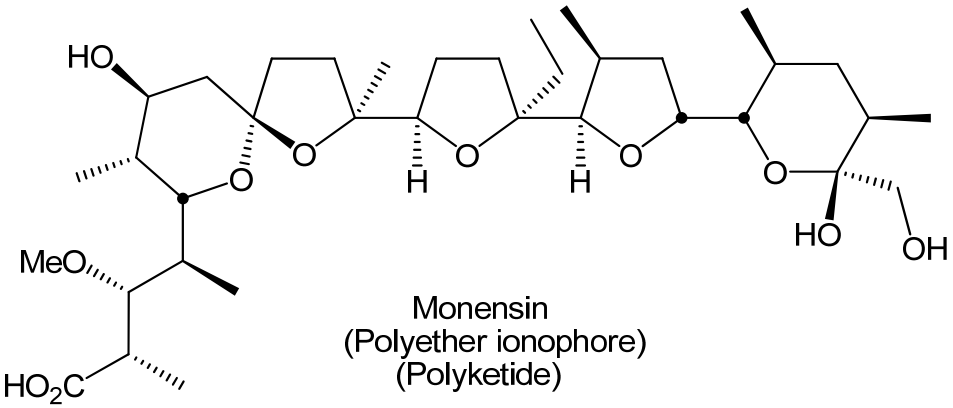
Orsellinic acid
(Polyketide)



PGF_2a
(Prostaglandin)

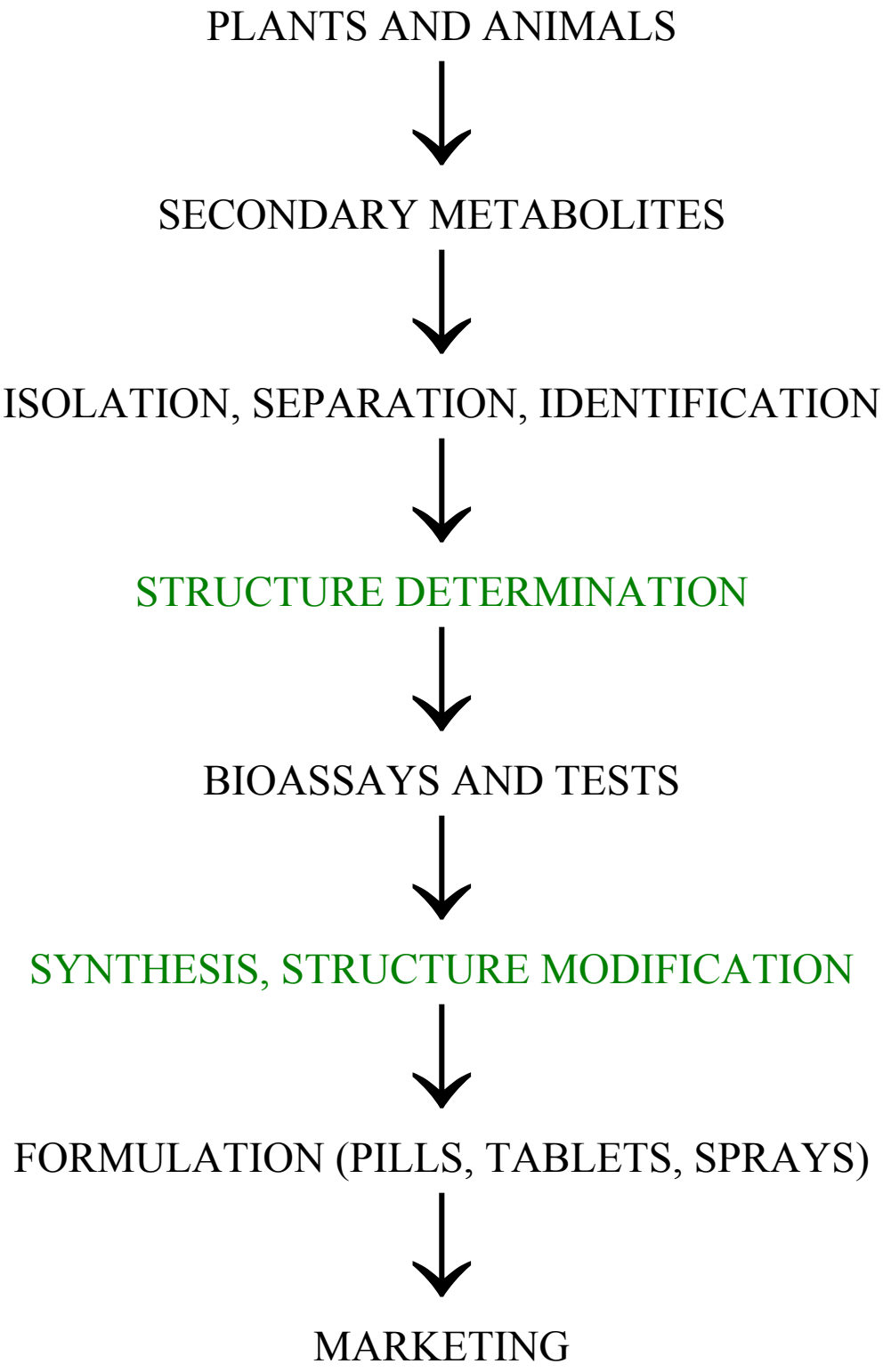


Morphine
(Alkaloid)



Monensin
(Polyether ionophore)
(Polyketide)

Figure 1.



Research is a vital part of almost each of the above areas of chemistry. Research helps develop new techniques, new products, and new ways of understanding our environment.

Applied research focuses on new and marketable products and is short to medium term oriented. Most industries do applied research to improve and accelerate production, find new products, etc. Fundamental research centers on the theory and understanding of the principles of the science in a more general and long term oriented way. While some industries also carry out fundamental research, it is for the most part concentrated in universities and government agencies.

RESEARCH
(Applied and fundamental)

Isolation and separation: find new ways to separate natural products (chromatography, gel permeation, electrophoresis...), new isolation techniques.

Structure determination: find new spectroscopic techniques, new and faster ways to elucidate structures of compounds.

Synthesis: discover new reactions, invent new methods for the synthesis of compounds, make new drugs, antibiotics, hormones, etc.

Biosynthesis: determine the mechanism by which nature makes natural products and how it regulates their formation and use.

Biological activity: find new biological activities, therapeutic properties for new or known compounds. Establish the mechanism of action of drugs and other agents. Uncover the relationship between the structure of the active compound and its biological activity.

Theoretical calculations: design computational methods to help predict molecular behavior (conformation, motion, etc.). Design software to help choose synthetic pathways or for information retrieval.

Multidiciplinary research: scientists in various research fields cooperate to rationally design more effective and less toxic drugs, to uncover the molecular basis of diseases, and to help understand in a global way how living being function.

Areas that benefited from applied and fundamental research are numerous and varied. For example, research brought pain killers that are now faster acting, more powerful and will not upset your stomach. New drugs with little addictive properties are used to treat drug addicts. Insecticides with lower mammalian toxicity are synthesized. Therapeutic drugs with higher potency and lower side effects are developed. The long term benefits of such product development are immeasurable.

I. Isolation and Structure Determination

I.1 Isolation and separation

The first steps in the study of a natural product are of course its isolation, purification and identification. Crushed samples of plants or animals are extracted with organic solvents such as methanol or chloroform for polar compounds, and ether or hydrocarbons for less polar ones. Such extracts contain mixtures of several natural products, sometimes dozens, and the targeted ones will have to be separated from the others. If the purpose of the isolation is to identify new compounds, then many different products will be separated, identified, and tested for biological activity. If only one compound was targeted, the rest will be discarded. The amount of products obtained from the extraction depends on its live source. It is not uncommon for kilos of crushed samples to yield milligrams of a particular product.

Separation techniques are numerous and varied. Perhaps the most powerful and common is liquid chromatography on a solid support. Different compounds pass through a column of solid absorbant, such as silica or alumina, and are eluted off the column with solvents, at different speeds, according to their structures. Often, several chromatographies are necessary to obtain pure fractions of compounds. Other methods include distillation, recrystallization, gel permeability, etc.

After it has been isolated, separated from other components, and purified, the natural product must be identified i.e. its physical characteristics and measurements are recorded and catalogued: melting point for solids, boiling point for liquids (if applicable), optical rotation, etc. These measurements will later be valuable in helping to identify a compound. Then a name is found for the natural product. The next step, perhaps the most difficult, is the structure elucidation of the compound (nowadays, compounds are named only after their structure has been elucidated).

I.2 Structure Elucidation

In elucidating chemical structures, many physical and spectral measurements of the unknown compound must be obtained. Known compounds can be identified from a comparison of these measurements with those reported in the chemical literature. For new compounds, the sum of all the different pieces of information provided by spectral data will enable the chemist to deduce its exact structure. This deduction is a true exercise in combining logic and intuition.

Structure determination used to be a long and arduous task. Fifty years ago, before the advent of nmr and other powerful tools, only spectral data with limited usefulness (UV, IR etc.) could be recorded and the chemist relied mostly on chemical transformations and degradations to prove the structure of a compound. Mistakes were common in determining structures or sometimes only a partial structure could be determined. Stereochemistry was difficult if not impossible to assign. The advent of single crystal X-Ray crystallography coupled with computer analysis improved greatly the situation since complete structures and relative stereochemistry could be determined. However, only a few structures could be deciphered because of the requirement to grow single crystals of very high purity and quality. Since most compounds are liquids, X-Ray crystallography had limited applications. To give you an idea of the difficulties involved in determining structures of compounds, the isolation and identification of cholesterol was done in 1812. Its structure was determined 120 years later in 1932. The stereochemical assignment had to wait until the late 1940's.

Today's modern spectroscopy has made available to the chemist a wealth of spectral information on a routine basis. The main physical tools utilized to determine a structure are:

LRMS - Low Resolution Mass spectrometry (molecular mass of molecules)
HRMS - High Resolution Mass spectrometry (exact mass)
EA - Elemental analysis (percent of atoms in molecule)
ir - Infrared spectroscopy (functional groups identification)
UV - U.V. spectroscopy (chromophores identification - limited use)
nmr - Nuclear Magnetic Resonance (extremely powerful - lots of information)
X-Ray - Single Crystal X-Ray Crystallography (extremely powerful - limited use)

Nearly all of the above methods give data that can now be Fourier-Transformed with computers and microprocessors which greatly diminish the time involved in getting the information. Of these method, none are more powerful than **nmr**. We will therefore spend more time exploring the different modern nmr experiments that can be performed to extract useful information on the structure of a compound.

When confronted with the problem of determining the structure of a chemical, the order in which one links the information together is important. One must proceed step by step and logically. For that reason, we will deal with structure elucidation in a **problem set** / **tutorial mode**. This exercise involves a great deal of memory and experience. Therefore, the student is

encouraged to try many problem sets and get familiar with the most important physical data in the various spectral fields.

Structure elucidation of **unknown 1**:

You have completed the isolation and purification of **unknown 1**. You want to determine its structure. You do not know if it is a known or a new compound. You decide to try and elucidate its structure.

First you will have to prepare many samples and submit them to have different spectrum recorded. You will need approximately 1 or 2 mg for mass spectrometry (both low and high resolution), ~ 5 mg for the infrared spectrum (which you can recover), ~ 30 to 50 mg for all of the nmr experiments (which you can also recover). You also keep a few hundred mg for possible derivatizations or chemical transformations.

This is what you get back from the spectral labs:

LRMS: M^+ : 198 m/z + many other fragments of lower mass.
HRMS 198.0892
EA H = 7.11 %, C = 60.57%, O = 32.32% Total = 100%
ir $\sqrt{}_{max}$: 3500, 1750, and 1710 cm^{-1}
^1H NMR (δ) 4.93 (1H, dt, *J*= 11.2, 6.1, 6.1 Hz) 3.79 (1H, s) 2.94 (1H, dd, *J*= 15.2, 8.0 Hz) 2.77 (1H, m) 2.64 (1H, dd, *J*= 15.2, 2.3) 2.61 (1H, dd, *J*=14.7, 6.1 Hz) 2.38 (1H, dq, *J*= 12.1, 7.0 Hz) 1.89 (1H, dd, *J* = 14.7, 11.2 Hz) 1.47 (3H, s) 1.26 (3H, d, *J*= 7.0 Hz)

^{13}C NMR (δ) 210.6 (s) 170.0 (s) 73.7 (d) 73.6 (s) 44.2 (d) 42.4 (t) 37.9 (d) 35.6 (t) 25.0 (q) 13.3 (q).

Figure I.2.1. and I.2.2. show computer drawings of the ^1H nmr and ^{13}C nmr of **unknown 1**.
One can break up the structure elucidation in three main categories:

- Determination of the Formula (number and identity of atoms)
- Elucidation of the general structure (connectivity of atoms)
- Assignment of Stereochemistry(exact 3D picture of the structure)

I.2.1 Molecular Formula

There are a few things you can do that are simple and that give very useful information that you will need in order to make certain deductions from other data sets. The first is to determine

the number and identity of the atoms present. EA shows that only C, H, and O are present since their relative percentages add up to 100%. Next, from the HRMS you can find the exact formula. The Silverstein book on "spectrometric identification of organic compounds" contains mass spectrometry tables that give the exact molecular masses for a series of molecular formulae (two different formulae cannot have the same exact molecular mass). From this table, we find that 198.0892 gives $C_{10}H_{14}O_4$ as our formula. One can then verify that $C_{10}H_{14}O_4$ gives indeed the right EA of H = 7.12 %, C = 60.59%, O = 32.29%.

{ (10 x 12.01115 = 120.1115) + (14 x 1.00797 = 14.11158) + (4 x 15.9994 = 63.9976) = **198.22**}

$$\left\{\frac{120.1115}{198.22} \times 100 = 60.59\% \quad \frac{14.11158}{198.22} \times 100 = 7.12\% \quad \frac{63.9976}{198.22} \times 100 = 32.29\%\right\}$$

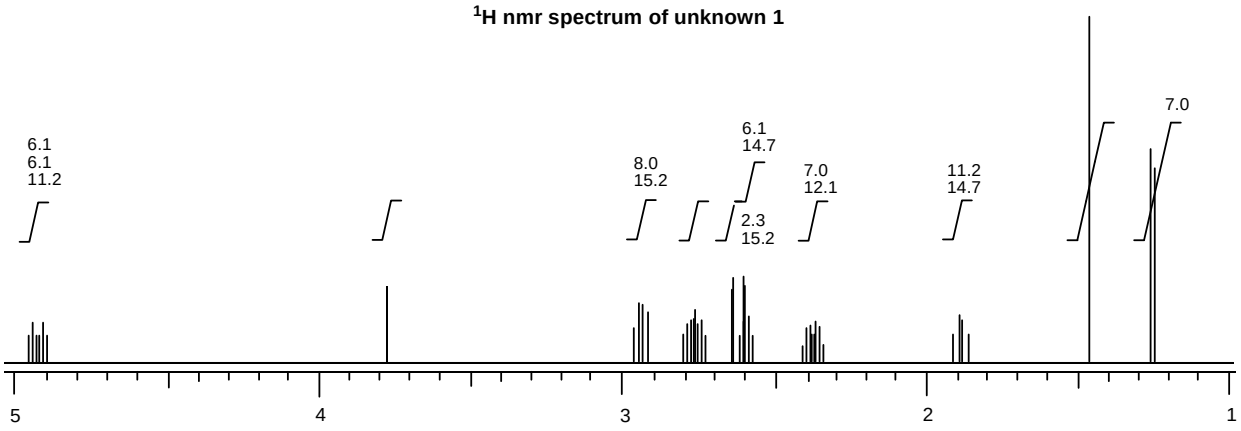


Figure I.2.1

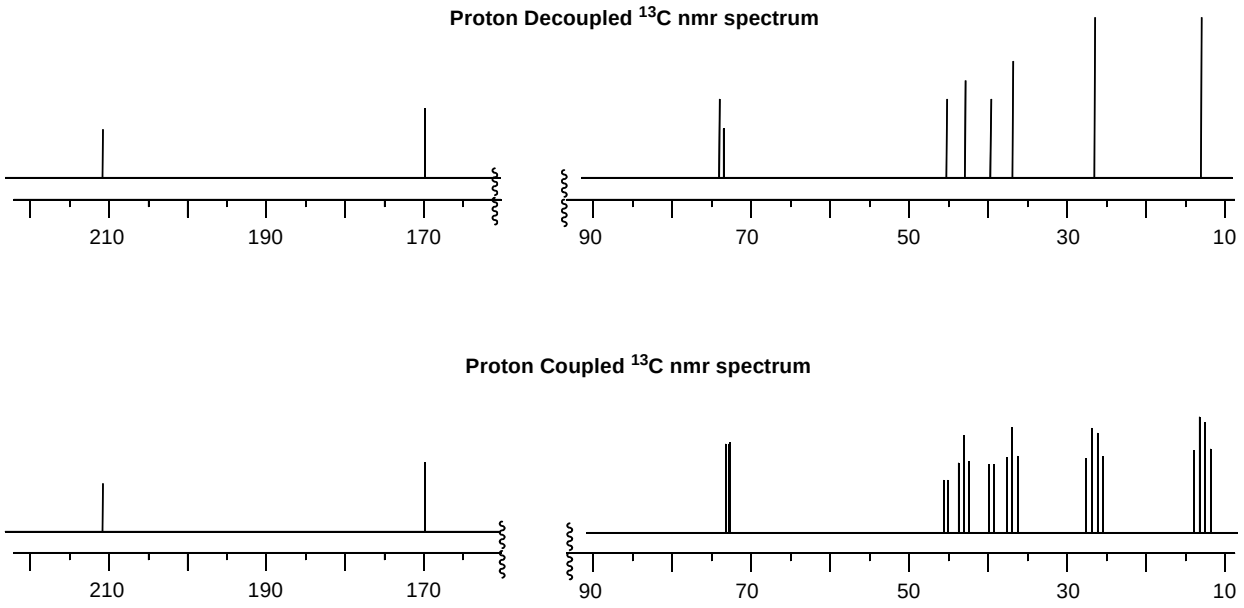


Figure I.2.2

These calculations are in accordance with the experimental EA. Now we are certain that **unknown 1** has the formula $C_{10}H_{14}O_4$. The second easy and very useful thing to do is to determine the degree of unsaturation in the molecule. Here is how:

$$\#C - \frac{\#H}{2} + \frac{\#N}{2} + 1 = \textbf{Degree of unsaturations}$$

#C : Carbon, silicon, tin, or any **tetravalent** atoms.

#H : Hydrogen, halogen, or any **monovalent** atoms

#N : nitrogen, phosphorus, or any **trivalent** atoms

Unsaturations are :

double bonds (including carbonyls, etc) = 1 unsaturation

rings = 1 unsaturation

triple bonds = 2 unsaturations.

In our case:

10 - 7 + 0 + 1 = 4 unsaturation sites

We must therefore keep in mind that **unknown 1** contains either 4 double bonds or carbonyls or rings or 2 triple bonds **or any combinations thereof**. Now let's look at the spectral information. It is a good practice to start with the spectrum that gives the least amount of information. That way, we can have that information already in hand when we analyze the more

complex spectra. Infrared is a very simple spectrum to analyze that gives information about functional groups in the molecule. Functional groups are reactive sites like carbonyls, double bonds, alcohols, amines, etc.

I.2.2 General Structure Elucidation

The ir indicates the presence of 2 or more carbonyls (1750, and 1710 cm^{-1}) and one or more alcohol (3500 cm^{-1}). Two carbonyls account for 2 of the 4 unsaturations. One carbonyl falls in the region of esters or lactones (1750 cm^{-1}) and the other in the region of ketones (1710 cm^{-1}). So we can draw the first portions of our molecule: -C—OH -O—C=O and $(C)_2$—C=O. If none of these atoms are the same we already have assigned 5 carbons and 4 oxygens to our unknown structure!

We are now prepared to attack the more difficult but more interesting part of the game: nmr analysis. Carbon 13 nmr indicates the presence of 10 different carbon atoms (the formula $C_{10}H_{14}O_4$ is accounted for in # of C's). Only two of the signals are in the carbonyl region. In fact, one is in the ketone region (210 δ) and the other is an ester or lactone carbonyl signal (170 δ). Most likely we have only two carbonyls in the molecule. Therefore the other two unsaturations could be alkenes, rings or one triple bond. IR and ^{13}C nmr indicate no signals corresponding to alkenes or alkynes (this will be confirmed by the proton nmr). Therefore the other unsaturation **must be two rings**. How these rings are connected or what size they are remains unknown for the moment.

Before we start analyzing the ^{1}H nmr, let's extract a few more clues from the ^{13}C nmr. The proton-coupled spectrum of **unknown 1** indicate that 3 signals are singlets, 3 are doublets, 2 are triplets, and 2 are quartets. Though it is not the case here, very often proton-coupled ^{13}C nmr spectra are blurred by extensive overlapping of signals since each signal can give rise to several lines. Therefore, it is now custom to use the DEPT experiment (from Distortionless Enhancement by Polarization Transfer) to elucidate the signals multiplicity (Figure I.2.3). In this experiment, a pulse of magnetization is send to the sample and the accumulation of data starts after a well defined time lap. During this time lap, each excited carbon will try to come back to its original energy state (relaxation). However, quaternary carbons take a different time lap in order to relax than do methine (CH), methylene (CH_2), or methyl protons. The end result is that in a DEPT spectrum the signals due to quaternary carbons have vanished, those due to methylene show as negative peaks, while methine and methyl signals are unaffected. By inspection of the DEPT spectrum of unknown 1, we can expect 3 carbons with no hydrogen attached to them (carbonyl, quaternary carbons...), 2-CH_2's and 5 -CH's and/or -CH_3's (Figure I.2.3). Since methine protons rarely resonate below 30 ppm, we can assume that the two signals at 25 and 13

ppm are methyl groups. This result is in agreement with the proton-coupled spectrum analysis. The signals at 73.7 and 73.6 δ indicate that an oxygen is single bonded to those carbons.

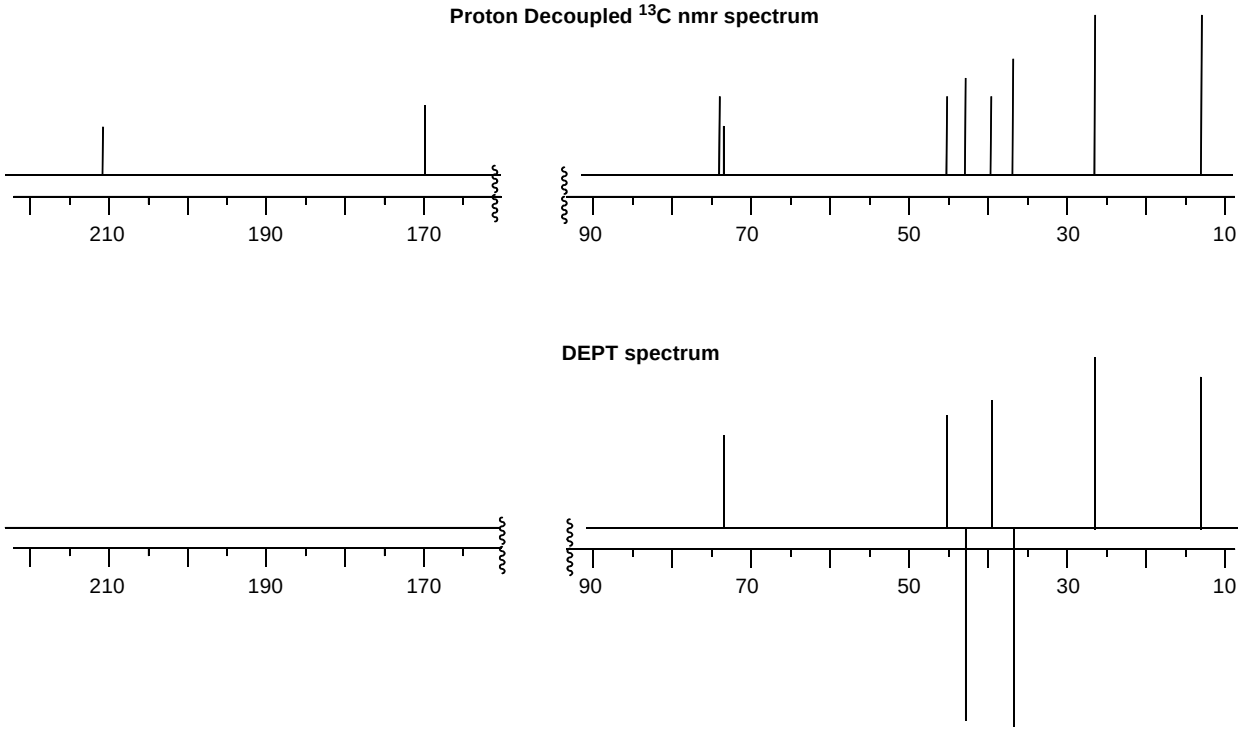


Figure I.2.3

Remember that we know the formula is $C_{10}H_{14}O_4$ and there are no double bonds present. 73.7 and 73.6 δ must therefore be due to C's attached to an oxygen. From the preliminary examination of data so far, we can deduce:

-**C**—OH	H**C**—O—**C**=O	$(C)_2$—**C**(O)=O	$(C)_2$—**C**=O
73.6	73.7	170	210

In addition, we have: 3 -CH's, 2-CH_2's, 2 CH_3's, 2 rings. We cannot start putting these together yet. We need more information on the chemical environment of each of these groups.

Now, let's look at the ^{1}H nmr. There are 10 distinct signals in the nmr that must account for 14 protons. Considering that we know of 2-CH_3's from the ^{13}C nmr, and that 2 of the ^{1}H nmr signals integrate for 3 protons, all of our 14 hydrogens are there. Remember that the ^{13}C nmr indicated the presence of 2-CH_2's. None of the signals in the ^{1}H nmr integrate for 2 protons. We must therefore conclude that **chiral centers** are present resulting in the **chemical shift**

inequivalence and therefore in the **magnetic inequivalence** of the methylene protons. That is each proton of the -CH_2 will give a distinct signal. In addition, they will couple to each other. A typical value for such a geminal coupling constant is 15 to 17 Hz. In cases where two methylene protons are chemical shift and magnetically equivalent, they will both give one signal that shows no geminal coupling (Figure I.2.4 left side). When they are not chemical shift equivalent (different environment) and thus necessarily not magnetically equivalent, then the two protons will give rise to two different ^{1}H nmr signals (Figure I.2.4 right side). Each of these signals will have the same J Ha-Hb coupling constant but the J Ha-H' will be different from the J Hb-H' coupling constant. Notice that in all cases we have a signal with four lines. However, one is a quartet (left side) and all the lines are even spaced. The other two signals are called doublets of doublets (dd) and the first two lines give the smallest J constant whereas the first and third lines give the largest J constant.

Now let's take each signal and try to deduce as much information as possible out of them. Let's start with the signal due to alcohol proton. To find its position is trivial. One has only to add a drop of D_2O to the nmr sample and rerun the spectrum. Because alcohols have the capacity to exchange their hydrogen with other alcohols and water, one deuterium of D_2O will replace the ROH, and thus the signal due to ROH will disappear completely. A new signal due to HOD should appear around 4 ppm. If we do this experiment, we find that the signal at 3.79 δ disappears and is therefore assigned to -O-**H**.

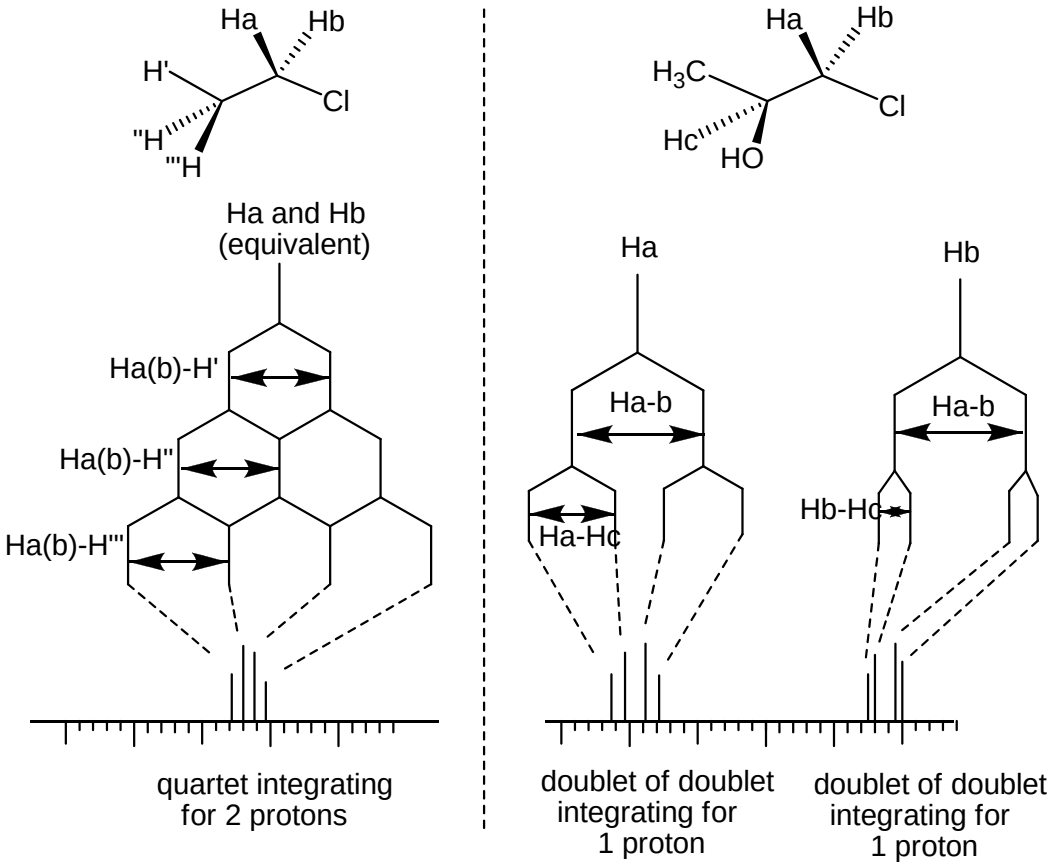


Figure I.2.4

We can see that one signal is more deshielded than all the others. The signal at 4.93 δ is likely due to a proton attached to a carbinol of an ester (**H**-C-O-C=O). Why? Because the chemical shift is in the right region, the ir indicates the presence of an ester, no other heteroatom is present in the molecule, there are no double bonds, and **H**-C-OH would give a signal around 3.5 ppm. Nothing else can account for such a low field signal. The proton at 4.93 δ is a ddd, with J = 11, 6, 6. Therefore the structure **S1** depicted in Figure I.2.5 is likely to be a partial structure of **unknown 1**.

Let's take the right hand side of the spectrum. Two signals are integrating for 3 H's; one doublet at 1.26 δ and one singlet at 1.47 δ. It is highly probable that both are the two methyl groups we have seen in the ^{13}C nmr. In addition the singlet at 1.47 δ is deshielded with respect to a methyl group attached only to carbons. It is thus likely that an electronegative atom be placed near this methyl group. Thus we have the situation depicted in **S2** and **S3** of Figure I.2.5. The doublet has a coupling constant of 7.0 Hz. Let's see what other signal in the ^{1}H nmr spectrum has also a J = 7.0 Hz. The only other signal with J = 7.0 Hz is the doublet of quartet at 2.38 δ. We know that this signal has to be at least a quartet (due to the coupling with the methyl group).

Now we realize that this proton is coupled to one more proton only and we conclude that the partial structure **S3** in Figure I.2.5 must be part of **unknown 1**. That doublet of quartet has J = 7.0 and 12.1 Hz. We cannot find another signal with a J = 12.1 Hz. However, the signal at 2.77 δ which is a "m" (multiplet) could possess this J = 12.1 Hz (multiplets are complex signals for which values for the coupling constants are difficult to assign).

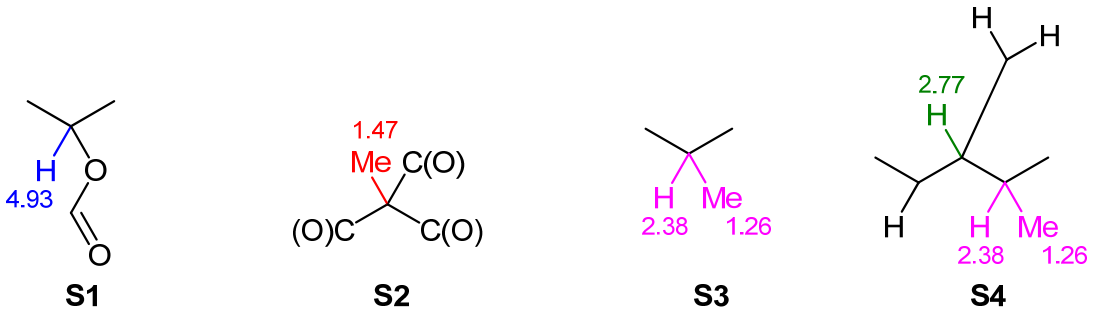


Figure I.2.5

We realize that it can become quite difficult and arduous to try and determine how each hydrogen is coupled to each other by looking only at the coupling constant. How can we obtain more reliable information on the connectivity of each proton with respect to each other? For relatively simple spectrum, this is done with **spin-spin decoupling** experiments. In this experiment, a specific proton signal is irradiated with a second rf while acquiring the spectrum. This causes the disappearance of that signal due to "saturation" i.e. the equalization of the populations of its higher and lower energy levels. But more importantly, any proton which normally couples with this irradiated hydrogen is now "decoupled" because it can no longer "sense" the energy levels of the irradiated proton. In fact, it "senses" the average of these energy levels which is 0 (equal populations) and therefore no longer couples with that hydrogen. The new spectrum is plotted, along with any other "decoupled spectrum", usually one over another for easy comparison. By looking at the altered (decoupled) signals in each decoupled spectrum, one finds the hydrogens which couple to the irradiated signal in the normal, non-decoupled, spectrum. Figures I.2.6, I.2.7, and I.2.10 show decoupled spectra of **unknown 1**.

The first irradiation at 2.38 δ (Figure I.2.6) shows two changes; the methyl at 1.26 δ is now a singlet confirming our hypothesis that these two were coupled as shown in **S3** of Figure I.2.5; the signal at 2.77 δ has also changed and now appears as a ddd with J = 2.3, 6.1 and 8.0 Hz) which indicates it is normally coupled to four protons. We now have structure **S4** in Figure I.2.5.

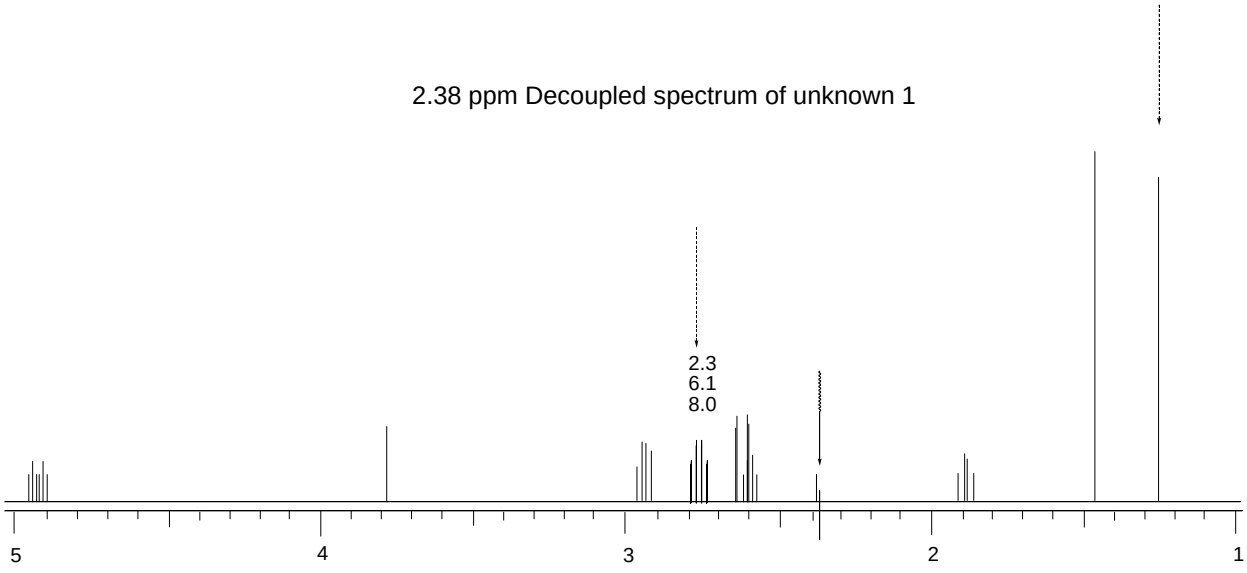


Figure I.2.6

Irradiation at 4.93 δ (Figure I.2.7) shows simplification of three signals at 2.61, 1.89, and 2.77 δ which is in accordance with our first hypothesis **S1** of Figure I.2.5. Those simplified signals are in turn easier to analyze. The two signals at 2.61 and 1.89 δ are now both doublets with J = 14.7 Hz i.e. they both couple to only one more proton besides 4.93 δ. We can conclude that they are on the same carbon and couple to each other since J = 14.7 Hz cannot be found anywhere else. In addition, they must be next to a carbon with no hydrogens. The ^{13}C nmr shows only three such carbons and they are carbonyls or have an oxygen attached to them. Therefore we must have the situation depicted in **S5** or **S6** in Figure I.2.8.

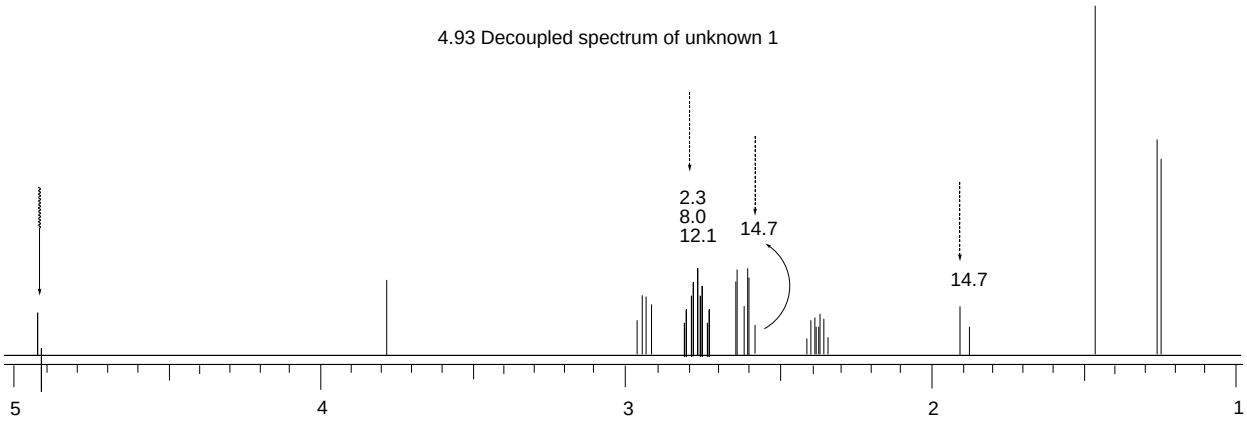


Figure I.2.7

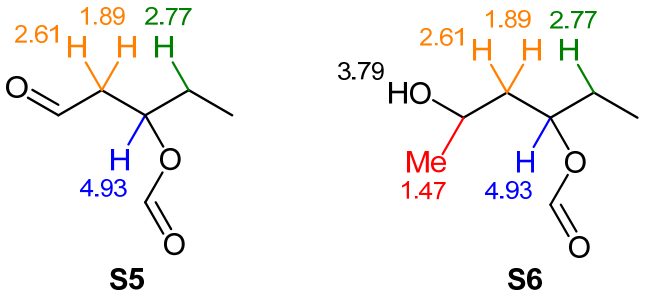


Figure I.2.8

The peak at 2.77 δ, which was quite complex, now displays a ddd signal with J = 2.3, 8.0 and 12.1 Hz. This hydrogen at 2.77 δ is therefore coupled with 3 hydrogens besides 4.93 δ. We know it is coupled to 2.38 δ from the first decoupling experiment. In addition, we know that 2.38 δ is coupled to no other proton but the ones on the methyl group and thus must be next to a carbon without hydrogens i.e. a ketone or a quaternary carbon. This is the third such quaternary carbon we have assigned. Since the ^{13}C nmr indicates only two carbonyls, we must assume that the third quaternary carbon is made from the remaining methyl group and alcohol. Therefore we obtain the partial structure **S7** as shown in Figure I.2.9. However, the methylene bearing the two protons that display signals at 2.94 and 2.64 is maybe near a carbonyl since the other methylene's signals are shielded in comparison. However, this is a little tenuous and we will not make that an affirmation yet. We do not know yet if the methylene at 2.64 and 2.94 is near the ester or the ketone or perhaps (but unlikely) the carbon bearing the methyl and alcohol as well. Therefore, that leaves us with structure **S7** (Figure I.2.9).

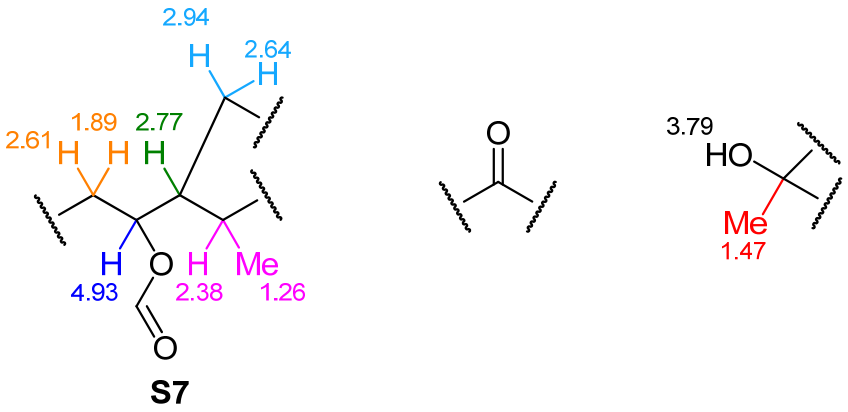


Figure I.2.9

From the last irradiation at 2.77 δ (Figure I.2.10), we of course obtain simplification at 4.93 δ which becomes a dd with *J* = 11.2 and 6.1 but we also obtain simplification of the signal at

2.38 δ which becomes a perfect quartet with J = 7.0 Hz. Therefore that methyne must be attached to one of the three chain ends shown in **S7** or **S8**. The other two signals that are coupled to 2.77 δ are thus 2.64 and 2.94 δ. These signals become doublets with the same J = 15.2 Hz and are therefore coupled to each other. In addition they must be next to a carbon with no hydrogen but we knew that already. All ten carbons and fourteen hydrogens are now accounted for.

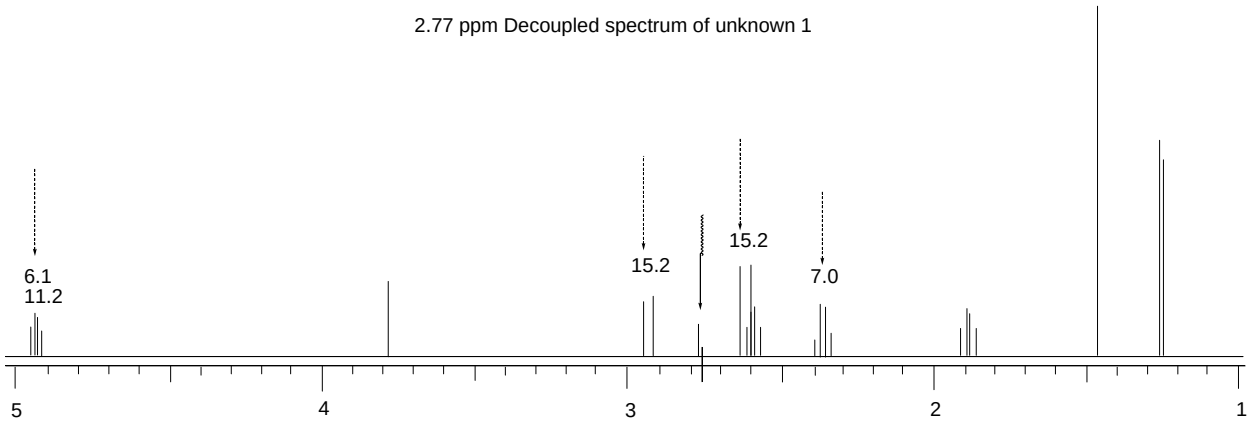


Figure I.2.10

It is worth mentioning that decoupling experiments, although very useful, are time consuming since they require acquisition of a spectrum for every irradiated signal. There exists a modern 2 dimensional nmr experiment which will give the entire connectivity pattern for each and every hydrogen in a single spectrum. This experiment is called the "COSY" experiment (from COrrelated SpectroscopY). The interested reader can find a detailed description of this experiment in A.E. Derome's "Modern NMR Techniques for Chemistry Research", Pergamon Press, 1987. The acquisition of the spectrum is complex and will not be detailed here. The interpretation of the spectrum is straightforward. A normal proton nmr is plotted on each axis of the 2D COSY spectrum (Figure I.2.11). Each and every **coupled signals** will show a cross peak which can be found by drawing a vertical line from either of the coupled signal to the diagonal of the spectrum and then drawing a horizontal line to the identical signal on the other axis. Any cross peak found on this horizontal line belongs to a hydrogen(s) coupled to the signal in question. To find which hydrogen the cross peak belongs to, just trace a vertical line from the cross peak to the signal on the horizontal spectrum. All of this is pictured in Figure I.2.11. Note how the connectivity that we found using the decoupling experiments correlates very well with the one found using the COSY spectrum. One disadvantage of the normal COSY experiment is that it is not possible to get the coupling constants of the coupled hydrogens.

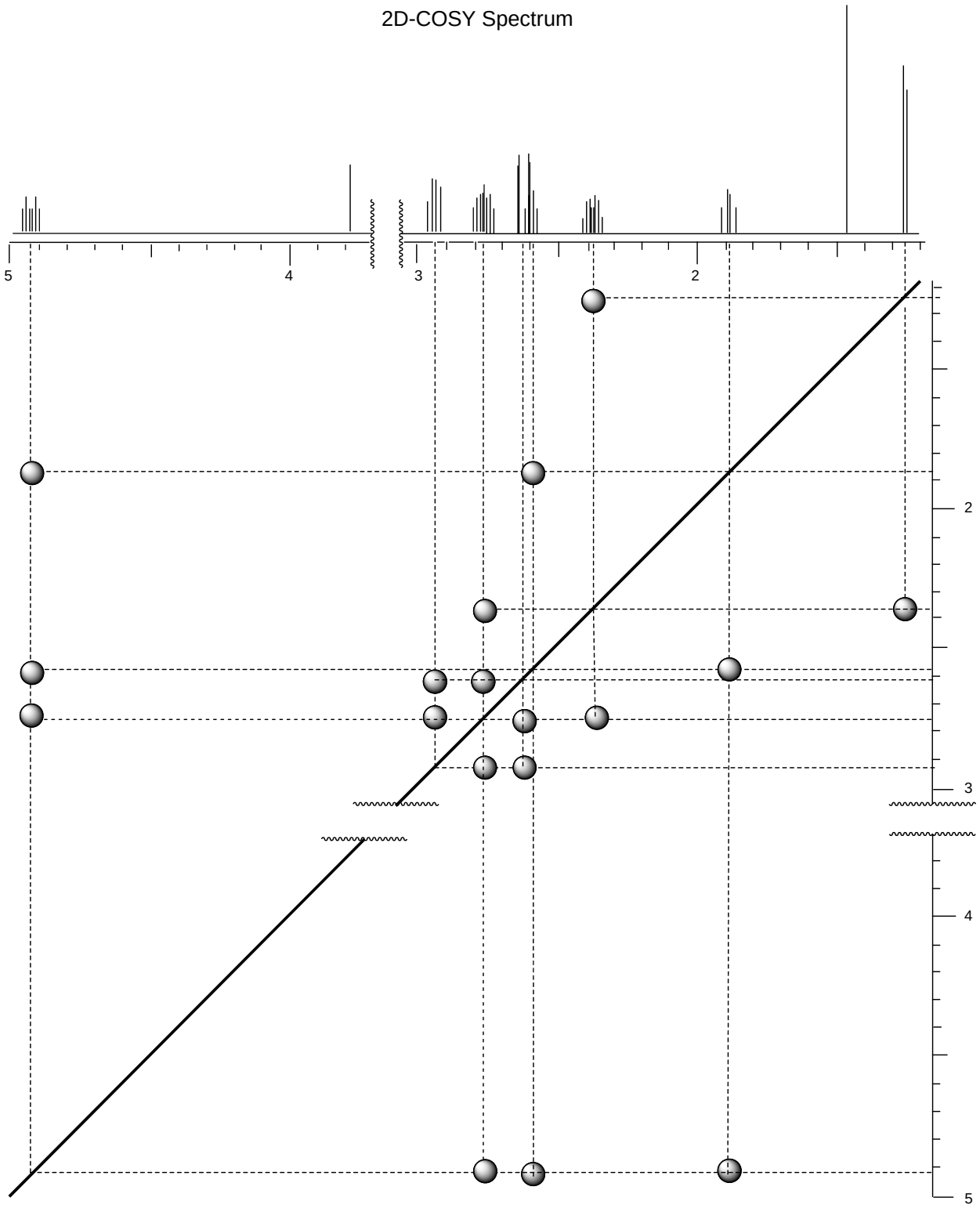
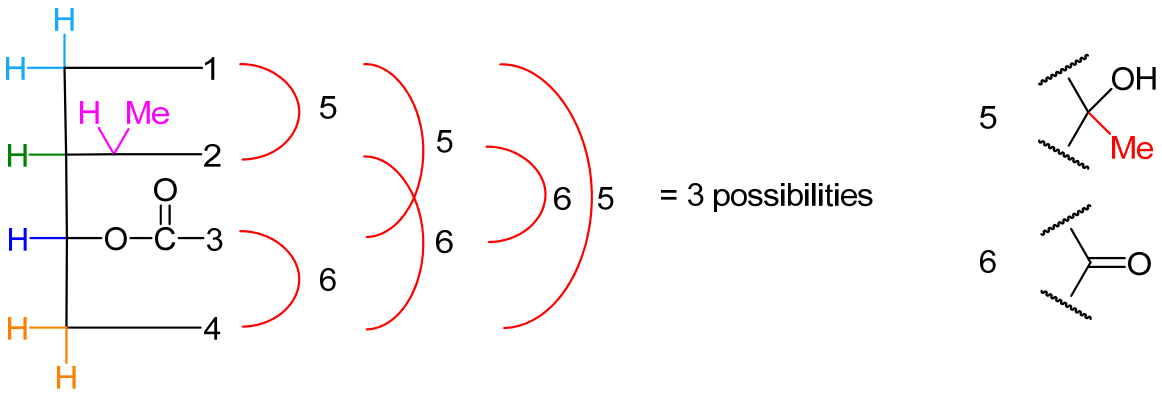
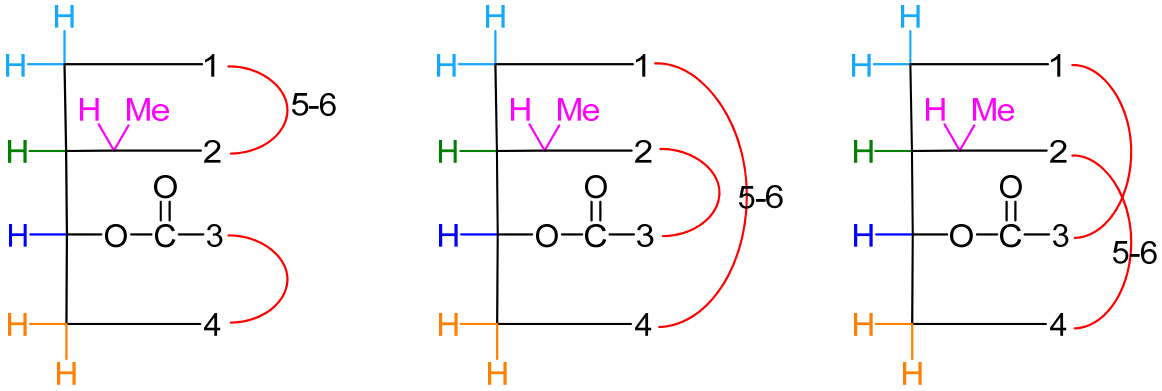


Figure I.2.11

We already have a pretty good picture of the molecule. Two rings remain to be constructed. Their size and connectivity must now be assessed. We will have to decide which chain ends are connected together. There are four chain ends, each of which can accept only one bond, for a total of nine possibilities (Figure I.2.12). The resulting structures are pictured in Figure I.2.13.

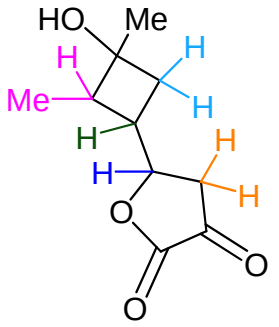


5 and 6 can be inverted. So, x 2 = 6 possibilities

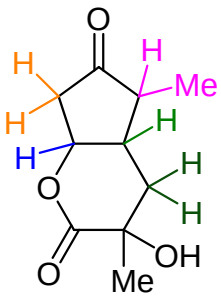


position 3 and any of position 1, 2 or 4 can be directly bonded
so 6 + 3 = 9 possibilities.
The last 3 are doubled because 5-6 or 6-5 can be binding, for a total of 12 possibilities.
There are none other.

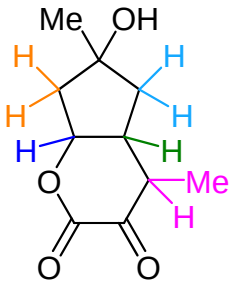
Figure I.2.12



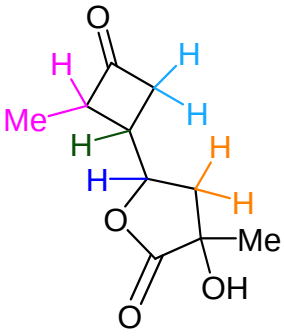
S8



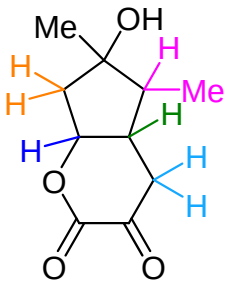
S9



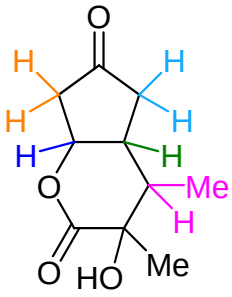
S10



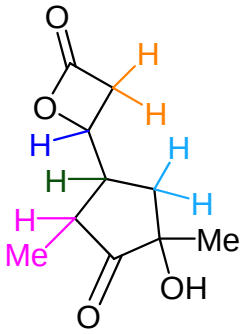
S11



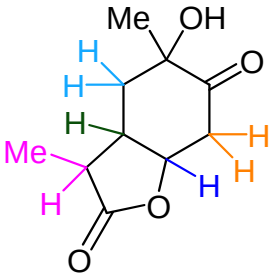
S12



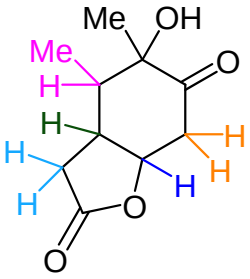
S13



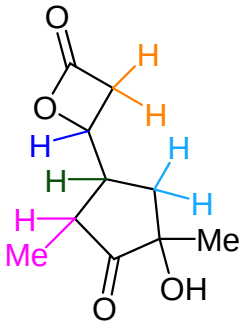
S14



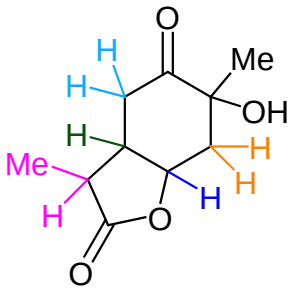
S15



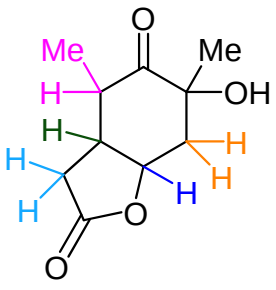
S16



S17



S18

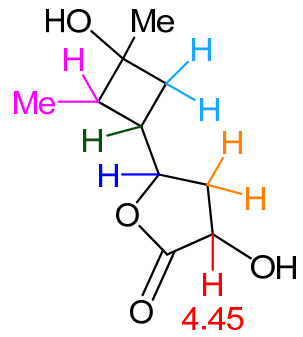


S19

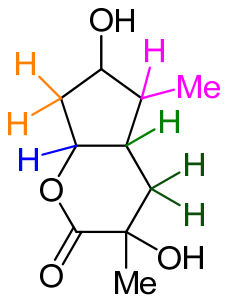
Figure I.2.13

Structure **S11** can be eliminated on the basis of the IR spectrum (a ketone in a four membered ring should absorb at 1775 cm^{-1}). It is for the moment very difficult to decide between the remaining eleven structures. We must rely on more powerful spectroscopic techniques or on chemical transformations and degradations to get more information about the structure.

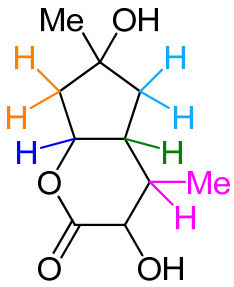
Deciding what chemical transformation to do requires advanced knowledge of available chemical reactions. For the purpose of this course, such chemical transformation will always be given to you. In the case of **unknown 1**, the ketone was selectively reduced with sodium borohydride ($NaBH_4$) without concomitant reduction of the ester. We then examined the ^1H nmr of the reduced product, and deduced the position of that ketone. Two isomeric alcohols were obtained from that reduction (establishing that the two products are isomeric is easy: mass spec gives the same MW and Formula). The newly introduced proton resonates at 4.45 δ and a decoupling experiment shows that it is coupled to the two signals that were previously assigned at 2.94 and 2.64 δ but not to the signals at 1.89 and 2.61 δ (Figure I.2.14). That implies that the two hydrogens at 2.94 and 2.64 δ were next to the ketone such that we can now eliminate structures **S8**, **S9**, **S10**, **S-14-18**, and **S-19** because the ketone in those structures is not next to the protons at 2.94 and 2.64 δ. We have now only four possibilities left: **S11-13** or **S18**. How can we differentiate between those compounds? Also, what is the relative stereochemistry of the molecule? Further chemical transformation of the unknown could help differentiate between compounds, however, we will see that there is a faster and more efficient way to determine the exact structure of **unknown 1** while assigning the stereochemistry of the product at the same time.



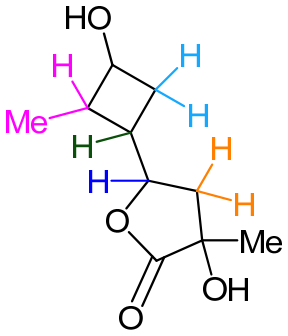
S8'



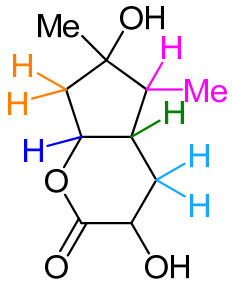
S9'



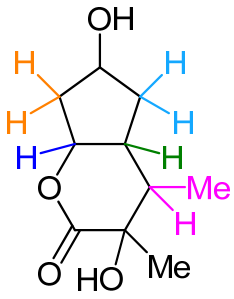
S10'



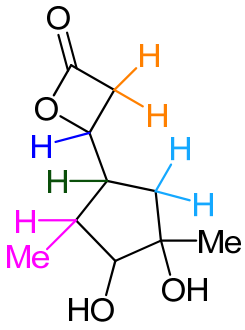
S11'



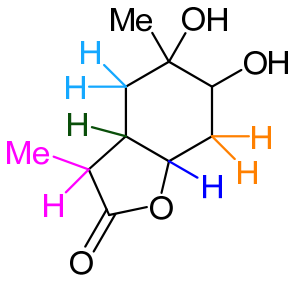
S12'



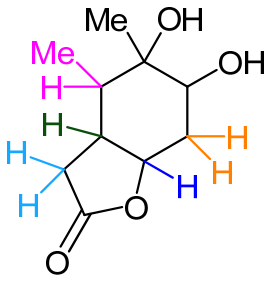
S13'



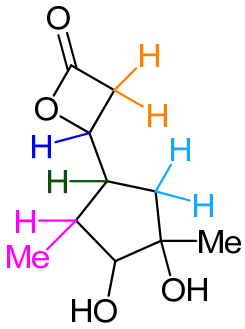
S14'



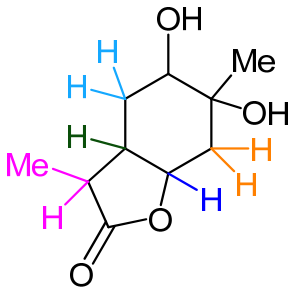
S15'



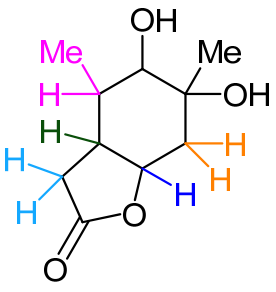
S16'



S17'



S18'



S19'

Figure I.2.14

I.2.3. Assignment of Stereochemistry

The question of stereochemistry is a very challenging one. In fact, this is where most of the mistakes are made in structure elucidation. But we will see that careful nmr experimentation can lead to extremely useful information about the stereochemistry of compounds. The nmr experiment we are referring to is the **nuclear Overhauser effect** or nOe. We will describe this experiment in detail.

First let us define the actual effect; the nOe is defined as the change in **intensity** of an nmr resonance when the transition of another resonance is perturbed. In practice, this perturbation is an irradiation with a weak radio frequency (rf) field of a specific nucleus while observing the resonances of the others. The very simple mathematical expression for the nOe of a nucleus is:

$$\eta\,(s) = \frac{(I-I_0)}{I_0} \times 100$$

If "I_0" is the **intensity** of the signal in the normal unperturbed nmr spectrum of a particular nucleus and " I " is the **intensity** of that same signal after irradiating another nucleus in an nOe experiment then "η" is the nOe effect expressed in percentage of increase in intensity. In most cases however we need not quantify the nOe and only a **qualitative** analysis of the effects is satisfactory. This simplifies greatly the interpretation of the spectrum.

The reason why the nOe phenomenon occurs is a complex one. Let us not attempt to explain all the complex mechanisms by which the phenomenon manifests itself but rather let us try to understand the basic principles behind the phenomenon. We are all familiar with representing the nucleus as a small magnet possessing a magnetic field of its own. When a molecule contains several nuclei, their respective magnetic fields will interact with each other through a plethora of different mechanisms. The familiar scalar or "*J* "coupling is caused by the interaction of an electron with a nucleus and is transmitted from electron to electron up to another nucleus. This is a through-bond interaction. In general the interaction decreases with the number of bonds and is normally lost after three bonds (the two hydrogens couple in **H**—C—C—**H** but not in **H**—C—C—C—**H**). Another, more complex, interaction is the *dipolar coupling* between nuclei. We said earlier that each nucleus has its own magnetic field much like a magnet has. In a magnet, the north and the south poles represent this *dipole*. If you bring two magnets to proximity, their magnetic fields will influence each other. This is also true of the magnetic fields of nuclei. We call this interaction a *dipole-dipole* or *dipolar* interaction. We can easily understand that this interaction will be a function of the distance between the two magnets or nuclei. The nOe is a direct consequence of this *dipolar* interaction and is therefore influenced by the distance between

the nuclei in question. Generally it is true that the nOe decreases with the 6^{th} power of the distance.

To describe briefly this interaction, we must look at how the spin system (nuclei) tries to go back to their original energy state (relax) when we perturb the system (irradiate). In a normal spectrum, irradiation causes the nuclei to "flip" (resonate) between the lower and the higher energy levels (if nuclear spin is $^1/_2$ = two energy levels). After the irradiation, the spin system (nucleus) tries to get back to its original energy state (relaxation). It can do this in many different ways, but we will describe only the mechanism of interest here. The *longitudinal* or *spin lattice* relaxation is a return to original energy state (a return to equilibrium conditions) which occurs when the perturbed (irradiated) nucleus is giving off its excess energy (the energy it absorbed when irradiated) to the other nuclei surrounding it, i.e the molecular lattice. A perturbed nucleus will give off its resonance energy (relax), through a magnetic field oscillating at the same frequency as its resonance (such an oscillating magnetic field can be induced by the surrounding nuclei). This is called *stimulated emission*. Because such oscillating magnetic fields are not abundant in the environment of the molecule, the relaxation time T_1 of a typical nuclei is long (millisecond to minutes). Because T_1 is long, we can measure and use it. Let us now look at what effect this relaxation has on the surrounding nuclei.

A nucleus relaxing through the molecular lattice will increase the higher energy level population of the surrounding nuclei which will absorb this energy (remember that energy absorption will decrease with the distance to the 6^{th} power (r^{-6}) . If this is true, then the nmr signal that we detect for those nuclei will **increase in intensity**. In the nOe experiment, we record an nmr spectrum while irradiating a specific nucleus with a second rf. The nucleus relaxes and increases the intensity of the nmr signals of the nuclei in proximity. It can be, however, difficult to detect these increases in intensity (which varies between 5 and 50% maximum) in the nmr spectrum. For that reason, we also record a normal spectrum (with no second irradiation) which we later **subtract** from the previous nOe spectrum. The result is a plot of only those signals which display an nOe since all the other signals are cancelled. This technique is called the **nOe difference experiment** or **nOeD**.

In rigid molecules (small ring systems), the distances between nuclei vary only slightly and the nOe experiment becomes meaningful and useful for calculating the internuclear distances. In more flexible systems, the nOe reflects an average distance which could be difficult to assess. Nevertheless this is possible. It is possible to calculate the nOe's and translate this into real interatomic distances. However, this is often time consuming and not always necessary. In most instances, a simple qualitative analysis is sufficient. To have a better understanding of the nOe experiment, let's look at a real example.

In determining the structure and stereochemistry of **unknown 1**, the nOe experiment can be extremely useful. Figure I.2.16 shows a partial NOESY spectrum of **unknown 1** and Figure I.2.15 shows some of the observed nOe enhancement deduced from the NOESY spectrum. This spectrum is interpreted exactly in the same way as the COSY spectrum. However, each cross peak signifies a nOe (through space) relationship and not a coupling relationship. This is important as nOe is not a through bon process. As long as the physical distance between two groups is within range (2-5 Å) they can give rise to a nOe.

From the NOESY spectrum, we detect a clear nOe between the signals from the peak at 4.93 δ and that of the protons at 2.38, 2.77 (green dots), and the methyl doublet at 1.26 δ (small red dot). From this we can eliminate structure **S11** because free rotation around the bond linking the two rings should not favor a nOe between 4.93 and 2.38 δ over a nOe between 4.93 and any of the signals 1.64 or 2.94. This correlation also indicates that the methyl at 1.26 is likely on the same side of the ring as the methane at 4.93. Moreover, the ring junction in any of the remaining structures must be *syn* because of the nOe between 4.93 and 2.77.

The methyl singlet resonance at 1.47 δ correlates with the signals at 1.89 and 2.61 δ (blue dots). This indicates that those protons are closer in space to the methyl group than any other protons. From these nOes we first deduce that structure **S13** does not correspond to the spectroscopic evidence because they are too far in distance to give rise to a cross peak. Structure **S12** can also be dismissed on the basis of the lack of cross peaks between the methyl signal at 1.47 and either of the signals at 1.26 or 2.38.

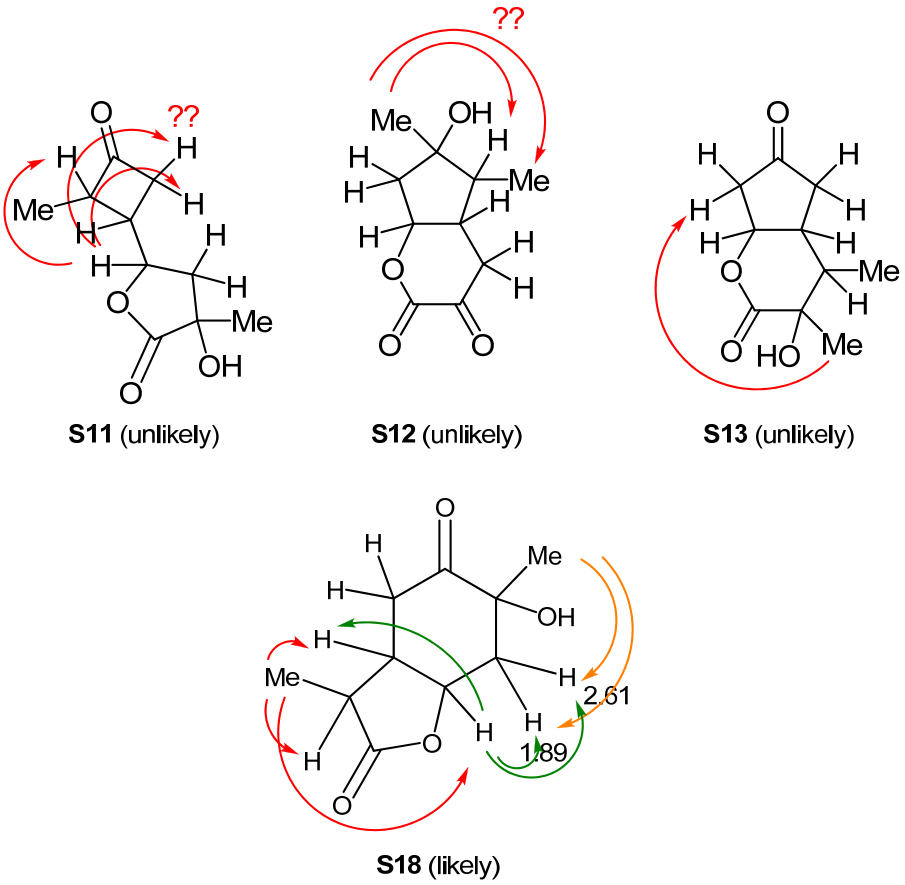

??
O
H
H
Me
H
H
H
H
O
Me
OH
O
S11 (unlikely)
??
Me
OH
H
H
Me
H
H
H
H
H
O
H
O
O
S12 (unlikely)
O
H
H
H
H
Me
H
O
HO
Me
S13 (unlikely)
O
H
Me
H
OH
H
Me
H
H
2.61
H
1.89
H
O
O
S18 (likely)

Figure I.2.15

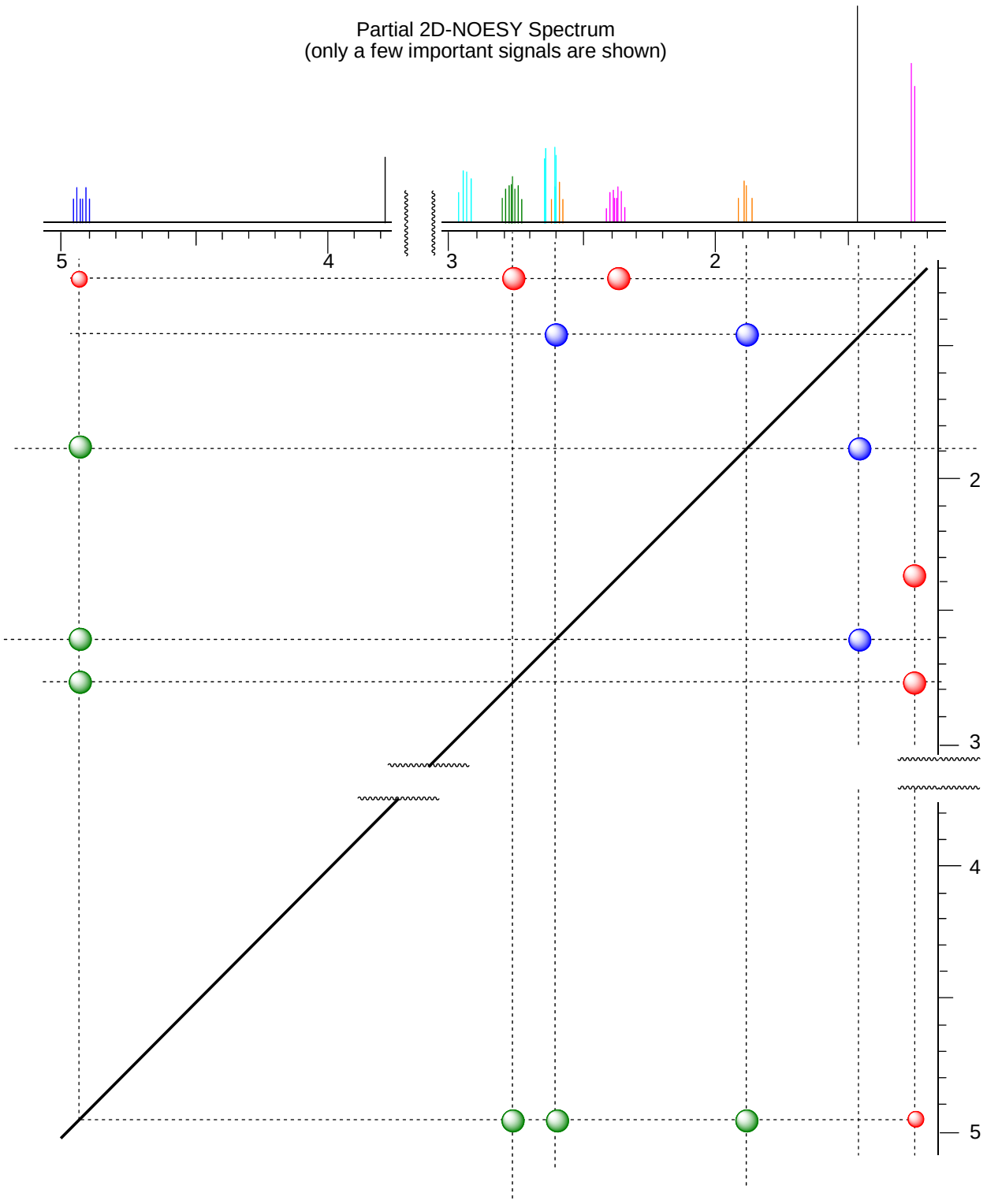

Partial 2D-NOESY Spectrum
(only a few important signals are shown)
5
4
3
2
2
3
4
5

Figure I.2.16

Stucture **S18** is thus that of the unknown **1**. We must now determine its stereochemistry. Figure I.2.17 shows the only 3-dimensional structure of the unknown that fits best with the observed nOes. The conformation chosen is that which should be the most stable, therefore the most prominent in solution. However, conformational equilibrium is fast and the NMR experiment will detect an average of the conformation. Other configurations or conformations would not explain satisfactorily the observed nOe data.

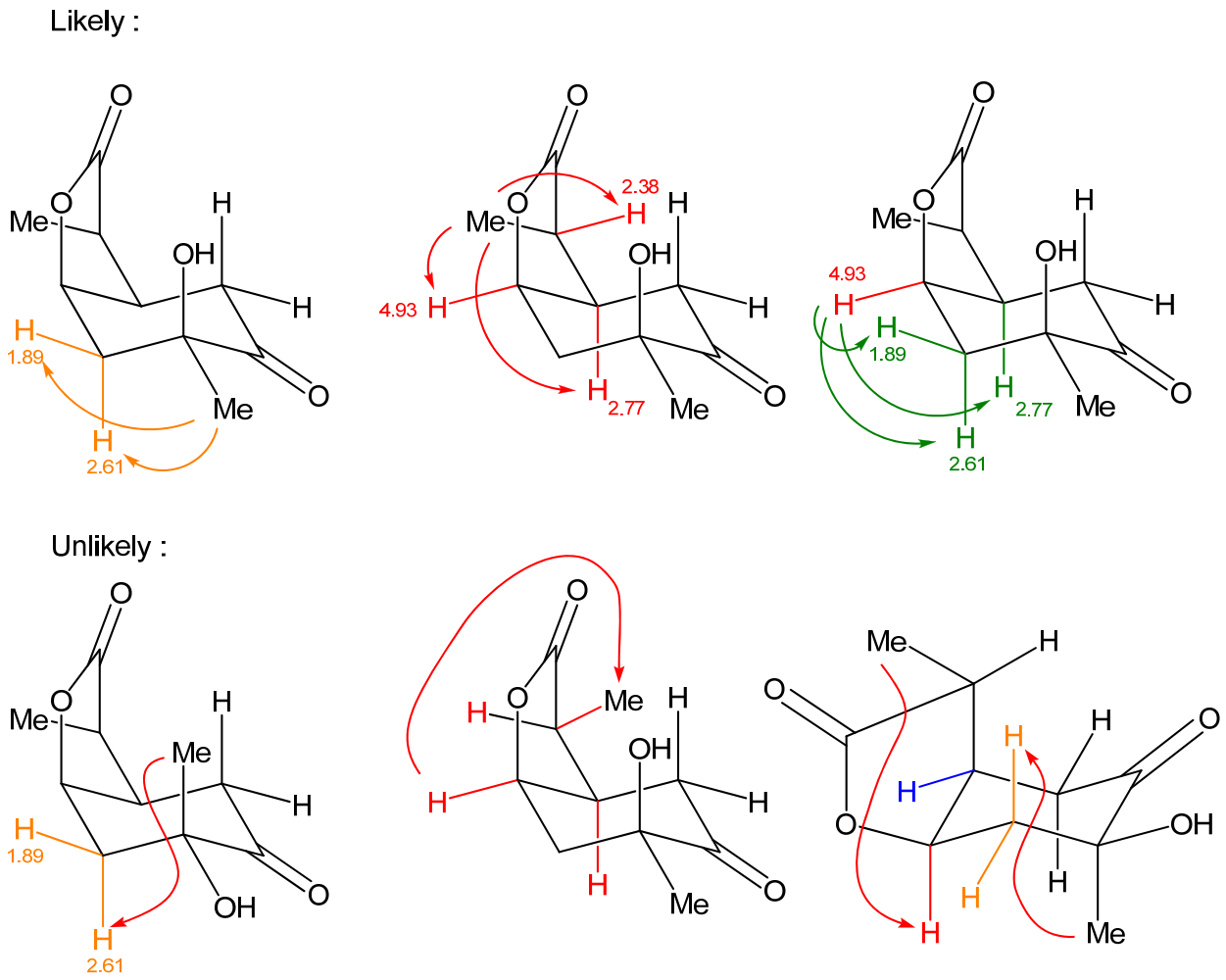


Figure I.2.17

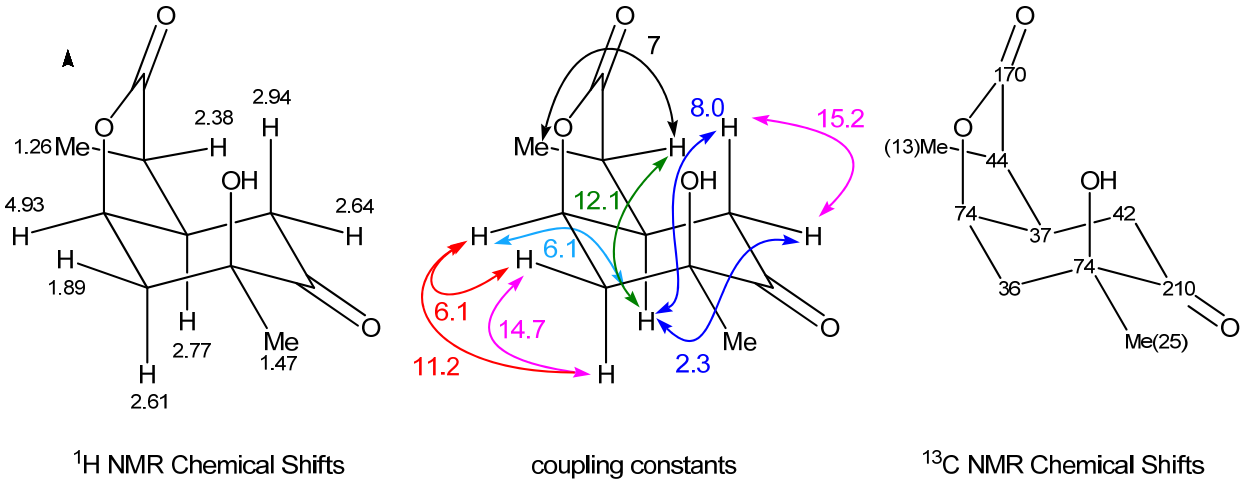
We now have a complete understanding of the original ^{1}H nmr spectrum and ^{13}C nmr spectrum and can go back and assign each signal and coupling constant in those spectra. This is pictured in Figure I.2.18. The coupling constants follow the Karplus rule of dependence on the torsional angle. The rather small constants between the proton at 2.38 δ and the two methylene

protons at 2.64 and 2.94 δ are due to the fact that the conformation must not be perfectly chair and the coupling constants are the mean average of all the conformations in solution.

Note that one enantiomer was arbitrarily drawn for clarity but it could be that the unknown has the opposite absolute stereochemistry. Since the nmr spectrum of enantiomers are identical, it is not possible to assign the absolute stereochemistry at this time. There are several ways to determine the absolute stereochemistry of chiral molecules. If it turns out that the "unknown" was in fact a known compound, we can measure the "optical rotation" of the compound and thus determine what is the absolute stereochemistry by comparing with the literature value. This is indeed the case here, and searching the literature will reveal the identity of **unknown 1** as **paeonilactone A**. In cases where the "unknown" is a new compound, however, then we must revert to chemical derivatization. Enantiomers have identical spectral properties (except for the optical rotation) but diastereomers do not. The latter can be differentiated by nmr spectroscopy. Therefore we could react paeonilactone A with an **optically pure** chiral auxiliary such as O-acetate mandelic acid of **known absolute configuration** and obtain **one** optically pure compound that could be either **S20** or **S21** (Figure I.2.19). To determine the absolute configuration **S20** or **S21** would be simple if the compound is crystalline. We would then grow single crystals, and run an X-Ray crystallographic analysis. This analysis would give us the structure of the derivative with the exact **relative** stereochemistry between the mandelic acid moiety and the paeonilactone part of the molecule. Because we already know the **absolute** configuration of the mandelic acid moiety, we can deduce the **absolute** stereochemistry of paeonilactone A. Note that the X-Ray data could give the wrong **absolute** stereochemistry but **not** the wrong **relative** stereochemistry of the whole derivative. Since we know the absolute configuration of the mandelic acid moiety, we could make the correction.

If the chiral derivative of paeonilactone A is not crystalline, then the determination of the stereochemistry gets complicated. One could do more nOe experiments to determine the interatomic distances in the derivative. However, note that any chiral auxiliary attached to the alcohol of paeonilactone A can freely rotate. This implies that the nOe would give an average distance between the irradiated hydrogen and the proximal protons. Making any deduction from such nOe experiments implies a good knowledge of the conformational behavior of the derivative and could otherwise lead to false deductions. One could try and make a rigid chiral derivative of the molecule with a conformational behavior that is better understood. Then, careful nOe experiments could lead to useful information that in turn could help us deduce the structure of the compound.

PAEONILACTONE A



^{1}H NMR Chemical Shifts coupling constants ^{13}C NMR Chemical Shifts

Figure I.2.18

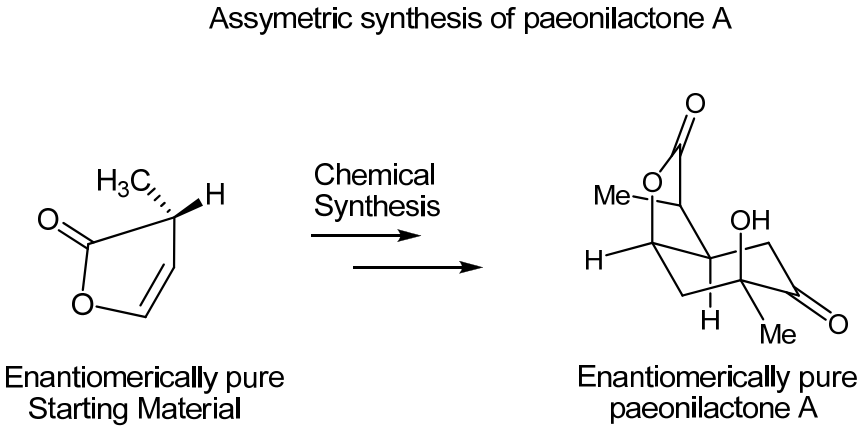
Possible chiral derivative of Paeonilactone A

S20 **S21**

Figure I.2.19

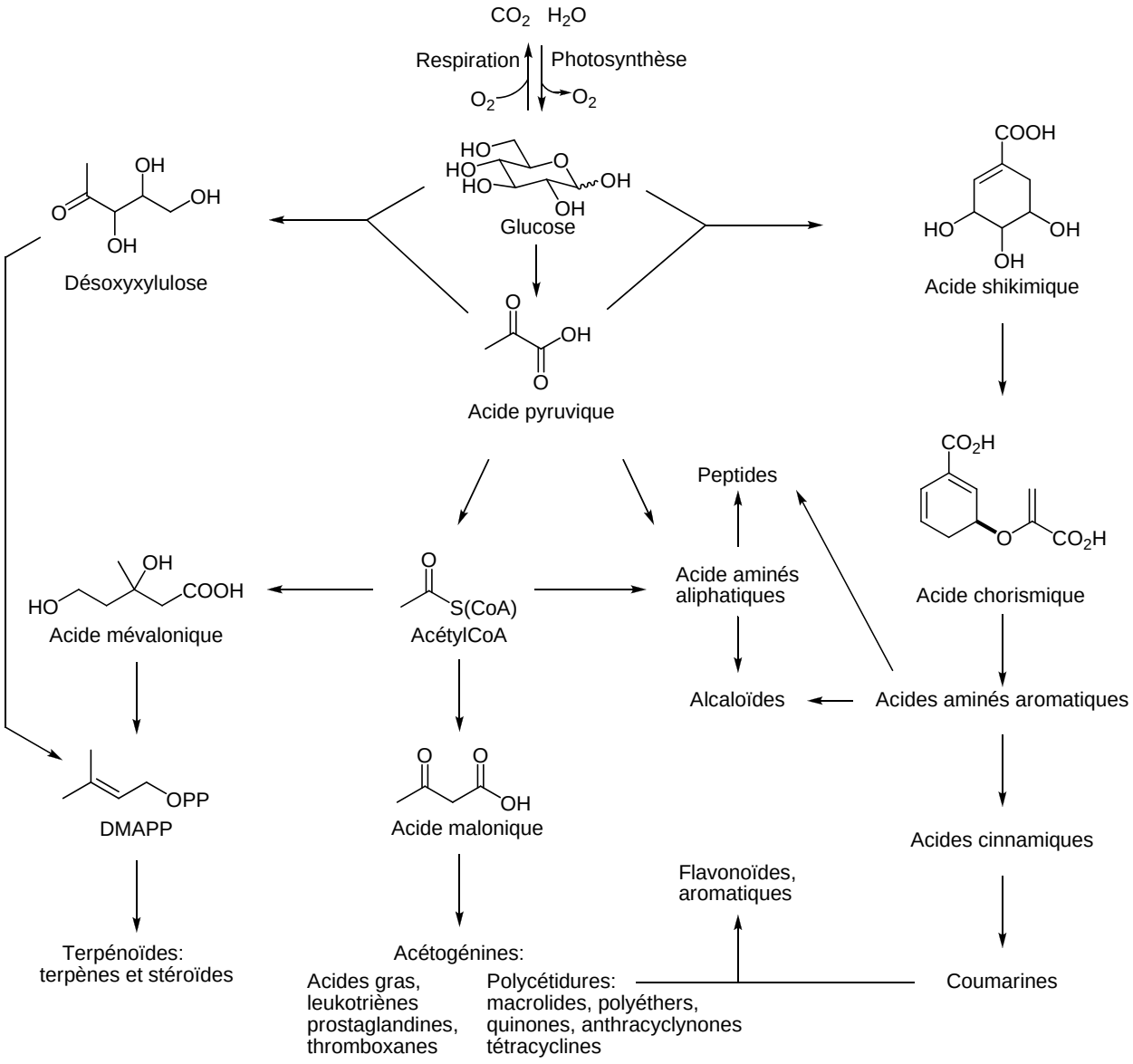
However, despite all these efforts, the conclusions would be thwarted by the uncertainties about the conformation of the molecule or by the necessity to derivatize the molecule in many chemical steps. Each step might have its own stereochemical uncertainty etc. Very often the assignment of the absolute stereochemistry of a particular compound has to await its **total synthesis** by synthetic organic chemists. To be able to assign the absolute stereochemistry of paeonilactone A by total synthesis, the starting material used must be chiral and of **known absolute configuration** (Scheme I.2.1). In addition, each reaction that creates new chiral centers must do it in either a predictable way or the stereochemistry of each new chiral center must be unambiguously proven. When the synthesis is completed, and both the synthetic and natural

compounds are proven to be identical by spectroscopic methods, the measurement of the optical rotation of both products would give the absolute configuration of paeonilactone A.



Scheme I.2.1

Biosynthèse (biogénèse)



Principales routes de biosynthèse de métabolites secondaires

II. Terpénoïdes

II.1 Introduction et historique:

Les terpénoïdes ou isoprénoïdes constituent une famille de produits naturels ayant en commun une origine biosynthétique résultant formellement d'un assemblage d'unités « isoprène » (5 carbones) selon un arrangement tête à queue (le plus répandu) ou mixte tête (t) à queue (q) et queue à queue. L'unité isoprène et les arrangements naturels les plus fréquents, soit tête à queue et queue à queue, sont illustrés à la figure II.1.1. Notez que dans les livres écrits dans la langue anglaise, la condensation queue à queue est souvent référée comme étant 'tête à tête' (head-to-head).

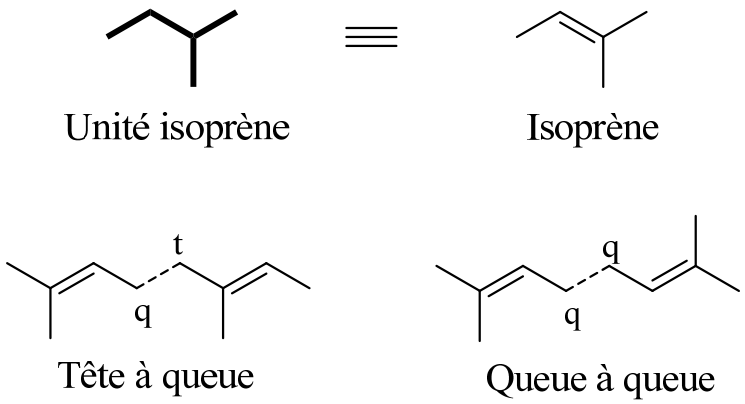


Figure II.1.1

Des exemples sont donnés aux figures II.1.2 et II.1.3. Le farnésol, le β-pinène, le β-cadinène et le caoutchouc ont des arrangements tête à queue. Le squalène et le β-carotène présentent un arrangement mixte. La vitamine E (α-tocophérol) possède une chaîne isoprénique avec assemblage tête à queue et une portion qui n'est pas isoprénoïde. Le β-pinène et le β-cadinène sont des exemples de mono- et sesquiterpène, respectivement, où l'arrangement des chaînes isopréniques est régulier (tête à queue) mais où il y a eu cyclization subséquente des intermédiaires. Dans le lanostérol, il y a eu plusieurs réarrangements du squelette initial car l'assemblage tête à queue est brisé dans une portion de la molécule. Ce dernier est formé à partir du saqualène et possède donc un assemblage queue à queue qui est disparu à cause des réarrangements.

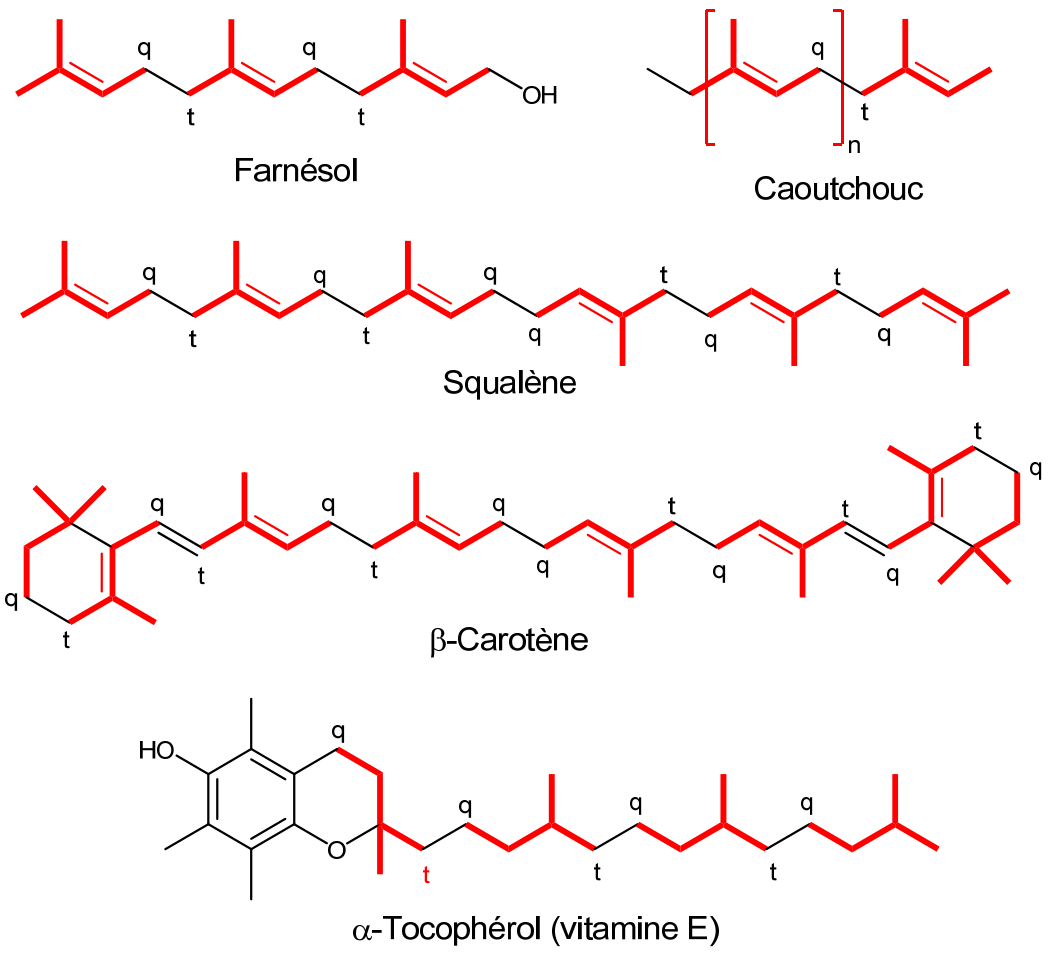

q
q
t
t
OH
Farnésol
q
t
n
Caoutchouc
q
q
q
t
t
t
t
q
q
q
Squalène
t
q
q
q
q
t
t
q
t
t
t
q
q
q
t
β-Carotène
HO
q
O
q
q
q
t
t
t
α-Tocophérol (vitamine E)

Figure II.1.2

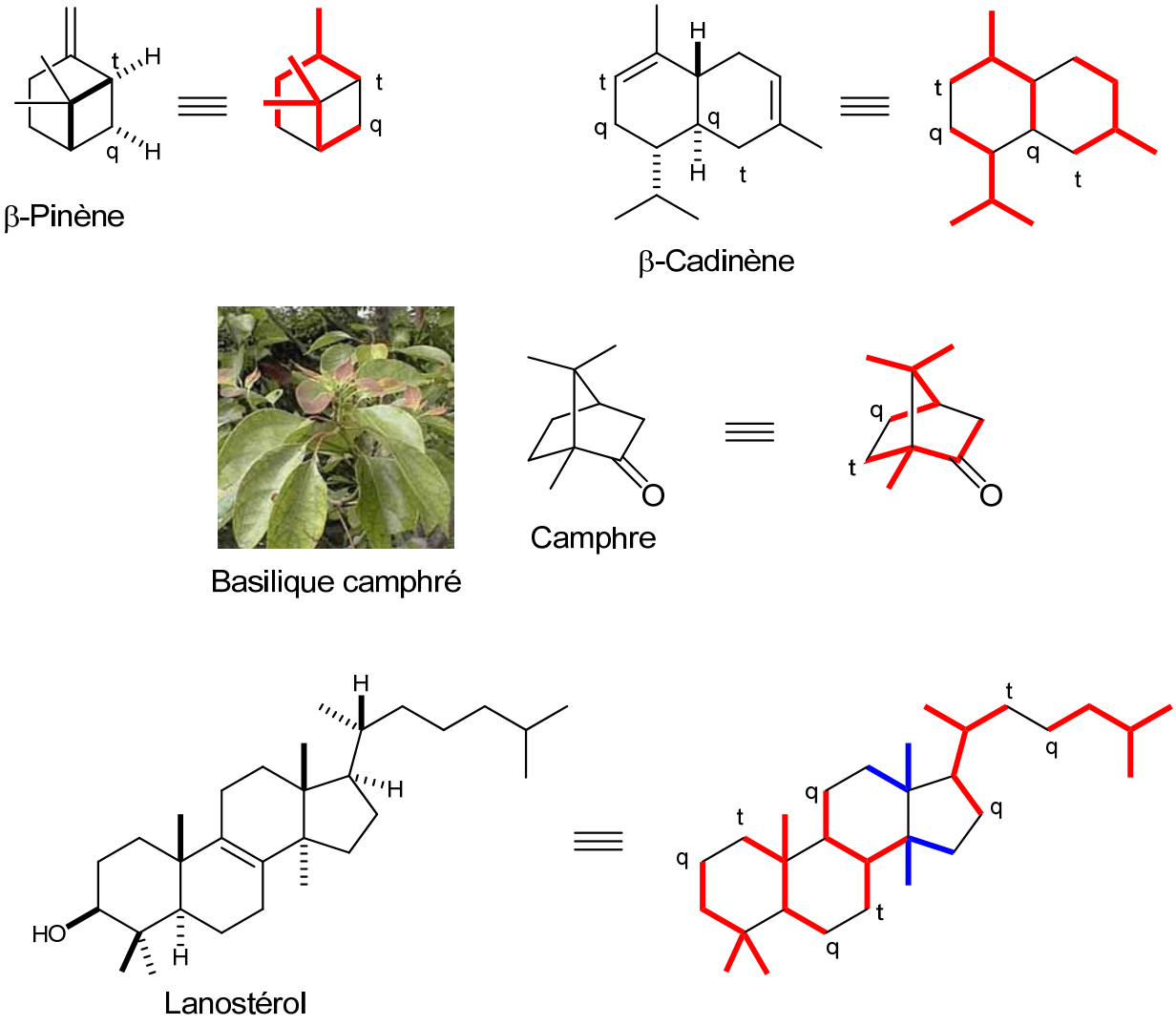


Figure II.1.3

Plusieurs terpènes se retrouvent fréquemment dans les huiles essentielles d'herbes, d'arbustes, d'arbres, de fleurs, bref dans les plantes en général. Par exemple, le camphre nous vient d'Arabie, extrait du camphre basil et connu sous le nom de *kafur*, qui est devenu *camfora* en latin, puis *camphre* en français (figure II.1.3). Le chimiste Al-Kindi, au 9^{ième} siècle, décrit la première recette pour extraire le camphre qui est depuis utilisé alors comme épice alimentaire dans des mets traditionnels comme le *tharid* (fait de pain dans bouillon de viande). Le processus pour obtenir des huiles essentielles par extraction 'graisseuse' est connu depuis le moyen âge. Arnaud de Villanosa a décrit au 12^{e} siècle la distillation d'huile de romarin et de sauge. Le *Dispensatorium valerii cordi*' de 1592 contient la description de plus de 60 huiles essentielles.

En 1818, le chimiste J. J. Houston de Billardière analyse le contenu de la 'térébenthine'. Le nom est en fait dérivé du grec *therebinth* qui est un arbre dont la sève contient plusieurs terpènes

et dont les huiles étaient connues de par le monde et peut-être la première source de ‘térébenthine’.

Les terpénoïdes se retrouvent partout dans la nature et dans les plantes en particulier. Leur rôle biologique, surtout celui des petits terpènes, n’est pas très bien compris. Très longtemps, nous avons cru qu’ils n’avaient aucune importance pour les organismes vivant. Goodwin a même postulé qu’ils étaient de simples sous-produits de l’évolution par lesquels passaient les différentes routes biosynthétiques vers les hormones essentielles. D’autres comme Bu'Lock pensaient que les terpènes étaient synthétisés pendant l’hibernation ou les périodes de dormance afin de garder l’appareil enzymatique de l’organisme fonctionnel et prêt pour les tâches plus importantes qui les attendaient au réveil. Il y a des terpénoïdes qui sont des vitamines comme certains carotènes, d’autres qui sont des hormones comme certains stéroïdes. D’autres jouent un rôle important dans les mécanismes de défense, de communication, etc.

Pour bien comprendre l’origine des terpénoïdes, il faut connaître leur biosynthèse. Les premières études remontent 19^{ième} siècle. En 1826, Faraday établissait que le rapport atomes de carbone sur atomes d’hydrogène dans le caoutchouc était de 5:8.

En 1860, Williams fit la pyrolyse du caoutchouc et obtînt un produit volatil ayant un rapport C:H de 5:8 qu’il nomma ***isoprène***. En 1884, Otto Wallach (prix Nobel de chimie 1910) montra que la composition élémentaire de plusieurs huiles essentielles était un multiple de C_5H_8. L’idée que ces substances puissent avoir une relation biosynthétique commençait déjà à circuler. En ce sens, Wallach serait le père de la « règle de l’isoprène » qui stipule que tous les terpénoïdes sont constitués d’un multiple d’une même unité isoprène.

FORMULA	OLD NAME	CURRENT NAME
C_5H_8	Pentenes	
$C_{10}H_{16}$	True terpenes	Monoterpenes
$C_{15}H_{24}$	Tripentenes	Sesquiterpenes
$C_{20}H_{32}$	Tetrapentenes	Diterpenes
$C_{25}H_{40}$		Sesterterpenes
$C_{30}H_{48}$		Triterpenes
$C_{40}H_{64}$		Tetraterpenes
$(C_5H_8)x$	Polyterpenes	Polyisoprenes

Cependant, comme la structure de l’isoprène et des isoprénoïdes était inconnue à cette époque, il a fallut attendre Ruzicka (prix Nobel de chimie 1939 conjointement avec Adolf Butenandt) dans les années 1920 pour l’énoncé de cette règle, soit 20 ans après la détermination de la structure de l’isoprène. Les avancées en biosynthèse ont été possibles par l’utilisation du

marquage isotopique, d'abord en utilisant des isotopes radioactifs, ^{14}C et ^{3}H (tritium). Des organismes (bactéries, plantes, animaux) ou des organes (par exemple le foie) sont nourris avec un composé marqué sur un atome spécifique. Si ce composé est impliqué dans la biosynthèse d'un métabolite, le métabolite sera marqué à des positions bien spécifiques. La dégradation chimique en petites molécules marquées de structure connue permet de déterminer la position exacte des isotopes dans le terpénoïde. C'est un processus long et lent. Les études biosynthétiques ont été considérablement accélérées et facilitées par l'utilisation de la RMN permettant d'utiliser le marquage au ^{13}C et ^{2}H.

II.1.1 Le cholestérol

L'histoire de la biosynthèse du cholestérol est fascinante et démontre le génie des chimistes de l'époque. Après avoir nourris des souris avec du D_2O pendant 3 mois de sorte que le contenu en deutérium atteigne 1.5%, Schoenheimer et son équipe ont remarqué un niveau d'incorporation si élevé (la moitié en poids) qu'il ne pouvait plus y avoir de doute que le cholestérol était biosynthétisé à partir de petites molécules, peut-être même celles impliquées dans le métabolime de gras et de sucre. Sonderhoff et Thomas ont ensuite remarqué un niveau d'incorporation de deutérium élevé dans des stérols lors d'expériences impliquant le trideutéro-acétate et ont conclu que ceux-ci devaient être fabriqués à partir d'unités acétates. Était-ce la seule unité impliquée? Comment une unité structurellement si simple pouvait-elle être le précurseur d'un stéroïde de 27 carbones et quatre cycles carbonés?

Konrad Bloch, un scientifique venu à New York, réfugié de l'Allemagne, en 1936 est devenu une sommité mondiale sur la biosynthèse des lipides et en particulier du cholestérol. Grâce à la nouvelle technique du marquage isotopique, il a pu démontrer, avec son collaborateur Rittenberg, que **tous** les carbones du cholestérol provenaient de l'unité acétate.

En utilisant de l'acétate (CH_3COONa) marqué sur le méthyle et sur le carbone du carbonyle, Robinson (Robert, prix Nobel de chimie 1947), Bloch, Conforth, Popják et d'autres ont montré que le cholestérol était fabriqué à partir d'unités acétates et ont pu déterminer quels carbones provenaient du méthyle et quels carbones provenaient du carbonyle (schéma II.1.1). Déjà, en 1937, Robinson avait postulé que le squalène, isolé du foie de requin, était possiblement un précurseur du cholestérol. Ils ont donc montré que l'acétate était converti d'abord en squalène et ce dernier converti en lanostérol puis le lanostérol converti en cholestérol.

Le chaînon manquant entre l'unité acétate et le squalène a persisté pendant plusieurs années. Bloch, en s'inspirant des travaux de Bonner et Arreguin qui proposait que le caoutchouc était un polymère d'unités isoprènes, a lui-même proposé que l'unité isoprène puisse être synthétisée à partir de trois unités acétates. Il a fallu plusieurs années pour déterminer la nature de l'unité

isoprène précurseur de tous les terpénoïdes et pour trouver comment cette unité était fabriquée à partir de l'acétate. C'est un groupe de Merck qui a fait la découverte de l'acide mévalonique, alors qu'ils cherchaient des substituants nutritionnels à l'acétate.

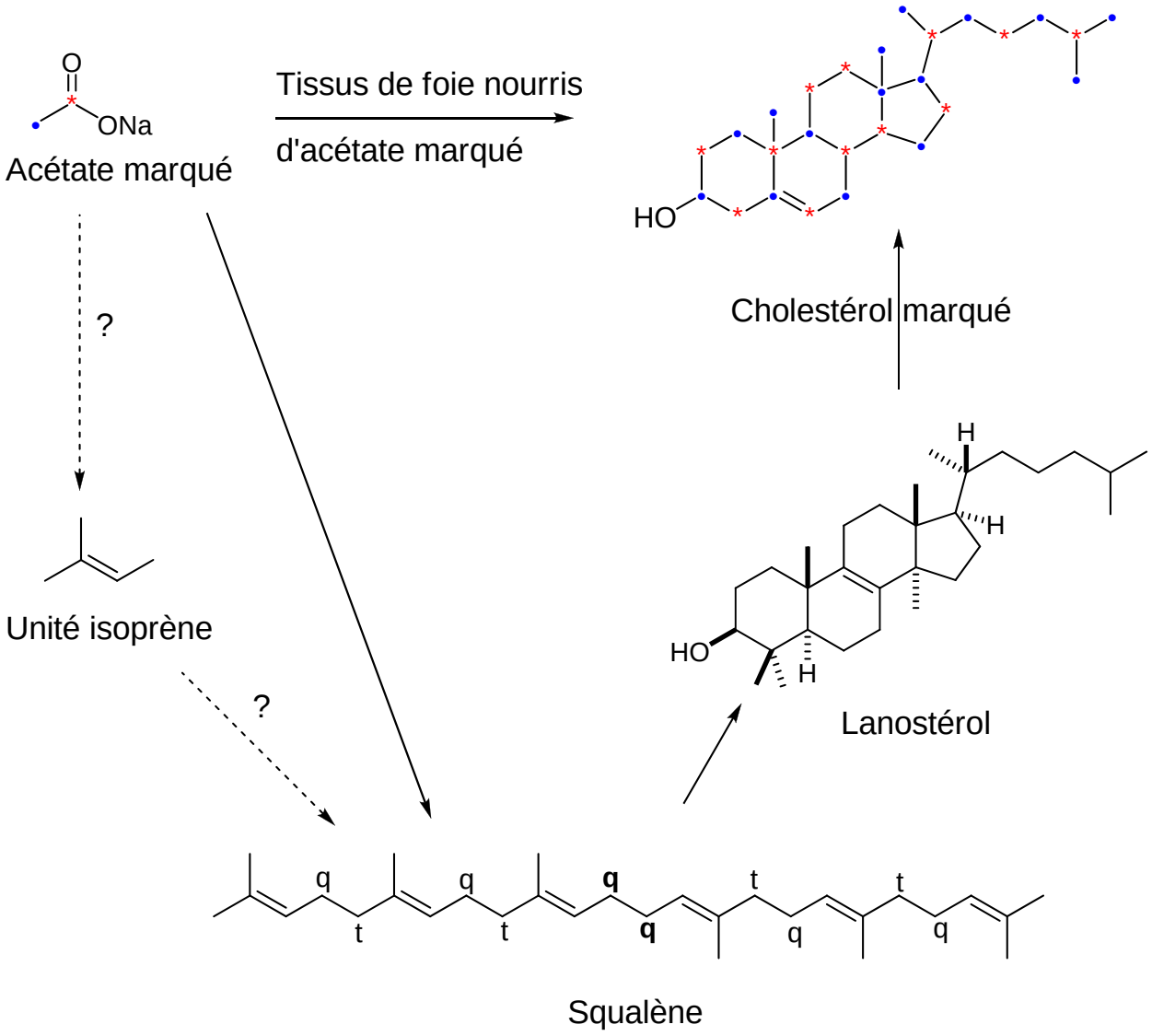


Schéma II.1.1

L'unité isoprène est le ***pyrophosphate d'isopentényle*** (IPP selon l'acronyme de « *i*sopentenyl*pyro*phosphate ») et son isomère le ***pyrophosphate de diméthylallyle*** (DMAPP selon l'acronyme de « *dim*ethyl*a*llyl*pyro*phosphate »). Leur conversion est catalysée par un enzyme (figure II.1.4) Ce sont les acronymes anglais que nous allons utiliser. Voyons plus en détails comment sont biosynthétisés les terpénoïdes ou isoprénoïdes.

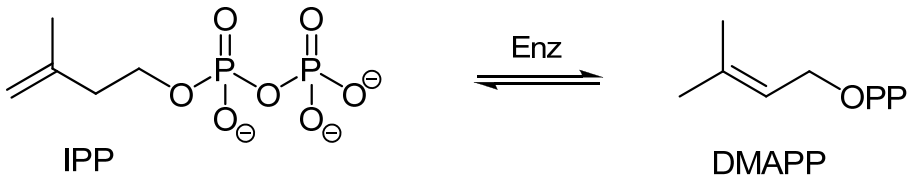


Figure II.1.4

II.2 Biosynthèse des pyrophosphates d'isopentényle (IPP) et de diméthylallyle (DMAPP)

II.2.1 Généralité sur la biosynthèse des terpènes.

L'IPP et le DMAPP sont les précurseurs de tous les terpénoïdes (au-delà de 30,000 composés). Il existe deux routes de biosynthèse de l'unité isoprène, la route de l'acide mévalonique (MVA pour « mevalonic acid ») et la route du 2-désoxyxylulose. Cette dernière n'a été découverte qu'au début des années 1980 et la raison de cette découverte tardive est le dialogue croisé ou promenade croisée entre les deux routes dans un même organisme (schéma II.2.1).[1] Les deux routes partent du glucose et utilisent l'acide pyruvique comme intermédiaire. Ce dialogue croisé entre les deux routes métaboliques serait dû à un échange de métabolites intermédiaires entre le cytoplasme et les chloroplastes dans certaines plantes. Voici quelques exemples : dans *Catharanthus roseus*, le phytol et le β-carotène sont biosynthétisés par la route désoxyxylulose et le sitostérol par la route MVA ; dans *Cyanidium caldarium*, le phytol l'est par la route désoxyxylulose et l'ergostérol par la route MVA ; dans *Daucus carota*, le phytol est biosynthétisé par la route désoxyxylulose mais le sistostérol et le stigmastérol le sont par la route MVA ; et dans *Heteroscyphus planus*, le phytol est biosynthétisé par les deux routes et le β-carotène l'est par la route MVA. Les structures de ces métabolites sont illustrées à la figure II.2.1. Dans les levures et le règne animal, on ne connaît que la route MVA. Dans les bactéries, les deux routes sont utilisées mais celles-ci se retrouvent rarement dans une même bactérie.

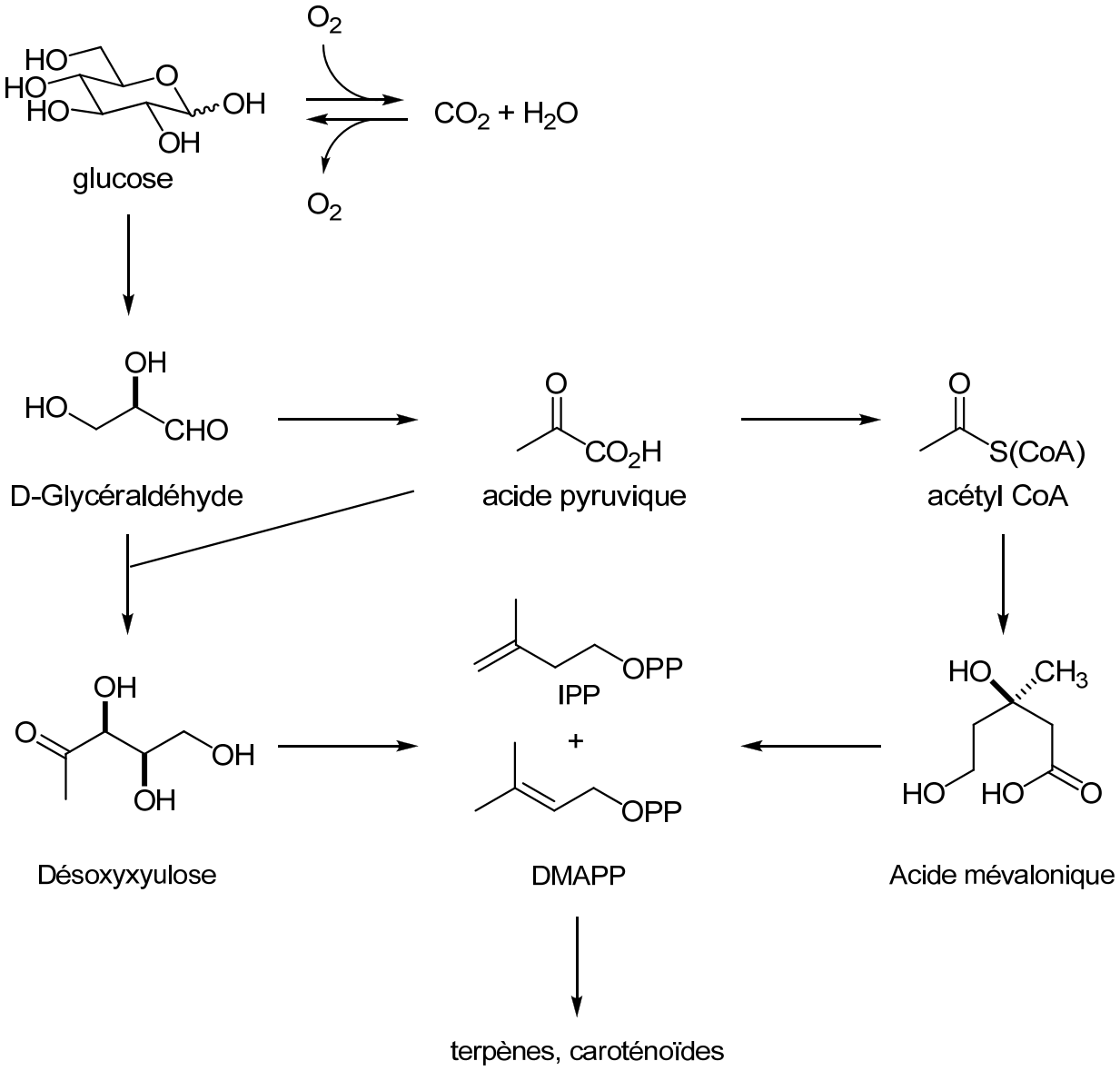

O2
HO
HO
HO
O
OH
OH
glucose
O2
CO2 + H2O
OH
HO
CHO
D-Glycéraldéhyde
O
CO2H
acide pyruvique
O
S(CoA)
acétyl CoA
OH
O
OH
OH
Désoxyxyulose
OPP
IPP
+
OPP
DMAPP
HO
CH3
HO
HO
O
Acide mévalonique
terpènes, caroténoïdes

Schéma II.2.1

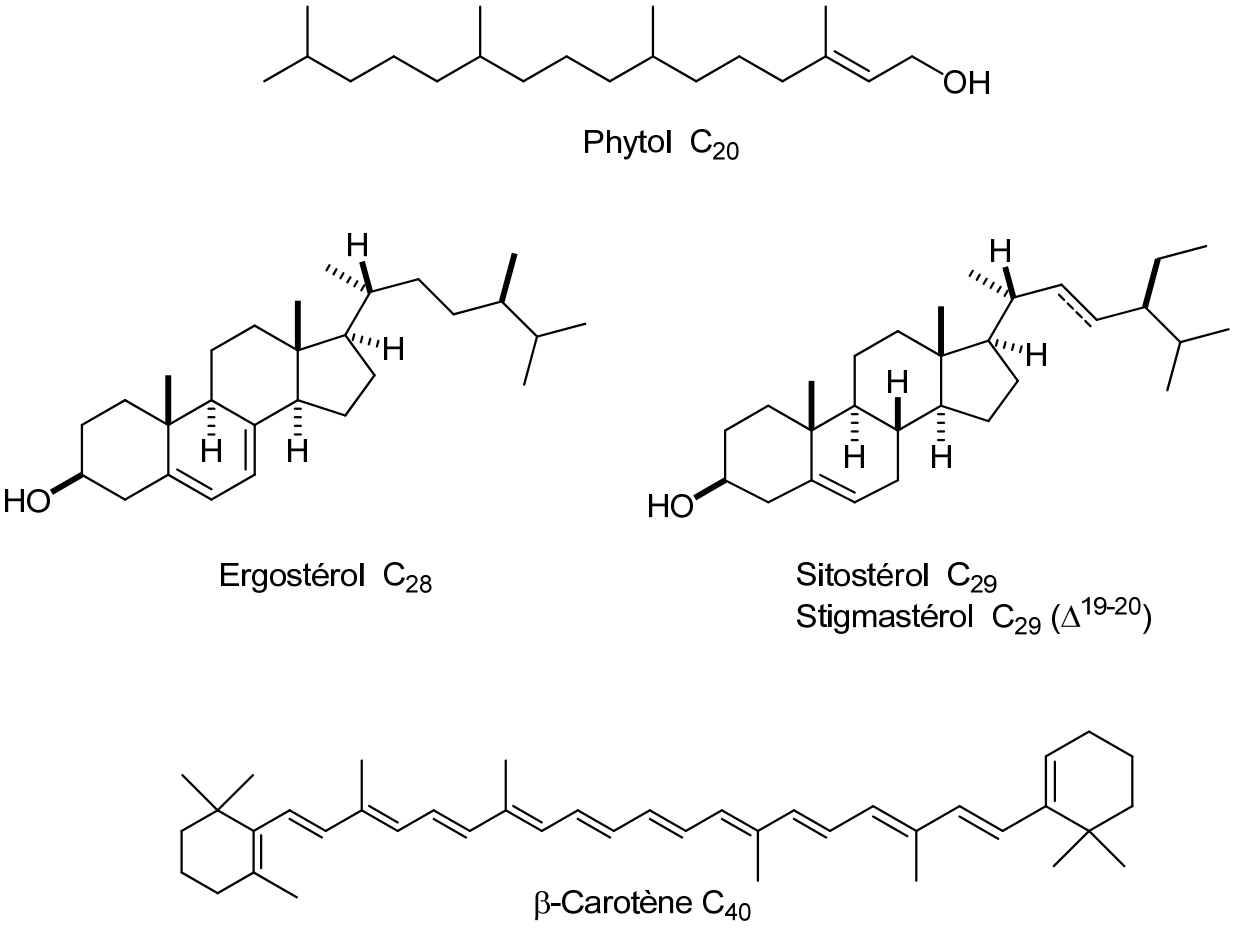
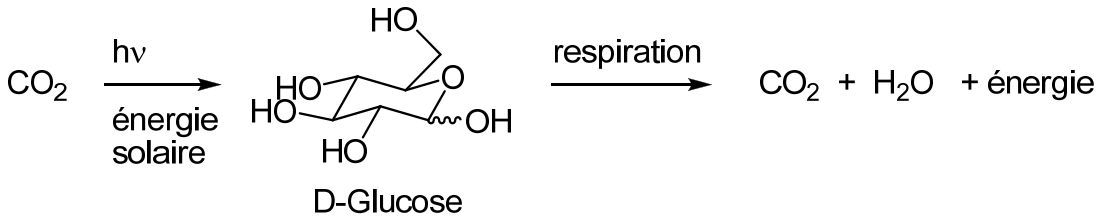


Figure II.2.1

II.2.2 Conversion du glucose en acide pyruvique

Les sucres sont fabriqués à partir du bioxyde de carbone et l'eau par photosynthèse dans les plantes vertes. La fixation biosynthétique du CO_2 est le processus inverse de la respiration (équation 1).



Les étapes de la conversion du D-glucose en acide pyruvique sont données dans les schémas II.3 et II.4. Le glucose est d'abord converti en fructose par tautomérie céto-énol. Ce dernier est

phosphorylé en dérivé 1,6-diphosphate qui est ensuite transformé en D-glycéraldéhyde-3-phosphate par une rétro-aldolisation (schéma II.2.2). Une oxydation de l'aldéhyde en acide via son hémithioacétal, une isomérisation du groupement phosphate de la position 3 à la position 2 et un élimination d'eau donne le phosphoénolpyruvate qui est hydrolysé en acide pyruvique. Le glucose marqué aux positions 1 et 6 conduit ainsi à l'acide pyruvique marqué à la position 3 (méthyle) (schéma II.2.3). Il faut noter que le D-glycéraldéhyde-3-phosphate est un intermédiaire clé dans la route désoxyxylulose (section II.2.4).

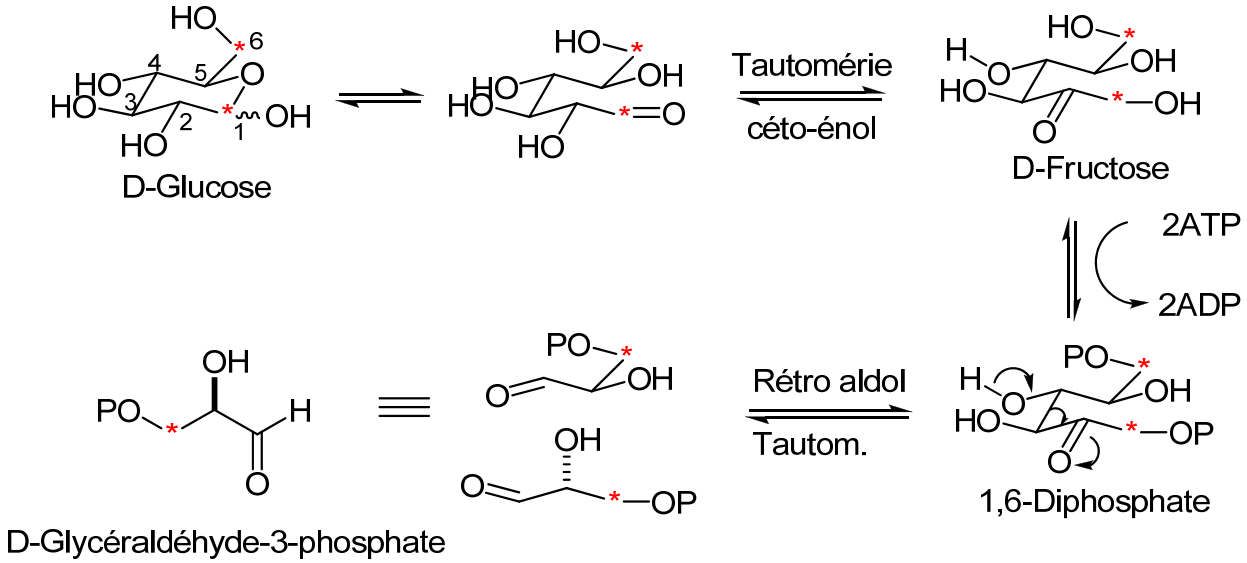


Schéma II.2.2

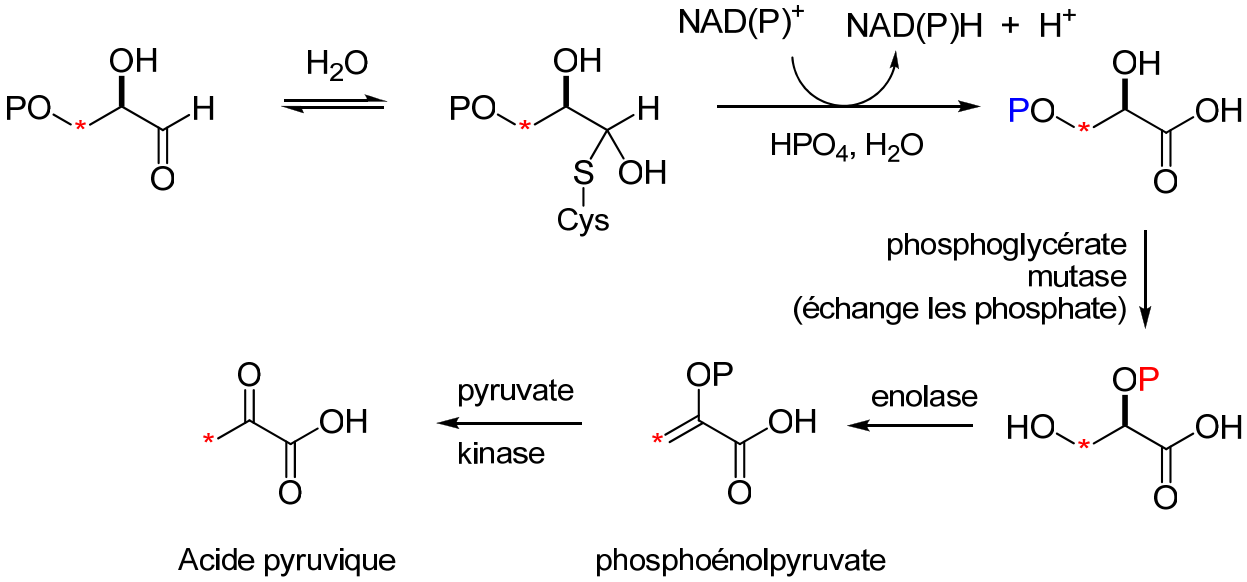


Schéma II.2.3

L'acide pyruvique est au centre de bien des chemins biosythétiques. Le schéma II.2.4 est très important pour bien comprendre l'ensemble des transformation biochimiques du CO_2 en métabolites secondaires. Pour les sections qui suivent, l'acide pyruvique servira de matière première pour la synthèse des 'building blocks' que sont l'IPP et le DMAPP.

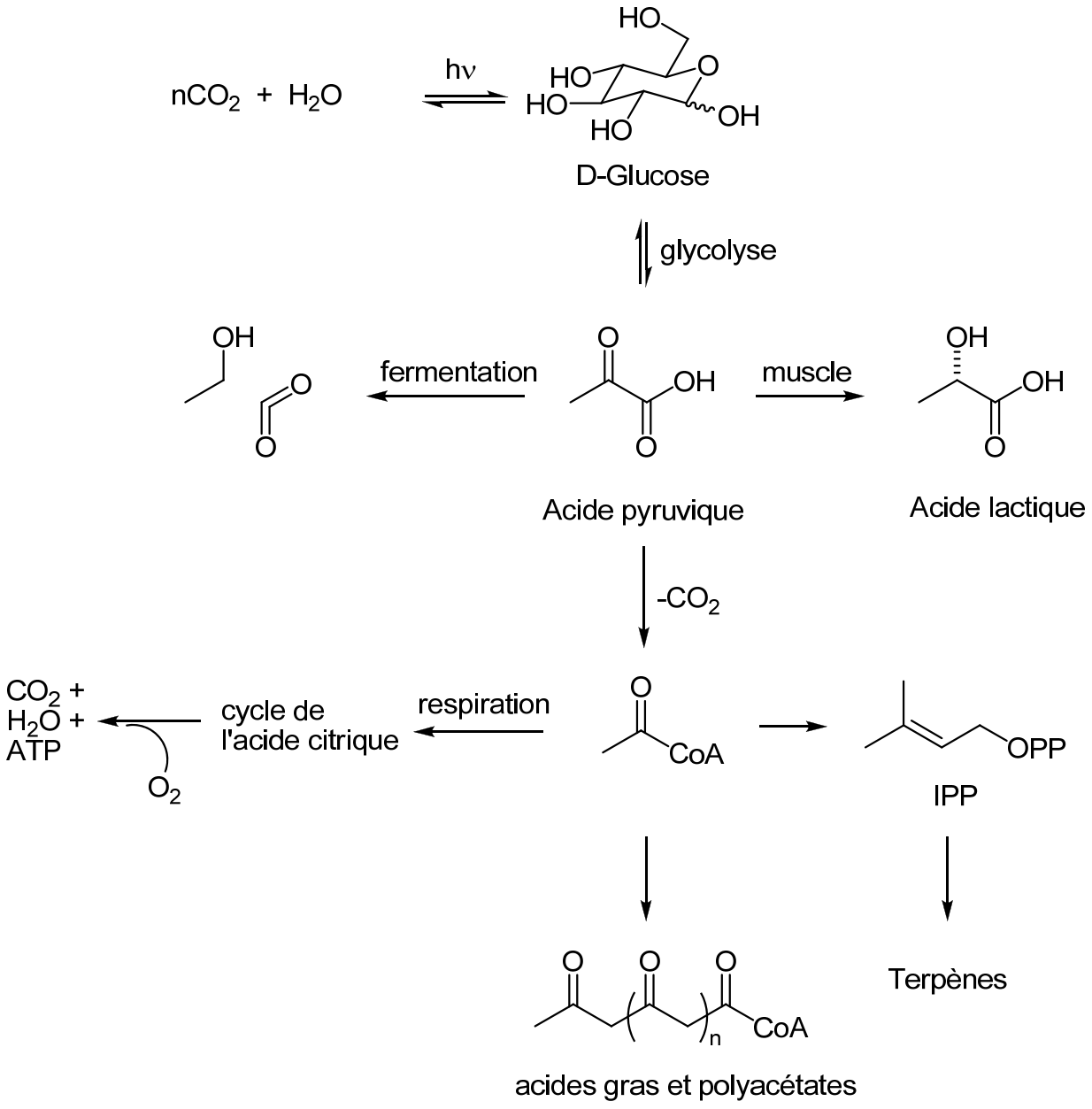


Schéma II.2.4

II.2.3 Route acide mévalonique

II.2.3a Conversion de l'acide pyruvique en acétylcoenzyme A (acétylCoA)

Trois coenzymes présentes dans le complexe enzymatique sont impliquées dans la conversion de l'acide pyruvique en acétylcoenzymeA (ou acétylCoA): la coenzyme pyrophosphate de thiamine (TPP), la coenzyme lipoate et la coenzyme A (CoA). La structure de la CoA est donnée à la figure II.2.2.

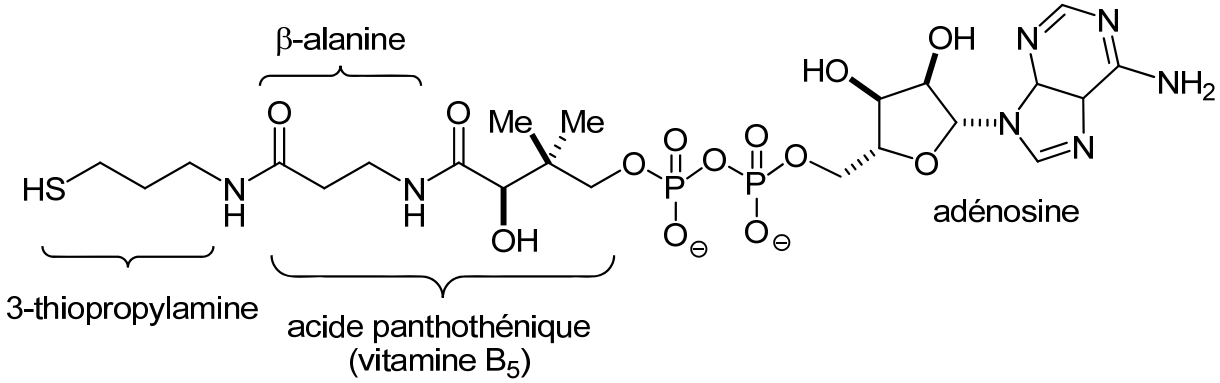


Figure II.2.2

Les étapes de la conversion de l'acide pyruvique sont décrites au schéma II.2.5. La coenzyme TPP est convertie en carbène nucléophile par arrachement de proton créant ainsi un accepteur (**I_1**) pour la décarboxylation, laquelle donne un nucléophile ou réducteur (**I_2**). Le nucléophile/réducteur **I_2** réagit avec la coenzyme lipoate (électrophile/oxydant) pour donner l'intermédiaire **I_3** (schema II.2.6). La coenzyme TPP est régénérée avec formation du thioester **I_4**. Cette transformation globale (de TPP à **I_4** avec regeneration de la TPP) possède un équivalent en synthèse au laboratoire. La réaction s'appelle la réaction de Stetter et utilise aussi un carbène nucléophile. Par contre, plutôt qu'une décarboxylation, comme dans la biosynthèse de l'acétylCoA, la réaction de Setter fait appel à une base pour générer l'équivalent de **I_2**. Bien sûr, les électrophile varient et ne sont pas restreint à des disulfures. La régénération du carbène, cependant, est identique.

La transestérification de **I_4** avec la coenzyme A donne l'acétylCoA (marquée sur le méthyle) et un dithiol qui est oxydé pour régénérer la coenzyme lipoate (disulfure cyclique).

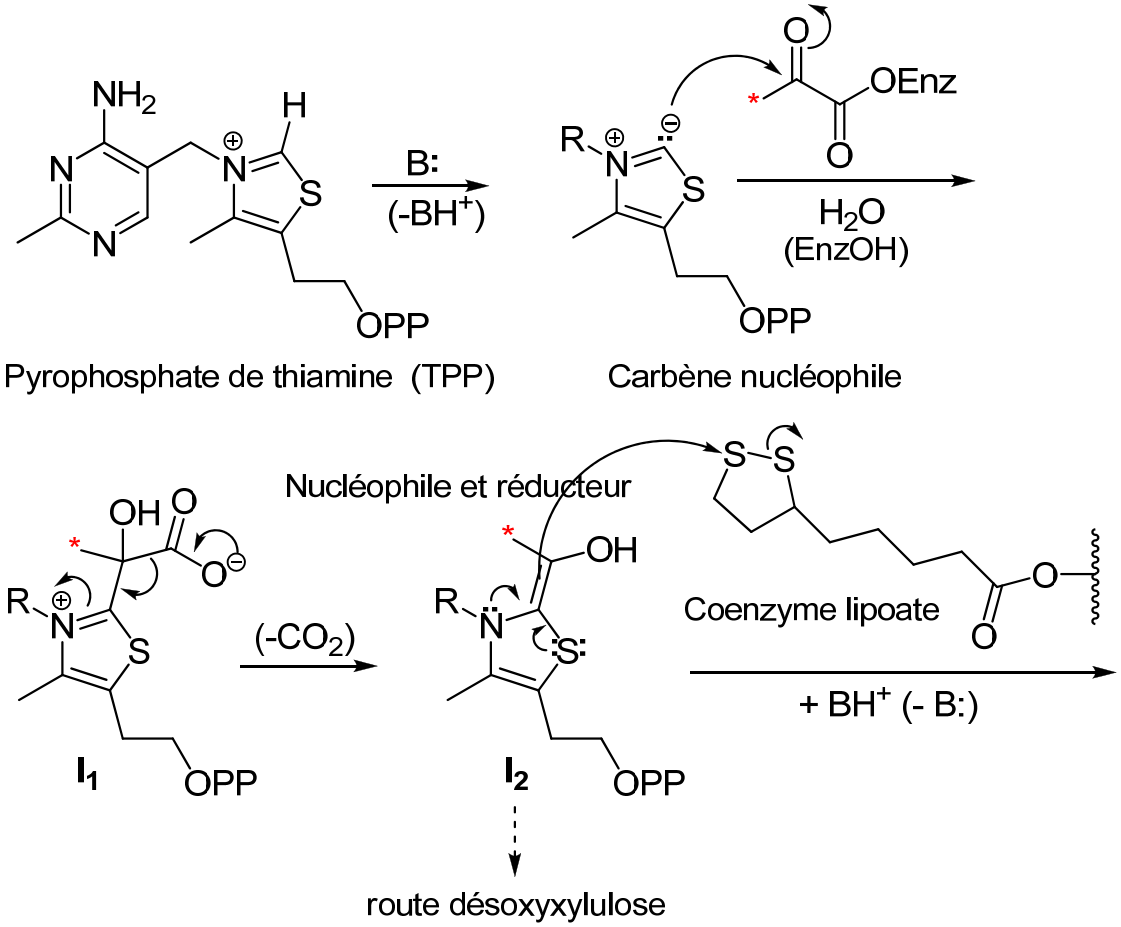


Schéma II.2.5

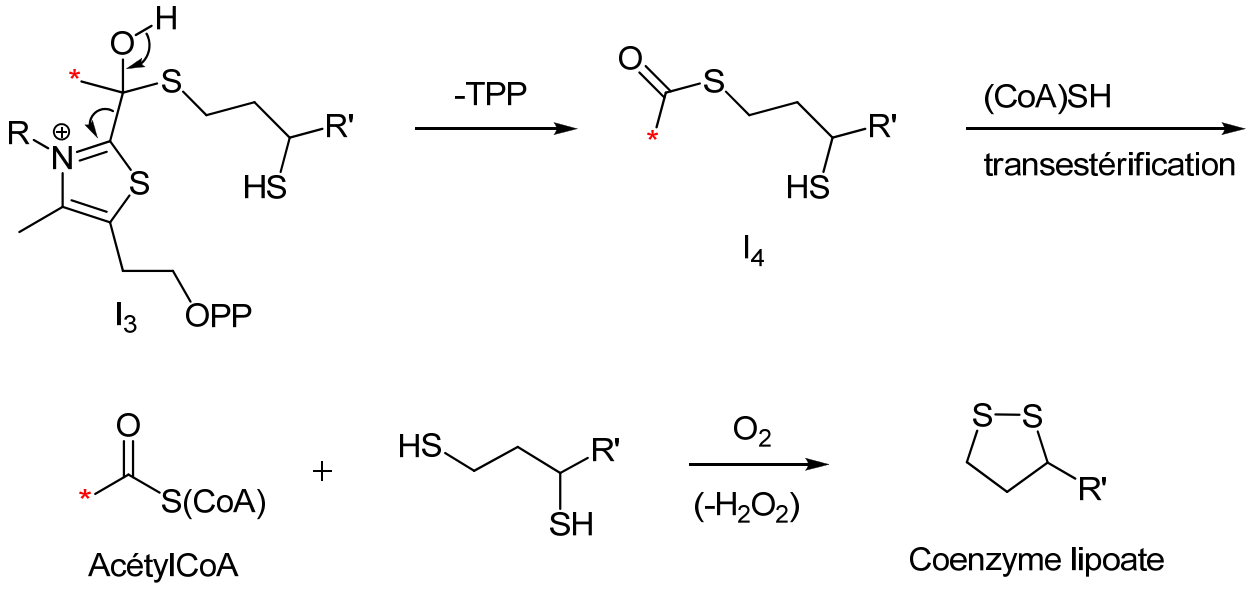


Schéma II.2.6

II.2.3b Conversion de l'acétylCoA en acide mévalonique (MVA)

La conversion décrite au schéma II.2.7 ci-dessous comprend les étapes suivantes : attachement du groupement acétyle à un système enzymatique par transestérification de la coenzyme A; condensation de Claisen avec la coenzyme A pour donner l'acétoacétylCoA; formation de la coenzyme A de l'acide 3-hydroxy-3-méthylglutarique (HMGCoA pour « *h*ydroxy*m*ethyl-*g*lutarylcoenzyme A ») par condensation aldolique; réduction du thioester en alcool pour donner l'acide 3R-mévalonique et la coenzyme A. La démonstration que l'acide 3R-mévalonique était un intermédiaire de la biosynthèse des terpénoïdes a été le point tournant de l'élucidation du mécanisme de cette biosynthèse.

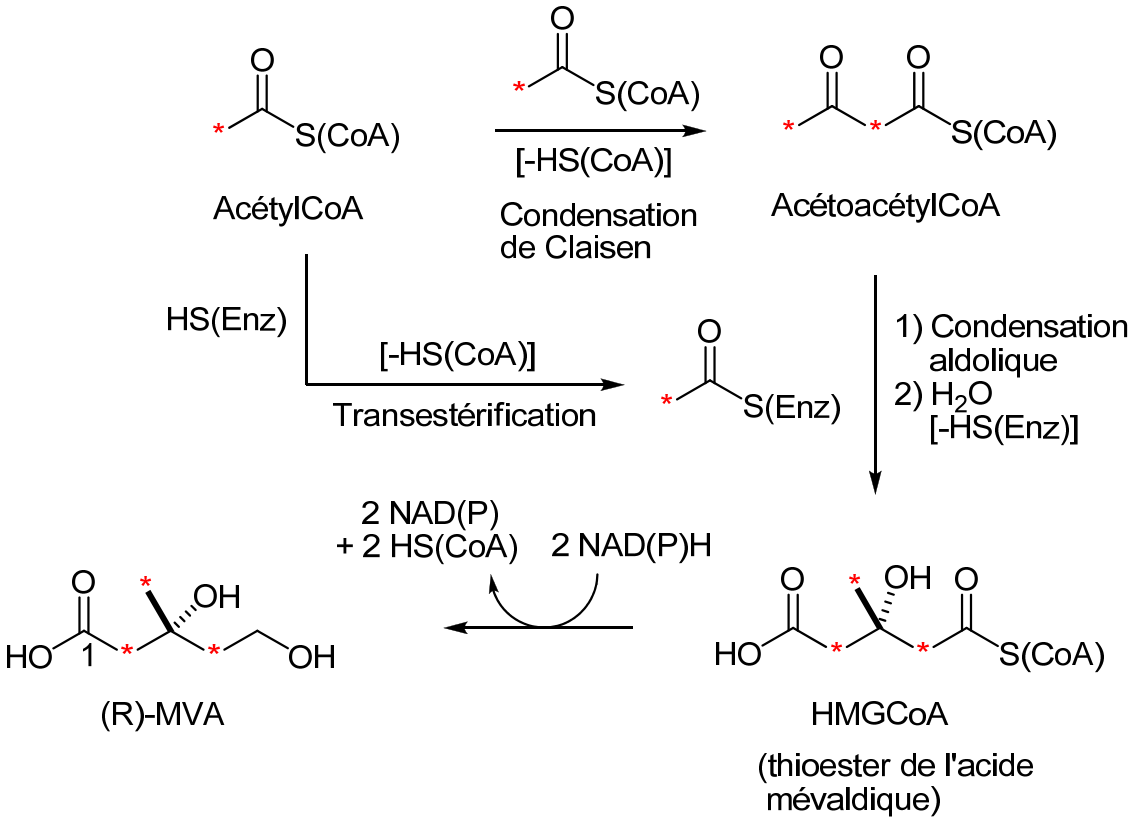


Schéma II.2.7

II.2.3c Conversion de l'acide mévalonique en unités isoprèniques : le pyrophosphate d'isopentényle (IPP) et le pyrophosphate de diméthylallyle (DMAPP)

L'acide mévalonique est d'abord converti en son dérivé 5-pyrophosphate (MVAPP) puis le groupement hydroxyle en position 3 est phosphorylé (schéma II.2.8). Une décarboxylation et élimination de l'ion phosphate (un très bon nucléofuge) par fragmentation de Grob donne l'IPP. L'IPP est ensuite isomérisé en DMAPP à l'aide d'une isomérase. Dans le site actif de l'enzyme, il y a protonation de la face *si* de l'IPP et un site basique vient enlever le proton pro-R uniquement, ce qui montre à quel point les réactions catalysées par les enzymes peuvent être stéréosélectives. À partir du glucose marqué aux positions 1 et 6, on obtient l'IPP et le DMAPP marqués sur les carbones 2 et 4 et sur le méthyle.

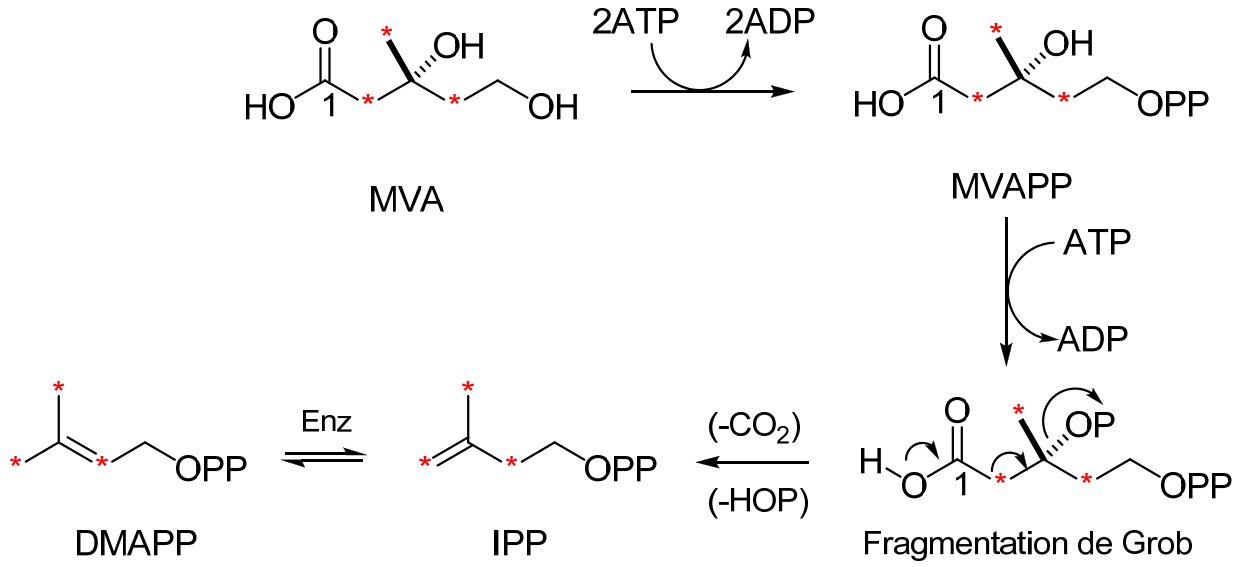


Schéma II.2.8

Contrairement à ce qu'on pourrait penser, bien que les molécules d'IPP et de DMAPP soient achirales, les processus enzymatiques qui mènent à leur formation sont hautement stéréospécifiques. Pour mieux comprendre ce qui se passe, voyons un peu le concept de **pro-chiralité**. La figure II.2.3 montre l'acide mévalonique avec chacun des hydrogènes étiqueté soit pro-R soit pro-S. L'atome pro-chiral, dans chaque cas, est l'atome qui porte le groupement pro-R ou pro-S. **ATTENTION** : les groupements pro-R et pro-S doivent être identiques et ils ne sont pas pro-chiraux, bien qu'ils portent l'étiquette; c'est l'atome qui les porte qui est pro-chiral. La définition de pro-chiralité est double. La vraie définition est : *un centre pro-chiral deviendra chiral à la suite d'une réaction*. Forcément, il est donc achiral avant la réaction. Pour un carbone

cela implique deux substituants identiques (2 H ou 2 Me ou 2 CH_2OH etc) ou une insaturation (double liaison). Cependant, il est possible que nous voulions identifier les groupements pro-R ou pro-S même si la réaction ne produit pas de centre chiraux. Alors, il faut inventer une réaction (fictive) pour pouvoir assigner les étiquettes pro-R et pro-S. On ne peut pas prendre n'importe quelle réaction pour éviter que le centre pro-chiral devienne *S* ou *R* dépendent de la réaction utilisée. Pour cette raison, on doit remplacer chacun des groupements par son isotopes le plus proche. On remplace donc chacun des H par D, chacun des CH_3 par $^{13}CH_3$, chacun des Br par ^{81}Br etc. De cette façon, le centre chiral deviendra *R* ou *S* de façon prévisible. Voyez comment les hydrogènes sont étiquettés pro-R ou pro-S pour l'acide mévalonique (figure II.2.3)

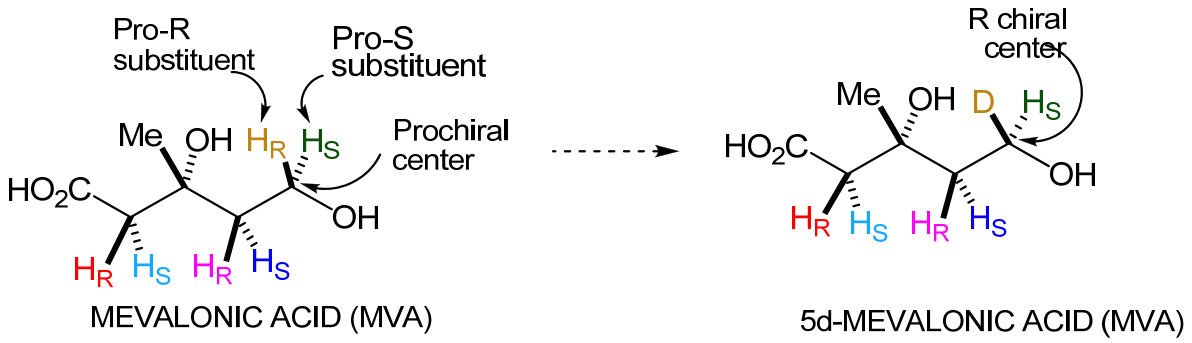


Figure II.2.3

Ce concept est extrêmement utile en biosynthèse. Par remplacement d'un groupement pro-R ou pro-S par son isotope (e.g. un H par un D) il est possible de savoir que la fragmentation de Grob se produit de façon *anti* (figure II.2.4). Lors de l'isomérisation de l'IPP en DMAPP, c'est l'hydrogène pro-R qui est perdu.

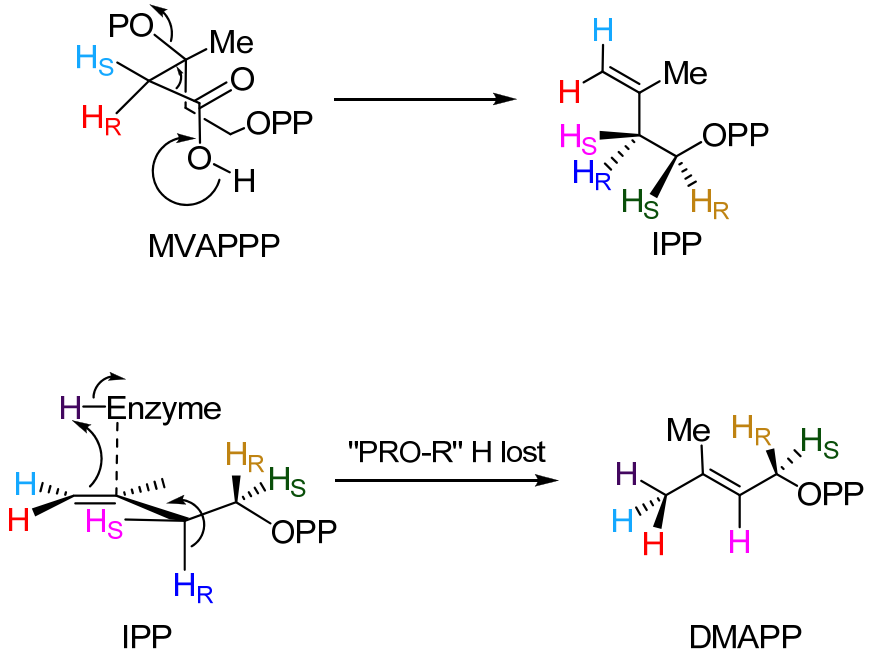


Figure II.2.4

II.2.4 Route désoxyxylulose

II.2.4a Conversion de l'acide pyruvique en phosphate de désoxyxylulose

Tel que vu au schéma II.2.5, la réaction de l'acide pyruvique avec la coenzyme TPP déprotonée conduit à l'intermédiaire **I_1** qui subit une décarboxylation pour donner le nucléophile **I_2**. Comme montré au schéma II.2.9 ci-dessous, l'intermédiaire **I_2** s'additionne au phosphate de glycéraldéhyde provenant du glucose pour donner l'intermédiaire **I_5**. Puis l'élimination (régénération) de la coenzyme TPP donne le phosphate de désoxyxylulose. À partir du glucose marqué aux carbones 1 et 6, on obtient le désoxyxylulose marqué aux carbones 1 et 5.

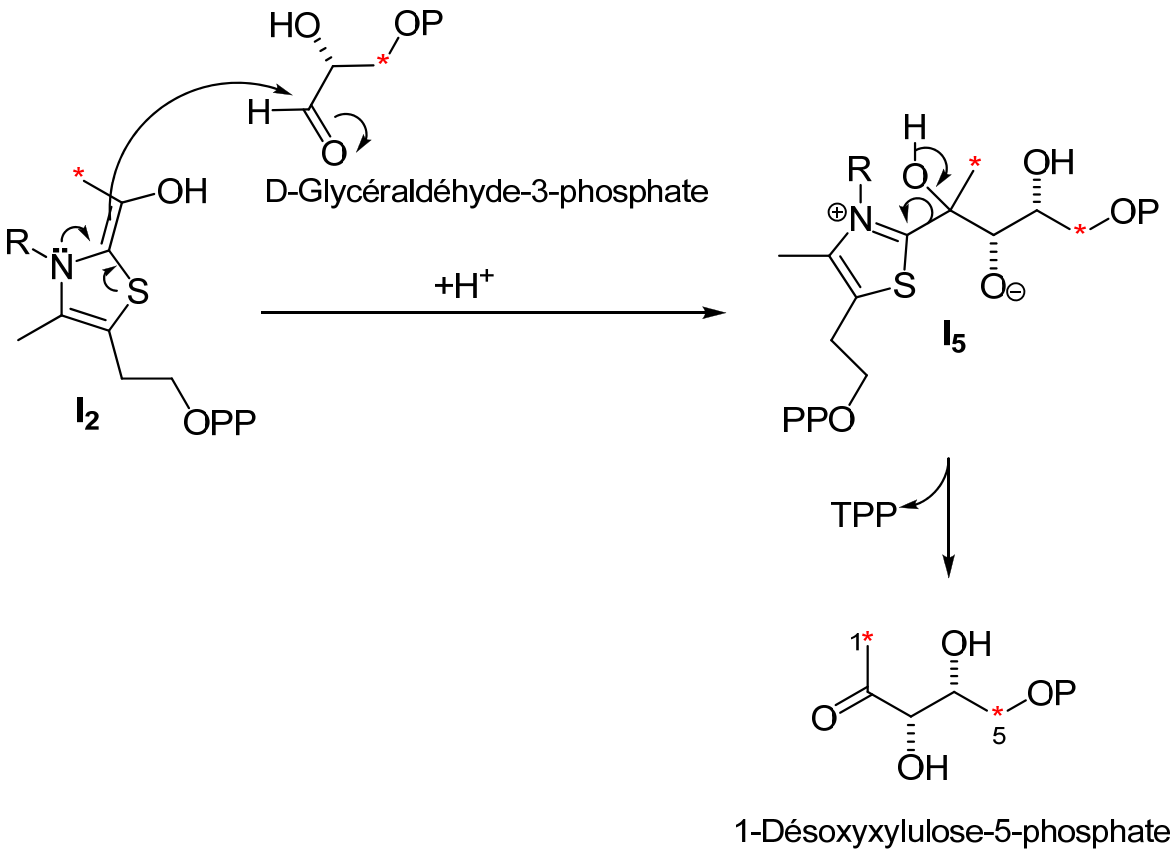


Schéma II.2.9

Le schéma II.2.10 ci-dessous montre la transformation du phosphate de désoxyxylulose en IPP. L'hypothèse courante est que la protonation du carbonyle du 5-phosphate de désoxyxylulose induirait un réarrangement sigmatropique [1,2] donnant le dihydroxy aldéhyde phosphate **I_6**. Ce dernier n'a pas été encore identifié hors de tout doute. Des expériences de marquage isotopique ont montré que l'IPP obtenu par incorporation du glucose marqué aux positions 1 et 6 donne l'IPP marqué sur le méthyle et sur le carbone 1 tel qu'indiqué. L'intermédiaire **I_6** pourrait subir une réduction et la cytidinetriphosphate (CTP) serait apparemment incorporé sur cette molécule tel qu'indiqué. Après phosphorylation, une cyclization produit le composé macrocyclique **I_7**. Une elimination produit le composé pyrophosphorylé **I_8** et le reste des étapes est incertain mais il manque encore une élimination de H_2O ainsi que la réduction de la cétone à l'alcane.

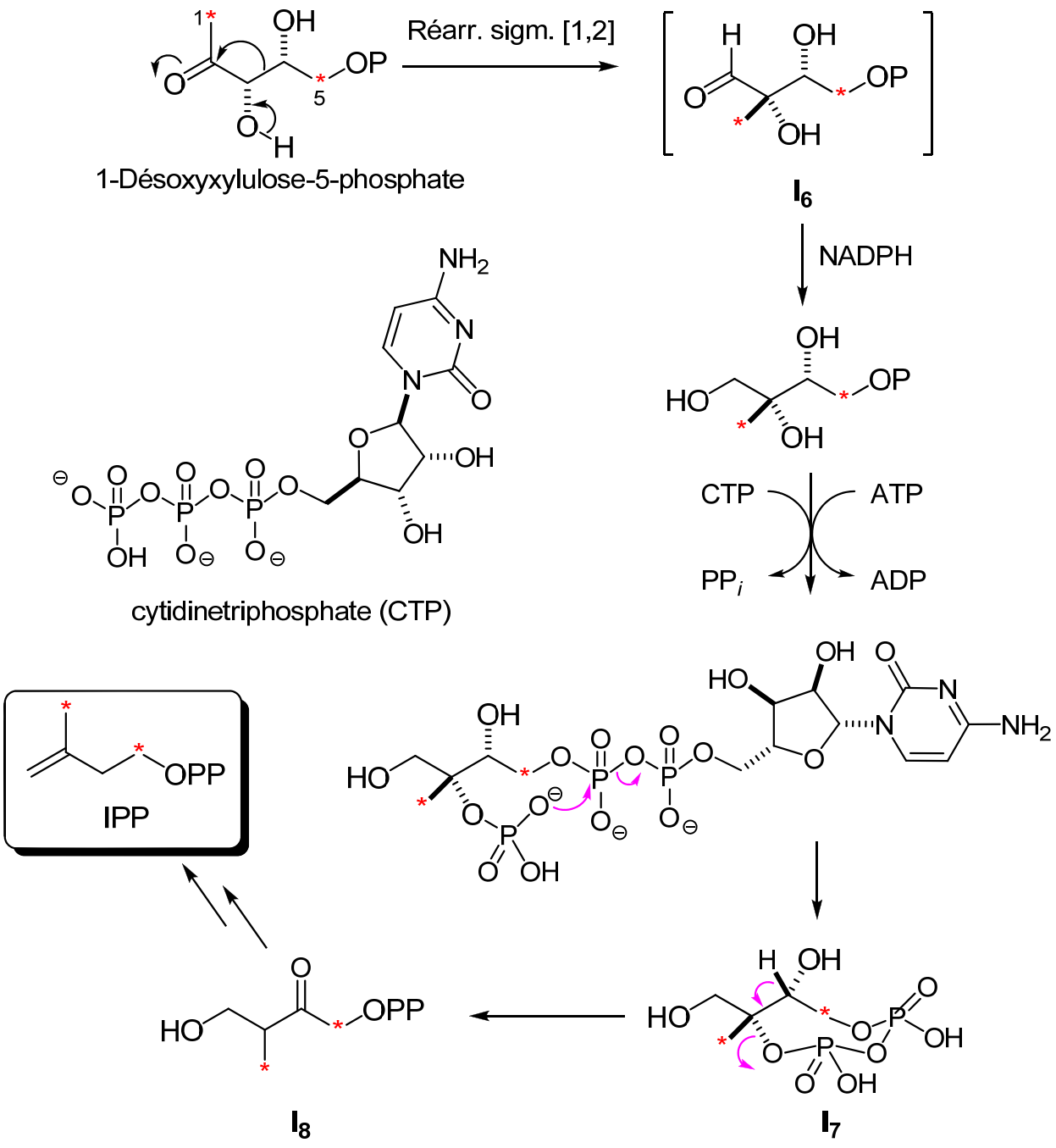


Schéma II.2.10

II.3 Condensations des unités isopréniques (IPP et DMAPP)

II.3.1 Condensations tête-à-queue

Tel que décrit au schéma II.3.1, la condensation du DMAPP et de l'IPP conduit au pyrophosphate de géranyle (GPP pour « géranylpyrophosphate »), le précurseur des monoterpènes constitués de deux unités isoprène reliées tête-à-queue (C_{10}). On ajoute une autre unité isoprène par condensation du GPP avec l'IPP, ce qui donne le pyrophosphate de farnésyle (FPP pour « farnesylpyrophosphate »), le précurseur des sesquiterpènes (C_{15}). L'ajout de l'IPP au FPP donne le pyrophosphate de géranylgéranyle (GGPP), précurseur des diterpènes (C_{20}) et finalement l'ajout de l'IPP au GGPP donne le pyrophosphate de géranylfarnésyle (GFPP), le précurseur des sesterterpènes (C_{25}). Toutes les unités isoprènes sont reliées tête-à-queue. Le mécanisme de ces condensations est discuté de façon plus approfondie au chapitre II.x.

Il y a trois possibilités de mécanisme de condensation entre l'IPP et le DMAPP : a) formation d'un carbocation classique suivie de l'élimination du proton pro-R (H_R); b) formation d'un carbocation non classique suivie de l'élimination du proton H_R; ou c) élimination du proton H_R concertée avec l'attaque du lien double et le départ de PPO^-. L'élimination 100% sélective du proton H_R a été clairement démontrée par marquage isotopique. Par analogie avec l'isomérisation de l'IPP en DMAPP, un genre de complexe avec la double liaison dans le site actif de l'enzyme bloque la face *si* de l'IPP de sorte que le DMAPP doit s'approcher du côté de la face *re* de l'IPP. Une base dans le site actif de l'enzyme vient arracher le proton H_R uniquement.

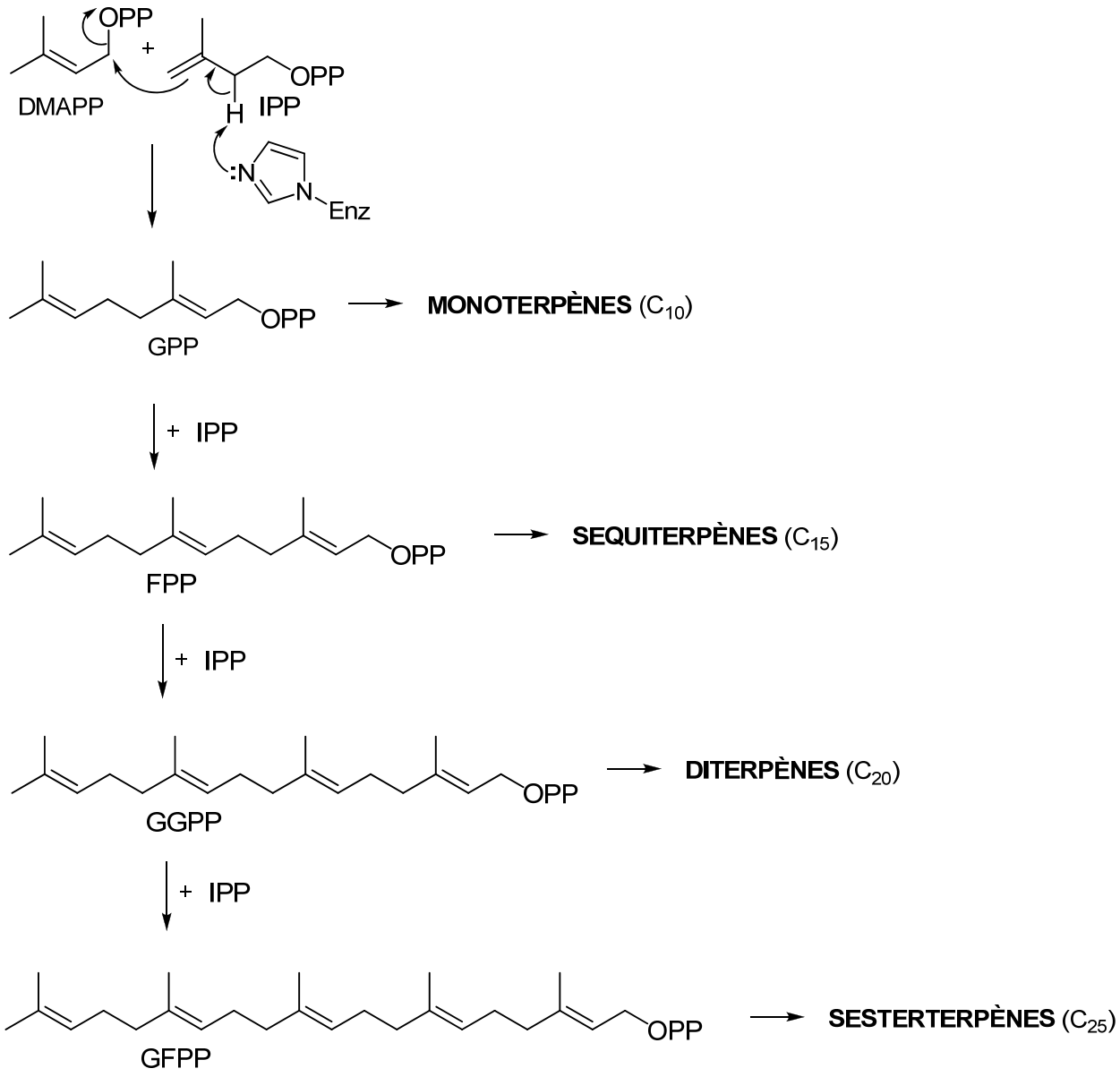

OPP
+
OPP
DMAPP
H
IPP
:N
N
Enz
OPP
GPP
MONOTERPÈNES (C10)
+ IPP
OPP
FPP
SEQUITERPÈNES (C15)
+ IPP
OPP
GGPP
DITERPÈNES (C20)
+ IPP
OPP
GFPP
SESTERTERPÈNES (C25)

Schéma II.3.1

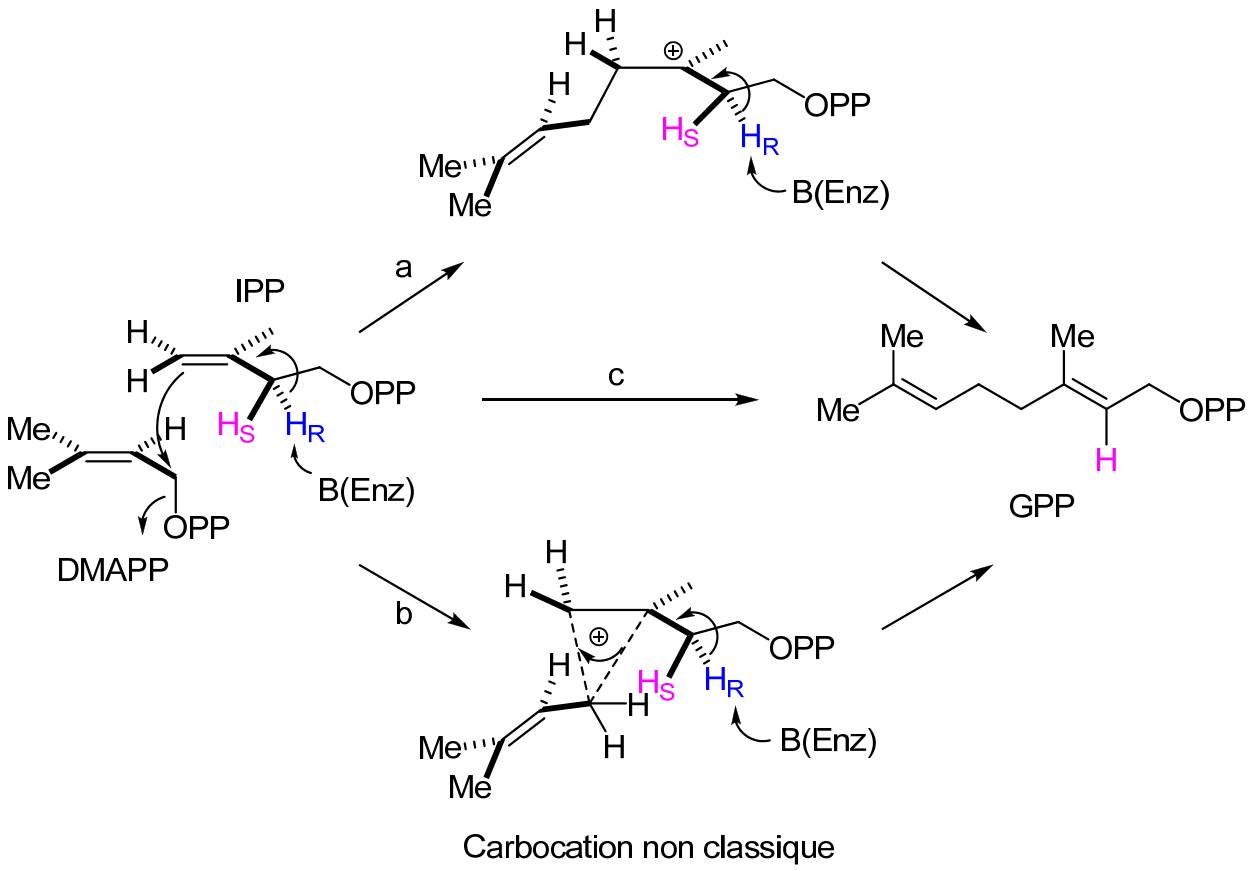


Schéma II.3.2

II.3.2 Condensations queue-à-queue

Les schémas II.3.3 et II.3.4 ci-dessous montrent la formation du squalène à partir de deux molécules de FPP et du phytoène à partir de deux molécules de GGPP. Ce sont des condensations qui correspondent formellement à un couplage queue-à-queue (voir le mécanisme détaillé dans la section II.8.2). Le squalène est le précurseur des triterpènes (C_{30}) et des stéroïdes et le phytoène est le précurseur des tétraterpènes (C_{40}) et des caroténoïdes.

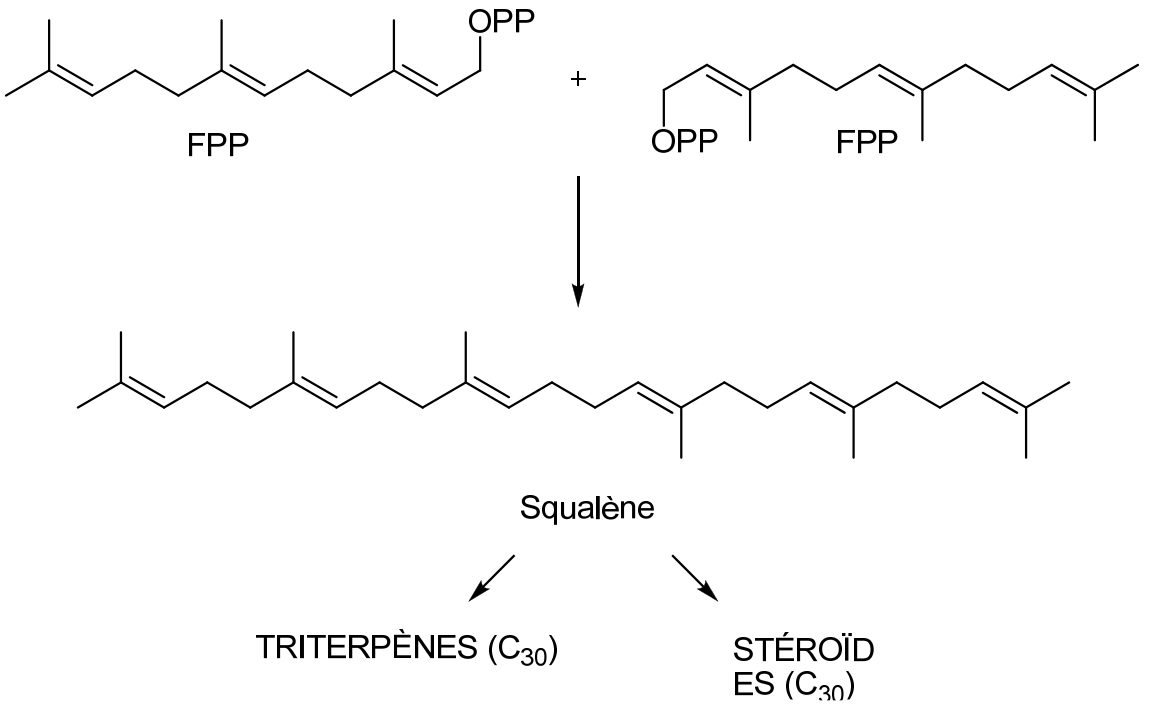


Schéma II.3.3

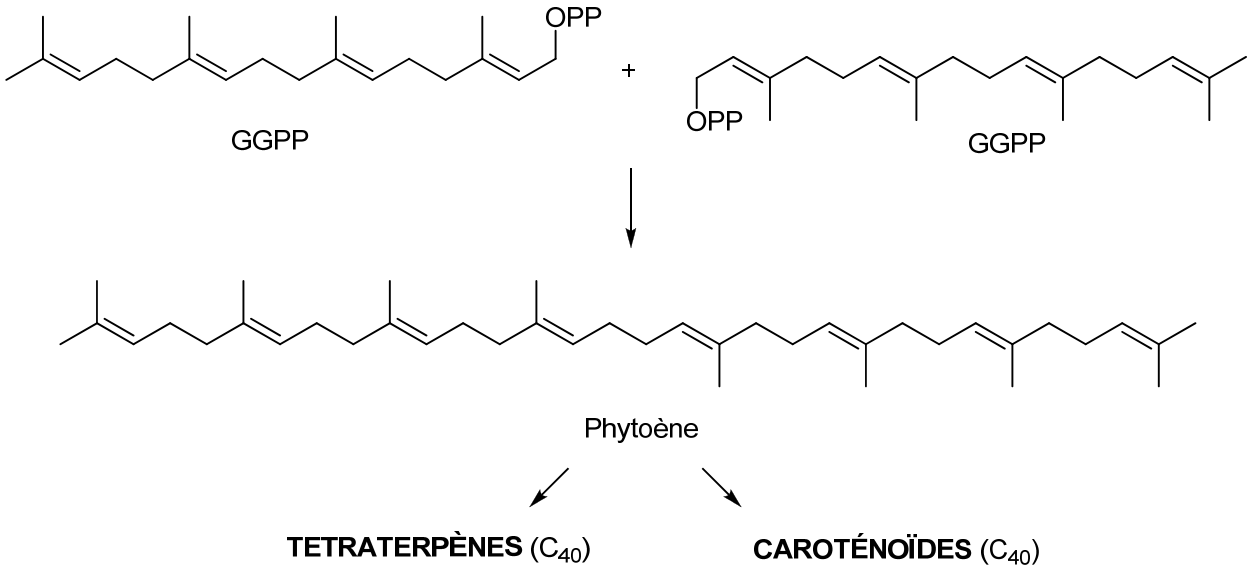


Schéma II.3.4

II.3.3 Autres condensations

Plusieurs monoterpènes sont fabriqués à partir d'une condensation de deux unités de DMAPP (Schéma II.3.5). Le carbocation non-classique mène aux unités cyclobutyle ou cyclopropyle, selon l'enzyme et l'organisme. Puis le cyclopropane est peut être ouvert de différentes façons pour donner lieu à deux squelettes monoterpéniques différents. Il existe des monoterpènes naturels possédant un des quatre squelettes colorés au schéma II.3.5.

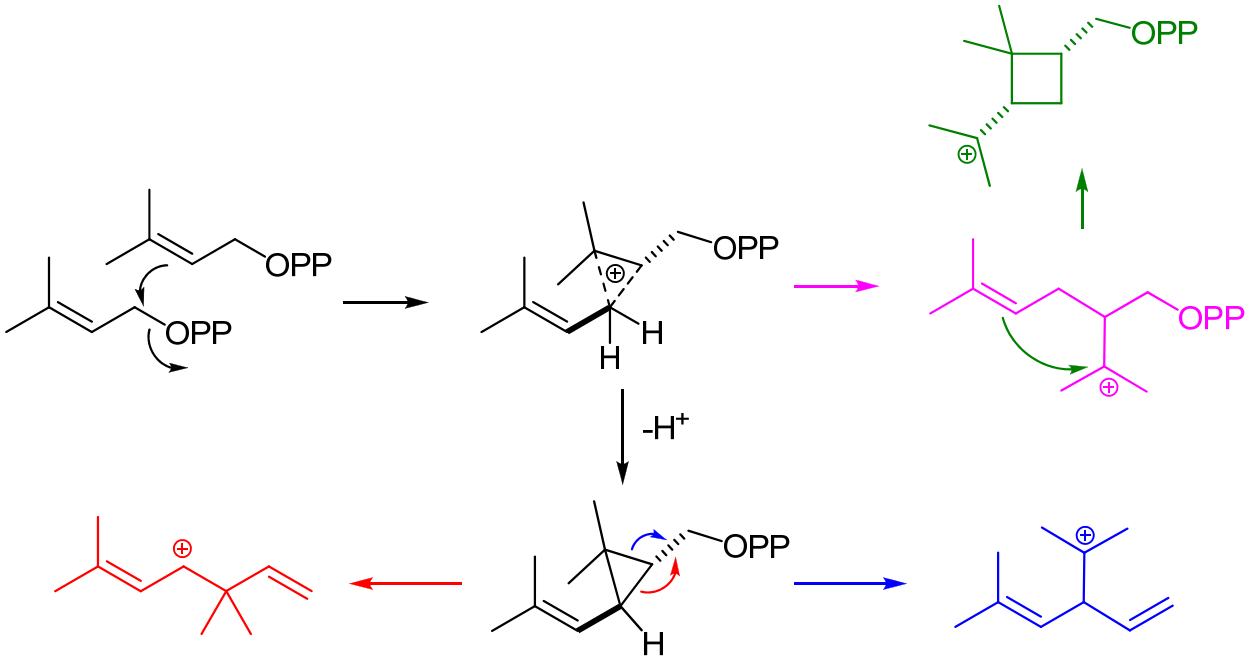


Schéma II.3.5

II.4 Biosynthèse des monoterpènes

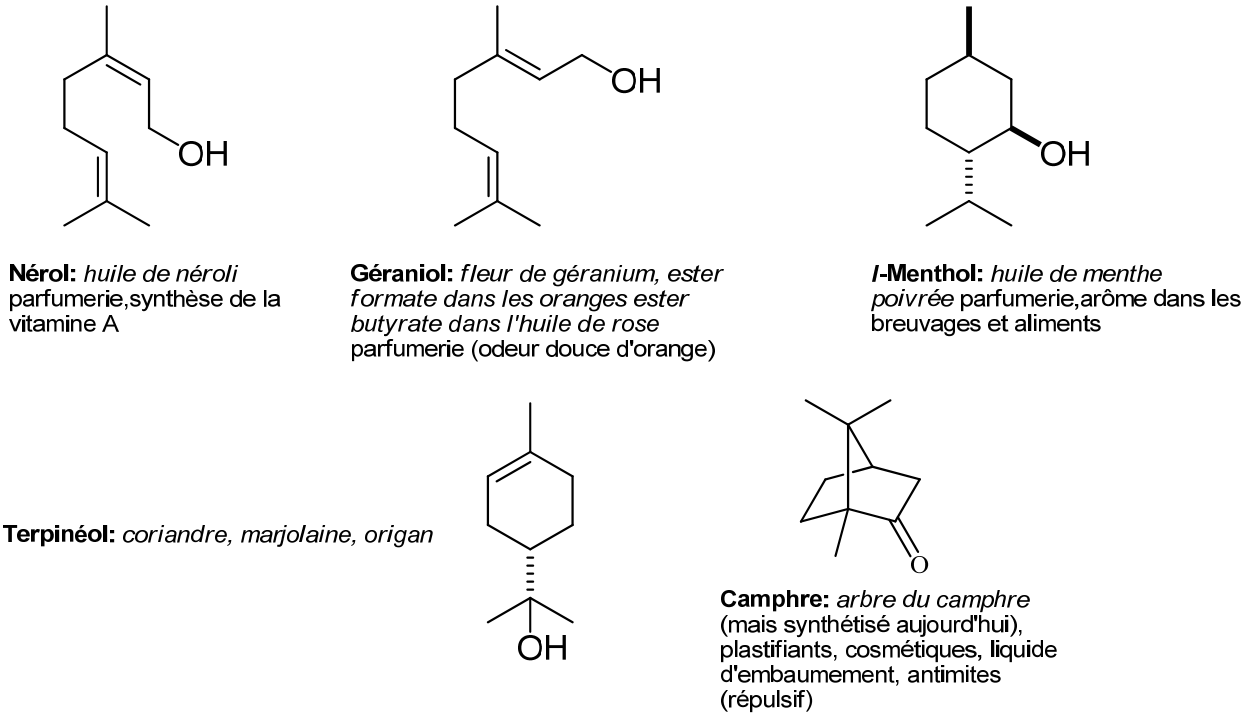
II.4.1 Généralités

Les monoterpènes sont constitués de deux unités isoprène (10 carbones = C_{10}). Ils proviennent principalement des arbres et des plantes plus évoluées (« higher plants »). Ce sont souvent des constituant « d'huiles essentielles » car assez volatils (leur arôme est donc facilement détectable. Ici, 'essentielle' signifie 'essence ou arôme'). Ils sont obtenus le plus souvent par distillation à la vapeur ou par extraction par un solvant organique non polaire.

Plusieurs sont utilisés en parfumerie ou comme arômes. Ils sont omniprésents dans les plantes évoluées mais se retrouvent aussi dans certains fongus (champignons) et moisissures, dans certaines plantes moins évoluées (en particulier dans les algues marines sous forme de terpènes halogénés le plus souvent), dans des insectes et dans quelques animaux (provenant surtout de leur diète).

Le rôle physiologique des monoterpènes n'est pas encore très bien connu et, actuellement, l'idée qu'ils peuvent être des agents chimiques de défense ou de communication dans les plantes et certains insectes (phéromones sexuelles, phéromones d'alerte, phéromones de pistage) est récente mais de plus en plus acceptée.

Les monoterpènes ont été utilisés depuis des siècles comme épices et herbes médicinales. Ils ont été parmi les premiers composés chimiques à être obtenus à l'état pur et caractérisés. La térébenthine (« turpentine » en anglais et en allemand, ce qui est à l'origine du mot « terpène ») est le distillat de conifères contenant plusieurs monoterpènes (hydrocarbures) de formule $C_{10}H_{16}$. Ils sont utilisés dans l'industrie de la peinture, des explosifs, des agents de protection du bois, etc. Les figures II.4.1 et II.4.2 montre quelques exemples de **monoterpènes**, leur *provenance* et leur utilisation.



Nérol: *huile de néroli* parfumerie,synthèse de la vitamine A

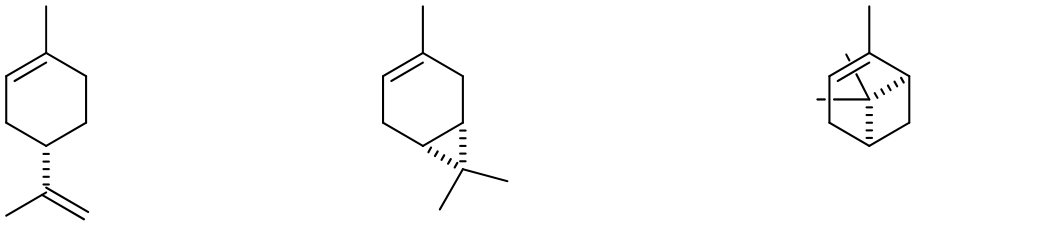
Géraniol: *fleur de géranium, ester formate dans les oranges ester butyrate dans l'huile de rose* parfumerie (odeur douce d'orange)

***l*-Menthol:** *huile de menthe poivrée* parfumerie,arôme dans les breuvages et aliments

Terpinéol: *coriandre, marjolaine, origan*

Camphre: *arbre du camphre* (mais synthétisé aujourd'hui), plastifiants, cosmétiques, liquide d'embaumement, antimites (répulsif)

Figure II.4.1



Limonène: *fruits des agrumes, aneth, cumin, racine de bergamote,* arôme à saveur de lime, parfurmerie

3-Carène: *sève de pin* térébenthine et autres diluants

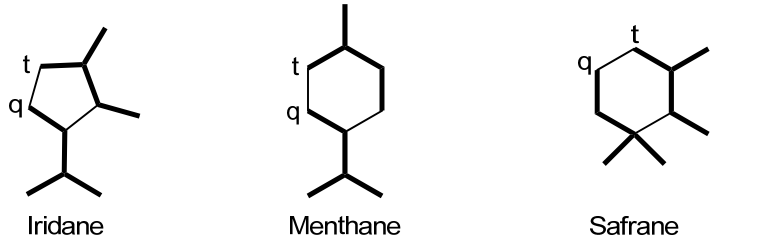
α-Pinène: *huile de cèdre et de pin* parfumerie, insecticides, fabrication d'huile de pin, synthèse du camphre

Figure II.4.2

Les monoterpènes sont classés en quatre catégories principales selon leur structure : les monoterpènes réguliers (acycliques, monocycliques et bicycliques) et les monoterpènes irréguliers. Dans les monoterpènes réguliers acycliques (figures II.4.3), monocycliques (figure II.4.4) et bicycliques (figure II.4.5), l'arrangement des deux unités isoprène est tête-à-queue (couplage du DMAPP avec l'IPP pour donner le GPP).

Diméthyloctane Néral

Figure II.4.3



Iridane Menthane Safrane

β-Skytanthine Menthone (odeur de menthe) Safranal

Figure II.4.4

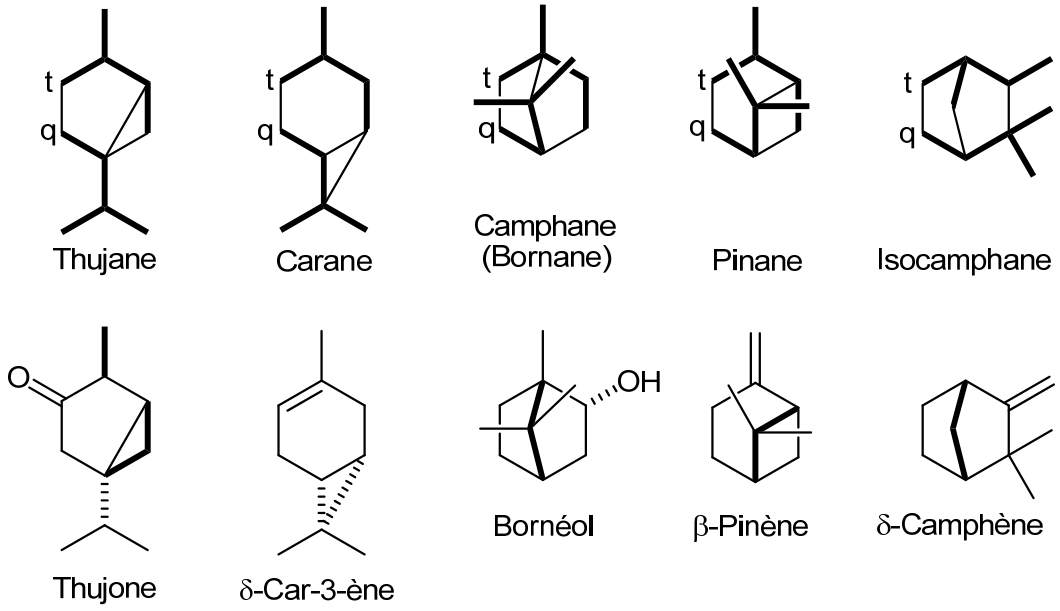


Figure II.4.5

Cet arrangement ne se retrouve pas dans les monoterpènes irréguliers (figure II.4.6) soit parce que leur biosynthèse fait intervenir des réarrangements suite à un couplage tête-à-queue des unités à cinq carbones (voir le fenchol et la nézukone) ou parce que ces unités se combinent d'une autre façon (voir les autres monoterpènes de la figure II.4.6). Dans chaque catégorie, les noms de familles de composés proviennent généralement du nom du premier membre de la famille qui a été isolé et étudié. Par exemple, la famille du menthane (monoterpènes monocycliques) tire son nom du menthol (odeur de menthe). En font partie le terpinéol, le limonène et plusieurs autres monoterpènes.

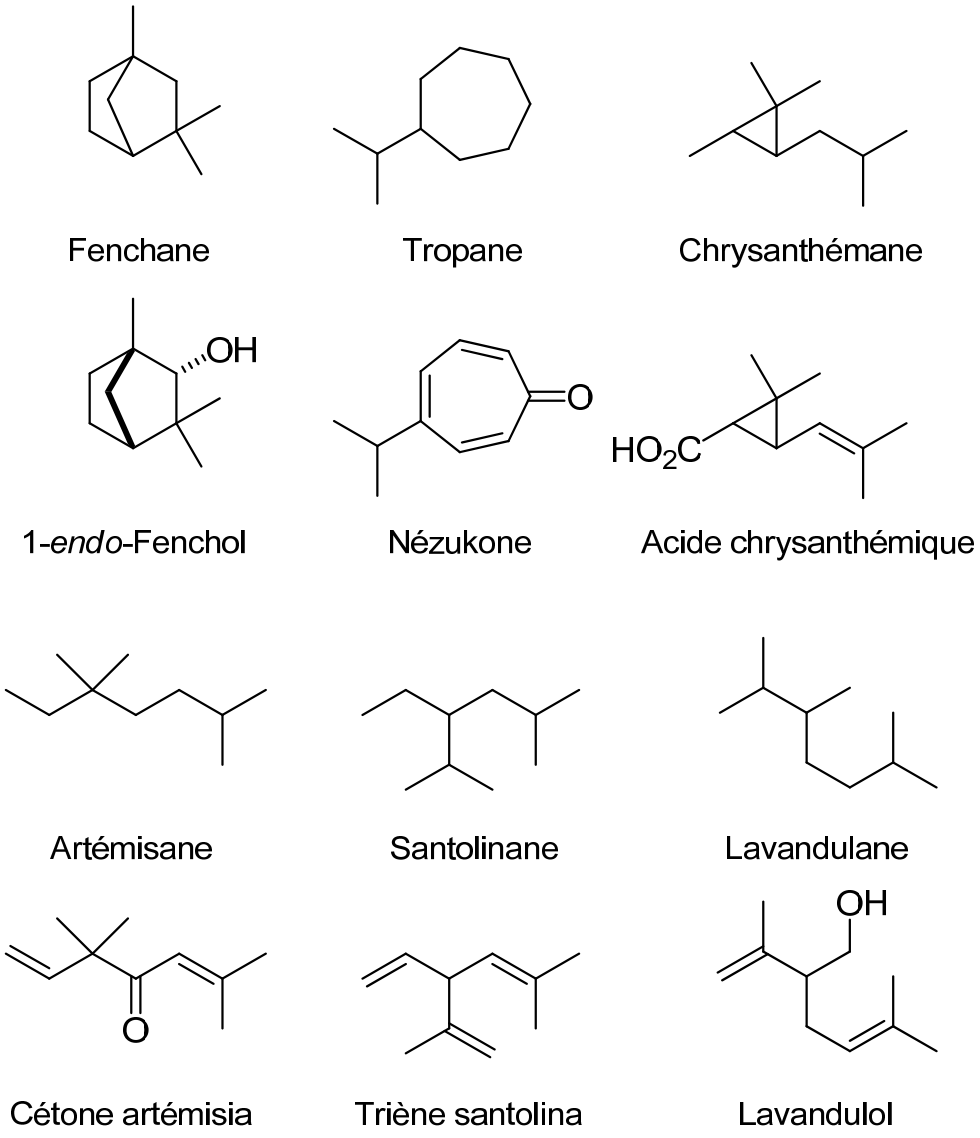


Figure II.4.6

II.4.2 Biosynthèse

Les réactions qu'emploi la nature pour transformer les intermédiaires de biosynthèse sont les mêmes que celles que les chimistes emploient à la différence près qu'elle sont fait dans l'eau à 37 ºC ! Les enzymes sont très bien équipés pour surmonter les barrières enthalpiques des réactions. Ils compensent par apport entropique, c'est-à-dire en rigidifiant les structures (rotation) et en maintenant les réactifs à proximité l'un de l'autres (translation). De plus, ils utilisent la technique du 'push-pull', c'est-à-dire qu'ils activent un groupement par chélation acide et en même temps assiste à la réaction par une chélation basique. Par exemple, l'élimination du

groupement pyrophosphate de l'IPP en myrcène, qui figure au schéma II.4.1 et reprise à la figure II.4.7, se fait par l'entremise d'une protonation (souvent par un acide carboxylique) du groupement phosphate et par déprotonation par un groupement basique (souvent un groupement histidine). L'enlignement des orbitales du substrats est assuré par l'ensemble des interactions entre l'enzyme et le substrats (ponts-H, interaction hydrophobe/hydrophile, interactions dipolaire ou de van der Waal, etc.). Nous ne répèterons pas cet aspect des réactions d'ici la fin du cours, mais vous pouvez supposez qu'il en est toujours ainsi.

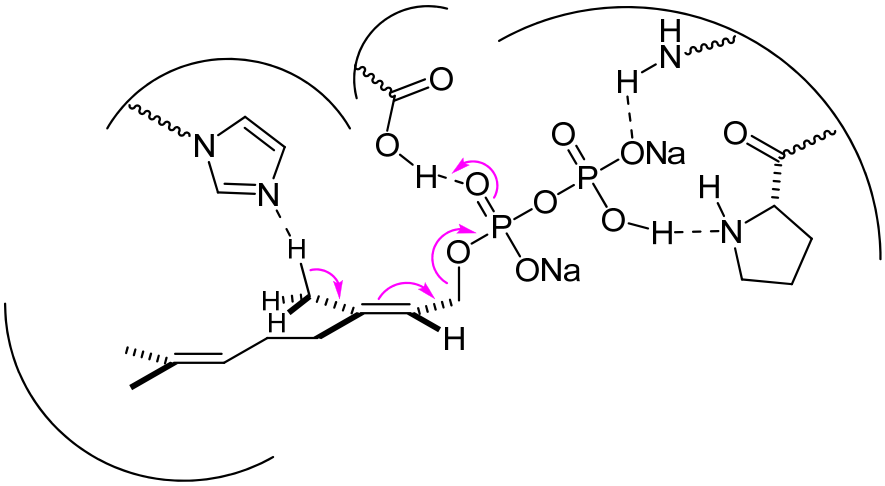


Figure II.4.7

Comme indiqué aux schémas II.4.1 et II.4.2, ils sont dérivés du pyrophosphate de géranyle (GPP) ou de son isomère géométrique, le pyrophosphate de néryle (NPP), ou encore de l'isomère de position, le pyrophosphate de linalyle (LPP). Des enzymes (isomérases) effectuent ces isomérisations. Les différents squelettes dérivés de ces précurseurs et le type de réactions qui interviennent dans les conversions enzymatiques sont indiqués dans ces mêmes schémas.

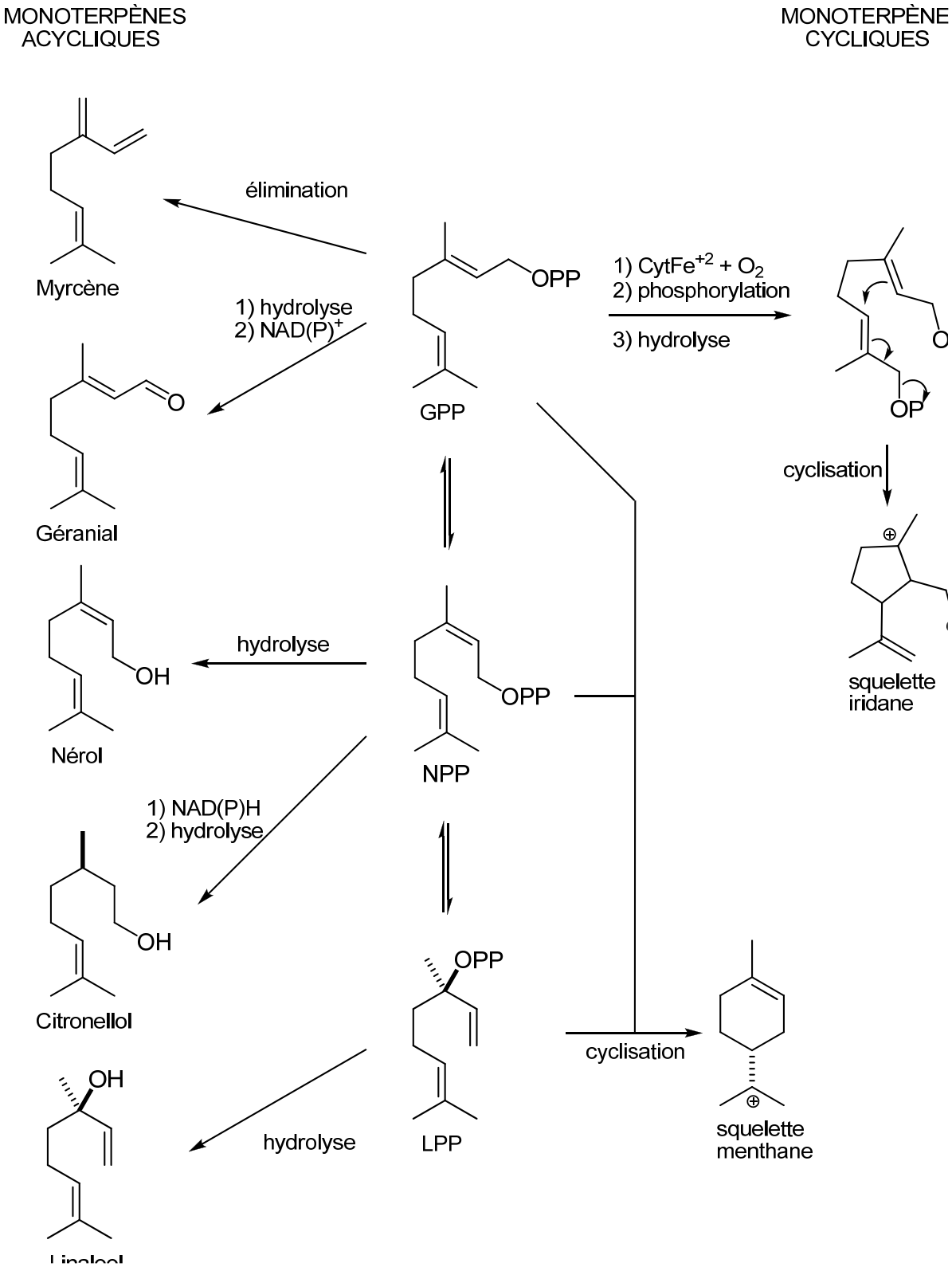

MONOTERPÈNES
ACYCLIQUES
MONOTERPÈNE
CYCLIQUES
élimination
Myrcène
1) hydrolyse
2) NAD(P)+
Géranial
GPP
OPP
1) CytFe+2 + O2
2) phosphorylation
3) hydrolyse
cyclisation
squelette
iridane
hydrolyse
Nérol
NPP
1) NAD(P)H
2) hydrolyse
Citronellol
OPP
cyclisation
squelette
menthane
hydrolyse
LPP
Linalool

Schéma II.4.1

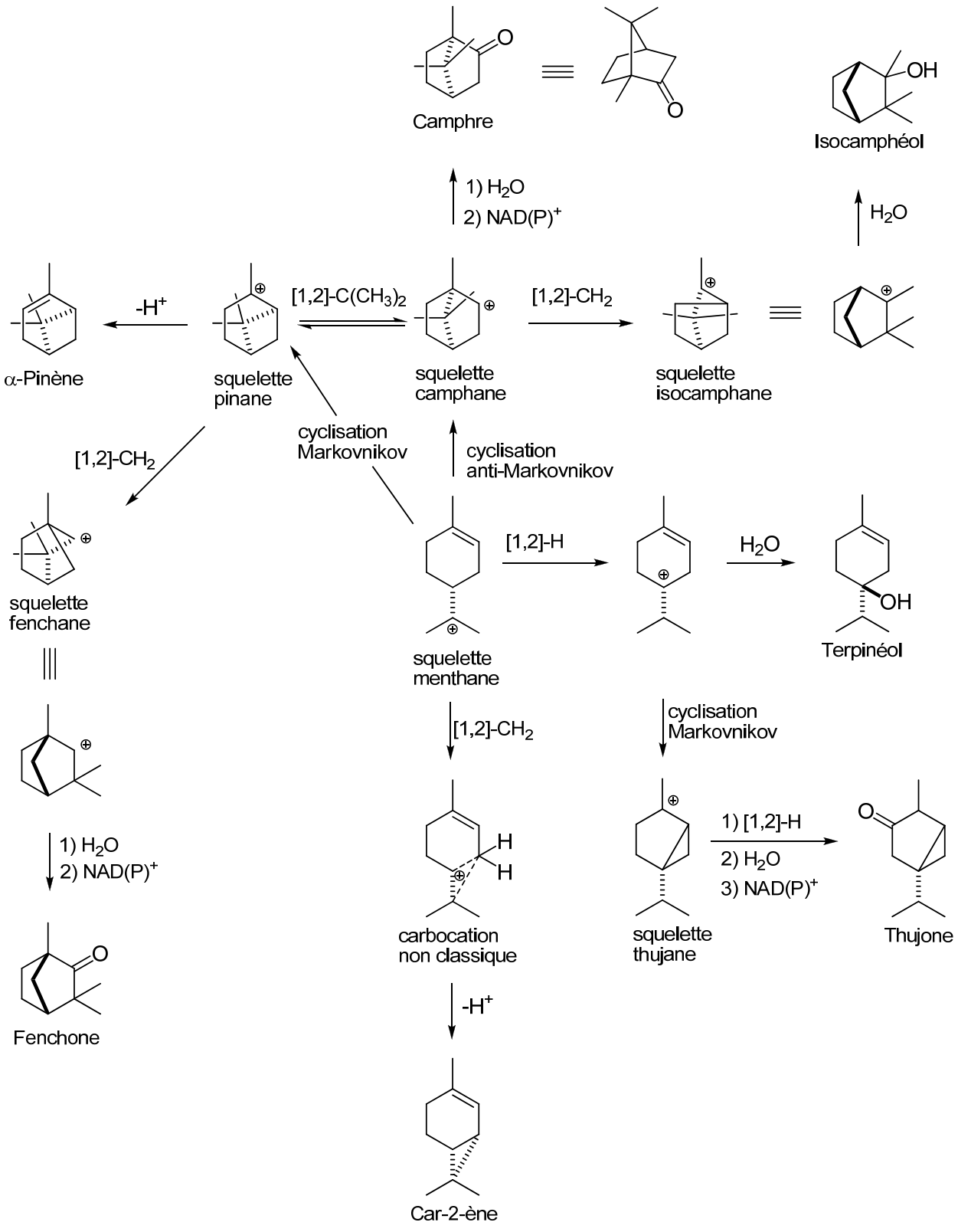

Camphre
Isocamphéol
1) H2O
2) NAD(P)+
H2O
α-Pinène
-H+
squelette pinane
[1,2]-C(CH3)2
squelette camphane
[1,2]-CH2
squelette isocamphane
[1,2]-CH2
cyclisation Markovnikov
cyclisation anti-Markovnikov
squelette fenchane
squelette menthane
[1,2]-H
H2O
Terpinéol
[1,2]-CH2
cyclisation Markovnikov
1) H2O
2) NAD(P)+
carbocation non classique
squelette thujane
1) [1,2]-H
2) H2O
3) NAD(P)+
Thujone
Fenchone
-H+
Car-2-ène

Schéma II.4.2

Il n'est pas rare de retrouver l'un ou l'autre des énantiomère des monoterpènes selon l'organisme ou l'espèce en question ou même pour une même espèces de provenance différente. L'enzyme de cyclisation contrôle le pliage du GPP, NPP ou LPP. Par exemple, le (+)-camphre se retrouve dans la sauge (*salvia officinalis*) et le (-)-camphre dans la tanaisie commune (*Tanacetum Vulgare*). Il en est de même pour la (+)-carvone que l'on retrouve dans le carvi (*Carum carvi*) et la (-)-carvone qui est produite dans la menthe verte (*Mentha spicata*).

Parfois les deux énantiomères se retrouve dans le même organisme. Le schéma II.4.3 montre comment la menthe poivrée prend le LPP et le pli pour donner le (+)-*R*-limonène et aussi le (-)-*S*-limonène. Le pin commun, lui, produit et (-)- et le (+)-α-pinène.

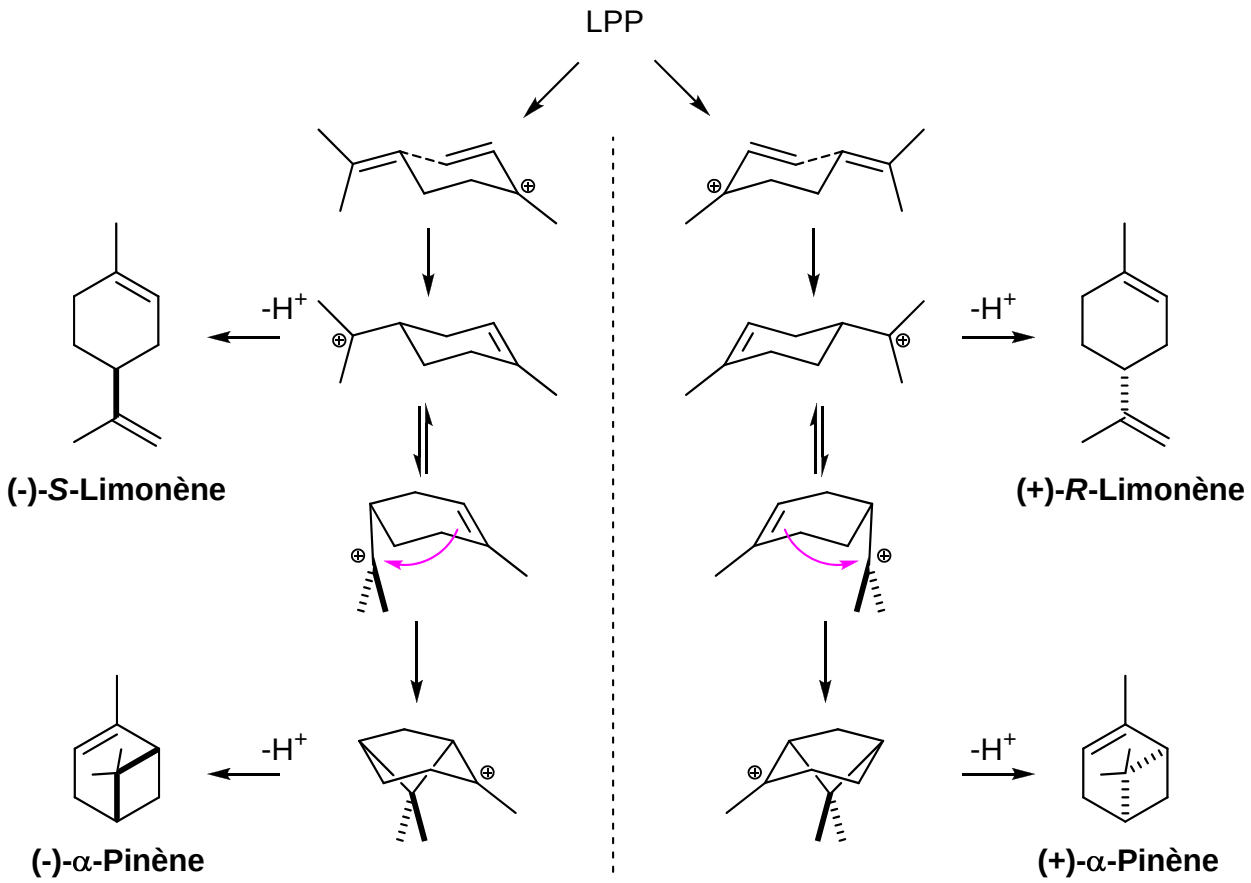


Schéma II.4.3

Les réactions sont des hydrolyses, des oxydations, des réductions, des phosphorylations et des réarrangements sigmatropiques [1,2] de groupement alkyle et d'hydrure (migration suprafaciale avec rétention de configuration à l'atome qui migre tel que permis par la symétrie des orbitales). Pour la formation de monoterpènes monocycliques et bicycliques, les cyclisations (addition intramoléculaire d'un carbocation sur une double liaison) peuvent avoir une orientation

Markovnikov ou anti-Markovnikov. L'orientation de la cyclisation est déterminée par le site actif de l'enzyme (cyclase). Ceci montre la sélectivité des réactions enzymatiques. Une hypothèse pour la conversion du cation menthyle (squelette menthane) en carène (schéma II.4.2) est la formation intermédiaire d'un carbocation non classique par migration [1,2] d'un CH_2 suivie par une déprotonation (réaction analogue à la formation du nortricyclène par déprotonation du carbocation non classique norbornyle (voir annexe VII.4 et chimie organique IV).

L'étudiant(e) doit pouvoir proposer une biosynthèse sur la base des schémas présentés dans les notes de cours et des remarques et discussions faites au cours. Dans le cas d'un monoterpène donné, l'unité précurseur à 10 carbones est le GPP et sa conversion dans les différents squelettes de monoterpènes est décrite dans les schémas II.4.4 et II.4.5. En proposant une biosynthèse, il faut tenir compte de toute l'information donnée et disponible. Par exemple, la biosynthèse du *l*-menthol à partir du GPP, si aucune donnée (aucune contrainte) autre que celles du schéma II.4.1 et II.4.2 n'est fournie, pourrait faire intervenir les étapes décrites dans schéma II.4.4.

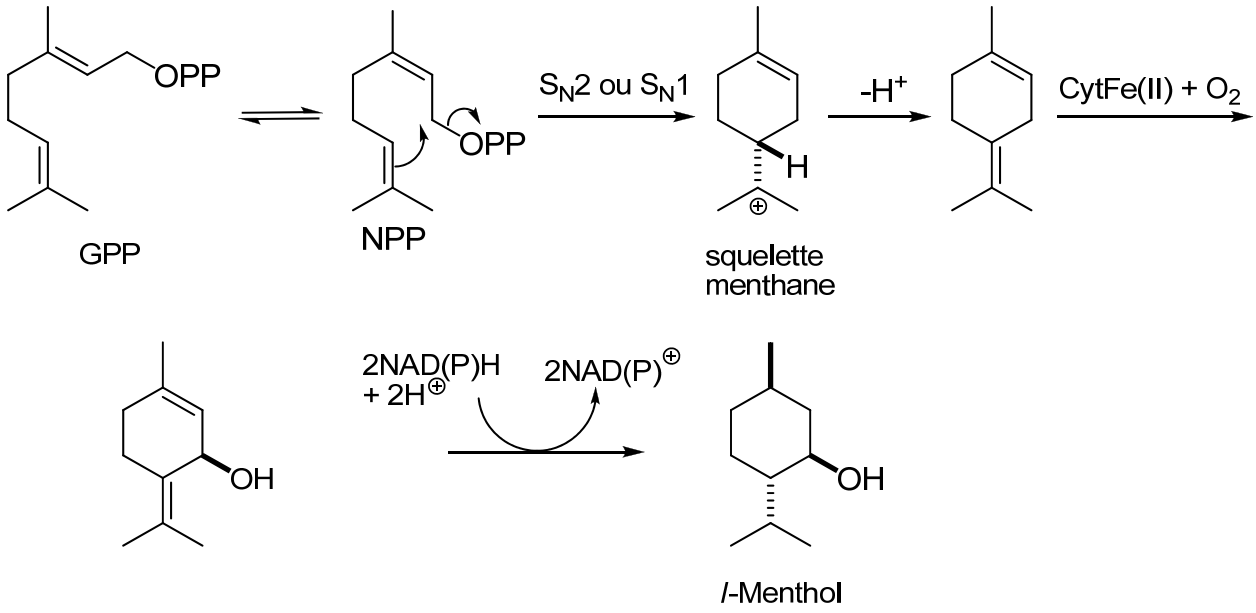


Schéma II.4.4

Mais en réalité, cette biosynthèse n'est pas celle que les plantes font puisqu'il a été démontré que la *l*-menthone est d'abord formée puis convertie en *l*-menthol par la menthe poivrée. La voie du schéma II.4.5 peut alors être proposée.

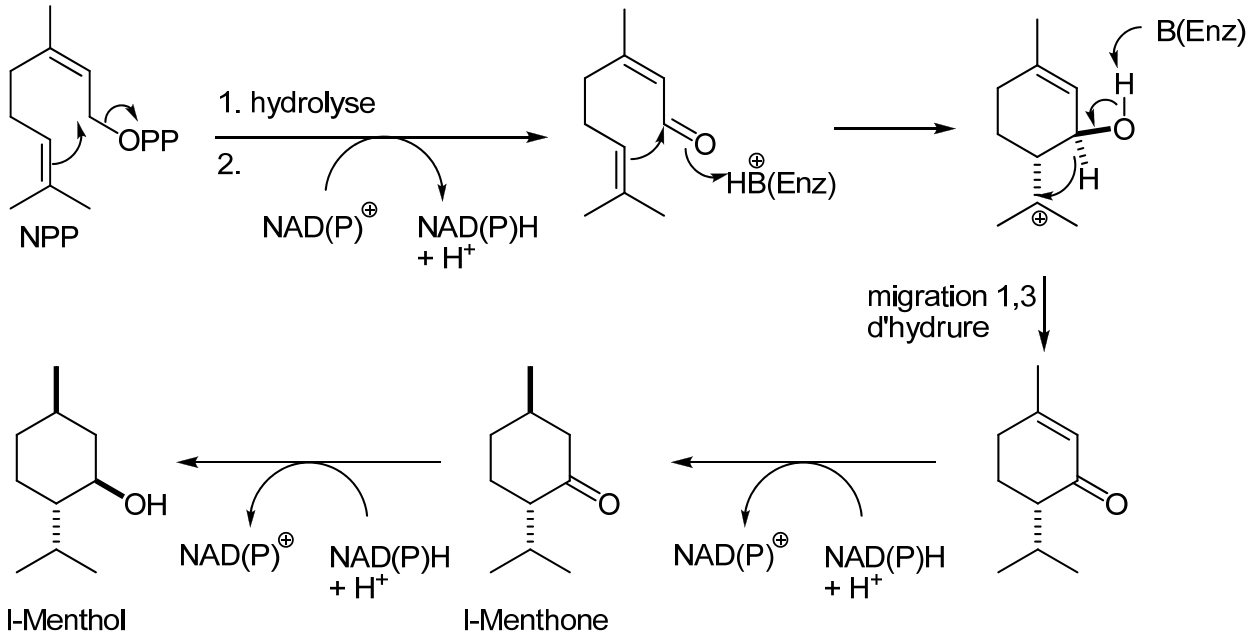


Schéma II.4.5

Cependant des études d'incorporation du GPP marqué au deutérium à la position 1 ont montré que le *l*-menthol ne contenait qu'un deutérium à la position 2. Comme la voie du schéma II.4.5 aurait donné le *l*-menthol avec le deutérium sur le carbone tertiaire de la chaîne isopropyle, cette voie biosynthétique n'est pas en accord avec les expériences de marquage isotopique. La voie du schéma II.4.6, par contre, est en accord avec les expériences de marquage isotopique.

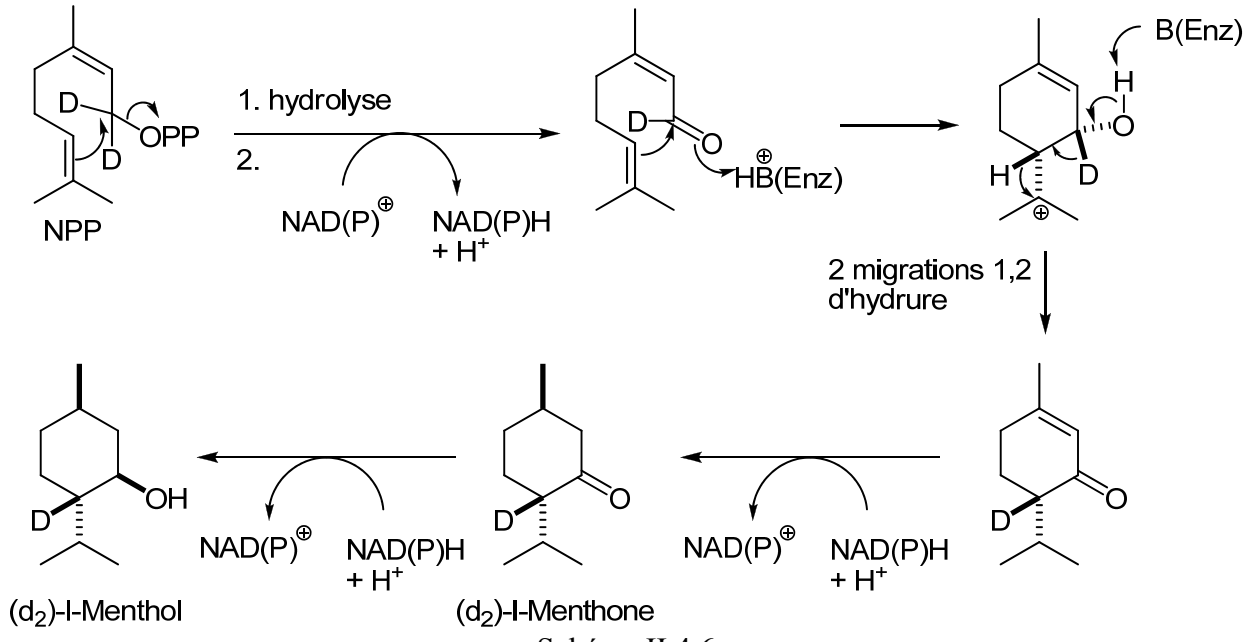


Schéma II.4.6

II.5 Biosynthèse des sesquiterpènes

II.5.1 Généralités

On retrouve les sesquiterpènes dans les plantes aussi bien les plus évoluées que les moins évoluées. Ils se retrouvent également dans certaines moisissures et champignons micros-copiques mais plus rarement dans les bactéries. Quelques bactéries produisent des sesquiterpènes qui possèdent de puissantes activités biologiques. Quelques exemples de sesquiterpènes sont illustrés à la figure II.5.1. Leur rôle biologique n'est pas très bien compris. Certains, comme le longifolène, ont un rôle semblable à celui des monoterpènes. Les sesquiterpènes polyoxygénés de la figure II.5.2 produits par des bactéries sont toxiques et possèdent des propriétés antibiotiques. Leur étude au tout début a contribué à la formulation de la règle de l'isoprène. Ils présentent une très grande diversité de squelettes et leur synthèse a suscité beaucoup d'intérêt de la part des chimistes organiciens de synthèse. Les efforts de ces chimistes ont conduit au développement de nouvelles stratégies de synthèse et de nouvelles réactions.

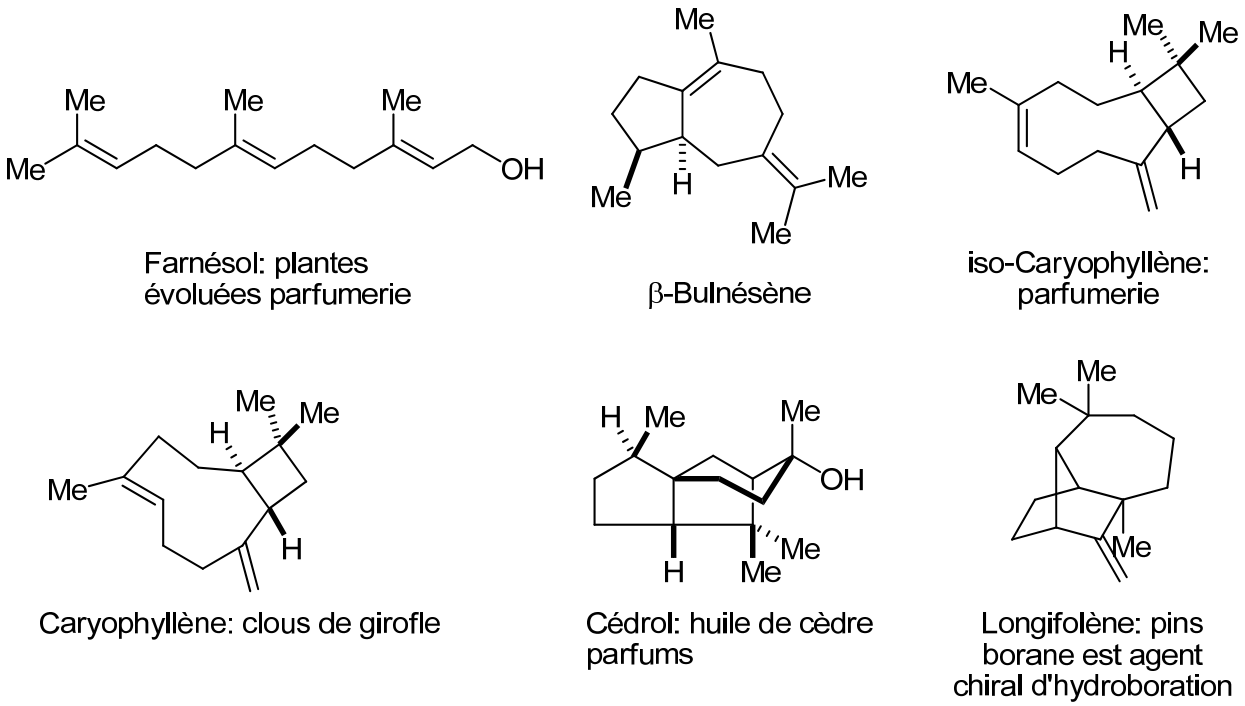
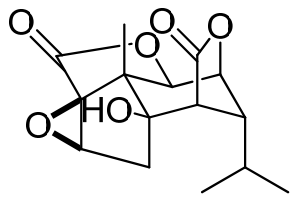
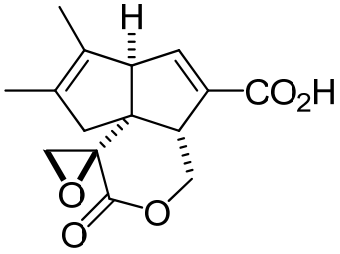


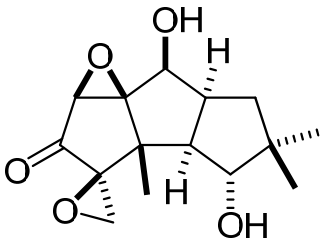
Figure II.5.1



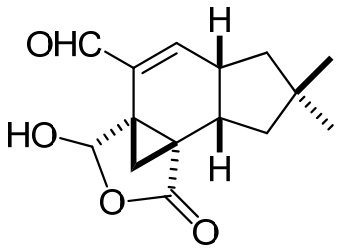
Picrotoxinine: *picrotoxine*



Pentalénolactone: *streptomyces* antibiotique



Corioline: *coriolus consor* antibiotique



Acide marasmique: *marasmius conigenus* antibiotique

Figure II.5.2

II.5.2 Biosynthèse de squelettes monocycliques

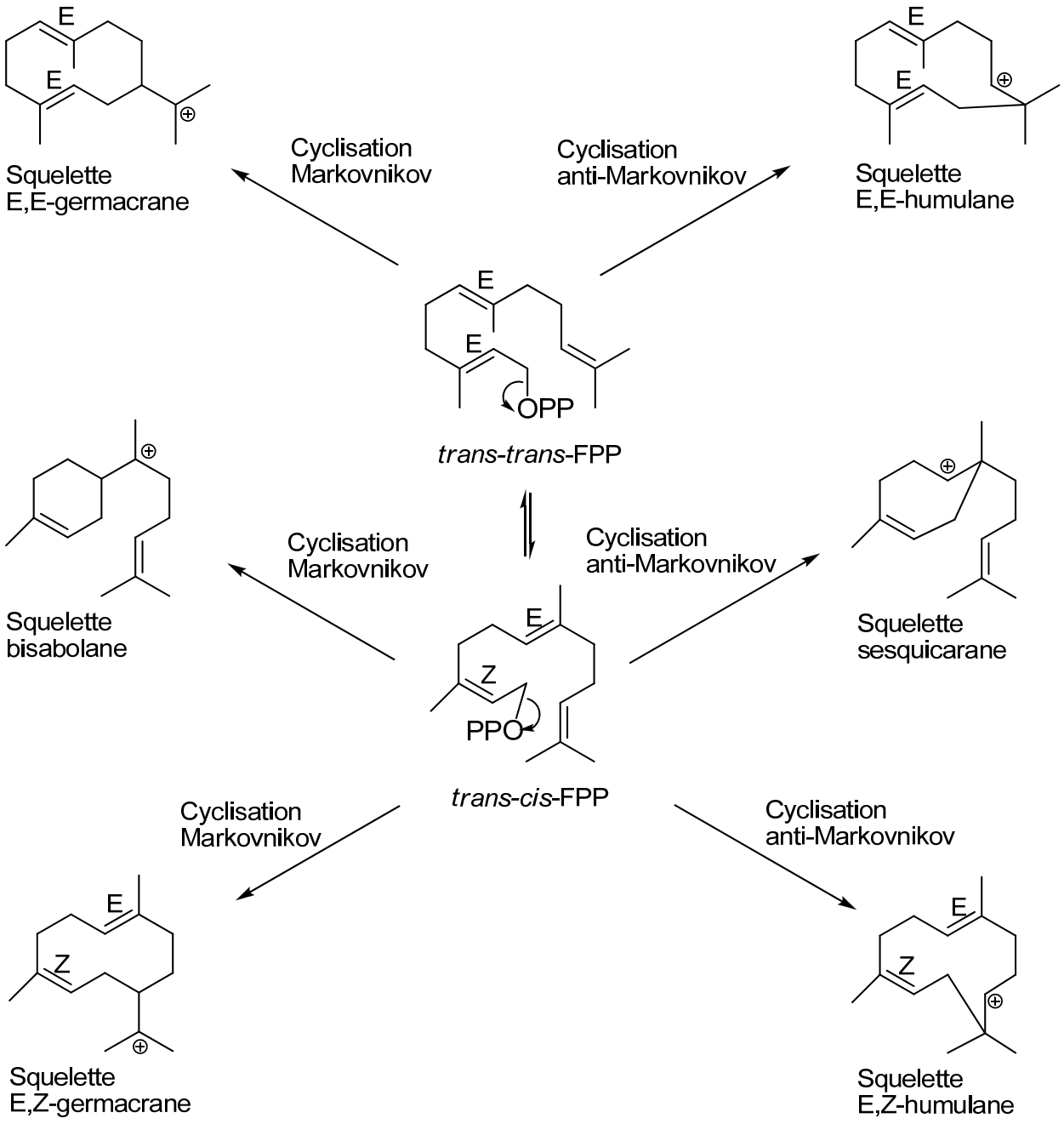


Schéma II.5.1

Les études sur la biosynthèse des sesquiterpènes n'ont pas été très poussées. Ils sont très probablement dérivés du *trans-trans* ou du *trans-cis* pyrophosphate de farnésyle (FPP) tel

qu'illustré au schéma II.5.1. La biosynthèse du FPP à partir du DMAPP et de l'IPP a été décrite au schéma II.3.1. Le schéma II.5.1 montre les différents modes de cyclisation du *trans-trans* FPP et du *trans-cis* FPP en cations des squelettes monocycliques *E*,*E*-germacrane, *E*,*Z*-germacrane, *E*,*E*-humulane, *E*,*Z*-humulane, bisabolane et sesquicarane. Les sesquiterpènes monocycliques sont obtenus à partir de ces carbocations. Les cyclisations de type S_N1 peuvent être d'orientation Markovnikov ou anti-Markovnikov.

II.5.3 Biosynthèse de squelettes bi- et tricycliques

Les schémas II.5.2, II.5.3 et II.5.4 montrent la grande diversité de squelettes de sesquiterpènes bicycliques et tricycliques obtenue par les transformations des cations monocycliques du schéma II.5.1. La biosynthèse des sesquiterpènes dérivés du cation *E*,*E*-germacrane est décrite au schéma II.5.2. Celle des sesquiterpènes dérivés du cation *E*,*Z*-germacrane est décrite au schéma II.5.3. Il y a des cyclisations de type Markovnikov et anti-Markovnikov, des réarrangements sigmatropiques [1,2] (migration d'hydrure, de CH_2 et de CH_3) et deux cas de transfert d'hydrure 1,3 que les systèmes enzymatiques effectuent assez fréquemment (les enzymes sont aussi capables de faire des transferts d'hydrure 1,4 et 1,5 comme nous le verrons plus loin).

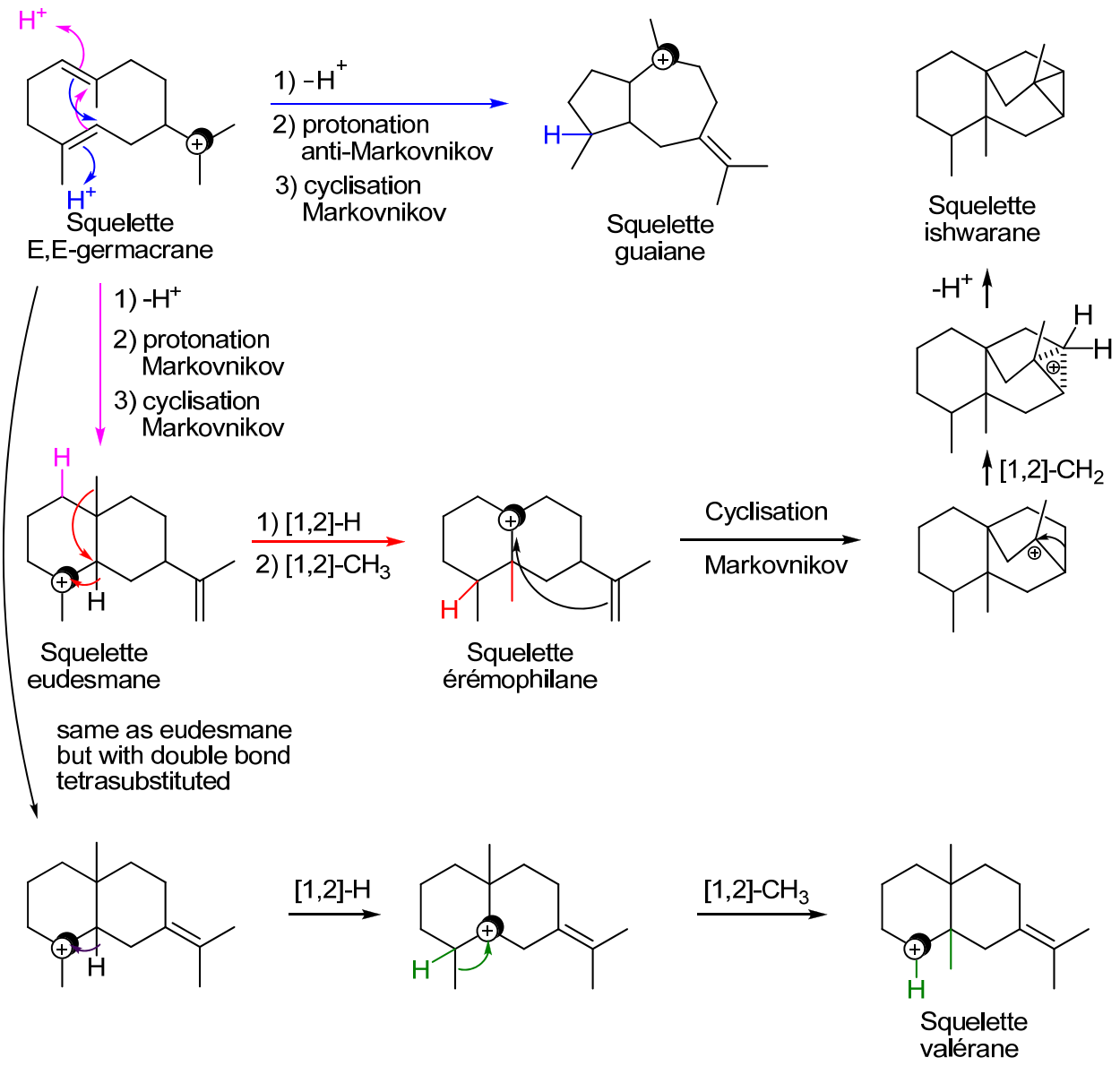


Schéma II.5.2

Probablement que la cyclisation du *trans-cis* FPP en cation *E*,*Z*-germacryle est effectuée par l'enzyme dans la conformation indiquée ci-dessous. La conformation du cation permet ensuite un transfert d'hydrure 1,3 (schéma II.5.3).

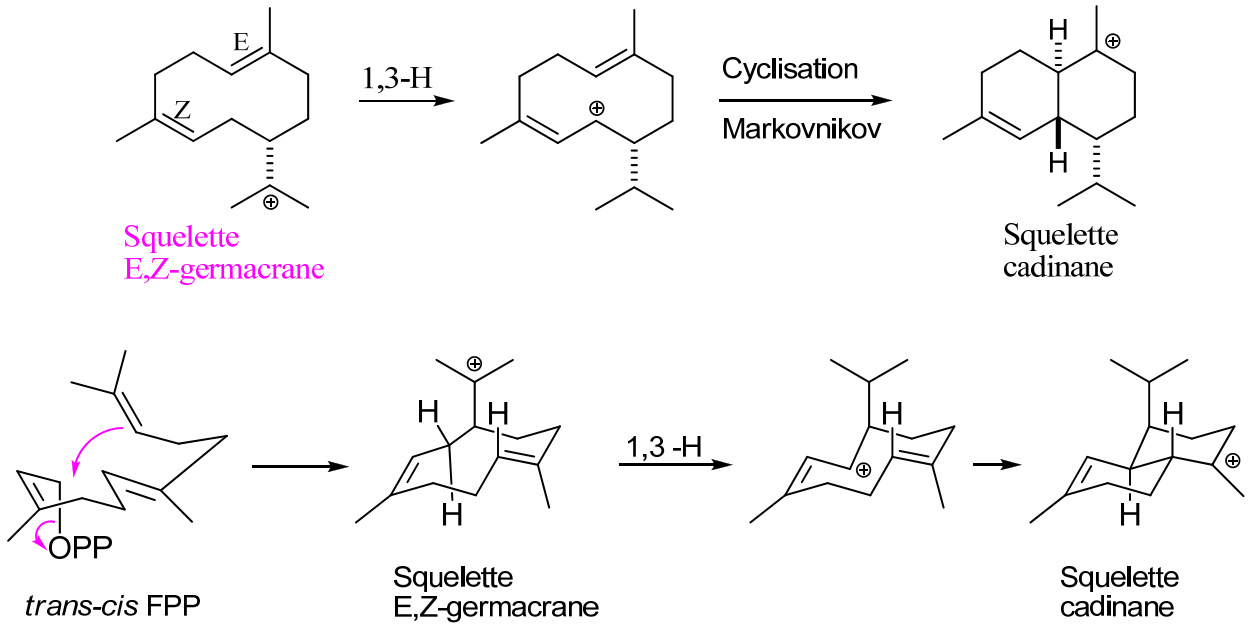


Schéma II.5.3

La biosynthèse des sesquiterpènes dérivés du cation bisabolane est décrite au schéma II.5.4 et celle des sesquiterpènes dérivés du cation *E,Z*-humulane est décrite au schéma II.5.5. Les squelettes des chamigranes, des acoranes et des cédranes sont accessibles par cette voie à partir du cation bisabolane, tandis que le squelettes des longifolanes est fabriqué à partir du cation humulane.

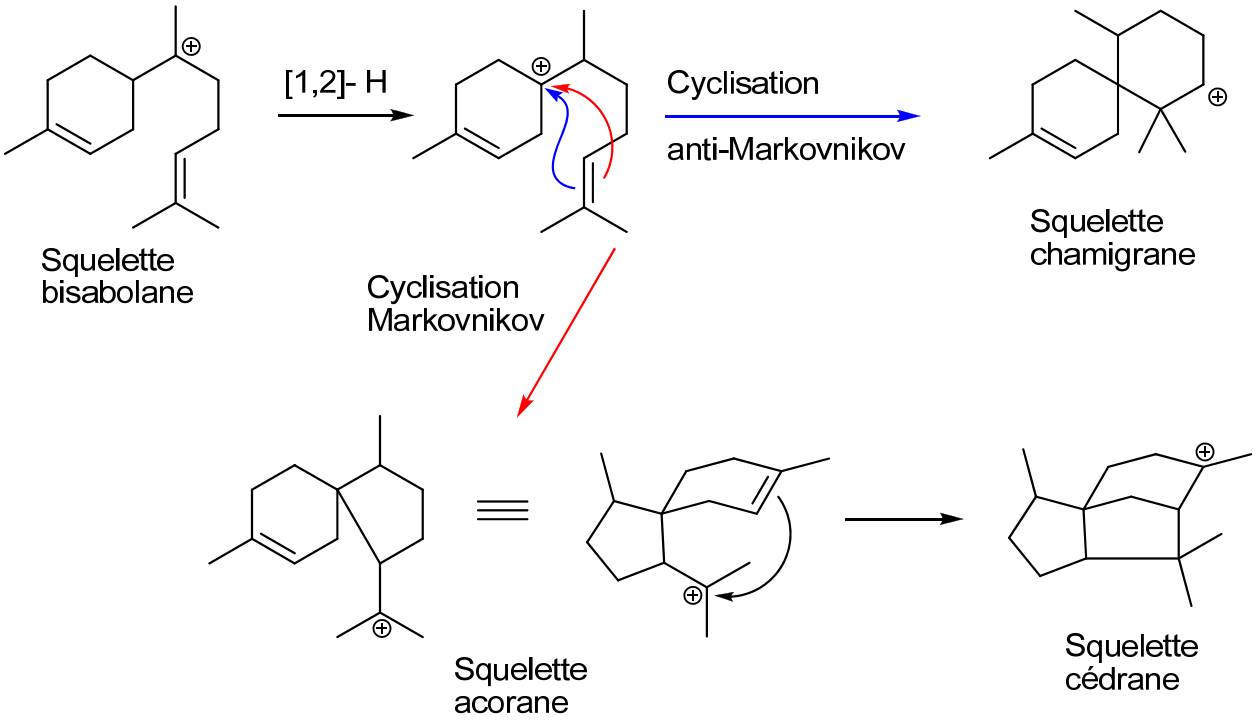


Schéma II.5.4

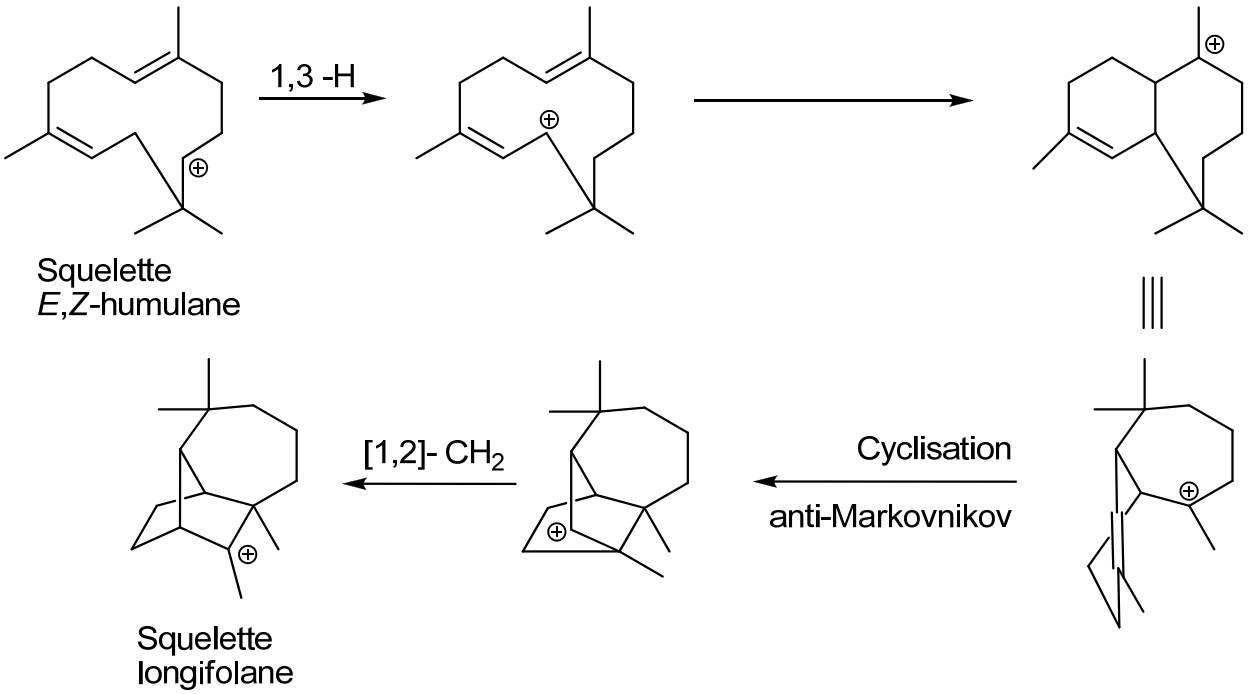


Schéma II.5.5

Finalement, le cation *E*,*E*-humulane mène aux sesquiterpènes de la famille des cariophyllanes et des hirsutanes (Schéma II.5.6). Les cariophyllanes sont nombreux et représente des défis synthétiques de tailles dû à leurs cycles à 4 et 8 membres fusionnés. Les terpènes de types hirsutane font parti d'une grande famille structurale qu'on appelle les triquinanes. Les hirsutanes font parti des triquinanes linéaires ou fusionnés. Il y a aussi des triquinanes angulaires que nous verront pendant les problèmes en classe. Pour former le noyau hirsutane, il faut plusieurs migration. D'abord la migation-1,2 d'un hydrogène. Puis, deux cyclisations, l'une Markovnikov, l'autre anti-Markovnikov. Enfin, deux autres migrations de carbones conduisent au squelette hirsutane. La figure II.5.3 montre le carbocation **A** en 3 dimensions pour une stéréochimie donnée. Les deux carbones impliqués dans la migration sont en effet à proximité l'un de l'autre.

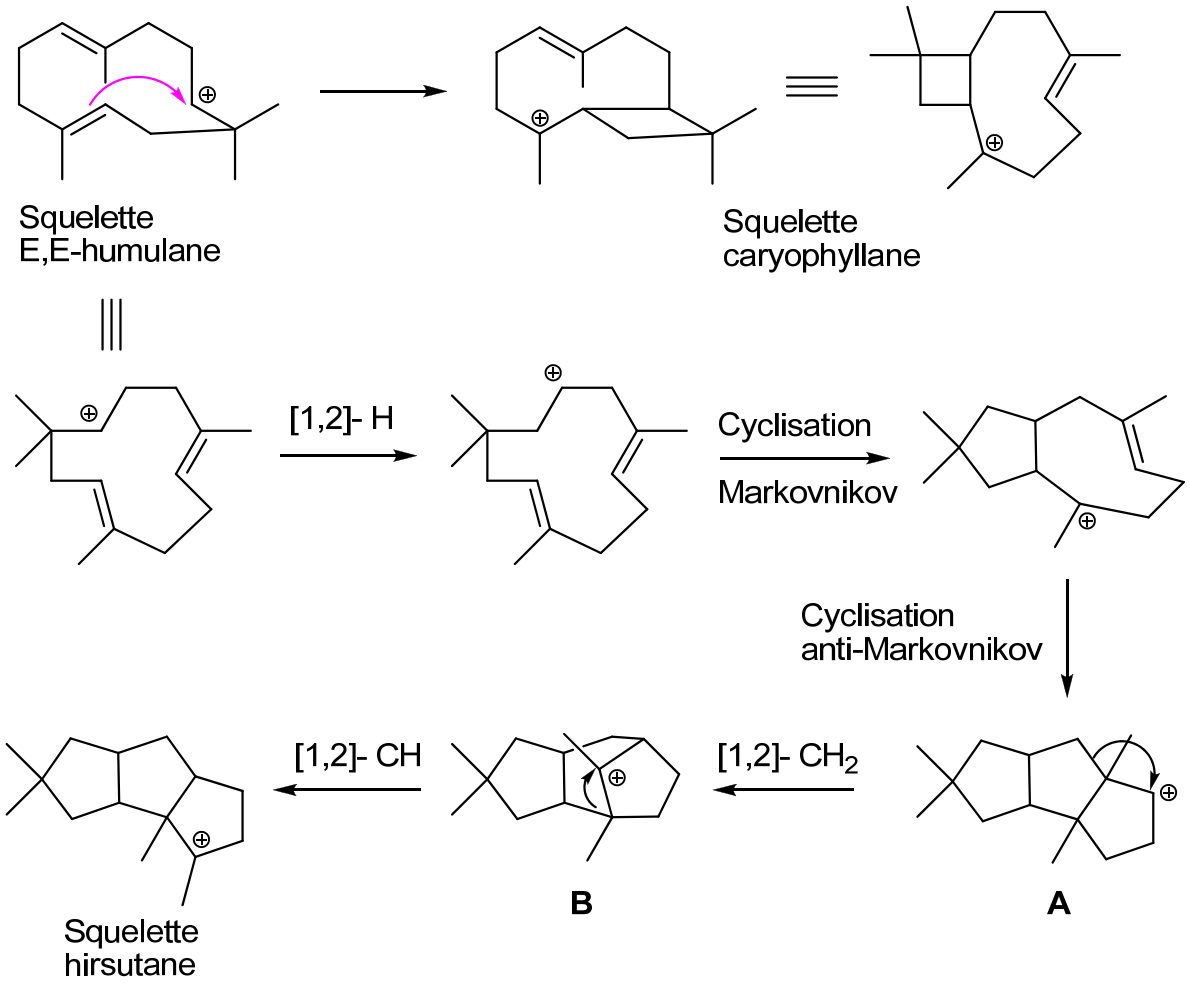


Schéma II.5.6

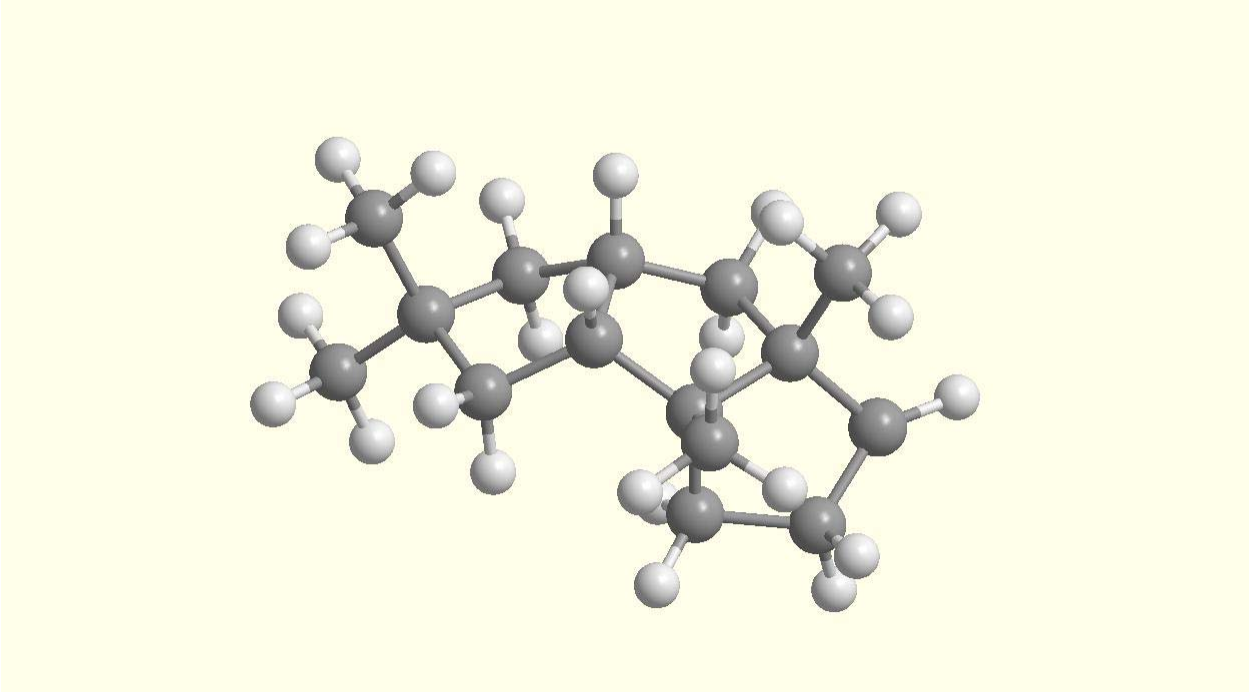


Figure II.5.3

II.5.4 Sesquiterpènes trans-décaliniques

Le *trans-trans* FPP peut être hydrolysé en farnésol qui est ensuite oxydé en acide. Ce dernier, relié probablement à un coenzyme, est cyclisé dans le site actif d'un enzyme et dans la conformation chaise-chaise indiquée ci-dessous par protonation Markovnikov de la double liaison C_{10}=C_{11} pour former un squelette *trans*-décalinique (schéma II.5.7).

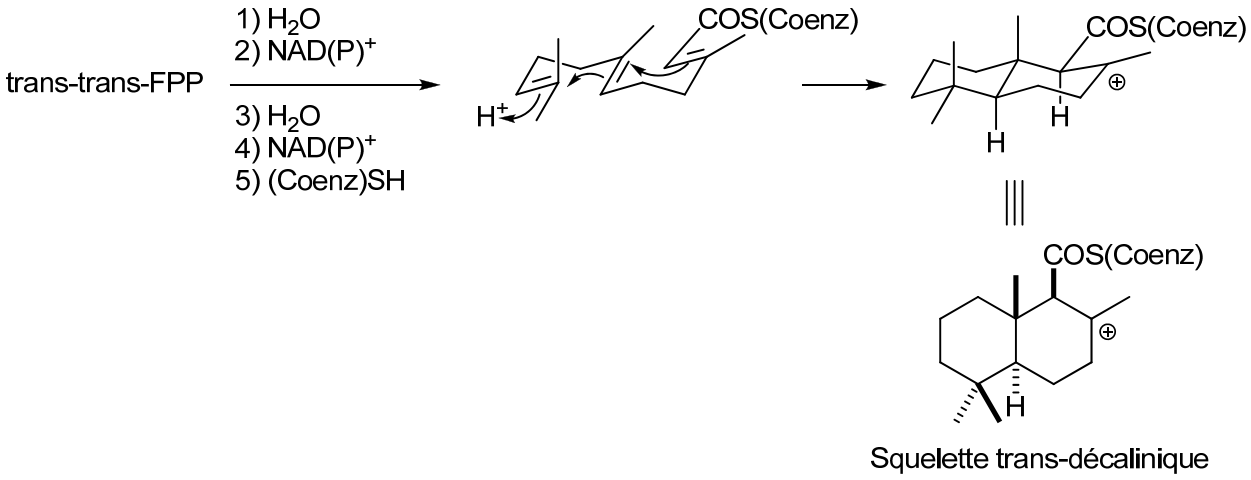
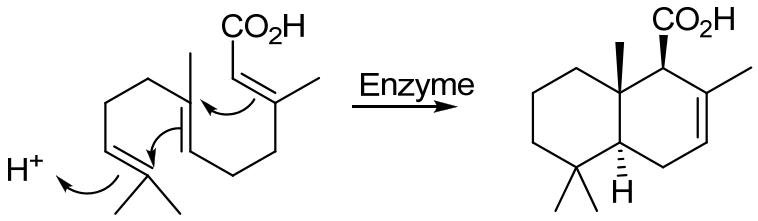


Schéma II.5.7

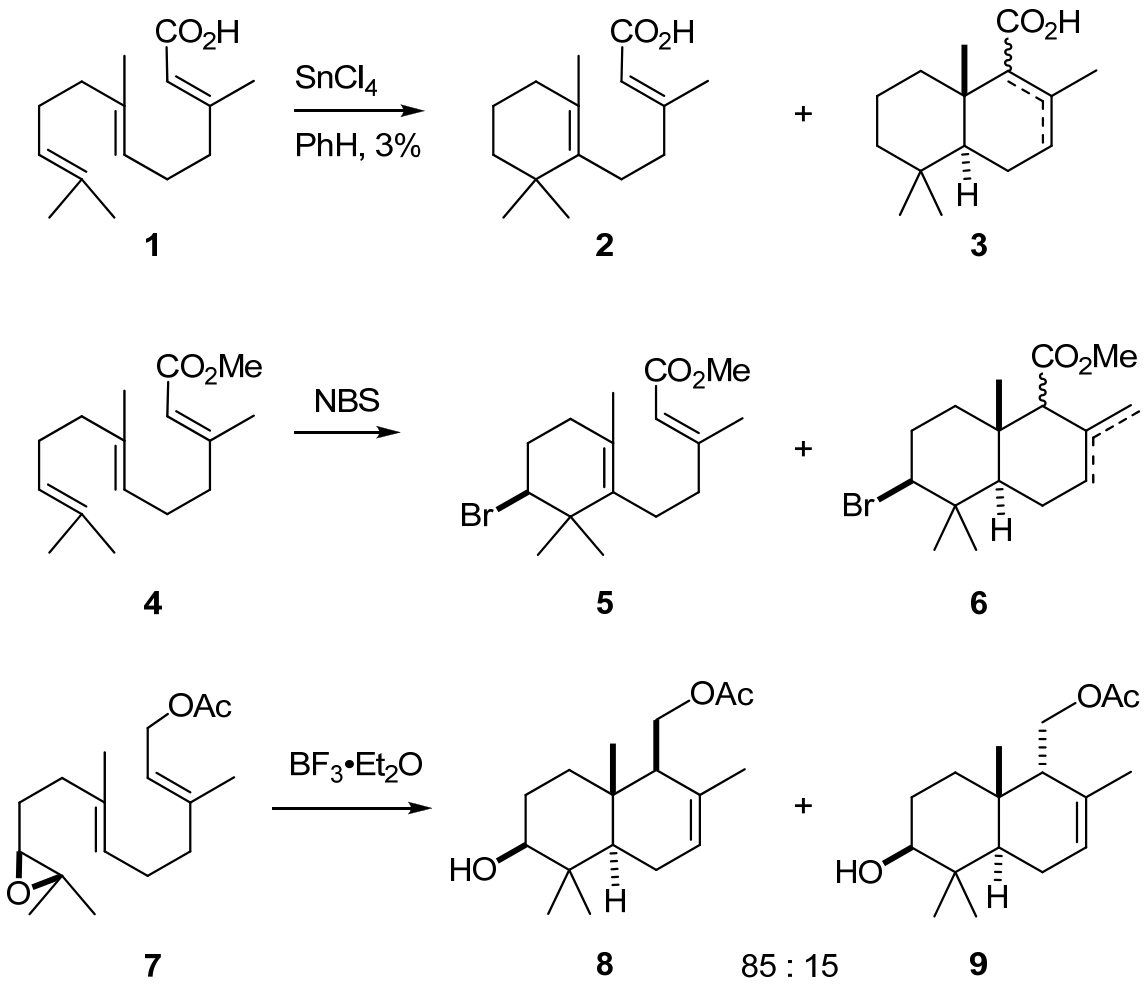
II.5.5 Synthesis and Chemistry of Sesquiterpenes.

In addition to the interest created by the novel and numerous carbon backbones, sesquiterpenes have spurred efforts in the development biomimetic reactions. A small group of sesquiterpenes possessing a decalin carbon skeleton reminiscent of the A/B ring system of steroids is known. These compounds are assumed to be derived biogenetically by protonation of the terminal double bond followed by cyclization of the other olefin (Scheme II.5.8).

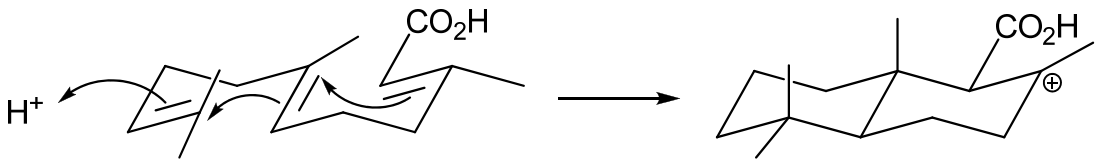


Scheme II.5.8

Reactions that mimic this biocatalyzed cyclization were conceived in the laboratory. Stereoselectivities were studied and useful synthetic starting materials were constructed. Stork and Bergstahler cyclized farnesic acid **1** with the Lewis acid tin tetrachloride and obtained a mixture of four compounds in only 3% yield! (Scheme II.5.9). Later efforts were more fruitful when it was found that *N*-bromosuccinimide (a source of Br^+) cyclized methyl farnesate **4** into compounds **5** and **6**. The monoepoxide **7** provided two compounds **8** and **9** upon treatment with BF_3 etherate (a Lewis acid). Notice that the chemist rarely uses Brønsted acids as the catalyst for cyclizations. That is because it is very hard to find a mild source of H^+ that will not also cause other unwanted side reactions. Of course biological systems use enzymes to catalyze such cyclizations and do not have that problem. In the laboratory, Lewis acids are often preferred because of their mildness, higher selectivity, and lower temperature requirements. Furthermore, notice that biomimetic cyclizations often lead to mixture of isomers. The formation of the major one can often be explained with conformational arguments. Frequently cyclizations occur via a six-membered ring transition state that adopts a **chair like** conformation (in preference to a boat for example) as shown in Scheme II.5.10. Enzymatic reactions usually produce one isomer with complete stereospecificity and stereoselectivity. The transition states are also thought to proceed via conformational control, though the enzyme could probably force the molecule into any desired conformation.

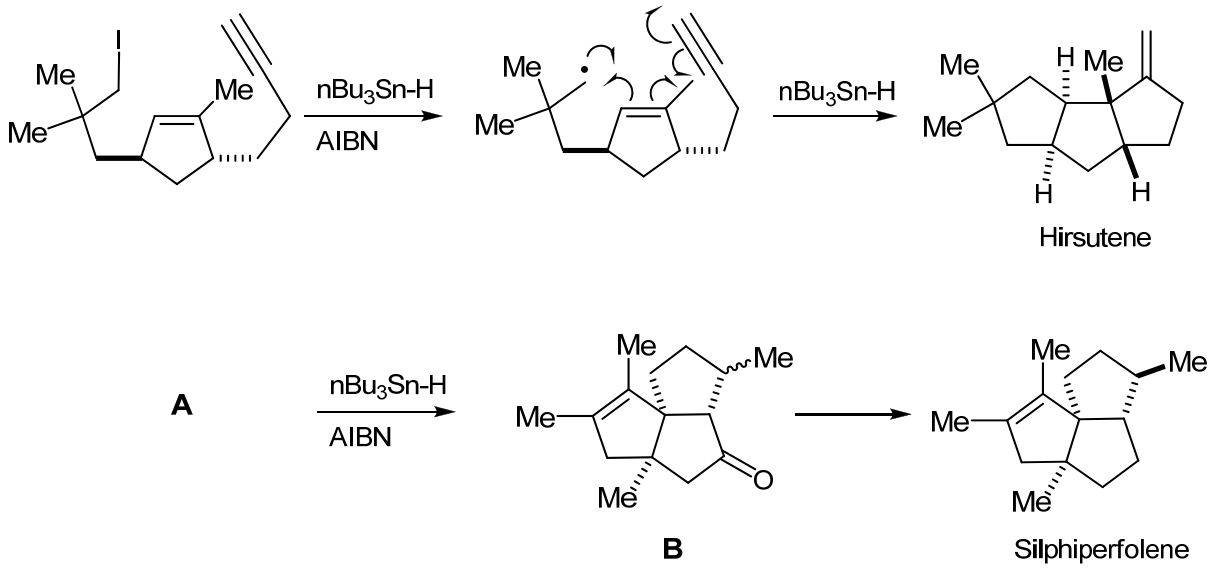


Scheme II.5.9



Scheme II.5.10

Radical polycyclization has also been used to make the tricyclic nucleus of some sesquiterpenes (Scheme II.5.11). Such tricyclic compounds were dubbed triquinanes. Radical cyclisation obey the **Baldwin rules**, an empirical model that predicts what cyclisation is favored or disfavored. Five-membered rings are by far the easiest ring size formed by such radical cyclisation but only when the alkene or alkyne is in the so-called 'exo' position (Scheme II.5.12). The terms 'trig' and 'dig' refer to 'trigonal' for double bonds and 'digonal' for triple bond. The term 'tet' for tetrahedral (not shown) refers to leaving groups on sp^3 carbons.

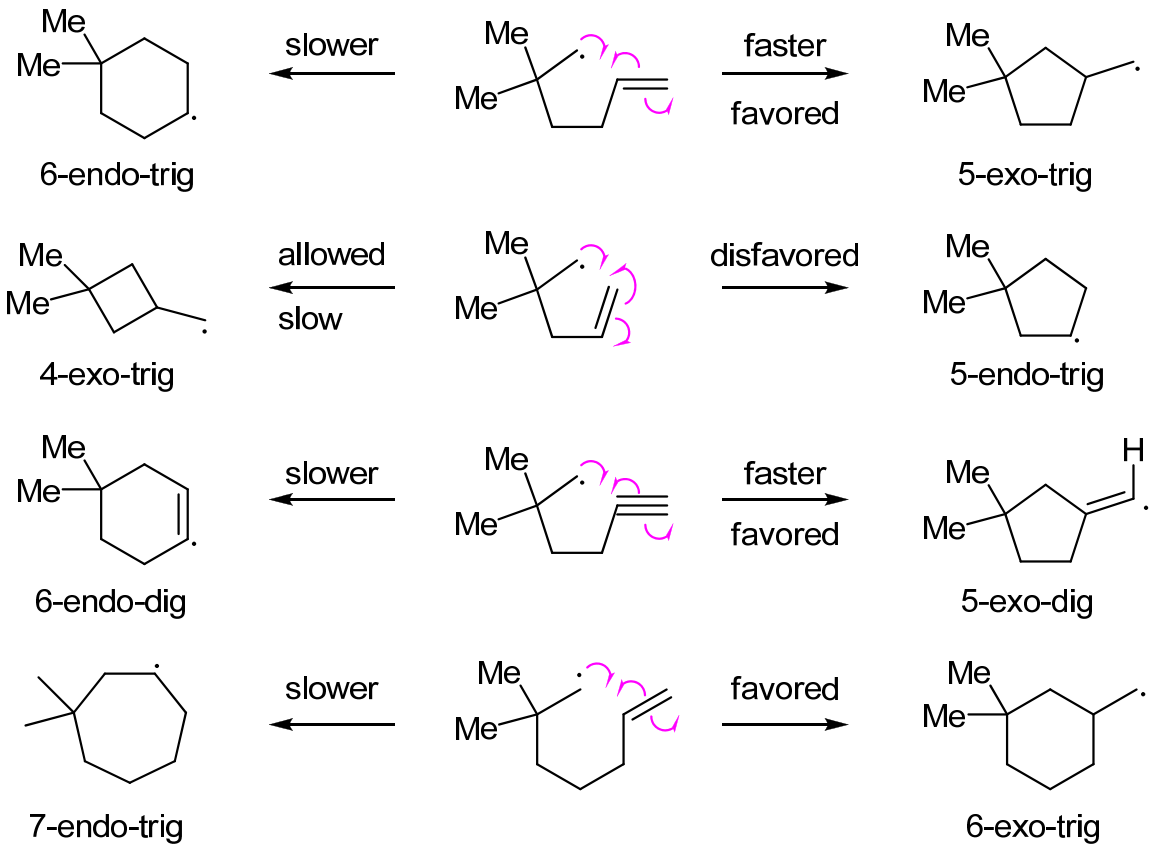


Scheme II.5.11

(Pouvez-vous proposer une structure du précurseur **A** qui polycyclise de façon radicalaire pour donner **B** tel que montré au schéma II.5.11?)

The ease with which the cyclisation takes place has to do with orbital alignment. Orbitals are best aligned when the cyclisation takes place in the 5-exo mode. The alignment is bad in the 5-endo mode, enough that most 5-endo cyclisation will not take place. However, the 6-exo and 6-endo mode are permitted but slower than the 5-exo mode. When a competition between a 5-exo or a 6-endo cyclisation occurs, the former usually takes place. The same is true between a 6-exo and a 7-endo, the former being usually faster. While the 4-exo-trig cyclisation is allowed, it is slow and reversible and such reaction are rarely seen.

The radical cyclisation is thus the realm of the 5-membered ring and it is most often use to construct cyclopentanes and somewhat used to make cyclhexanes but rarely used to make other ring size.



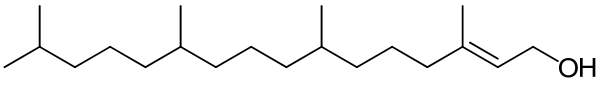
Scheme II.5.12

II.6 Biosynthèse des diterpènes

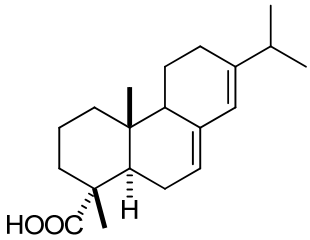
II.6.1 Généralités

Les diterpènes sont normalement constitués de quatre unités isoprène (C_{20}) mais peuvent contenir plus ou moins de 20 carbones selon les modifications apportées au cours de la biosynthèse par les enzymes. La structure de plus de mille diterpènes a été déterminée et la majorité de ces diterpènes sont des molécules cycliques constituées de un à cinq cycles. La plupart on été isolés de plantes terrestres, des moins évoluées aux plus évoluées. Certains insectes en fabriquent et ils servent soit d'agents de défense, soit de phéromones. Des plantes et organismes marins en fabriquent également. La figure II.6.1 montre quelques exemples de diterpènes. Il est intéressant de noter les formes énantiomères du kaurène et de l'*ent*-kaurène, deux produits naturels. Le glycoside du stéviol (aglycone) avec deux unités sucre, l'une attachée à la fonction alcool et l'autre au groupement carboxyle, est 350 fois plus sucré que le glucose. Les gibérellines sont responsables d'une maladie des jeunes plants de riz appelée « bakanae »

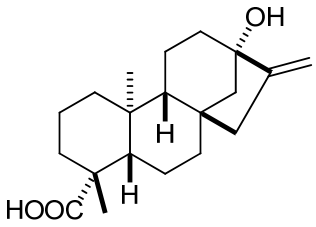
(surdéveloppement du jeune plant) et causée par le champignon *giberella fujikoroi.* La figure II.6.2 montre des exemples de métabolites contenant une sous structure diterpénique. Par exemple, le phytol est la partie alcool du groupement ester de la chlorophylle qui se retrouve dans toutes les plantes et c'est probablement l'isoprénoïde les plus abondant sur terre.



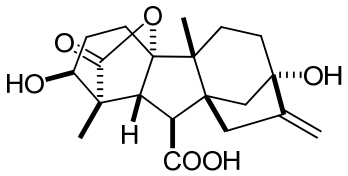
Phytol: plantes (produit d'hydrolyse de la chlorophylle) synthèse des vitamines E et K_1



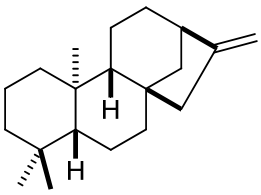
Acide abiétique: conifères (constituant majoritaire de l'oléorésine) synthèse d'esters utilisés dans les vernis et les lacques



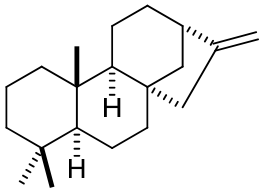
Stéviol: feuilles et tiges de Stevia rebaudiana (contrôle la croissance)



Gibbérelline A_3: champignon Giberella fujikuroi et plantes terrestres évoluées (hormone de croissance)



ent-Kaurène



Kaurène

Figure II.6.1

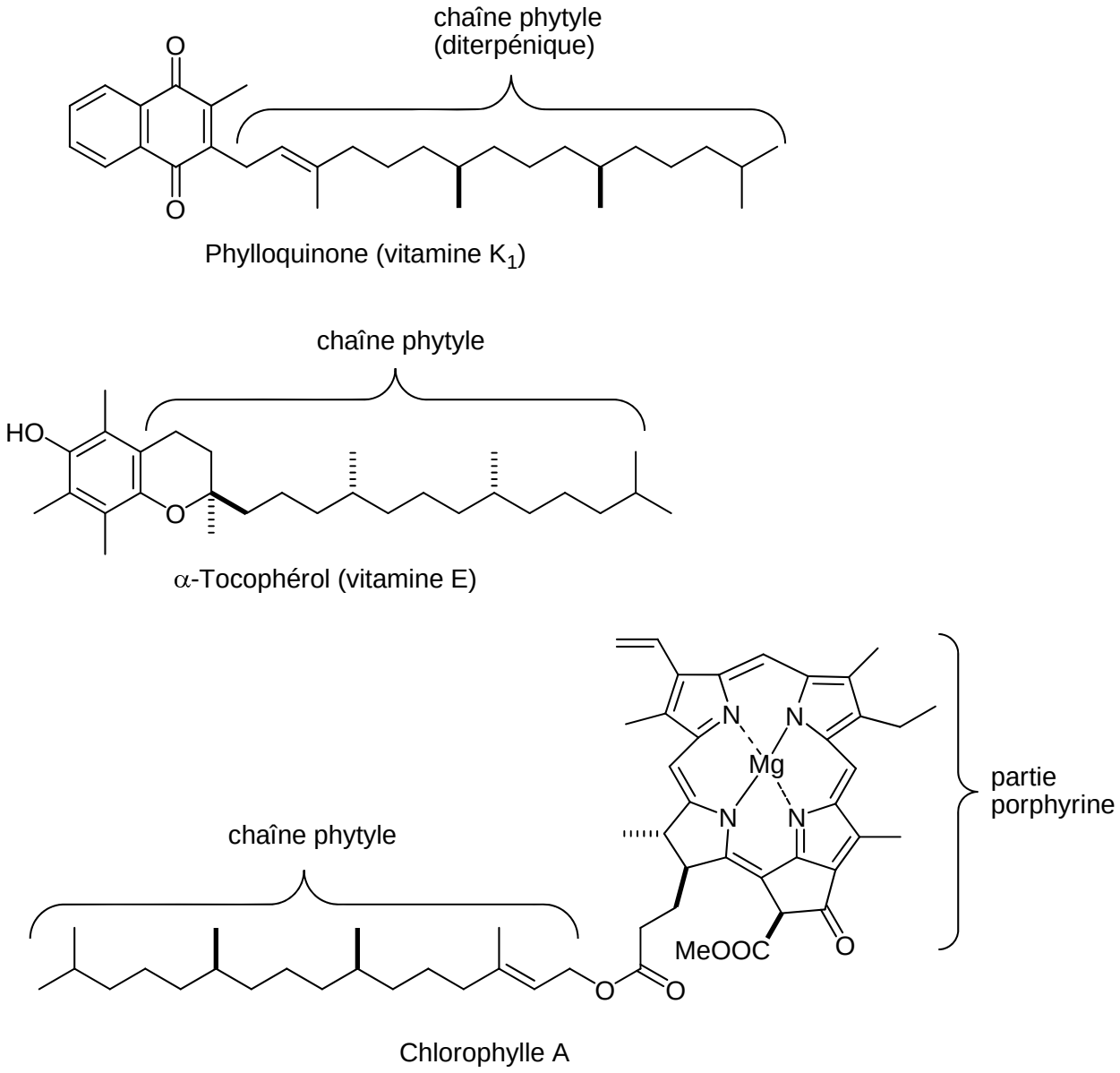


Figure II.6.2

II.6.2 Biosynthèse

II.6.2a Cyclisation du pyrophosphate de géranylgéranyle (GGPP)

Le GGPP est le précurseur de pratiquement tous les diterpènes mais aussi des tétraterpènes (C_{40}) ou carotènes comme il sera vu plus loin (chapitre II.9). Il occupe un rôle central dans la biosynthèse des polyterpènes. Le schéma II.6.1 montre trois façons de cycliser le GGPP dont

deux où le GGPP est replié dans des conformations chaise-chaise-chaise images de miroir l'une de l'autre. Les enzymes contrôlent ce pliage.

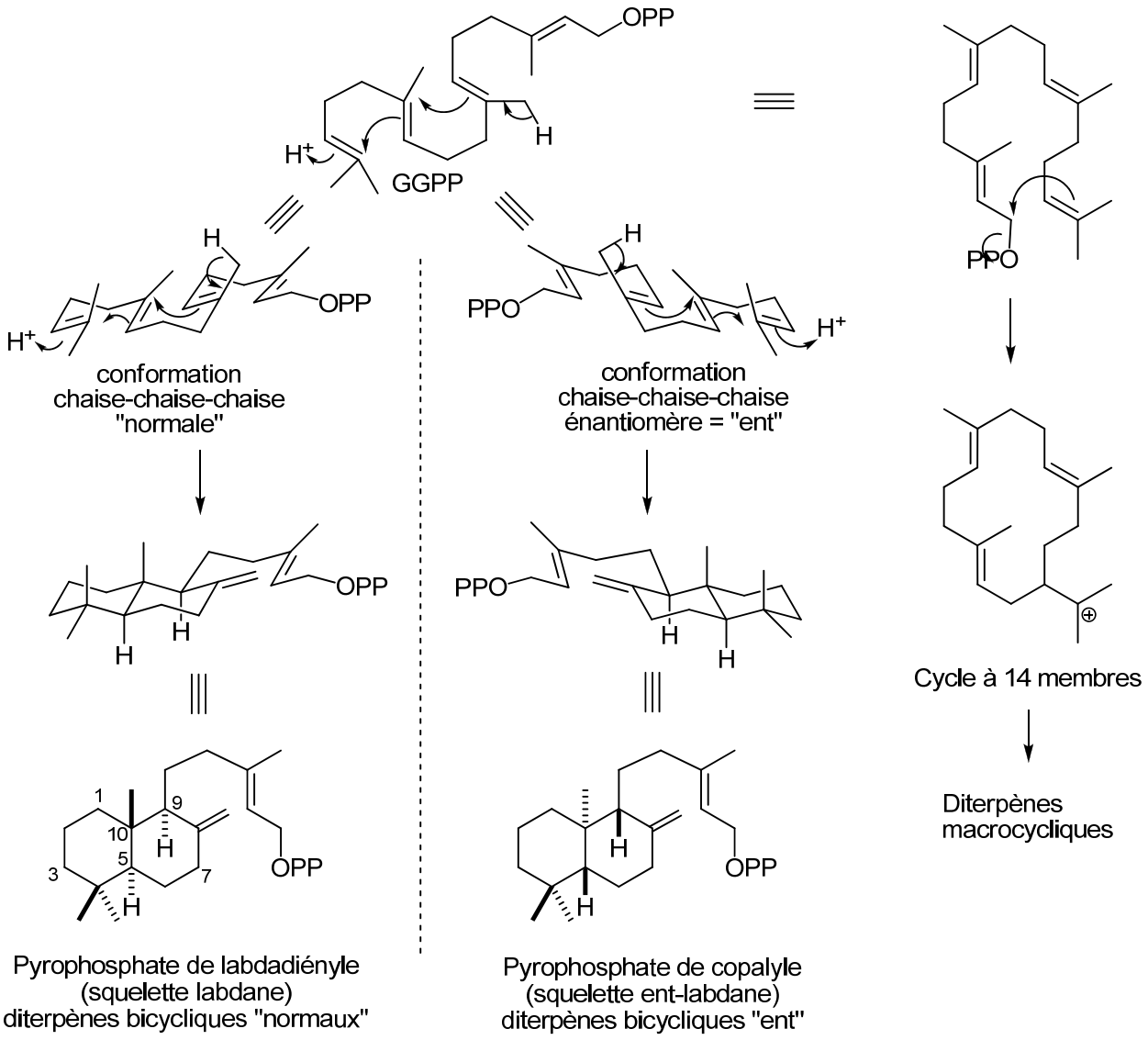


Schéma II.6.1

Les enzymes impliquées dans la cyclisation du squalène (biosynthèse des triterpènes et des stéroïdes) plient le squalène et l'époxyde de squalène (voir section II.8) dans une seule des deux conformations chaise-chaise-chaise énantiomères possibles, celle correspondant à la conformation qualifiée de « normale » dans le schéma II.6.1. La cyclisation du GGPP dans cette conformation conduit donc à la même stéréochimie relative en C5, C10 et C9 (numérotation utilisée pour les stéroïdes) que celle des stéroïdes et des triterpènes, soit l'arrangement *trans-anti*. Le fait de retrouver quelques diterpènes dans les deux série énantiomériques dans la nature est

une caractéristique qui leur est pratiquement unique. Il arrive même qu'une plante produise les deux énantiomères, bien que ce phénomène soit plus rare. La cyclisation est initiée par la protonation de la double liaison C14=C15. La troisième façon de cycliser le GGPP est celle conduisant au macrocycle à 14 membres, précurseur des sesquiterpènes macrocycliques de la famille des cembranes, par attaque Markovnikov de la double liaison C14=C15 sur C1 (mécanisme S_N2 ou S_N1).

II.6.2b Cyclisation du pyrophosphate de labdadiényle

Le schéma II.6.2 illustre la cyclisation du pyrophosphate de labdadiényle en squelette tricyclique pimarane, série « normale ». Le pimaradiène, l'acide abiétique et l'acide pimarique (non montré, aussi un constituant de l'oléorésine des conifères) sont trois représentants de ces diterpènes tricycliques « normaux ». Le kaurène émane du cation pyraménYle via le carbocation kaurénYle par cyclisation, migration-[1,2] d'un carbone et perte d'un proton.

La cyclisation du cation 8-pimarénYle conduit aux diterpènes tétracycliques de la série « normale » kaurane dont le kaurène est un exemple. Comme le montre le schéma II.6.3, le cation 8-pimarénYle est aussi le précurseur de la rosénonolactone, un diterpène tricyclique, et de l'amphidocolène, un diterpène tétracyclique.

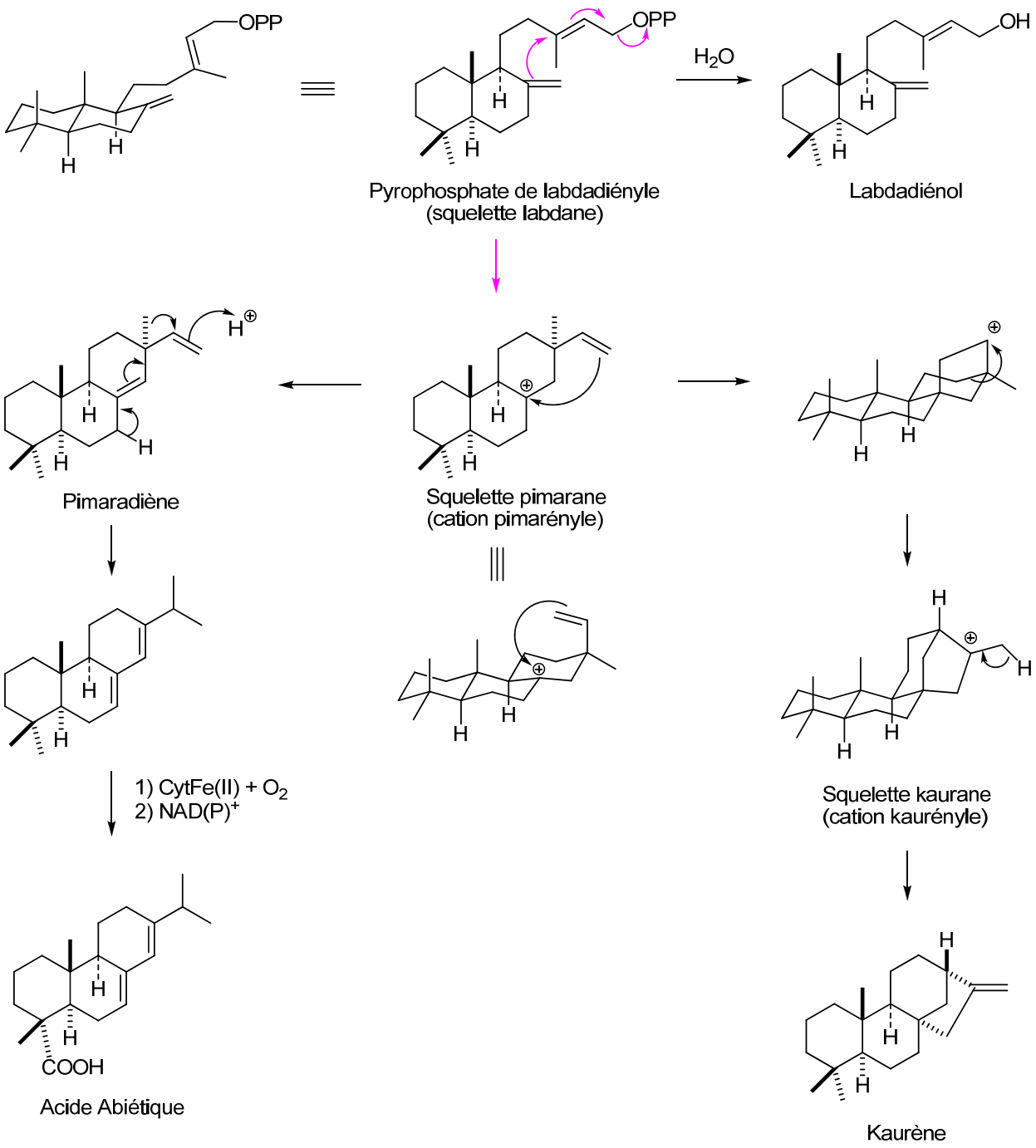

OPP
H
H
OPP
H
H
H2O
OH
H
H
Pyrophosphate de labdadiényle
(squelette labdane)
Labdadiénol
H
H
H
Pimaradiène
H
H
Squelette pimarane
(cation pimarényle)
H
H
H
H
H
H
H
H
1) CytFe(II) + O2
2) NAD(P)+
COOH
Acide Abiétique
Squelette kaurane
(cation kaurényle)
H
H
H
Kaurène

Schéma II.6.2

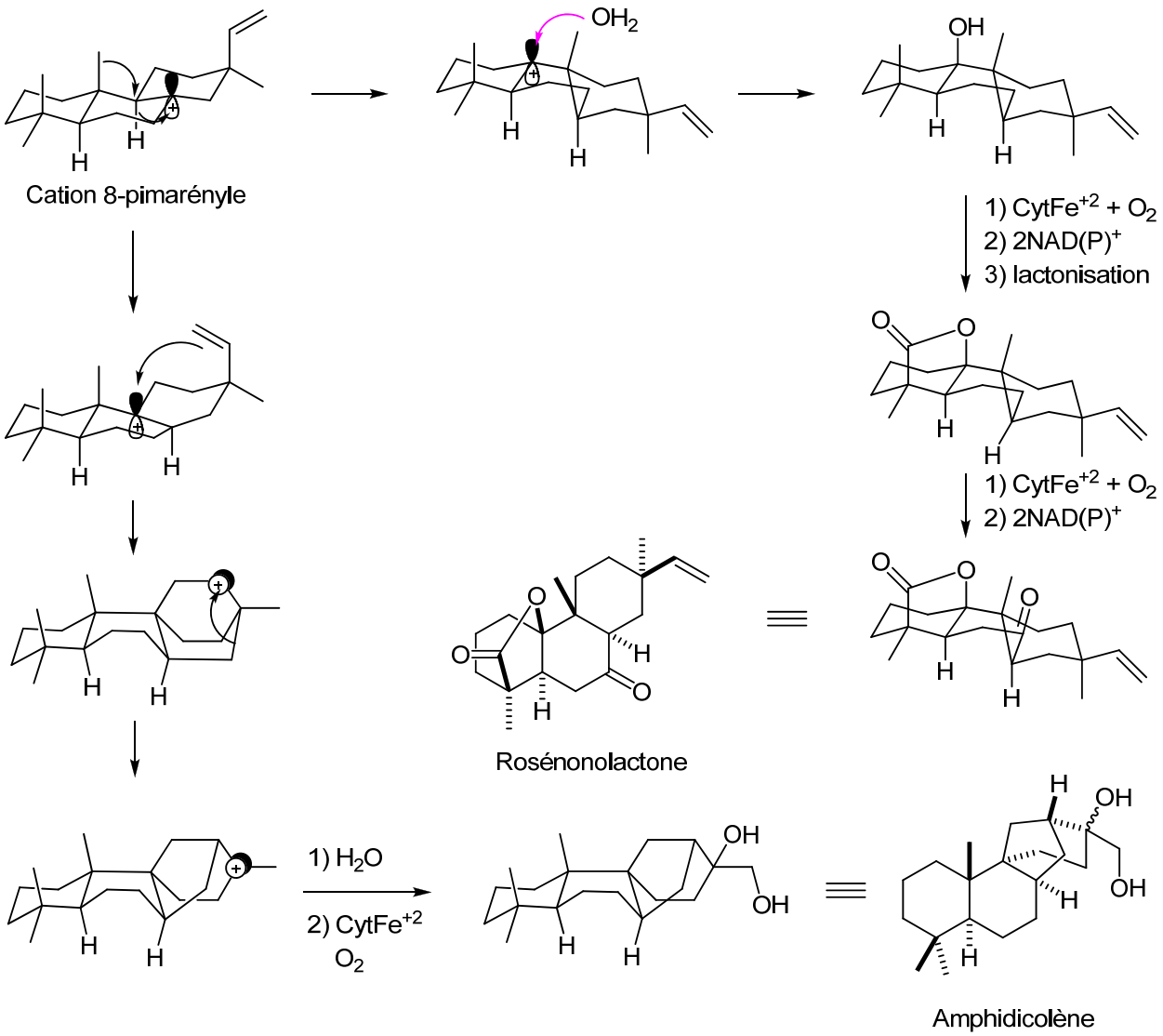


Schéma II.6.3

Une autre façon de cycliser le pyrophosphate de labdadiényle est illustrée au schéma II.6.4. Celle-ci conduit d'abord à un cation isomère à la position 13 (numérotation des stéroïdes) du cation 8-pimarényle. Ce dernier, comme le cation 8-pimarényle lui-même, peut aussi être converti en acide abiétique car au cours de cette conversion la configuration du centre asymétrique en C_{13} est détruite. Cependant sa cyclisation conduit au squelette tétracyclique stachane, isomère aux positions 8 et 13 du squelette kaurane. L'époxyde d'hibaène est un membre de la famille des diterpènes tétracycliques stachanes.

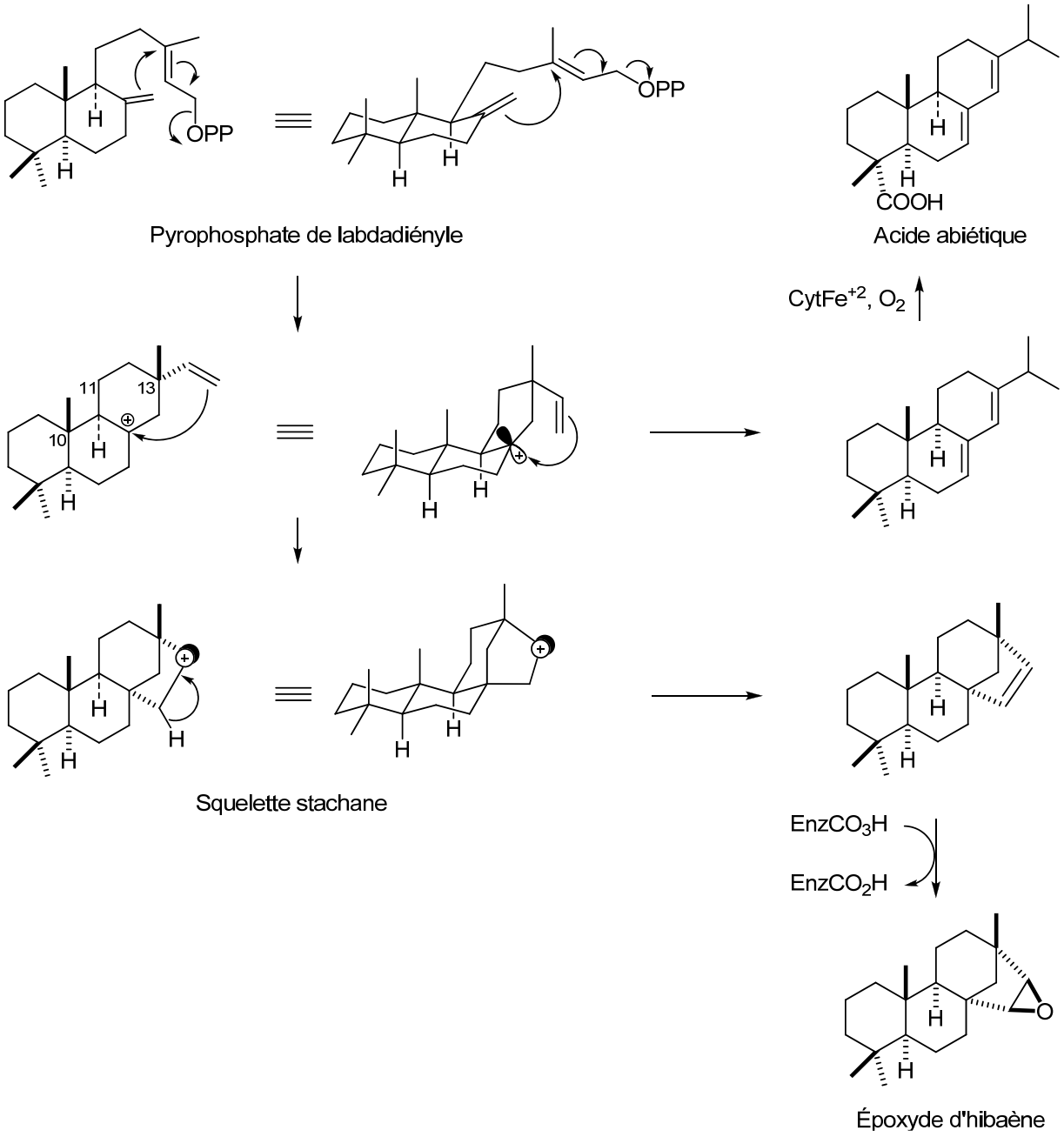


Schéma II.6.4

II.6.2c Cyclisation du pyrophosphate de *ent*-labdadiényle

La cyclisation décrite au schéma II.6.5 conduit au squelette tétracyclique ent-kaurane (C_{20} toujours) et le *ent*-kaurène est un représentant de cette famille de diterpènes tétracycliques. Le *ent*-kaurène est le précurseur des diterpènes tétracycliques de la famille des gibérellanes.

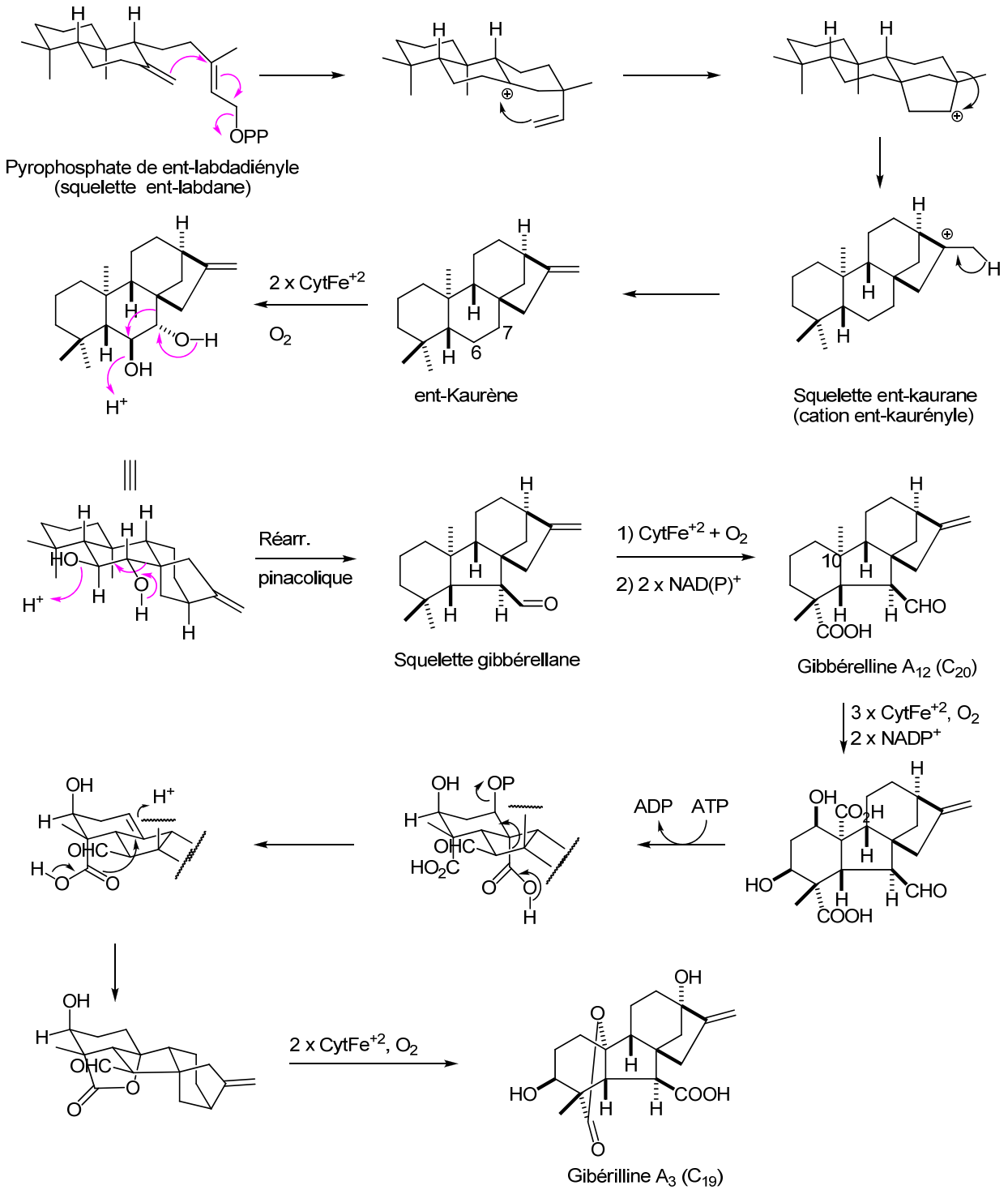


Schéma II.6.5

Ainsi, l'oxydation du *ent*-kaurène aux positions 6 et 7 suivie du réarrangement pinacolique du diol conduit au squelette gibérellane. L'oxydation du méthyle α en position 4 donne la gibérelline A_{12}, un diterpène tétracyclique en C_{20}. Certains membres de cette famille n'ont que

19 atomes de carbone (C_{19}) comme le montre la biosynthèse de la gibérelline A_3 : le méthyle en position 10 (numérotation des stéroïdes et aussi celle du squelette gibérellane pour ce carbone) est éliminé au cours de la transformation de la gibérelline A_{12} en gibérelline A_3.

II.6.2d Diterpènes macrocycliques et diterpènes polycycliques dérivés

Tel que déjà vu, la biosynthèse des diterpènes macrocycliques à partir du GGPP fait intervenir une attaque Markovnikov de la double liaison sur C_1 porteur du nucléofuge pyrophosphate (mécanisme S_N2 ou S_N1). Les schémas II.6.6a-c décrient plus en détails la biosynthèse du casbène et de dérivés du squelette cembrane dont les squelettes taxane, lathyrane et tigliane. L'attaque du pyrophosphate par la double liaison C_{14}=C_{15} du GGPP conduit à un carbocation non classique qui peut être en équilibre avec un carbocation classique ayant le squelette macrocyclique cembrane. Ce dernier donne le squelette tricyclique du taxane après deux cyclisations. La première cyclisation peut aussi être faite directement à partir du carbocation non classique tel qu'indiqué par les flèches. La déprotonation de ce dernier donne le casbène qui est le précurseur des squelettes lathyrane et tigliane. Le phorbol, un diterpène tricyclique de la famille des tiglianes, est un agent co-carcinogène puissant, surtout si ingéré dans une diète incluant aussi beaucoup d'acides gras et d'hydrocarbures. Le taxol, un membre de la famille des taxanes isolé de l'écorce d'un if du pacifique, est un agent anti-cancérigène pour plusieurs types de cancers chez l'humain.

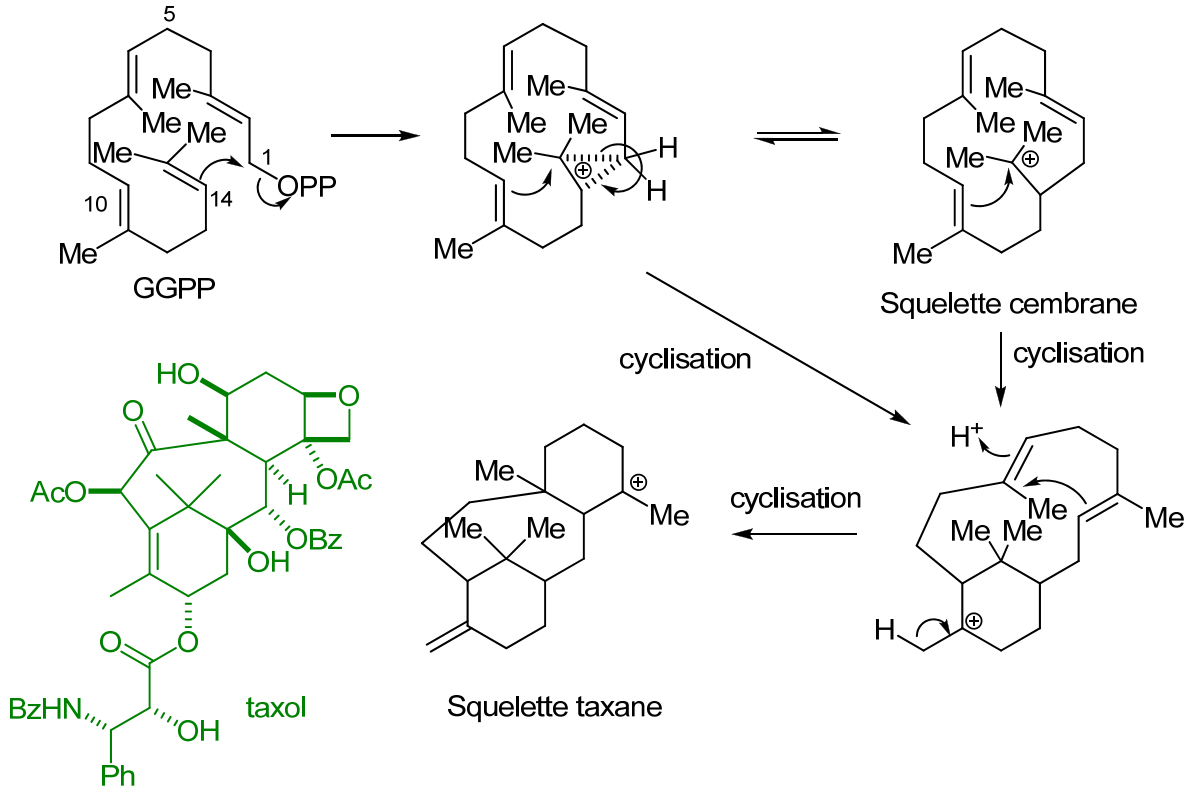


Schéma II.6.6a

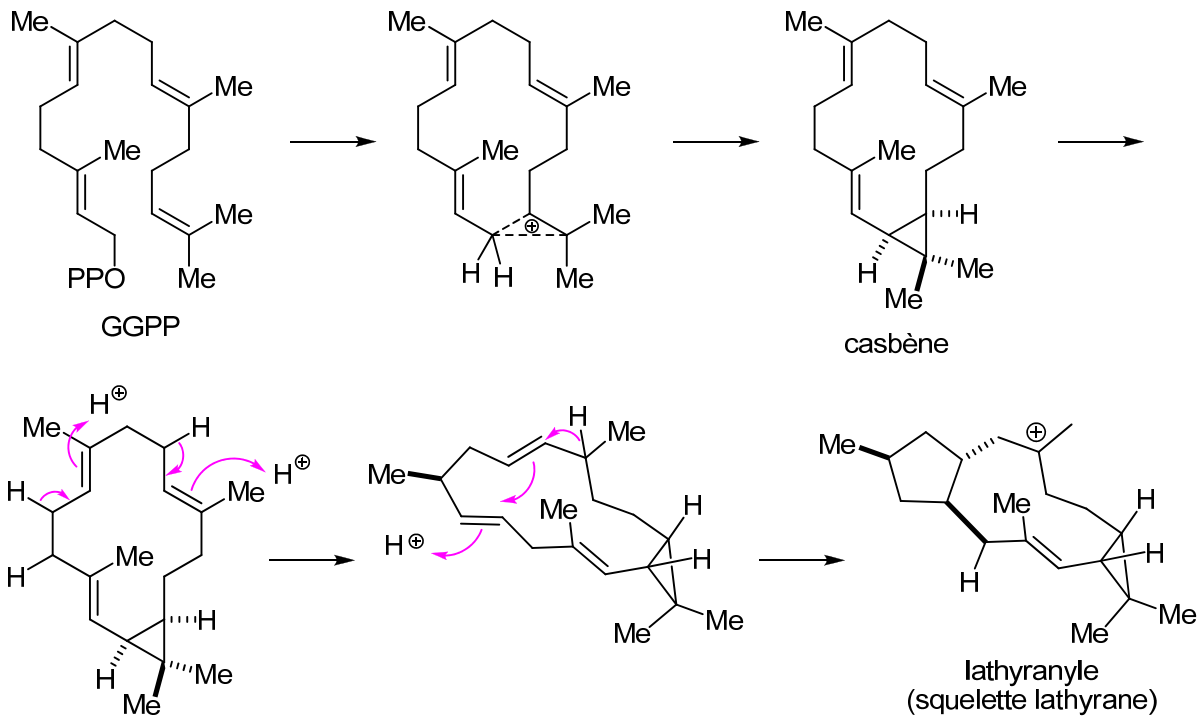


Schéma II.6.6b

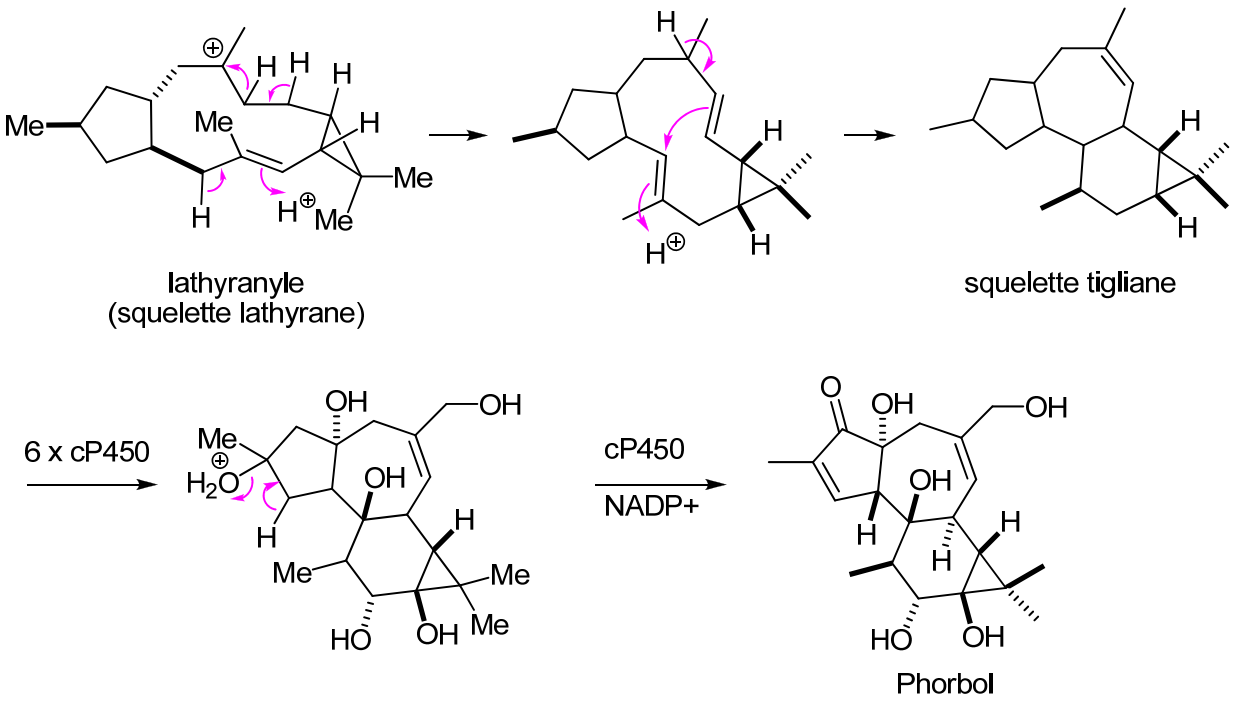
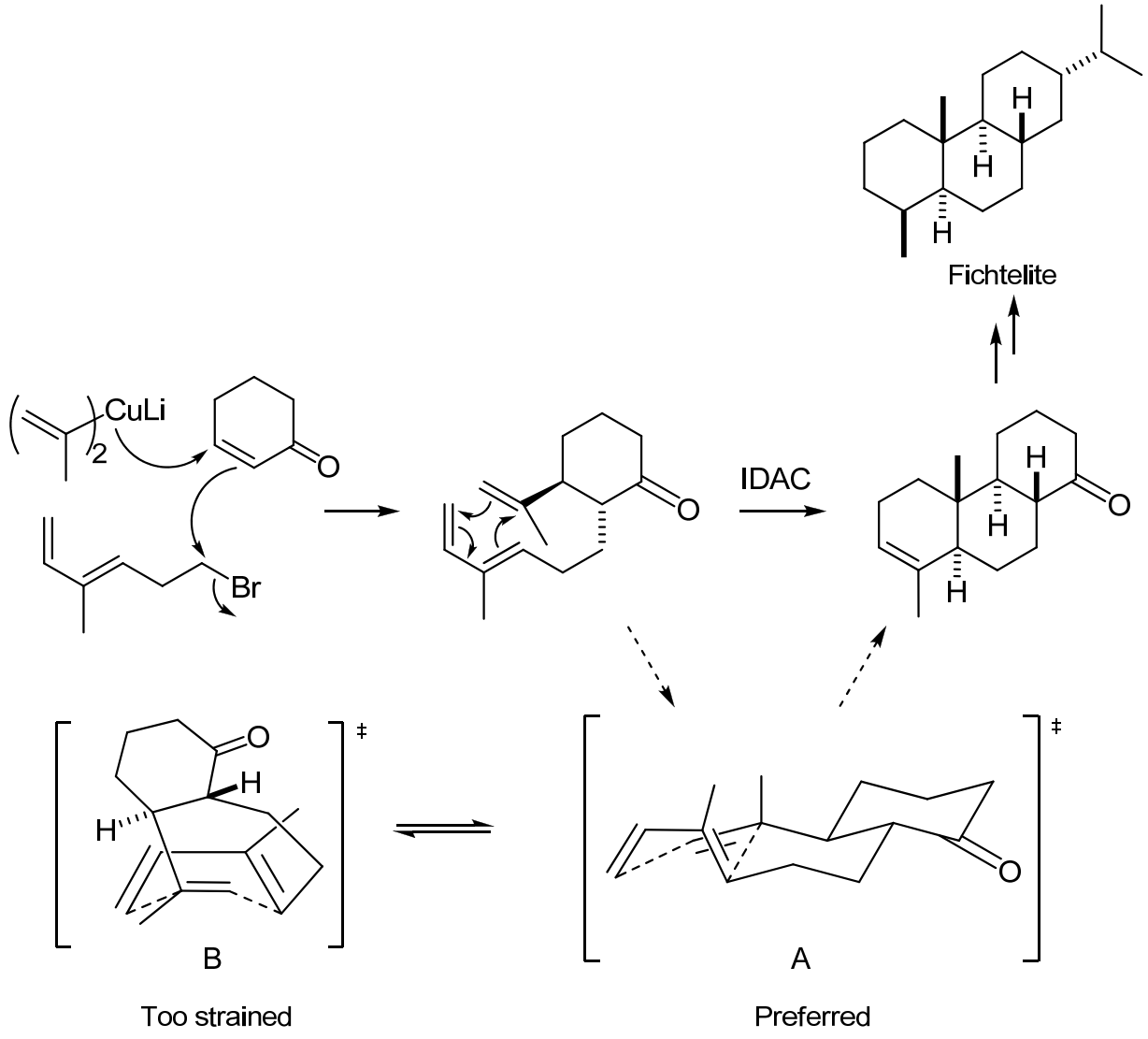


Schéma II.6.6c

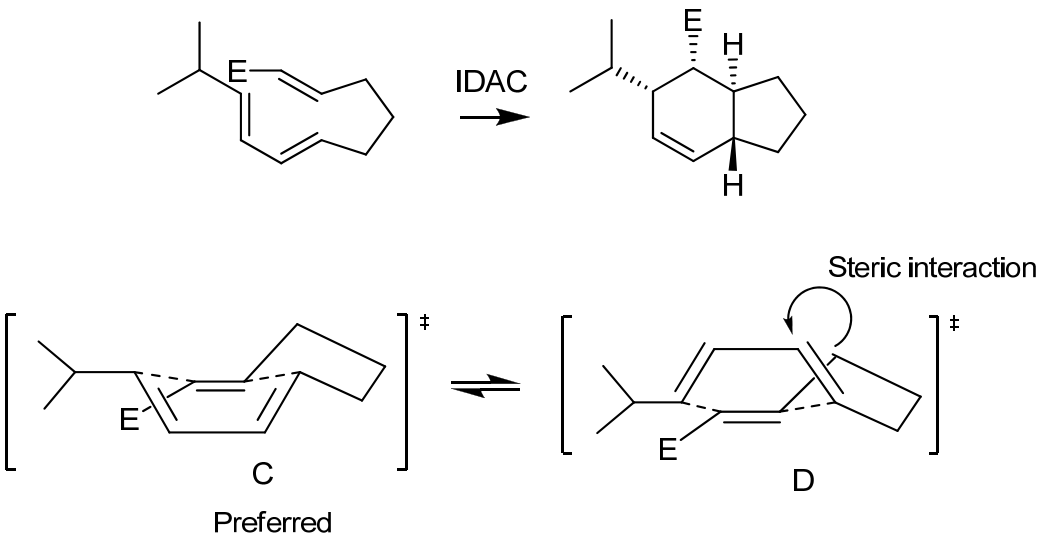
II.6.3 Synthesis and Chemistry of Diterpenes.

The tricyclic nucleus of many di- and higher terpenes, was often assembled using the Robinson annulation or, more recently, the intramolecular Diels-Alder cycloaddition (IMDAC). In the latter case, studies demonstrated that a high level of regio- and stereochemistry could be achieved. It became quickly one of the most powerful method for constructing that type of tricyclic molecule. Let's look at some examples and at the reasons behind the high selectivity of that reaction. An excellent review of the IMDAC reaction by Alex Fallis can be found in Can. J. Chem. **1984**, *62*, 183. The IMDAC requires that at least three carbons link the diene to the dienophile. Moreover, depending on the regiochemistry of the reaction, two types of molecules can result: fused or bridge bicyclic molecules. Fused rings are more abundant in nature and will be the focus of this discussion. Scheme II.6.7 shows a short synthesis of a tricyclic diterpenoid nucleus which was later converted to fichtelite, a diterpene belonging to the abietane family. The regiochemistry of the IMDAC in this case arises from a net preference of the transition state **A** over **B**, the latter having too much strain in the carbon tether (Scheme II.6.7). The *trans* stereochemistry of the two fused six-membered rings is derived from a preference of the tether to adopt a chair conformation (like a six membered ring as shown in **A**) in which all the substituents are at a position of minimum energy (equatorial in this case). There are of course

exceptions to these rules, and any IMDAC's stereochemical outcome must be assessed independently. Tethers of 3 carbons will give fused 5,6-membered rings with primarily a *trans* stereochemistry. The transition states involved in this reaction are shown in Scheme II.6.8. The slight preference for *trans* stereochemistry is due to the steric interaction between the carbons indicated by the arrow. Substitution on the tether can change or reverse completely the sense of selectivity. Predicting the stereochemical outcome of any IMDAC reaction requires careful analysis of all electronic and steric factors involved which is beyond the scope of this discussion. The student is referred to the review article cited above for further reading.



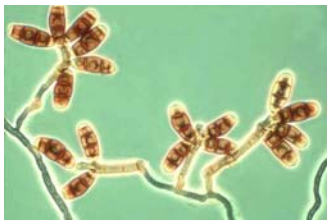
Scheme II.6.7



Scheme II.6.8

II.7 Biosynthèse des sesterterpènes

Les sesterterpènes (C_{25}) sont constitués de cinq unités isoprène provenant du couplage tête à queue entre le GGPP et l'IPP pour donner le pyrophosphate de géranylfarnésyle (GFPP). Il y a peu de familles structurales et ils proviennent presque toutes de champignons et d'organismes marins. Le géranylfarnésol a été isolé des sécrétions (cire) de l'insecte *ceroplastes albolineatus*. Les sesterterpènes sont les membres les plus nouveaux et les moins nombreux de la famille des terpénoïdes. Le premier, l'ophioboline A, un sesterterpène tricyclique, a été isolé en 1965 du champignon pathogène *cochliobolus miyabeanus*. Ce champignon de l'espèce ascomycètes est responsable d'une des plus graves maladies du riz appelé helminthosporiose. L'ophioboline A est une phytotoxine qui stimule la perte d'électrolytes et de glucose des racine de maïs et de riz. Depuis, un certain nombre d'ophiobolanes ont été isolés de ce type de champignons. D'autres sesterterpènes ont été isolés de fougères, de lichens et d'éponges marines. Environ cinq types différents de squelettes sesterterpéniques sont connus. Les schémas II.6.9 à II.6.12 montrent la biosynthèse de quatre types de squelette à partir du GFPP : les ophiobolanes (schéma II.6.9); deux sesterterpènes tétracycliques, l'acide rétigéranique (isolé d'un lichens, schéma II.6.10) et la scalarine (isolée d'une éponge marine, schéma II.6.11); et un furanosesterterpène (isolé d'une éponge marine, schéma II.6.12).



cochliobolus miyabeanus

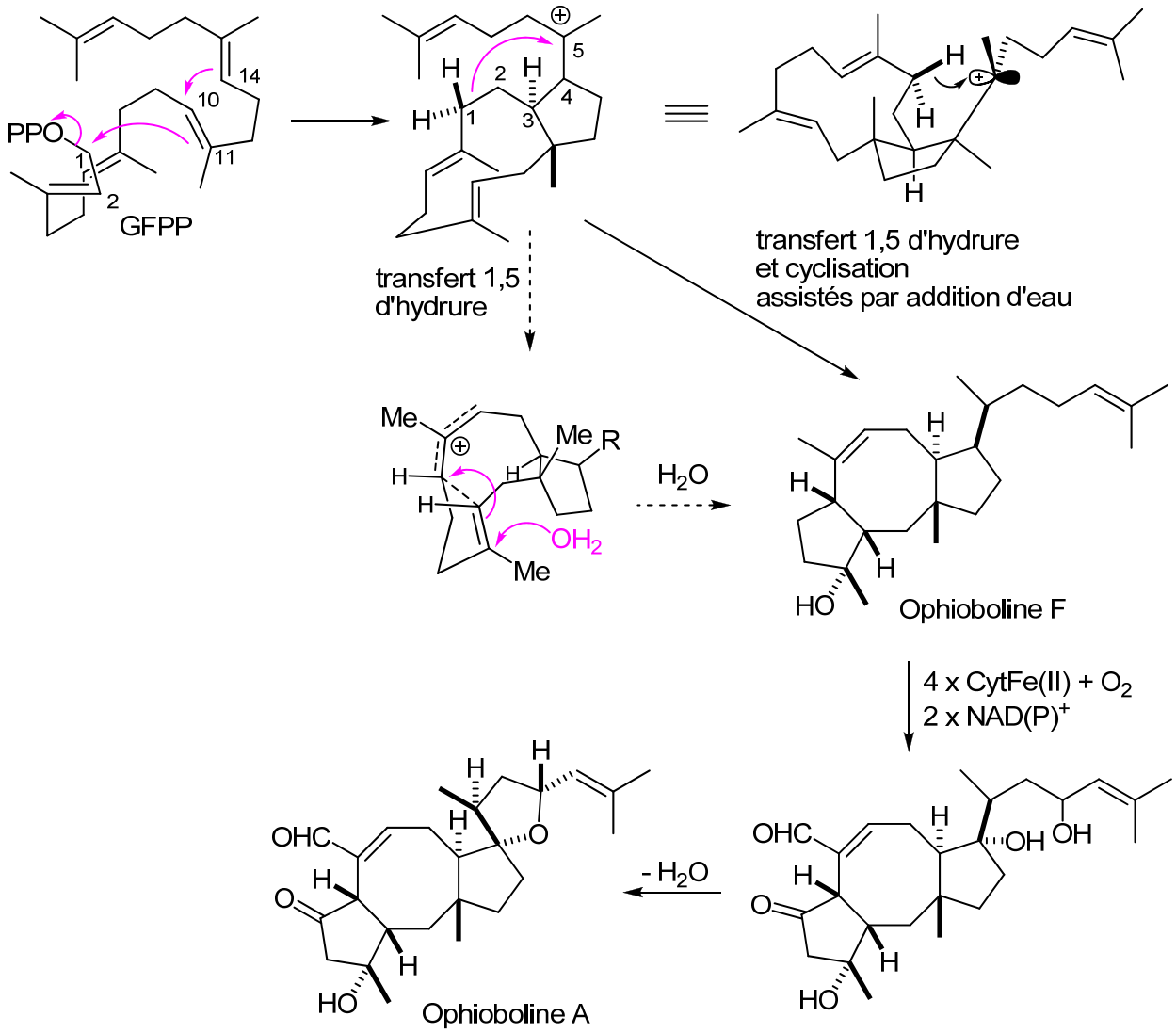


Schéma II.6.9

La biosynthèse des ophiobolanes (schéma II.6.9, page précédente) et celle de l'acide rétigéranique (schéma II.6.10) illustrent de nouveau qu'un enzyme peut placer le substrat dans une conformation propice à la réaction dans son site actif. C'est le cas pour le pliage du GFPP par une cyclase de façon à diriger la cyclisation dans ces deux biosynthèses. Dans la biosynthèse

des ophiobolanes, il y a double cyclisation, soit entre les positions 10 et 14 (addition Markovnikov à la double liaison C_{14}=C_{15}) et entre les positions 1 et 11 (addition anti-Markovnikov à la double liaison C_{10}=C_{11}), pour donner un système bicyclique [9.3.0] (fusion entre un cycle de 11 et un cycle de 5). Le cycle de 11 peut facilement accommoder deux doubles liaisons de configuration E et une fusion *trans* des deux cycles. Une transférase donne ensuite la bonne conformation du système permettant le transfert 1,5 de l'hydrure probablement concerté avec la cyclisation et l'addition d'eau (cette conformation est facilement obtenue avec les modèles moléculaires) pour donner le squelette tricyclique 5-8-5 des ophiobolanes.

Dans la biosynthèse de l'acide rétigéranique (schéma ci-dessous), le pliage du GFPP conduit à une cyclisation entre les doubles liaisons C_{18}=C_{19} (Markovnikov) et C_{14}=C_{15} (anti-Markovnikov) avec déplacement de l'anion pyrophosphate pour donner un squelette bicyclique [13.3.0] (fusion entre un cycle de 15 et un cycle de 5). Il y a encore dans cette biosynthèse une migration 1,5 d'hydrure probablement assistée par l'attaque de la double liaison (la cyclisation) dans une conformation favorable du bicycle.

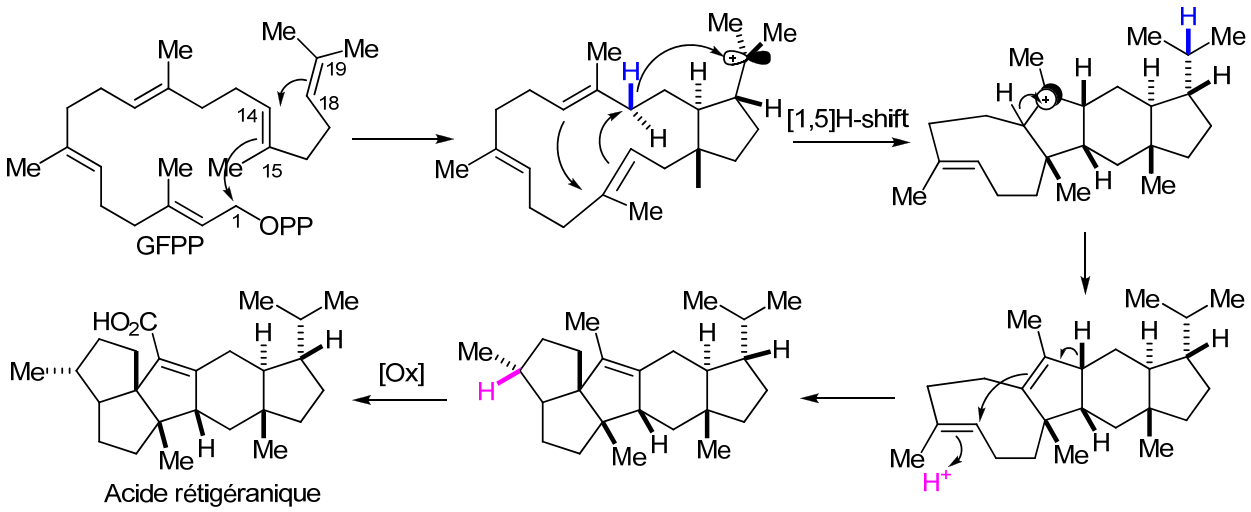


Schéma II.6.10

La biosynthèse de la scalarine est décrite au schéma II.6.11 de la page suivante. Le GFPP est plié dans une conformation chaise-chaise-chaise. La polycyclisation initiée par la protonation Markovnikov de la double liaison C_{18}=C_{19} donne un squelette tétracyclique avec une fusion *trans* de tous les cycles (arrangement *trans-anti-trans-anti-trans*). Les étapes qui suivent sont classiques en biosynthèse : des migrations sigmatropiques [1,2] d'hydrure et de méthyle, des oxydations ($[O]_a$ = CytFe(II) + O_2; $[O]_b$ = $NAD(P)^+$ ($-NAD(P)H$, $-H^+$), des éliminations et

additions. La dernière étape est une acétylation d'un groupement hydroxyle et les enzymes utilisent l'acétylcoenzyme A pour acétyler un alcool, un phénol ou une amine.

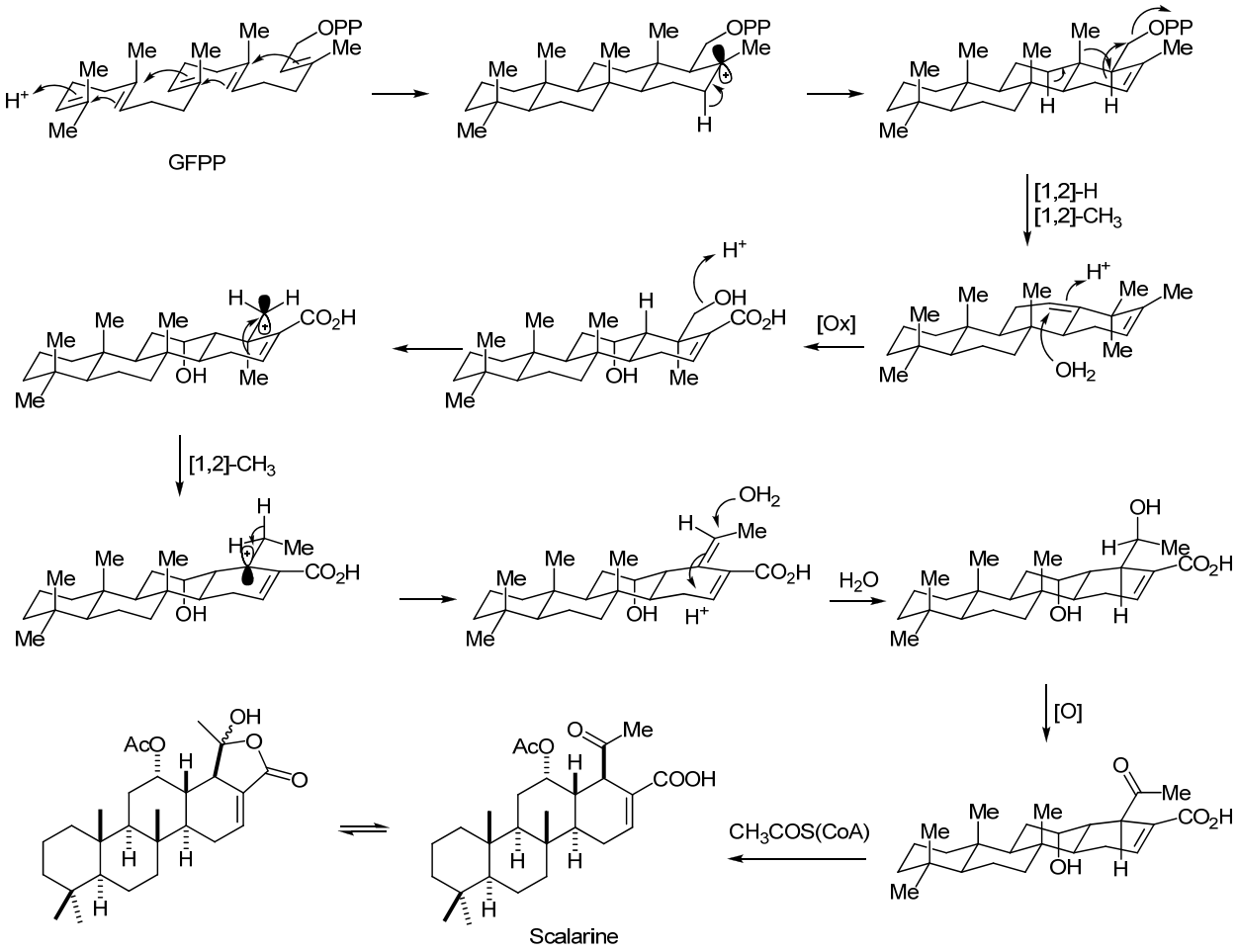


Schéma II.6.11

La biosynthèse des furanosesterterpènes ircinin-1 et irnicin-2 est décrite au schéma II.6.12. Tous les oxygènes, ceux des deux noyaux furanne et ceux de la lactone insaturée (buténolide) proviennent de l'oxygène de l'air. Ils sont tous introduits par oxydation avec CytFe(II) + O_2. L'oxydation subséquente des alcools secondaires en cétones et de l'alcool primaire en acide se fait avec $NAD(P)^+$.

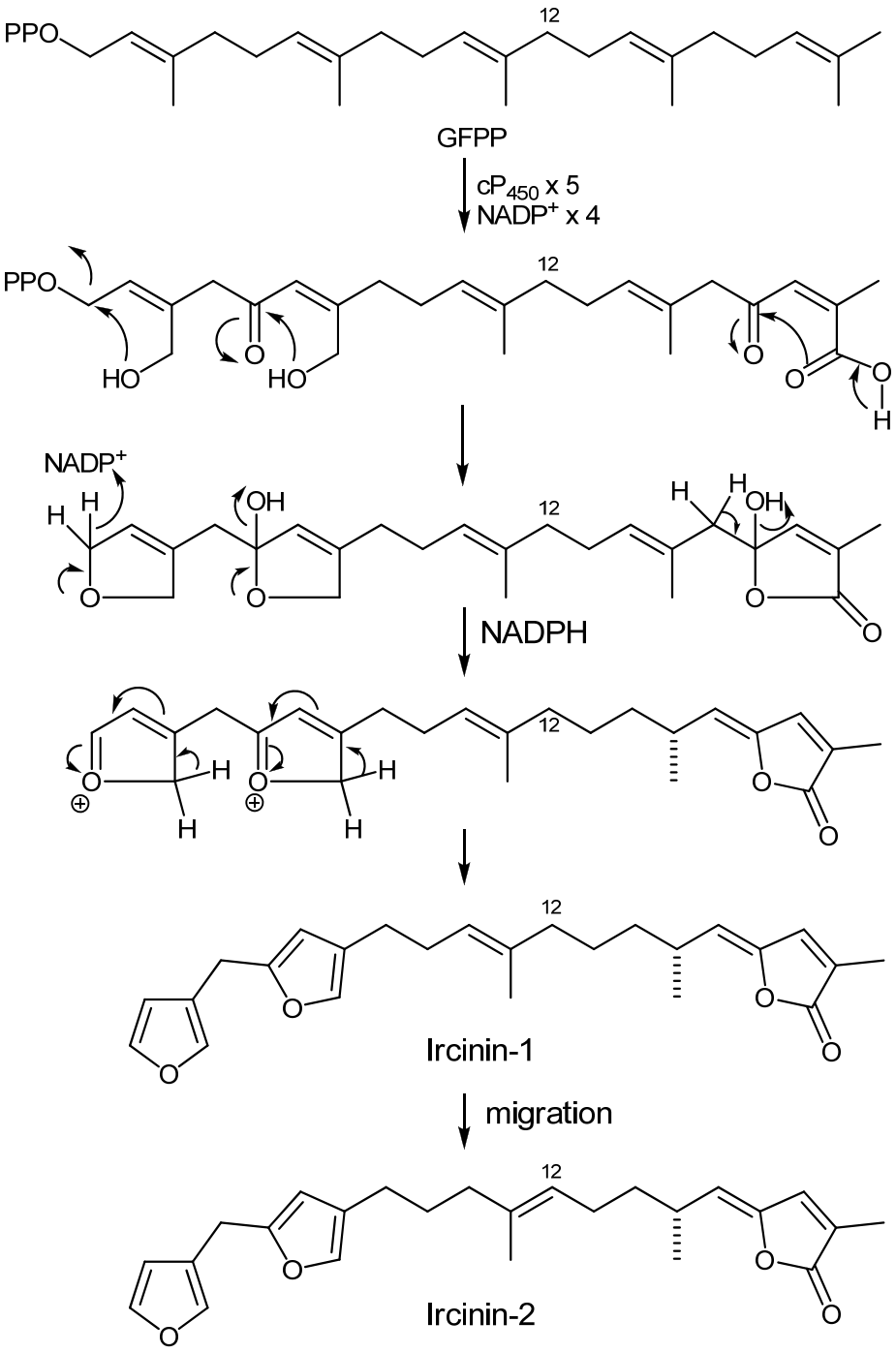


Schéma II.6.12

Il existe toute une famille de furanosesterterpènes qui possèdent pour la plupart des propriétés cyctotoxiques aigües. Par exemples, l'éponge marine *Sarcotragus sp.* (famille Thorectidae, order Dictyoceratida, retrouvée dans la mer méditerranée) produit une panoplie de

furanosesterterpènes qui sont actives contre 5 types de cellules tumorales humaines (Figure II.6.3).

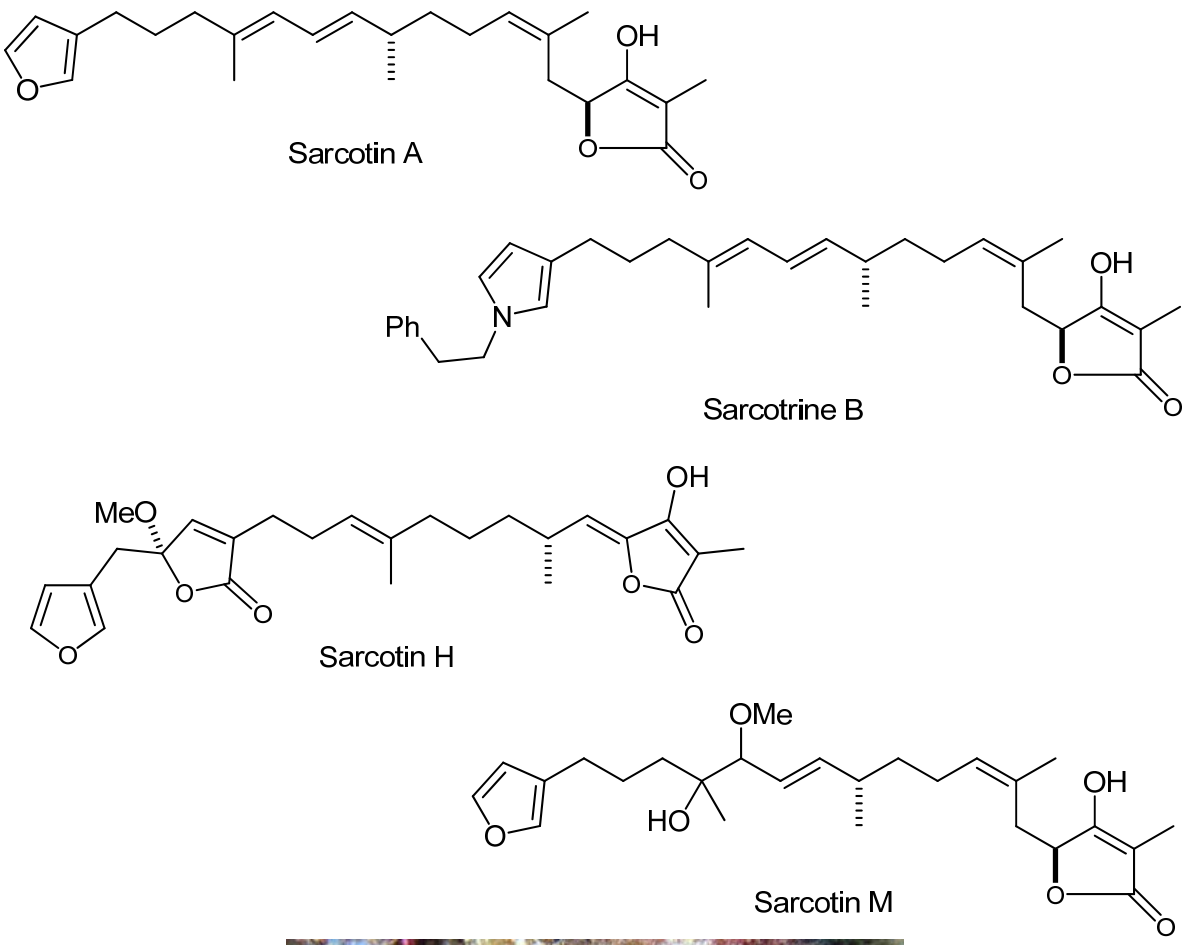


Figure II.6.3

II.8 Biosynthèse des triterpènes et des stéroïdes

II.8.1 Généralités et historique

On a proposé dès 1926 que le squalène (isolé de l'huile de requin mais qui se retrouve un peu partout dans la nature) pouvait être le précurseur du cholestérol dans la biosynthèse de ce stéroïde constitué de 27 carbones (C_{27}). En 1934, Robinson montrait qu'on pouvait plier le squalène de façon à surperposer 27 de ces carbones aux 27 carbones du cholestérol tel que montré à la figure II.7.1. Cependant, du fait que certains groupes méthyle du squalène ne se retrouvent pas dans le cholestérol, il a fallu attendre que soit démontré, par marquage isotopique, qu'il y avait incorporation d'unités acétate non seulement dans la biosynthèse du cholestérol mais aussi dans celle du triterpène lanostérol (C_{30}). La distribution du marquage déterminée par dégradation ne correspond cependant pas à celle qui aurait été obtenue selon le pliage proposé par Robinson. Selon ce pliage, le carbone 13 du cholestérol proviendrait du carbonyle de l'acétate. Or ce carbone provient du groupement méthyle de l'acétate comme déjà vu. Un peu plus tard, Woodward proposait le pliage qui s'est avéré être le bon comme illustré à la figure II.7.1.

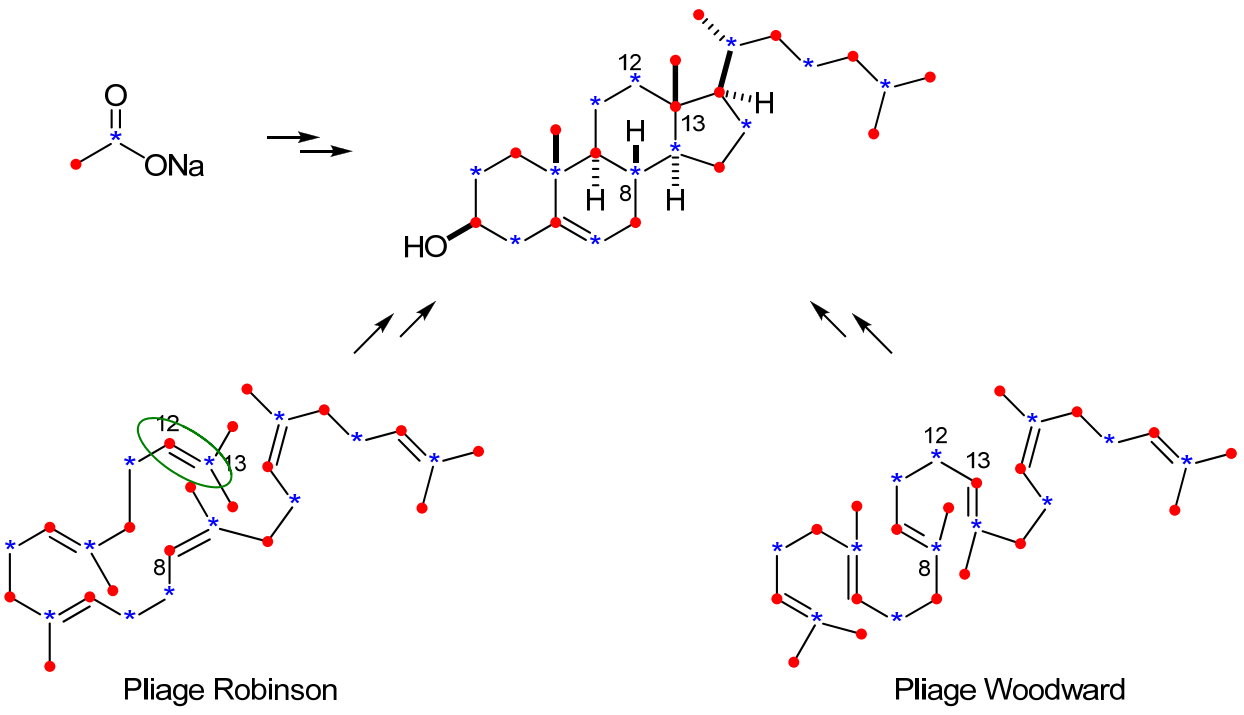


Figure II.7.1

Après la découverte de l'acide mévalonique comme intermédiaire de la biosynthèse du cholestérol entre l'acétate et le squalène, l'identification des autres intermédiaires s'est faite rapidement jusqu'au pyrophosphate de farnésyle (FPP). Les intermédiaires intervenant dans la transformation du FPP en squalène ont été identifiés plus tard.

II.8.2 Transformation du FPP en squalène

Les expériences de marquage isotopique ont révélé: 1) que deux molécules de FPP se condensent; 2) que le NAD(P)H est nécessaire pour cette condensation et que l'atome H du NAD(P)H est incorporé dans le squalène; 3) qu'en l'absence de NADPH, un intermédiaire ayant un noyau cyclopropane, le pyrophosphate de présqualène, s'accumule; 4) qu'une seule des deux molécules de FPP perd un proton; et 5) qu'il y a inversion de configuration au carbone 1 de l'autre molécule de FPP. Les transformations du schéma II.7.1 sur la page suivante rendent bien compte de toutes ces observations expérimentales. La condensation implique d'abord l'attaque de la double liaison $C_{2'}$=$C_{3'}$ de l'un des FPP sur le carbone 1 (C_1) de l'autre (mécanisme S_N1 ou S_N2) pour former un carbocation non classique, lequel conduit au dérivé cyclopropane, le pyrophosphate de présqualène, par élimination de l'un des protons de C_1. Il a été établi que seul le proton qui était pro S dans le FPP est éliminé. Ensuite, il y a départ de l'anion pyrophosphate sur le $C_{1'}$ assisté par (ou suivi de) la migration [1,2] du C_1 pour former un carbocation cyclobutyle. L'ouverture de ce dernier donne un carbocation allylique avec formation d'une double liaison de configuration E. Le carbocation allylique est finalement réduit par NAD(P)**H**, avec une stéréosélectivité de 100%, pour donner le squalène qui correspond formellement à un couplage queue à queue de deux FPP.

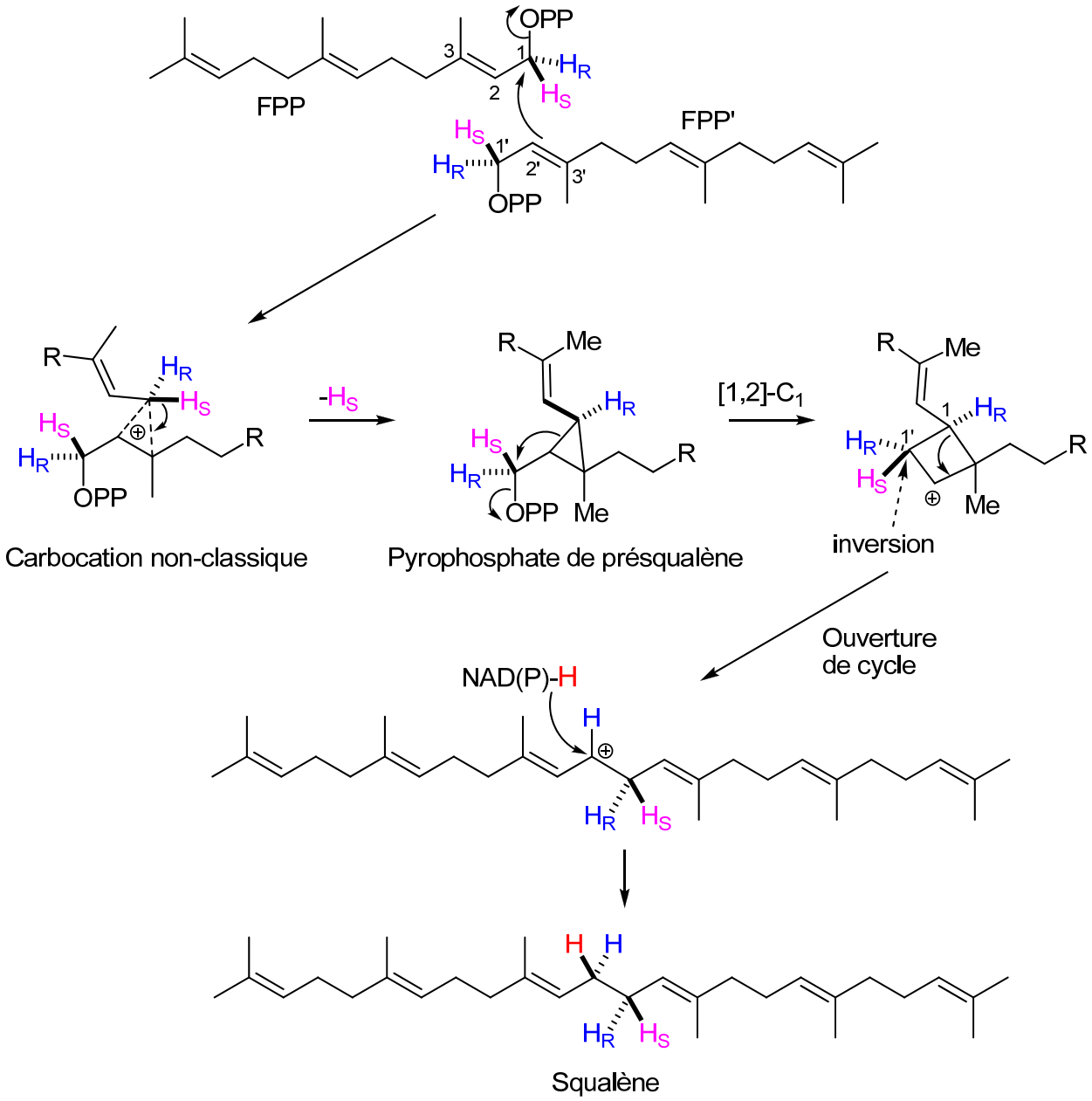


Schéma II.7.1

II.8.3 Biosynthèse des triterpènes

II.8.3a Cyclisation de l'époxyde de squalène

Les triterpènes sont divisés en deux classes principales : les triterpènes tétracycliques et les triterpènes pentacycliques. Ils sont tous formés à partir du squalène via une séquence oxidation-cyclisation-réarrangements sigmatropiques [1,2] (migration d'hydrure et de méthyle). Les

premières évidences du mécanisme de cyclisation du squalène sont venues de l'isolement de deux types d'enzymes impliquées dans la conversion du squalène en triterpènes et de l'isolement et caractérisation d'un intermédiaire dans la biosynthèse, l'époxyde de squalène. Les deux types d'enzymes isolées sont : 1) une mono-oxygénase qui oxyde la double liaison C_2=C_3 du squalène en époxyde pour former l'époxyde chiral correspondant, l'époxyde de squalène, ayant la configuration absolue *S* à la position 3 (Schéma II.8.1); 2) deux cyclases impliquées dans la cyclisation de l'époxyde de squalène, l'une oeuvrant dans les plantes (systèmes photosynthétiques) pour mener au cycloartenol, l'autre dans les animaux, les moisissures et les bactéries (systèmes non-photosynthétiques) qui mène au lanostérol. Si l'organisme vivant est placé dans des conditions anaérobiques, l'époxyde de squalène s'accumule et il s'accumule aussi si les microsomes (où a lieu sa cyclisation) sont désactivés par chauffage. Il faut noter que presque tous les triterpènes polycycliques ont un groupement hydroxyle orienté β en position 3.

Le premier processus de cyclisation de l'époxyde de squalène est décrit dans le schéma II.8.1 et la conformations de pliage sont contrôlées par la cyclase. La conformation chaise-bateau-chaise-enveloppe conduit au carbocation protostérol I (schéma II.8.1). La conformation chaise-chaise-bateau-enveloppe conduit, elle, au carbocation protostérol II (schéma II.8.2). Le carbocation protostérol I ne conduit qu'à des triterpènes tétracycliques tandis que le carbocation protostérol II est le précurseur de triterpènes aussi bien tétracycliques que pentacycliques. Aucun triterpène tétracyclique ayant le squelette du carbocation protostérol I n'a été isolé jusqu'à aujourd'hui. Ce n'est pas le cas pour le carbocation protostérol II. Le dammarènediol est un triterpène tétracyclique résultant de l'addition d'eau à ce carbocation (voir schéma II.8.2).

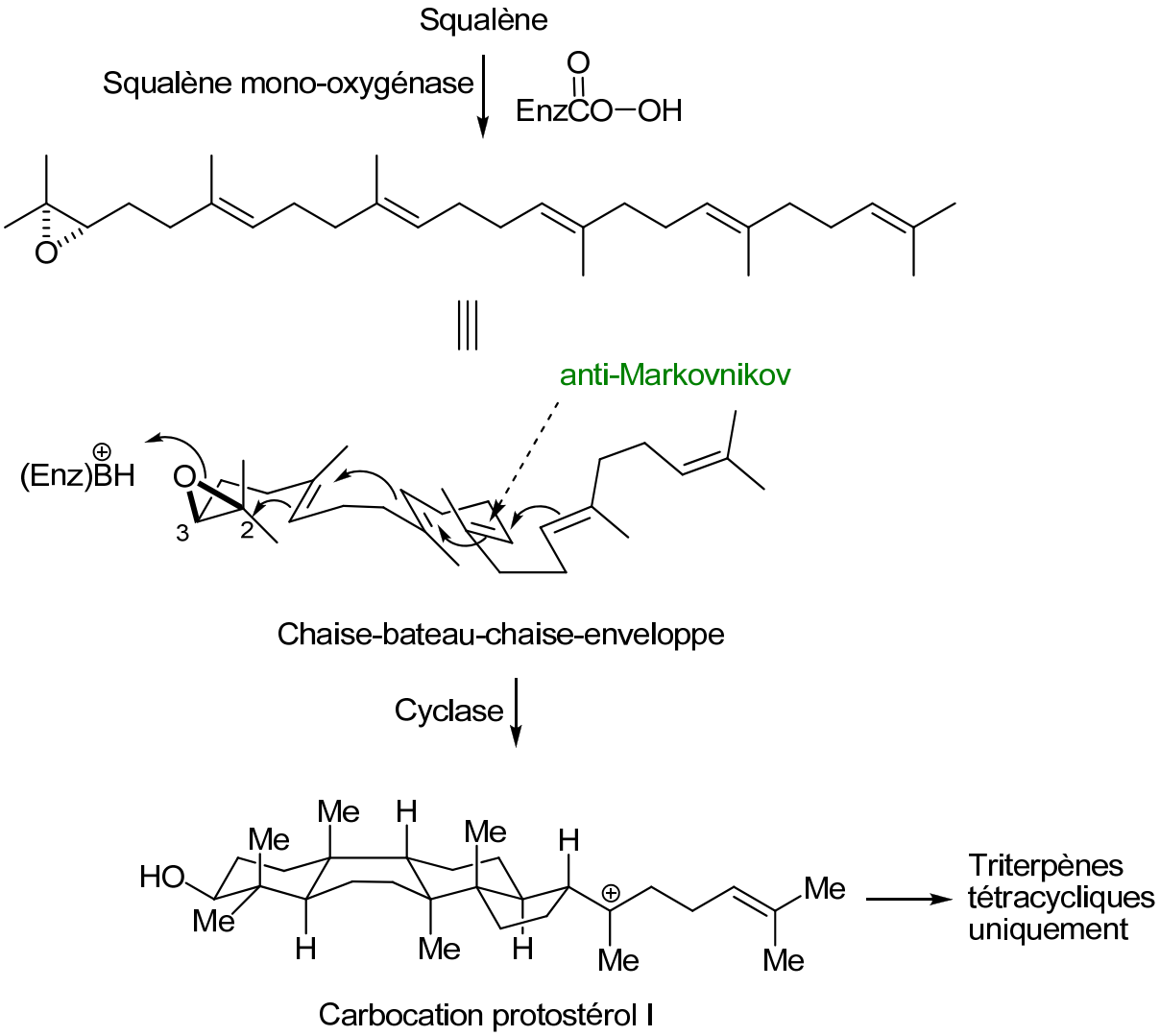

Squalène
Squalène mono-oxygénase
O
EnzCO—OH
anti-Markovnikov
(Enz)BH
O
3
2
Chaise-bateau-chaise-enveloppe
Cyclase
Me
H
Me
Me
H
HO
Me
H
Me
H
Me
Me
Triterpènes
tétracycliques
uniquement
Carbocation protostérol I

Schéma II.8.1

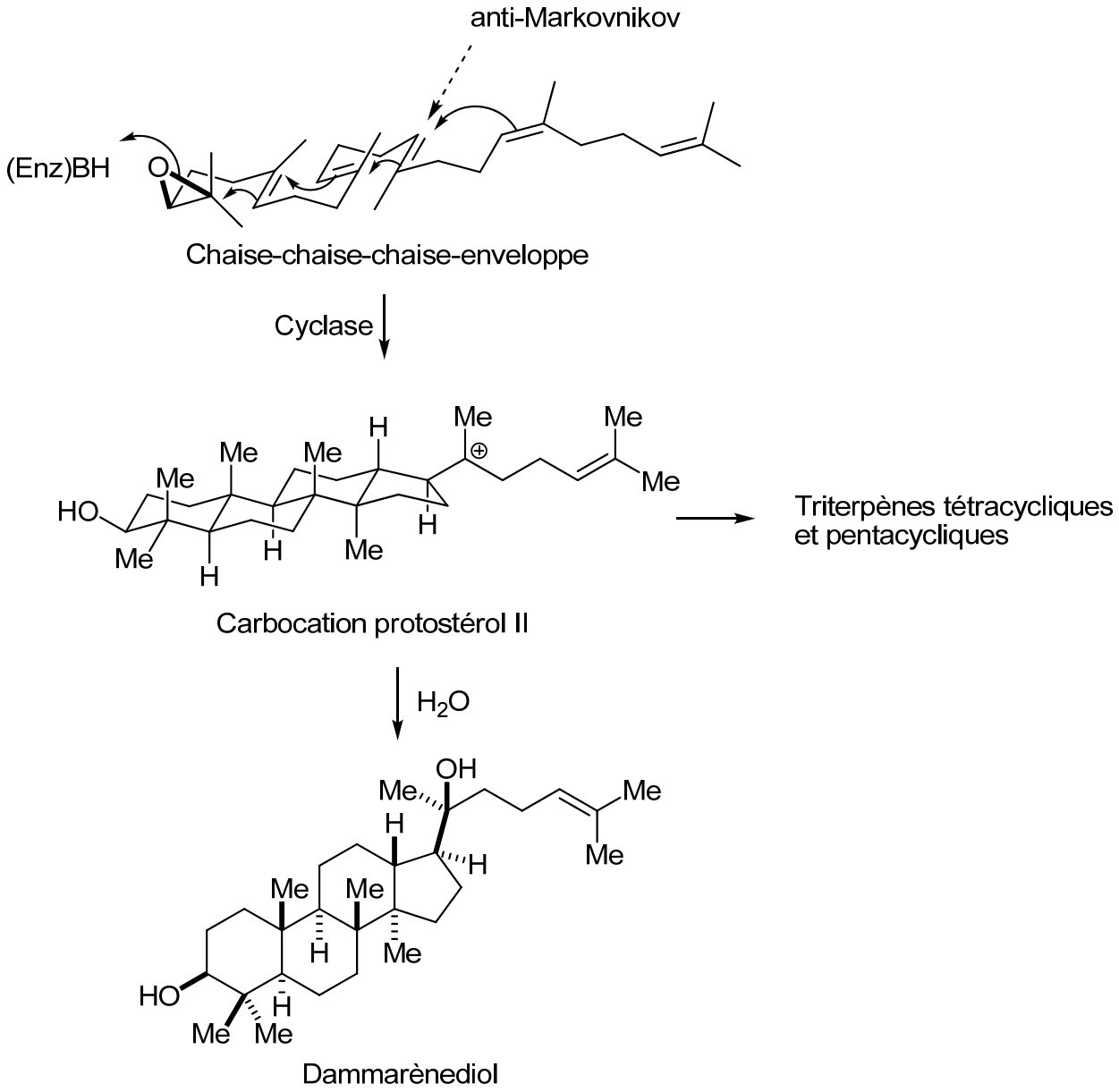


Schéma II.8.2

II.8.3b Conversion des carbocations protostérols I et II en triterpènes

Conversion du carbocation protostérol II. - Le schéma II.8.3 montre la conversion du carbocation protostérol I en lanostérol alors que le schéma II.8.4 montre la conversion du même carbocation en curcubitacine E et en cycloarténol. Le carbocation protostérol I est d'abord converti en un autre carbocation intermédiaire via une série de migrations sigmatropiques [1,2] d'hydrure et de carbone. L'élimination du proton en position 9 donne le lanostérol (schéma II.8.3). Le lanostérol est un triterpène très important puisqu'il est le précurseur d'une classe de

produits naturels de haute importance, les stéroïdes. Le lanostérol n'est produit que par les organismes n'utilisant pas la photosynthèse (non-photosynthétiques).

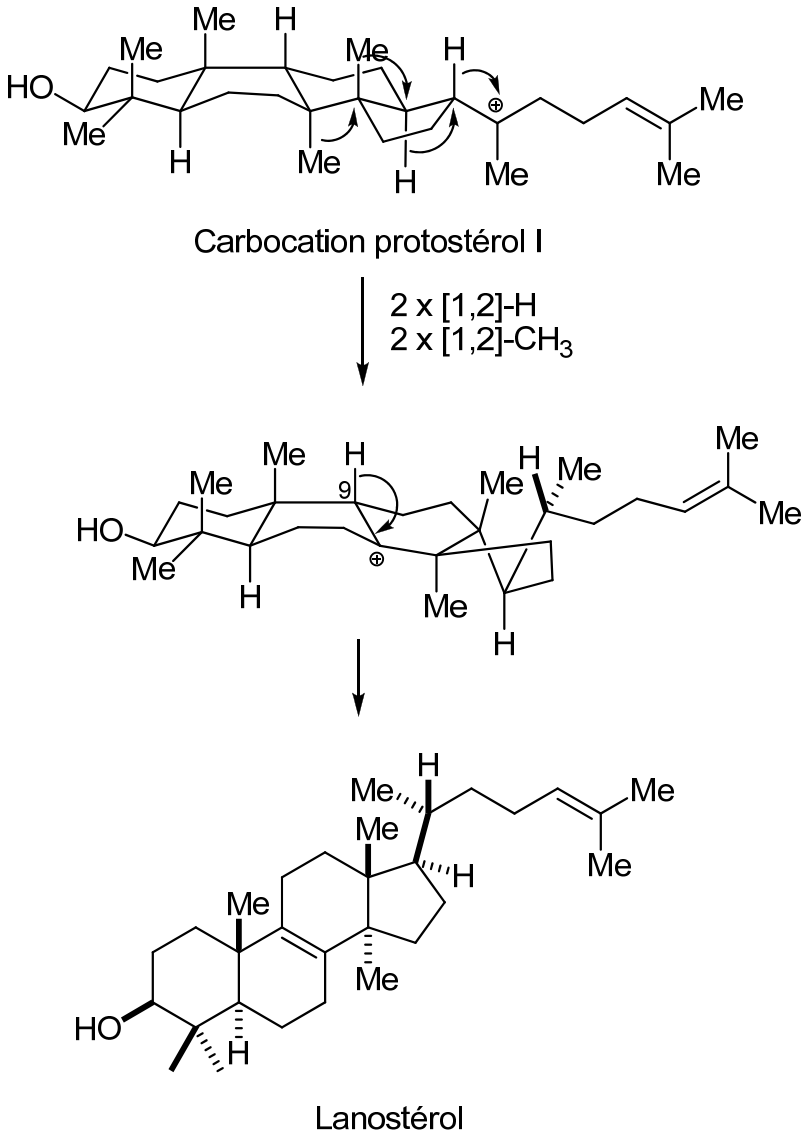


Schéma II.8.3

Dans les organismes photosynthétiques, il y a plutôt migration [1,2] de l'hydrogène en position 9. De ce troisième carbocation intermédiaire sont fabriqués selon l'organisme : 1) le squelette cucurbitane par transferts [1,2] successifs de méthyle et d'hydrure suivis de l'élimination du proton en position 5; ou 2) le cycloarténol par élimination 1,3 de l'un des protons du groupe méthyle (ou par élimination de l'un de ces protons à partir d'un carbocation non classique résultant d'une migration [1,2]-CH_3). Le cycloartenol est précurseur de la plupart

des phytostérol (plantes) comme le lanostérol l'est pour les organismes non-photosynthétique. La réouverture du cyclopropane pour mener à d'autres phytostérols est montré plus loin au schéma II.8.8.

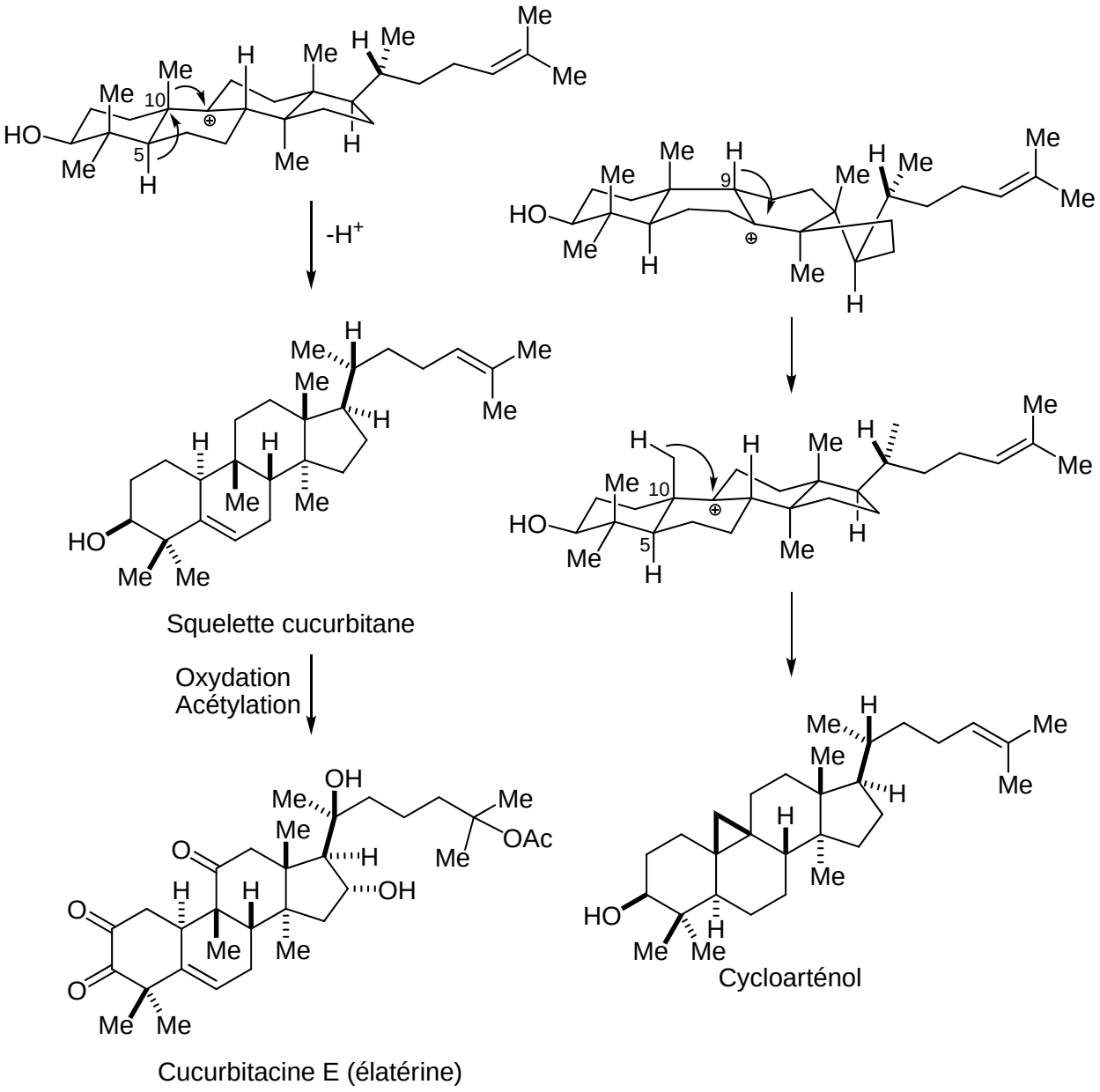


Schéma II.8.4

Conversion du carbocation protostérol II. – La conversion du carbocation protostérol II en triterpènes tétracycliques (euphol et apotirucallol) est décrite au schéma II.8.5. L'euphol et l'apotirucallol résultent encore ici de migrations sigmatropiques successives d'hydrure et de

groupement(s) méthyle(s) suivies d'une déprotonation. Dans le site actif des enzymes catalysant ces transformations du carbocation protostérol II, l'orientation de la chaîne latérale en position 17 (porteuse du carbocation en position 20) pour la conversion en euphol est différente de celle conduisant à l'apotirucallol. C'est deux triterpènes sont des intermédiaires biosynthétiques de triterpènes hautement dégradés (oxydation avec perte de carbone) faisant parti de la famille des limonènes et des quassinoïdes (voir plus loin). Ces deux familles comportent des membres hautement biologiquement actifs, en particulier dans l'arène virale et de l'oncologie.

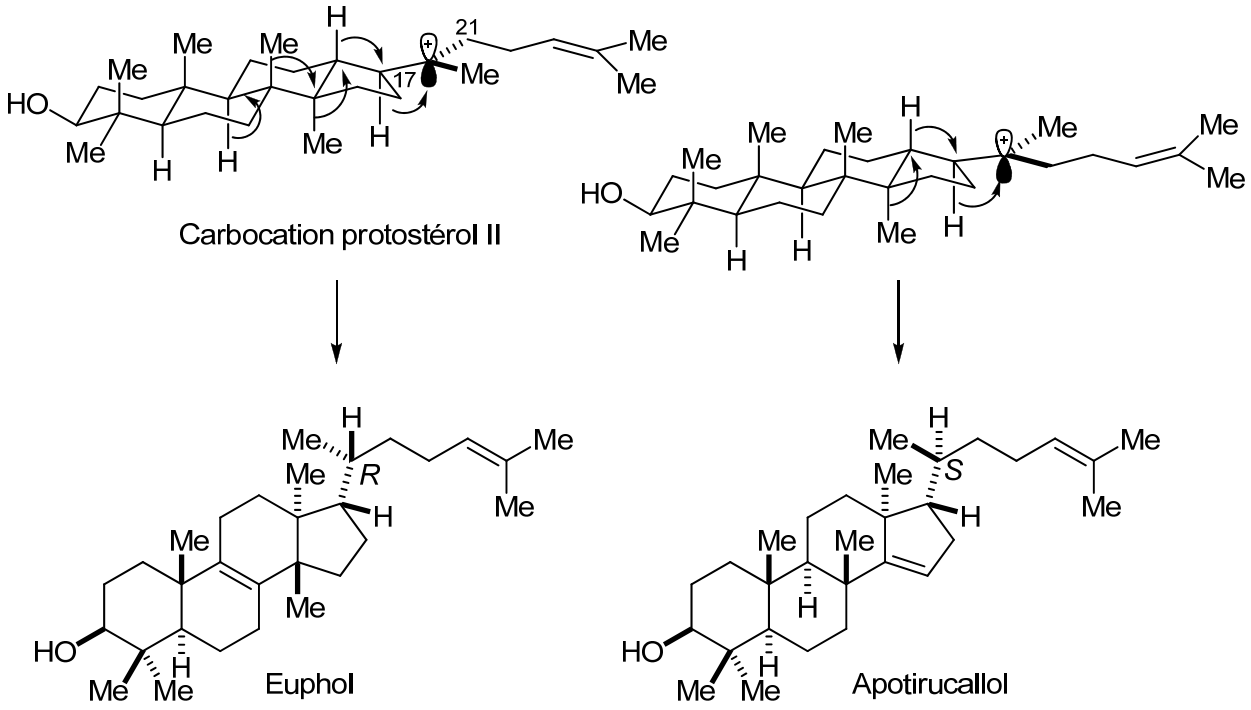


Schéma II.8.5

Pour la formation des triterpènes pentacycliques à partir du carbocation protostérol II (schéma II.8.6), il y a d'abord agrandissement du cycle D par réarrangement sigmatropique [1,2] puis addition Markovnikov du carbocation ainsi formé sur la double liaison de la chaîne latérale. La perte d'un proton mène à la formation du lupéole, un agent anti-cancer prometteur, inhibiteur de l'oncoprotéine *Ras* (pour rat sarcoma virus) impliquée dans la pathogénèse de tumeurs pancréatiques (figure II.8.1). L'acide bétulinique, qui provient du lupéol par oxydation du groupement méthyle, est un composés avec un large spectre d'activités biologiques, incluant l'anti-inflammation, des propriétés anti retro-virales et anti-malariales. On le retrouve dans plusieurs espèces d'arbres et d'arbustes, comme le boulot blanc et même la plante carnivore africaine *Triphyophyllum peltatum* (figure II.8.1). L'acide bétulinique induit aussi l'apoptose

dans les cellules du carcinome humain et est donc un agent potentiel pour le traitement du cancer de la peau.

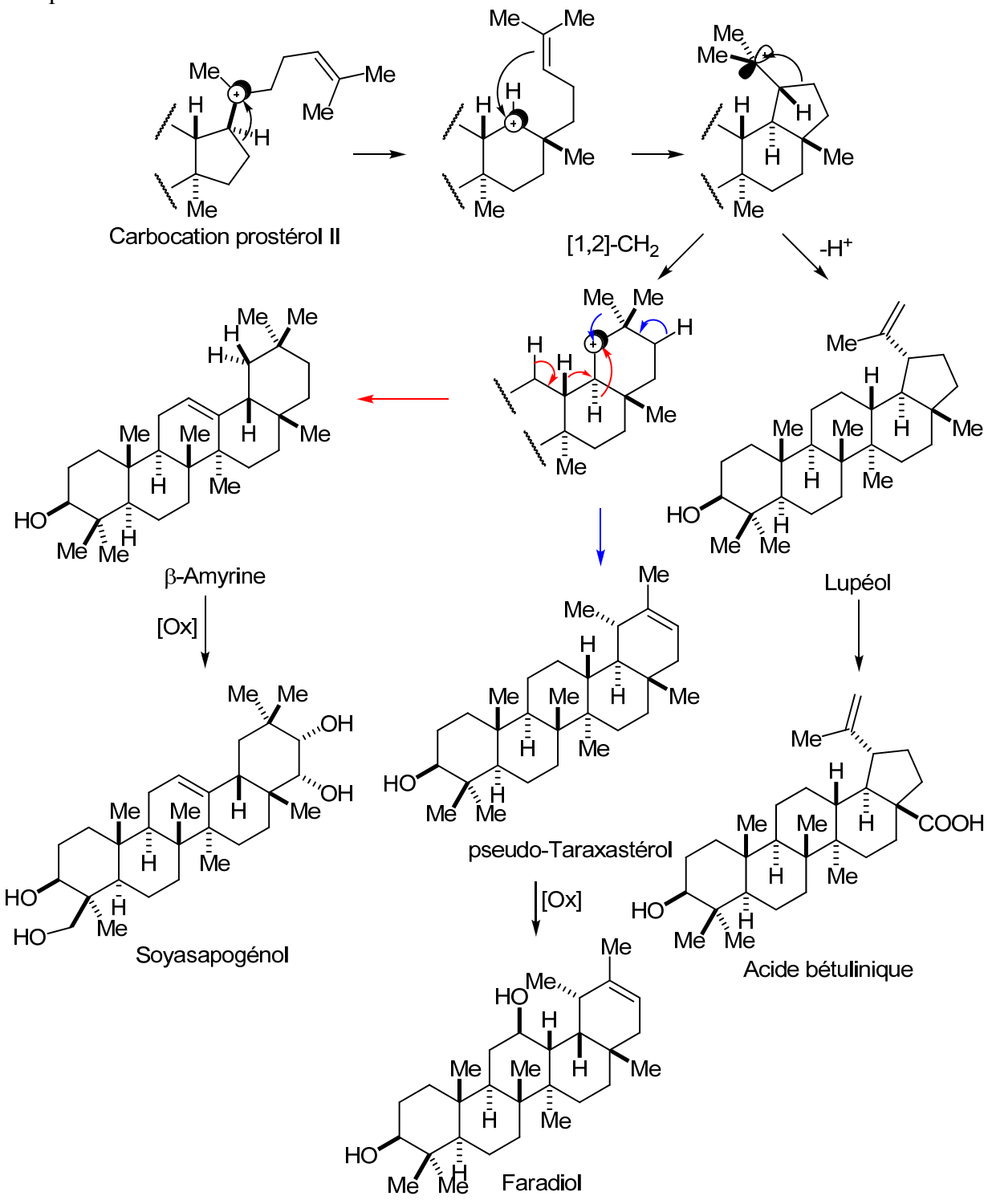
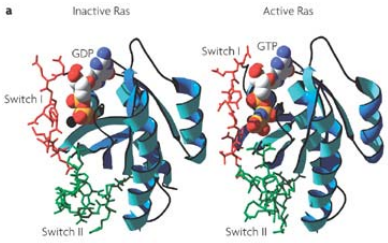
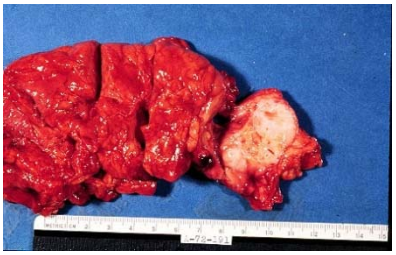


Schéma II.8.6





Oncoprotéine *Ras* Pancréas cancéreux *Triphyophyllum peltatum*

Figure II.8.1

Un autre agrandissement de cycle, le cycle E cette fois, causé par une migration-[1,2] d'un carbone, produit un intermédiaire qui subit la perte la migration-[1,2] d'un méthyle et la perte d'un proton pour mener au pseudotaraxastérol, une composante de l'arbuste *Artemisia monosperma* qui pousse sur le long des plages et des banquises de sable (figure II.8.2). La petite fleur jaune ou orange appelé *Souci officinal* (marigold) (figure II.8.2) produit le faradiol par oxidation du pseudo-taraxastérol (schéma II.8.6). Le faradiol possède des propriétés anti-inflammatoires.

Artémisia monosperma *Souci officinal*

Figure II.8.2

Une série de migration-[1,2] d'hydrogène terminé par la perte d'un proton produit la β-amyrine et, par oxydation de celle-ci, le soyasapogénol, une triterpène de la famille des saponines qu'on retrouve dans le soya (figure II.8.3).

Figure II.8.3

II.8.3c Biosynthèse du fern-9-ène

Le fern-9-ène est un triterpène pentacyclique isolé des fougères communes (voir figure II.8.4) qui n'est pas formé de la cyclisation de l'oxyde de squalène mais de la cyclisation du squalène lui même dans la conformation chaise-chaise-chaise-chaise-enveloppe initiée par protonation de la double liaison C_2=C_3 comme le montre le schéma II.8.7. La transformation du carbocation intermédiaire en fern-9-ène implique là encore une succession de transferts [1,2] d'hydrure et de méthyle et se termine par une déprotonation. La fougère contient plusieurs fernènes qui sont peut-être et en partie responsable de son action astringente et expectorante. On utilisait des infusions de fougère pour contrer la toux. On peut encore aujourd'hui utiliser des infusions de cette plante contre la toux légère.

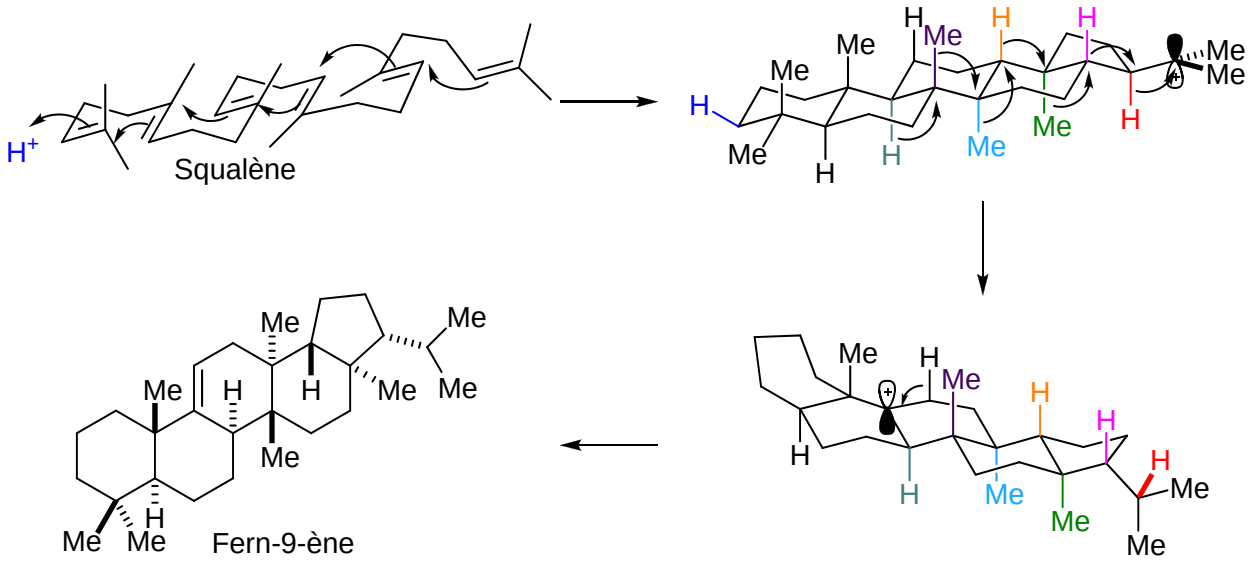


Schéma II.8.7

Figure II.8.4

II.8.3d Modifications subséquentes (secondaires) des triterpènes

Après la construction des squelettes triterpéniques via la cyclisation de l'oxyde de squalène (ou du squalène) et une série de réarrangements Wagner-Meerwein (migrations [1,2]-H et [1,2]-C), ces squelettes peuvent être modifiés par des oxydations (suivies ou non de réductions), par des réductions de liaisons doubles, par l'introduction de liaisons doubles (oxydation en alcool et déshydratation), par des déméthylations (par oxydation en acide et décarboxylation entre autres), des méthylations (le groupement méthyle provient la *S*-adénosylméthionine (SAM)) et/ou par coupure oxydante de liaisons carbone-carbone. Il peut y avoir des modifications par oxydation (et autre) sans que le squelette initial soit touché comme dans la conversion de la β-amyrine en soyasapogénol et du taraxastérol en faradiol où un seul site est oxydé (schéma II.8.6) ou comme dans la conversion du lupéol en acide bétulonique où deux sites sont oxydés (schéma II.8.6) ou encore comme dans la conversion du squelette cucurbitane en cucurbitacine E où plusieurs sites sont oxydés (schéma II.7.4). La figure II.8.5 montre quelques exemples d'une dégradation importante du squelette initial accompagnée de l'oxydation de plusieurs sites. Dans le cas de la limonine (les limonines sont isolées des agrumes) et de l'azadirone, la chaîne latérale a été convertie en noyau furane avec perte de quatre carbones (passage de C_{30} à C_{26}). Le squelette de la quassine et celui de la glaucarubinone n'ont plus que 20 carbones au lieu de 30 et plusieurs sites ont été oxydés. La quassine est un membre de la famille des quassinoïdes, une famille

importante du point de vue médicinal dû au large spectre de leur activité biologique allant de antivirale à néoplastique.

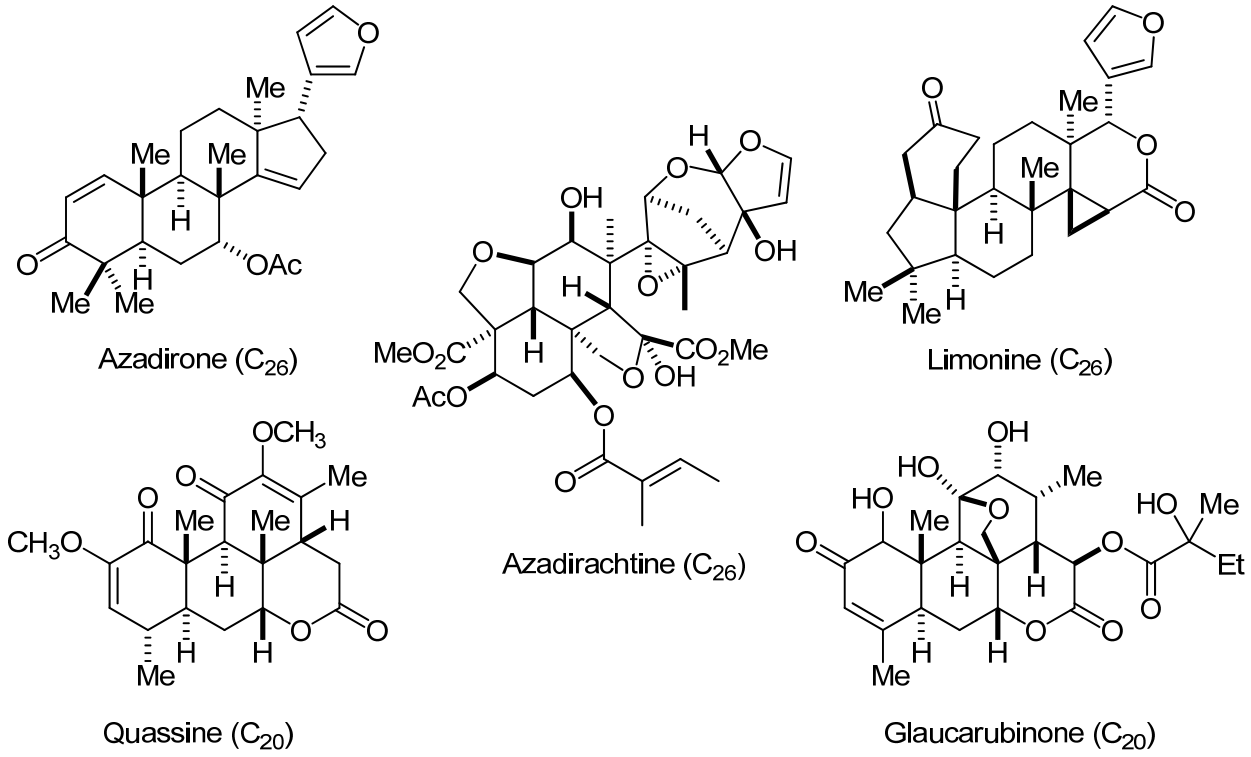
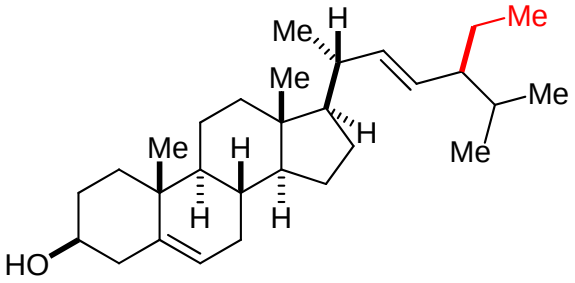


Figure II.8.5

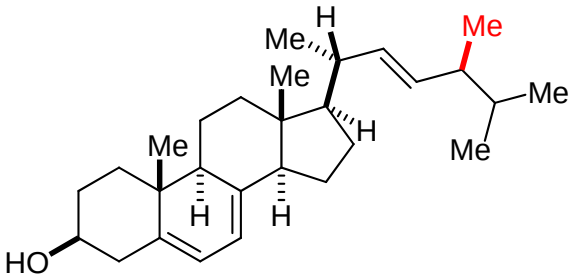
II.8.4 Biosynthèse des stéroïdes

II.8.4a Considérations générales sur les stéroïdes

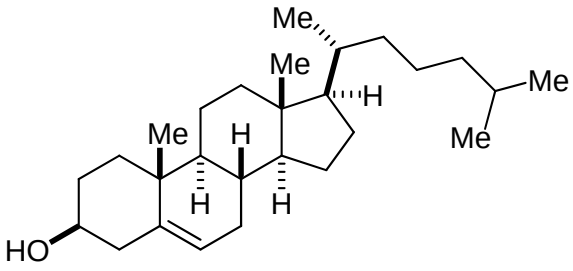
Les stéroïdes constituent une grande famille de composés largement répandus dans la nature, aussi bien dans les mondes animal et végétal que dans le monde des insectes et des microorganismes. Les stéroïdes d'origine animale sont des zoostérols (comme le cholestérol, figure II.8.6), ceux d'origine végétale des phytostérols (comme le stigmastérol) et ceux provenant de moisissures et levures des mycostérols (comme l'ergostérol). Les insectes font la biosynthèse des ecdysones classifiées sous « phytostérols d'insecte » parce qu'ils les fabriquent à partir de phytostérols qu'ils convertissent d'abord en cholestérol.



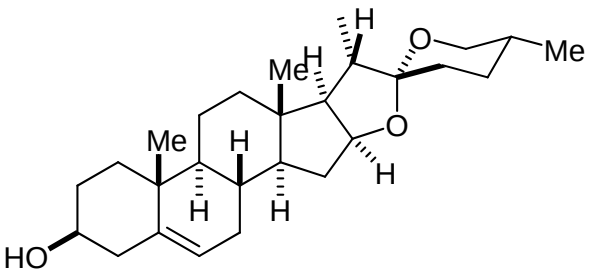
Stigmastérol (phytostérol)
$C_{27} + C_2$



Ergostérol (mycostérol)
précurseur de la vitamine D
$C_{27} + C_1$



Cholestérol (zoostérol)
C_{27}



Diosgénine (sapogénine)
C_{27}

Figure II.8.6

Les stéroïdes sont divisés en six grandes catégories basées surtout sur la nature de la chaîne latérale (figures II.8.6 à II.8.10) : les ***stérols***, par exemple le stigmastérol, un phytostérol, le cholestérol, un zoostérol, l'ergostérol qui est obtenu principalement de levures et est le précurseur de la vitamine D et les ecdysones (hormones de la mue des insectes); les ***sapogénines*** comme la diosgénine; les ***aglycones des glycosides cardiaques*** (cardiotoniques) tels la digitoxigénine, l'aglycone de la digitoxine qui est un cardénolide, et la proscillaridine A, l'aglycone du scyllarène A qui est un bufadiénolide (les bufadiénolides sont moins répandus ou abondants que les cardénolides); les ***acides biliaires*** tels les acides cholique et tauro-cholique; les ***corticostéroïdes***, par exemple l'aldostérone et la cortisone (hormones des surrénales (« adrenal hormones »)); et les ***hormones sexuelles***, le β-oestradiol et la progestérone, hormones sexuelles de la femme, et la testostérone et l'androstérone, hormones sexuelles de l'homme. Les hormones sexuelles sont biosynthétisées dans les organes sexuels.

La plupart des phytostérols trouvent leur origine dans la transformation du cycloarténol (section II.8.3). L'ouverture du cyclopropane se fait par protonation pour conduire à un carbocation (Schéma II.8.8). Le cyclopropane est un groupement fonctionnel tendu dont la

chimie ressemble beaucoup à celle des doubles liaisons. Sa protonation est donc facile. La biosynthèse de l'obtusifoliol au schéma II.8.8 donne un exemple de la conversion d'un dérivé du cycloarténol (nous verrons dans les sections qui suivent comment la nature enlève un groupement méthyle et alkyle la chaîne latéral des stéroïdes).

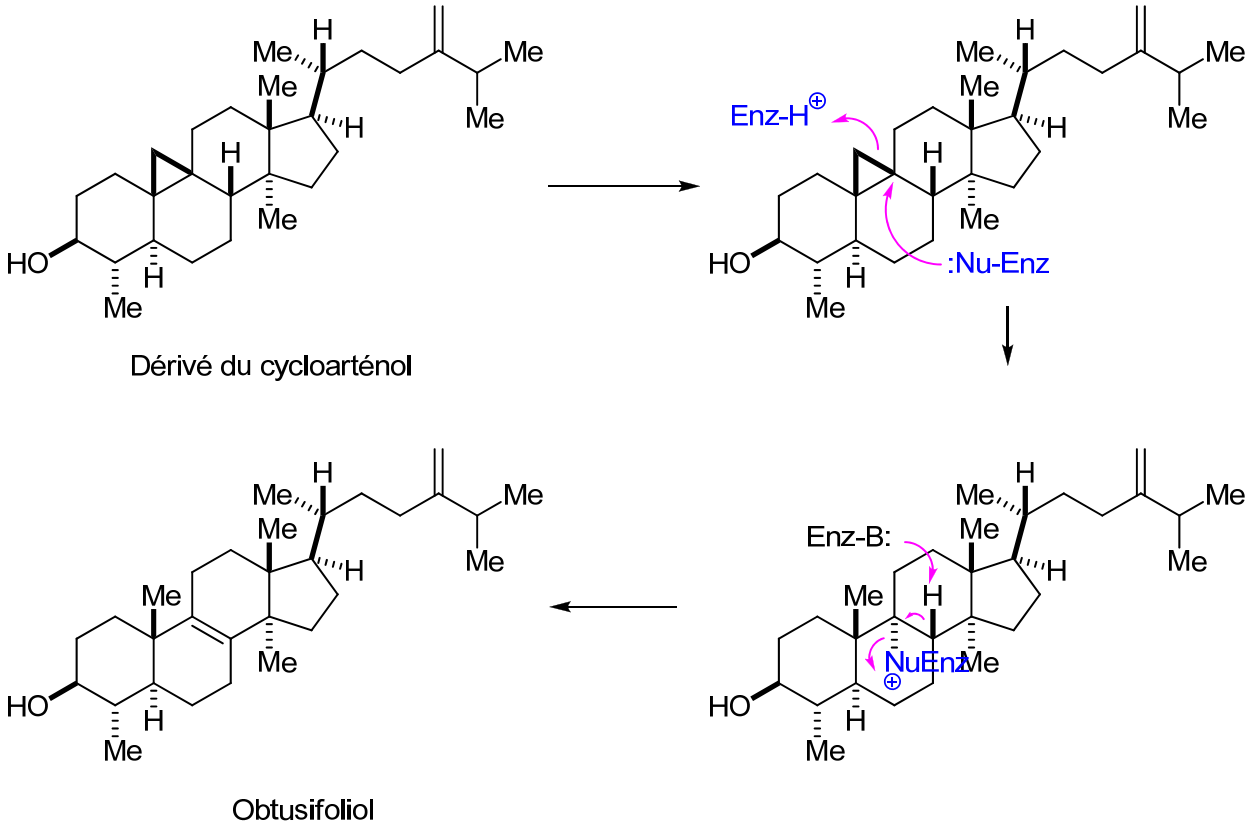


Schéma II.8.8

Le cholestérol est le stéroïde le plus universel du règne animal et se retrouve surtout dans les cellules du cerveau et des nerfs. Les pierres de la vésicule biliaire sont constituées de 66% à 99% de cholestérol. Il sert d'intermédiaire pour la biosynthèse de nombreux autres stéroïdes destinés à des fonctions aussi variées qu'hormones sexuelles, acides biliaires et agents cardiotoniques (figure II.8.7).

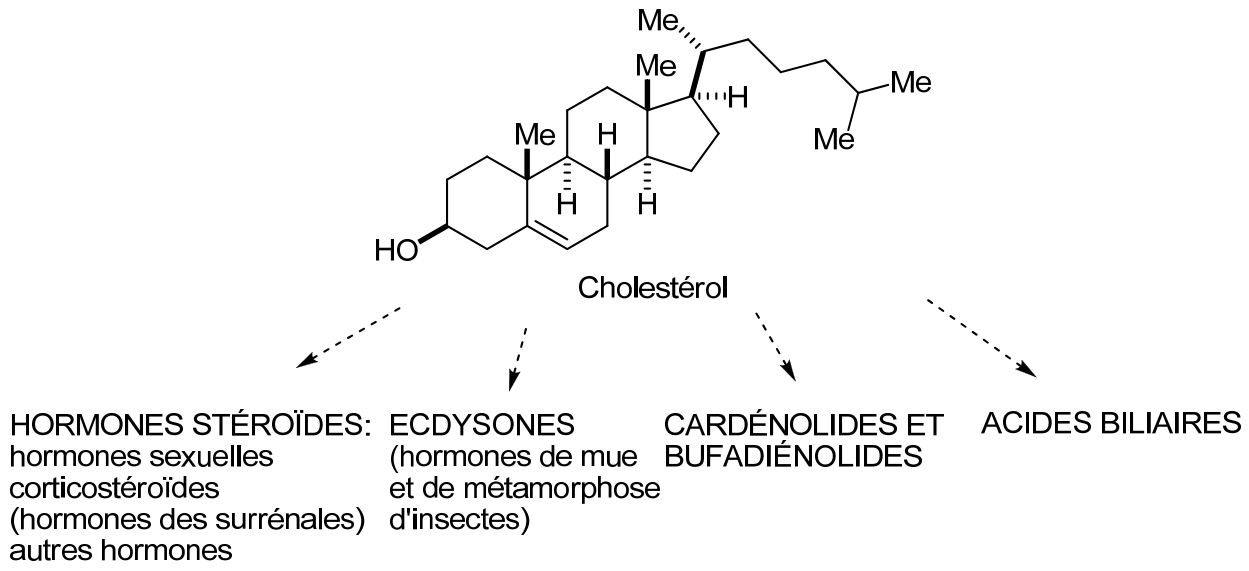


Figure II.8.7

Les hormones sexuelles féminines, comme la progestérone et le β-oestradiol, influencent la puberté et contrôlent le cycle menstruel (figure II.8.8). Les hormones sexuelles masculines, comme l'androstérone et la testostérone, influencent la puberté et sont impliquées dans la fabrication du sperme. Elles sont responsables des signes de virilité (barbe, mue de la voix, etc.) et sont des agents anaboliques.

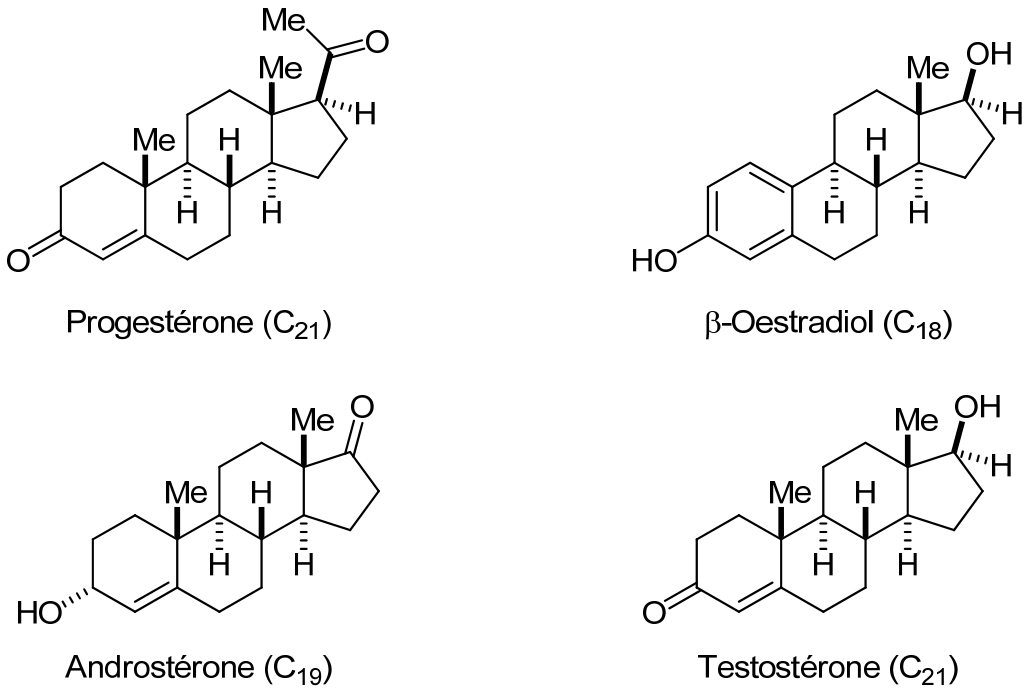
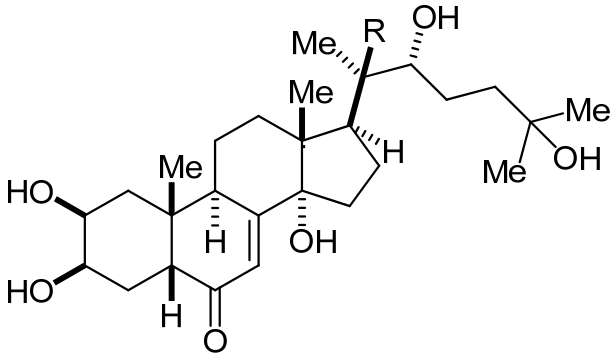
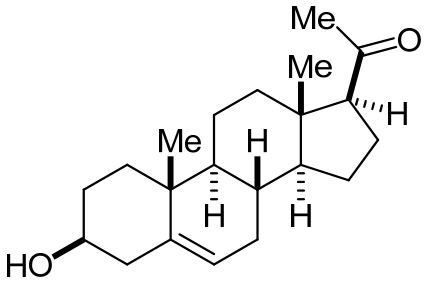


Figure II.8.8

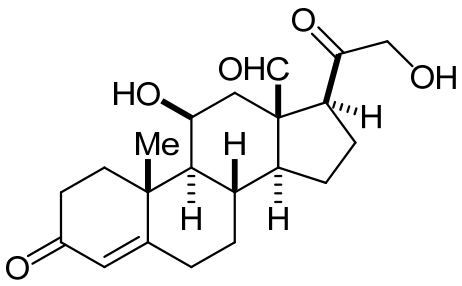
L'ecdysone et l'ecdystérone sont des hormones d'insectes et induisent la mue dans les larves ou interviennent dans la métamorphose (figure II.8.9). D'autres hormones sont présente à la fois chez les plantes et chez les animaux, comme la pregnénolone, et induisent la floraison dans les plantes ou sont des agents anaboliques chez les animaux. Les corticostéroïdes, comme l'aldostérone et la cortisone sont des hormones biosynthétisées dans les glandes surrénales et contrôlent le métabolisme des sucres, des protéines et des gras. De plus, elles régulent la rétention et l'excrétion du sel et de l'eau, ou peuvent être des agents anti-inflammatoires.



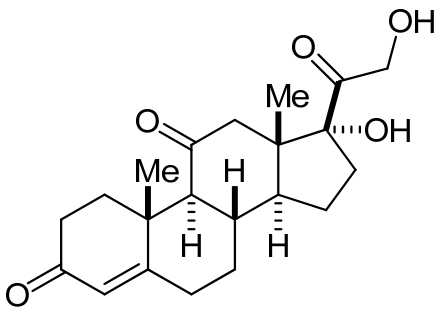
Ecdysone: R = H (C_{27})
Ecdystérone: R = OH (C_{27})



Pregnénolone (C_{21})



Aldostérone (C_{21})

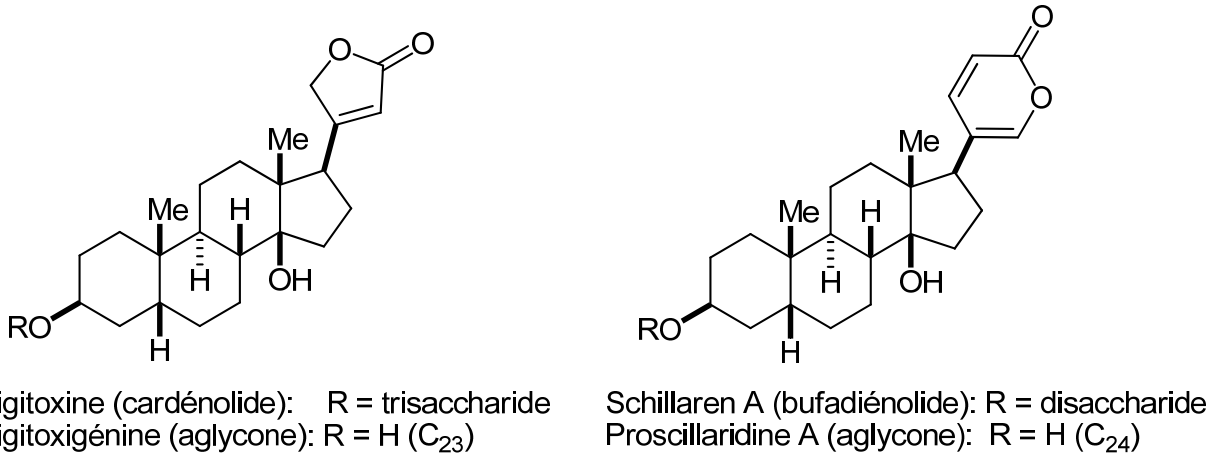


Cortisone (C_{21})

Figure II.8.9

Les stéroïdes cardio-actifs (cardénolides et bufadiénolides) sont des stimulants cardiaques (figure II.8.10). Ils peuvent être d'efficaces poisons induisant le vomissement et causant du dommage neurologique. Ils sont utilisés par les grenouilles ou les crapauds africains contre les prédateurs et par certaines plantes contre les animaux brouteurs. Les acides biliaires sont

biosynthétisés dans le foie et emmagasinés dans la vésicule biliaire. Ils favorisent la digestion en émulsifiant les gras en fines gouttelettes.



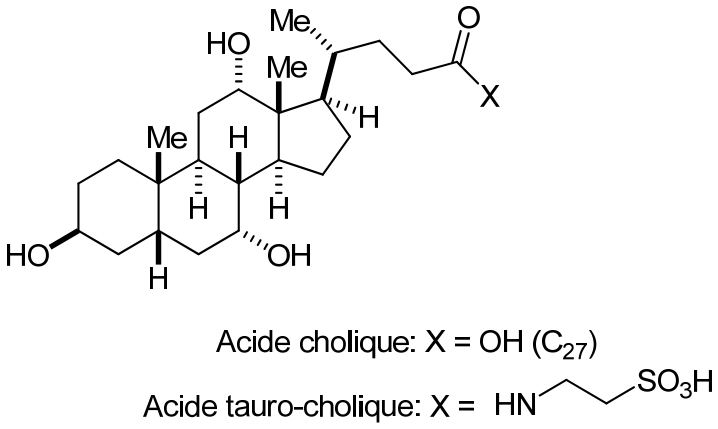


Figure II.8.10

Pratiquement tous les stéroïdes sont constitués de quatre cycles avec différentes chaînes latérales en position 17. La figure II.8.11 indique la numérotation conventionnelle des stéroïdes et le tableau II.8.1 donnent les noms standards de stéroïdes avec les particularités structurales qui les caractérisent ou les différencient. La figure II.8.8 illustre la terminologie utilisées pour décrire certaine modifications du squelette des stéroïdes : ***nor*** pour la perte d'un carbone; ***séco*** pour un cycle ouvert; ***homo*** pour un agrandissement de cycle; ***cyclo*** pour l'ajout d'un cycle de trois; ***anhydro*** pour la perte d'eau; ***déshydro*** pour la perte de d'hydrogènes; ***dihydro*** pour l'addition d'hydrogènes; ***désoxy*** pour la réduction d'un alcool en C−H; et ***désoxo*** pour la conversion d'un carbonyle en CH_2.

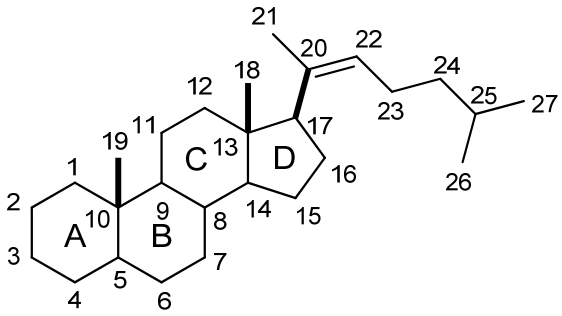
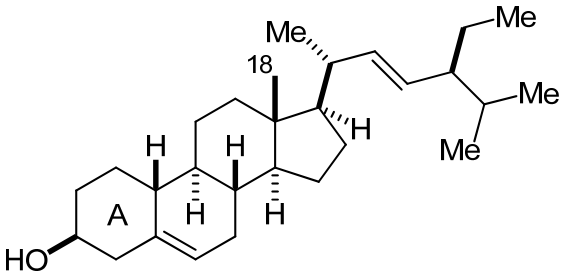


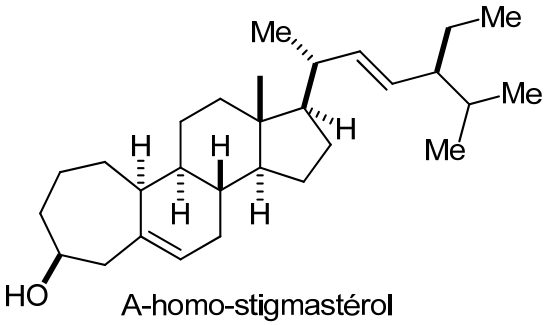
Figure II.8.11

Tableau II.8.1. Noms standards des squelettes stéroïdiens

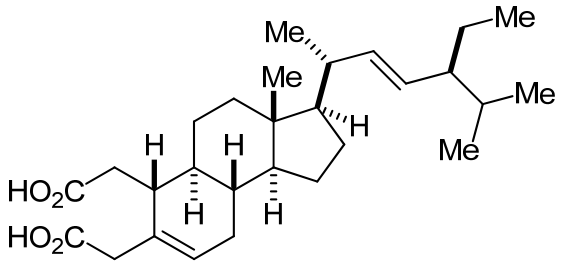
Nom standard du squelette :	Particularités structurales :	Exemples :
Cholestane	≥ C_{27} 5α-H ou pas de 5-H	Cholestérol, stigmastérol, ergostérol, diosgénine
Coprostane	≥ C_{27} 5β-H (fusion A/B *cis*)	Ecdysone, ecdystérone
Cholane	C_{24} 5β-H, chaîne latérale de C_5	Acides cholique et taurocholique
Pregnane	C_{21} pas de 5-H, chaîne latérale de C_2	Aldostérone, cortisone, pregnénolone, progestérone
Androstane	C_{19} pas de 5-H ni de chaîne latérale	Androstérone, testostérone,
Oestrane	C_{18} pas de 5-H ni de C-19 méthyle et pas de chaîne latérale	β-Oestradiol,



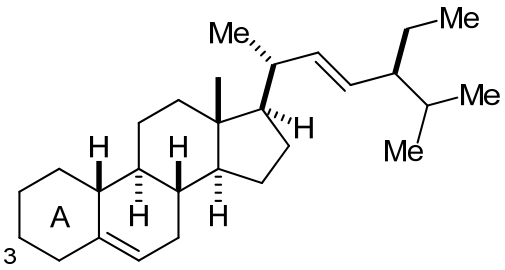
19-nor-stigmastérol
(sans carbone 19)



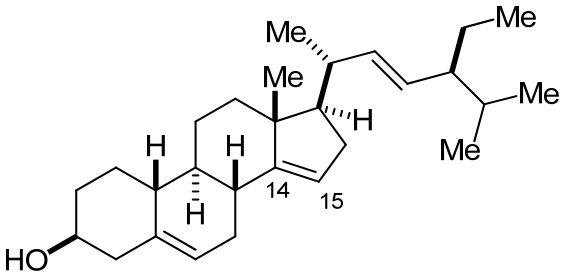
A-homo-stigmastérol



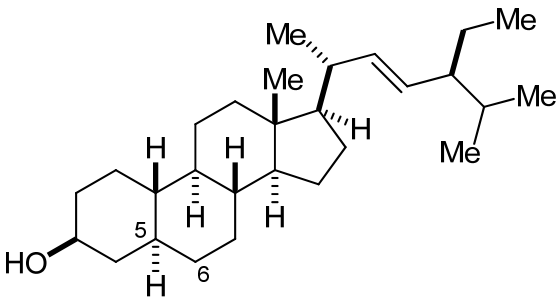
A-seco-stigmastérol



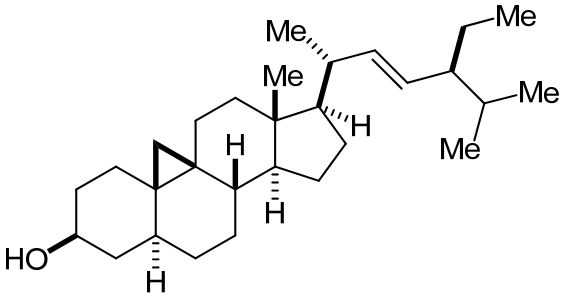
3-désoxy-stigmastérol



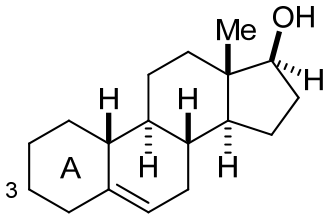
14,15-déshydro-stigmastérol



5,6-dihydro-stigmastérol



Cycloarténol



3-Désoxo-testostérone

Figure II.8.12

II.8.4b Conversion du lanostérol en cholestérol

Lors de la conversion biosynthétique du lanostérol (C_{30}) en cholestérol (C_{27}), trois groupes méthyles sont enlevés : le méthyle 30 (sur le carbone 14) et les deux méthyles 28 et 29 (sur le carbone 4). Ils sont tous les trois enlevés par oxydation. Les étapes de la première conversion sont présentées dans le schéma II.8.9. Une première oxydation par le Cyt Fe(II) + O_2 transforme le groupement méthyle 30 en groupement hydroxyméthyle (alcool primaire) et elle est suivie par une oxydation avec $NAD(P)^+$ pour donner l'aldéhyde. En fait, ce sont les même deux premières étapes pour les trois groupements méthyles (voir schémas II.8.10 et II.8.11). Une hypothèse pour l'enlèvement du carbone 30 est l'introduction d'un groupement OH en position 15 suivie de l'élimination de l'acide formique par addition-élimination d'eau. Ce mécanisme est en accord avec l'incorporation de ^{18}O dans l'acide formique lorsque la biosynthèse est effectuée sous atmosphère de $^{18}O_2$. Lors de l'oxydation de la position 15, c'est le H_α qui est oxydé selon les expériences de marquage au tritium.

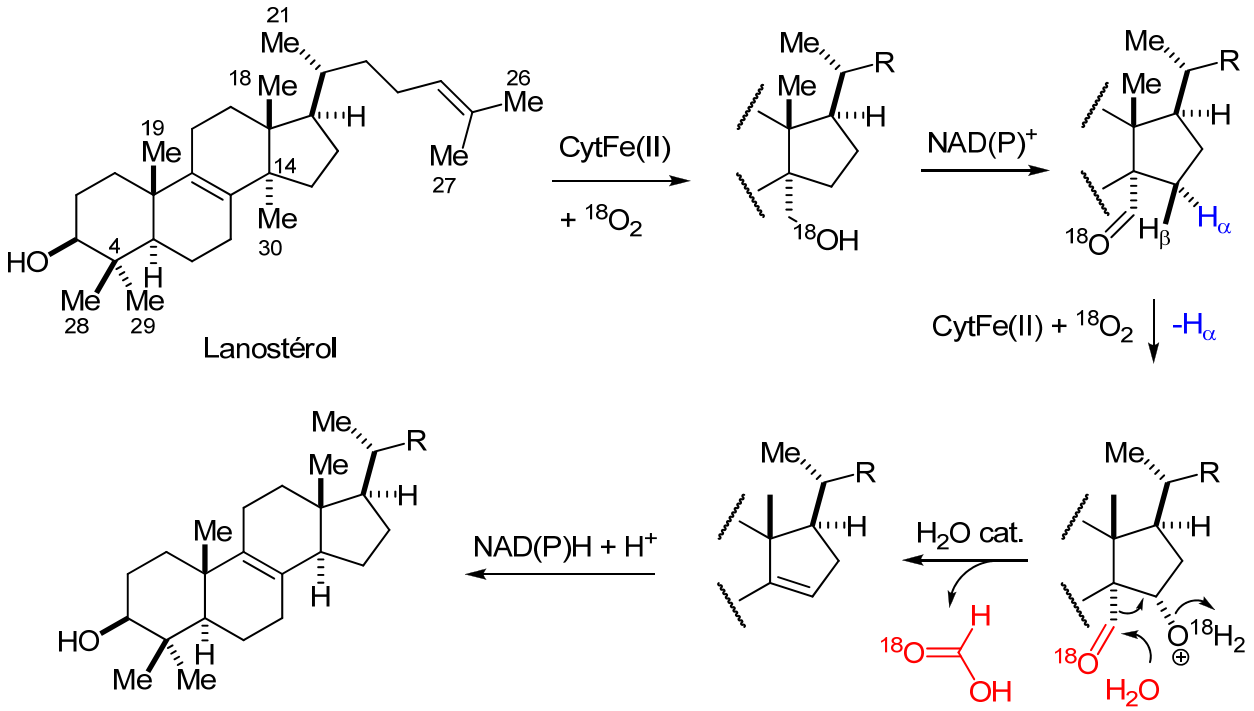


Schéma II.8.9

Notez qu'il existe un mécanisme d'oxydation alternatif par le CytFe(II) + O_2 et qui s'applique à l'oxydation d'alcool primaire en peroxyde d'hydrate d'aldéhyde. Ce mécanisme est montré au schéma II.8.10. Dans ce cas-ci, la déformylation se fait en même temps que

l'hydrogène en 15-α est enlevé sous forme de proton. Les deux mécanismes présentés sont conforment aux expériences de marquage isotopique à l'oxygène ou deutérium mentionnées ci-haut.

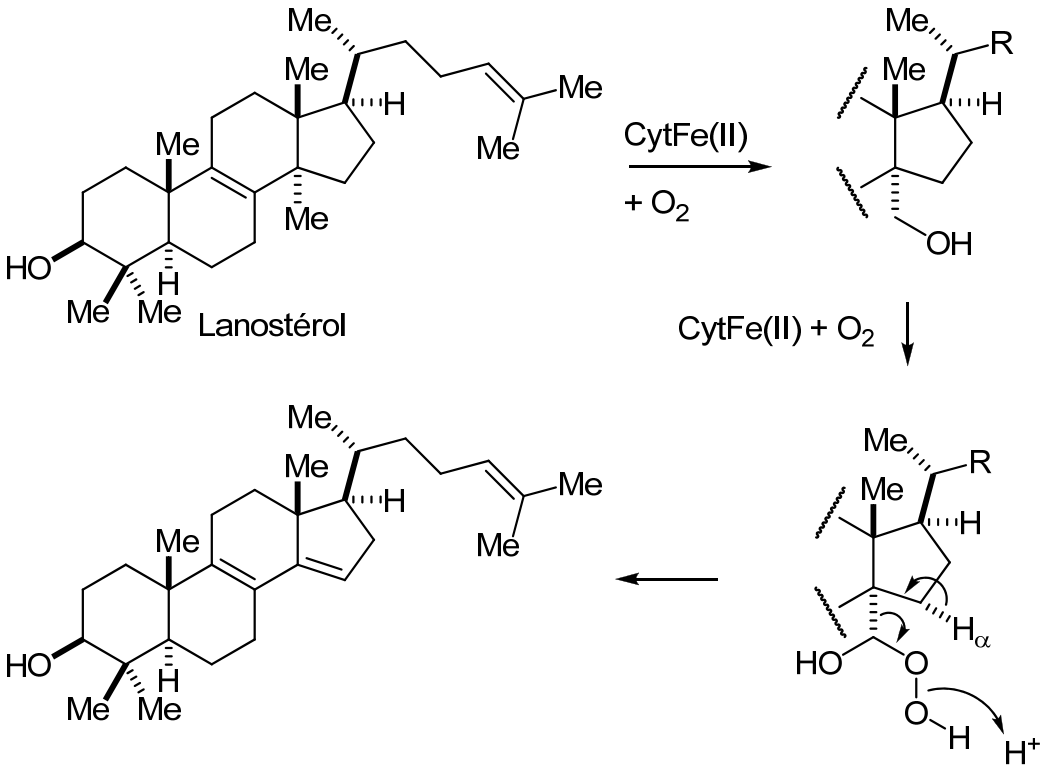


Schéma II.8.10

Pour les deux autres méthyles, la séquence commence par l'oxydation par le cytochrome du méthyle 29 (schéma II.8.11). Plutôt que de déformyler avec la perte de l'alcool en position 3, ce dernier est oxydé et cétone après que l'aldéhyde ait été oxydé en acide carboxylique. Il s'ensuit une décarboxylation facile puis le carbonyle est réduit de nouveau en alcool β. Notez la stéréochimie du méthyle 28 restant.

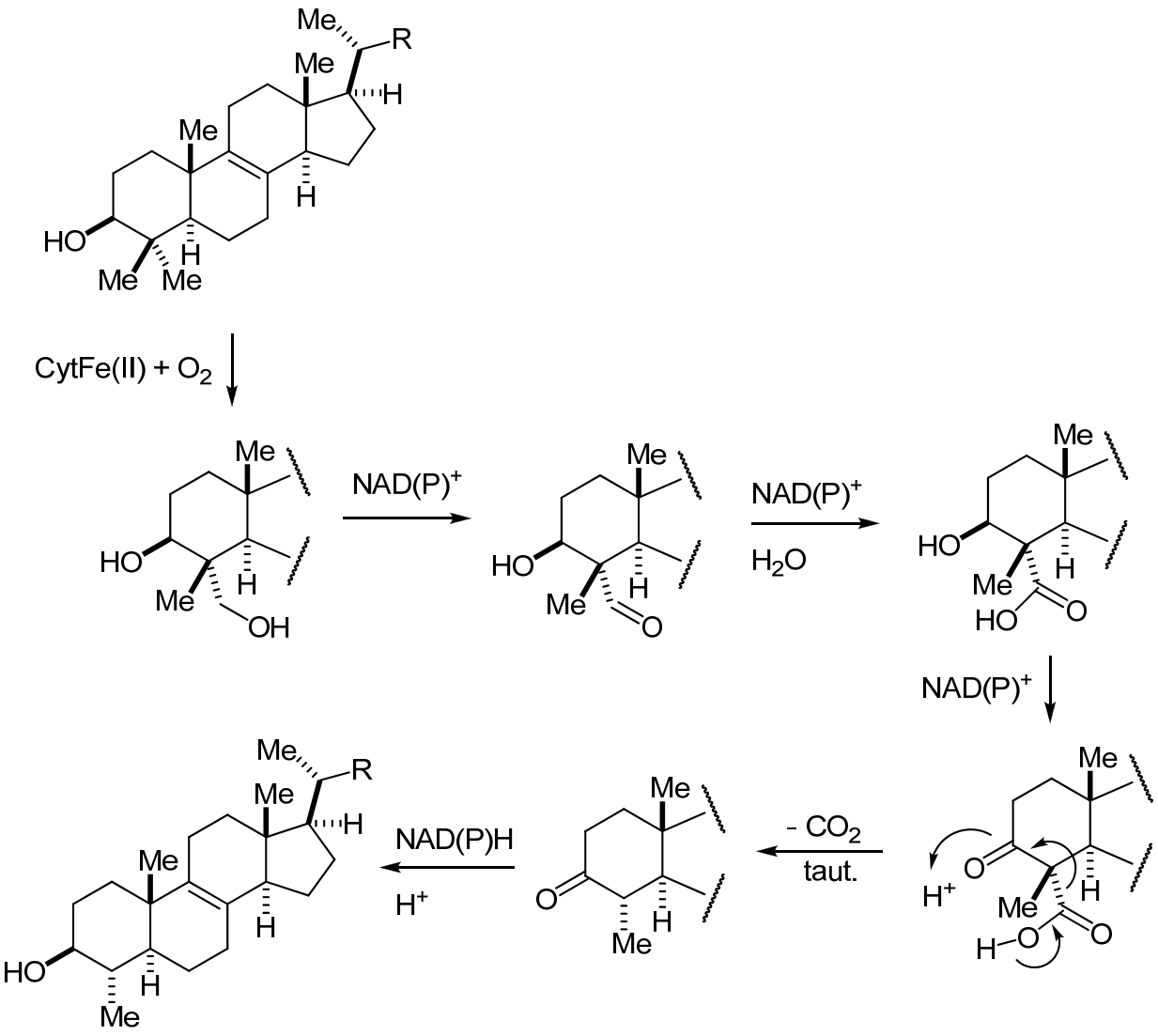


Schéma II.8.11

La même séquence se répète ensuite pour le méthyle 28. Enfin, la double liaison en C8-C9 est isomérisée en C7-C8 (schéma II.8.12). Une oxydation du carbone 6 suivit de l'élimination de l'alcool secondaire résultant introduit une double liaison en C5-C6. L'alcène en C7-C8 est réduit, ce qui établit la stéréochimie du carbone 8, et finalement la double liaison de la chaîne est réduite pour conduire au cholestérol.

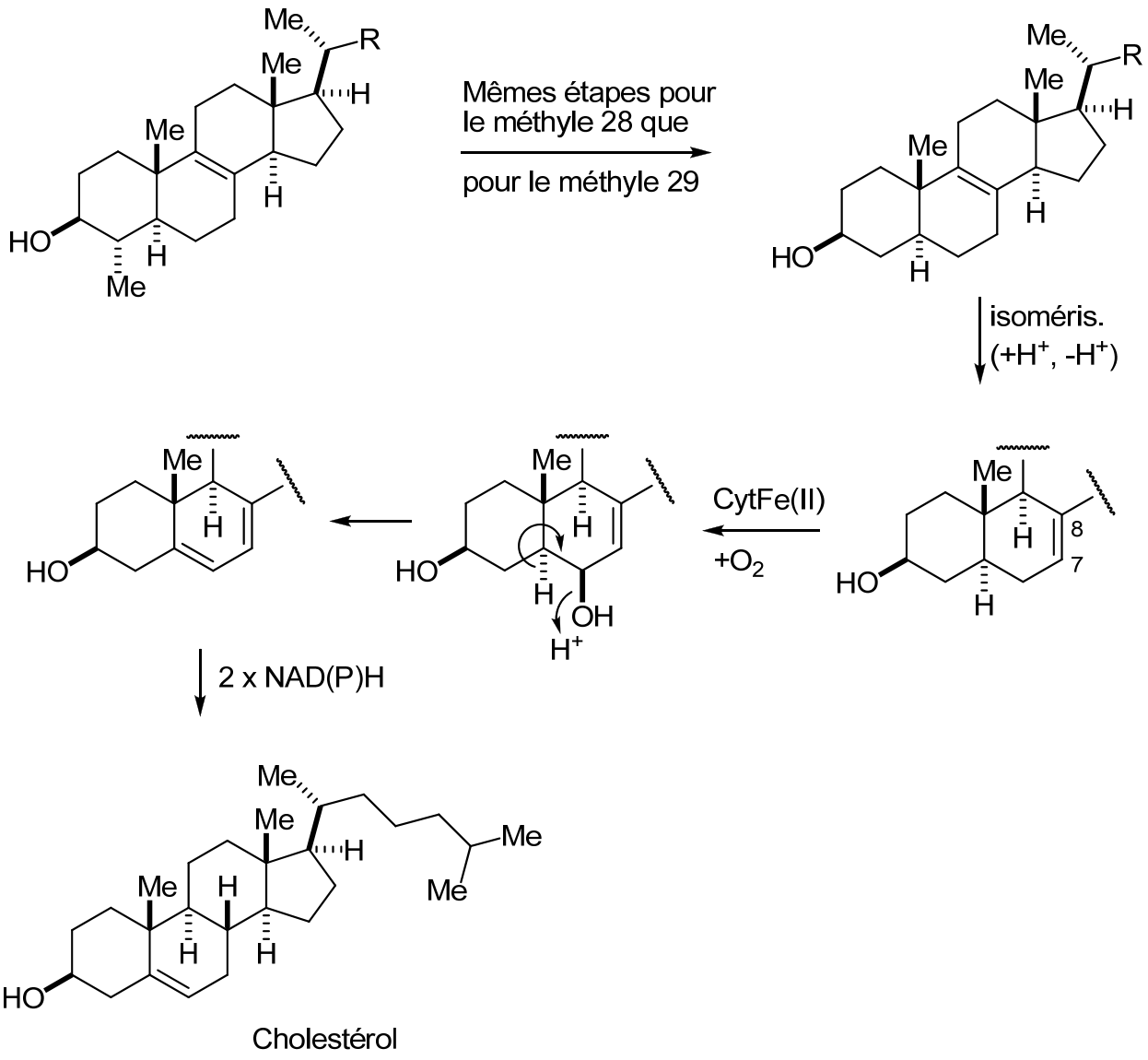


Schéma II.8.12

II.8.4c Méthylation et éthylation à la position 24 de la chaîne latérale des stéroïdes

L'agent de méthylation en biosynthèse est la *S*-adénosyle méthionine (SAM). Elle est formée par alkylation de la méthionine par l'ATP. Il s'agit d'une des quelques occasions ou l'ATP ne sert pas d'agent phosphorylant mais bien de source d'adénosyle. La structure de SAM est montrée au schéma II.8.13.

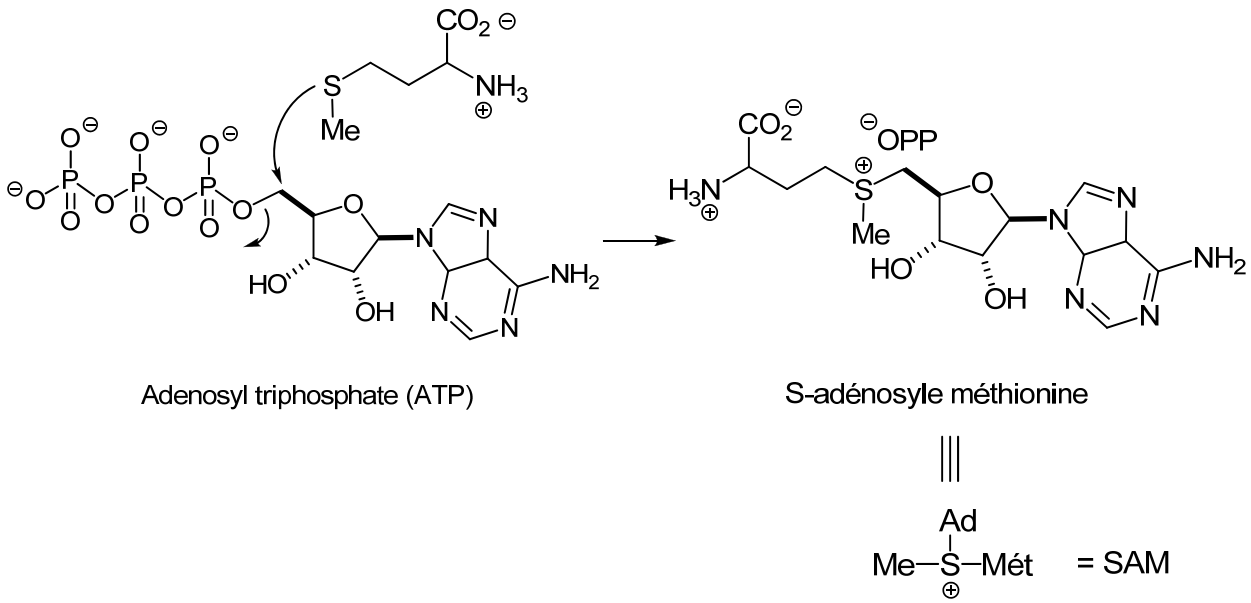


Schéma II.8.13

Les mécanismes de méthylation sont indiqués au schéma II.8.14. Ils ont été proposés à la suite d'expériences de marquage isotopique. Les enzymes sont capables de méthyler une double liaison en utilisant la SAM. Un groupement éthyle est introduit par une deuxième méthylation de la double liaison terminale (groupement méthylène) formée par une première méthylation suivie de la déprotonation du groupement méthyle. Il faut bien noter que les routes diffèrent selon que la configuration du carbone à la position 24, qui devient chiral suite à l'alkylation, est S ou R. Ce sont principalement les noyaux lanostérol et cycloarténol qui subissent des méthylations de la chaîne latéral dans les organismes végétaux.

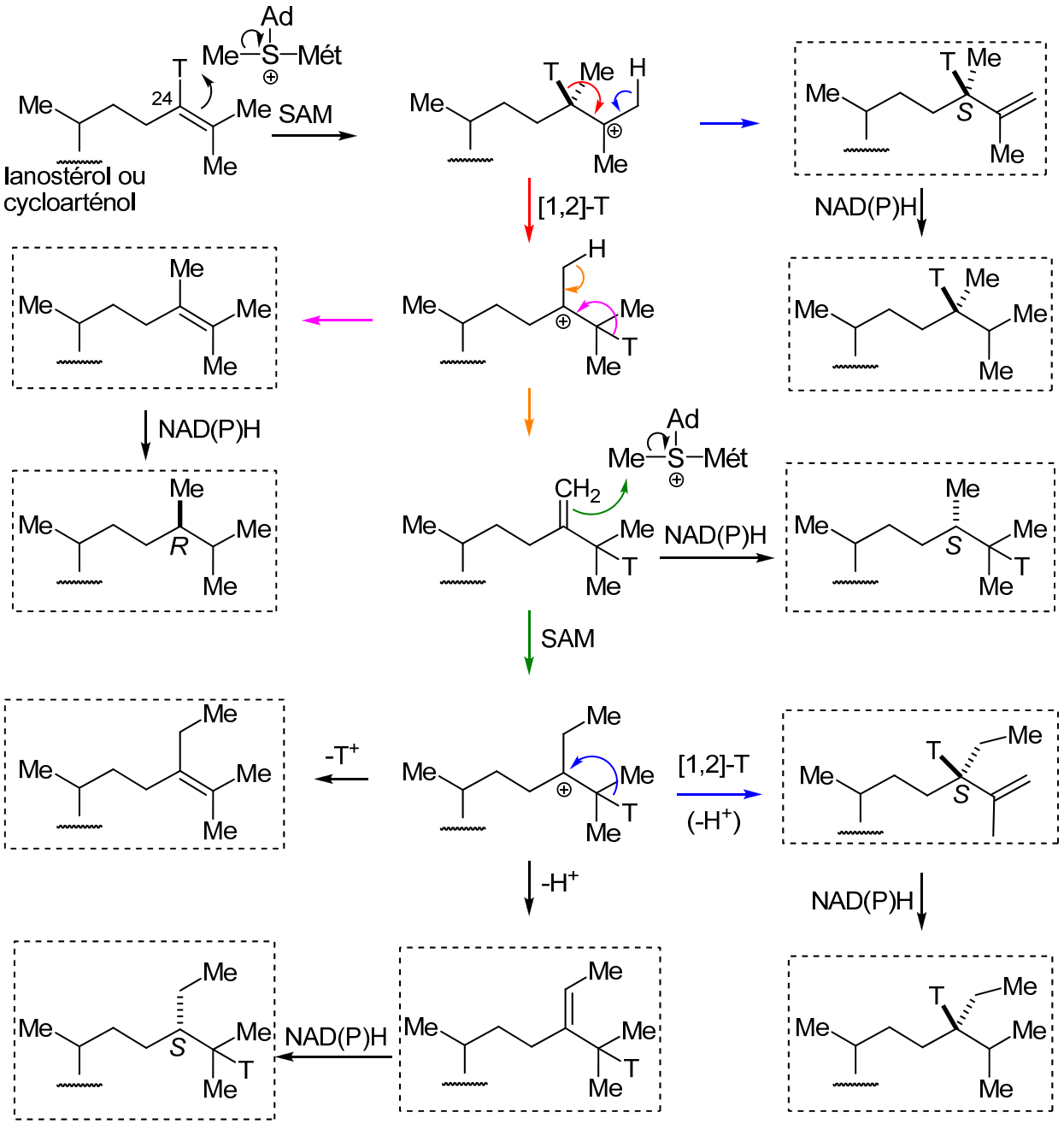


Schéma II.8.14

II.8.4d Modification et dégradation de la chaîne latérale des stéroïdes par oxydation

Le schéma II.8.15 montre la transformation du cholestérol en sapogénine par oxydation du carbones 15 du cycle D et des carbones 22 et 26 de la chaîne latérale. Deux voies possibles sont illustrées. Les deux voies sont différentiables par la source de l'atome d'oxygène sur l'acétal.

Dans un cas, l'oxygène du carbonyle attaque un carbocation généré à partir de la double liaison. Dans l'autre cas, l'hydroxyle sur la position 26 cyclise directement pour donner l'acétal.

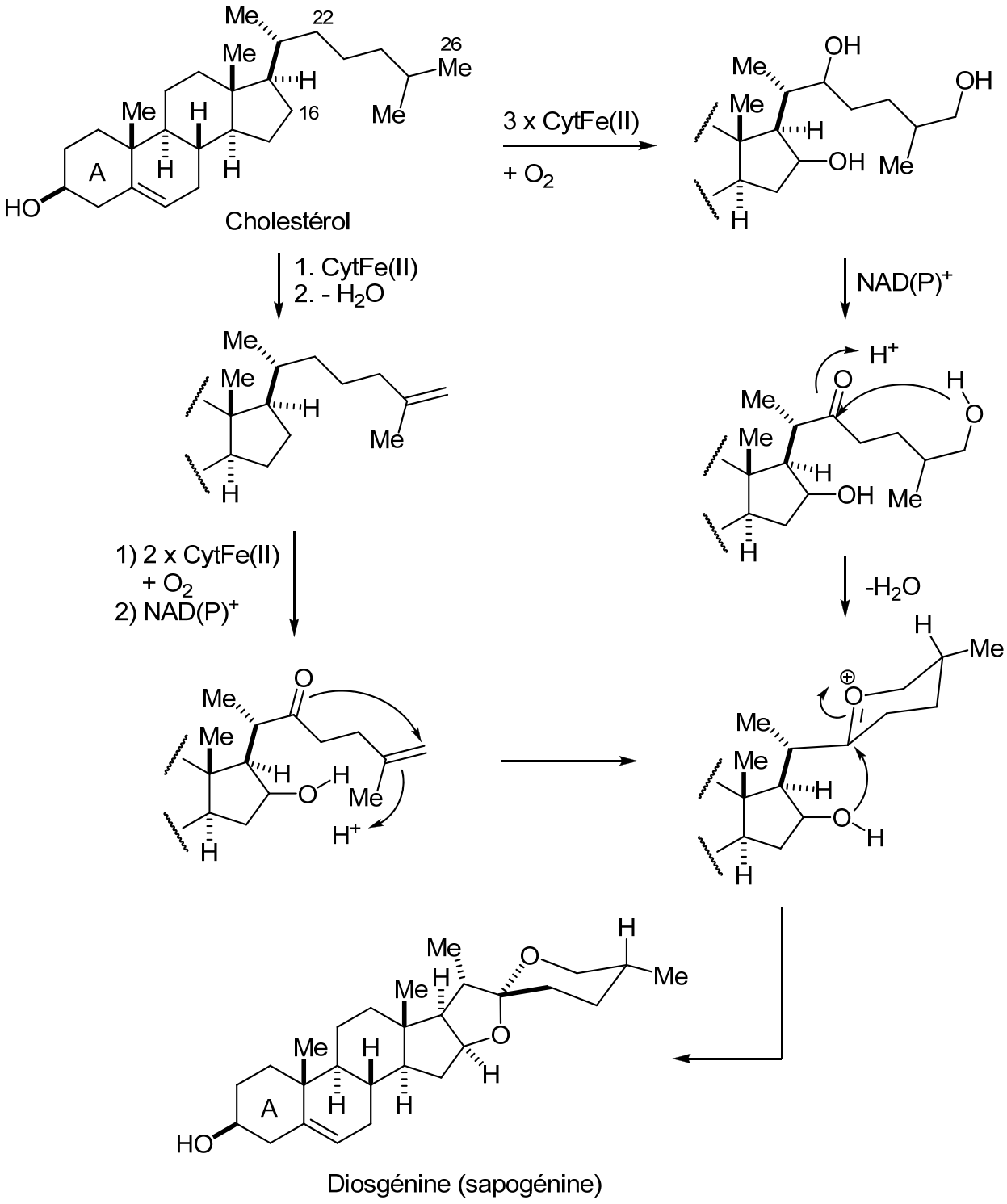


Schéma II.8.15

Le schéma II.8.16 montre deux voies possibles de conversion de la chaîne latérale du cholestérol en chaîne latérale de l'acide cholique (perte de 3 carbones). Dans la première voie, deux oxydations et une phosphorylation créer le précurseur requit pour une fragmentation de Grob avec une perte de trois carbones. L'autre alternative utilise plutôt un peracide pour déclencher une réaction de Baeyer-Villiger.

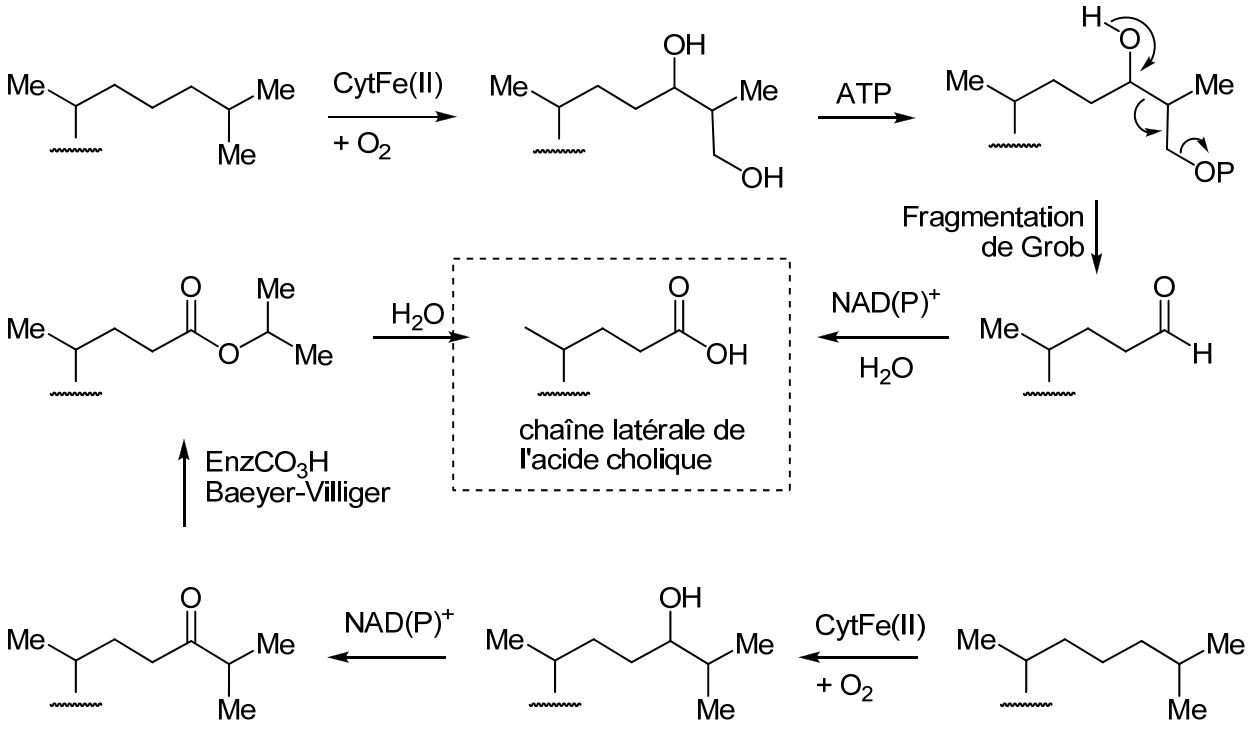


Schéma II.8.16

Les hormones sexuelles sont aussi fabriquées par oxydation de la chaîne latérale du cholestérol mais par une voie différente.

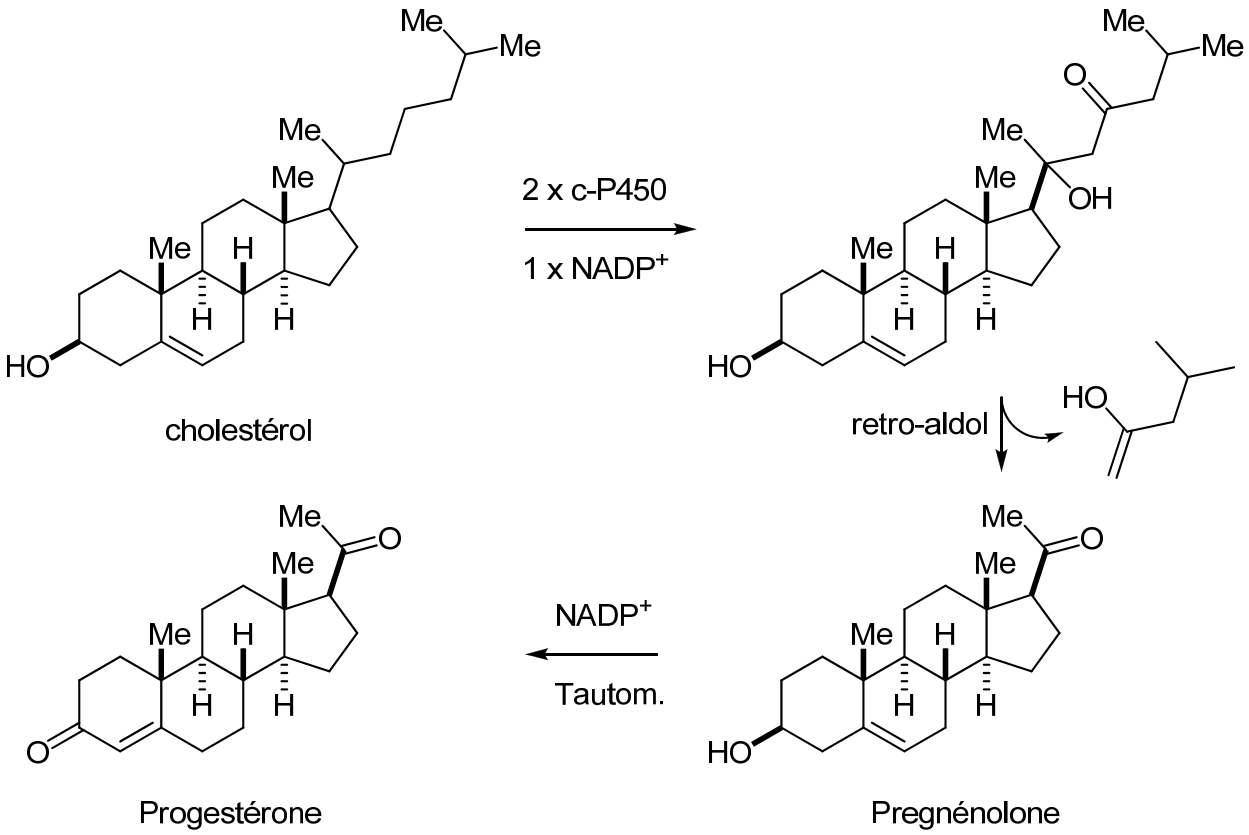


Schéma II.8.17

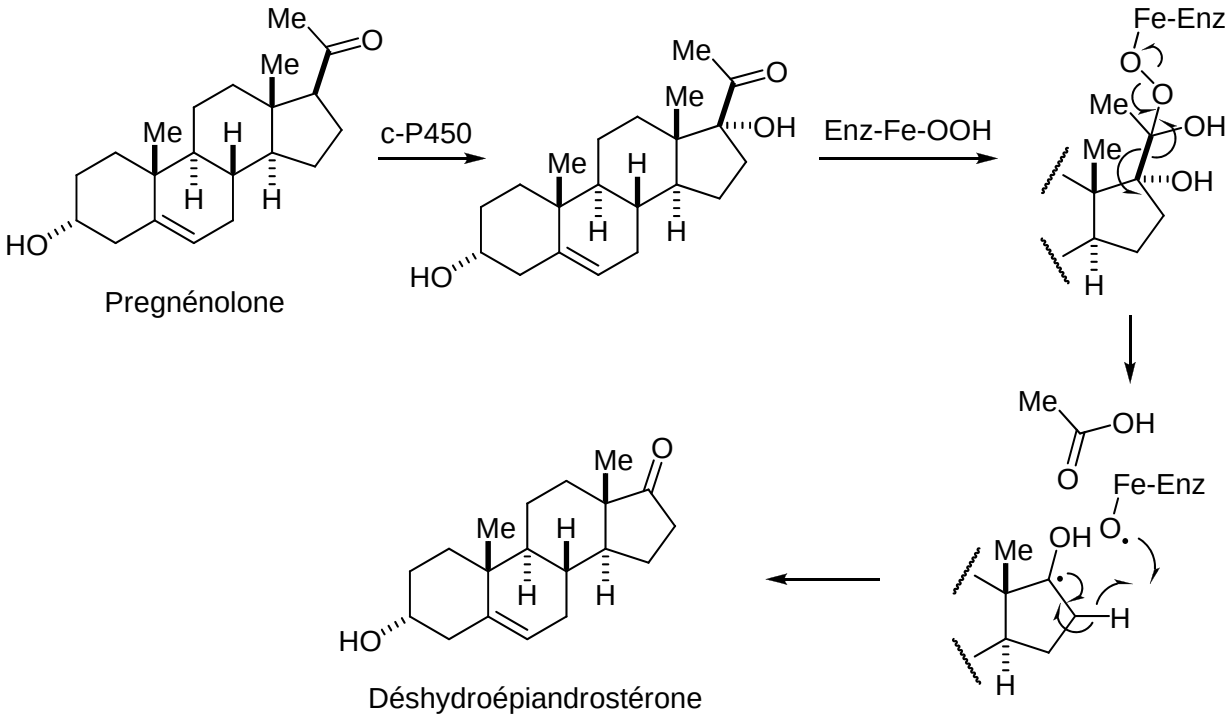


Schéma II.8.18

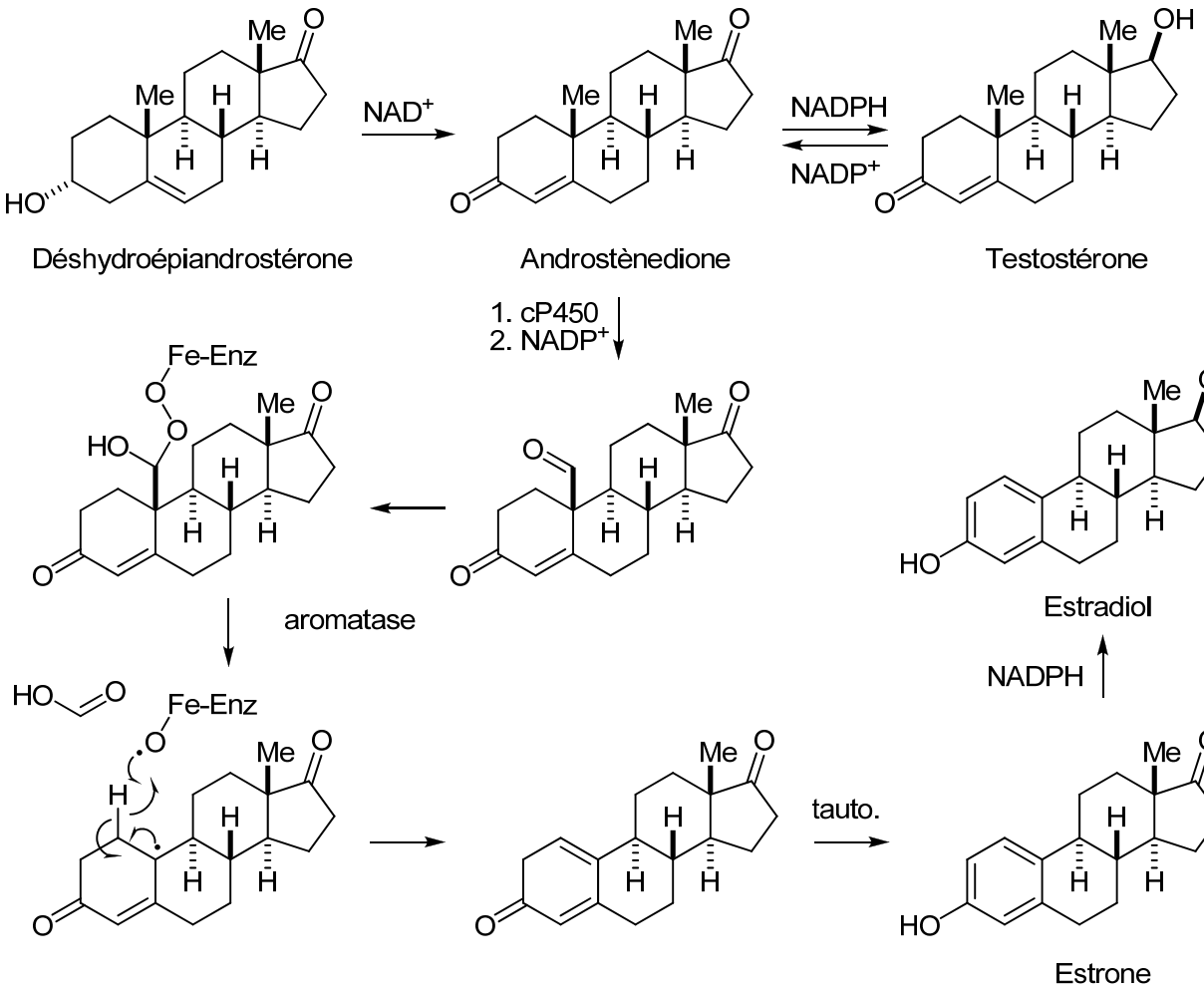
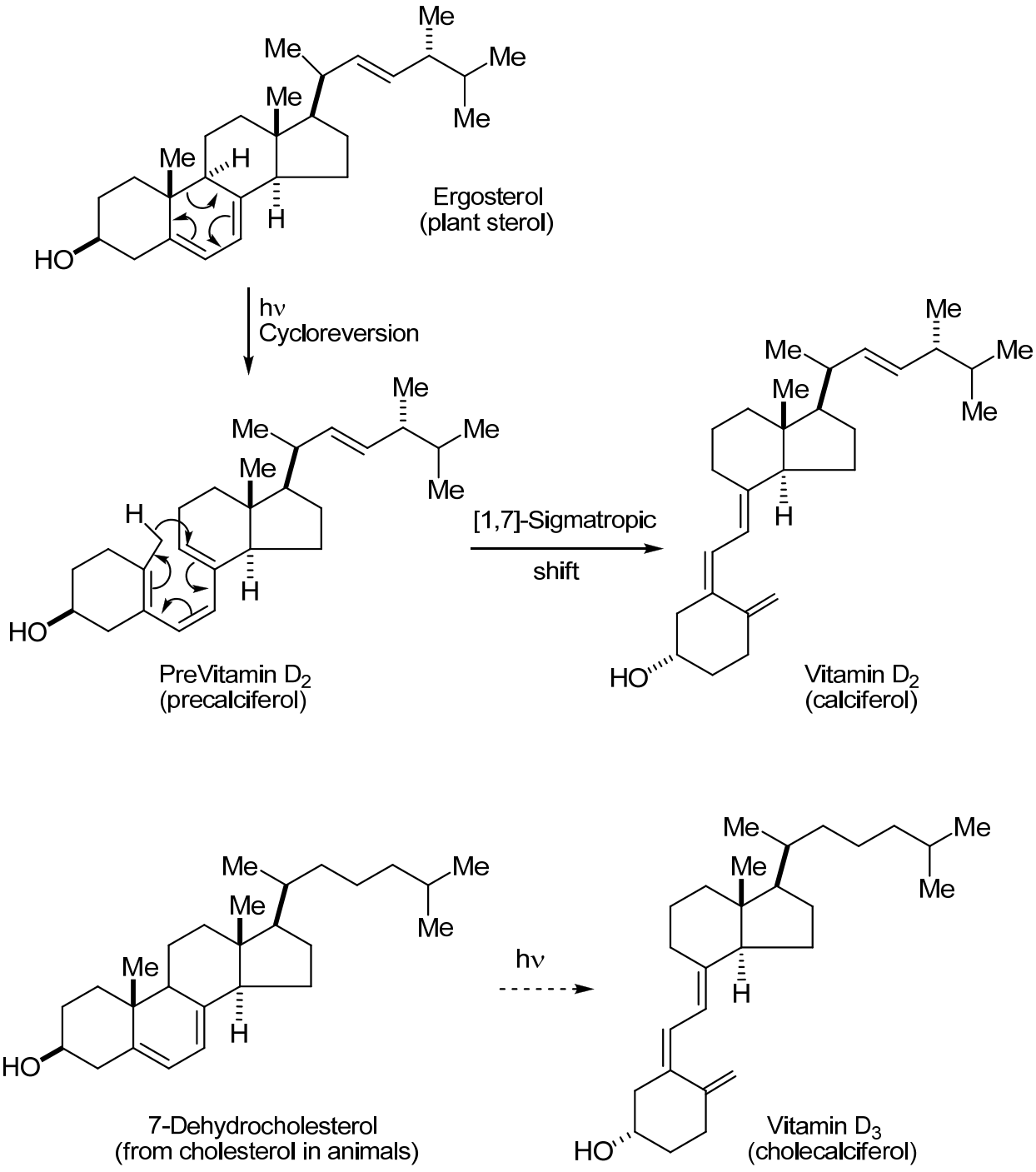


Schéma II.8.19

II.8.5 Vitamin D Chemistry.

Vitamin D's have a B-secosteroid structure and are formed from a photochemical reaction of steroid molecules (Scheme II.8.20). Their role in the control of calcium regulation has long been recognized. A deficiency in Vitamin D leads to "rickets" in which the calcification of bones is insufficient resulting in a weakening of the bones. This condition can be corrected by ingestion of fish liver oil where Vitamin D_3 is abundant. 7-Dehydrocholesterol is present in the skin where the ultraviolet irradiation of the sun transforms it first into previtamin D_3 and then into vitamin D_3. It is thus no surprise that "rickets" was more common in areas where winter is long. Vitamin D_2 is formed in plants by the action of uv light on ergosterol. Ergosterol is first transformed into precalciferol (preVitamin D_2) in a photochemical reverse pericyclic reaction called a [4π + 2 σ]-

cycloreversion. Then, precalciferol suffers a [1,7]-sigmatropic shift to afford vitamin D_2. The same process occurs with 7-dehydrocholesterol and previtamin D_3.



Scheme II.8.20

II.9 Biosynthèse des tétraterpènes et caroténoïdes

II.9.1 Biosynthèse du lycopène et du β-carotène

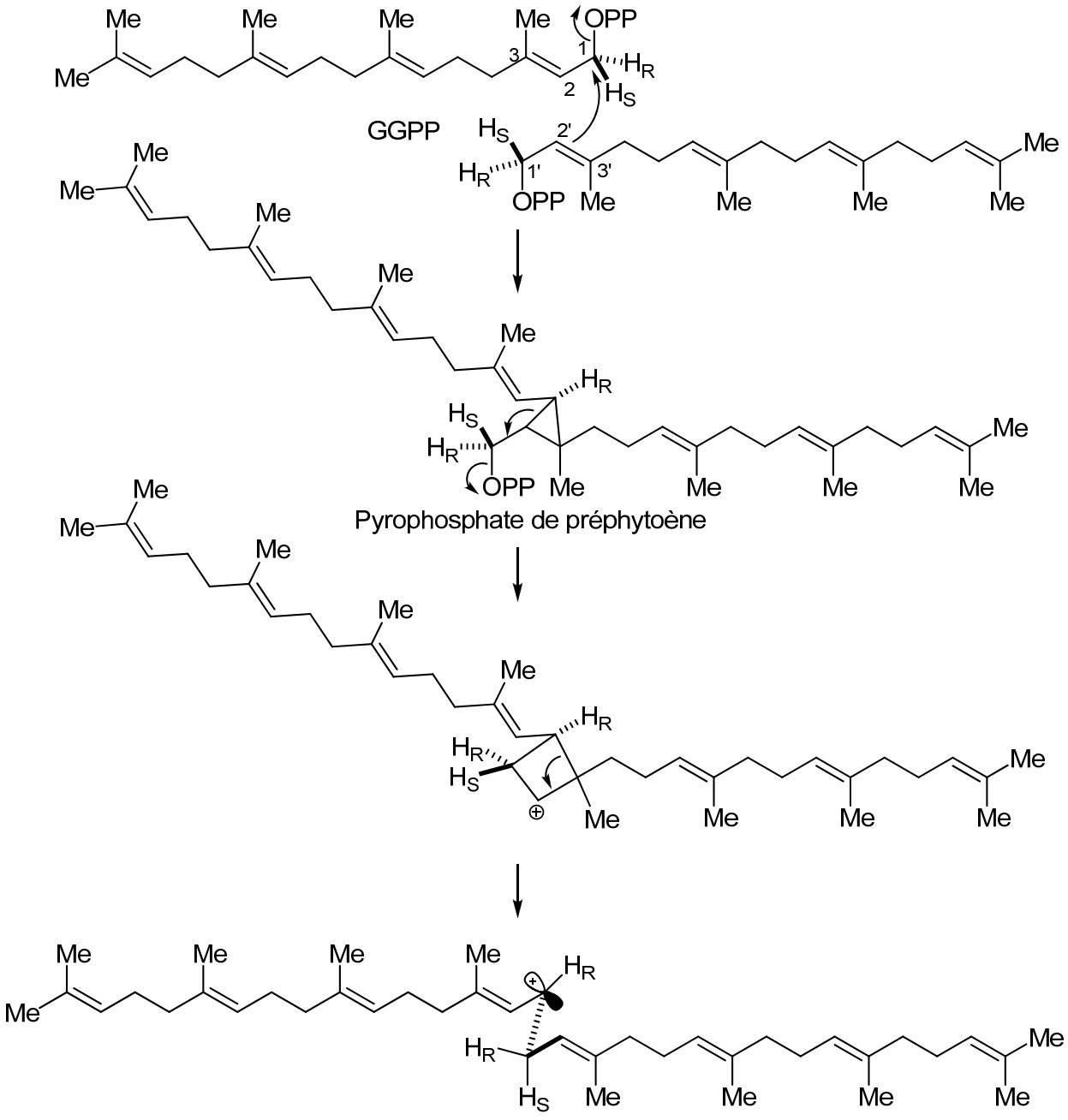


Schéma II.9.1

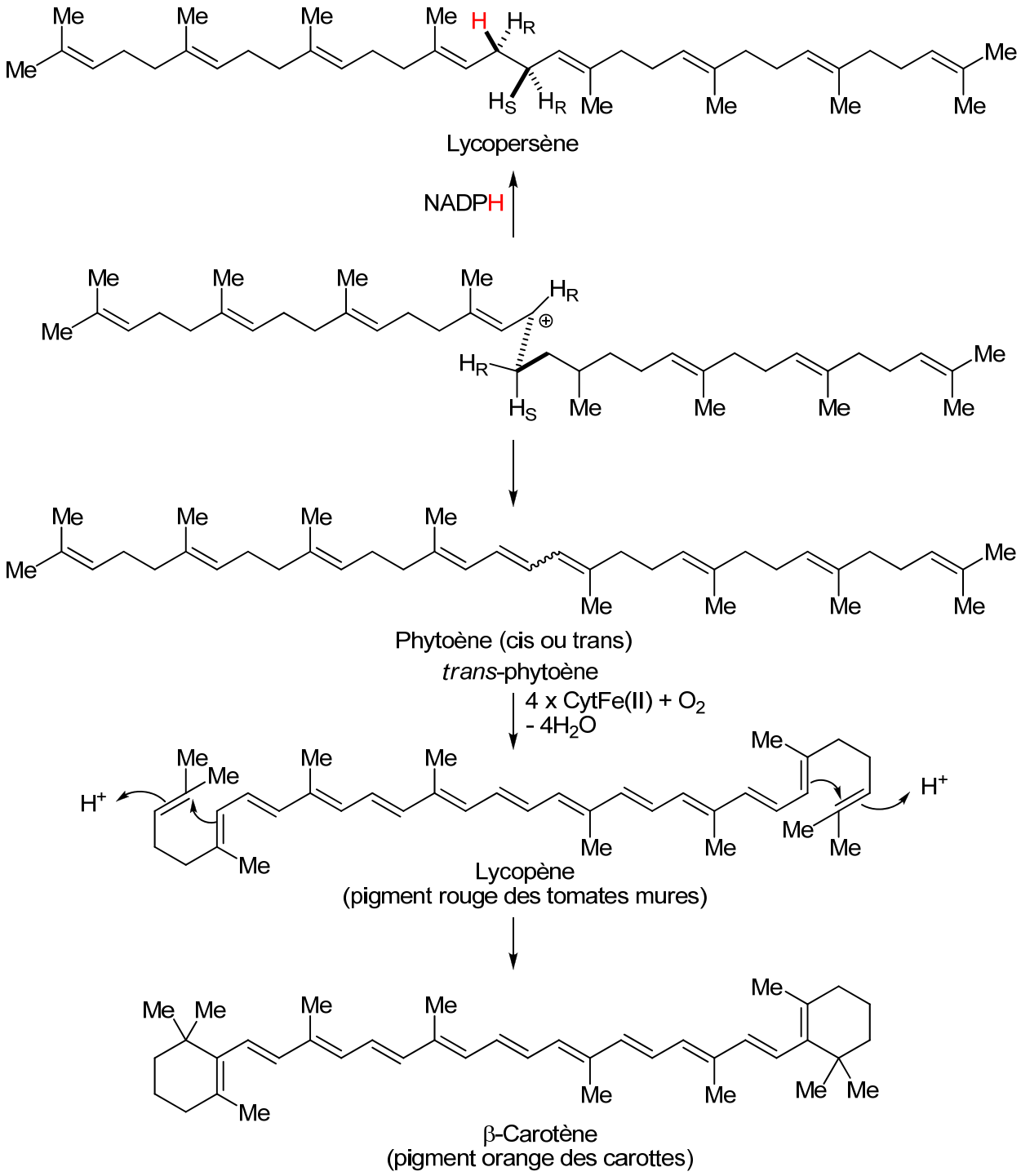


Schéma II.9.2

Tel qu'indiqué au schéma II.9.1, il y a condensation queue à queue de deux molécules de GGPP pour donner d'abord le pyrophosphate de préphytoène (cyclopropane). Le mécanisme de couplage est exactement le même que celui du couplage queue à queue du FPP dans la

biosynthèse des triterpènes. La réduction de l'intermédiaire cation allylique donne le tétraterpène (C_{40}) lycopersène (schéma II.9.2, haut) tandis que l'élimination d'un proton conduit aux tetraterpènes *cis*-phytoène (élimination de H_S) et *trans*-phytoène (élimination de H_R). L'introduction de doubles liaisons par oxydation dans le *trans*-phytoène donne le lycopène, un polyène conjugué de couleur rouge (le pigment rouge des tomates mûres). Le lycopène est le précurseur du β-carotène (le pigment orange des carottes) lui-même précurseur d'autres pigments caroténoïdes (schéma II.9.2).

II.9.2 La vitamine A

La vitamine A est essentielle à la croissance et à la vision chez les animaux. Elle provient de la dégradation du β-carotène par des enzymes du foie. Cette conversion implique des oxydations avec coupure de liaison C−C. Le mécanisme détaillé n'est pas connu. Le schéma II.9.3 montre comment la vitamine A intervient dans le mécanisme de vision. Elle est d'abord oxydée en *trans* rétinal qui est isomérisé en 11-*cis*-rétinal par les enzymes du foie. Parvenu dans les cellules de type bâtonnets situées dans la rétine (qui contient aussi des cellules de type cônes), le 11-*cis*-rétinal réagit avec un phospholipide porteur d'un groupement NH_2 (une phosphatidyl-éthanolamine par exemple) pour former une base de Schiff protonée (un ion iminium) qui réagit avec la protéine opsine pour donner la rhodopsine qui est sensible à la lumière dans le visible. Lorsque la lumière atteint les bâtonnets de la rétine, il y a isomérisation de la double liaison C_{11}=C_{12} de la configuration *cis* à la configuration *trans* pour former la métarhodopsine II. En absence de lumière, l'isomérisation *cis*-*trans* prendrait environ 1100 ans. En présence de lumière, elle se fait en 20 picosecondes. Le changement de géométrie qui accompagne l'isomérisation cause l'activation d'une protéine Gs, la transducine, laquelle active une phosphodiestérase qui transforme le c-GMP en GMP. Finalement, le processus cause une fermeture des canaux Na^+, ce qui provoque une hyperpolarisation de la membrane cellulaire ganglionique et une impulsion nerveuse qui est transmise au cerveau par le nerf optique. Il y a ensuite coupure de la métarhodopsine II en opsine et trans-rétinal qui est reconverti en rhodopsine (en 40-60 minutes chez l'humain). Une déficience en vitamine A résulte en une difficulté à voir dans la pénombre.

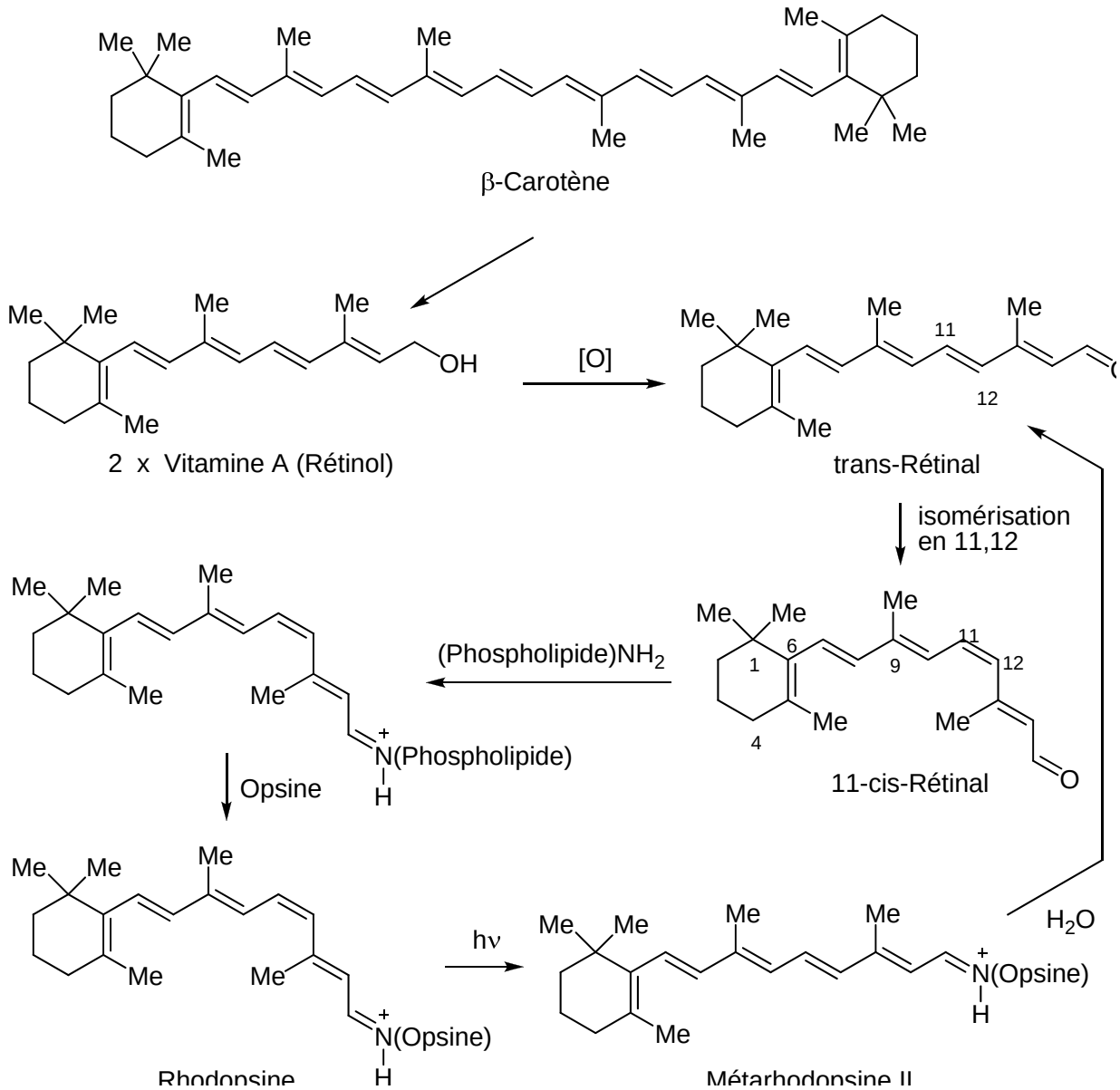

β-Carotène
2 x Vitamine A (Rétinol)
[O]
trans-Rétinal
isomérisation
en 11,12
11-cis-Rétinal
(Phospholipide)NH2
N(Phospholipide)
H
Opsine
Rhodopsine
hν
Métarhodopsine II
H2O
N(Opsine)
H

Schéma II.9.3

Bibliographie

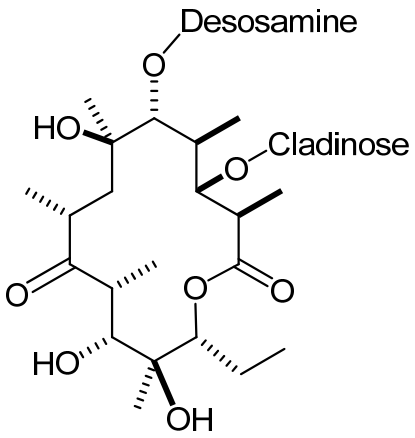
1. W. Eisenreich, M. Schwarz, A. Cartayrade, D. Arigoni, M.H. Zenk et A. Bacher, *Chem. Biol.*, 1998, **5**, R221-R233)

III. Acetogenins

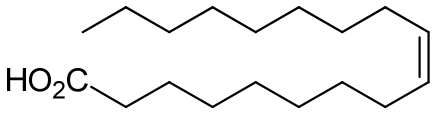
III.1. Introduction.

Acetogenins are natural products derived from acetate units assembled in a linear fashion. Acetogenin compounds are varied in structure. Figure III.1.1 lists some examples. The macrolide antibiotics, such as erythromycin, are produced by bacteria. Erythromycin is often prescribed to patients allergic to penicillin. Wax and fatty acids like oleic acid have many roles, which include an energy storage mechanism, cell membrane functions and providing a water proof layer on skin, feather, leaves, etc. Isoflavones are a large family of flavourous and colorful compounds while xanthones, found in fruits, are used as insecticides. Prostaglandins and leukotrienes are involved in vasodilatation and constriction, platelet aggregation, and many other vital roles.

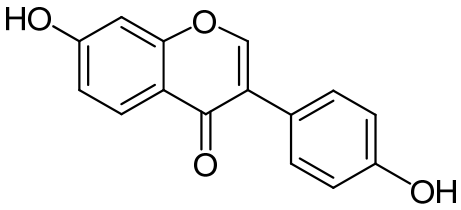
The acetogenins are divided into two main categories. The fatty acids, include fatty acids, wax, phospholipids, prostaglandins, leukotrienes, and others. The polyketide derived metabolites, include aromatics, heteroaromatics, and others. Although both classes find their origin in long chains of acetate units, they differ in the way the chain is grown. Their respective mechanisms are similar but distinct, the main difference being that fatty acids have each unit reduced before elongation takes place and polyketides have the units preserved in the elongation awaiting proper transformation into aromatic compounds. Unlike the terpenoids, mevalonic acid is not an intermediate. Scheme III.1.1 shows the pathway leading to fatty acids, polyketides, and terpenoids, from acetyl coenzyme A. Let's see the two classes of acetogenins separately.



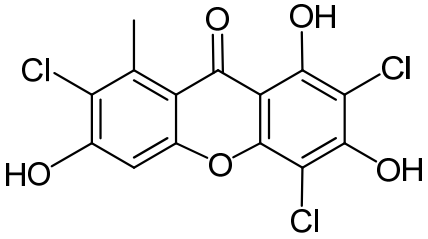
Erythromycin A



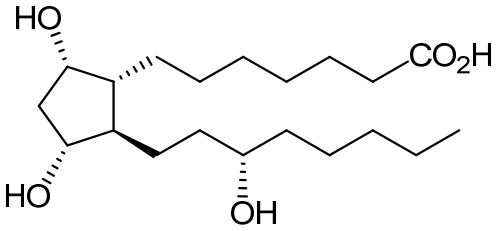
Oleic acid



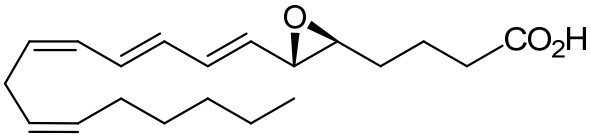
Isoflavone



Xanthone

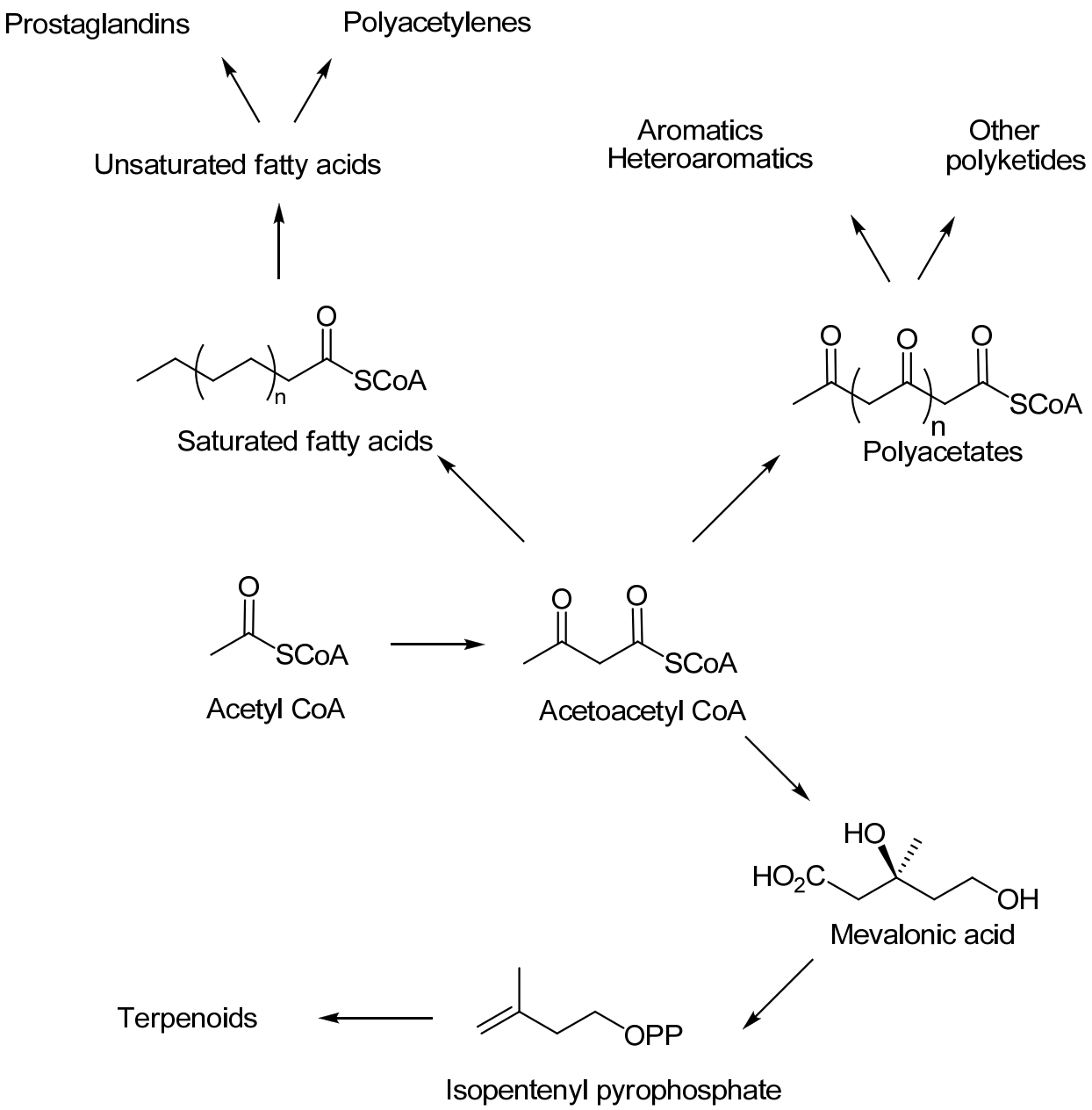


PGF_1
(prostaglandin)



LTA_4
(leukotriene)

Figure III.1.1



Scheme III.1.1

III.2. Fatty Acids.

III.2.1. Biosynthesis of Fatty Acids.

The fatty acids are abundant in all cellular organisms since the major constituents of cell membranes are lipids of which fatty acids are principal components. In fact the term lipid,

although it encompasses all cellular components soluble in non-polar solvent such as the fatty acids, the steroids, and others, is often used to refer to long chain hydrocarbons. They are most abundant in the form C_8, C_{10}, C_{12}, C_{14}, C_{16}, and C_{18}. Saturated acids are of the form $MeCH_2(CH_2CH_2)_nCH_2COOH$ with the most common being (in order of increasing number of carbons) caprylic (C_8, n=2), capric (C_{10}, n=3), lauric (C_{12}, n=4), myristic (C_{14}, n=5), palmitic (C_{16}, n=6), and stearic (C_{18}, n=7).

In vivo, they form glycerides, i.e. one, two, or three fatty acids condense with glycerol to form esters (Figure III.2.1). They also form phospholipids, for example phosphatidyl choline, phosphatidyl *myo*-inositol, and platelet aggregation factor (PAF) (Figure III.2.1). The phospholipids are vital for the structural integrity of cell membranes by controlling several properties of the membrane such as contractility and permeability to ions. PAF is involved in immunoreactions such as allergies and blood clothing (thrombosis) and related phenomena. The most ubiquitous fatty acids have such an important role in the integrity of life that they are classified as primary metabolites. The uncommon fatty acids are considered as secondary metabolites. Fatty acids also play a major role in the storage of energy for cellular organisms. Excess energy may be stored in making fatty acids and it can be regained when needed by a degradation method called the ß-oxidation of fats. We shall look at that in more detail in section III.2.2.

Fatty acids with chain length of 20 C or greater are rare but do occur in natural waxes, usually as esters of sterols or hydroxy fatty acids. Beeswax contains palmitic (C_{16}) and cerotic acids (C_{26}) esterified with melissyl alcohol (C_{30} fatty acid derived). The lower acids occur mainly in animal fats: cow's milk fat contains considerable amount of butanoic acid (C_4) together with smaller amounts of C_6 to C_{12}.

Fatty acids are used to make soaps and detergents. The sodium salt of lauric and myristic acid are used in domestic soaps, while the potassium salts are employed in more specialized preparations such as shaving cream and liquid soap. Sulfates of fatty acids are synthetic soaps with added advantages to natural soaps in that they have an enhanced resistance to hard water (water rich in calcium and magnesium salts). Figure III.2.1 shows sodium glyceryl monolaurate sulfate derived from lauric acid, glycerol, and sulfate.

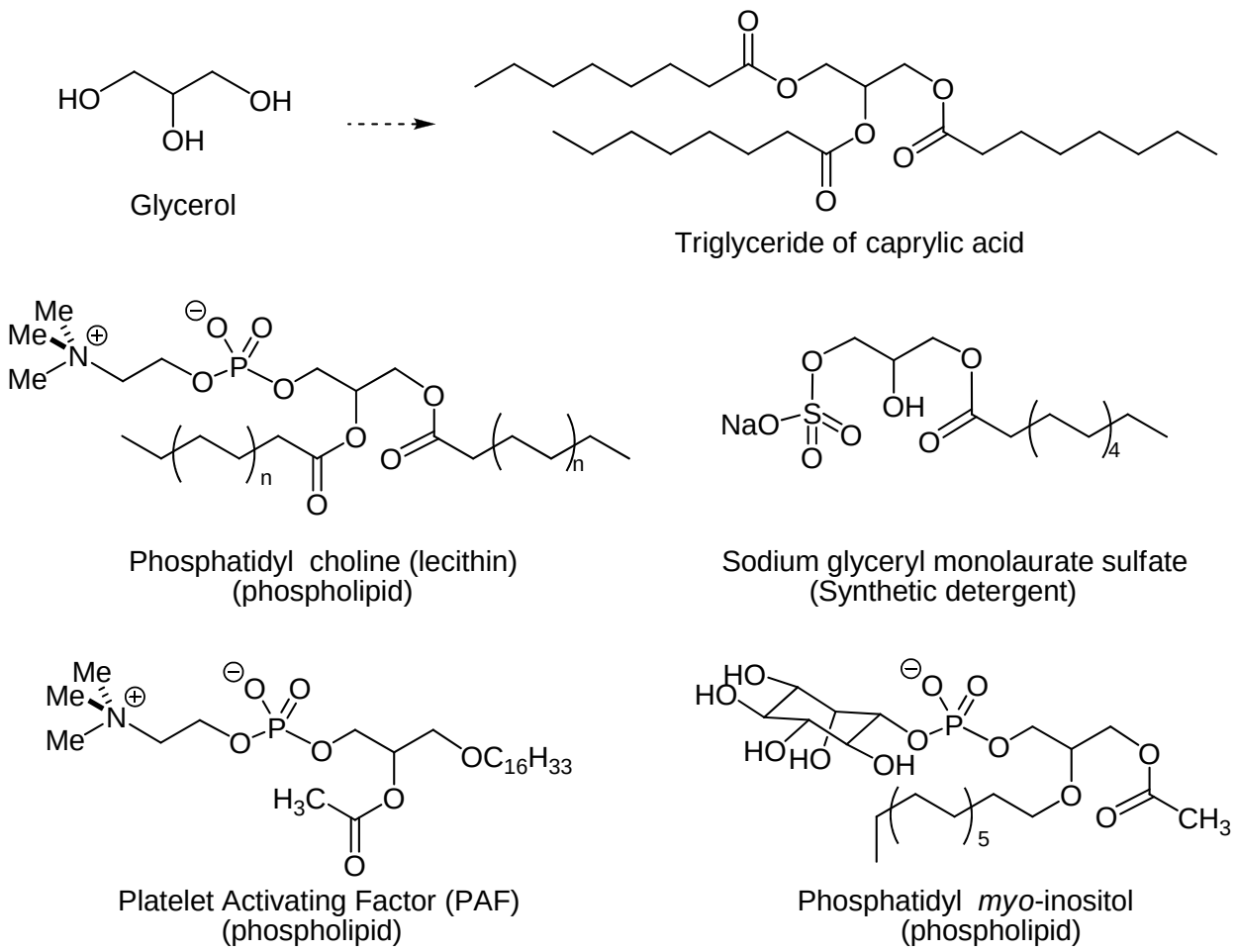


Figure III.2.1

In bacteria and plants, fatty acid biosynthesis is accomplished by a series of monofunctional proteins in a dissociated type II fatty acid synthase (FAS) system. In contrast, the type I FASs of fungi and animals are huge multifunctional polypeptides that integrate all steps of fatty acid synthesis into large macromolecular assemblies. The crystal structure of this type I FAS has recently been published (Science **2008**, *321*, 1315) with a resolution of 3.5Å (Figure III.2.2).

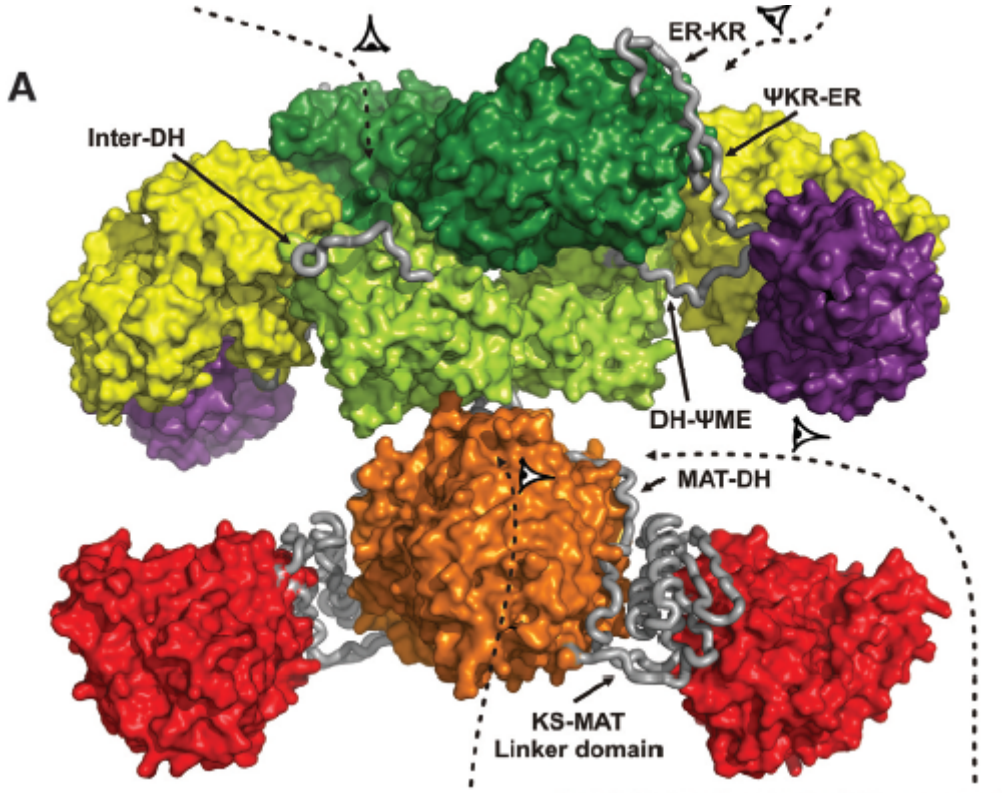
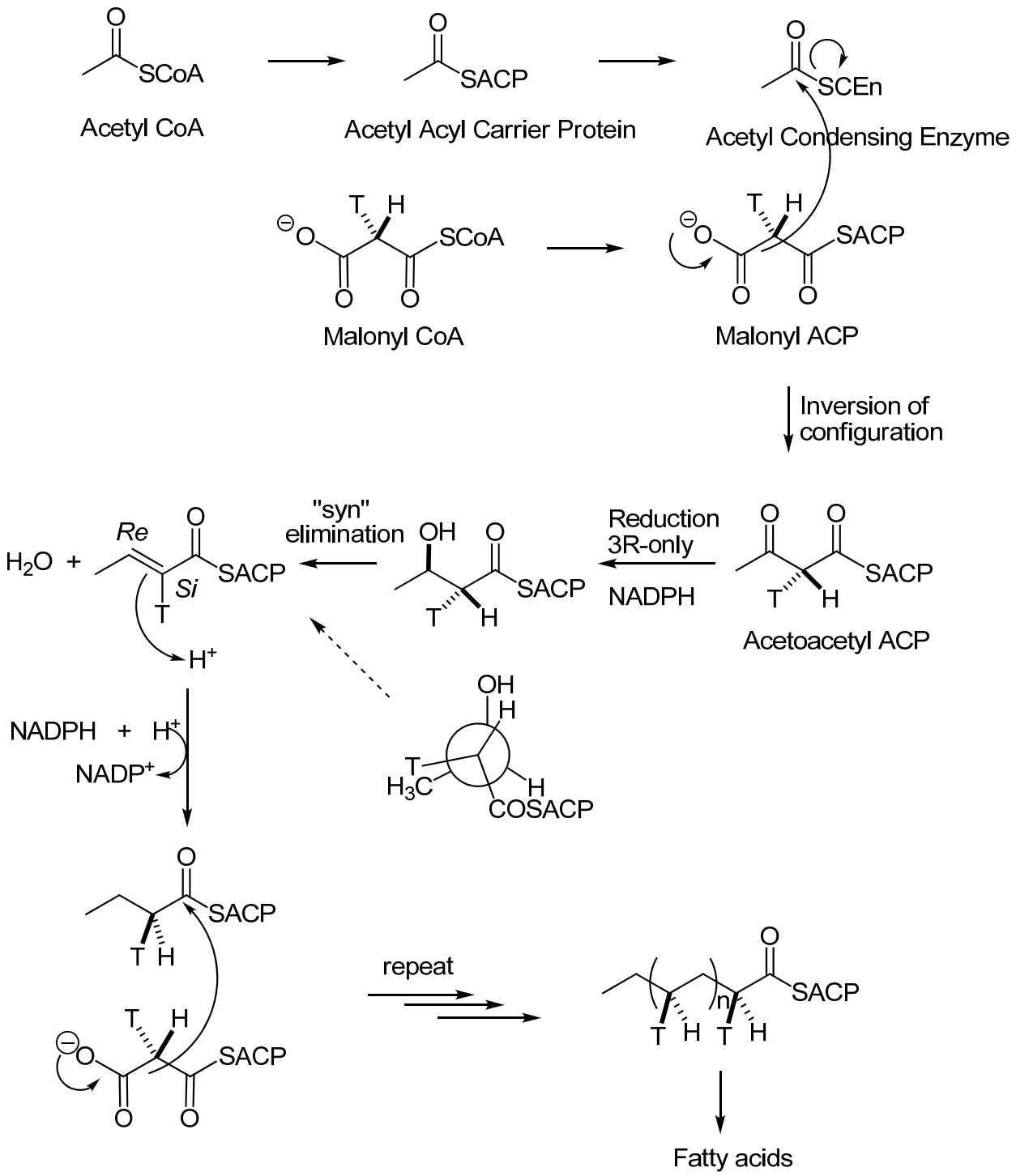


Figure III.2.2

The biosynthesis of fatty acids normally starts with a molecule of acetyl coenzyme A called the starter molecule (Scheme III.2.1). Sometimes the starter is a molecule of **propionyl CoA** (3 carbons) or **butyryl CoA** (4 carbons) and can lead to fatty acids with odd number of carbons and/or to complex acetogenins incorporating such units like the macrolide and the polyether antibiotics. It appears that the same enzyme system is at play whatever the starter molecule. The abundance of the different starters may determine what fatty acid will be synthesized.

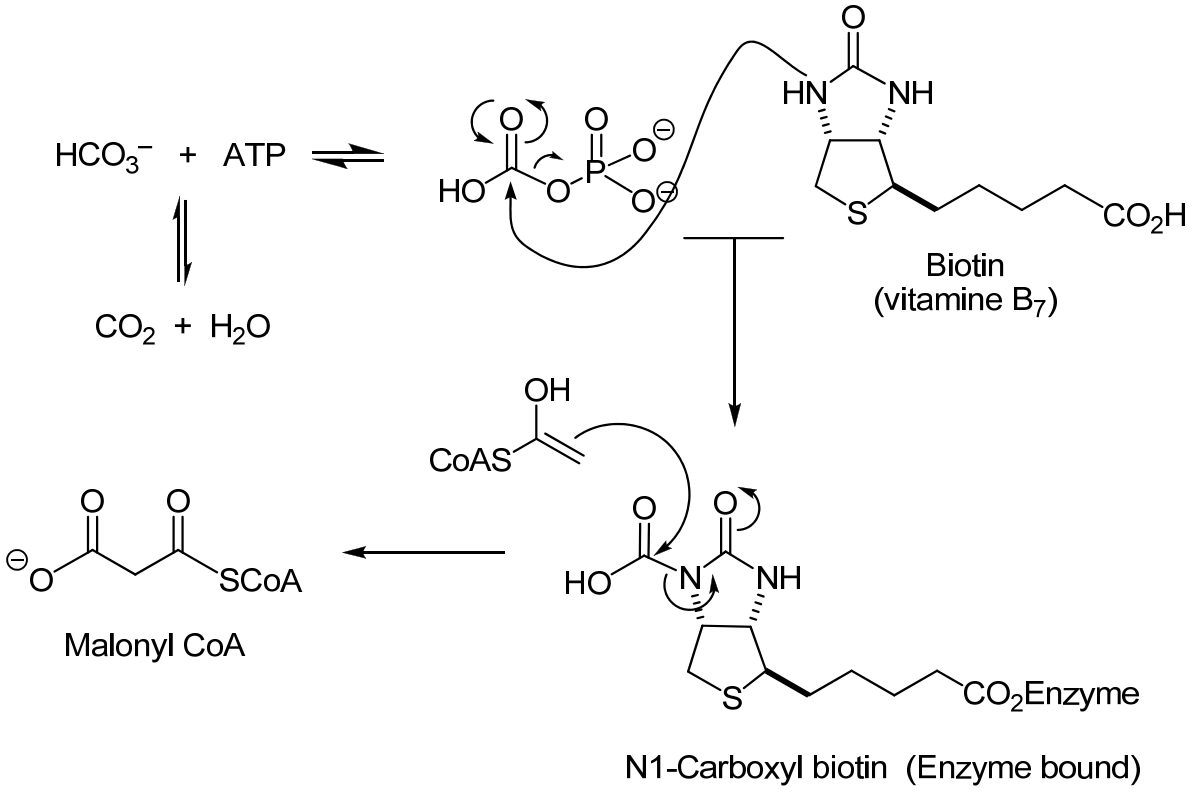
The condensation of acetyl CoA onto another molecule of acetyl CoA (Claisen condensation) is not directly catalyzed by the enzymes responsible for fatty acid synthesis. Instead, one molecule of acetyl CoA is first **activated** by adding an extra electron withdrawing carboxyl group to form malonyl CoA (Scheme III.2.1). This carboxyl group comes from addition of carbon dioxide but is **not** incorporated into the fatty acids as proved by ^{13}C labelling experiments. Rather, CO_2 is lost in the condensation of malonyl CoA with acetyl CoA. Thus, the sole purpose of the formation of malonyl CoA is to activate one molecule of acetyl CoA and thus facilitate its condensation with another molecule of acetyl CoA (Scheme III.2.1). This carboxyl unit is brought about by CO_2 in solution i.e. HCO_3^- (bicarbonate ion) via a molecule of ATP and biotin. The transfer of CO_2 is depicted in Scheme III.2.2. What follows after the carbonylation of acetyl CoA is described in separate steps in Scheme III.2.1.



Scheme III.2.1

The condensation of malonyl CoA takes place followed by the stereoselective reduction of the ketone to the 3R-alcohol with nicotinamide adenine dinucleotide-2-phosphate (NADPH). Note that it is the pro-S hydrogen of NADPH that is transferred in this reduction. The alcohol is

then enzymatically dehydrated stereospecifically in a *cis* elimination to give the *trans* double bond. The last step involves the reduction of this double bond by NADPH and H^+. The reduction leads to an achiral compound but it is still stereospecific. Protonation occurs specifically on the *Re* face of the double bond in algae and on the *Si* face in yeast. Tritium labelled olefins prove this stereospecificity (Scheme III.2.1).



Scheme III.2.2

The long chain ester thus formed is ready for incorporation of another molecule of malonyl CoA and the cycle repeats itself until the right number of carbon atoms are in the chain. The fatty acid is then released from the enzyme complex. So far we have discussed all biosynthetic steps for fatty acid synthesis separately, but in reality, the whole synthesis is part of a multienzyme complex where the long chains are retained by the carrier protein and each enzyme comes to play its role and then leaves the complex. The ACP carries a panthoteine fragment linked to the protein by a cysteine and a phosphate group (Figure III.2.3). This long pantotheine chain allows the acetate molecule to be transferred from the ACP to the condensing enzyme and the chain elongation to be carried out via the Claisen condensation reaction.

The fatty acid is released from the multienzyme complex only after the required length is achieved. The multienzyme system includes all of the necessary enzymes for each steps. The

first of these enzymes is the condensing enzyme which fixes a new molecule of acetate (or the elongated chain as it applies) and catalyzes its condensation with malonyl CoA (Scheme III.2.3). Then the reductase catalyzes the reduction of the ketone to the 3R-alcohol. The dehydrase then catalyzes the elimination of water and the following reduction of the double bond is catalyzed by another reductase. Then the transacylase moves the elongated chain to the condensing enzyme and catalyzes the transfer of another molecule of malonyl CoA so the cycle can continue.

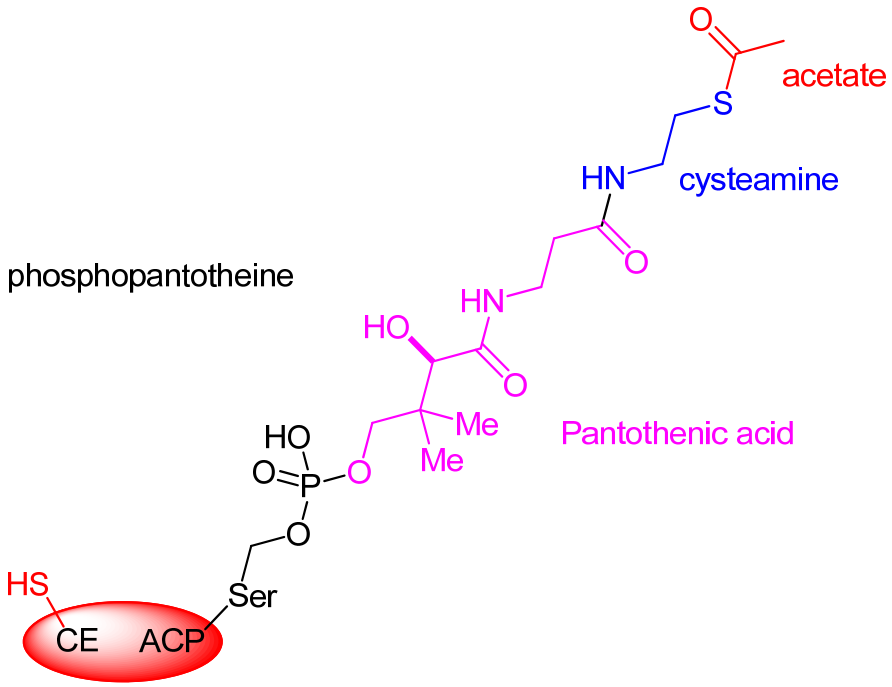
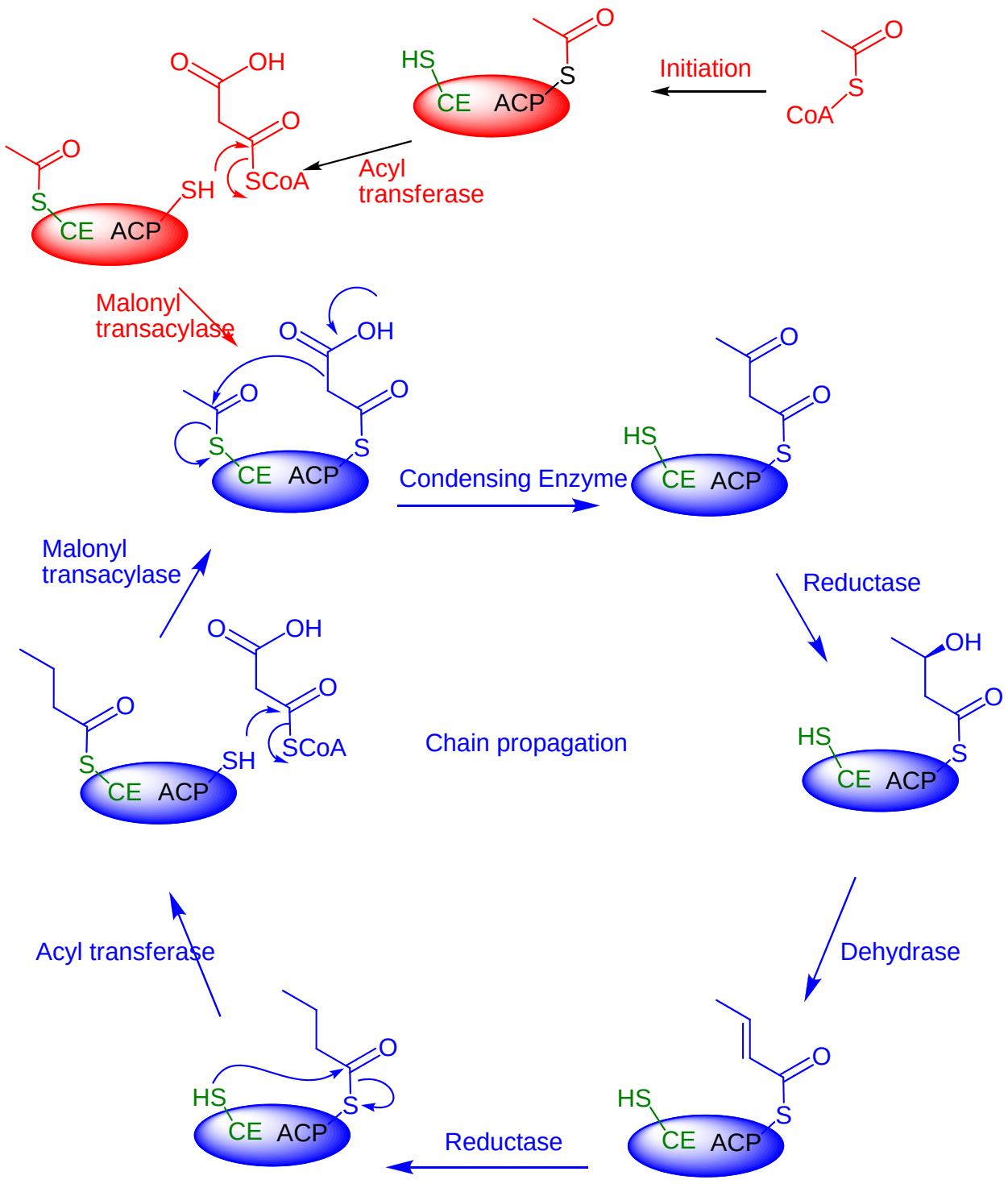


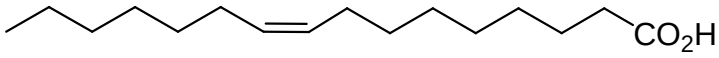
Figure III.2.3



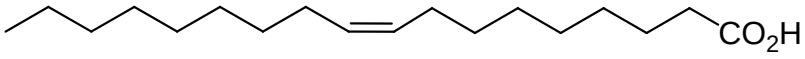
Scheme III.2.3

From these saturated fatty acids are derived the unsaturated fatty acids, the acetylenic fatty acids, the prostaglandins, the leukotrienes, and others, by further modification of the saturated chain. A series of enzymes are able to activate the saturated chain by first dehydrogenating at

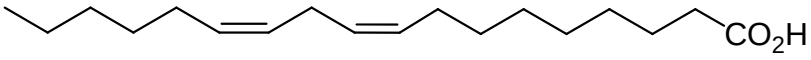
specific sites and further oxidation reactions transform the fatty acid into complex acetogenins. When only dehydrogenation reactions take place, the so-called unsaturated fatty acids are produced. The majority are occurring as C_{18} chains and are rarely longer than C_{22} or shorter than C_{14}. Representative examples are given in Figure III.2.4. Palmitoleic acid occurs in nearly all fats, especially those of marine origin; oleic acid represents 83% of total fatty acids present in olive oil and 60% in peanut oil; linoleic acid is the second most abundant fatty acid in peanut oil (21%); arachidonic acid is found in mammalian adrenal glands and is the common precursor of all the prostaglandin hormones (see section III.3.). Arachidonic acid itself is biosynthesized from linoleic acid which in turns comes from the saturated fatty acid stearic acid.



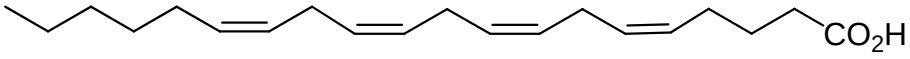
Palmitoleic acid (C_{16})



Oleic acid (C_{18})



Linoleic acid (C_{18})

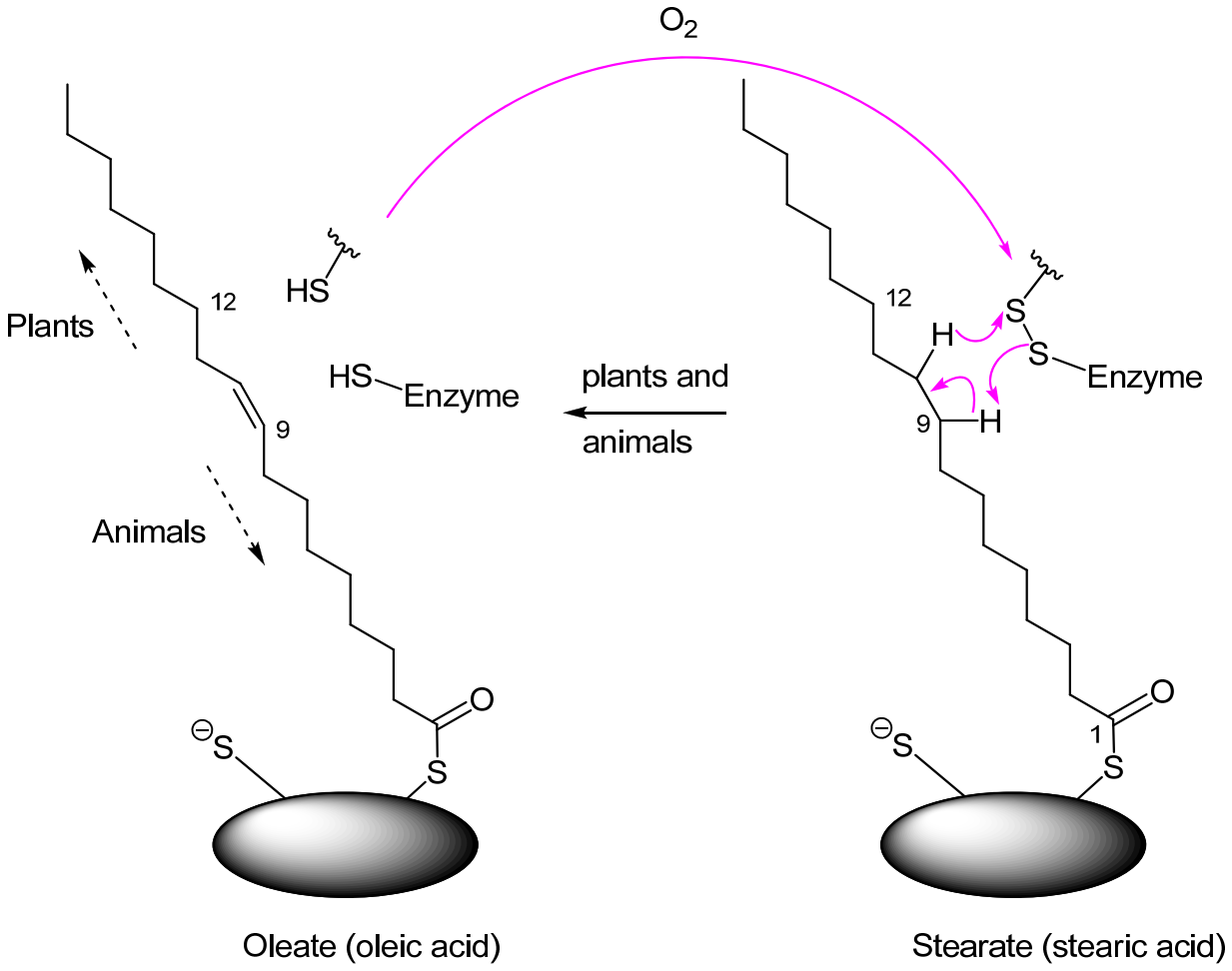


Arachidonic acid (C_{20})

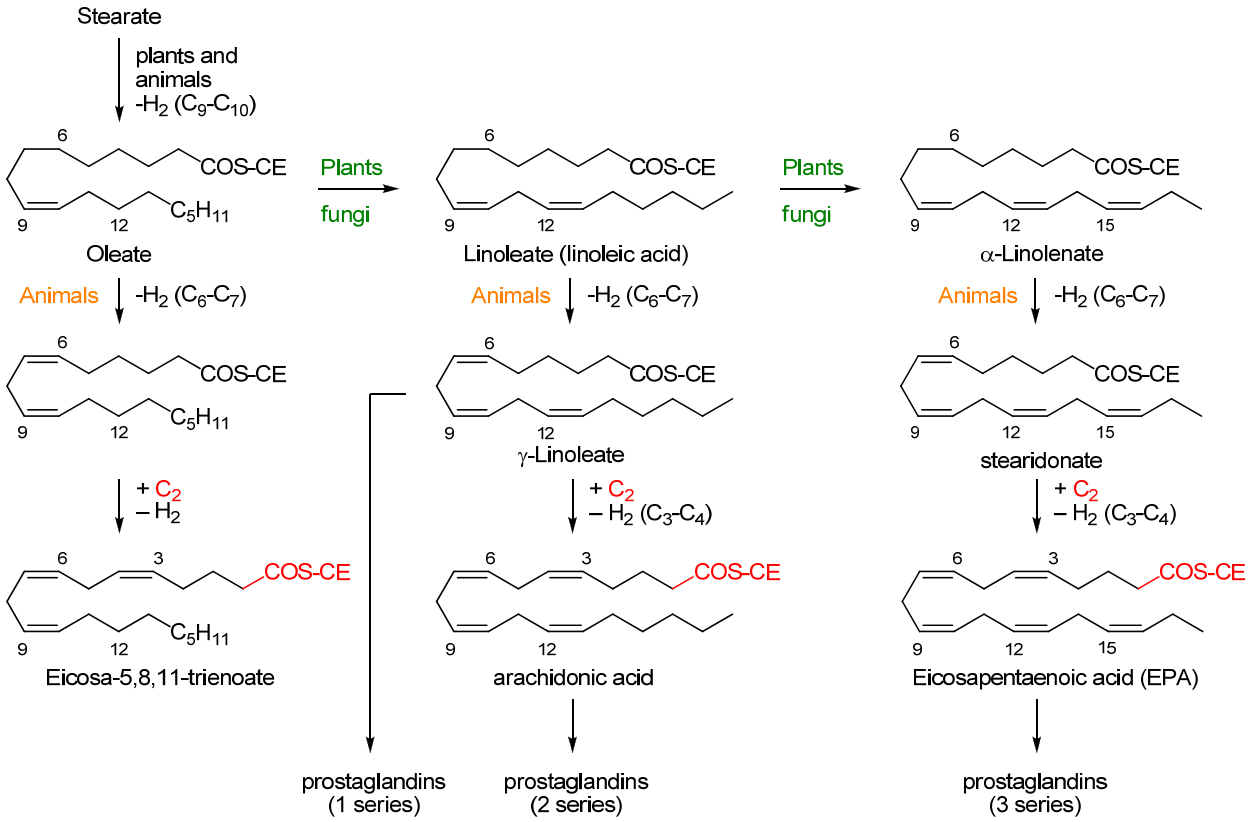
Figure III.2.4

The dehydrogenation of saturated fatty acids in animals and plants occurs by the mechanism shown in Scheme III.2.4. The exact mechanism is not known but the proposal that a disulfide bond of the enzyme is broken while effectively removing two *syn* vicinal pro-R hydrogens is widely accepted. Oxygen is the ultimate oxidizing agent in that it is needed to reform the sulfide bond of the enzyme for further dehydrogenations. Animals and plants alike first dehydrogenate

the C_9-position of stearic acid to afford oleic acid (Scheme III.2.4). The positions of subsequent dehydrogenations are not random and differ in the enzyme systems of animals and plants. The animals are not capable of desaturation at C_{12} of oleic acid or beyond. Plants are capable of dehydrogenation in positions closer to the end methyl group. This metabolic block in animals results in their incapacity to synthesized arachidonic acid **unless** they obtain linoleic acid from dietary sources. Arachidonic acid is essential to the well-being of animals. This is schematized in Scheme III.2.5. Finally it should be noted that bacteria living in anaerobic conditions synthesize monoenoic unsaturated fatty acids (only one double bond) by inhibiting the reductase enzyme at the required position. Instead an isomerase enzyme isomerizes the *trans* olefin to a *cis* olefin before further acetate units are added to the chain.



Scheme III.2.4



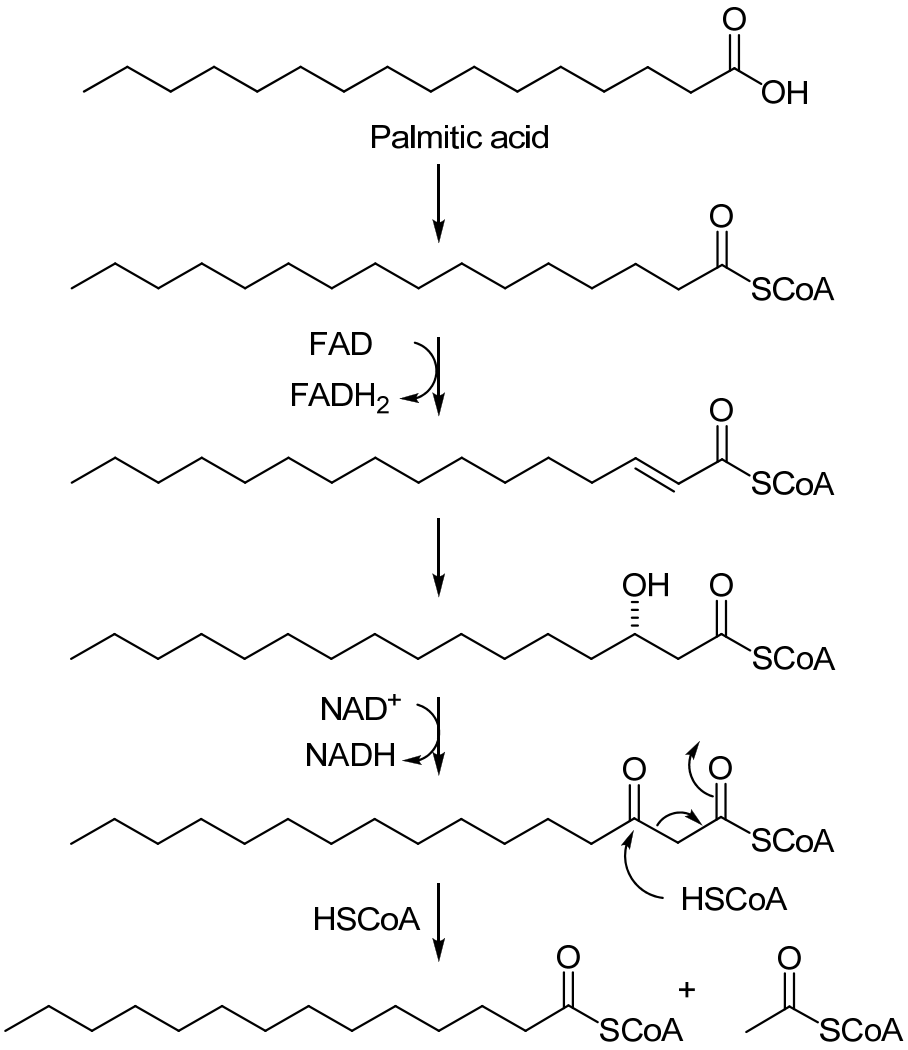
Scheme III.2.5

III.2.2. ß-Oxidation of Saturated Fatty Acids.

Fatty acids serve both as the primary material of biological lipid membranes and as energy storage. This stored energy can be released back into the system by a process called ß-oxidation. In this process, the fatty acid is cleaved back into units of acetyl coenzyme A (for example palmitic acid C_{16} gives 8 molecules of acetyl CoA), which in turn enters the Krebs cycle where it is transformed to CO_2 with release of free energy. The form in which this free energy is trapped is by oxidative phosphorylation or in other words by production of adenosyl triphosphate (ATP). The latter serves to trap and transport the energy where it is needed, for example in the production of steroid hormones, etc.

ß-Oxidation is formally the reverse of the biosynthesis of fatty acids. The first step involves esterification of the free acid to acyl CoA (Scheme III.2.6). Thus activated, the fatty ester suffers a dehydrogenation mediated by a molecule of FAD (see annexe VII.1). Note that the reverse reaction, i.e. the hydrogenation of the olefinic bond in the fatty acid synthesis, was not mediated by flavin adenine dinucleotide (FADH) but by NADPH. The *trans* olefin produced in the dehydrogenation is then hydrated to the *S*-alcohol (the reverse process involved the *R*-alcohol).

The *S*-alcohol is then oxidized to the ketone by NAD^+ and the resulting ß-ketoester is cleaved by a molecule of coenzyme A to acetyl CoA and the fatty acyl CoA shortened by 2 carbons.



Scheme III.2.6

III.3. Leukotrienes, Prostaglandins, and Thromboxanes.

These three natural products are biosynthesized from arachidonic acid or from eicosatrienoic acids found in animals. Representative examples are listed in Figure III.3.1. The so-called arachidonic acid cascade is the most well studied and most important. From this acid, two enzyme systems determine the production of all prostanoids; the lipoxygenase leads to leukotrienes and the cyclooxygenase affords prostaglandins and thromboxanes (Scheme III.3.1).

Let us see the two different pathways separately (interested readers can consult *Ang. Chem. Int. Engl.* **1983**, *22*, 805 and 858 and *Chem & Eng. News* **1989**, Oct 16 issue).

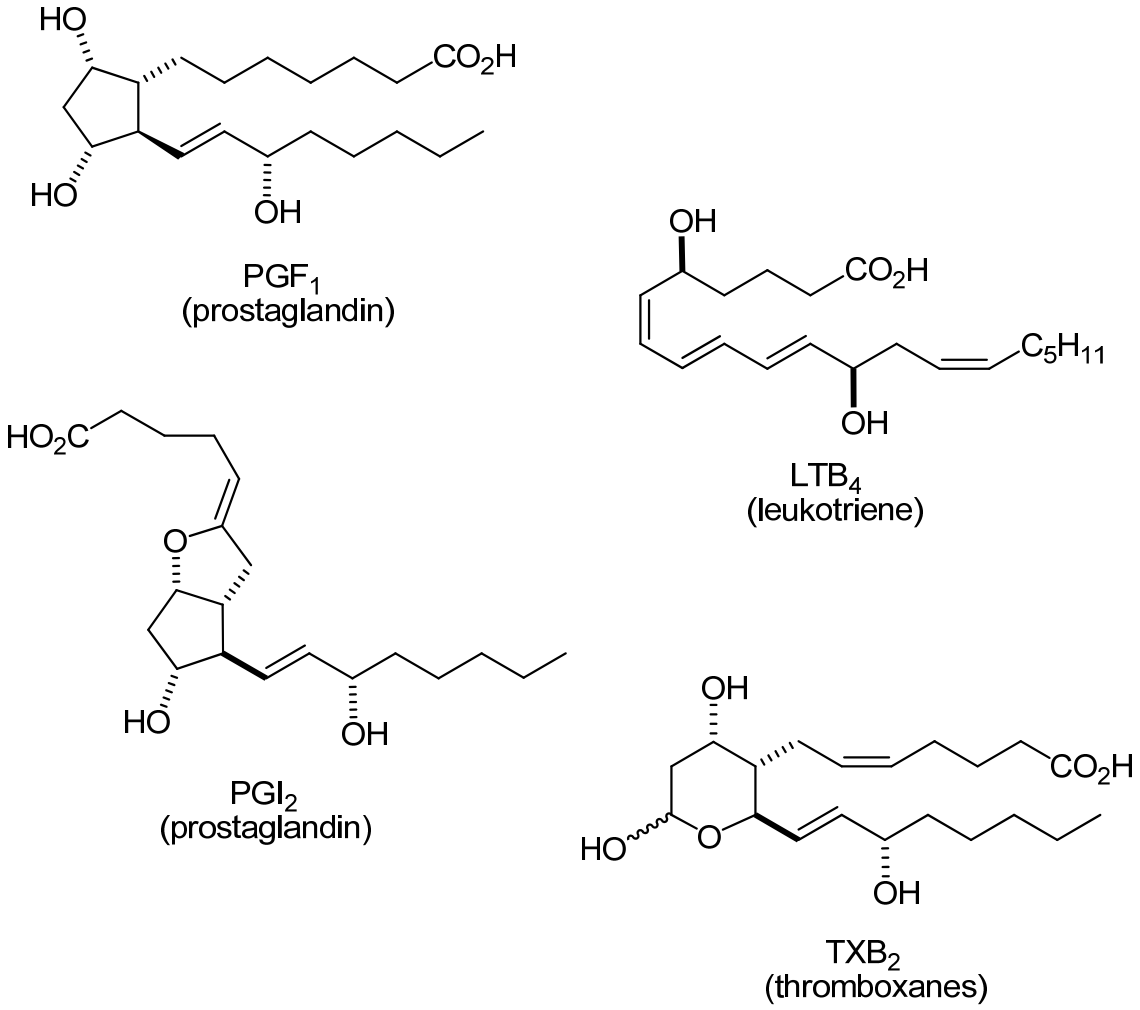


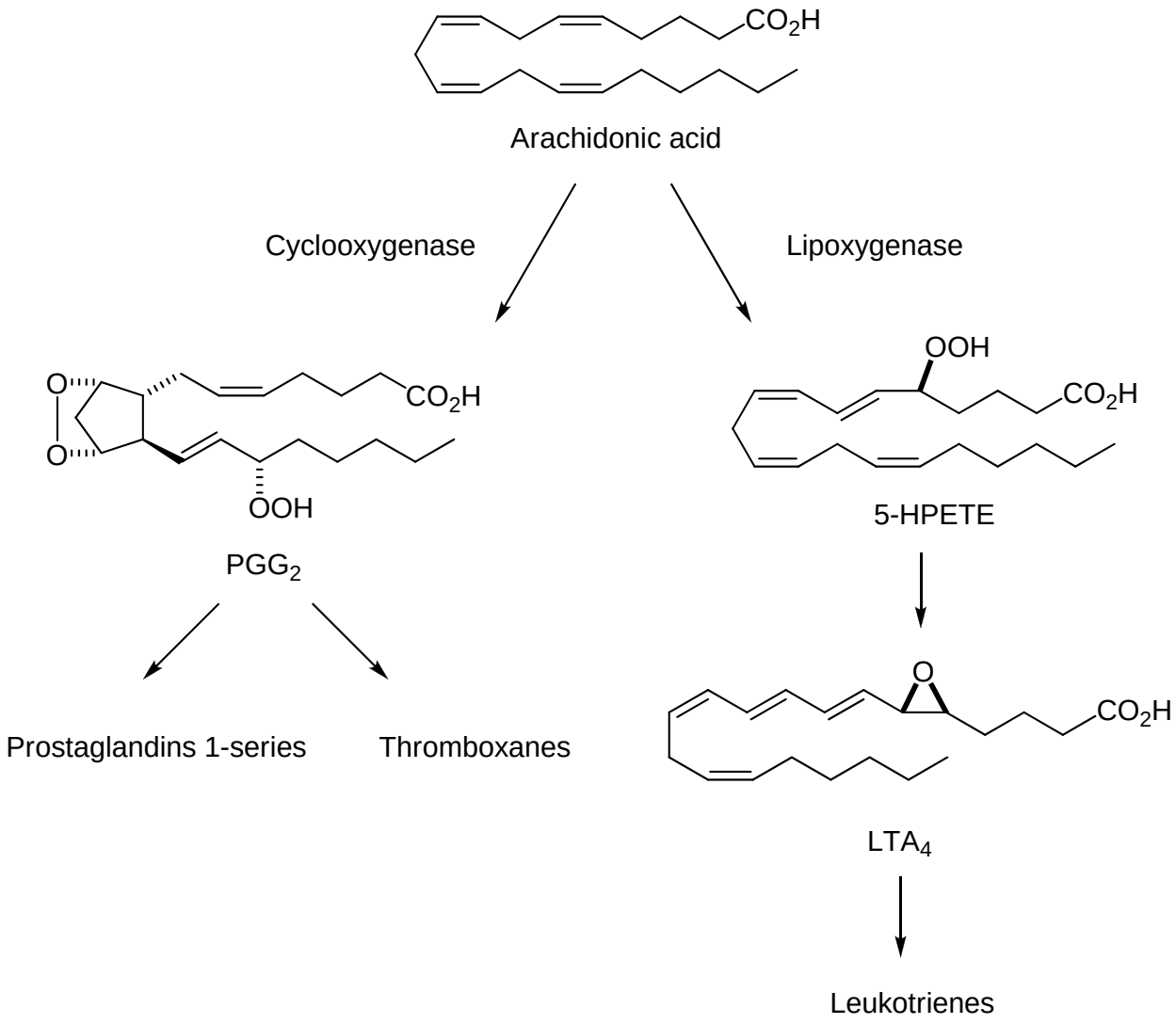
Figure III.3.1

III.3.1. Lipoxygenase Pathway and the Leukotrienes.

The arachidonic acid released from cellular lipids into the cellular milieu is submitted to the action of 5-lypoxygenase which catalyzes the oxidation of position 5 by oxygen to form the hydroperoxy eicosatetraenoic acid (HPETE) (Scheme III.3.1). Leukocytes (white blood cells) can metabolize arachidonic acid to produce a series of eicosatetraenoic or ecosatrienoic acids (the term leukotriene was coined from the discovery of the latter metabolites). The arachidonic acid metabolites produced are named as LT (leukotrienes) followed by a letter from A to E identifying the 5 types of metabolites so far isolated and a subscript "number" indicate the number of double bonds in the active leukotriene (Figure III.3.2, top). HPETE was isolated and characterized and shown to be converted in vivo to a series of leukotrienes via an intermediate that could not be isolated. With appropriate labelling studies it was found that this unstable intermediate was the epoxide LTA_4 (Scheme III.3.2). It is the pro-R hydrogen of position 10 in HPETE that is lost in its conversion to LTA_4. The fate of LTA_4 is then dictated by the different enzymes. LTA_4 can suffer nucleophilic attack either by water at C_{12} or by the protein glutathione at C_6 to afford LTB_4 or LTC_4 respectively. LTC_4 is further converted to LTD_4 and LTE_4 by successive hydrolysis of the amide bonds in the protein. A more complete arachidonic acid cascade is shown in Scheme III.3.3 (Taken in part from N.A. Nelson et al. *C&EN* **1982**, August 16, 30). The nomenclature of the leukotrienes is shown in Figure III.3.2

The leukotrienes were first detected in 1938 and found to have a contracting effect on smooth muscles. This was before modern chemical analysis permitted their isolation and structural characterization. Like the prostaglandins, they were shown to be derived from arachidonic acid. It was found that the contractions of intestine muscle cells (jejunum) were slow at first and increased with repeated contact with the substances. Thus, they were called the Slow Reacting Substances (SRS). Later, the same substances were found to be present in experiments with allergy causing agents. Anaphylaxis is an hypersensitivity caused by repeated contact with an allergen. The SRS were thought to be, at least in part, responsible for the ensuing exaggerated immune response. Leukotrienes are active at very low concentration in various immune response functions. LTB_4 is an extremely potent chemotactic factor (i.e. it causes the migration of cells by a concentration gradient) and promotes adhesion and aggregation of leukocytes to areas of inflammation. Certain allergens lead to anomalies in their biosynthesis and thus to an overreaction of the immune system. For that reason intense research of inhibitors to leukotriene synthesis was being carried out at pharmaceutical research centers such as Merck-Frosst in Montreal, in the hope of controlling allergic response of the immune system. Prostaglandins are also involved in allergies such as asthma. High concentrations of some prostaglandins cause inflammation. Corticosteroids are anti-inflammatory agents which function by blocking the

release of arachidonic acid from the phospholipids, thus shutting off both the lypoxygenase (LT's) and cyclooxygenase (PG's) branch of the cascade. Other non-steroidal anti-inflammatory agents, such as aspirin, block only the cyclooxygenase pathway and are thus less effective.



Scheme III.3.1

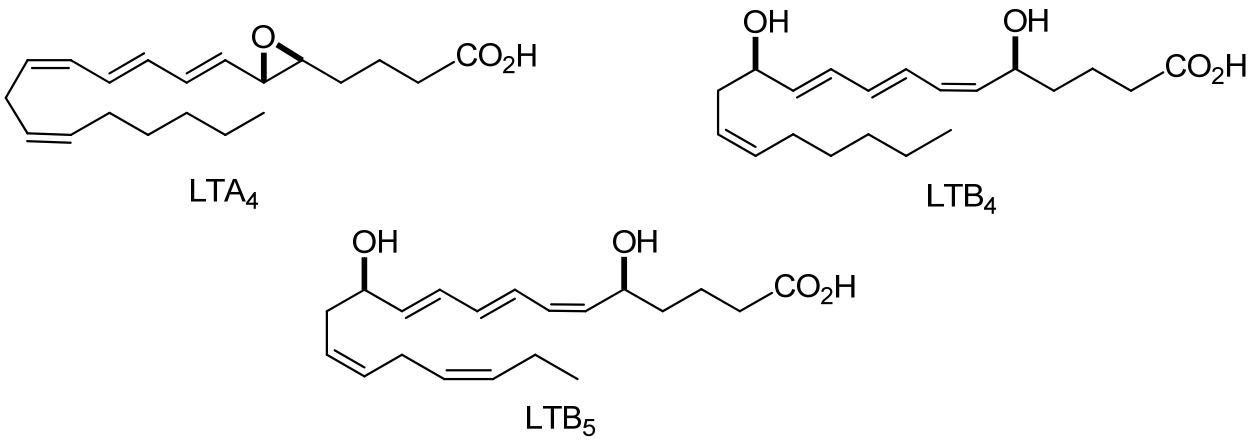
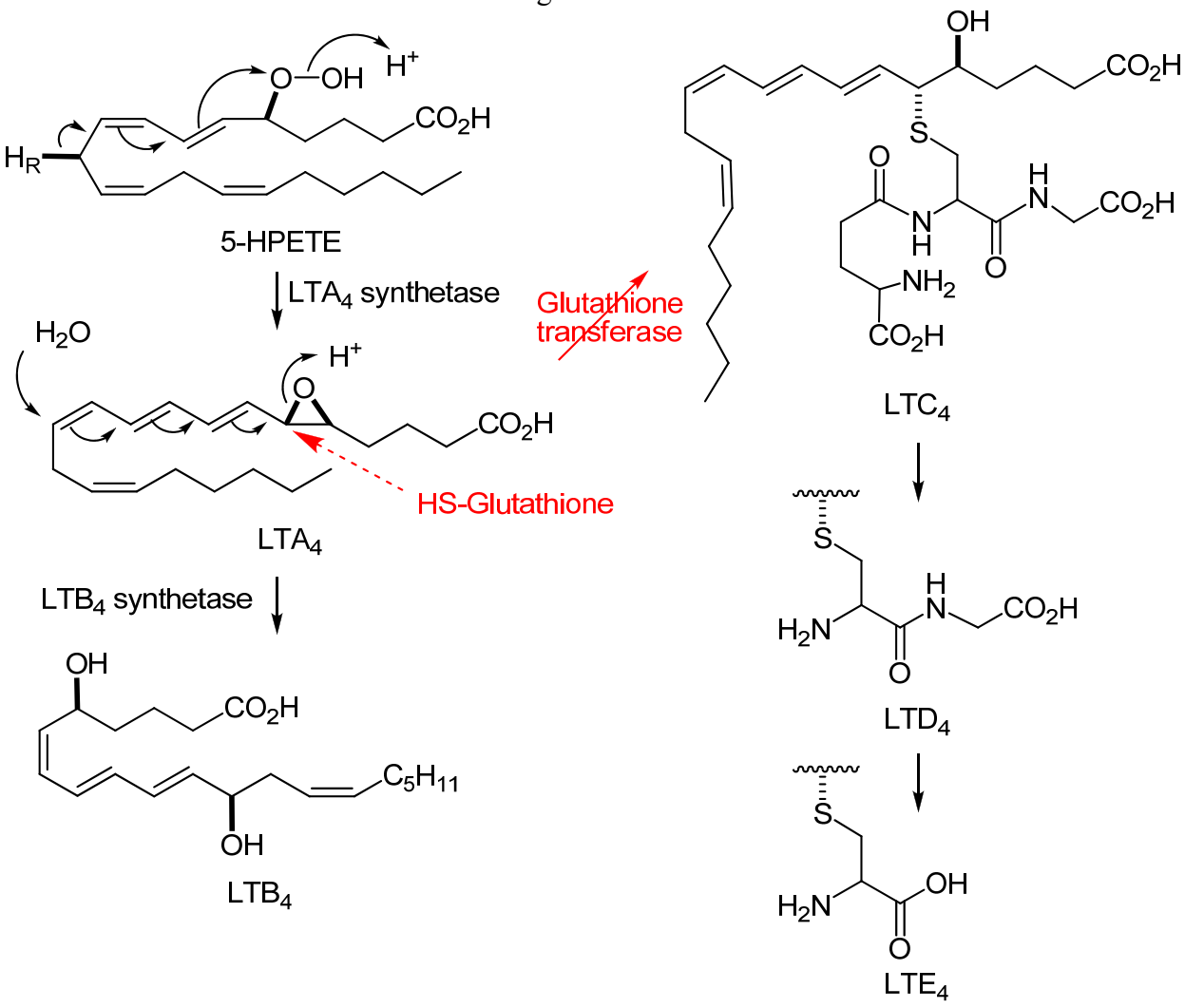
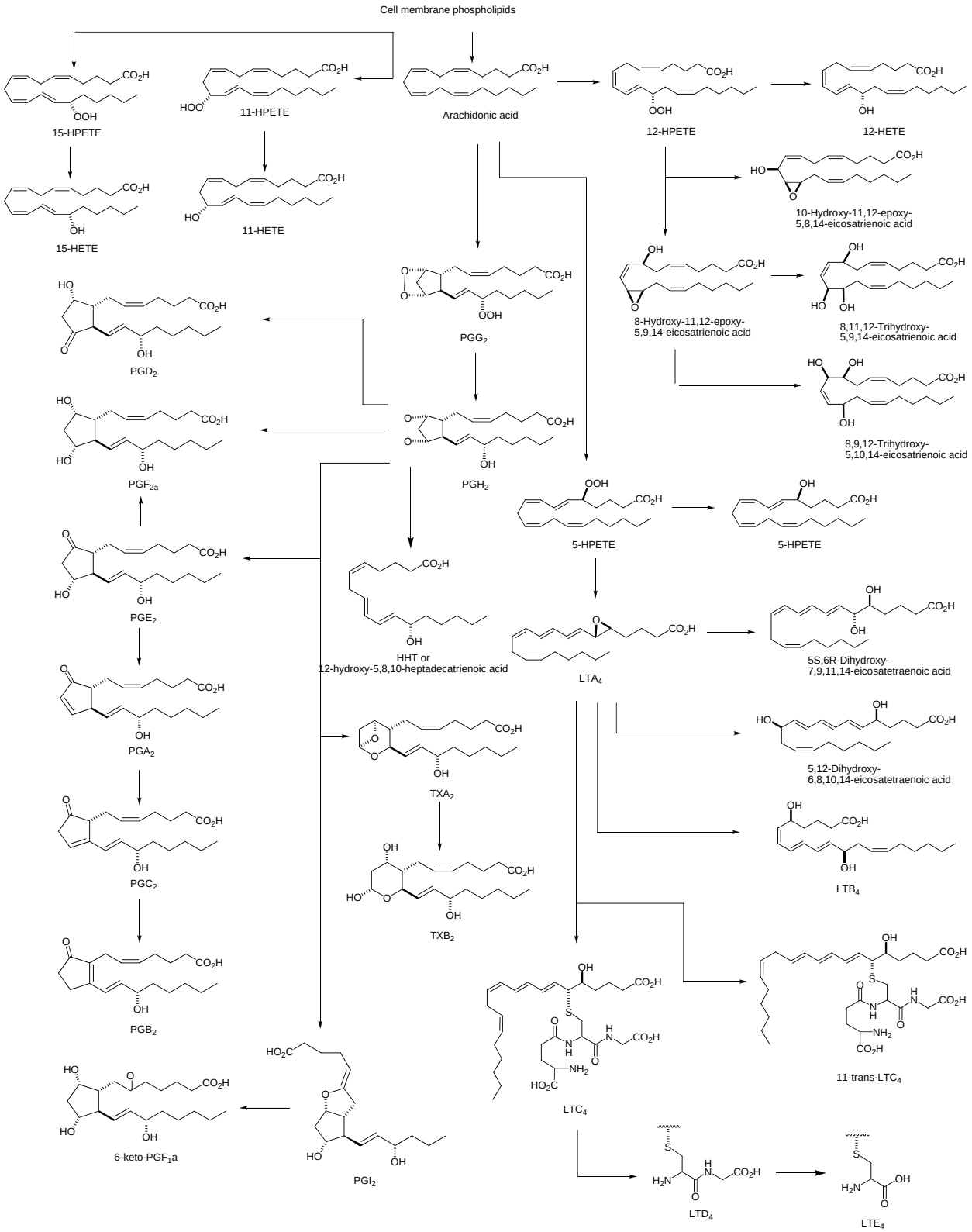


Figure III.3.2



Scheme III.3.2



Scheme III.3.3

III.3.2. Cyclooxygenase Pathway and the Prostanoids.

Prostaglandins were first encountered in 1930 (before leukotrienes) when two gynecologists found that an unidentified substance in human semen would cause smooth muscle contraction (i.e. muscle of the uterus, blood vessels, spleen). These substances were later found to lower blood pressure of animals and the term prostaglandins was coined from the fact that it was believed that the prostate gland of animals were producing them. The prostaglandins are not encountered in the plant kingdom (with only one exception) but are relatively widely distributed in animals. Their nomenclature is depicted in Figure III.3.3.

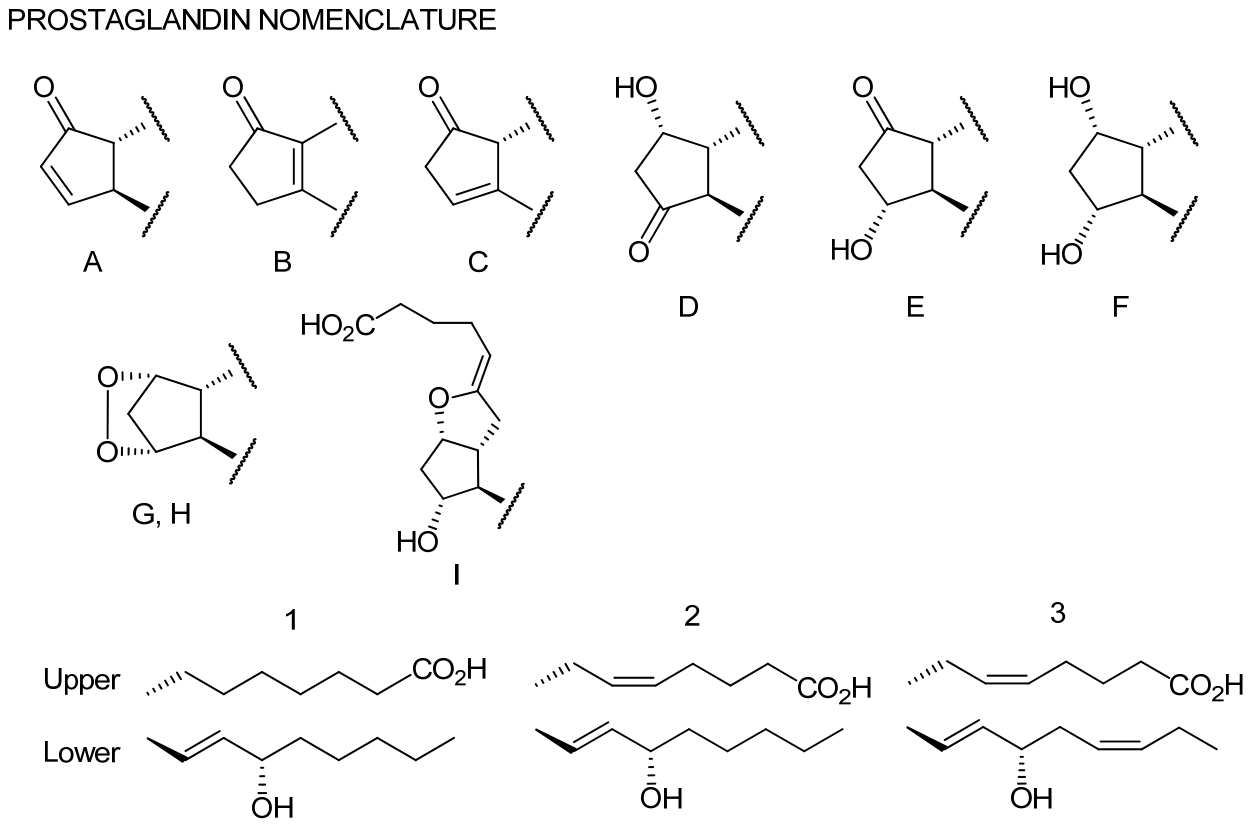
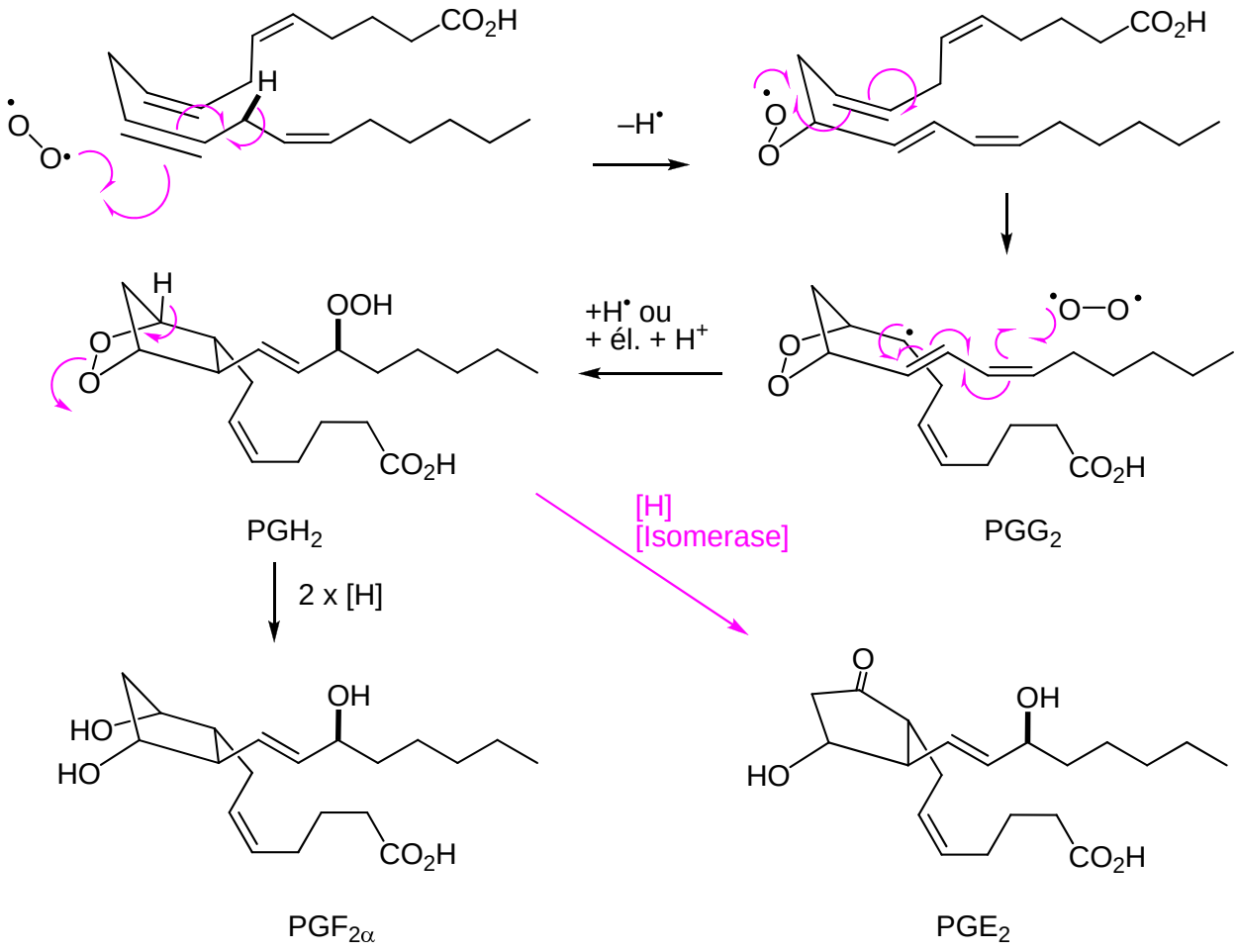


Figure III.3.3

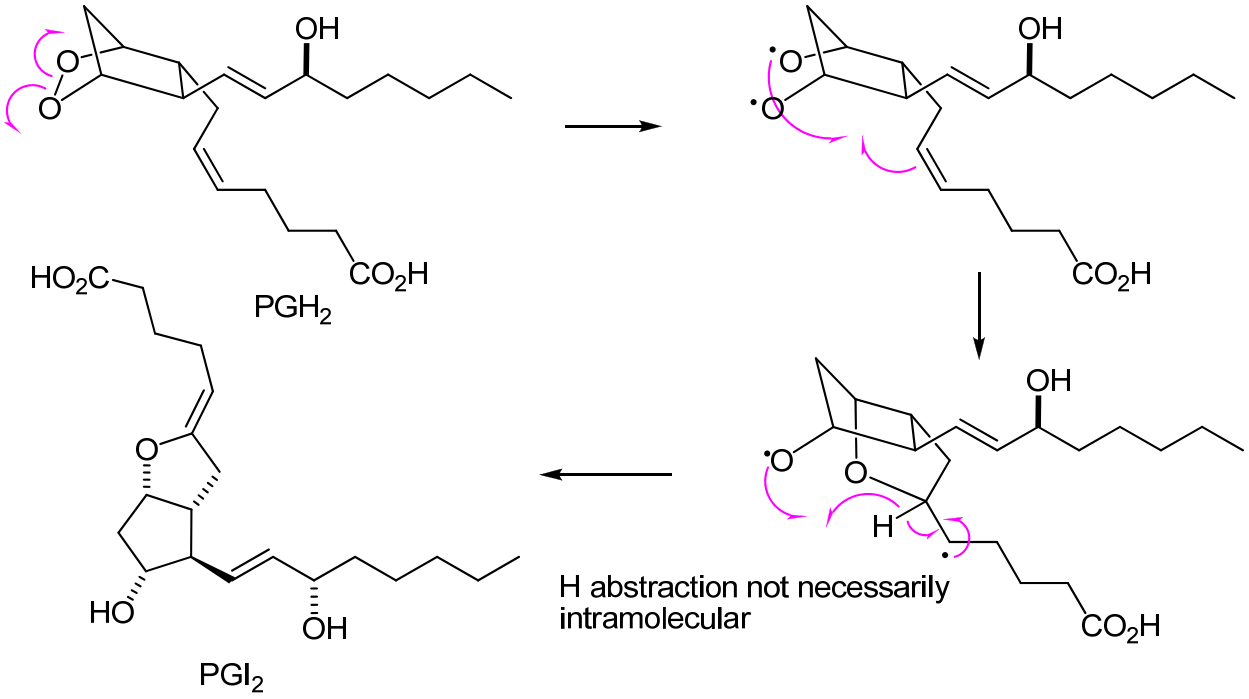
In the case of prostanoids, arachidonic acid is oxidized to PGG_2 by the action of oxygen catalyzed by the enzyme cyclooxygenase. The reaction proceeds via three peroxy-radical intermediates and is stereospecific as shown in Scheme III.3.4. In the first step, molecular oxygen (possibly the diradical triplet form, •O-O•, which is the ground state form of oxygen) is attacked by the C_{11}-C_{12} double bond. Hydrogen radical is lost to the enzyme in the process.

Then a cyclization cascade takes place which is initiated by the peroxy-radical and terminated by addition of molecular oxygen to yield the hydroperoxide PGG_2 which is reduced to the alcohol (PGH_2) and subsequent reduction of the peroxo-bridge leads to $PGF_{2\alpha}$ or isomerization of the peroxo-bridge leads to PGE_2.



Scheme III.3.4

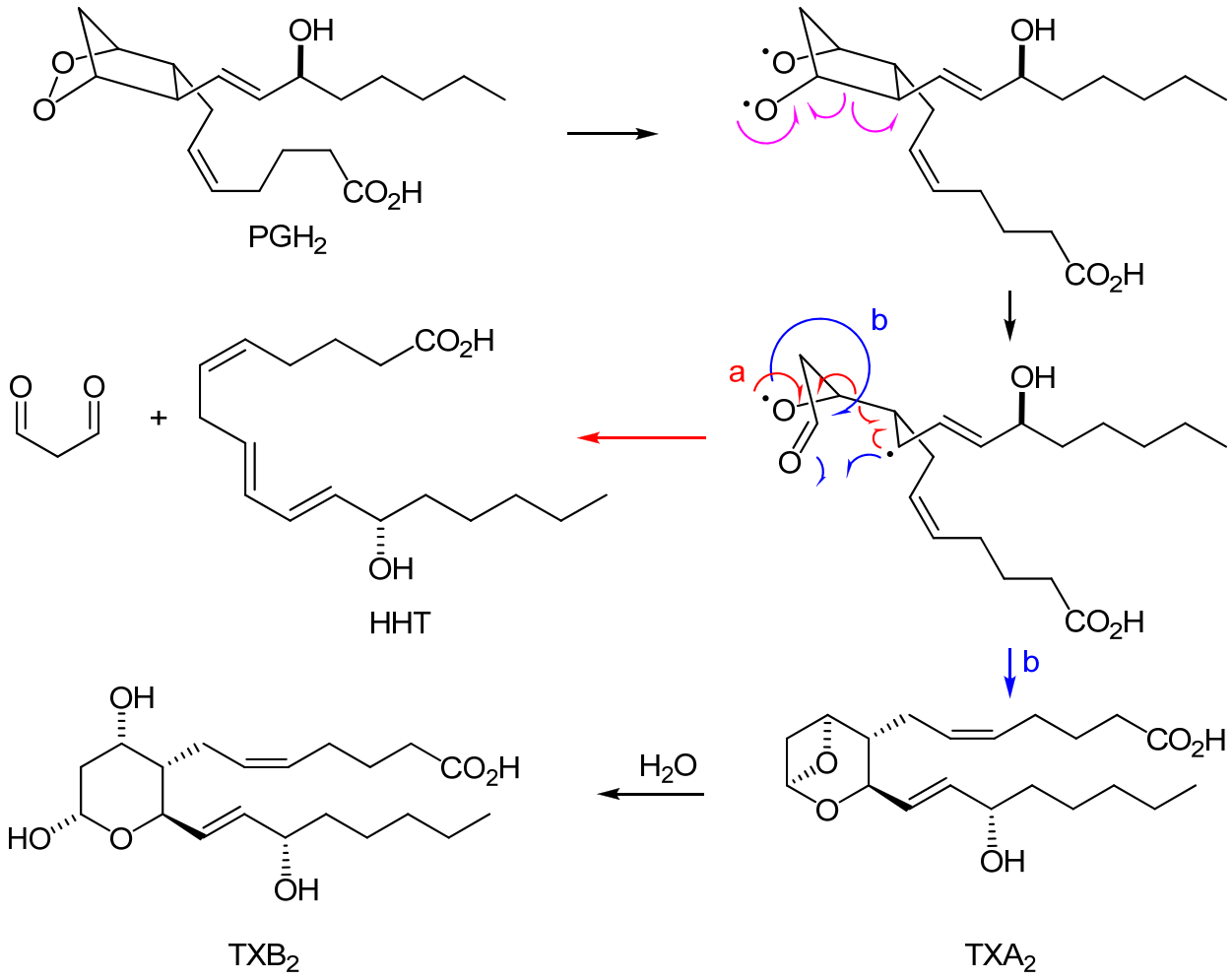
There are other alternate modes of opening of the peroxo-bridge as depicted in Scheme III.3.5 and III.3.6. These lead to the production PGI_2, of thromboxanes or 12-L-hydroxy-5,8,10-heptadecatrienoic acid (HHT). If homolytic opening of the peroxo-bridge is followed by reaction with the C_5-C_6 double bond, a tetrahydrofuran is formed. Abstraction of hydrogen can occur intramolecularly as shown in Scheme III.3.4 but this is not necessarily the case and hydrogen abtraction as well as loss of hydrogen could occur with the involvement of FAD or FADH.



Scheme III.3.5

If, on the other hand, a fragmentation occurs has shown in Scheme III.3.5, then the formation of thromboxanes or HHT ensues. TXA_2 is formed when the carbon and oxygen radicals (blue arrows) bind to form a rather strained acetal function. The latter is easily opened with water to transform TXA_2 into TXB_2. On the other hand, loss of malonaldehyde leads to the formation of a C_{17} acid called 12-hydroxy-6,8,10-heptadecatrienoic acid or HHT. Thromboxane nomenclature is depicted in Figure III.3.4.

Figure III.3.4



Scheme III.3.6

Prostaglandins have a wide range of biological activities. In the brain they are involved in temperature and circulation control and neuromodulation. PGE_2 and $PGF_{2\alpha}$ are used to induce **abortion** during the first / second trimesters of pregnancy and to induce **labour** at full term. The PGE's are involved in the control of blood pressure, and are used as antihypertensive agents. They are also effective in decreasing gastric secretion and the ability of aspirin to inhibit PGE biosynthesis may explain why it promotes stomach ulcers. PGA_1 and the PGE's are agents for the potential treatment of gastric ulcers. An orally active gastric antisecretory agent "deprostil" was developed for that purpose (Figure III.3.5). PGE_1 and PGD_2 are implicated in the inhibition of the aggregation of blood platelets, presumably via an increase in the concentration of cyclic AMP. PGG_2 and PGH_2 have the opposite effect of inducing platelet aggregation and in addition induce smooth muscle contraction.

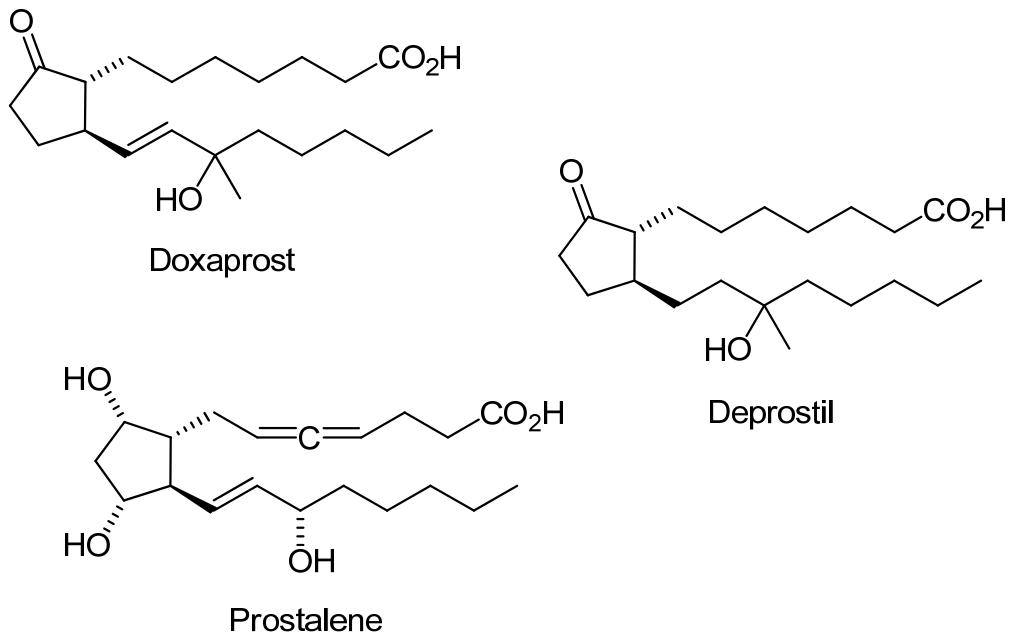
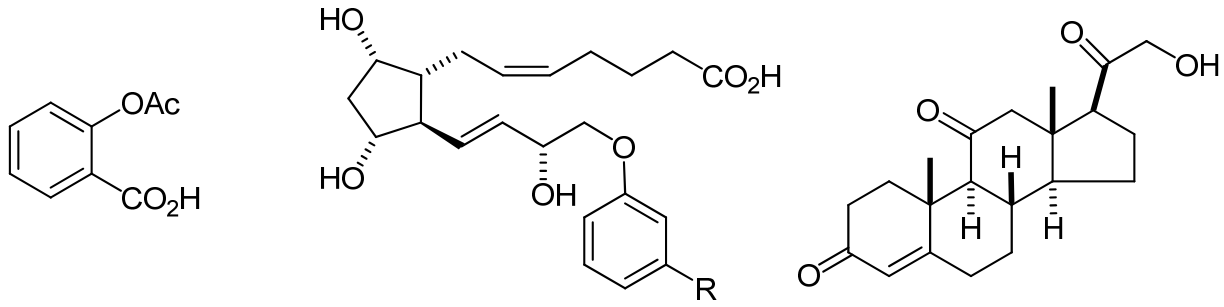


Figure III.3.5

Asthma, in particular, is closely related to prostaglandin biosynthesis because of their vasoconstriction ability. In allergen-provoked asthma attacks, $PGF_{2\alpha}$, a strong bronchoconstrictor, is released in abnormal quantities. PGE_1 and PGE_2 are bronchodilators and can be used as an emergency treatment. Other more orally active drugs have been developed for asthma control, such as doxaprost and prostalene (Figure III.3.6). Other drugs inhibit prostaglandin biosynthesis by competing for the proper enzyme receptors, for example, indomethacin inhibits cyclooxygenase. An abnormally high concentration of prostaglandins may cause inflammation of the joints (rheumatoid arthritis). Corticosteroids induce the synthesis of lipomodulin or macrocortin which inhibits the release of arachidonic acid from phospholipids. Thus there is no available starting material for the prostaglandins and it results in the **relief** of inflammation (pain). The reason why certain PG's cause inflammation when in excess is thought to be related to their peroxy-radicals which can cause damage to cartilage. Evidence of this theory is provided by the fact that radical scavengers are known to relieve arthritic pain. Table III.3.1 lists the biological activities of various prostaglandins. Treatments for the adverse effects of high concentration of prostaglandins are numerous. Aspirin will block the enzyme cyclooxygenase and thus prostaglandin biosynthesis altogether. As a result, it will affect your stomach by lowering the concentration of PGE's but will ease arthritic pain (anti-inflammatory) by lowering concentration of PG's. Corticosteroids will prevent arachidonic acid from being released from the phospholipids by stimulating the synthesis of lipomodulin and macrocortin (inhibitors of the release of arachidonic acid). Thus both LT's and PG's biosynthesis will be shut

off. Figure III.3.5 shows the structures of three therapeutic agents : doxaprost is an orally active bronchodilator; deprostil is an orally active gastric anti-secretory agent; prostalene is an hypotensive agent. Luteolytic agents (labour inducer): fluprostenol and chloprostenol (several hundred times more potent than PGF2a).



Aspirin (acetyl salicylic acid)

Fluoprostenol R = CF_3
Chloprostenol R = Cl

Cortisone (corticosteroid)

Figure III.3.6

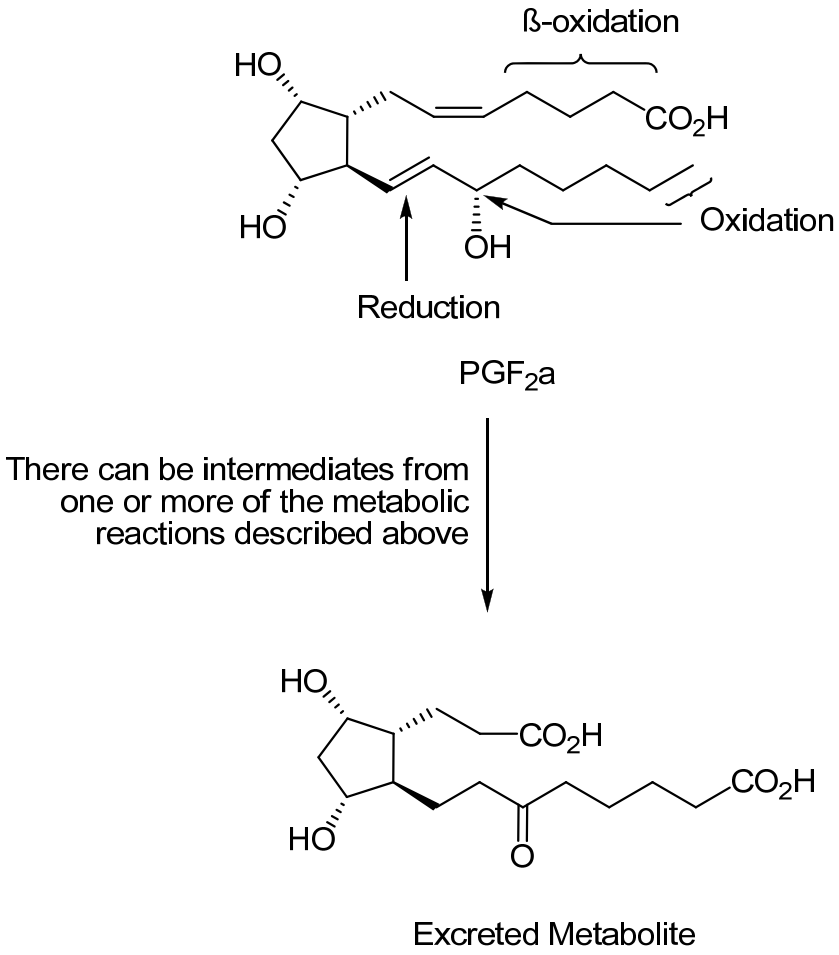
Table III.3.1. Biological Activities of Prostaglandins

PGA_1	Anti-secretory activity (potential treatment of gastric ulcers).
PGD_2	Inhibition of aggregation of blood platelets (prevents clotting).
PGE_1	Bronchodilator.
	Inhibition of aggregation of blood platelets (prevents clotting).
	Vasodilator (lowers blood pressure).
PGE_2	Anti-secretory activity (potential treatment of gastric ulcers).
	Bronchodilator (emergency treatment for asthma).
	Labour / abortion inducer (end and first term of pregnancy respectively).
	Vasodilator (lowers blood pressure).
$PGF_{2\alpha}$	Bronchoconstrictor (causes asthma).
	Labour / abortion inducer (end and first term of pregnancy respectively).
PGG_2	Induces blood platelet aggregation (blood clotting).
PGH_2	Induces blood platelet aggregation (blood clotting).
PG's (general)	High concentration causes inflammation (arthritis and others).

As investigations of prostaglandins in biological systems gained momentum, it soon was realized that the natural compounds were relatively short-lived in the body. Rapid metabolism

meant a short duration of activity for many of the desirable biological properties of the prostaglandins. Studies leading to the understanding of this metabolism were crucial for the design of active analogs with longer life span. PGE's and PGF's are subject to three major modes of attack by metabolizing enzymes. These are much the same in all species that have been studied, including humans. Occurring most rapidly is dehydrogenation at C-15 followed by reduction of the 13,14-double bond to give the corresponding 13,14-dihydro-15-oxo product (Scheme III.3.7). Also occurring rapidly is ß-oxidation of the carboxylic acid side chain (see section III.2.2). This process removes sequentially the C_1-C_2 and C_3-C_4 of the side chain. The final metabolic change occurs at the C_{20} or C_{19} positions to give a C_{20} carboxylic acid or C_{19} hydroxy metabolite, respectively.

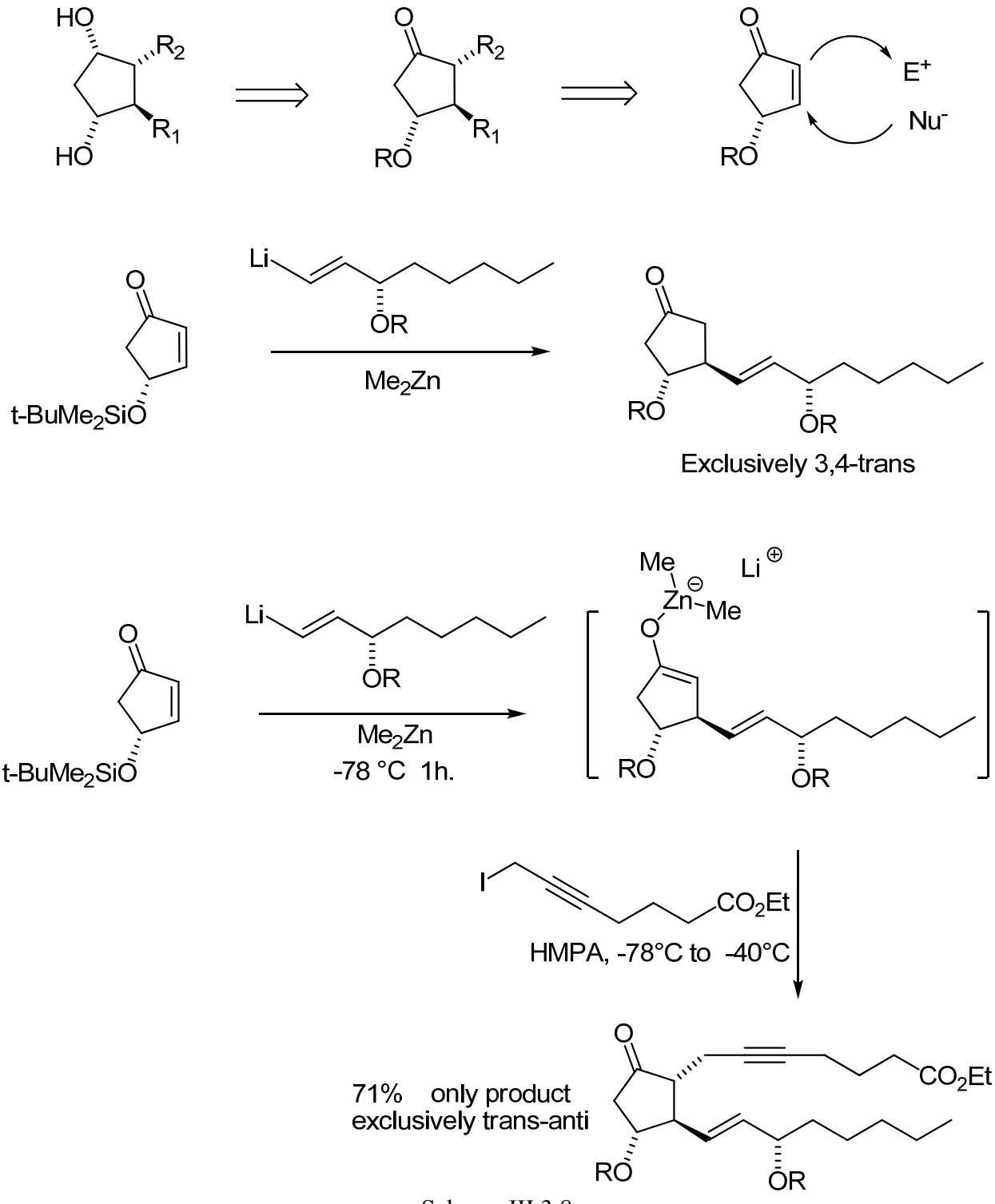
Rapid Metabolism of Prostaglandins



Scheme III.3.7

III.3.3. Total Synthesis of Prostanoids.

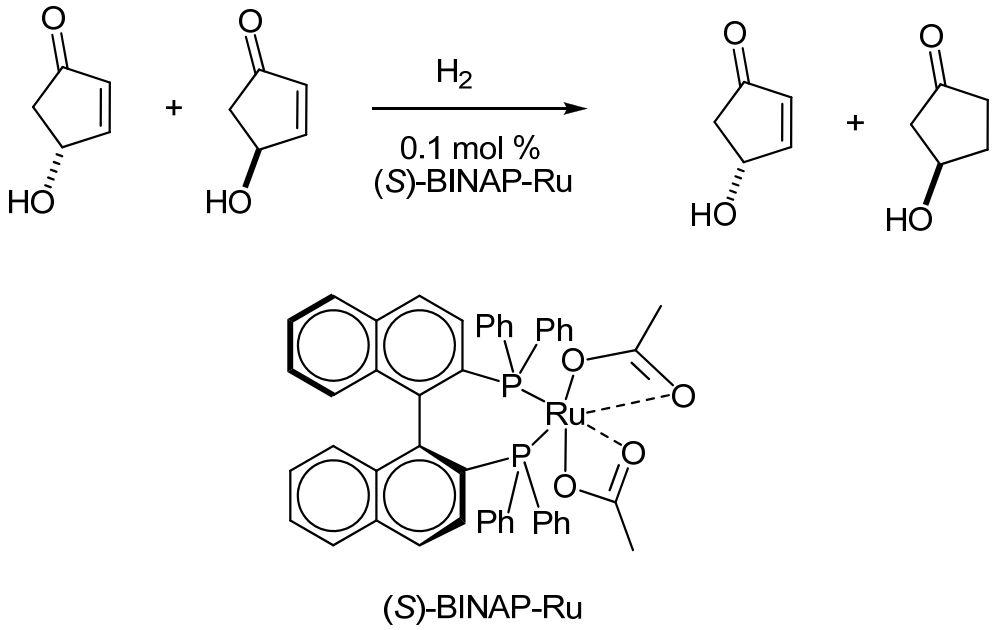
The low availability of prostaglandins made their total synthesis very attractive for biological testing. The major difficulty in the synthesis is the introduction of the four chiral centers in the cyclopentane ring. The best approach to date proceeds via a tandem 1,4-addition-alkylation reaction sequence onto a cyclopentanone ring as shown in Scheme III.3.8. The choice of the proper nucleophile and electrophile is of course crucial. Carbanions having lithium or magnesium as the counter ion (Grignard reagents) add in a 1,2-fashion to α,ß-unsaturated ketones. However, when copper or zinc is used as the counter ion, the addition proceeds in a Michael fashion (1,4-) only. In addition, the resulting copper or zinc enolate is stable at low temperature and can be mono-alkylated with reactive alkyl halides such as allylic iodides. The incoming nucleophile (zinc carbanion) is directed to the α-face of the molecule by the shielding of the ß-face by the neighboring ether group. Furthermore, the alkylation of the resulting enolate proceeds anti to the newly introduced chain to give the product shown having the proper 2,3-trans-3,4-anti stereochemistry (Scheme III.3.8). This approach is powerful so long as the desired nucleophile chain and electrophile chain are both available by synthesis (R. Noyori et al. *J. Org. Chem.* **1989**, *54*, 1787).



Scheme III.3.8

In addition, the starting cyclopentenone should be optically pure to lead to non-racemic prostaglandins. The latter is obtained most efficiently by "kinetic resolution" of the racemic mixture of 4-hydroxycyclopent-2-en-1-one. Using a chiral catalyst complex of ruthenium (*S*-

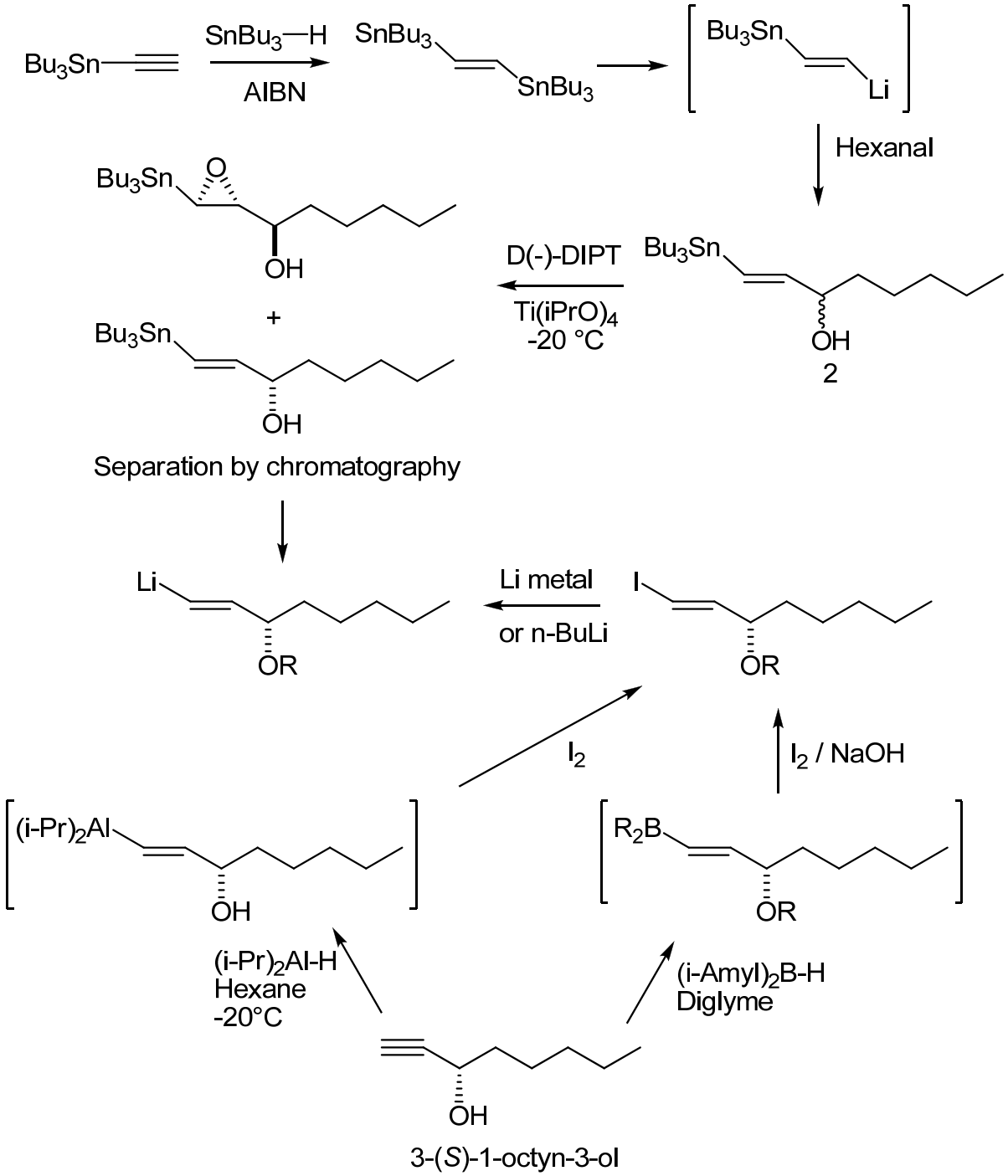
BINAP-Ru), the hydrogenation of the undesired *S*-hydroxycyclopentenone is much faster than that of the desired *R*-hydroxycyclopentenone (Scheme III.3.9). The reduced product is separated from the unreduced one after the alcohol is protected as the *t*-butyldimethylsilyl ether. At 68% conversion, one reduction gives the desired product in 98% ee (enantiomeric excess = ([R] - [S])/([R] + [S]) x 100). That means the *S*-product gets reduced ~ 11 times faster than the *R*-product. This is called a "kinetic resolution" where the higher reactivity of one enantiomer is used to separate it from the other (see R. Noyori et al. *J. Org. Chem.* **1988**, *53*, 710).



Scheme III.3.9

The required open chain organolithium reagent (nucleophile in the tandem reaction) was synthesized using modern chemistry. We are familiar with the preparation of Grignard reagents from the corresponding halide. Lithium reagents can also be formed from the corresponding halide but the reaction can be sluggish. A modern alternative to using the halide is trialkyltin chemistry. The trialkyltin derivative on a double bond will exchange the tin metal for a lithium metal when treated with a commercially available alkyllithium such as *n*-butyllithium (*n*-BuLi) (Scheme III.3.10, top route). Thus the synthesis of **1** starts with acetylenyltributyltin which is reacted with tributyltin hydride under free radical conditions. The resulting bis(tributyltin) ethene is transmetallated with *n*-butyllithium where one of the tributyltin moieties is replaced by lithium. Such a lithium species is highly nucleophilic and will add to hexanal. The racemic mixture of alcohol **2** is kinetically resolved by a selective Sharpless epoxidation reaction. Only the unwanted *R*-alcohol gets epoxidized rapidly and the mixture of epoxide (unwanted) and

alkene (desired) is chromatographically separated. After protecting the alcohol, the tributyltin moiety is transmetallated again using *n*-BuLi (see Y. Kitano et al. *Chem. Lett.* **1987**, 1523).



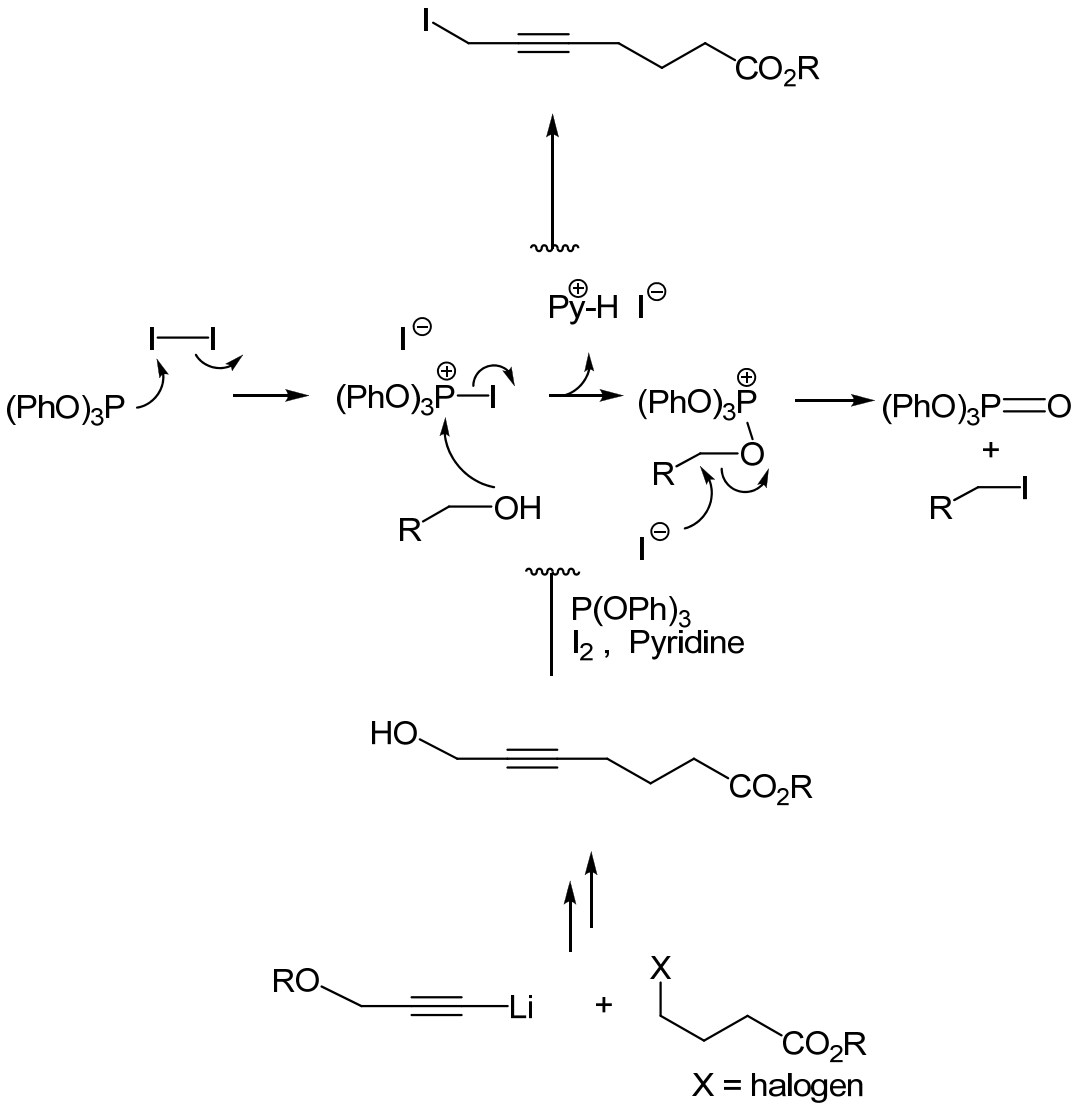
Scheme III.3.10

Another method starts with optically active 3*S*-1-octyn-3-ol. Either hydroalumination or hydroboration can be used to obtain a *trans* vinylic iodide **3** (Scheme III.3.8, bottom routes). The iodine is then replaced by lithium in a reaction analogous to the tin-lithium exchange reaction. Thus treatment of the vinyl iodide with *n*-BuLi exchange the iodine for lithium. In the reaction, iodobutane is formed. Alternatively, lithium metal can be used as effectively in a reaction analogous to the preparation of Grignard reagents (see C.J. Sih et al. *J. Am. Chem. Soc.* **1972**, *94*, 3643 and A.F. Kluge et al. *J. Am. Chem. Soc.* **1972**, *94*, 7827).

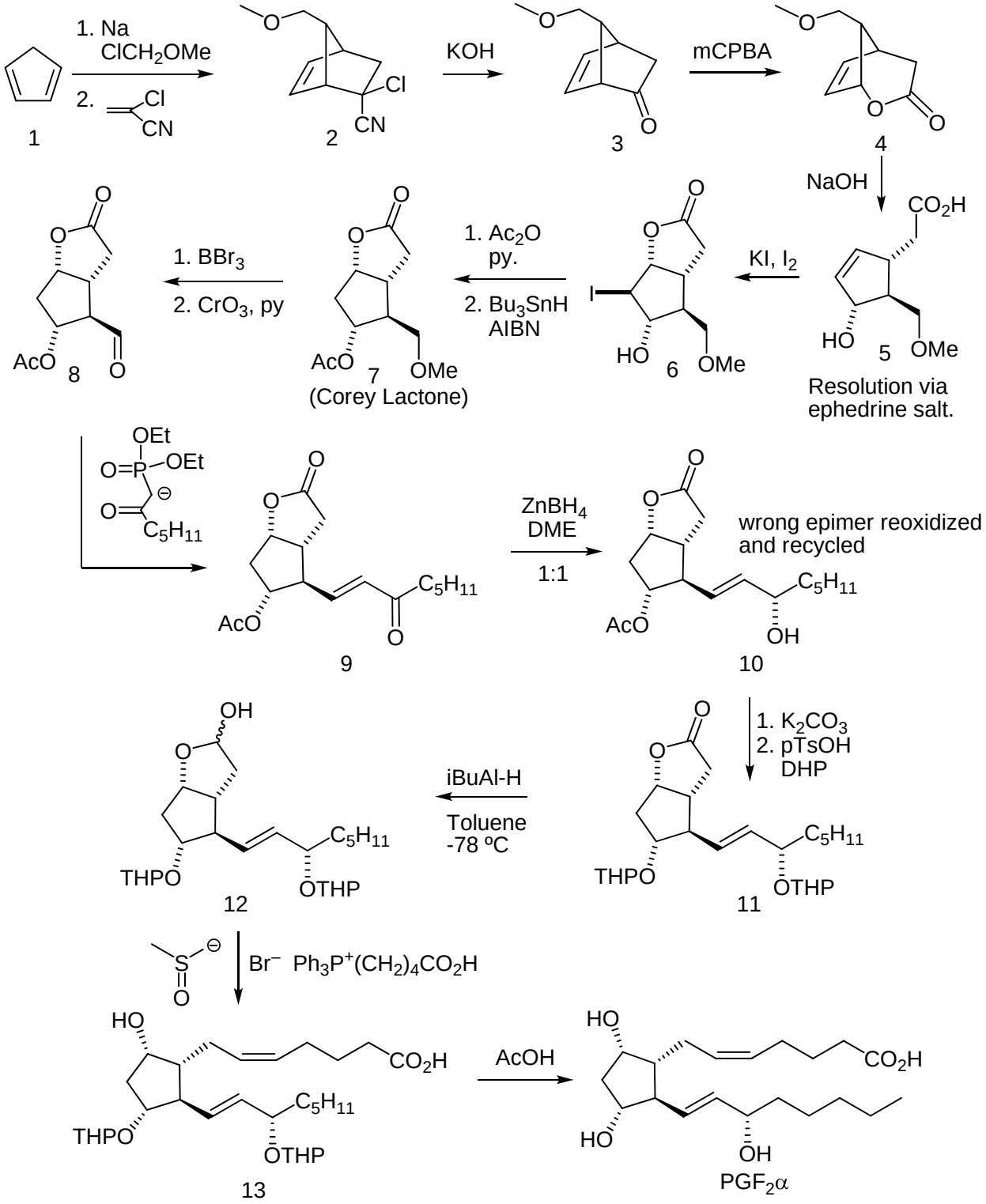
The electrophile (upper chain) was synthesized from the iodination of alcohol **1** (Scheme III.3.11). This alcohol comes from the reaction of the dianion of propargyl alcohol with methyl bromobutanoate. The iodination reaction proceeds by the formation of the phosphonium salt and further reaction with the alcohol (Scheme III.3.11). Displacement of triphenoxyphosphine oxide by iodide ion is favored by the high affinity of phosphorus for oxygen (see M. Suzuki et al. *J. Am. Chem. Soc.* **1985**, *107*, 3348).

A different approach that is still extensively used is to make bicyclic lactones like **7** dubbed the "Corey Lactone" (scheme III.3.12). The upper chain of prostaglandins can be easily introduced from the Corey lactone which already possesses the four contiguous chiral centers. Many syntheses of the Corey lactone have been devised and the original (by Corey) in shown in Scheme III.3.12. The synthesis is stereocontrolled and leads to a racemic mixture of $PGF_{2\alpha}$. In the first step, the anion of cyclopentadiene is formed by reaction with sodium and it is reacted with the electrophile chloromethoxymethane. The cyclopentadienyl anion is aromatic. Then a Diels-Alder reaction followed by hydrolysis gives the ketone **3**. Note that the dienophile is an equivalent of ketene which is an unstable molecule (it dimerizes rapidly). A selective Baeyer-Villager and hydrolysis leads to the cyclopentene **5**. Note that three of the four chiral centers were introduced with the correct relative stereochemistry. Compound **5** is resolved via the diastereomeric salts of ephedrine. Then, an iodolactonization is effected to give **6** with the fourth chiral center having the correct stereochemistry. The iodine is removed via a radical reaction after protection of the alcohol as its acetate. Then the methyl ether is selectively cleaved with boron tribromide (S_N2 reaction proceeds faster at primary methyl group to give bromomethane). Then the resulting alcohol is oxidized using the Collins reagent and the aldehyde thus obtained is reacted with the anion of a phosphonate reagent (Horner-Wadsworth-Emmons reaction) to give olefin **10** with the *E* geometry. Stabilized phosphonates are known to give *E* olefins. Reduction with zinc borohydride is unselective and give a mixture of two epimeric alcohols. They are separated and the wrong epimer is oxidized and recycled. The acetate protecting group is removed, and the two alcohols reprotected as their tetrahydropyranyl ether (THP). Partial reduction of the lactone to the lactol is effected with diisobutylaluminium hydride and the lactol **13** is reacted with a phosphonium bromide to give olefin **14** this time having the *Z* geometry.

The success of the reaction is due to the equilibrium of the lactol and the open chain alcohol-aldehyde. The use of the dimsyl anion in dimethyl sulfoxide is also important in this reaction. Cleavage of the THP ethers with acid liberates the target molecule $PGF_{2\alpha}$.

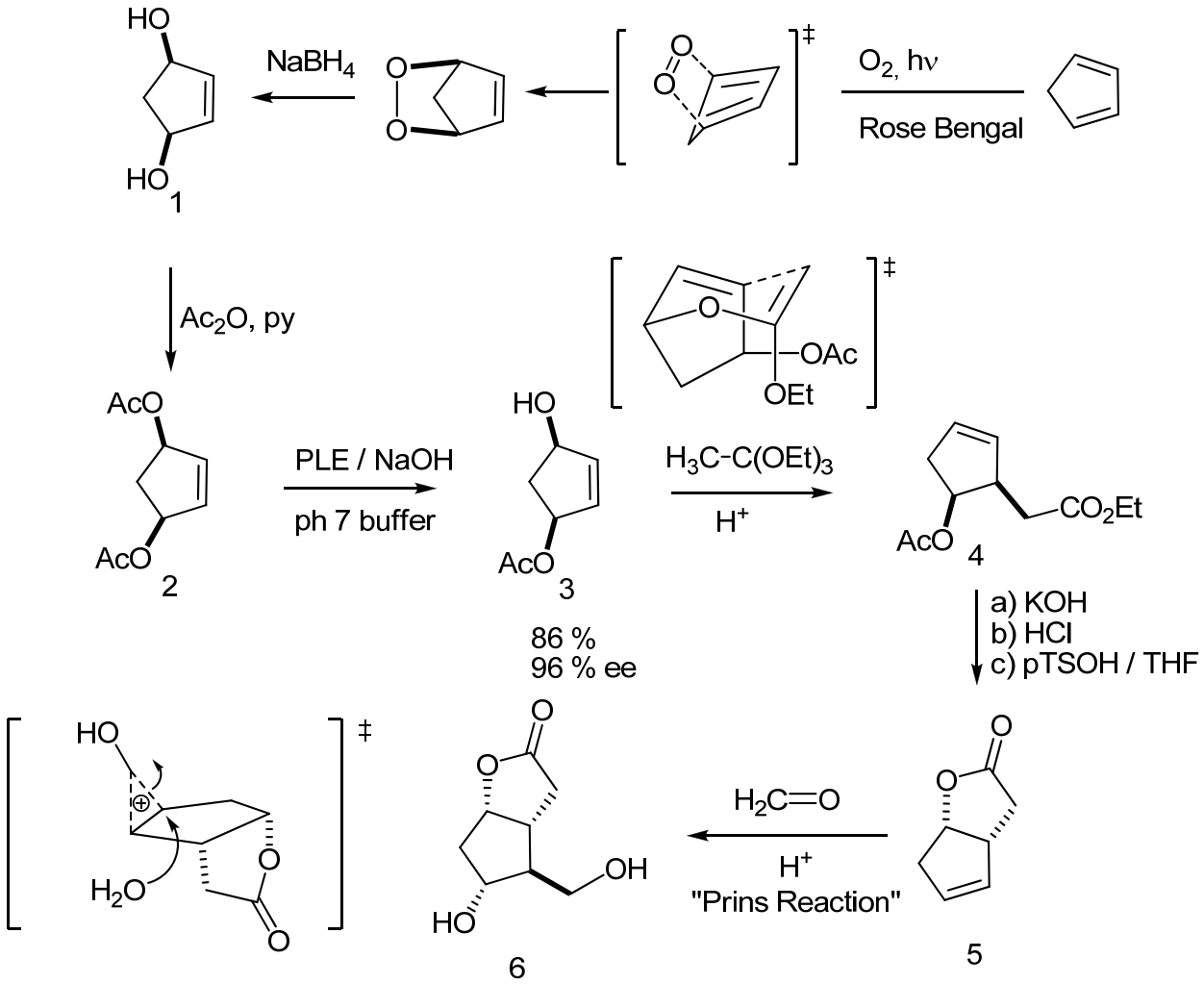


Scheme III.3.11



Scheme III.3.12

Another synthesis of the Corey lactone is summarized in Scheme III.3.13. It also starts with cyclopentadiene which is oxidized with singlet oxygen to the peroxo-bridged compound. In situ reduction gives diol **1** as non-chiral meso-type compound. After acetylation, the mixture is submitted to hydrolysis by an enzyme system. The enzyme stereospecifically hydrolyzes only one of the acetate groups to give optically pure monoacetate **3**. This one undergoes a Claisen rearrangement with acid and triethylorthoacetate. Lactonization followed by the Prins reaction gives a Corey lactone **6**. The Prins reaction gives the correct stereochemistry because of opening of the cyclopropylcation bridge from the least hindered side of the molecule (Scheme III.3.13).



Scheme III.3.13

III.4. Polyketides and Aromatics.

III.4.1. Biosynthesis of Polyketides.

In considering this large family of phenolic, polyphenolic, and macrocyclic compounds, ultimately derived from acetate units, it is not at all obvious how it was possible to imagine a biogenetic connection between the various types of structures. A look at Figure III.4.1 will convince you of that. Unlike the fatty acid derived metabolites, the polyketides suffer little or no reduction of the carbonyl group during biosynthesis. Because of that, almost all polyketides have an oxygen or equivalent oxidation level at every second carbon. It was this fact that lead J.N. Collie in 1907 to propose that several natural products of aromatic nature could be biogenetically related. As was often the case in the early days of natural product chemistry, his proposal was given little attention due to the lack of experimental proof and the lack of elucidated structures. Birch, in 1953, revived the hypothesis and proposed a viable theory of biosynthesis by which each acetate unit is linked in a head to tail manner without reduction of the carbonyl before suffering a series of aldol and/or Claisen condensation to afford the polyaromatic natural products.

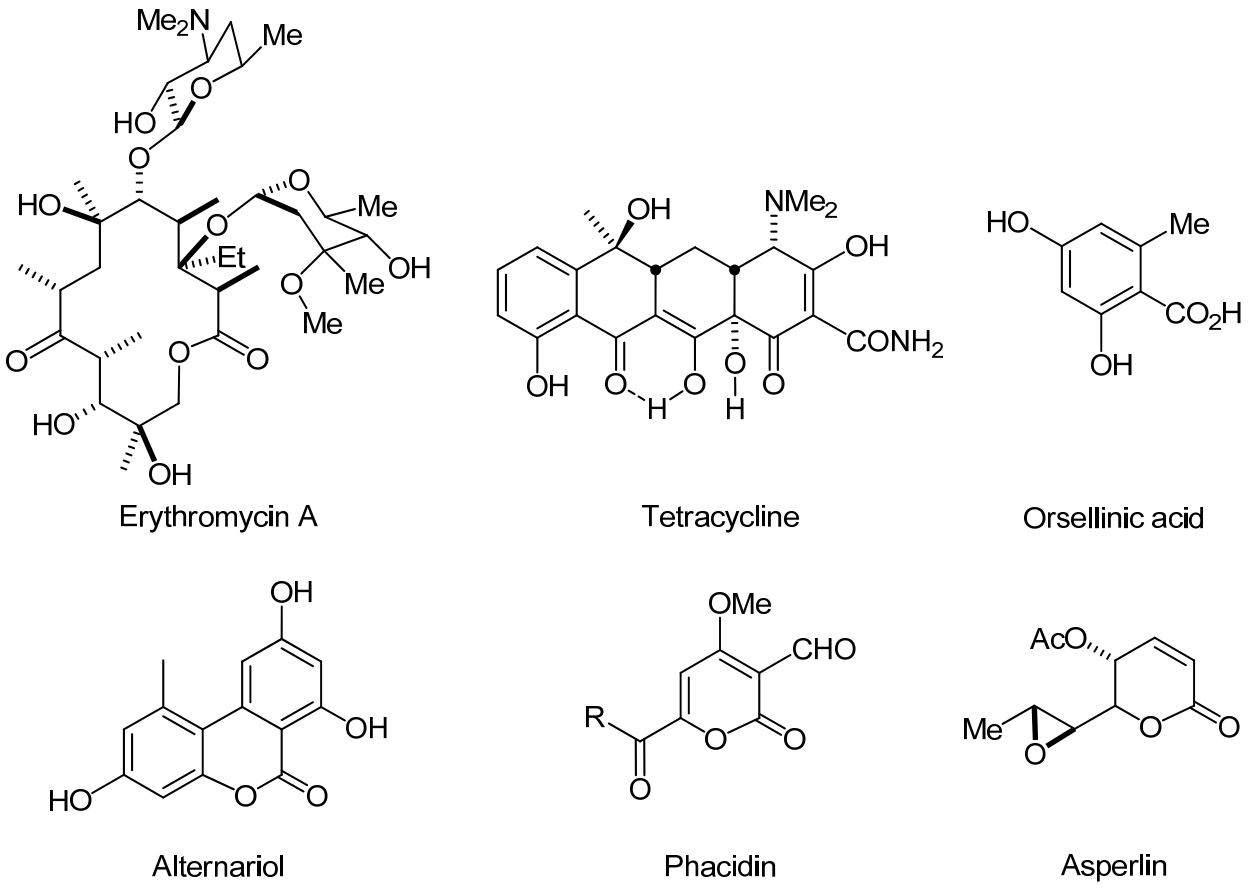
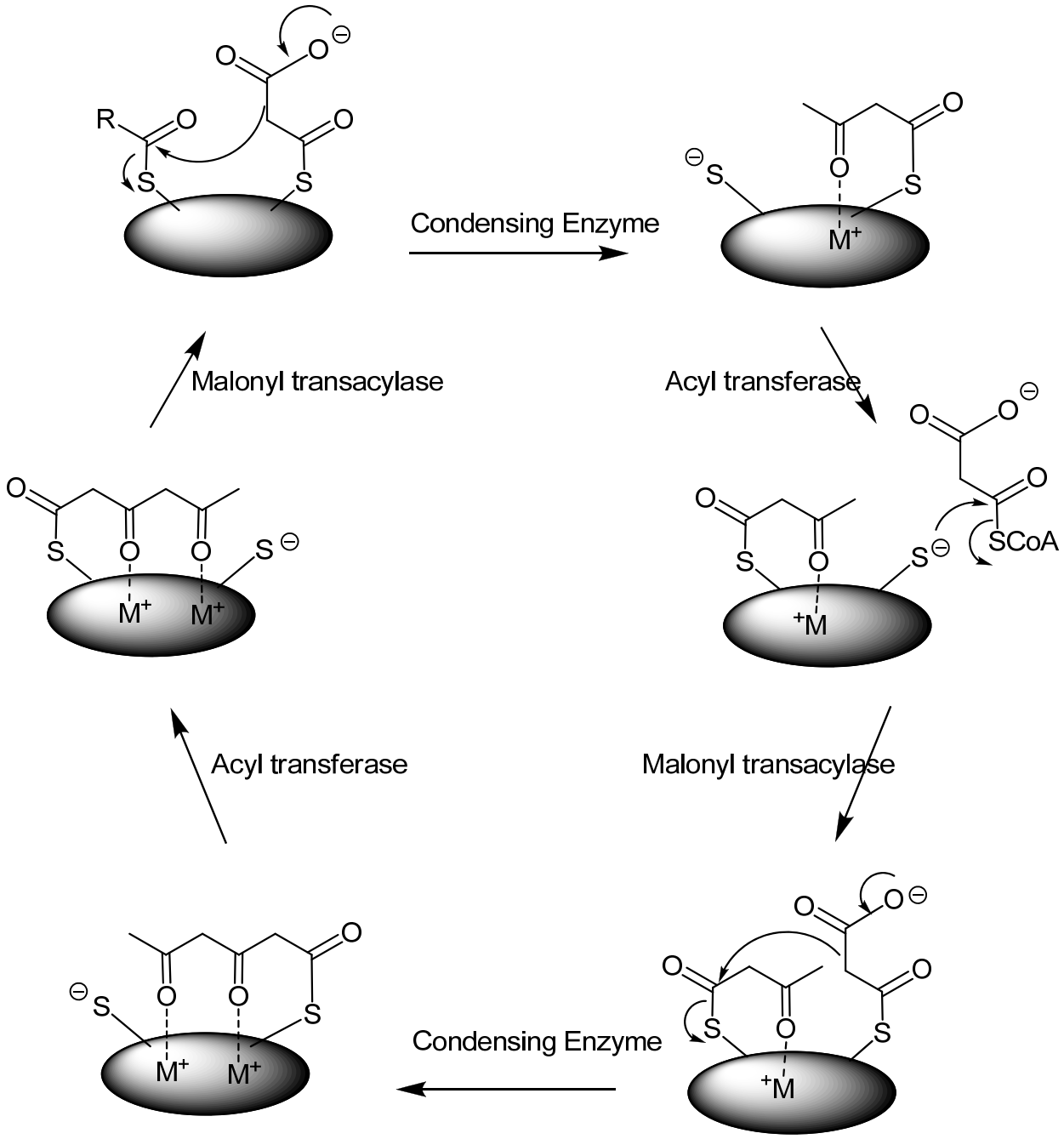


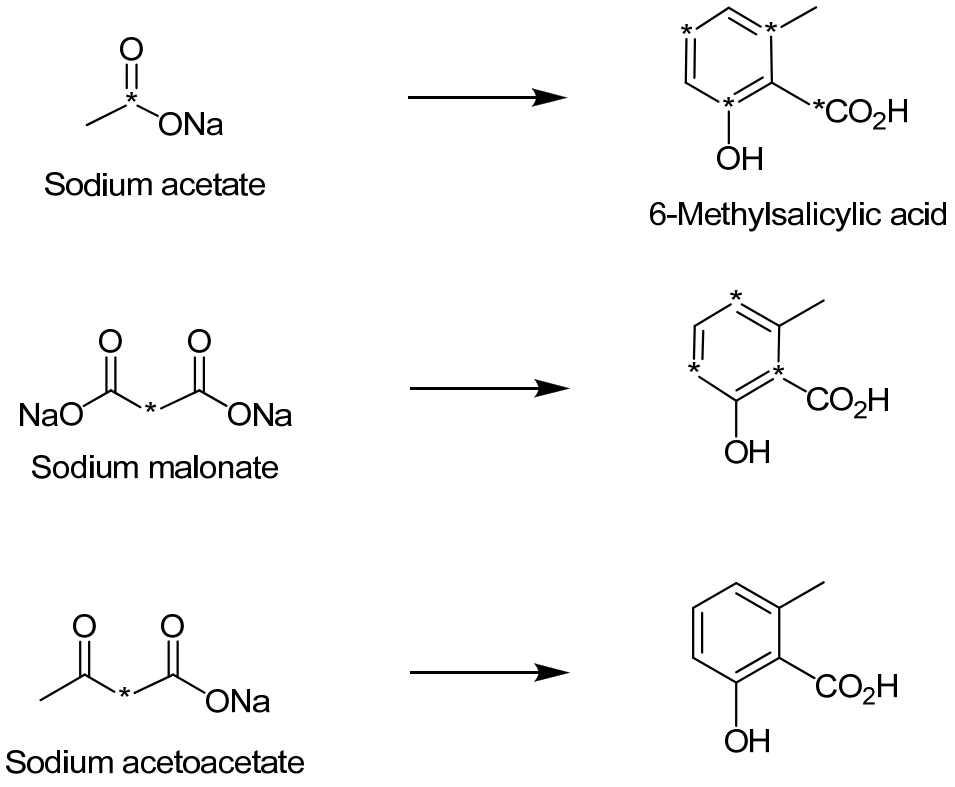
Figure III.4.1



Scheme III.4.1

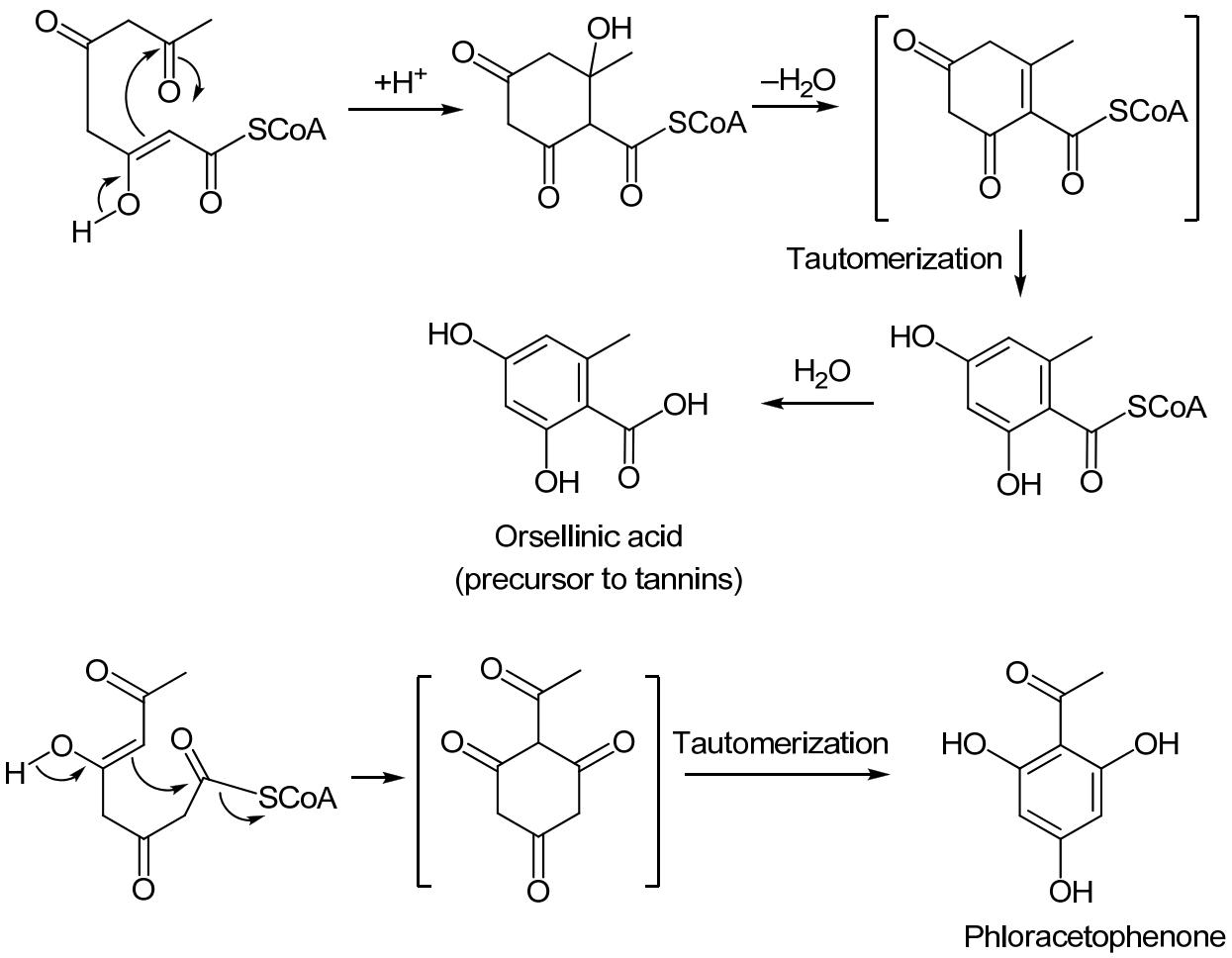
The joining of the acetate units is thought to occur in a multienzyme system much like the one involved in fatty acid synthesis. Again, acetyl CoA is the starter unit for most polyketides, although propionate and butyrate units are much more commonly encountered than in the fatty acid family. The unit is transfered to the carrier protein where it is acylated with malonyl CoA via the condensing enzyme (Scheme III.4.1). Transfer of the condensed unit to the carrier site

frees up the condensing site for incorporation of a second malonyl CoA unit and subsequent elongation of the chain. Labelling studies conducted with ^{13}C labelled sodium acetate concluded that the acetate was readily incorporated everywhere in the molecule (Scheme III.4.2). Labelled sodium malonate however was shown to be incorporated in the elongated chain but not in the starter units. Labelled sodium acetoacetate showed no incorporation at all in the natural product, therefore demonstrating that the polyketide chain is not released from the multienzyme unit until the required length and secondary transformations are completed (vide infra).

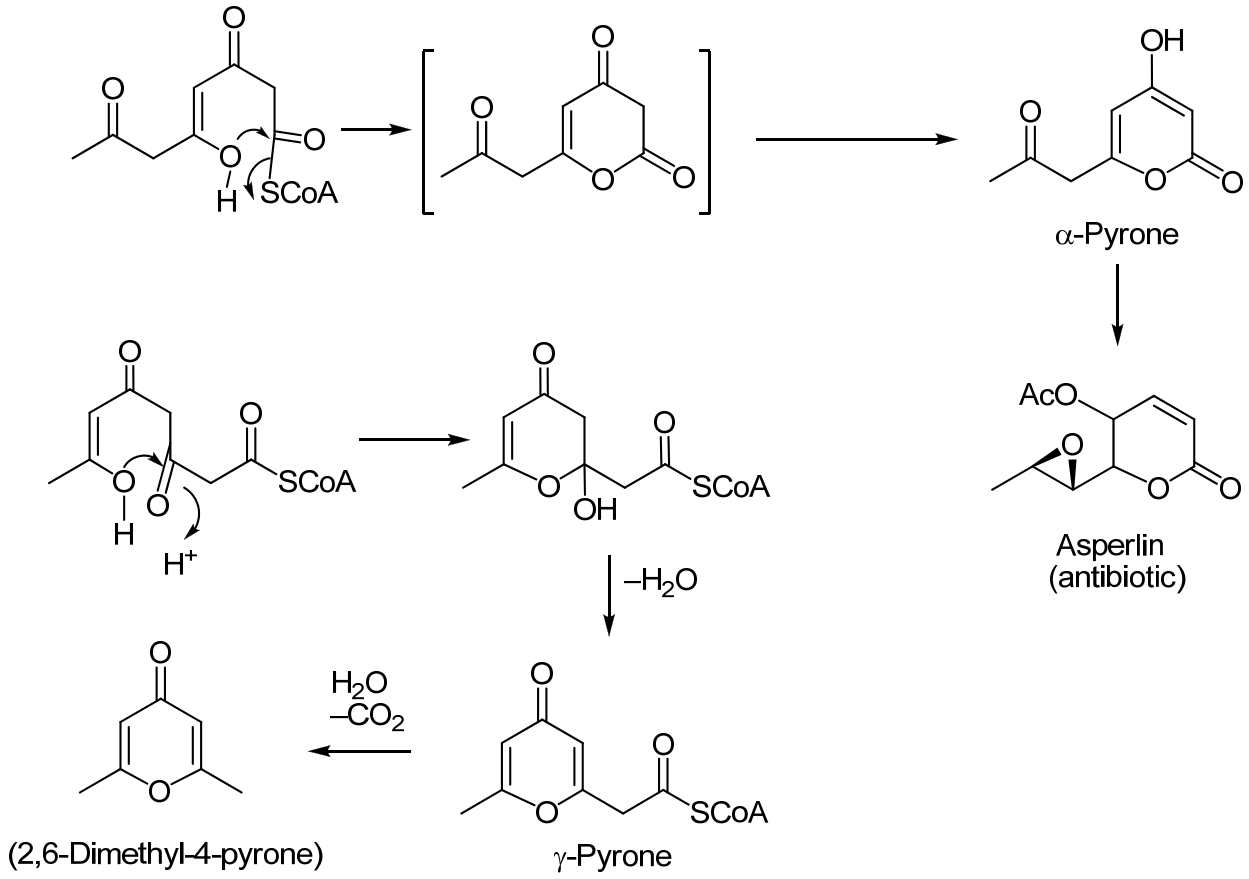


Scheme III.4.2

When the polyketide chain is complete a series of aldol and/or Claisen condensations and dehydrations occur, after proper folding of the chain, to give phenolic compounds. For example orsellinic acid is formed from an aldol condensation between C_2 and C_7 of the 8 carbon polyketide chain (Scheme III.4.3). Phloracetophenone may be produced by a Claisen-type condensation of C_6 onto C_1. Other types of cyclization are also possible as demonstrated in the biosynthesis of asperlin (Scheme III.4.4). The C_5 carbonyl condenses via its oxygen atom onto the C_1 carbonyl, resulting in the formation of a lactone. Furthermore, a similar condensation can lead to γ-pyrones when it occurs at a ketone carbonyl. The resulting alcohol dehydrates to give the pyrone ring as exemplified in the biosynthesis of 2,6-dimethyl-4-pyrone (Scheme III.4.4).

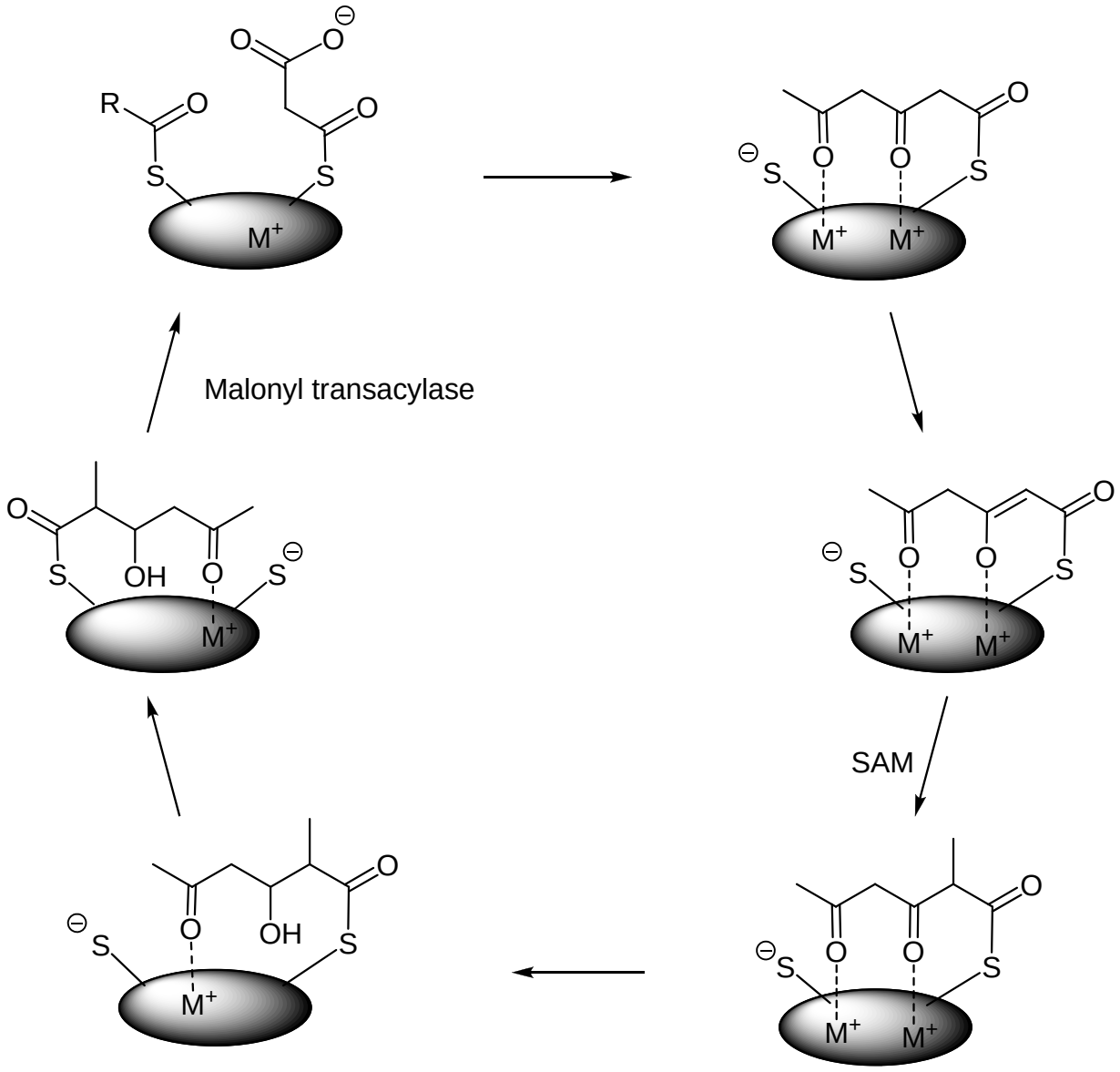


Scheme III.4.3

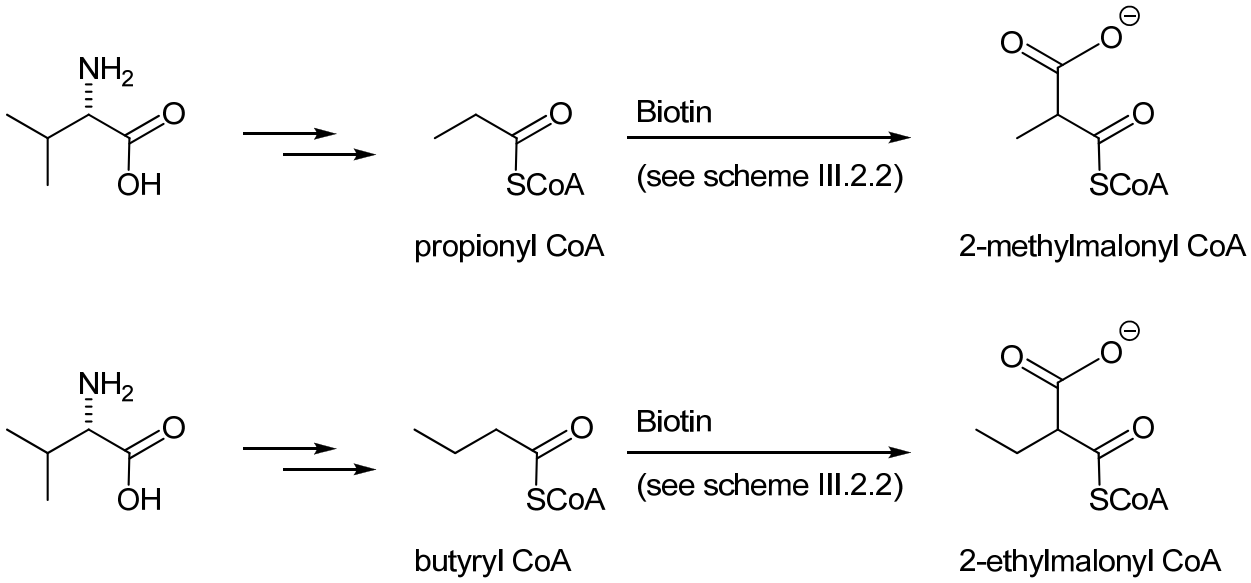


Scheme III.4.4

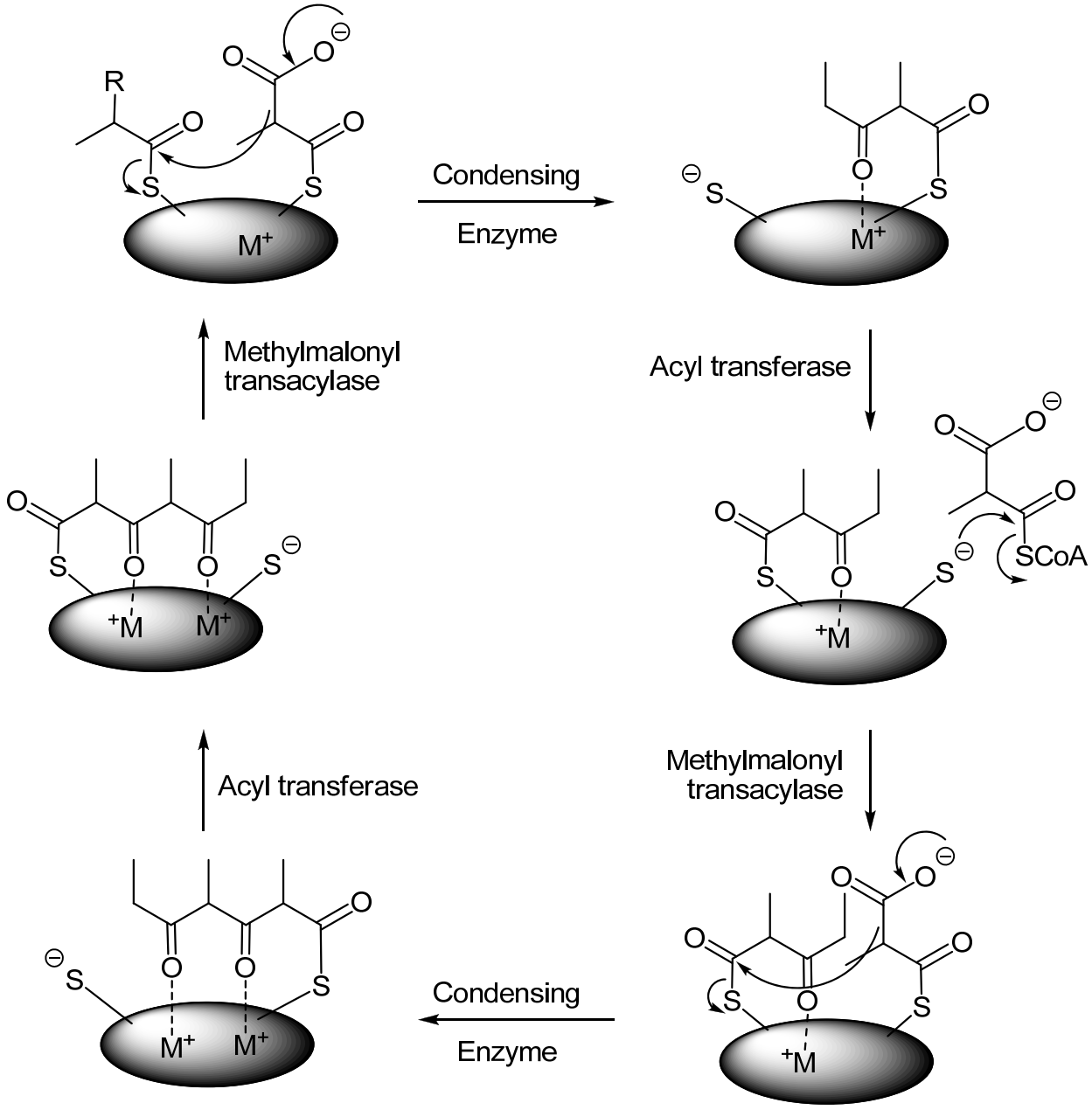
Some polyketides are subjected to a number of secondary modifications which include reduction of the carbonyl, oxidations, alkylation with S-adenosylmethionine (SAM), and introduction of halogens or other heteroatoms. In most cases, the reductions and methylations take place within the multienzyme complex before the polyketide chain is cyclized. Scheme III.4.5 shows how the methylation reaction and the reduction of a carbonyl, as examples, can arise during or after chain elongation. Other modifications, like oxidations, halogen incorporation, or amination, are more likely to occur after the cyclization stage outside the multienzyme system. Furthermore, groups other than acetate can serve as the starter or the elongation unit. This leads to products having propionate, butyrate, malonate, or aminomalonate and other units incorporated in them. Propionate and butyrate units are derived from the amino acid valine (Scheme III.4.6). This pathway must be distinguished from the propionate unit arising from methylation of acetate units after elongation as taken place. Scheme III.4.7 shows the biosynthesis of polyketides with propionate units. Let's see in more details the chemistry of specific families of polyketides in which some of these transformations will arise.



Scheme III.4.5



Scheme III.4.6

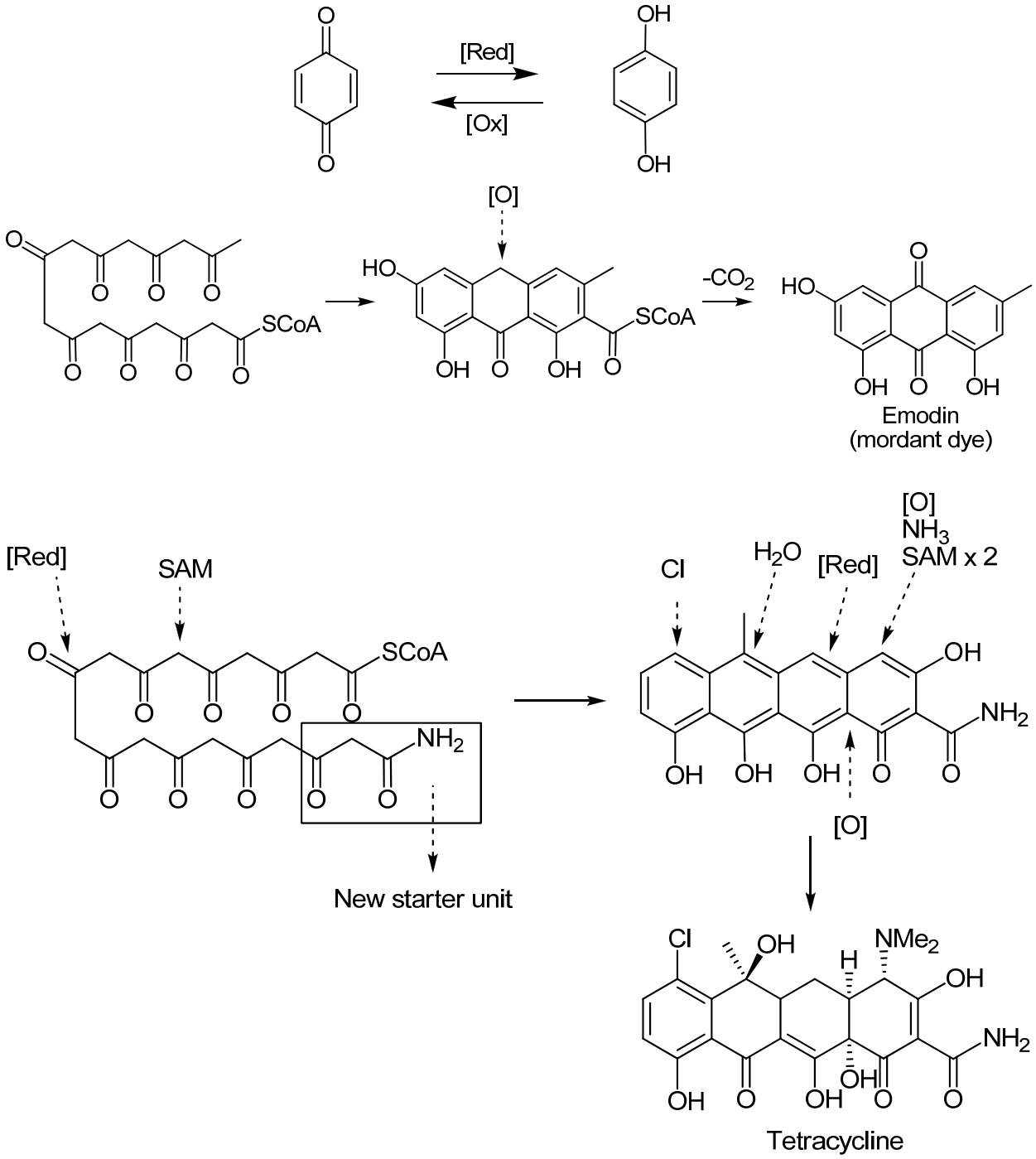


Scheme III.4.7

III.4.2. Quinones.

A quinone group is the oxidized form of hydroquinone (Scheme III.4.7, top). The quinones, as a family of natural products, are ubiquitous and comprise among others, the benzoquinones, the naphthoquinones, the anthraquinones, the anthracyclinones, and the tetracyclines. Anthraquinones are the largest group of quinones. They are widely spread among lower and

higher plants and have been used as dyes, e.g. alizarin and emodin. They lost their importance, like so many other natural dyes, with the advent of synthetic dyes.



Scheme III.4.7

The biosynthesis of quinones is difficult to systematize because of their diversity in carbon skeletons. Some, like many benzo- and naphthoquinones, are actually metabolites of shikimic acid (section IV) and mevalonic acid (section II) or of mixed origins (one ring may be derived from polyketide and the other from mevalonate). However, the polyketide pathway leading to anthraquinones and tetracyclines is well documented. Scheme III.4.7 (center) depicts the biosynthetic origin of emodin, an anthraquinone that has served as a mordant dye for many years. Above, in Scheme III.4.7, is the biosynthesis of tetracycline, a strong antibiotic which has served in human therapy. Note the advent of a malonamide as a starter unit different from acetate, propionate, or butyrate. A reduction and methylation with SAM takes place on the nonaketide before cyclization. Then a series of oxidations, amination, and chlorination leads to tetracycline.

III.4.3. Flavonoids.

The flavonoids are the major players in the coloration and taste of flowers and fruits. They are in large part responsible for the flavors from food of plant origins (fruit or others). The flavones give yellow or orange colors, the anthocyanins give red, violet or blue colors. The occurrence of this family of compounds is restricted to higher plants although some mosses contain a few flavonoid compounds. They comprise the chalcones, flavanones, flavones, flavanonols, flavonols, isoflavones, aurones, anthocyanidins and others (Figure III.4.2). The flavonoids play a prominent role in insect pollination, attracting them with bright colors and nice taste. However, some have very bitter taste and serve the opposite purpose of repelling predators such as caterpillars that feed on leaves, etc.

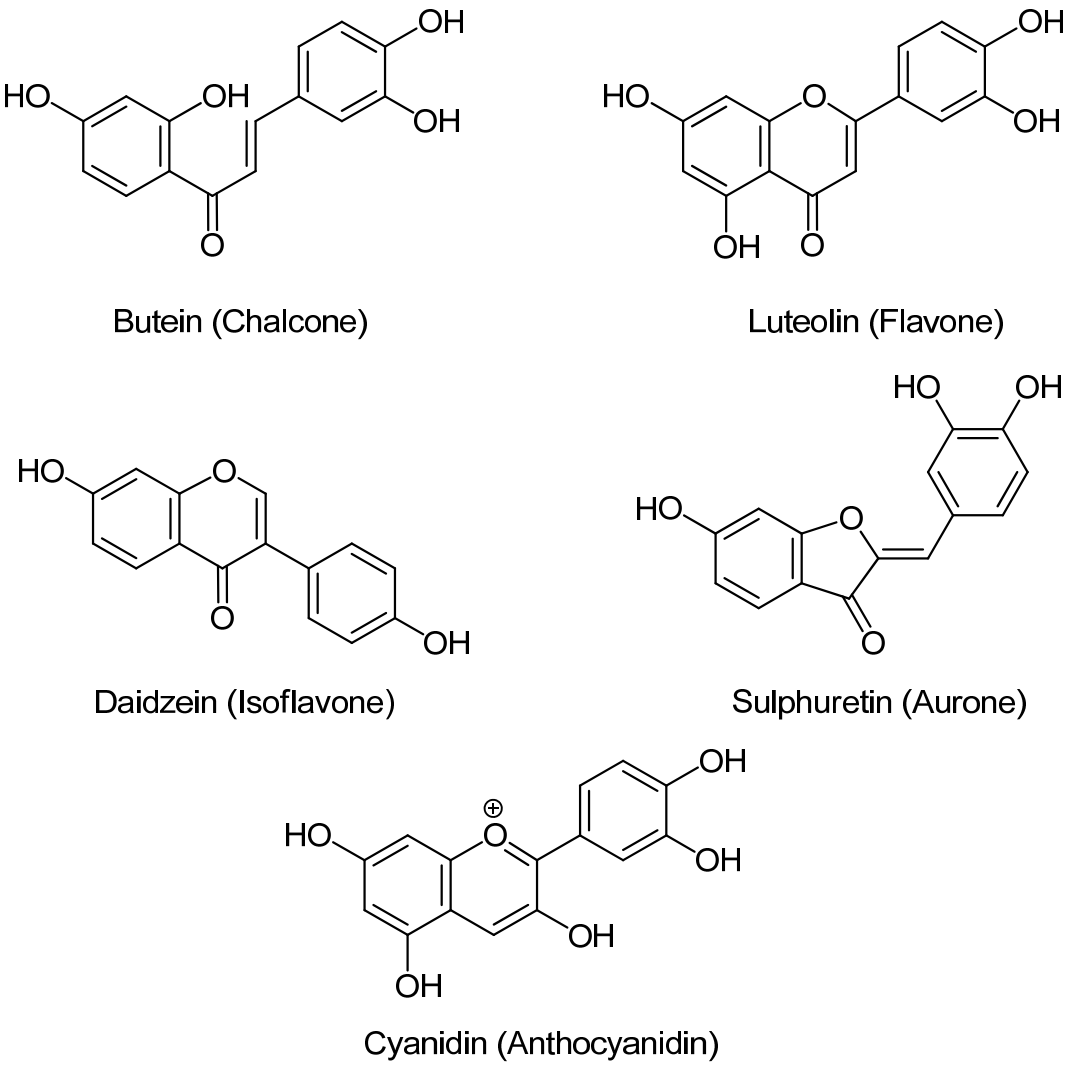
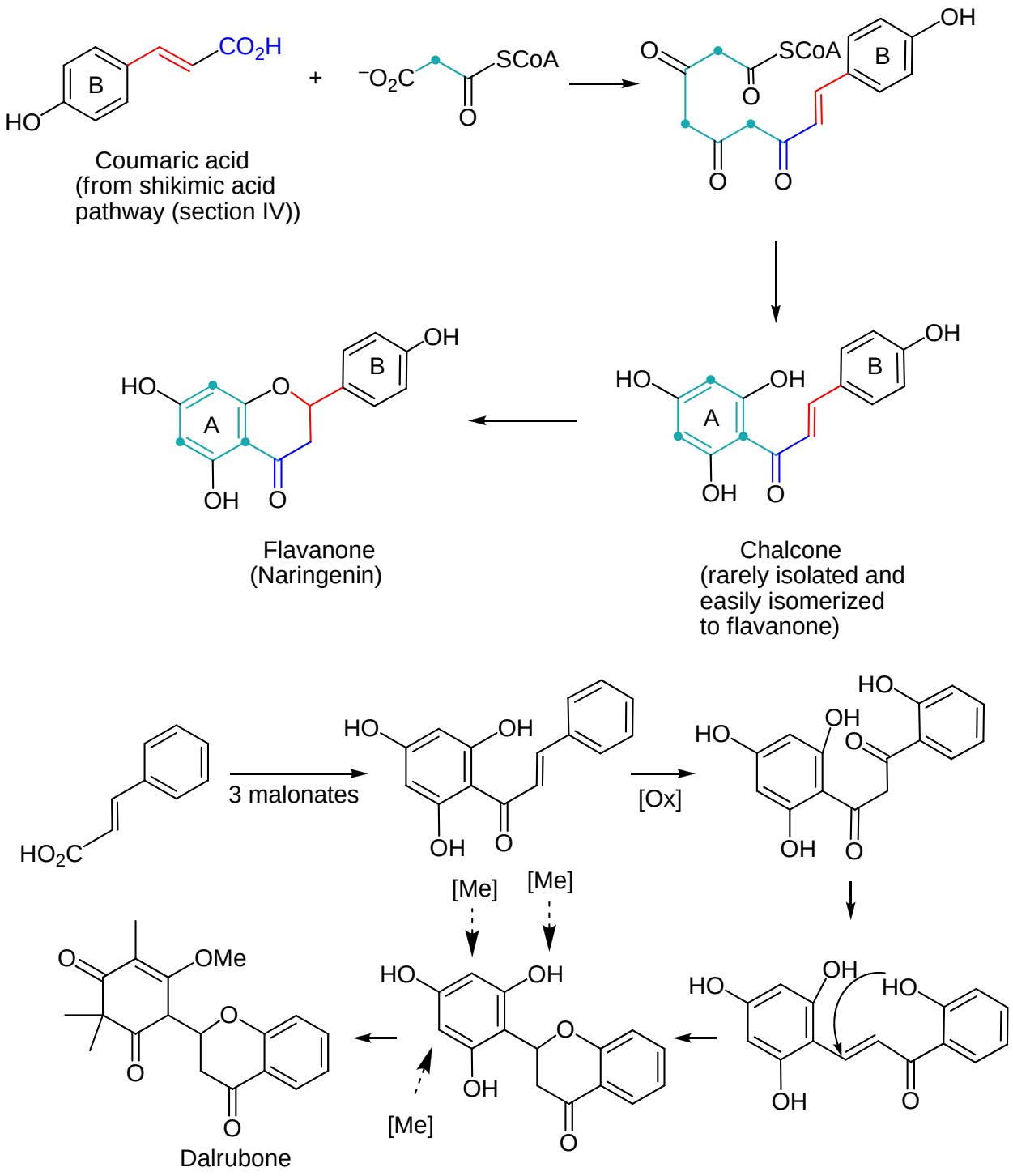
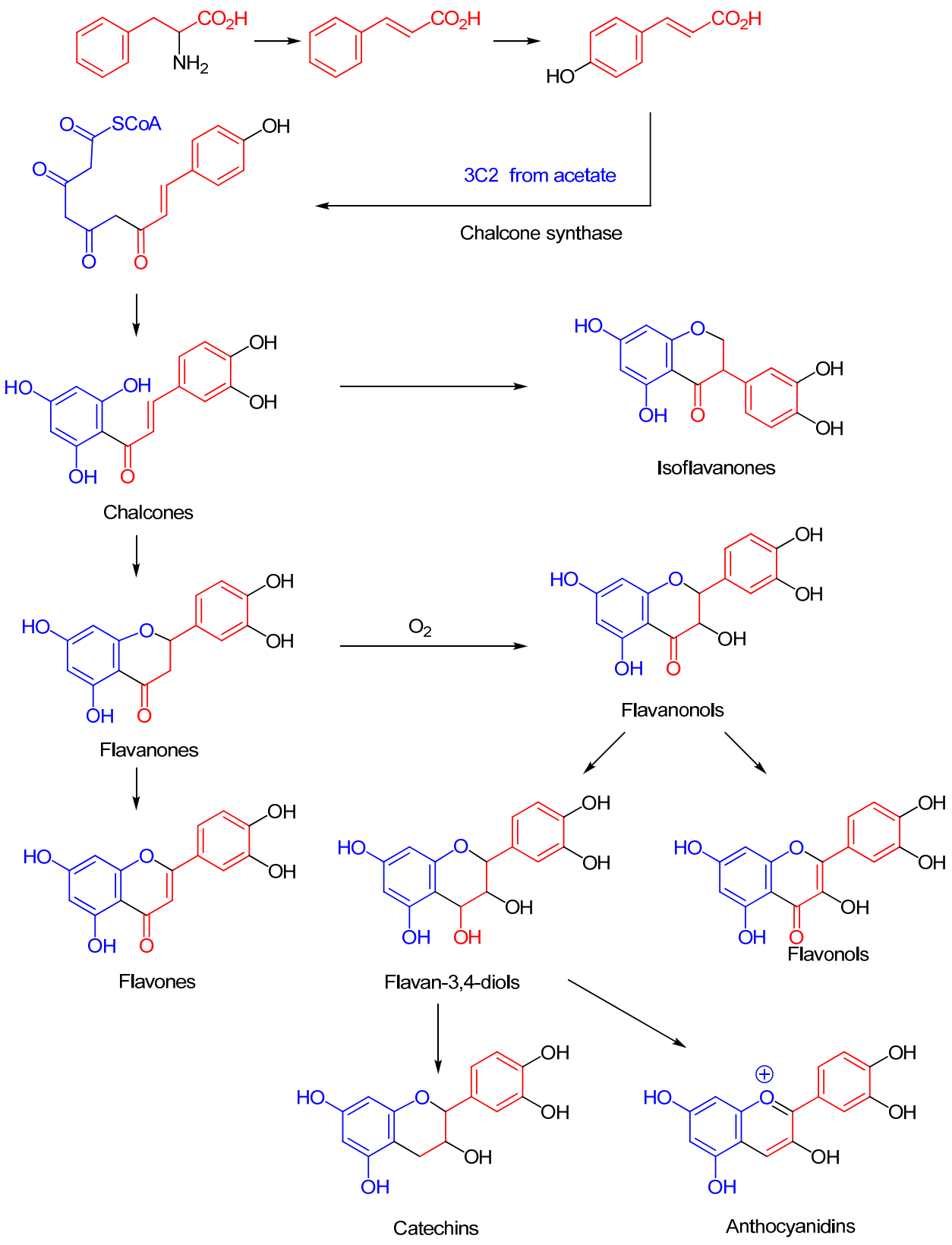


Figure III.4.2

This family of natural products is characterized by having two hydroxylated aromatic rings, A and B, joined by a three carbon fragment. One hydroxyl group is often linked to a sugar molecule. Hydroxyl groups on the chain can link the two rings together. Scheme III.4.8 shows the possible biosynthesis of some flavonoids. Nearly all flavonoids are of mixed biosynthetic origin, ring B emanating from the shikimic pathway (section IV) and the ring A coming from the polyketide pathway. The shikimate always serves as a starter unit different from acetate, propionate, or butyrate in this case. They are more or less all related in that isoflavones are usually derived from flavones which are in turn derived from chalcones, etc. (Scheme III.4.9). Narengenin is a strong antioxidant and radical scavenger found in grapefruits and other citrus.



Scheme III.4.8


CO2H
NH2
CO2H
CO2H
HO
SCoA
O
OH
3C2 from acetate
Chalcone synthase
Chalcones
Isoflavanones
Flavanones
O2
Flavanonols
Flavones
Flavan-3,4-diols
Flavonols
Catechins
Anthocyanidins

Scheme III.4.9

Chalcones are quite rare, due in part to their ready isomerization to flavanones (Scheme III.4.8) but one that is of some interest is a yellow pigment in the petals of safflower (Figure III.4.3). As the flower ages, a red pigment carthamone, is produced. The flavanones are also relatively uncommon, but occasionally accumulate in fruits, flowers, leaves, and wood. Two typical examples are the soluble, bitter-tasting compounds from citrus fruits, naringenin (grapefruit peel) and hesperitin (orange peel) (Figure III.4.3). Both flavones and flavonols are very widely distributed, and luteolin (flavone) and quercetin (flavonol) are particularly common in leaves, but also found as constituents of rinds, barks, clover blossom, and ragweed pollen. The isoflavones, occur most commonly in plants of the Leguminosae (peas) family. They are believed to be formed from flavones via an unusual rearrangement involving a diradical species (Scheme III.4.10). Finally, the anthocyanins are primarily responsible for the red, blue, and violet colors of flower and fruits. They are believed to derive from dihydroflavanols, but the final stages of the biosynthetic pathway remain to be elucidated. The colors of petals in flowers almost certainly act as attractive stimuli for pollinating insects.

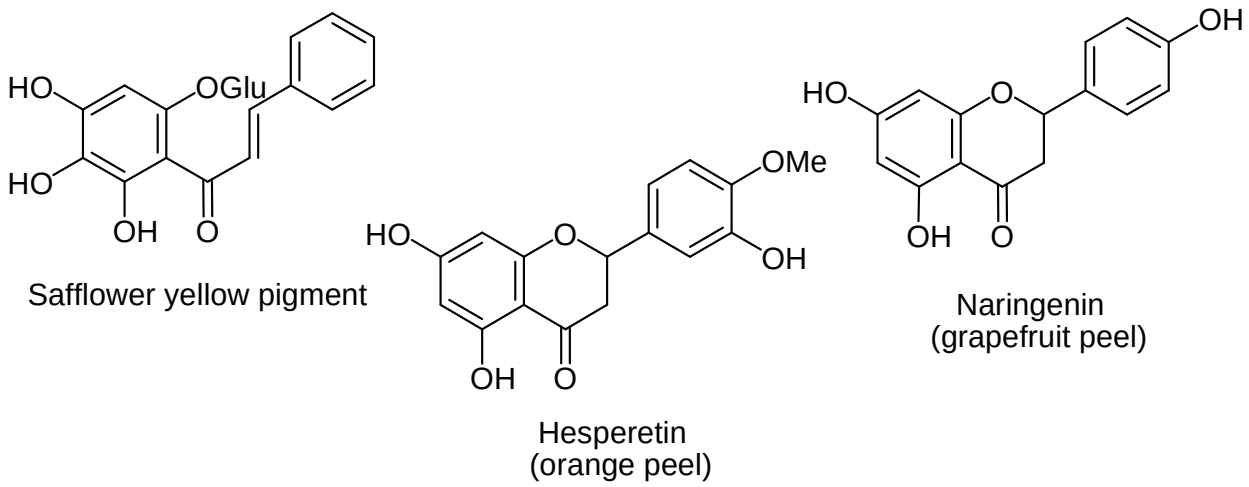
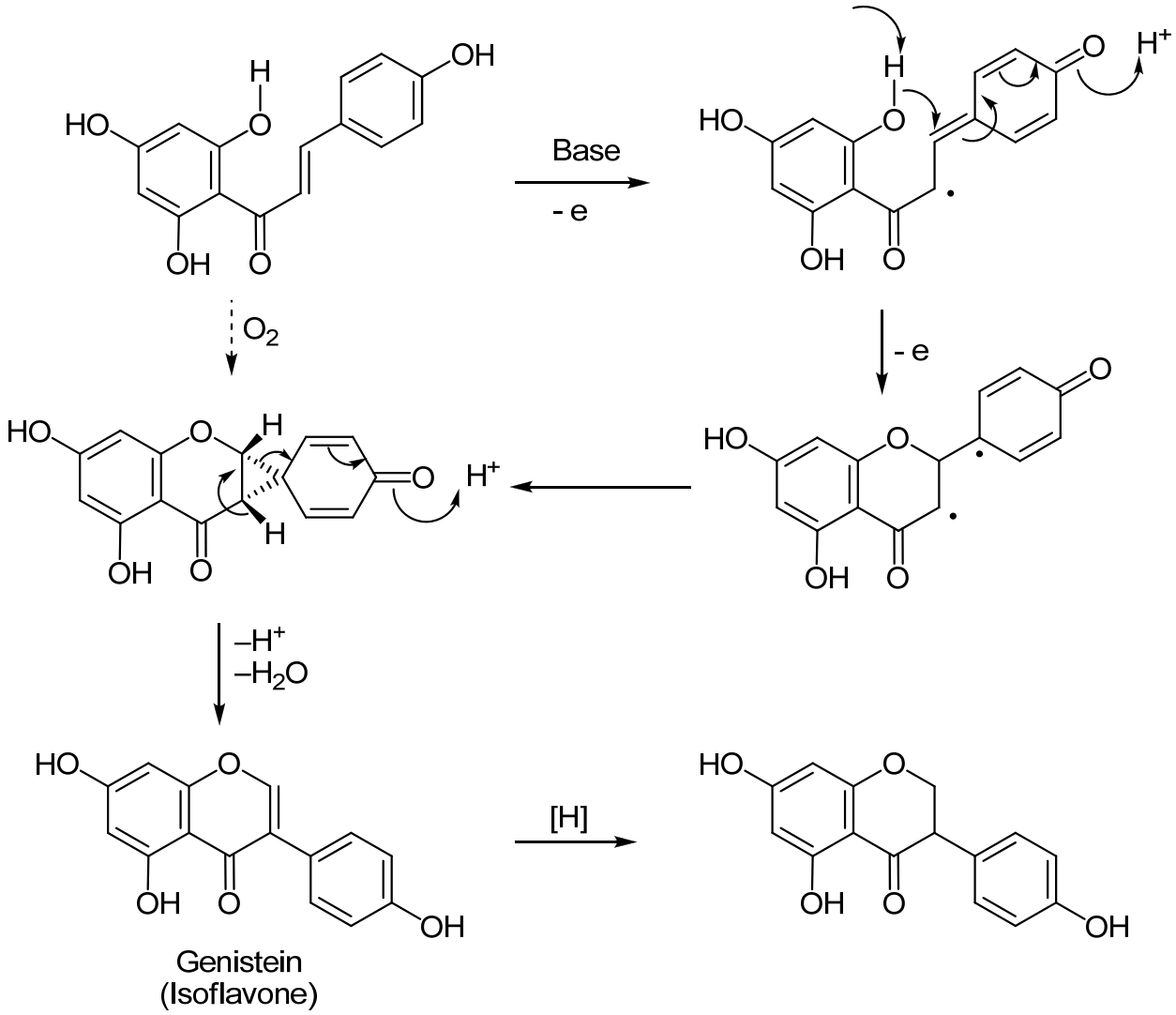


Figure III.4.3

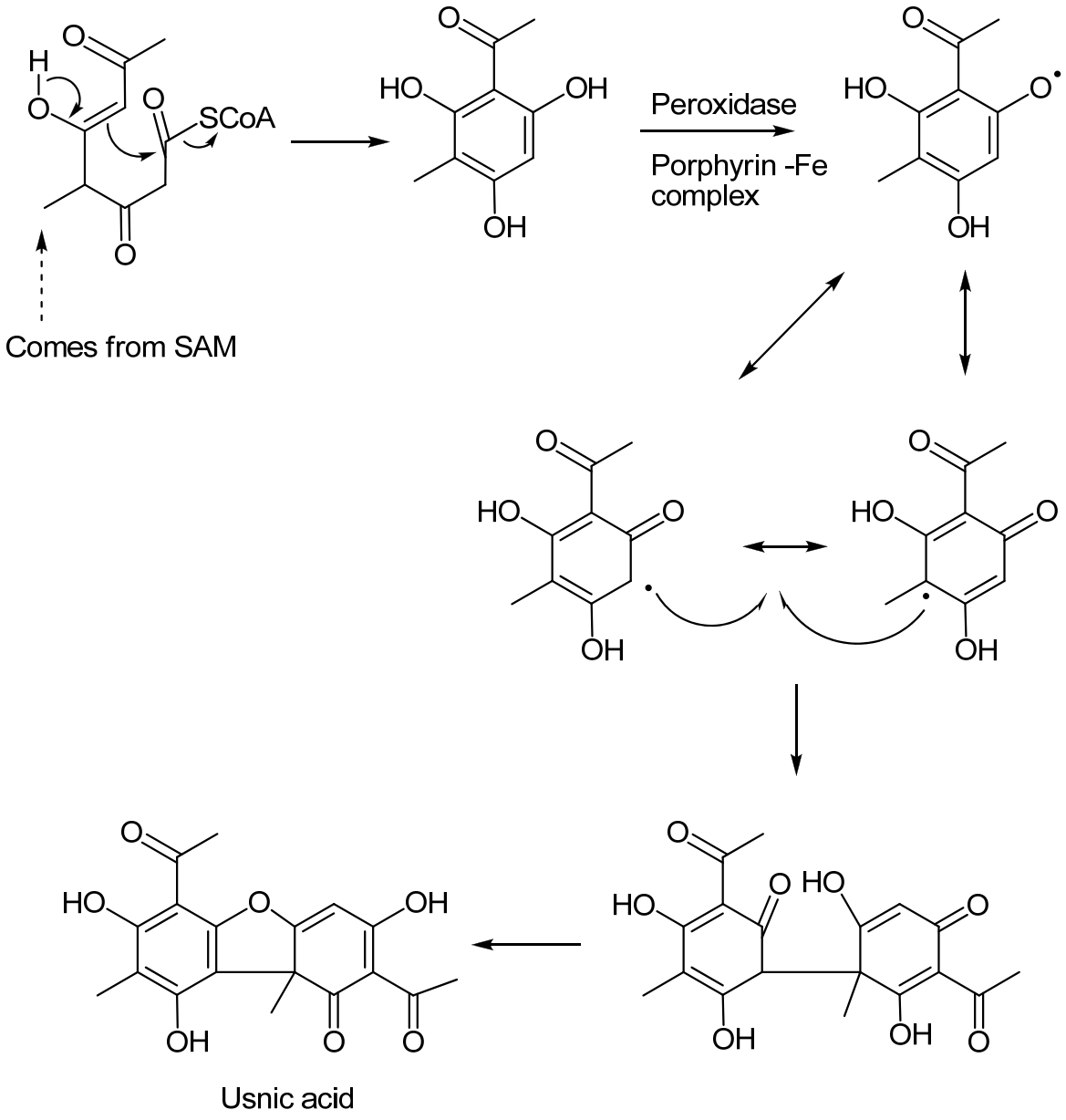


Scheme III.4.10

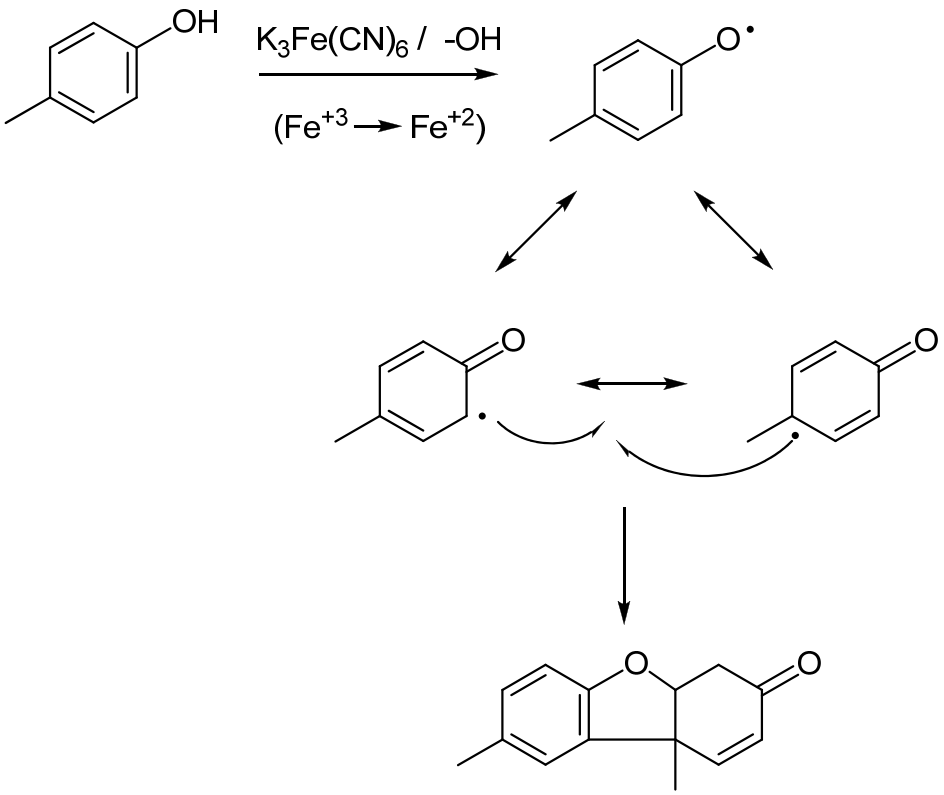
III.4.4. Oxidative Phenolic Coupling.

Several groups of enzymes are capable of catalyzing oxidative phenolic coupling. They have iron or copper as a prosthetic group and are all able to effect one-electron transfers. Hydrogen peroxide and molecular oxygen, used as oxidants, are ultimately reduced to water, while the transition metal is reoxidized to its higher state. A particularly rich source of such polyphenolic compounds are the lichens, an alga and fungus growing together in symbiosis. Usnic acid is a rather common lichen metabolite which inhibits tumor growth, and its probable mode of biosynthesis is shown in Scheme III.4.11. The same oxidative phenolic coupling reaction can be performed in the laboratory using iron (potassium ferricyanide) and it is thought that the

mechanism of operation is likely to be very similar to the in vivo peroxidaze phenolic coupling reaction (Scheme III.4.12).



Scheme III.4.11



Scheme III.4.12

III.5. Macrolides and Polyether Antibiotics.

III.5.1. Biosynthesis of Macrolides and Polyethers.

The macrocyclic lactones such as erythromycin and the polyether ionophores typified by monensin and lasalocid are two important classes of antibiotics produce by actinomycete and streptomyces bacteria and fungi (Figure III.5.1). Erythromycin is one of the most common antibiotics used in human medicine. It replaces the penicillins where the latter are ineffective or produces allergic response. Many polyether ionophores are used in the control of bacterial-induced diseases in livestock. The two classes have been compared because of the similarity in structural features and biogenetic origin. However their biochemical mode of actions and their antibacterial spectrum of activity are very different. Whereas the macrolides are known to inhibit protein synthesis in bacterial cells, polyethers disrupt the permeability and ion transport ability of the cell membrane. In the latter case, the polyether usually chelates the metal cation (Ca^{++}, K^{+}, Na^{+} and others) with its oxygens in such a way as to expose its outer hydrophobic (lipophilic) shell of alkyl groups to the membrane and thus it is able to transport the metal ion across

biological membranes. Some 70 polyether antibiotics are known today whereas over 150 macrolides have been characterized since the discovery of pikromycin in 1951.

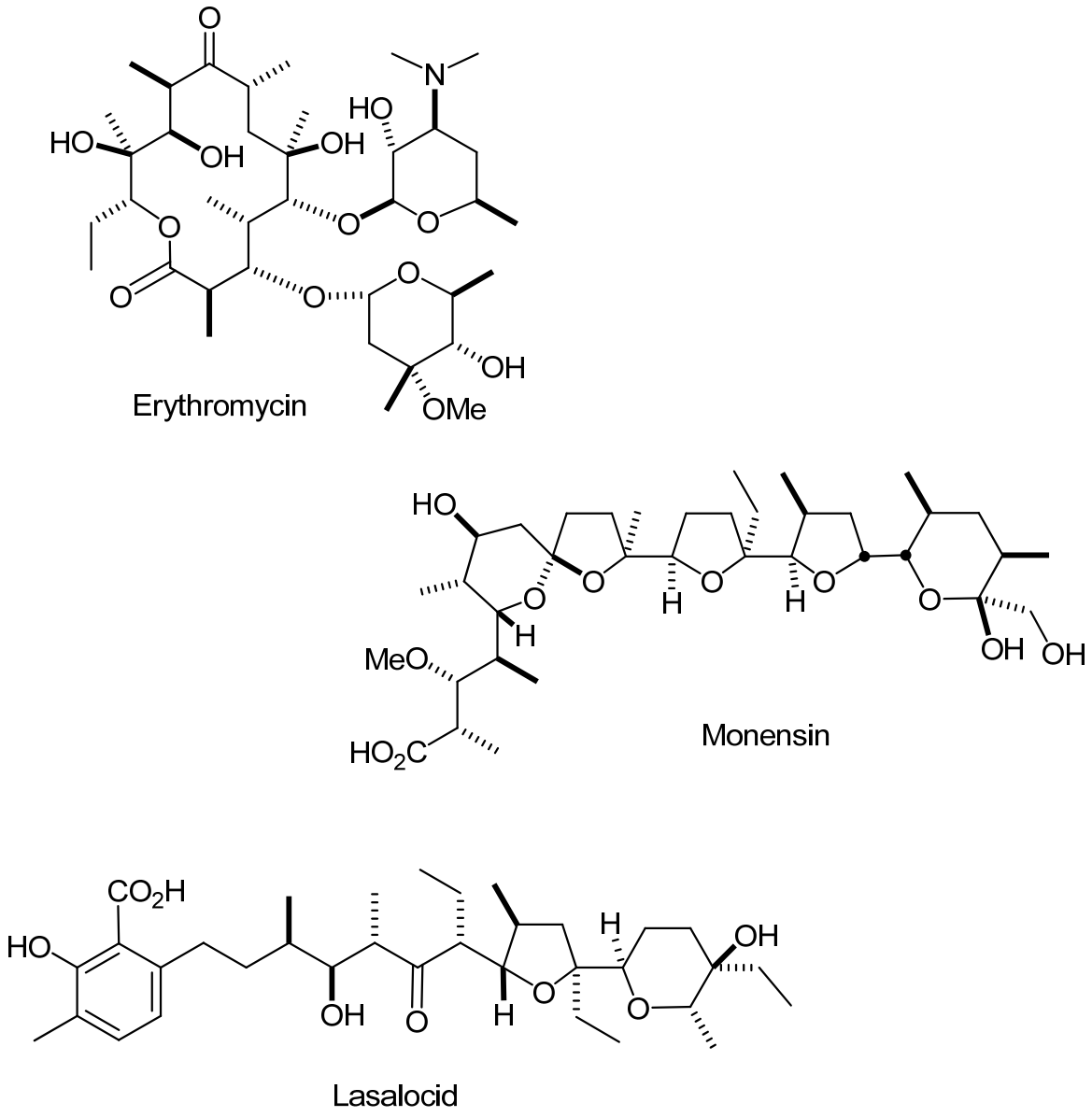
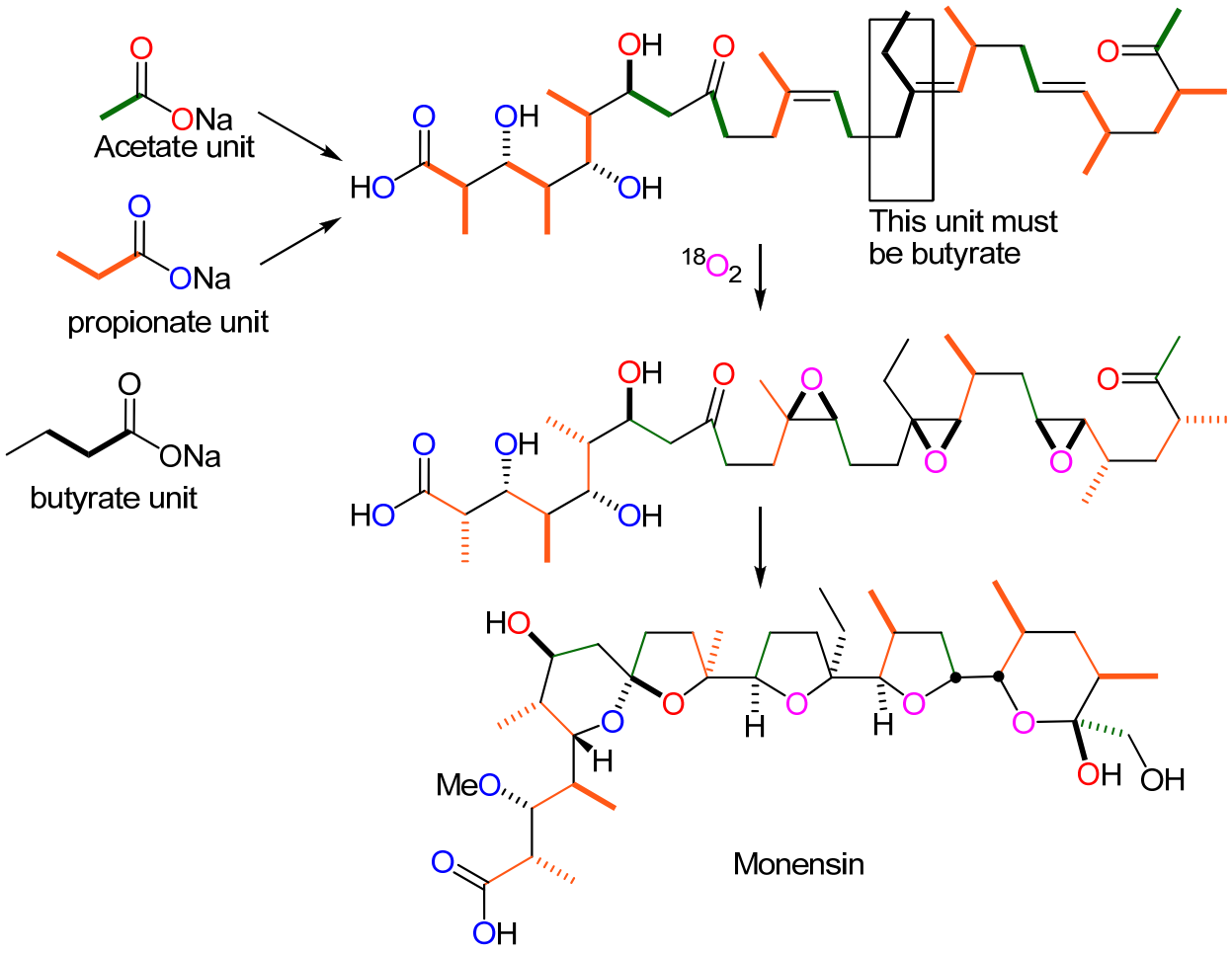


Figure III.5.1

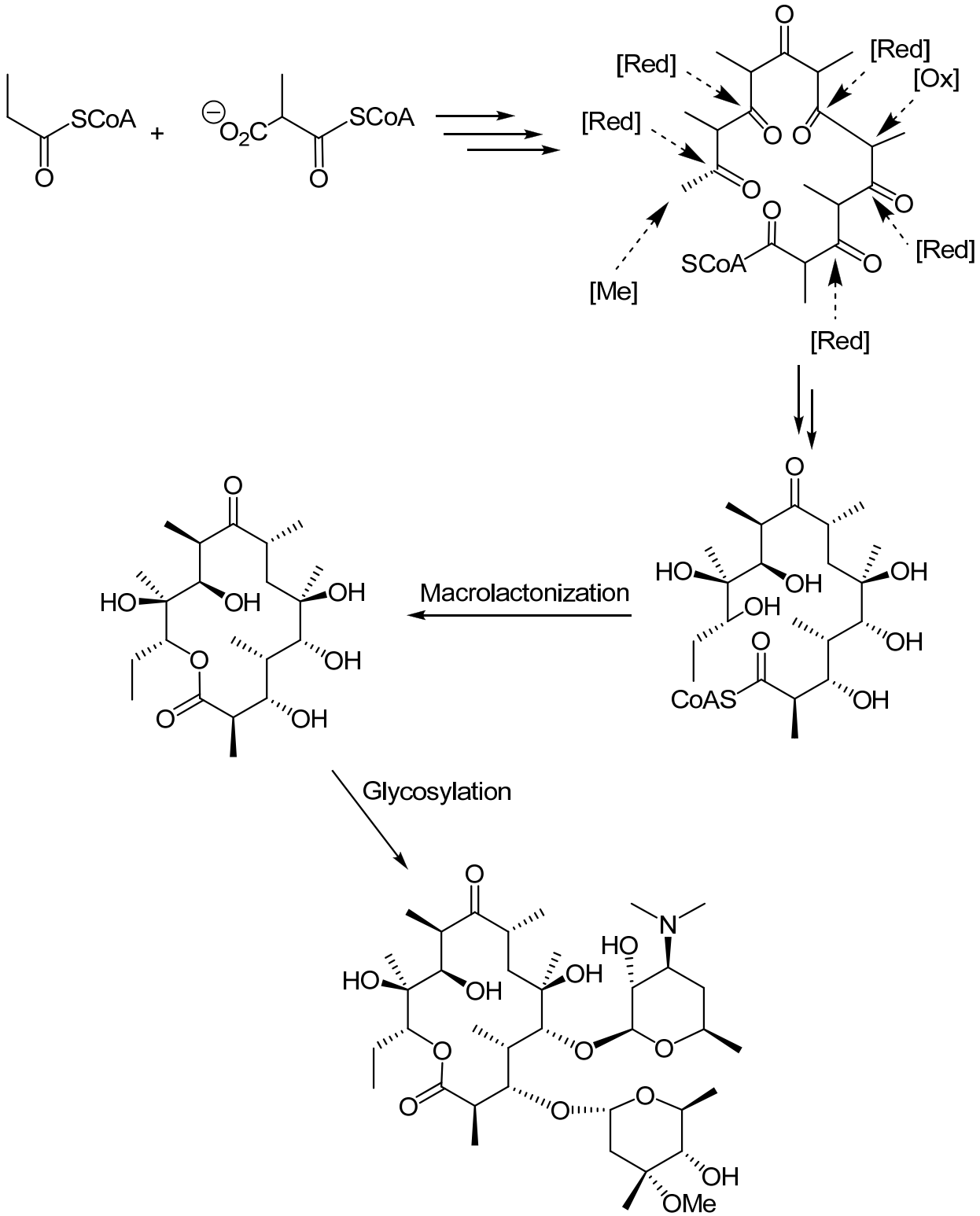
Their biogenetic origin is very much similar to the fatty acid biosynthesis with, however, some important differences. Starter and adding units can be acetate, propionate, or butyrate and most molecules contain a mix of these three units. In addition, oxidation by peroxidase is common and whereas some units suffer reduction of the olefin as in the normal fatty acids, other units are intact or retain the alcohol. The polyethers suffer a series of epoxidations of their

alkenes followed by internal cyclization or "polyepoxide cyclizations". The macrolides, on the other hand, have the acid moiety internally closed onto a terminal alcohol to form their unique 12, 14, or 16 membered rings. Scheme III.5.1 depicts the probable biological pathway to monensin. Note that five (5) units come from acetates whereas seven units come from propionates. The unmarked unit probably comes from a butyrate. Not all the oxygens of the acyl units were retained as some came from the oxidation of olefins by O_2 oxygen. A polyepoxide cyclization, akin to the polyolefin cyclization that we discussed in the terpenoid sections, is responsible for the formation of the tetrahydrofurans and -pyrans. Scheme III.5.2 shows a possible biogenesis of erythromycin having all propionate units, except the last unit which was acetate derived and has been extended with SAM to a propionate unit. Labelling experiments demonstrated that this unit did not come directly from the incorporation of a butyrate unit in the elongation of the polyketide chain. It is possible that the oxidation sequence takes place after or before the macrolactonization.



Scheme III.5.1

Biosynthetic pathway to erythromycin A



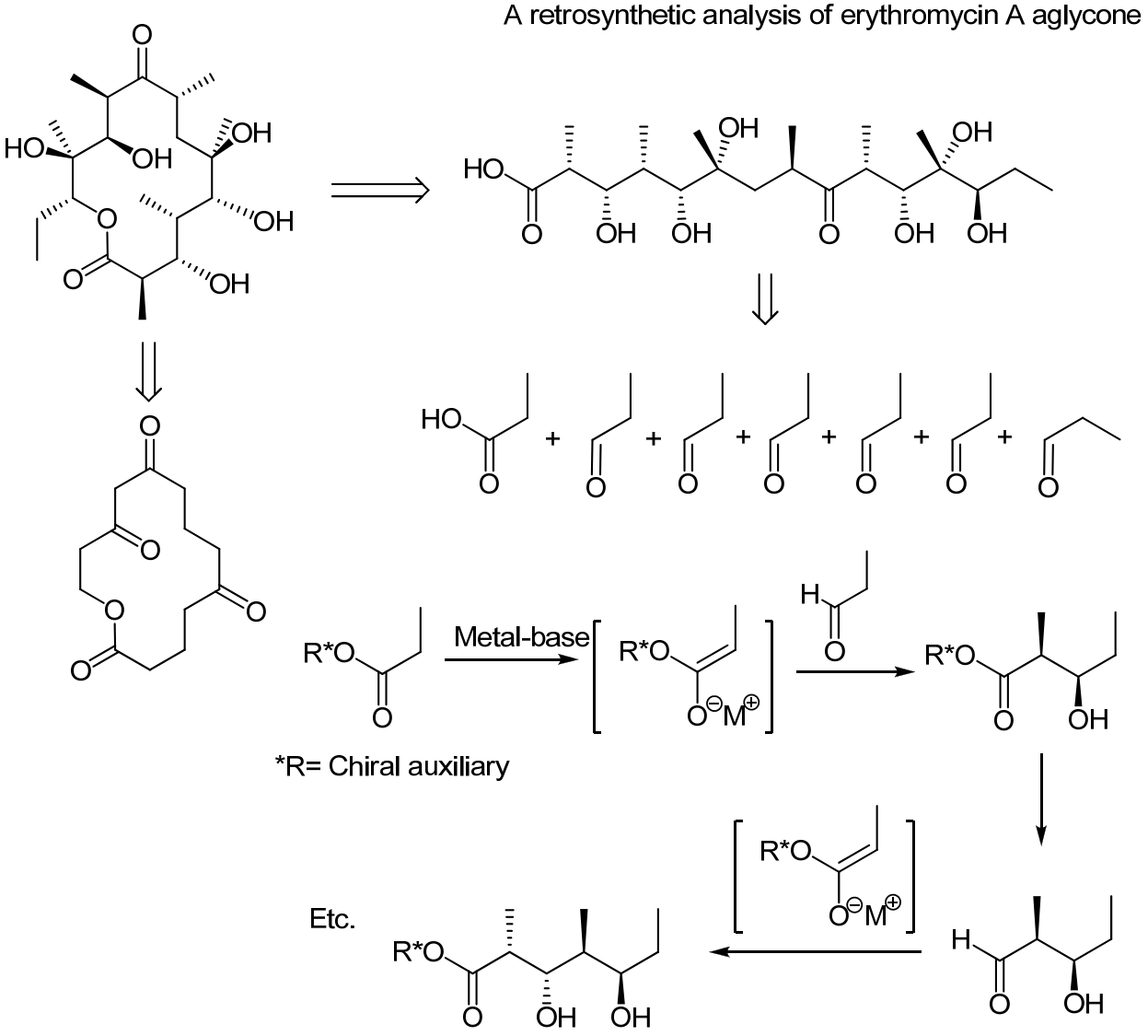
Erythromycin A

Scheme III.5.2

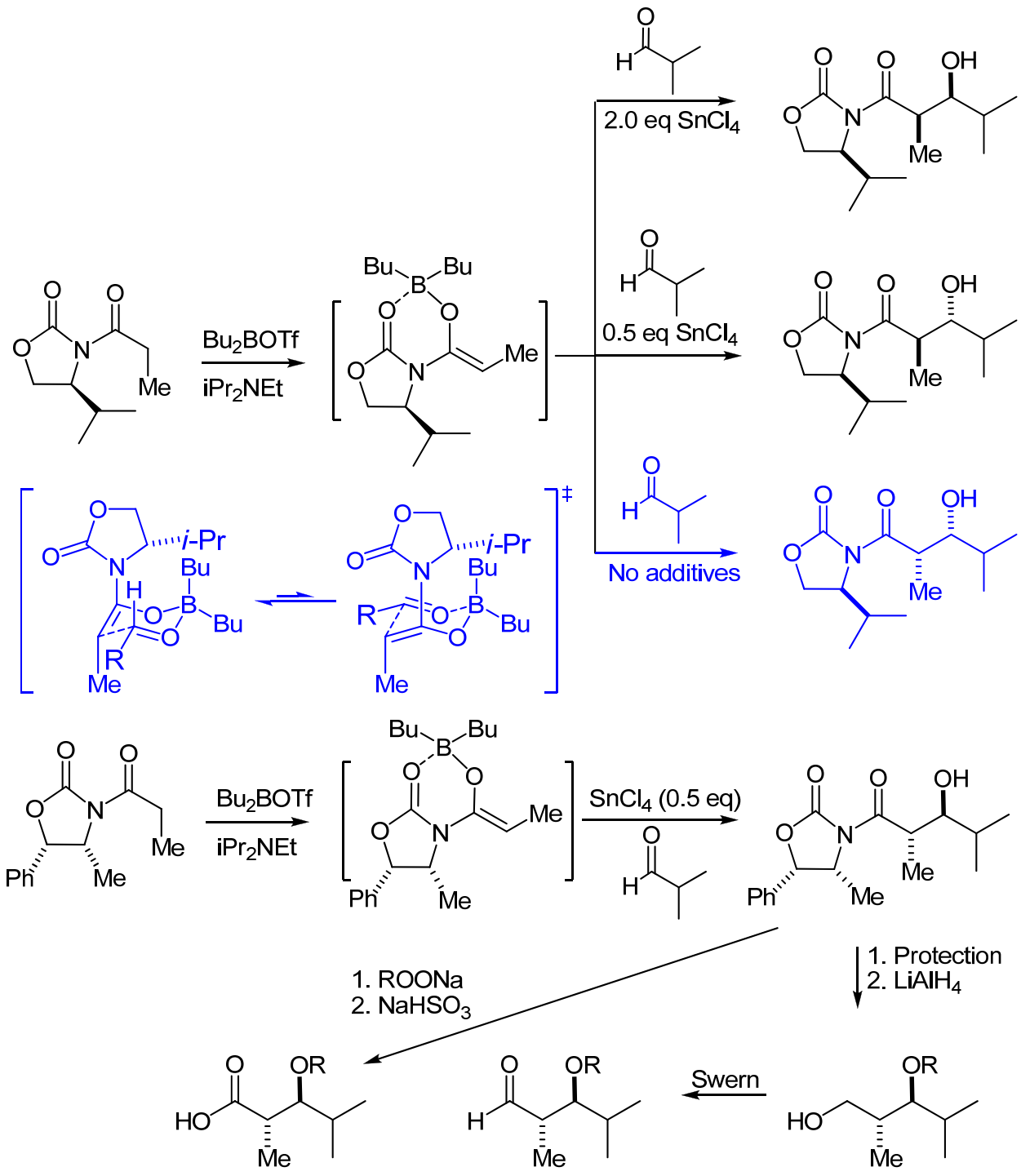
III.5.2. Chemistry of Macrolides and Polyether Ionophores.

Due to their importance in medicine and their challenging structures as synthetic targets, much effort has been devoted to the total synthesis of this family of compounds. In the macrolides, two main approaches were designed (Scheme III.5.3). In the first and most common approach the molecule is assembled as an acyclic compound and macrocyclized in the last steps. A second approach consists of cyclizing to the macrolactone early in the synthesis and using the conformational bias of the macrocycle to stereoselectively introduce all the required functionality. This latter approach is not without its caveat and difficulties because of the floppiness of the macrocycle. For that reason it has been pursued with limited success. It was recognized very early that the acyclic propionate units though, could be assembled most successfully via the aldol reaction (or Claisen condensation), a sort of mimic of their biosynthesis. In fact, the construction of erythromycin could be envisaged as a series of six aldol reactions (Scheme III.5.3). While the aldol condensation is one of the oldest chemical reaction known to chemists, the asymmetric aldol reaction was not developed before the mid 1970's early 1980's. Research in the asymmetric aldol reaction was spurred from the desire to synthesized natural products containing chiral propionate units like the macrolides and the polyether antibiotics. There are several possible approaches to induce chirality in the aldol reaction. The most successful and most utilized to date consists in fixing a chiral auxiliary molecule on the ester and condensing its metal enolate with an appropriate aldehyde under carefully defined conditions. Scheme III.5.4 shows one example of such a chiral auxiliary. Today, all four possible diastereomers of the chiral 1-hydroxy-2-methyl propionate unit can be stereospecifically prepared by careful adjustment of the reaction conditions. The chiral auxiliary is cleaved by reduction/oxidation or by hydrolysis by the nucleophilic peroxides and reduction. Aldehydes can be obtained and used iteratively in a sequence of aldol reactions. After completion of the acyclic chain, the free hydroxy acid is internally cyclized to afford the macrolactone. This step is particularly difficult in view of competing side reactions such as dimerization. Today, typical yields for such a step are still low but some success has been achieved by activating the acid group as amide or thioester groups (much like SCoA activates a carbonyl toward nucleophilic attack).

A retrosynthetic analysis of erythromycin A aglycone



Scheme III.5.3

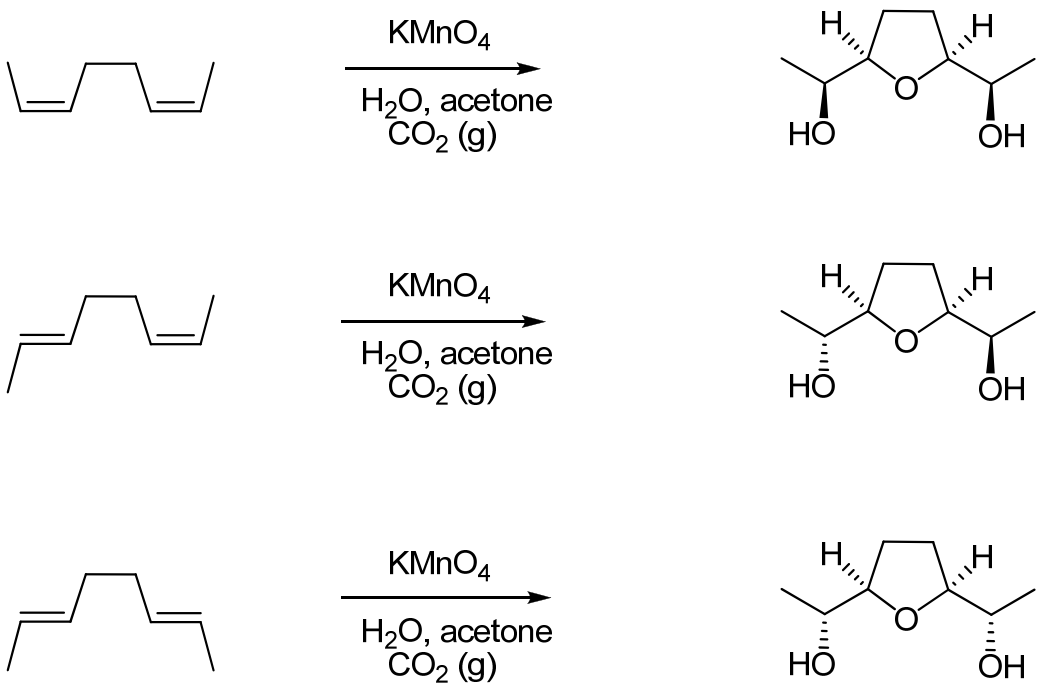


Scheme III.5.4

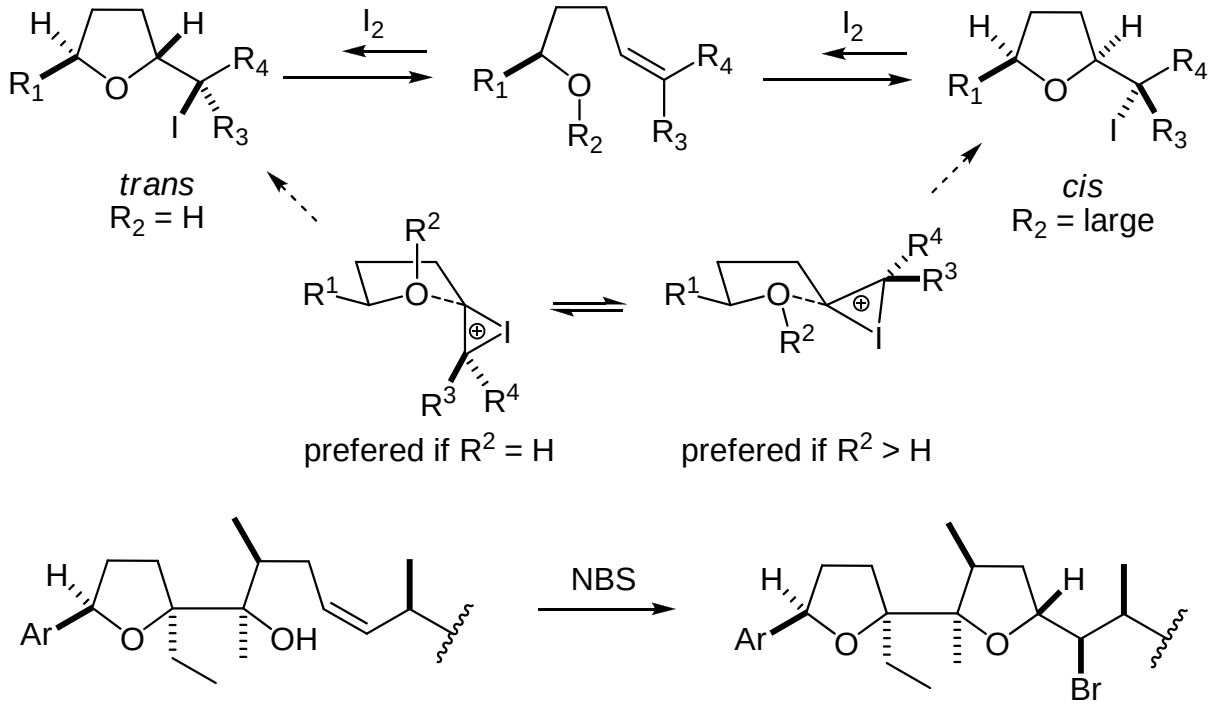
Polyethers have two main structural features: long acyclic acetate, propionate, or butyrate chains and tetrahydrofuran or -pyran rings. The asymmetric aldol condensation has been used

extensively to build the acyclic chains in the same way as for the macrolides. Much of the challenge in the synthesis of the polyethers came from the stereoselective construction of the heterocycles, and in particular the more common tetrahydrofuran. Schemes III.5.5 to III.5.7 list a few of the most common approaches. Still today, this is a challenging aspect of organic synthesis and other more stereoselective approaches with wider scope are needed. The permanganate oxidation of 1,5-dienes is a powerful one because four chiral centers are created in one step (Scheme III.5.5). Moreover the hydroxy group is often found in polyether ionophores. The mechanism is complex and not well understood and leads stereoselectively to *cis*-tetrahydrofurans only. In addition the yields are often modest because permanganate is a strong oxidant and side reactions such as overoxidation often complicate the reaction. The electrophile-promoted cyclization of alkene-alcohol is used for the construction of both tetrahydrofurans and -pyrans. In the case of tetrahydrofurans, *cis* compounds are formed preferentially when a bulky R_2 group, such as benzyl, is present as shown (Scheme III.5.6). When a simple alcohol is cyclized *trans* tetrahydrofuran rings are preferentially formed. This selectivity is rationalized in terms of steric effects in the transition state leading to cyclization: R_1 and R_2 will be in an anti-periplanar conformation when R_2 is large forcing the double bond to be also trans with R_2 and hence cis with R_1 (1,2-interactions are worse then 1,3-interactions in five-membered rings). However, when R_2 is small (e.g. hydrogen), it is the steric repulsion of the largest groups, R_1 and the double bond substituents, that controls the selectivity and, of course, they will adopt a *trans* orientation. Iodine, bromine, or other forms of electrophilic halogens such as *N*-bromosuccinimide (NBS) are possible electrophiles.

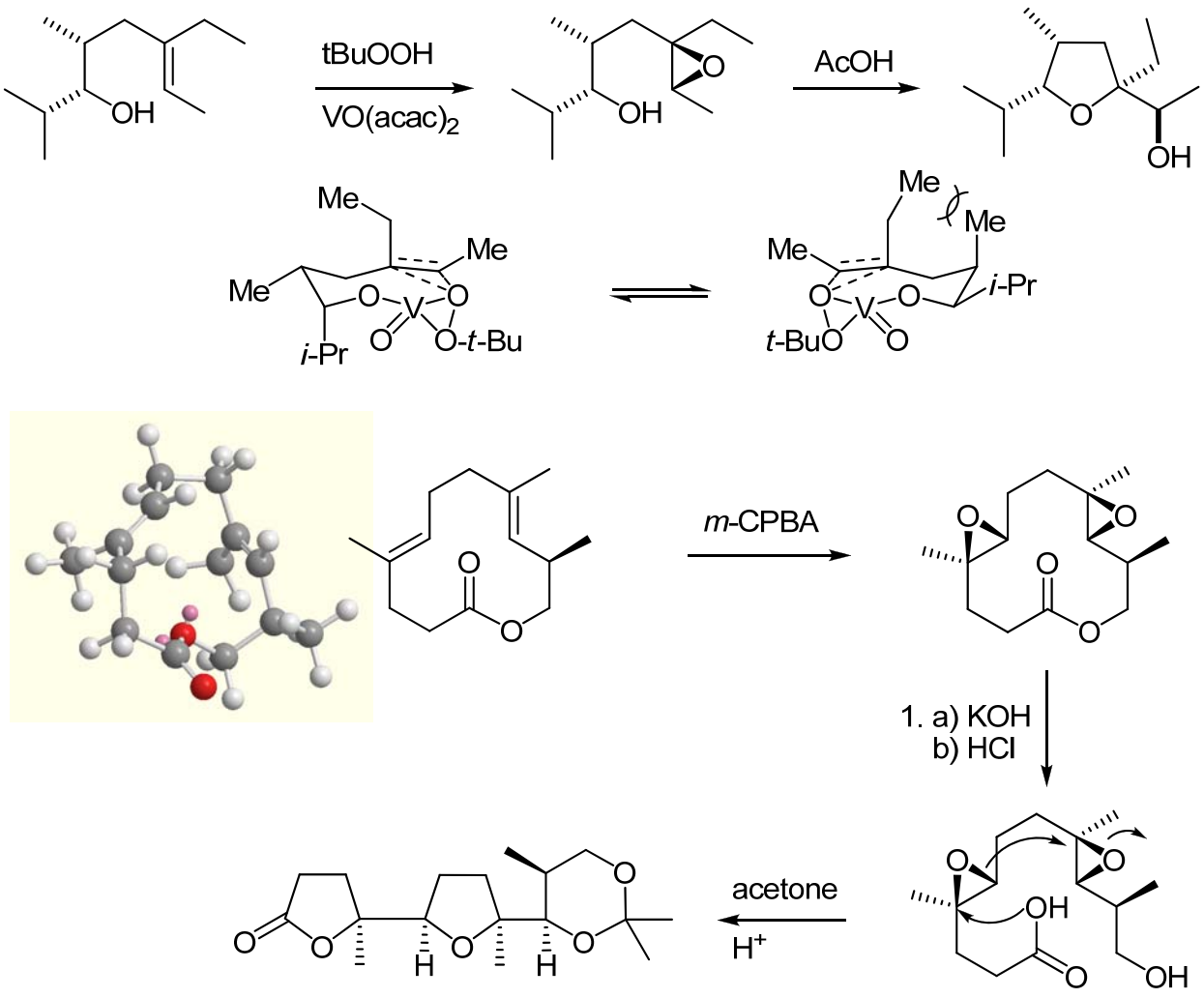
On the other hand, prior epoxidation of the alkene leads to an epoxy-alcohol which can also be cyclized with acids or other electrophiles (Scheme III.5.7). The stereochemistry of the epoxide, which is itself derived from the chiral induction of the asymmetric alcohol, determines the stereochemistry of the final tetrahydrofuran. Such epoxide cyclization can be performed on polyepoxide substrates in which each of the epoxide is sequentially opened in a cascade reaction much like in the biosynthesis of polyethers (see Scheme III.5.1). For example, at the bottom Scheme III.5.7, a macrocycle is stereoselectively epoxidized, thanks to its preferred conformation that exposes one face of each double bond. The macrolactone is then opened and the diepoxide is cyclized under acidic conditions. This completes our discussion on the acetogenins.



Scheme III.5.5



Scheme III.5.6



Scheme III.5.7

IV. Shikimic Acid Metabolites

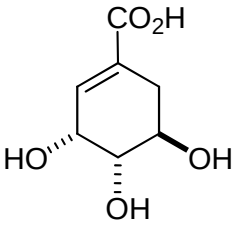
IV.1. Biosynthesis of Shikimic Acid Metabolites.

IV.1.1. Biosynthesis of Shikimic and Chorismic Acid.

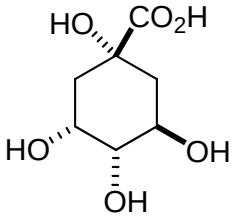
Shikimates, or shikimic acid metabolites, form a widely spread family of aromatic compounds, of which cinnamic acid, coumarin, lignin (hard consistency of wood), alkaloids, phenylalanine and aromatic amino acids are members. Moreover, many shikimates find their way into important biomolecules of mixed origin where part of the structure is derived from shikimic acid and other carbons come from the incorporation of molecules derived from other pathways, amino acids, isoprenoids, polyketides and others. For example, the vitamin K and E contain oxidized aromatic rings derived from shikimic acid and aliphatic amino acids as well as an isoprenoid side chain. Figure IV.1.1 gives representative examples of shikimates. Shikimic acid was first isolated in 1886 from a Japanese plant of the same name, long before its involvement in the biosynthesis of natural products was known. It was discovered that shikimic acid in the diet could replace essential amino acids (phenylalanine, tyrosine and tryptophan). Later, shikimic acid was found to be the biogenetic precursors of these essential amino acids.

The shikimic acid pathway is most important in higher plants and algae where its precursor D-erythrose-4-phosphate (DEP) and phosphoenol pyruvate (PEP) are found in large quantities. These two compounds are of utmost importance in the photosynthetic process of higher plants because they are intermediates in the recycling of ribulose-1,5-diphosphate, the molecule up-taking the carbon dioxide. Some of the DEP and PEP molecules get diverted for the synthesis of shikimates and lignin. The process by which this happens is summarized in Scheme IV.1.1. The first step involves an aldol condensation of DEP with a molecule of phosphoenol pyruvate to give compound **1**. This is a stereospecific process which is followed by the elimination of a molecule of phosphate to give directly the enol form of compound **2** which can undergo an intramolecular aldol condensation affording dehydroquinic acid. Then two routes are possible, one leading directly to quinic acid by an NADH mediated reduction. Quinic acid is not further metabolized but is used in the esterification of phenols and other shikimates (thus appearing as their quinate ester). The other route leads to shikimic acid via an elimination of water prior to NADH reduction. Shikimic acid is the precursor of almost all shikimates. The ways in which shikimic acid is transformed into various shikimates are both numerous and complex. Many of the shikimates have important biological roles as such, others are part of important vitamins,

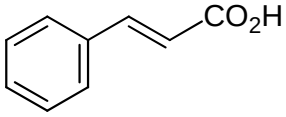
proteins, or other metabolites. In all biosynthetic pathways, shikimic acid is first transformed to chorismic acid which occupies a branching point in shikimate biogenesis.



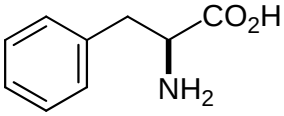
Shikimic acid



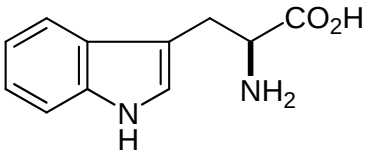
Quinic acid



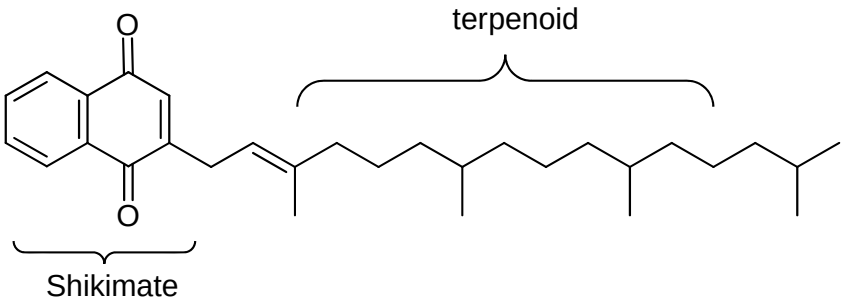
Cinnamic acid



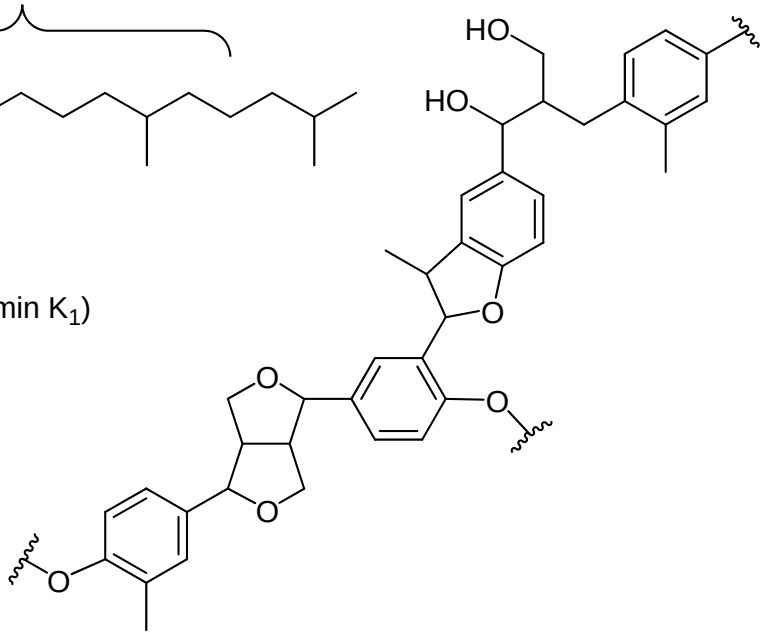
Phenylalanine



Tryptophan



Phylloquinone (Vitamin K_1)

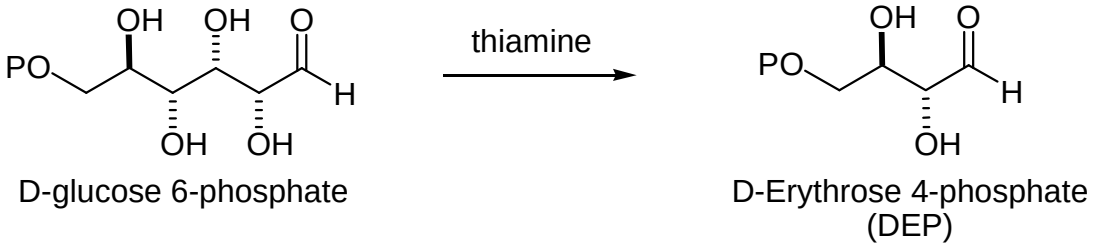


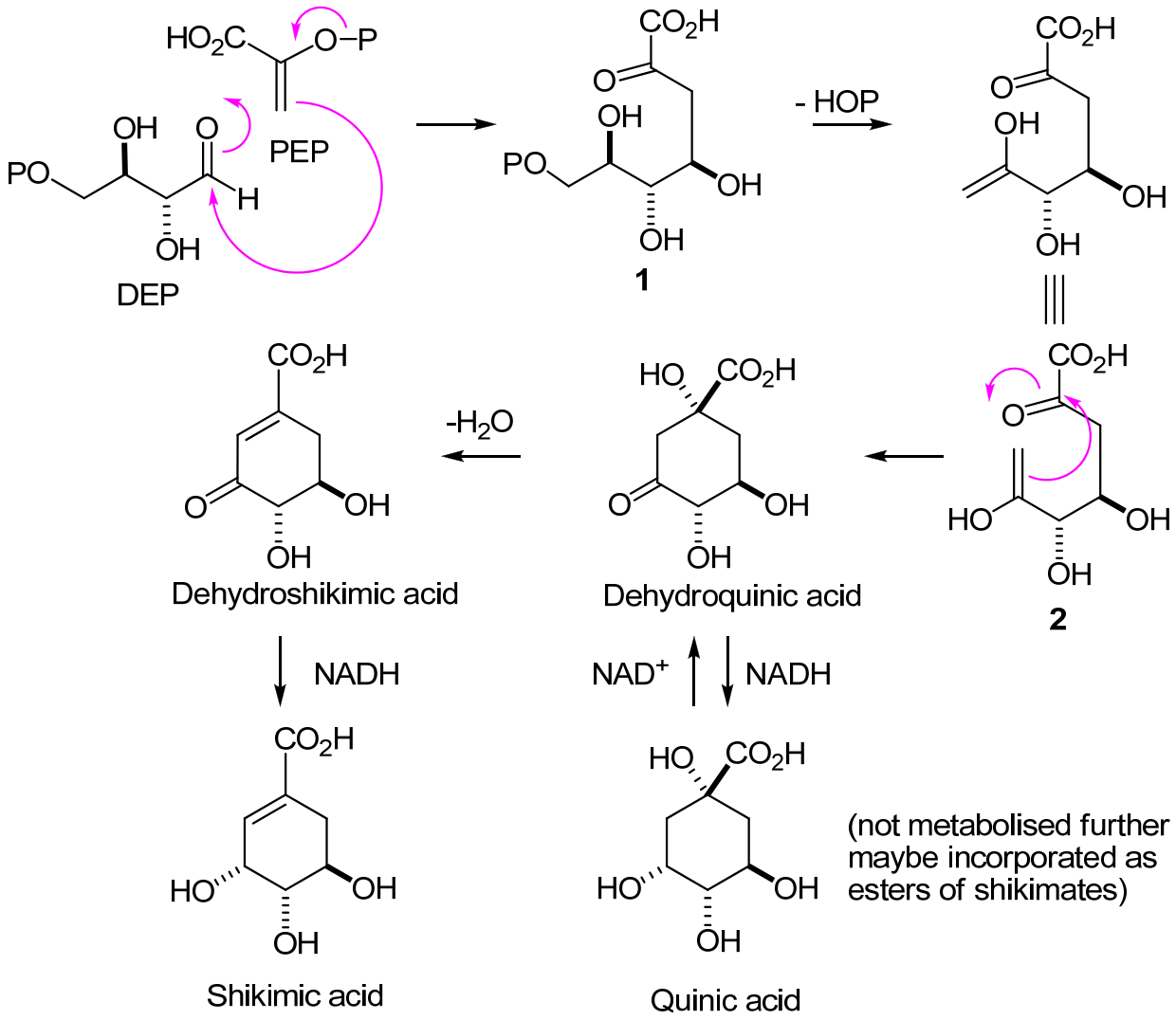
Component of lignin

Figure IV.I.1

APP.IV.1

Proposez un mécanisme pour la transformation du D-glucose 6-phosphate (issu de la photosynthèse) en D-Erythrose-4-phosphate.





Scheme IV.1.1

The principal pathways are described in Figure IV.1.2. In turn each of these steps is detailed in subsequent schemes. In Scheme IV.1.2, the C5-alcohol of shikimic acid is condensed with a molecule of PEP to give compound **4** and the subsequent elimination of a molecule of phosphate gives the enol ether **5**. A second elimination of a phosphate molecule (via a E_2' mechanism) affords chorismic acid. Labelling studies showed that only the pro-R hydrogen is eliminated. Chorismic acid is then transformed into important amino acids such as phenylalanine (which is also the precursor of other shikimates), tyrosine, tryptophan (which is the precursor for indoleacetic acid, an important growth hormone that controls cell elongation), the neurotransmitter serotonine (responsible for chemical communication between synapses), and *p*-aminobenzoic acid.

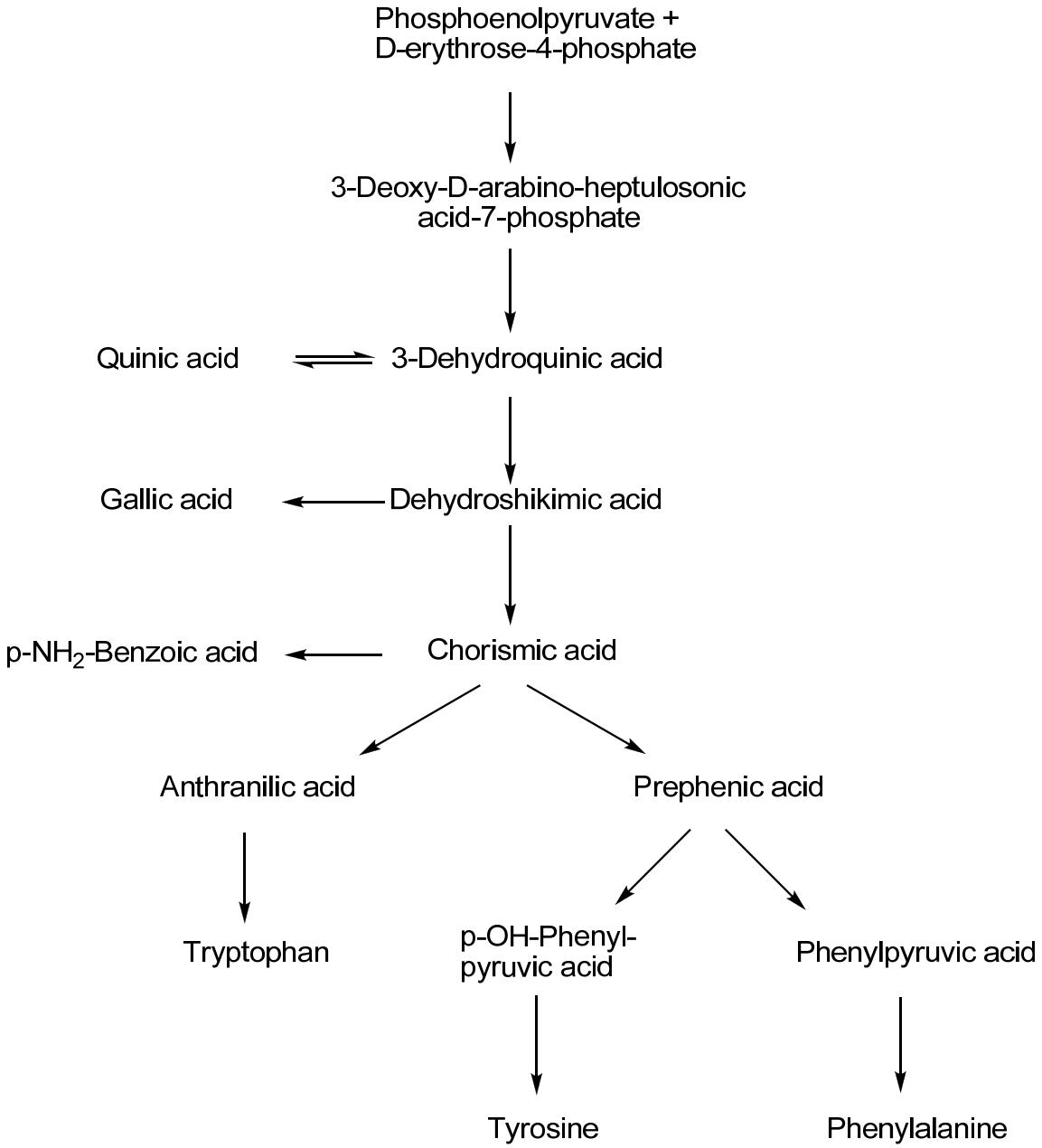
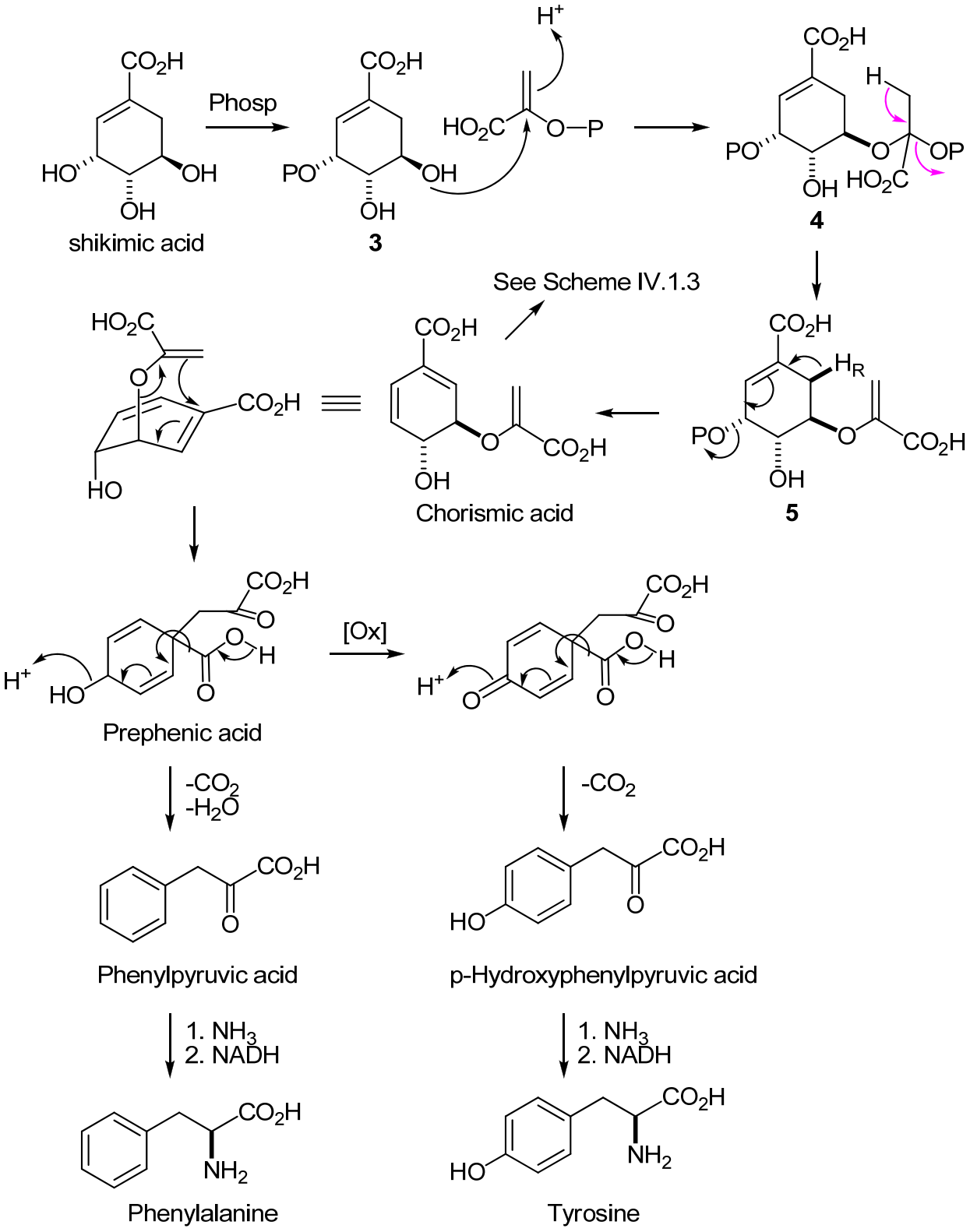

Phosphoenolpyruvate +
D-erythrose-4-phosphate
3-Deoxy-D-arabino-heptulosonic
acid-7-phosphate
Quinic acid
3-Dehydroquinic acid
Gallic acid
Dehydroshikimic acid
p-NH2-Benzoic acid
Chorismic acid
Anthranilic acid
Prephenic acid
Tryptophan
p-OH-Phenyl-
pyruvic acid
Phenylpyruvic acid
Tyrosine
Phenylalanine

Figure IV.1.2

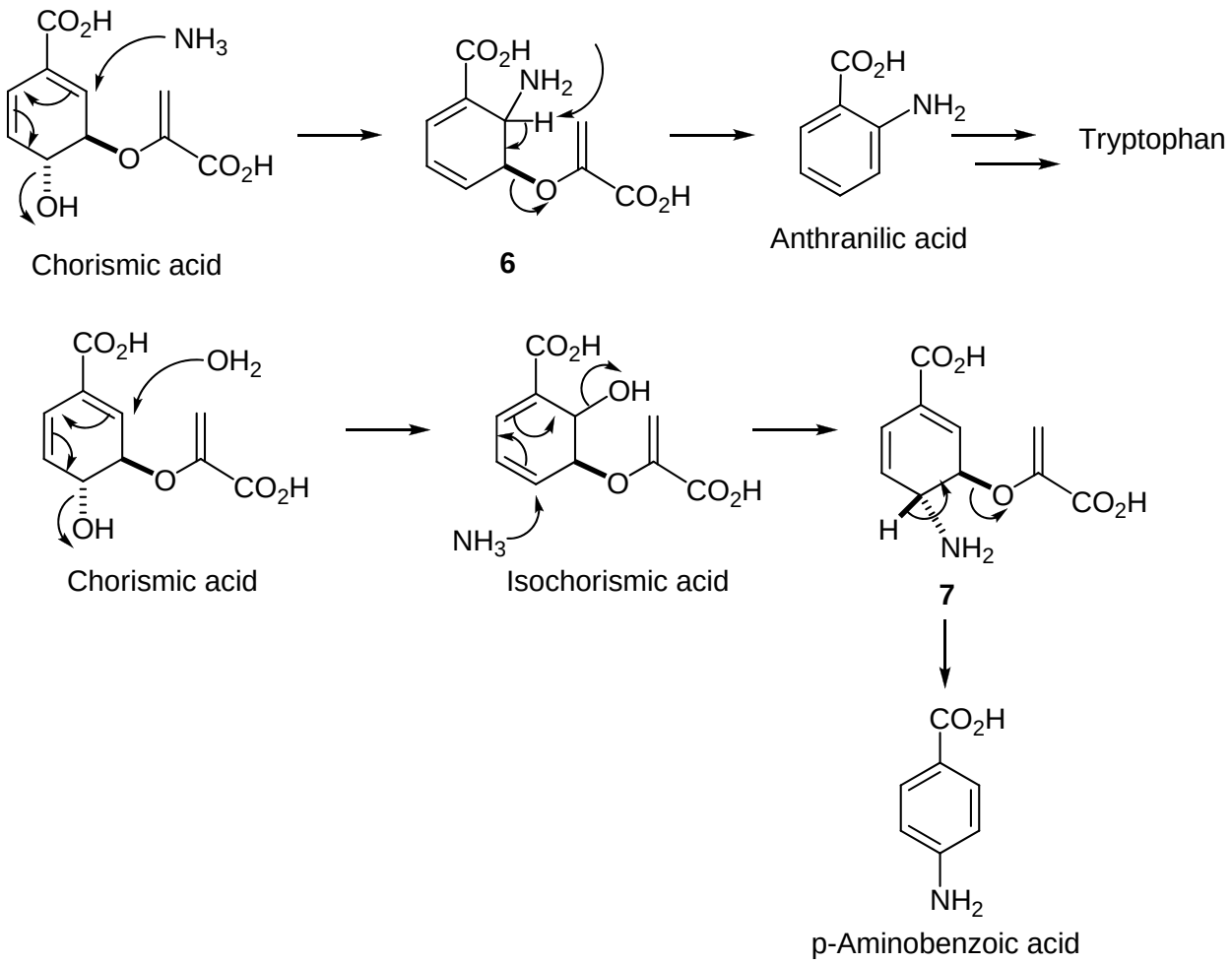

shikimic acid
Phosp
3
4
See Scheme IV.1.3
Chorismic acid
5
Prephenic acid
[Ox]
-CO2
-H2O
-CO2
Phenylpyruvic acid
p-Hydroxyphenylpyruvic acid
1. NH3
2. NADH
1. NH3
2. NADH
Phenylalanine
Tyrosine

Scheme IV.1.2

IV.1.2. Biosynthesis of the Aromatic Amino Acids.

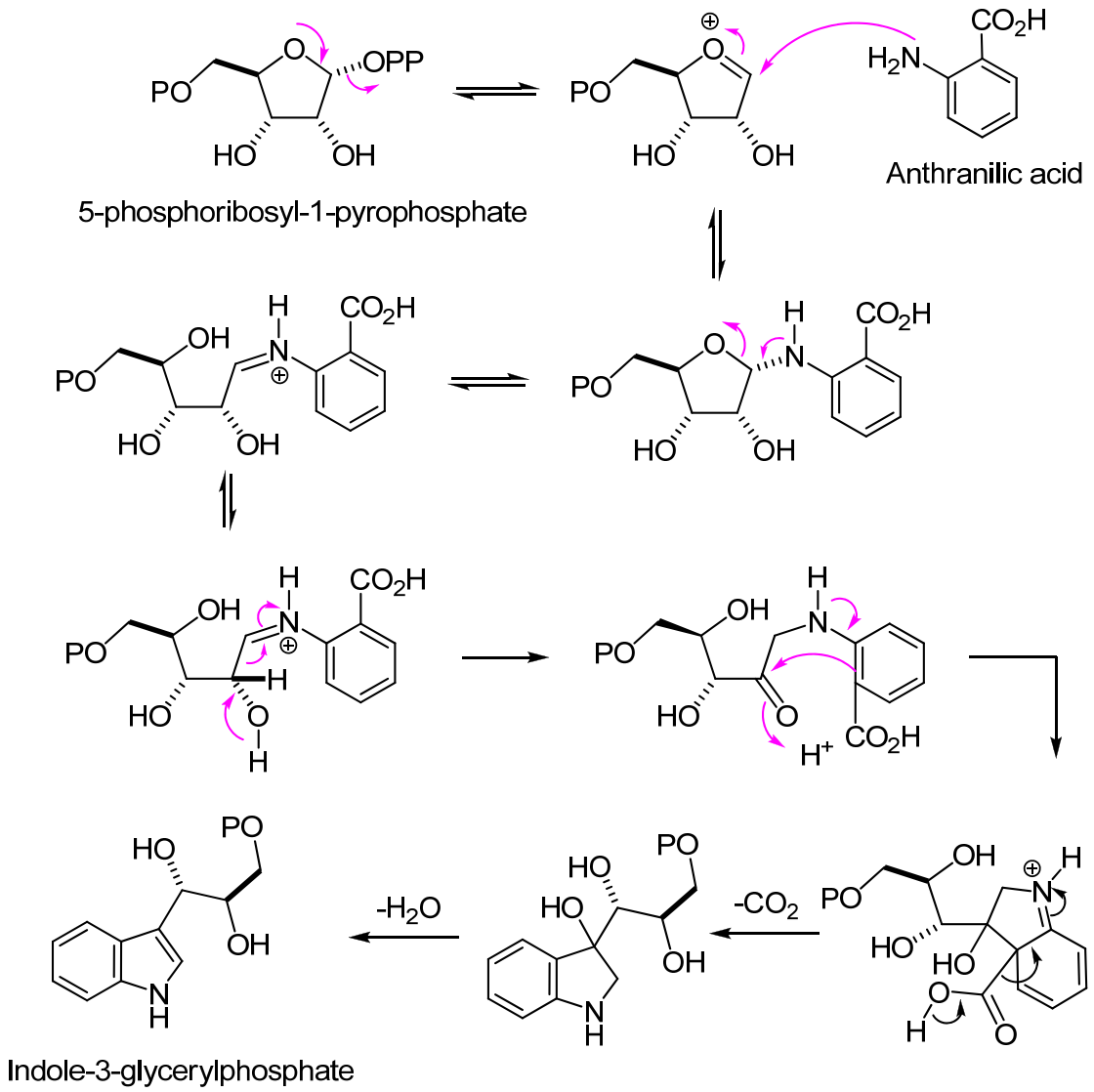
The first step in the biosynthesis of phenylalanine and tyrosine is a biological Claisen rearrangement or a 3,3-sigmatropic rearrangement (Scheme IV.1.2). The existing 6-membered ring must adopt a 1,2-diplanar conformation to bring the two double bonds in close proximity. However, the pseudo-6-membered ring of the transition state of the Claisen rearrangement adopts a pseudo-chair conformation. Nevertheless, the end result is the formation of prephenic acid (sometimes called pre-phenylalanine). From there, two possible biosynthetic pathways are available. If carbon dioxide is lost with concomitant elimination of water, then phenylalanine is produced. If, on the other hand, oxidation precedes elimination of CO_2, then water is not lost, but instead a phenol ring is produced affording tyrosine. Phenylalanine and tyrosine are amino acids and thus take part in the biosynthesis of essential proteins and enzymes. However, in addition to that role, they are precursors of many different shikimates having an aromatic ring with a 3 or 2 or 1 carbon chain (herein named shikimates of ArC_3 or ArC_2 or ArC_1 type). Moreover, together with aliphatic amino acids, they also are the precursor of numerous alkaloids which are the topic of the next chapter.

The separate pathway that leads to tryptophan via anthranillic acid is shown in Scheme IV.1.3. Firstly, ammonia adds in a Michael fashion to position 2 of chorismic acid. Concomitant elimination of water gives compound **6** which can suffer elimination of the PEP molecule to give anthranilic acid (or *o*-aminobenzoic acid). How the latter is transformed into tryptophane is shown in Scheme IV.1.4. When water adds at position 2 of chorismic acid, instead of ammonia, isochorismic acid is produced. 1,6-Addition of ammonia with elimination of water in isochorismic acid (exactly analogous to the formation of anthranilic acid) gives compound **7** which can also suffer elimination of the PEP molecule to give *p*-aminobenzoic acid. The latter is the precursor of some ArC_1 type shikimates. However, will shall not discuss further its implication in the biosynthesis of shikimates.



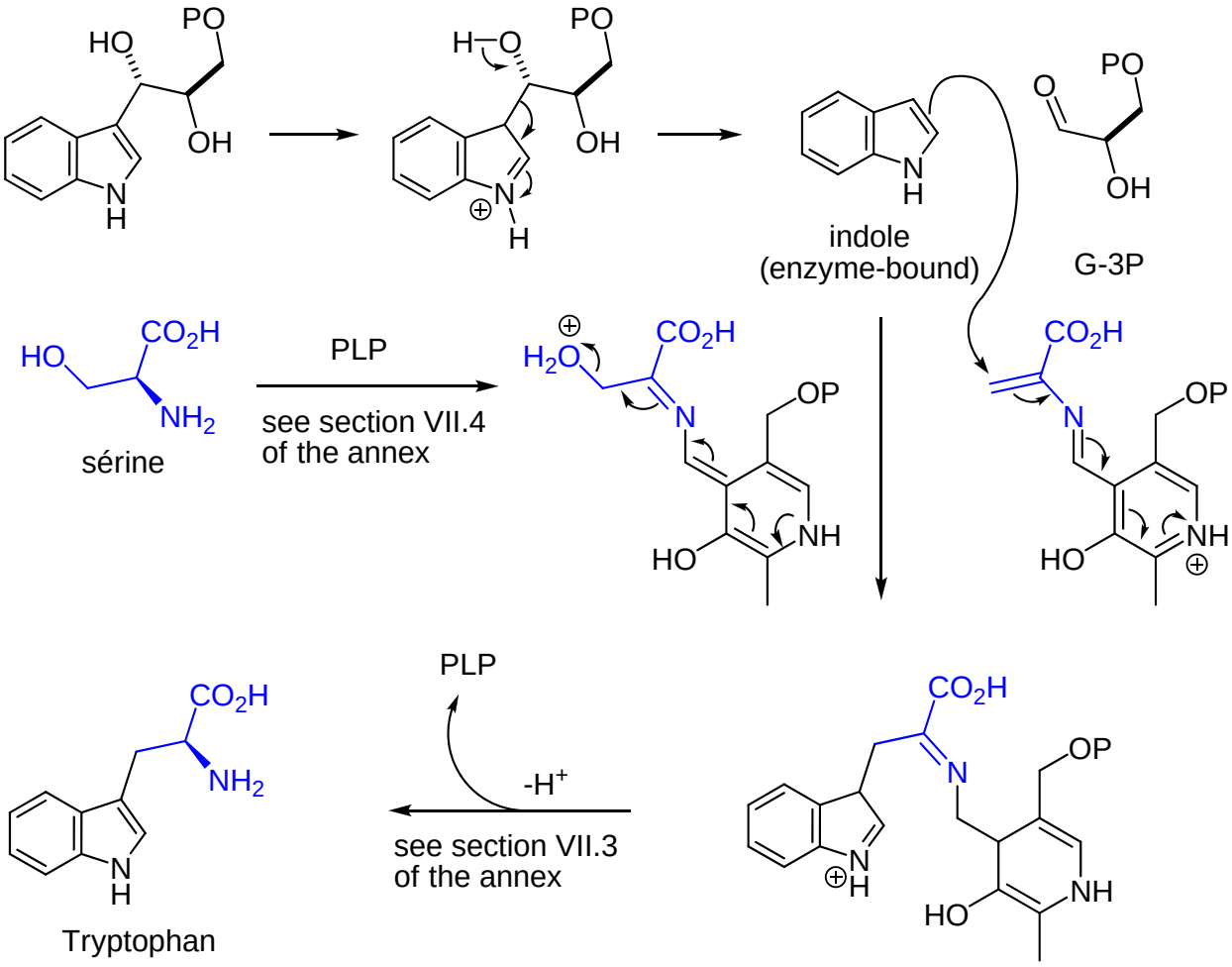
Scheme IV.1.3

The synthesis of tryptophan is interesting in that it involves three separate sources of carbons, namely anthranilic acid, a ribose, and serine (Scheme IV.1.4). First a ribosyl tetraphosphate is aminated with anthranilic acid via a S_N1 reaction. Then, an enzyme-catalyzed opening of the ribose followed by a hydride shift creates an α-amino ketone that undergoes a Friedel-Craft alkylation (enzyme-catalyzed of course) the product of which, in turn, suffers the loss of CO_2 to give the indoline ring. Elimination of water sets up the nucleophilic indole ring system, indole 3-glycerylphosphate.



Scheme IV.1.4

Then, protonation and a retro-Mannich reaction takes place and extrude glyceraldehyde 3-phosphate (G-3P) from the molecule indole, which remains enzyme-bound. Almost invariably, the indole rings attack at the 3-position (larger electronic density). This is no exception and the PLP (pyridoxal phosphate, see section VII.3 of the annexe) derivative of serine serves as electrophile (Scheme IV.1.5). Their reaction creates an iminium ion that looses a proton to give the indole nucleus of tryptophan, after enzymatic hydrolysis of PLP (see section VII.3 of the annex)


HO
PO
OH
N
H
H–O
PO
OH
N
H
indole
(enzyme-bound)
O
PO
OH
G-3P
HO
CO2H
NH2
sérine
PLP
see section VII.4
of the annex
H2O
CO2H
N
OP
HO
NH
CO2H
N
OP
HO
NH
PLP
-H+
see section VII.3
of the annex
CO2H
NH2
N
H
Tryptophan
CO2H
N
OP
N
H
HO
NH

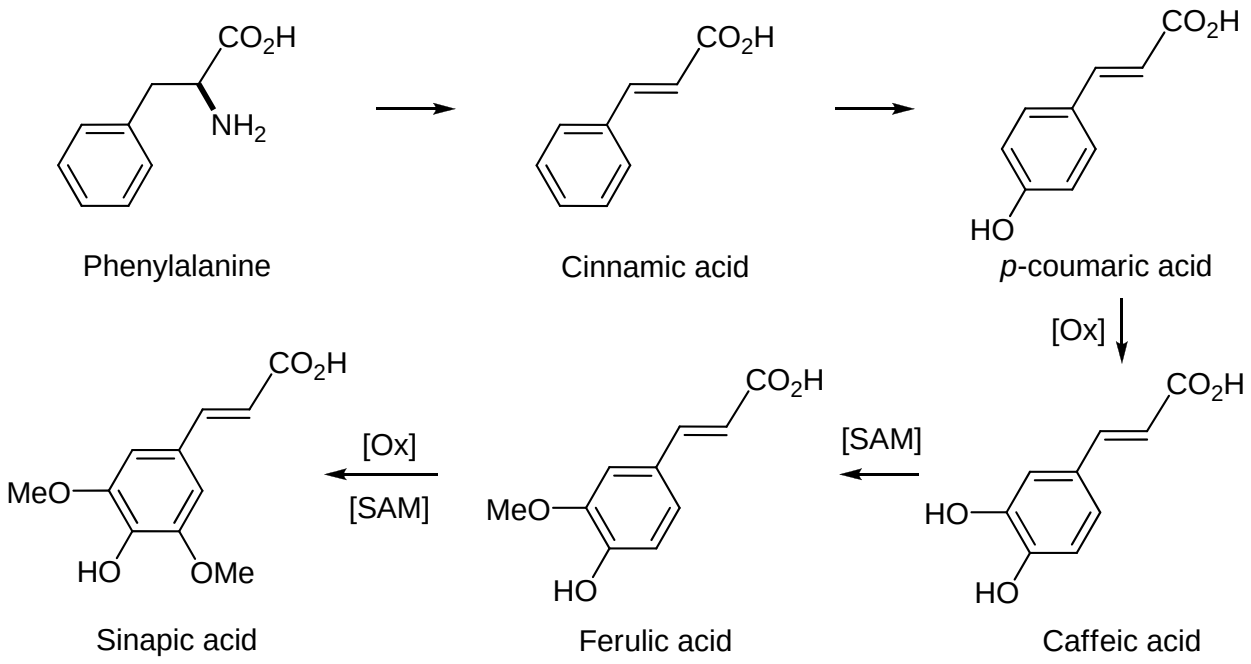
Scheme IV.1.5

IV.2. ARC_x-Type Shikimates.

IV.2.1. Biosynthesis of ArC_3-Type Shikimates and Coumarins.

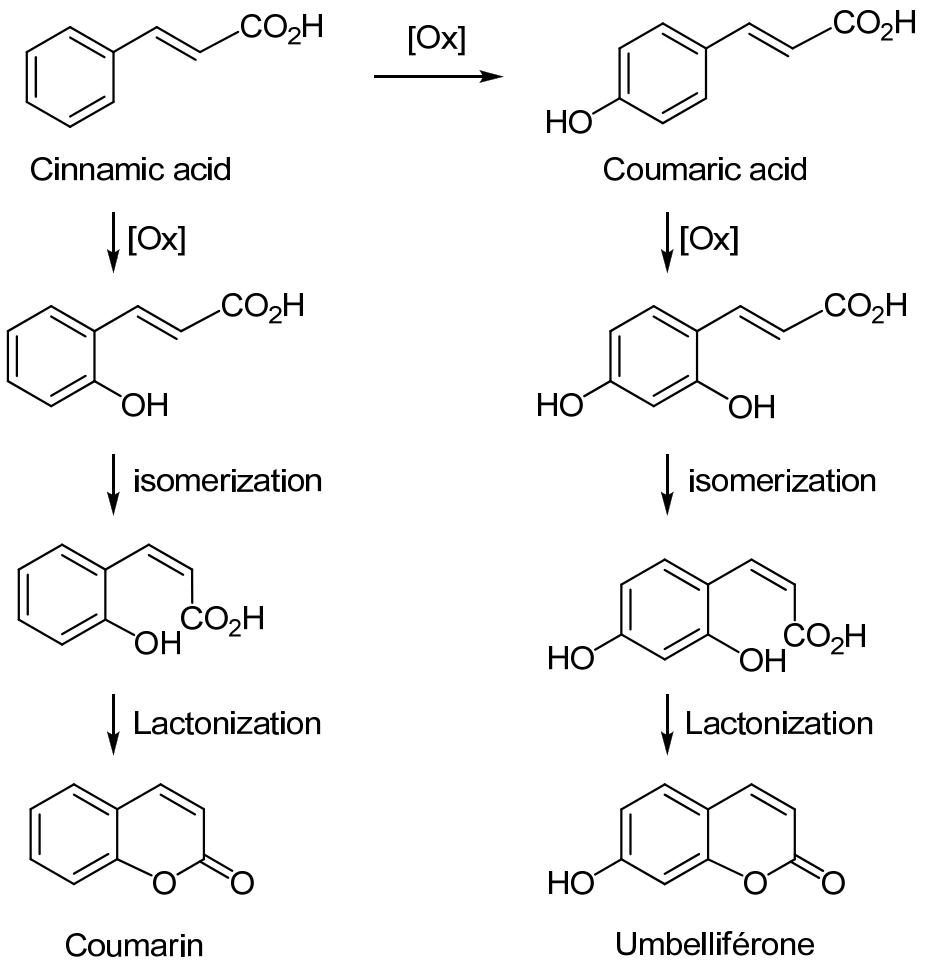
IV.2.1a. Coumarins.

The ArC_3 and coumarins are the most commonly encountered shikimates. They are derived from phenylalanine via cinnamic acid by loss of ammonia (Scheme IV.2.1). Oxidation of the aromatic ring can take place at different stages of the biosynthesis of shikimates. The enzymatic machinery of the particular organism dictates which pathway is going to be followed. The biosyntheses of caffeic acid, ferulic acid, and sinapic acid are shown in Scheme IV.2.1. Each of these natural products are found in higher plants. Caffeic acid is found as caffeyl quinate in coffee beans and is actually widely distributed in the plant world. Some, like sinapic acid, are incorporated in lignin (see section IV.2.4).



Scheme IV.2.1

Coumarins are widely distributed in the plant world where they serve a variety of different biological functions. For example, umbelliférone is found in lemons and German chamomille whereas coumarin found in sweet clover is responsible for the smell of freshly mown hay (Figure IV.2.1). Their respective biogeneses are shown in Scheme IV.2.2. As stated earlier, oxidation can take place at various stages and the pathways were proven by labelling studies.



Scheme IV.2.2

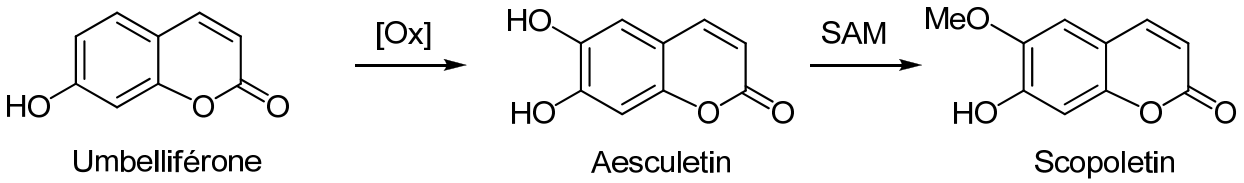
Sweet clover in hay

German Chamomille

Figure IV.2.1

Scopoletin is a germination stimulant and it lowers the blood pressure. It is derived from umbelliférone via the diol aesculetin (Scheme IV.2.3). The Hawaiian mengkudu fruit is an excellente source of scopoletin. The biological activity of many coumarins seemed to be derived

from their ability to inhibit blood clothing. Inhibition of vitamin K, which is responsible for the clothing, seems to be the major pathway.

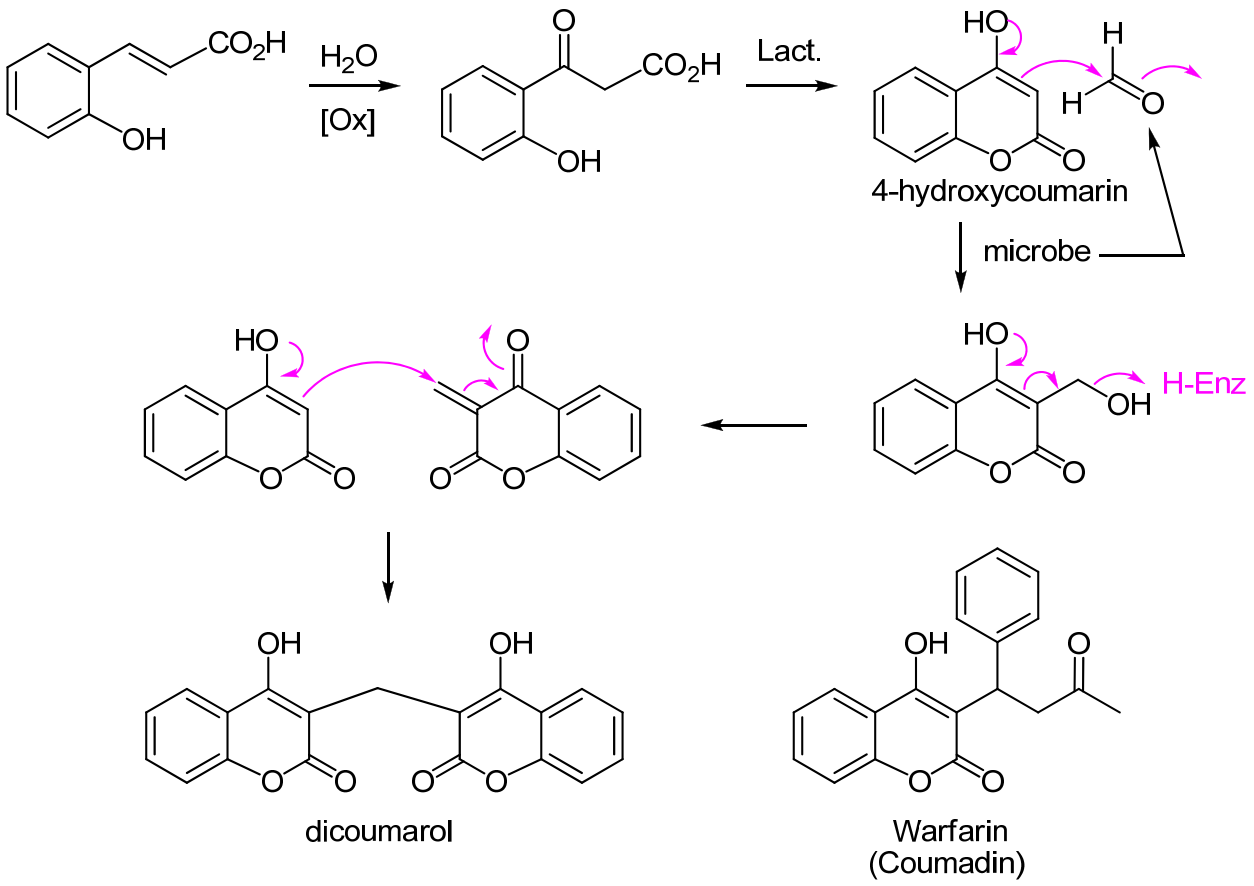


Umbelliférone Aesculetin Scopoletin

Mengkudu (*Hawaii*)

Scheme IV.2.3

Coumarin is transformed into dicoumarol in fermenting hay (Scheme IV.2.4). *m*-Hydroxycinnamic alcohol is hydrated, oxidized and lactonized. Micro-organisms produce formaldehyde in their digestion, which reacts with the 4-hydroxycoumarin produced. After elimination of the alcohol to procure a highly electrophilic Michael acceptor, a second 4-hydroxycoumarin molecule reacts to give dicoumarol. The latter compound is an anticoagulant so potent that livestock eating rotten hay will die of hemorrhage. Similarly, the anticoagulant warfarin (Coumadin^{TM}) is a strong rodenticide that has been used successfully in controlling rat populations until they developed resistance to the compound. It is now used clinically to help lower the risks of cardiac arrests from thrombose by inhibiting vitamin K_1.

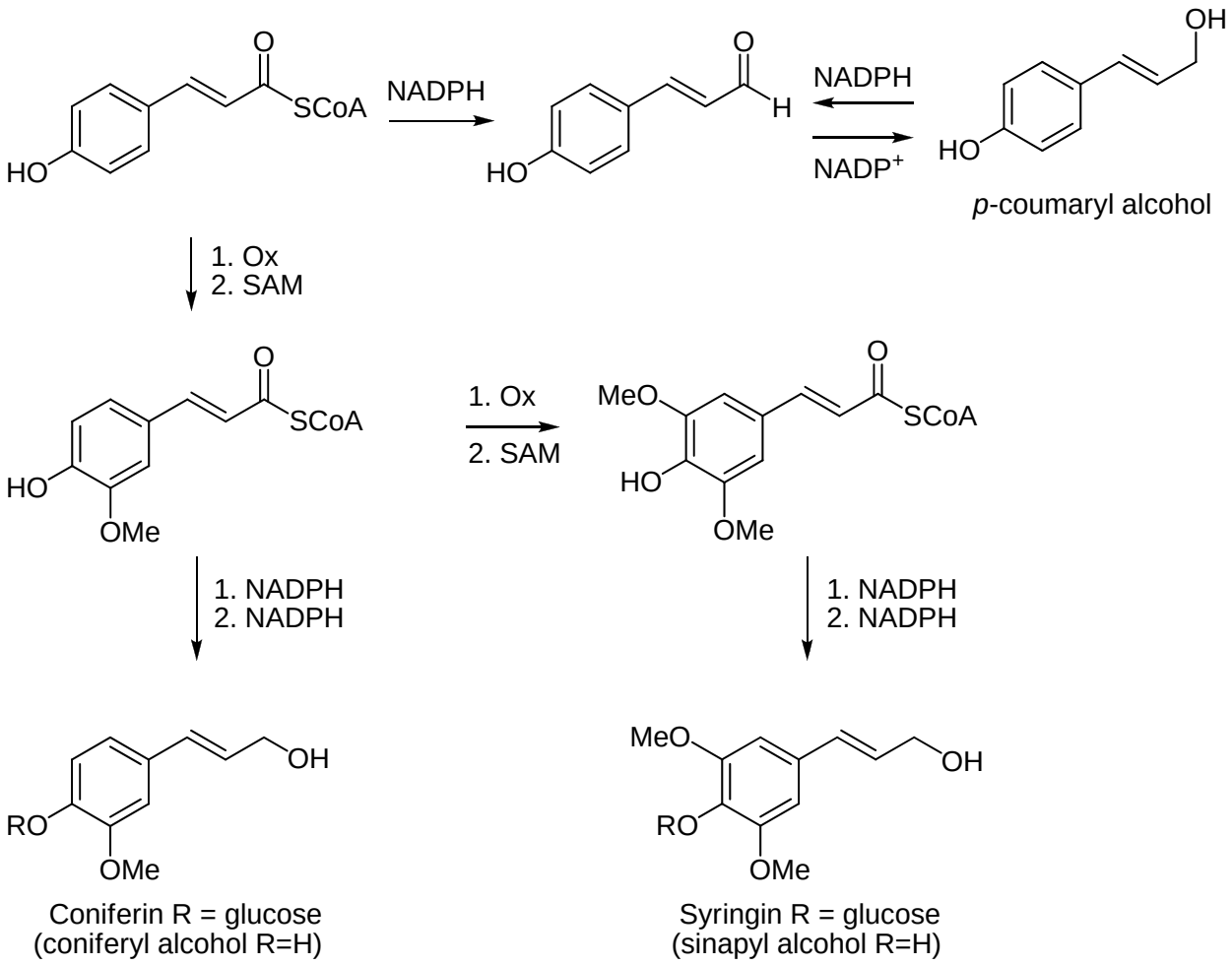


Scheme IV.2.4

IV.2.1b. Lignins and Lignans.

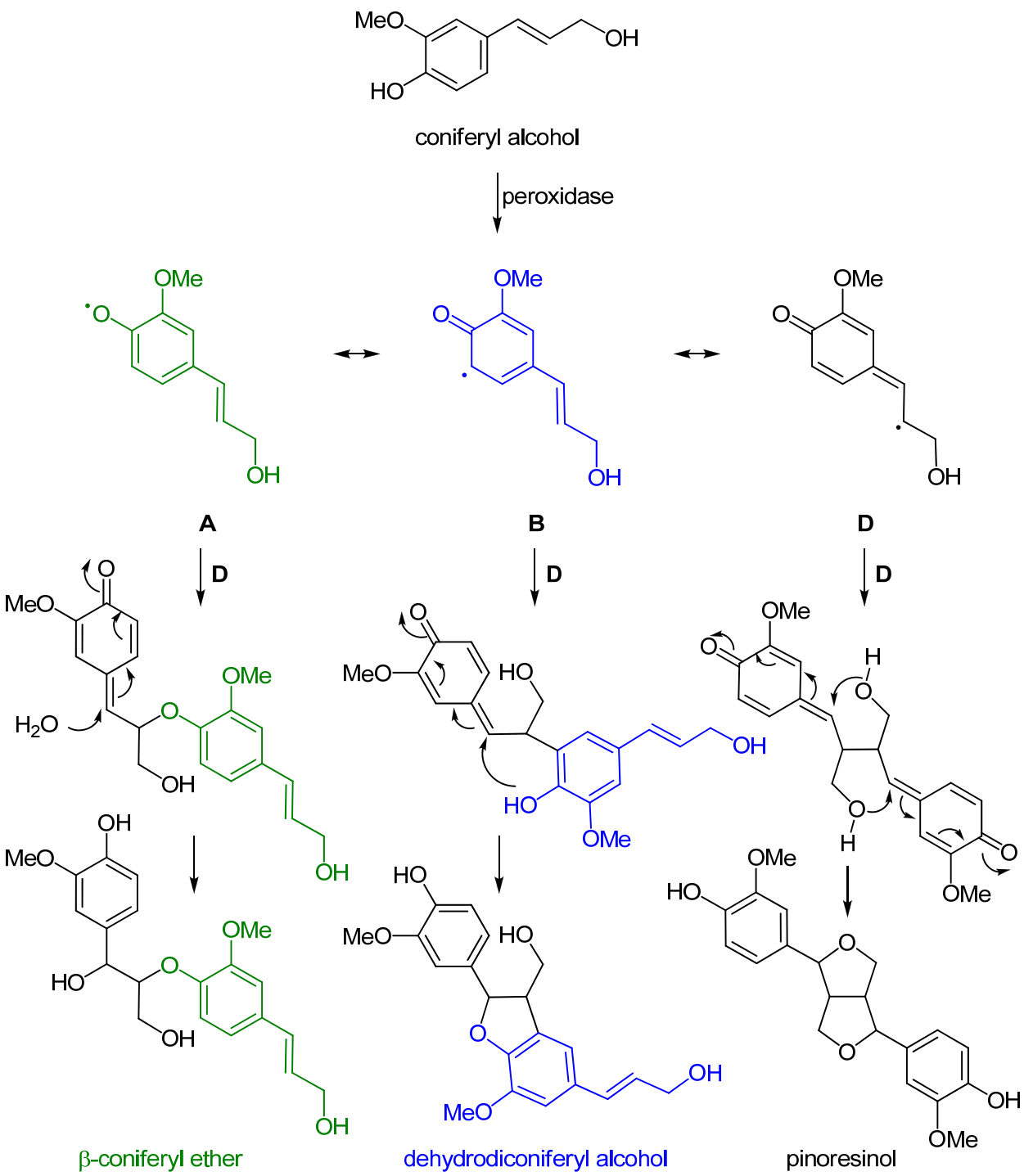
In woody plants a large proportion of all ArC_3 are incorporated in lignins, a sort of cement that keeps the cells in a rigid framework we call wood. Indeed, lignins account for a large proportion of all aromatic rings in the biosphere. They are polymeric materials held within a matrix of cellulose microfibrils, and appear to be utilized in strengthening the cell wall of the plant against external physical and chemical stresses. The three most important monomeric species utilized in the polymerization process are *p*-hydroxycinnamyl alcohol, coniferyl alcohol, and synapyl alcohol (Scheme IV.2.5). Their mode of biosynthesis is shown in the same scheme. The compounds accumulate as their ß-glucosides, and these are later hydrolyzed and the monomers oxidatively polymerized. Carbon-carbon bonds are formed by oxidative phenolic coupling (see section III.4.4) and carbon-oxygen bonds are formed by reductive coupling and other reactions. The main pathways leading to three principal dimers (lignans) involving different types of coupling are represented in Scheme IV.2.6. β-coniferyl ether comes from a

coupling between the oxygen of the phenol radical and the C3-chain, dehydrodiconiferyl alcohol involves a coupling between the ortho position and the chain radical, while two chain radicals come together to give pinoresinol.



Scheme IV.2.5

Polymerisation following a mixture of these coupling types (the β-coniferyl type coupling is usually the major one) appears to be statistically driven rather than enzymatically controlled since lignin is always racemic (optically inactive). In contrast lignans, which are dimers of the same monomers are invariably optically active. A part structure for hardwook lignin is shown in Figure IV.2.2 while that for softwood lignin is shown in Figure IV.2.3.


coniferyl alcohol
peroxidase
A
B
D
D
D
β-coniferyl ether
dehydrodiconiferyl alcohol
pinoresinol

Scheme IV.2.6

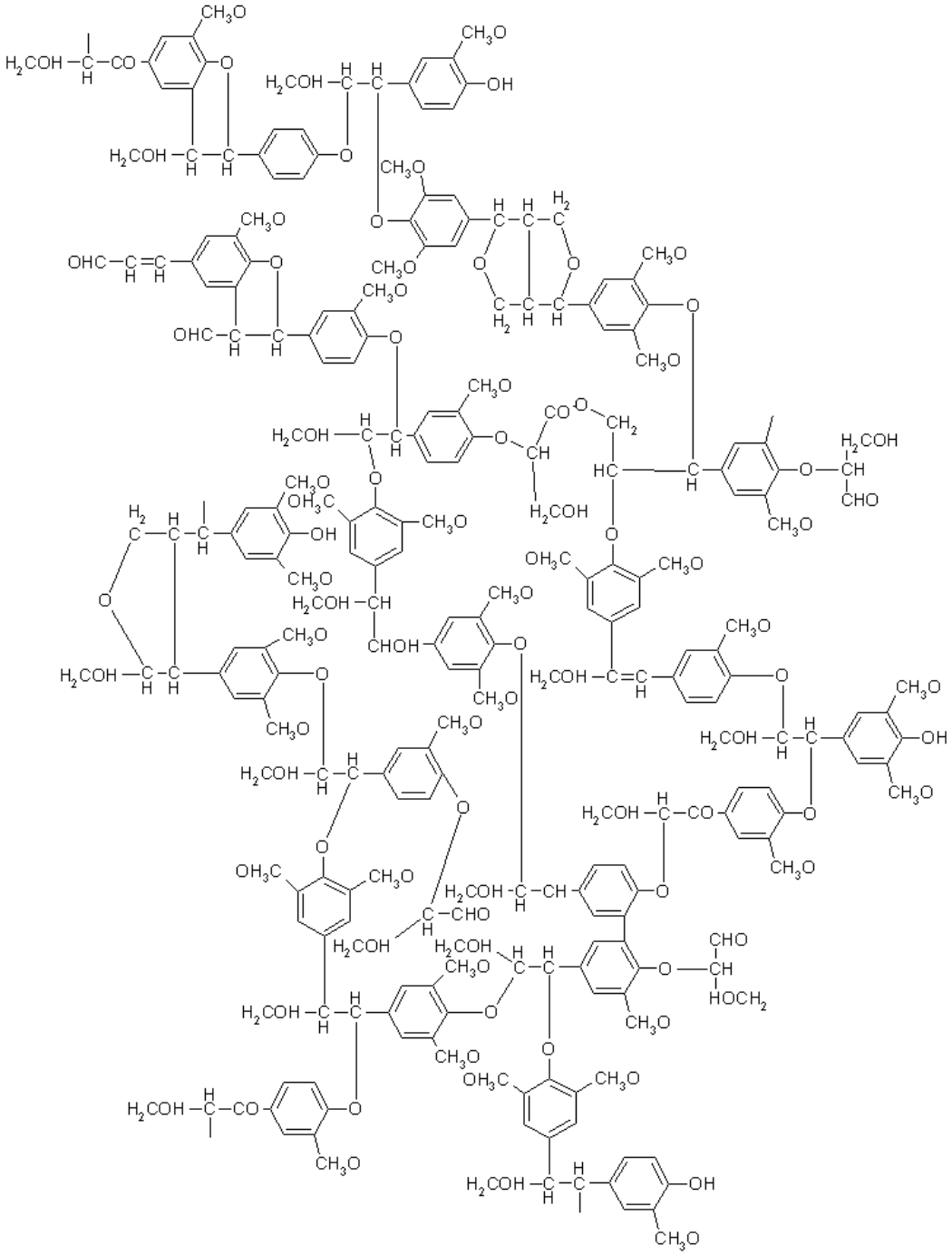


Figure IV.2.2 (Harwood lignin)

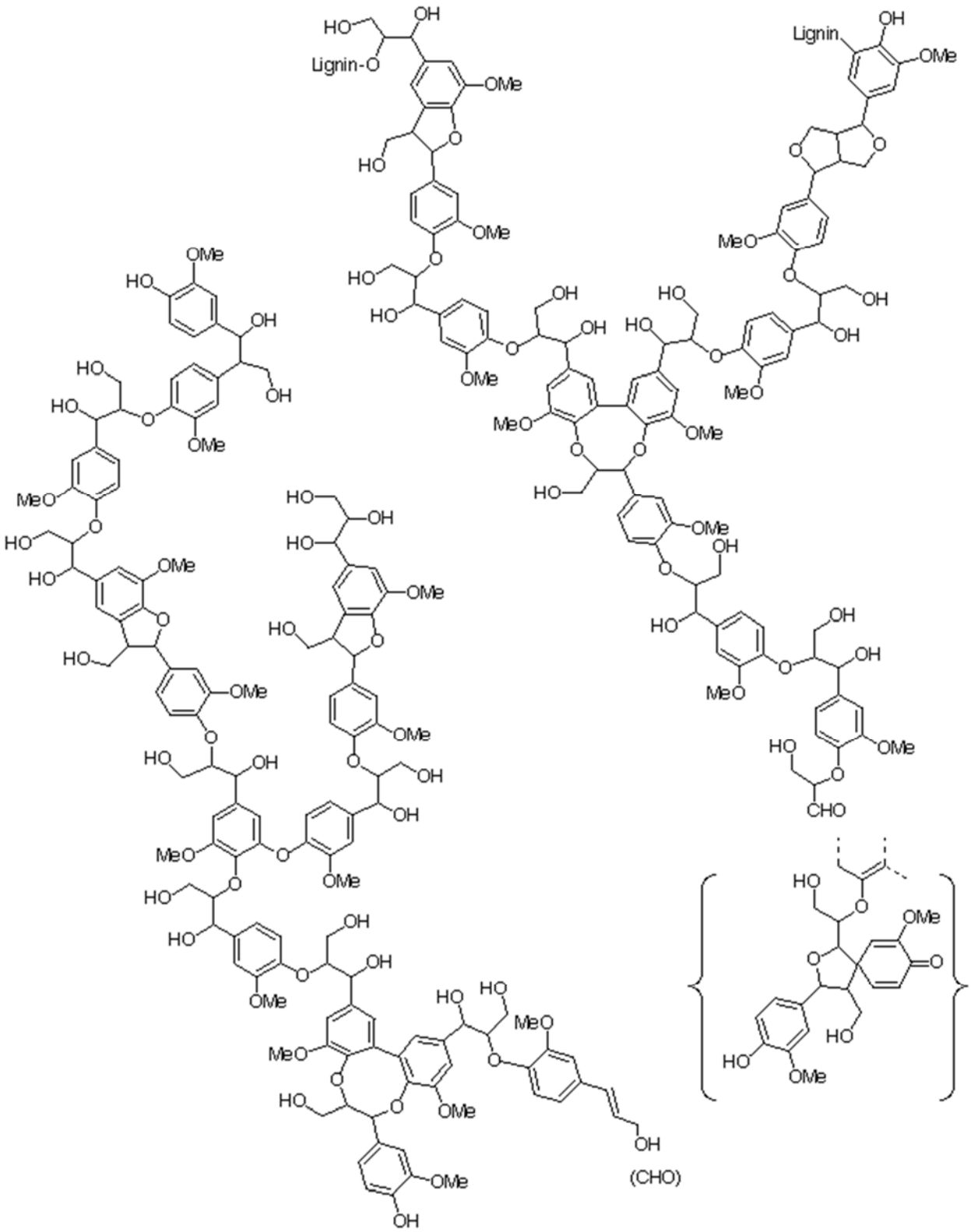
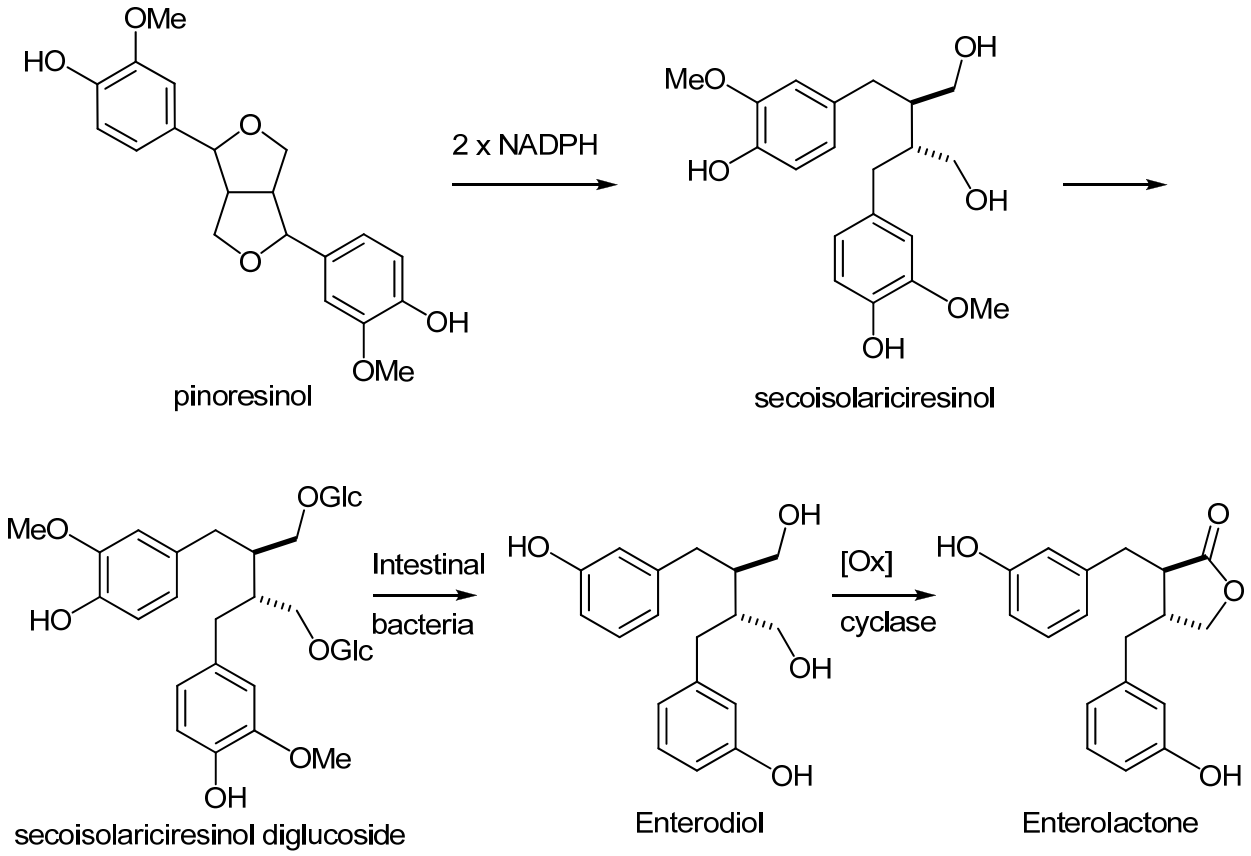


Figure IV.2.3 (softwood lignin)

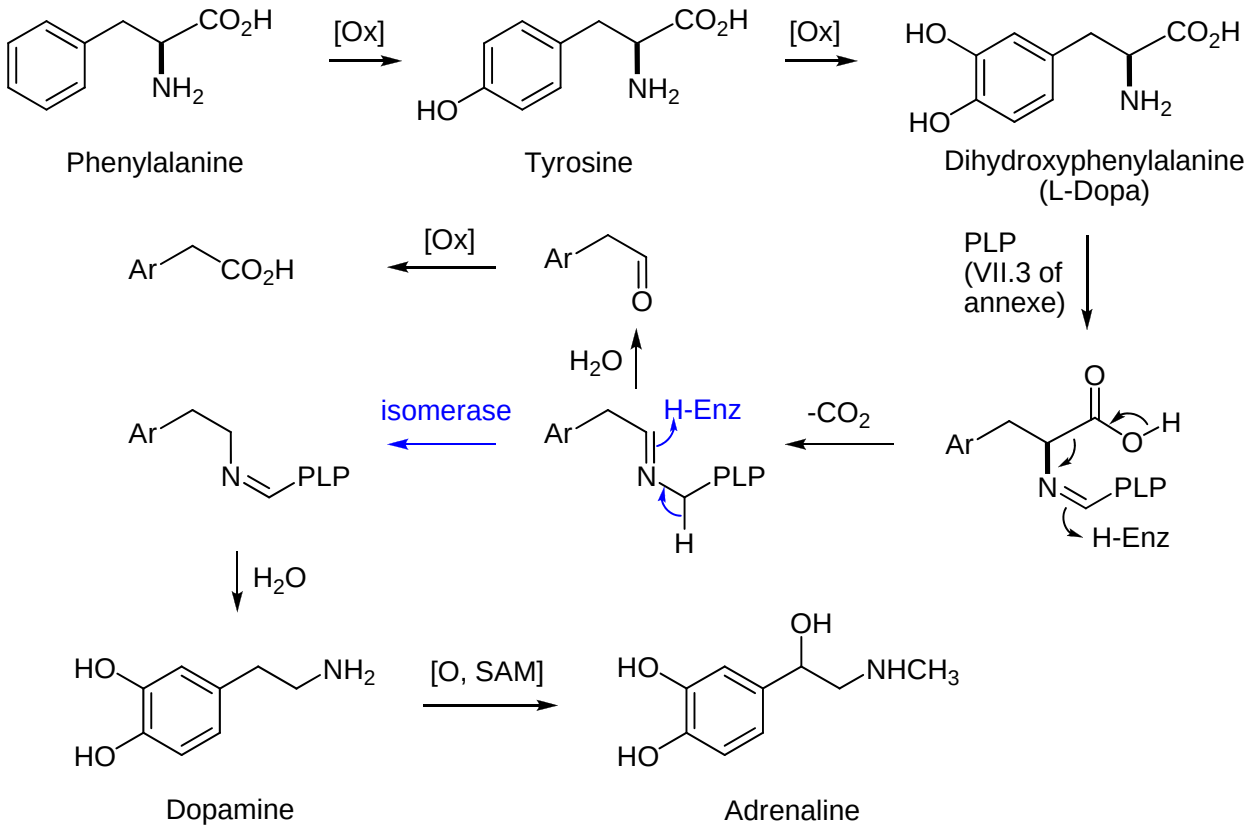
Enterolactone is derived from the lignan pinoresorcinol by the action of intestinal bacteria on ingested fibre-rich food source such as flaxweed (Scheme IV.2.7). Enterolactone (as a racemic mixture) has been found in human blood, bile and urine at level comparable to steroid hormones. This molecule is oestrogenic, meaning that it mimics the biological activity of oestrogenic hormones. The same is true of many flavonoids. Women who ingest a diet rich in flavonoids and/or lignans experience diminished effect during menopause and have a reduced risk of breast cancer. Women experience a change in levels of enterolactone during the second half of the menstrual cycle and during pregnancy. No explanation for the change in levels of enterolactone has been offered to date. Other lignans such as podophyllotoxin and its analogues are in clinical trial for human therapy against several types of cancers and in particular testicular tumors and lung cancer.



Scheme IV.2.7

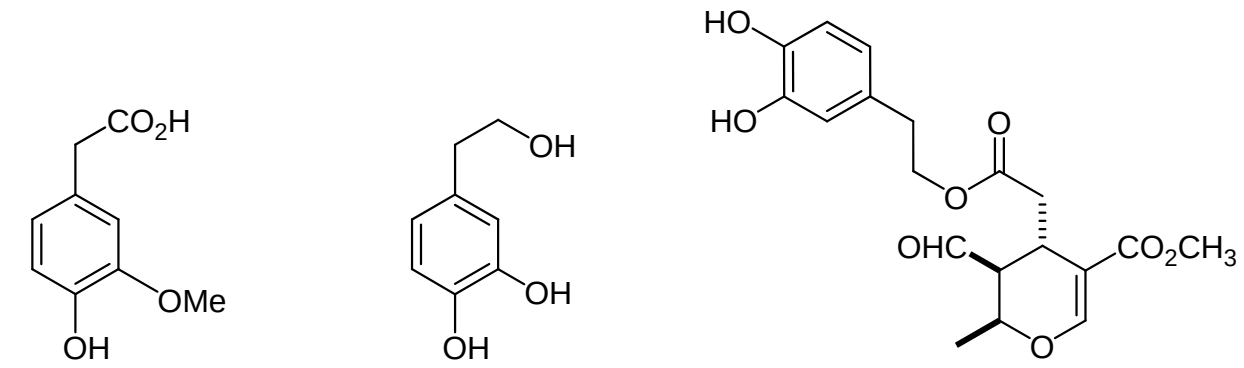
IV.2.2. Biosynthesis of ArC_2-Type Shikimates.

The ArC_2-type shikimates are rare as only a handful of compounds have been isolated and proven to be derived from shikimic acid. The great majority of ArC_2 metabolites are derived from acetate via the polyketide pathway. However, the ArC_3 amino acids may loose one carbon via a decarboxylative process shown in Scheme IV.2.8. Pyridoxal phosphate (PLP) is involved and the detail of this decarboxylative pathway is described in detail in section VII.3 of the annexe. Dopamine is a neurotransmitter of synapses (permits the transfer of nerve impulses from one neuron to another). Adrenaline is a cardiac stimulant.



Scheme IV.2.8

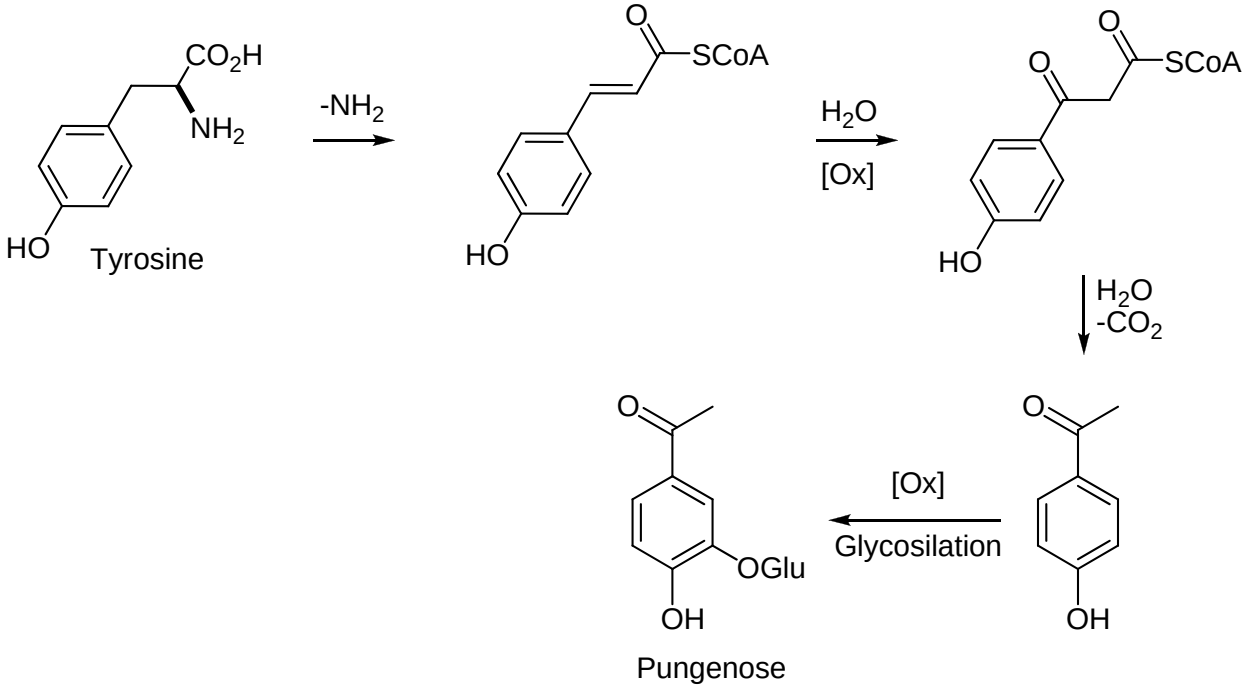
Hydroxytyrosol and homovanillic acid are metabolites of dopamine (Figure IV.2.4). Hydroxytryrosol exists in olive oil as its elenolic acid ester and is one of the most powerful anti-oxydant after gallic acid and its oxygen radical absorbance is twice that of coenzyme Q_{10} (see section IV.2.12). Levels of homovanillic acid has been used to monitor dopaminergic acitivity in the brain as a measure of stress.



Homovanillic acid Hydroxytyrosol oleuropein

Figure IV.2.4

Other ArC_2 may be formed by a decarboxylation from the β-ketoester obtained from phenylalanine, tyrosine or their derivatives. The biosynthesis of pungenoside is shown in Scheme IV.2.9. Most ArC_2 shikimates occur as their quinate esters (from esterification with quinic acid) or as glycosides (ether linkage with sugar molecules).

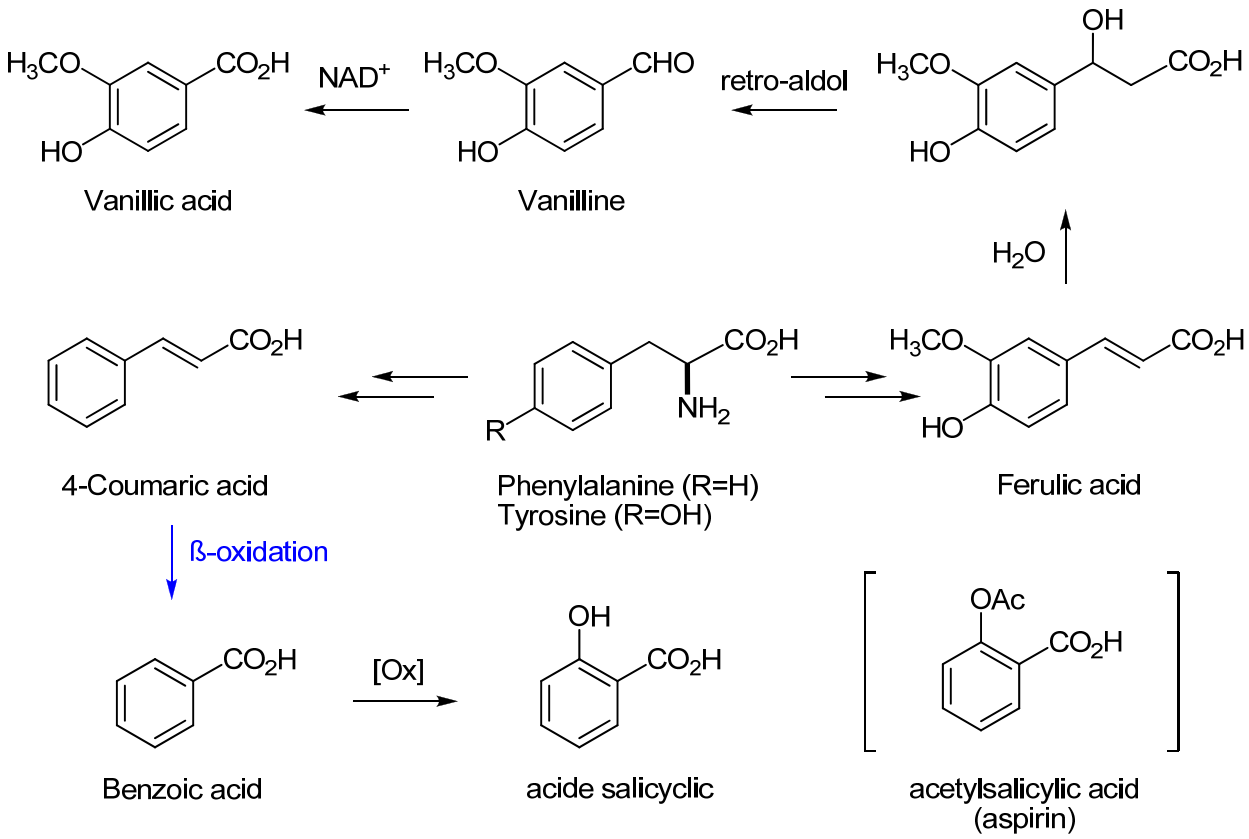


Pungenose

Scheme IV.2.9

IV.2.3. Biosynthesis of ArC_1-Type Shikimates.

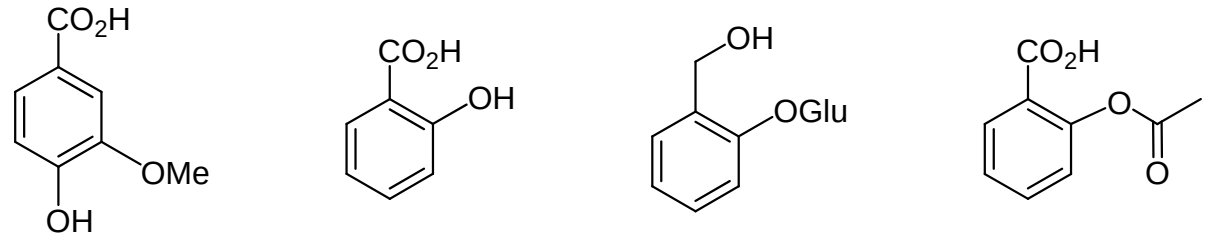
ArC_3 shikimates may suffer a β-oxidation of the chain to give ArC_1 shikimates. Oxidation of the aromatic ring can take place before or after ß-oxidation, though the former is more often the preferred pathway. Some of the ArC_1 shikimates are shown in Scheme IV.2.10. vanilline and vanillic acid are partly responsible for the taste and odor of vanillina extract while salicylic acid is a naturally occurring analgesic. Aspirin (acetylsalicylic acid) is a man-made derivative of salicylic acid prepared for its similar analgesic properties at lower dosage.



Scheme IV.2.10

Most compounds are involved in scents or flavors of plants and flowers but some were found to have certain healing powers. Extracts from willow plants were used for centuries in healing preparations, and the demonstration that salicin was at least in part responsible for those effects led to the synthesis of a more potent analogue acetylsalicylic acid sold under the brand name aspirin™ (Figure IV.2.5). In addition many shikimates are **allelopathics**, which are

chemical compounds produced by plants that 'leak' into the environment and suppress the growth or germination of other plants. These phenols are leached by rain from the foliage or shrubs which produce them, and make the underlying and surrounding soil barren. Thus salicylic acid is present in the soil beneath oak trees and suppresses undergrowth. Similarly, vanillic acid, is also a broad spectrum allelopathic compound in that it will inhibit the growth or germination of a wide range of plant species.

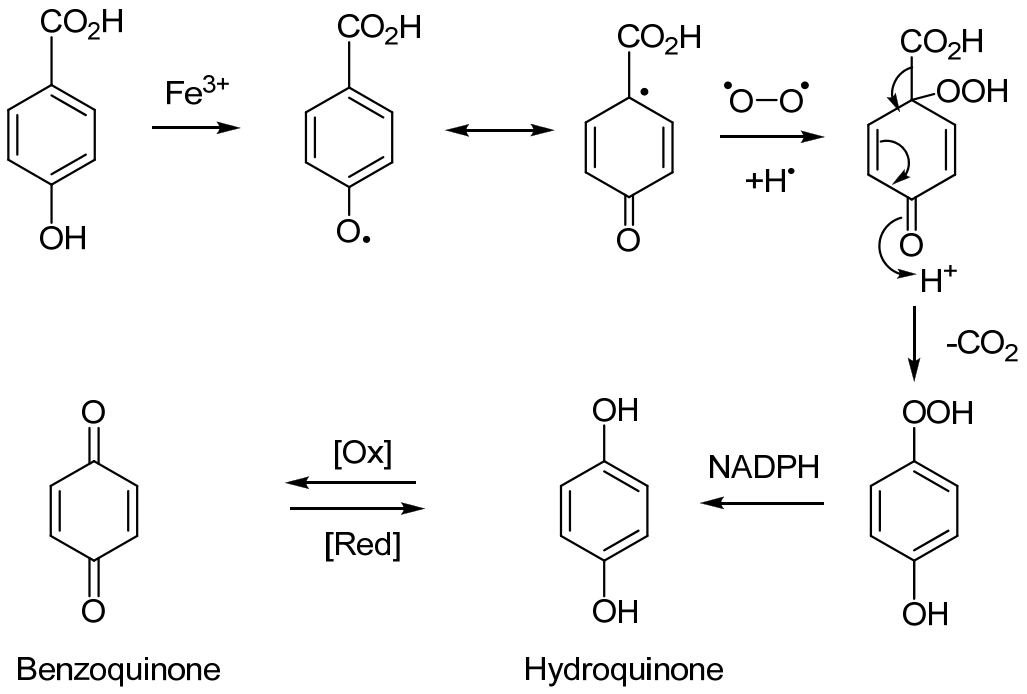


Vanillic acid (Allelopathic) | Salicylic acid (Allelopathic) | Salicin | Acetylsalicylic acid (aspirin)

Figure IV.2.5

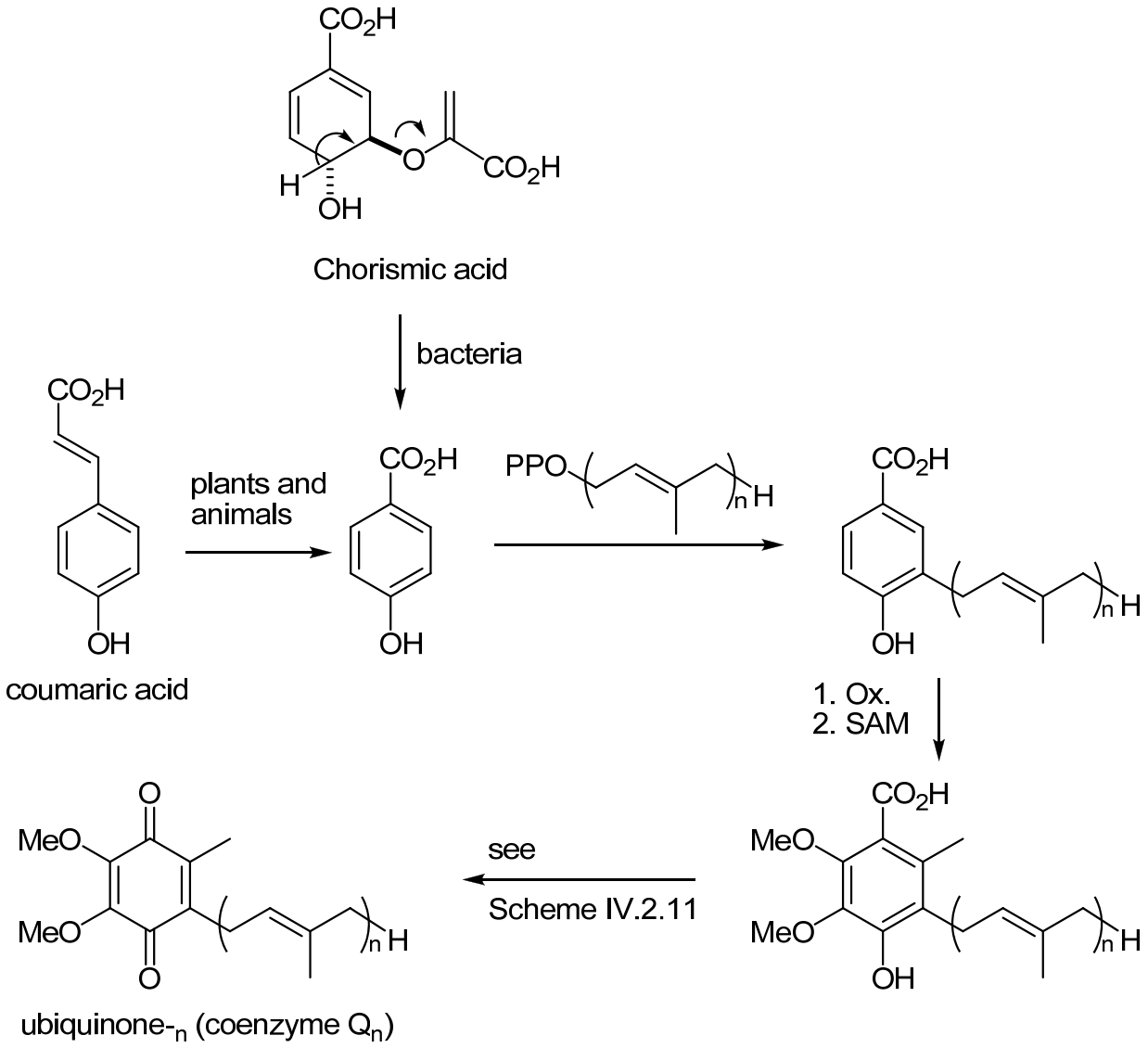
IV.2.4. ArC_0-Type Shikimates.

Quinones (ArC_0) are ubiquitous in nature where they encompass pigments, vitamins, and co-enzymes. The majority of them are formed via the polyketide pathway as shown in section III.4.2. Some, however, are shikimates or have mixed origine and are derived from shikimic acid by an oxidation of the benzoate (ArC_1) side chain (Scheme IV.2.11). Oxidative decarboxylation affords hydroquinone (after the peroxide is reduced with NADPH). Hydroquinones are easily oxidized to quinones and the latter are easily reduced to hydroquinone. This is true for enzymatic processes but it is also true in the laboratory.



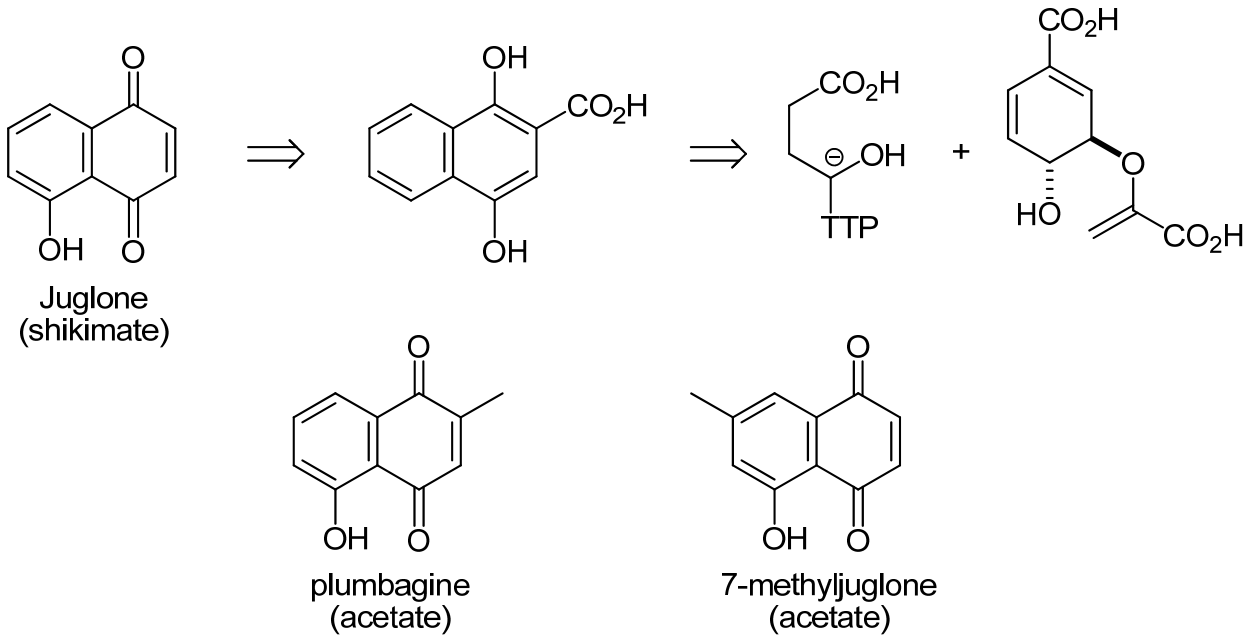
Scheme IV.2.11

To illustrate this process, the biosynthesis of coenzyme Q is shown in Scheme IV.2.12. In plants and animals, the synthesis starts from coumaric acid, which is transformed into *p*-hydroxybenzoic acid. In bacteria, the latter is made directly from chorismic acid. After alkylation with a terpenic chain of appropriate length (n = 1-12 but more often 7-12), oxidation, methylations, and oxidative decarboxylation (as per Scheme IV.2.11) furnishes the desired ubiquinone. The human redox carrier is a coenzyme Q_{10}.



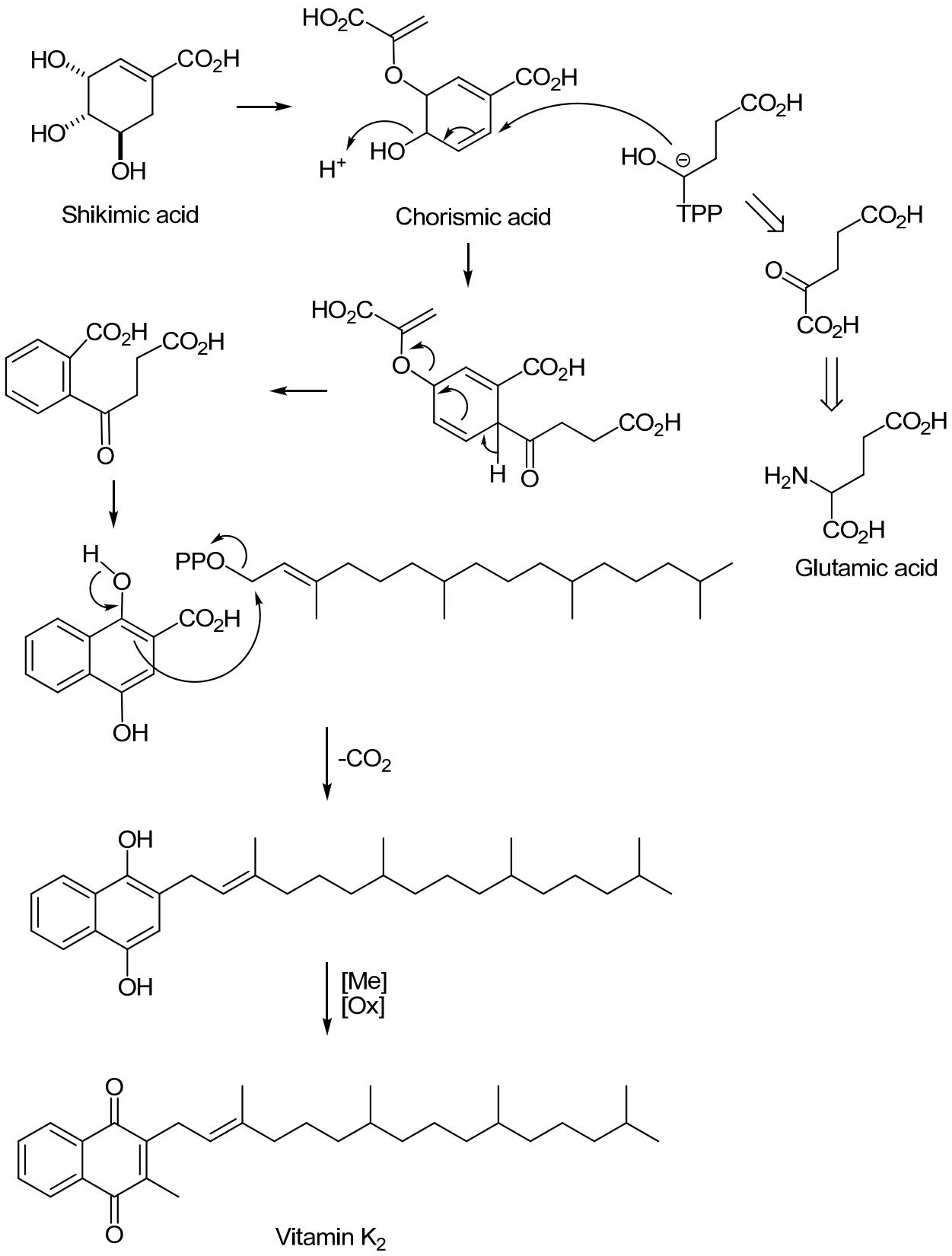
Scheme IV.2.12

One must be very careful when suggesting a biogenesis of quinine and benzoquinone natural products. For example, whereas plumbagin and 7-methyljuglone are derived from the polyketide pathway, juglone, isolated from the *Juglans regia* plant, is derived from shikimic acid via chorismic acid (Scheme IV.2.13).



Scheme IV.2.13

Ubiquinones and plastoquinones have mixed biogenesis (shikimic acid and other pathways) where the **benzoquinone** ring of these compounds comes from shikimic acid. For example, Scheme IV.2.14 gives a possible biogenesis of vitamin K where part of the benzoquinone ring is shikimate in origin. The molecule incorporates in its structure three carbons from an aliphatic amino acid and an isoprenoid chain. Their biosynthesis is therefore much more complex to elucidate.



Scheme IV.2.14

V. Alkaloids

V.1. Introduction.

V.1.1. Biological Activities of Alkaloids.

The term alkaloid, or "alkali-like" was first coined in 1819 by Meissner, an apothecary from Halle. He had noticed that a great many substances from plants were basic in nature (alkali) and contained nitrogens. The simple amines were not included in this family for some obscure reason, but for practical ones, this classification is retained today. Moreover, some alkaloids that contain the amide group are not basic at all, e.g. colchicine. Alkaloids are one of the most studied families of natural products since they have been used for thousands of years as medicinal and therapeutic agents. They occur abundantly in the plant kingdom from higher plants to microorganisms and are found only sparsely in animals. Surprisingly, the alkaloids are nearly nonexistent in marine organisms.

The alkaloids have a very potent effect in man and animals in general. The effects range from simple pain killers to hallucinogens and extremely potent poisons. Deadly poisons include strychnine (Figure V.1.1), coniine used by ancient Greeks for state executions (Socrates's death was caused by the alkaloid coniine from an injection of Hemlock extracts), and tubocurarine (curare alkaloids) used by South Americans as arrow poison. Strychnine has been used as rodenticide and vermin killer. Others were used as hallucinogens such as opium, lysergic acid, mescaline found in the Mexican Peyote cactus, and psilocybin from the Mexican "magic" mushroom used by the upper priesthood of the Mayas to gain spiritual contact with their gods. Several alkaloids are used as valuable drugs in medicine e.g. morphine as pain reliever, reserpine in psychiatry as tranquilizer, curare alkaloids in anaesthesia, atropine in eye surgery, ergonovine to induce labor and quinine as an antimalarial compound. Table V.1.1 lists some chronological use of drugs in ancient and modern medicine that are thought to be alkaloids or derived from them. Alkaloids are among the most potent medicinal drugs on Earth and in general plant metabolites were used thousands of years ago in diverse concoctions to heal, induce spiritual state of mind (narcotic), relieve pain, or for many other purposes. Table V.1.2 will convince you of the medicinal importance of plant metabolites in today's medicine.

Before all the properties attributable to alkaloids were fully established, especially the addictive properties of certain narcotic alkaloids, it was not unusual to find them advertised and sold as common remedies. For example, heroin was marketed as a sedative for coughs by the Bayer company along side such medicine as aspirin. Opium, the dried, resinous juice of the

unripe poppy seeds, was used in many of the patent medicines of the nineteenth century: Ayer's Cherry pectoral, Jayne's expectorant, Pierce's Golden medical discovery, and Mrs Winlows' soothing syrup. Figure I.1.2 shows such advertisements of the past.

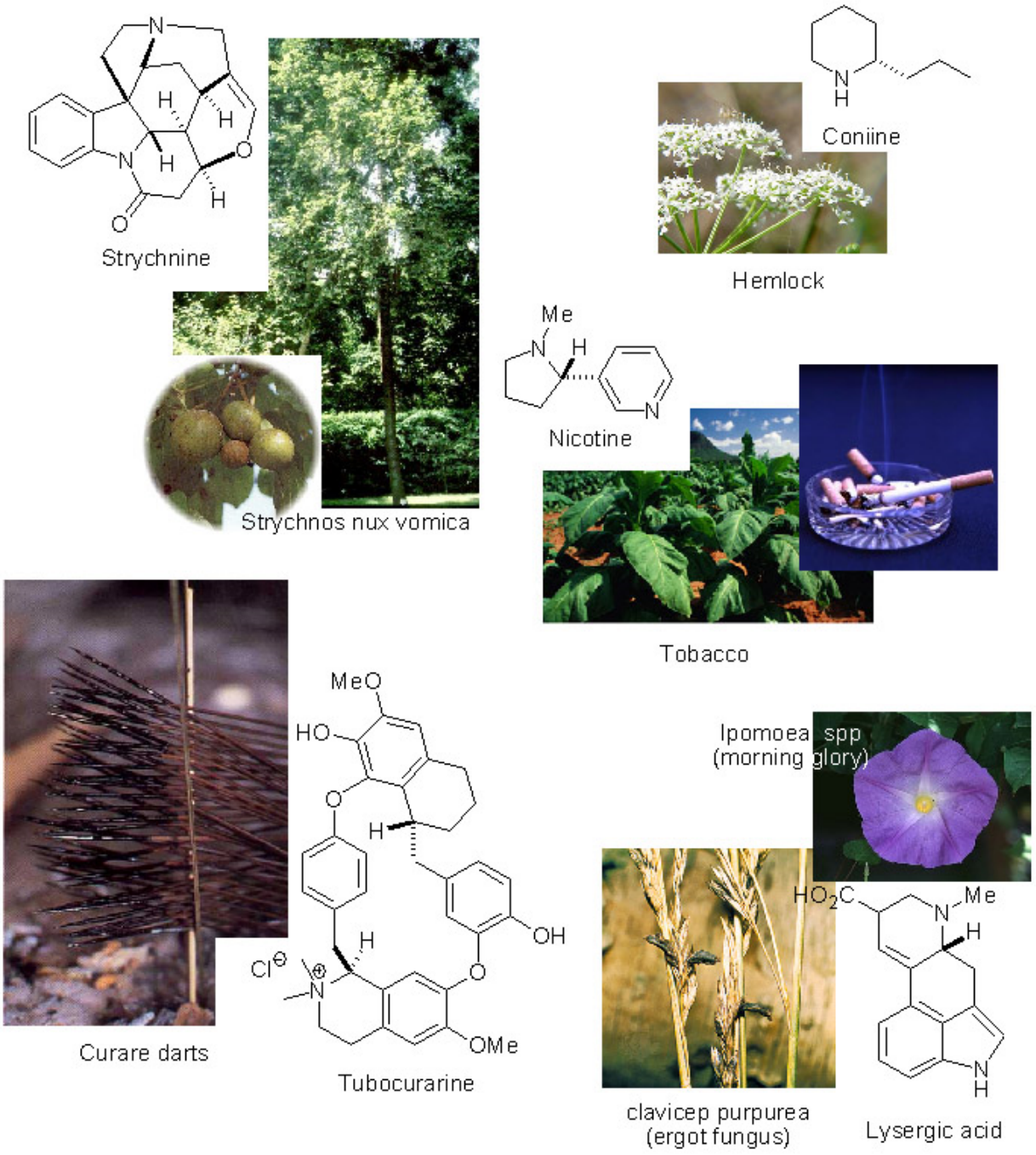
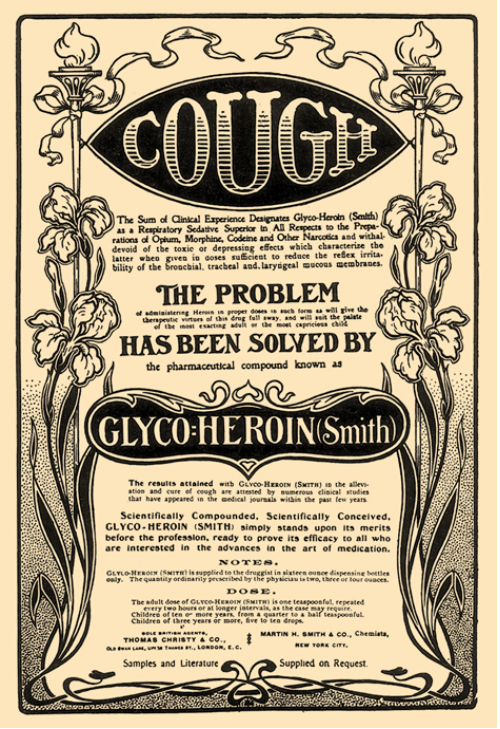


Figure V.1.1


Mrs Winslows
Soothing
Syrup.
For Children Teething.
Hagley Museum and Library


COUGH
The Sum of Clinical Experience Designates Glyco-Heroin (Smith) as a Respiratory Sedative Superior in All Respects to the Preparations of Opium, Morphine, Codeine and Other Narcotics and withal devoid of the toxic or depressing effects which characterize the latter when given in doses sufficient to reduce the reflex irritability of the bronchial, tracheal and laryngeal mucous membranes.
THE PROBLEM
of administering Heroin in proper doses in such form as will give the therapeutic virtues of this drug full sway, and will suit the palate of the most exacting adult or the most capricious child
HAS BEEN SOLVED BY
the pharmaceutical compound known as
GLYCO-HEROIN (Smith)
The results attained with Glyco-Heroin (Smith) in the alleviation and cure of cough are attested by numerous clinical studies that have appeared in the medical journals within the past few years.
Scientifically Compounded, Scientifically Conceived, GLYCO-HEROIN (SMITH) simply stands upon its merits before the profession, ready to prove its efficacy to all who are interested in the advances in the art of medication.
NOTES.
Glyco-Heroin (Smith) is supplied to the druggist in sixteen ounce dispensing bottles only. The quantity ordinarily prescribed by the physician is two, three or four ounces.
DOSE.
The adult dose of Glyco-Heroin (Smith) is one teaspoonful, repeated every two hours or at longer intervals, as the case may require. Children of ten or more years, from a quarter to a half teaspoonful. Children of three years or more, five to ten drops.
THOMAS CHRISTY & CO.,
MARTIN H. SMITH & CO., Chemists, NEW YORK CITY.
Samples and Literature Supplied on Request.


FARBENFABRIKEN
FRIEDR. BAYER & Co.
ELBERFELD
Heroin

Figure V.1.2

Table V.1.1 Native medicinals as drug sources.

2500 BC	Emperor Shen Nung anti-febrile effects of ch'ang shan.	Anti-malarial alkaloids.
ca 600 BC	Homer's "Odyssey" refers to wholesome and toxic herbs.	Hallucinogens.
900's	Rhazes - *Papaver somniferum* ; opium pills (coughs).	Morphine (1805).
1000's	Avrienna - extracts of wild autumn crocus (gout).	Morphine (1805).
1569	Chinese dispensary describes Ma Huang (*Ephedra* spp) as bronchodilator.	Ephedrine.
1600's	Jesuit missionaries - extracts of cinchona bark (chills, fever treatment).	Quinine (1920).
1758	Withering - extracts of fox-glove (dropsy).	Digitalis (1874).
1876	*Rauwolfia serpentina* extracts used (India) for snakebites, anxiety, tension.	Reserpine (1932).
	Vinca rosea extracts used as oral hypoglycemic. Now as anticancer agent.	VCR, VLB (1960).
	Infusions, concoctions of Datura plants by central & south American Indians (Shamanistic, religious rituals; narcotics).	Atropine (1927). Cocaine (1885) Scopolamine (1921)

Table V.1.2 **Distribution of Prescriptions by Therapeutic Category**

Therapeutic Category	*% of total prescriptions (1.05 billions)*	*% of products containing plant constituents*
Antibiotics	12.31	–
Hormones	11.46	86.76
Progesterone	4.75	100
Corticoids	4.71	100
Estrogens	1.60	8.88
Androgens	0.27	79.67
Anabolics	0.09	98.66
Others	0.05	87.34
Ataraxics	7.93	–
Cardiovascular Agents	7.23	73.24
Hypotensive Agents	3.27	89.59
Coronary Vasodilators	1.70	8.05
Miscellaneous	2.26	85.73
Antitussives–Decongestants	7.21	57.79
Analgesics	7.06	39.94
Sedatives–Hypnotics	5.28	0.37
Vitamins, Geriatric	0.30	77.47
Vitamins, Other	4.53	–
Diuretics	3.66	–
Antispasmodics	3.30	44.83
Antiobesity	3.08	–
Antihistamines	2.51	–
Sulfonamides	2.29	–
Thyroid Therapy	1.98	0.12
Antiarthritics (nonsteroidal)	1.96	7.91
Diabetic Therapy	1.93	–
Hematinics	1.64	1.82
Bronchodilators	1.61	44.68
Psychostimulants	1.41	–
Muscle Relaxants (Oral)	1.32	4.40
Antidiarrheals	1.18	96.56
Antinauseants	1.13	2.74
Dermatologicals	0.08	3.72
Antacids	0.71	4.56
Fungicides	0.70	0.02
Enzymes	0.65	41.10
Laxatives	0.62	44.24
Miotics	0.53	76.31
Hemorrhoidals	0.19	53.67
Hemostatics	0.19	6.23
Bile Therapy	0.14	29.92

Table V.1.2 (contd)

Local Anesthetics	0.13	8.71
Mydriatics	0.07	89.13
Hypcholesteremics	0.06	13.95
Oxytocics	0.06	3.11
Antihemics	0.06	0.15
Antimalarials	0.05	80.84
Miscellaneous	4.58	1.20

The wide range of biological effects of the alkaloids on animals probably stems from their different mode of action on the nervous system. Basically the nervous system is divided in two parts, the central nervous system (CNS) and the autonomous nervous system (ANS). The systems are not totally independent. The CNS is primarily concerned with integrating information from both internal and external sources, and it also processes appropriate responses to muscles and glands. The ANS on the other hand is primarily a motor system and is concerned with visceral functions and balance. The ANS, whose control center is the hypothalamus, is classically divided into two separate systems, the sympathetic and parasympathetic. The effects of these on a given function or organ are usually opposite. The nervous systems communicate with each other and the peripheral organs and muscles through specialized cells called neurons. Each neuron is linked to another neuron via synaptic gaps in which specialized chemicals flow from the neural receptors at one end to the neural receptors at the other. These chemicals are called neurotransmitters. The four known neurotransmitters are acetylcholine, noradrenaline, dopamine, and serotonin. Therefore, although the nerve impulse is transmitted electrically inside the neuron, it is transmitted chemically in between neurons.

A drug may mimic, facilitate, or antagonize a normal process. It may also increase, decrease, or intermittently disrupt the normal activity of the cell. Many of them will block, stimulate, or mimic neurotransmitters. For example, the curare alkaloids will occupy the acetylcholine receptor and therefore prevent transmission altogether. Acetylcholine transmits between nerve cells and skeletal muscles. Thus curare alkaloids will cause involuntary muscle contraction, convulsions, paralysis and ultimately death. The atropine alkaloids are capable of occupying the acetylcholine receptor at the postsynaptic stage and thus may act only on the peripheral nervous system which translates to a local anaesthetic. For example, atropine and cocaine are useful local anaesthetics used in surgery. Nicotine on the other hand, has a direct stimulating effect on the cholinergic effector cells (cells responsible for the synthesis of acetylcholine) and thus will increase gland and skeletal muscles and heart activity controlled by

this system. Many other alkaloids have a direct or indirect action on the nervous system and thus they produce a wide spectrum of effects ranging from stimulants and hallucinogens to venoms and poisons.

V.2. Aliphatic amino acids.

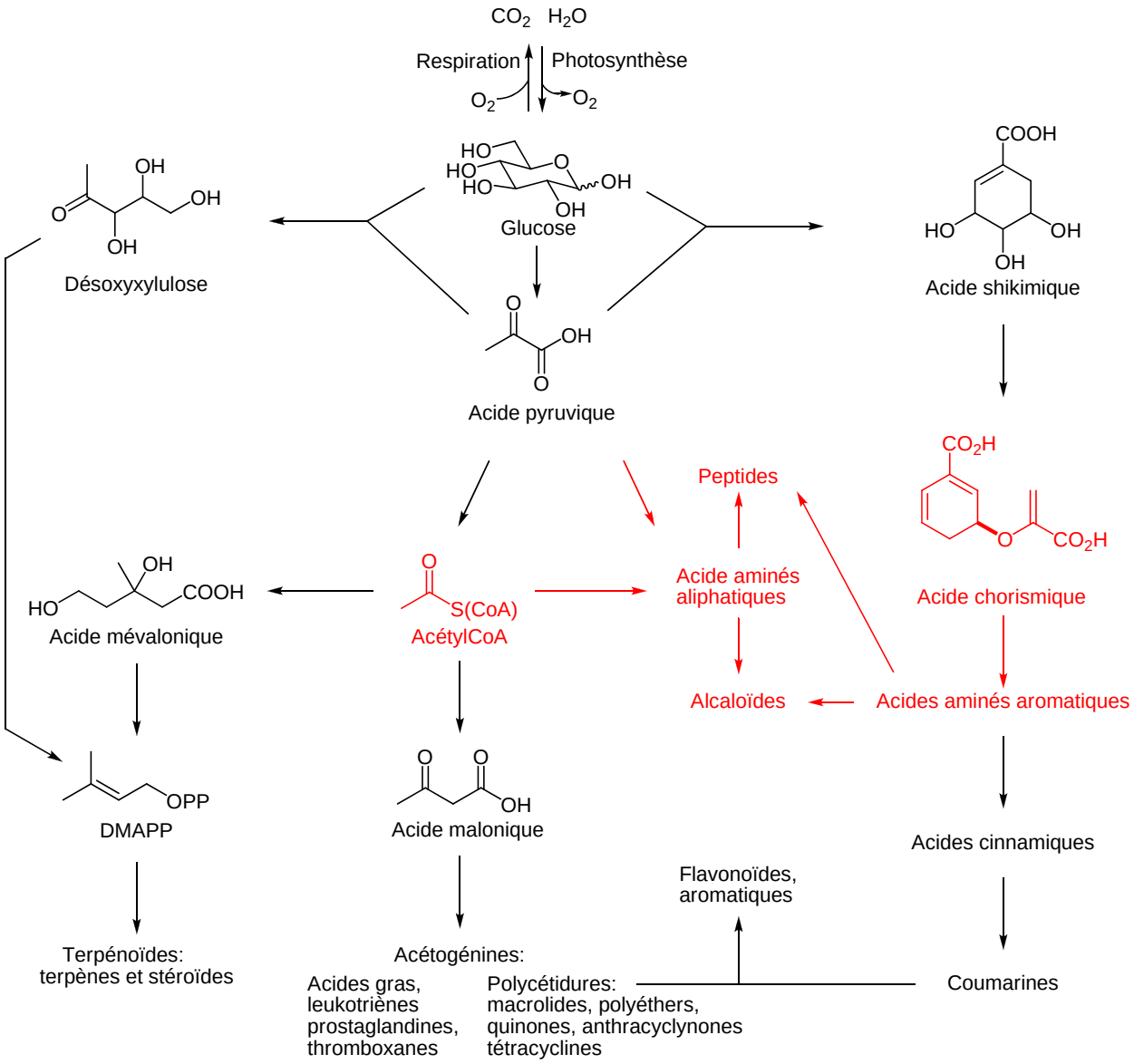


Figure V.2.1

You will remember Figure V.2.1, the same as from chapter 2 that summarizes the synthesis of secondary metabolites. The aromatic amino acids are derived from shikimic acid as seen in chapter 4. Aliphatic amino acids, many of which are biosynthetic precursors to alkaloids, are

derived from pyruvic acid or from acetate (in red). Figure V.2.2 summarizes the synthesis of amino acids from both. As can be seen, the Krebs cycle (respiration) is essential in producing oxaloacetic acid and 2-oxoglutaric acid. We will not see the details of the Krebs cycle but it can be found in any good biochemistry textbook. But we will study the details of the synthesis of some amino acids that are relevant to this chapter, i.e. alkaloid biosynthesis.

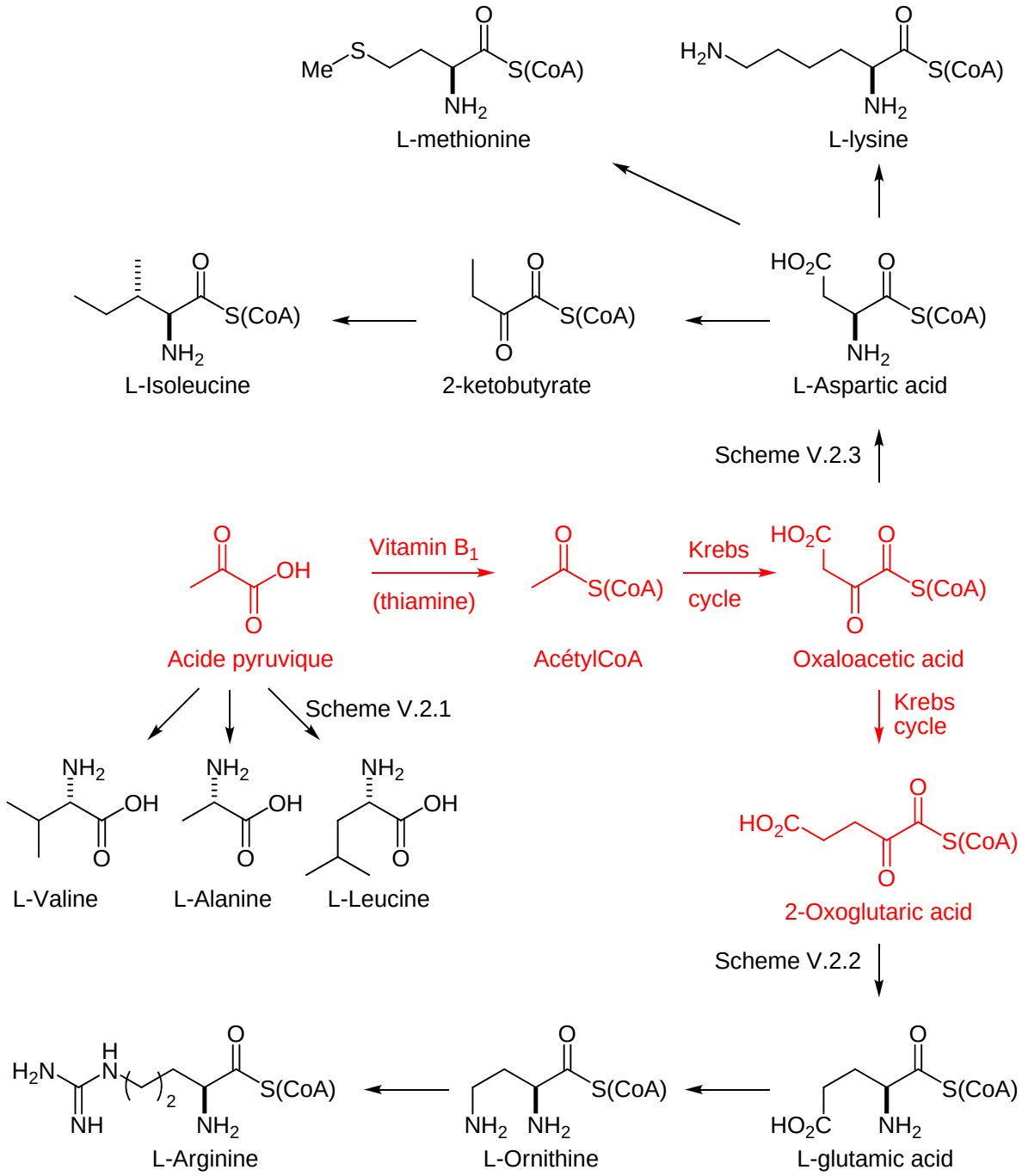
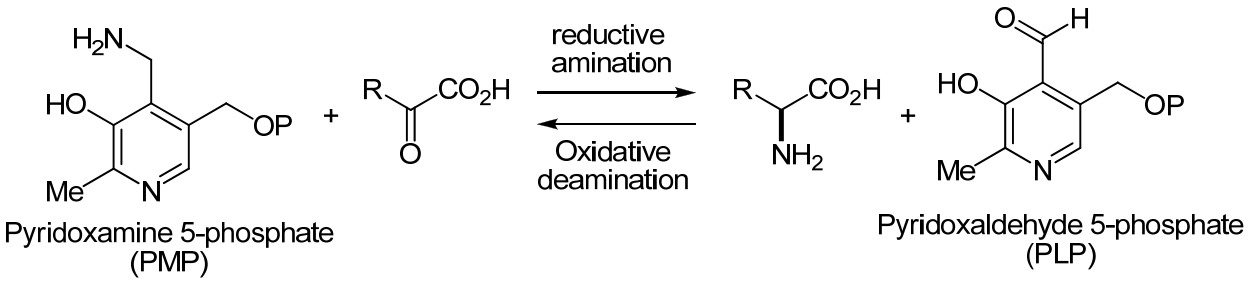


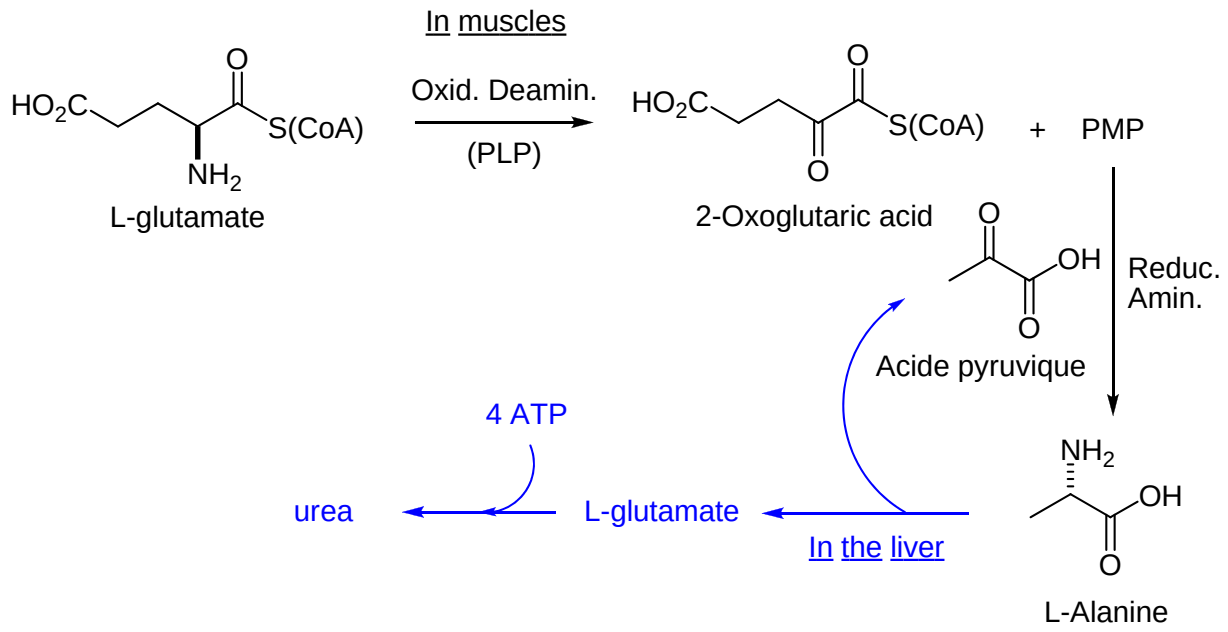
Figure V.2.2

The main transformation responsible for the transfert of amino groups (NH_2) are the reductive amination and oxidative deamination, which transforms ketones into amines or amines into ketones, respectively. This is illustrated in Scheme V.2.1 but the details are revealed in section **VII.3 of the annex**.



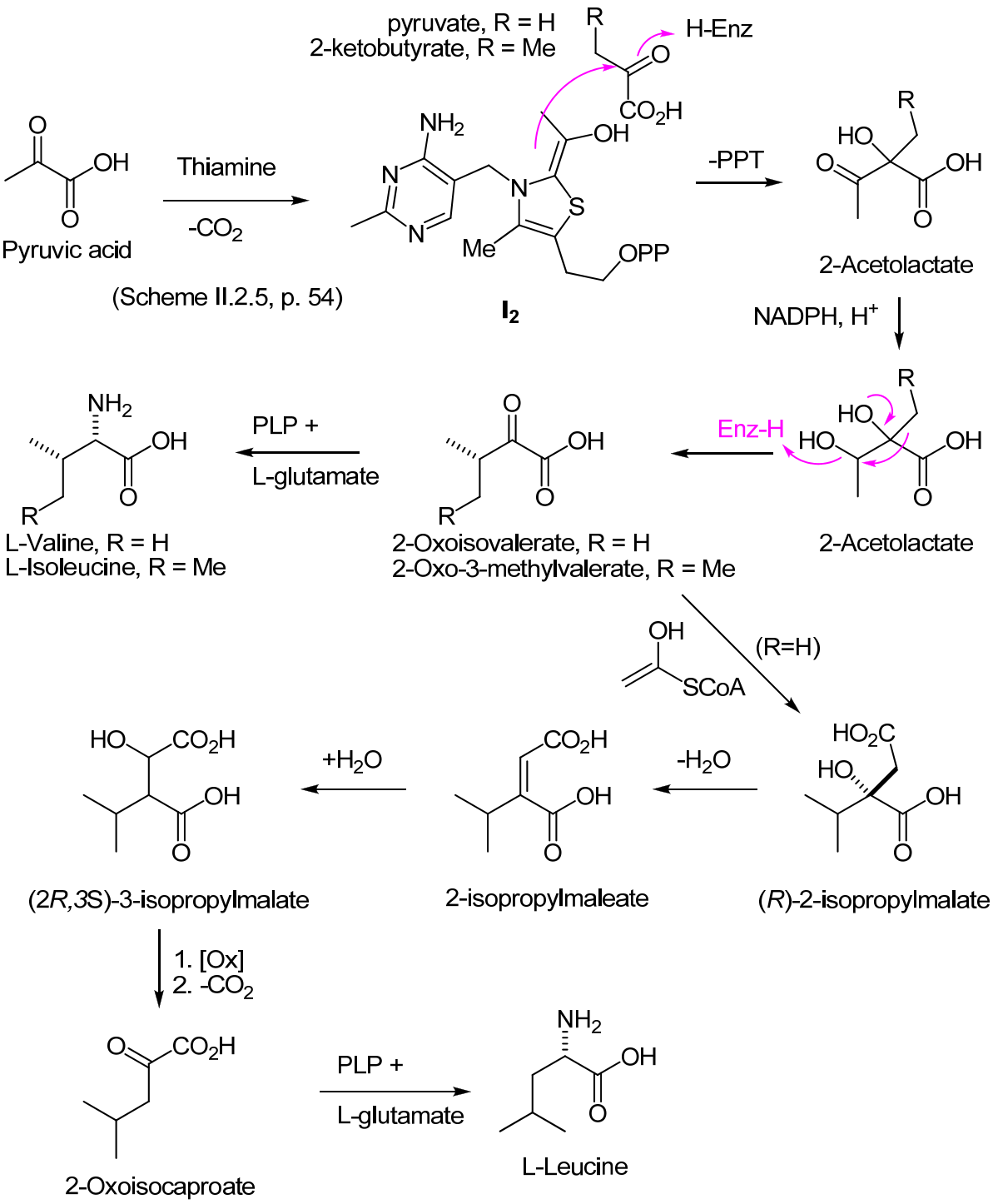
Scheme V.2.1

For example, the biosynthesis of alanine, in muscle tissues, involves the reductive amination of pyruvic acid by L-glutamate, which is itself oxidatively deaminated and produces PMP in the process (Scheme V.2.2). It is that PMP that is involved in the reductive deamination of pyruvate into L-alanine. In the liver, this process is reversed where glutamate is transformed by the urea cycle into urea, which is then excreted, thus maintaining the nitrogen balance in the body.



Scheme V.2.2

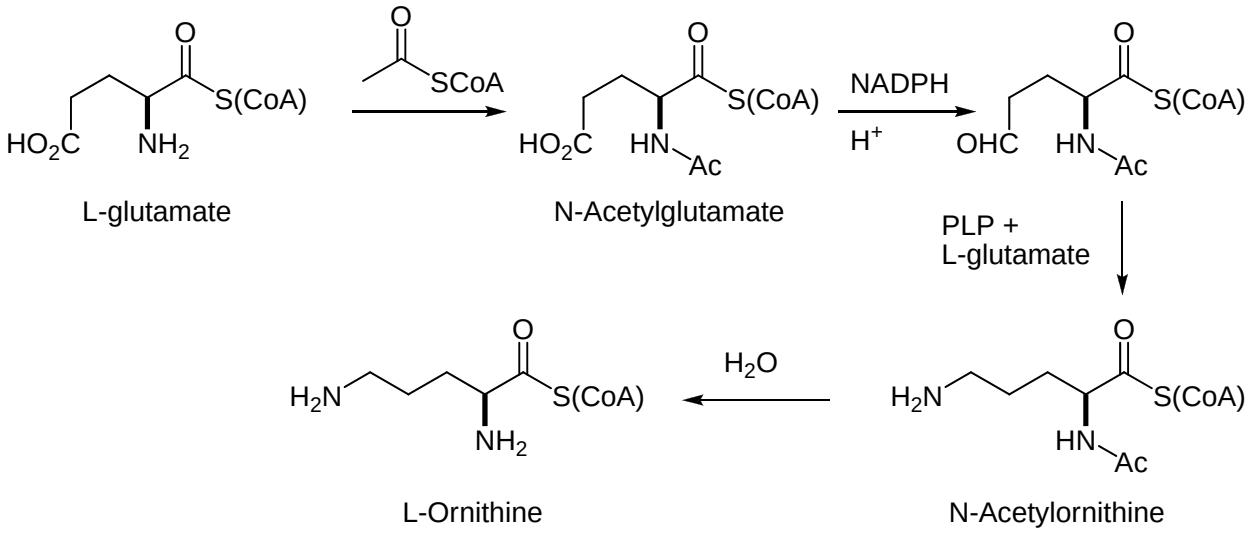
L-Valine's biogenesis starts from two molecules of pyruvic acid, one of which is transformed into a nucleophile (**I_2**) via decarboxylation, in the same way pyruvic acid is decarboxylated to acetate as seen in chapter 2 (see Scheme II.2.5 on p. 54). Instead of picking up a lipoate coenzyme, the intermediate **I_2** attacks another molecule of pyruvate to make 2-acetolactate (Scheme V.2.3). Reduction and methyl migration gives 2-oxoisovalerate, which is then reductively aminated with L-glutamate (or another amino acid) as a source of amino group. If **I_2** attacks 2-ketobutyrate instead of pyruvate, the same sequence will lead to L-isoleucine. We don't know everything about the biosynthesis of L-leucine, but we know it is made from 2-oxoisovalerate by condensation with acetyl co-enzyme A (via an aldol reaction). The resulting 2-isopropylmalate is dehydrated and rehydrated by the 3-isopropylmalate isomerase to the product of the same name. Normal oxidation (NADPH) and decarboxylation give 2-oxoisocaproate, which is converted to L-Leucine.



Scheme V.2.3

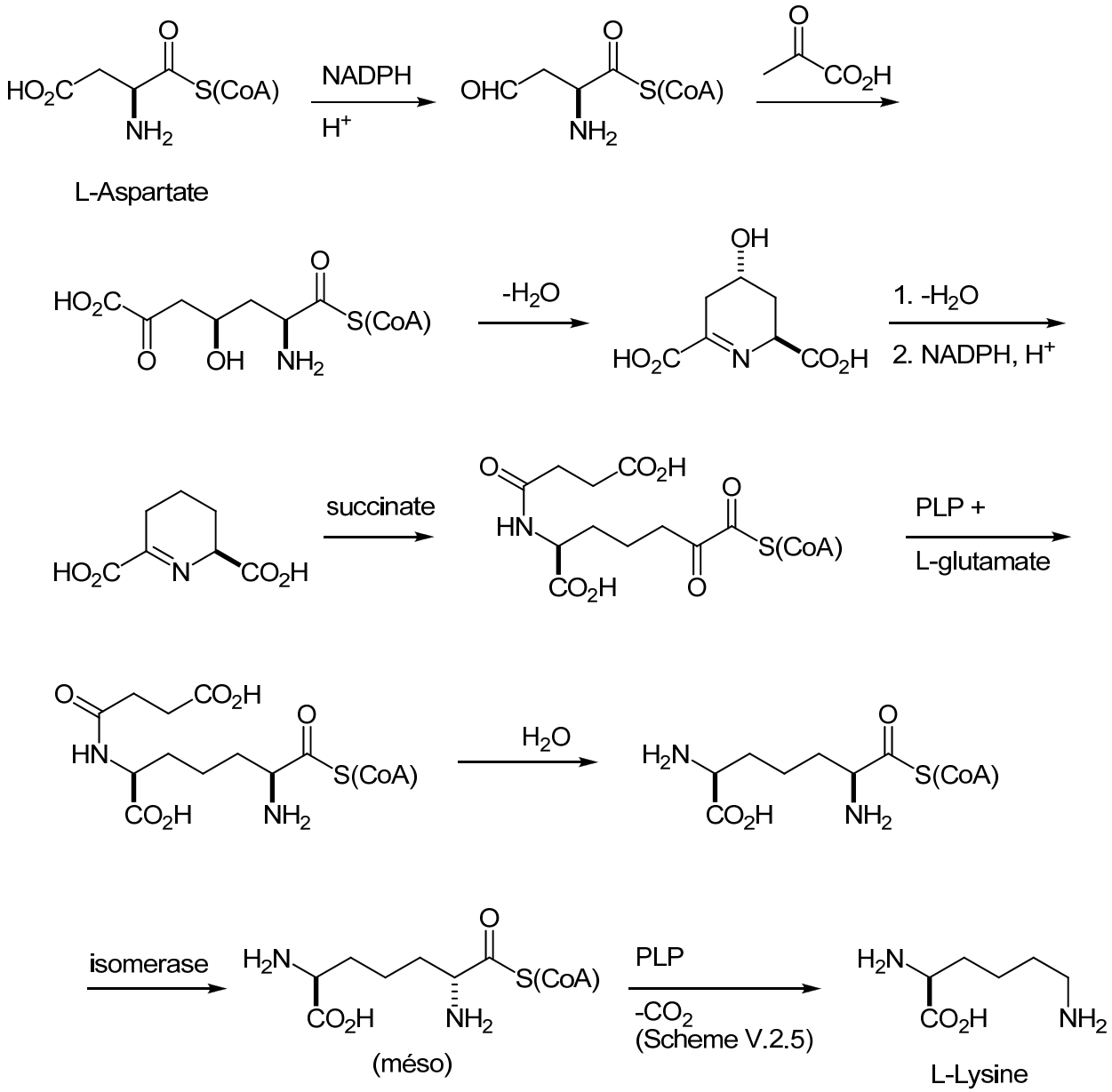
The biosynthesis of L-ornithine starts from L-glutamate. Acetylation of the amino group of glutamate followed by reduction (NADPH via the phosphate ester) and reductive amination,

where glutamate serves also as the source of the amino group, produces N-acetalyornithine (Scheme V.2.4). A simple hydrolysis affords L-ornithine.



Scheme V.2.4

The biogenesis of L-lysine is more complicated and proceeds via the amino acid L-aspartate (Scheme V.2.5). The latter is formed from 2-oxaloacetic acid by reductive amination. Reduction of L-aspartate (NADPH via the phosphate ester) produces the aldehyde, which participate in an aldol reaction with pyruvic acid. Cyclisation, dehydration and reduction of the double bond (NADPH, H^+) gives a cyclic intermediate that is opened with succinate. This reaction is akin the acetylation of L-glutamate in the biosynthesis of ornithine, except that a larger ester (succinate) is used here. Reductive amination, removal of the succinate is followed by an isomerization of the stereocenter bearing the amino group such that a meso (achiral) symmetrical bis-aminoacid is produced. It is the enantioselective decarboxylation that produces L-lysine. The latter is a common reaction in alkaloid synthesis and it is described in Scheme V.3.1.



Scheme V.2.5

V.3. Alkaloids Derived from Ornithine and Lysine.

V.3.1. General Biosynthesis of Alkaloids.

Although alkaloids come in an incredible variety they can be traced back to relatively few biogenetic sources. All alkaloids can be biosynthesized with building blocks from the shikimic

acid, mevalonic acid, and acetate pathways in **combination with an amino acid**. In fact only a few amino acids are involved in alkaloid biosynthesis. These are lysine, ornithine, glycine, glutamic acid, phenylalanine, tyrosine, tryptophan, and anthranillic acid (Figure V.3.1). The alkaloids can therefore be classified according to which amino acids they are derived from. Since each starting amino acid will give rise to an alkaloid with a specific nitrogen-heterocycle (N-heterocycle), the alkaloids have been classified by the skeleton of the main N-heterocycle. A few non amino-acid compounds are at the origin of a number of steroidal alkaloids in which ammonia (NH_3) is the source of nitrogen.

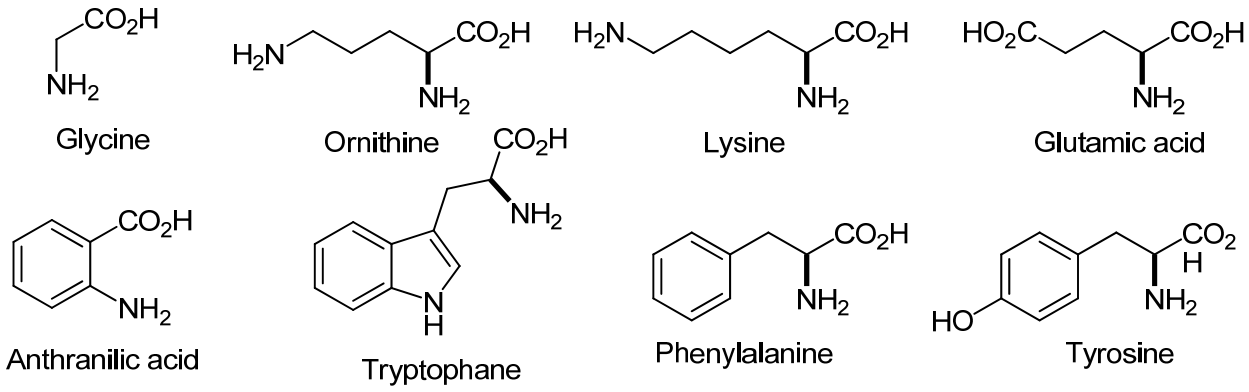
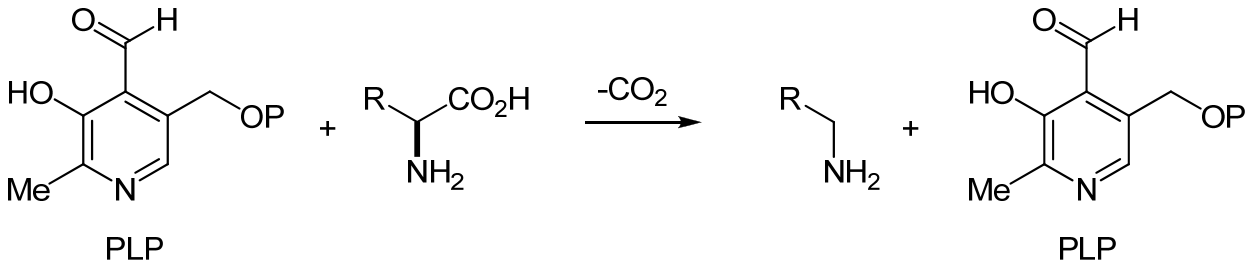


Figure V.3.1

As in the case of the other biogenetic pathways, a number of simple transformations contribute to alkaloid (and other) biosynthesis. We have seen reductive amination and oxidative deamination (Scheme V.2.1 and section VII.3 of the annex). Decarboxylation is also a common transformation that takes an amino acid and converts it to the corresponding amine (Scheme V.3.1). The details of this transformation is given in section VII.3 of the annex.



Scheme V.3.1

Before we start investigating the biosynthesis of alkaloids it is useful to review the names of several heterocyclic and heteroaromatic compounds containing nitrogen only. Figure V.3.2 gives a survey of the most encountered heterocyclic ring systems of alkaloids. Often the classification of alkaloids will follow the type of ring contained, e.g. the pyrrolidine alkaloids, the piperidine alkaloids, the indole alkaloids, etc. In the non-aromatic compounds, the sp^3 hybridized amines can undergo rapid inversion of configuration. Note that the non-bonded electron pair of pyridine are not involved in the aromaticity of the ring. In the case of pyrrole the nitrogen adopts an sp^2 hybridization to allow aromaticity by involving its lone pair of electrons to participate (Figure V.3.2). The chemical and physical properties of these compounds are dealt with in section V.4.

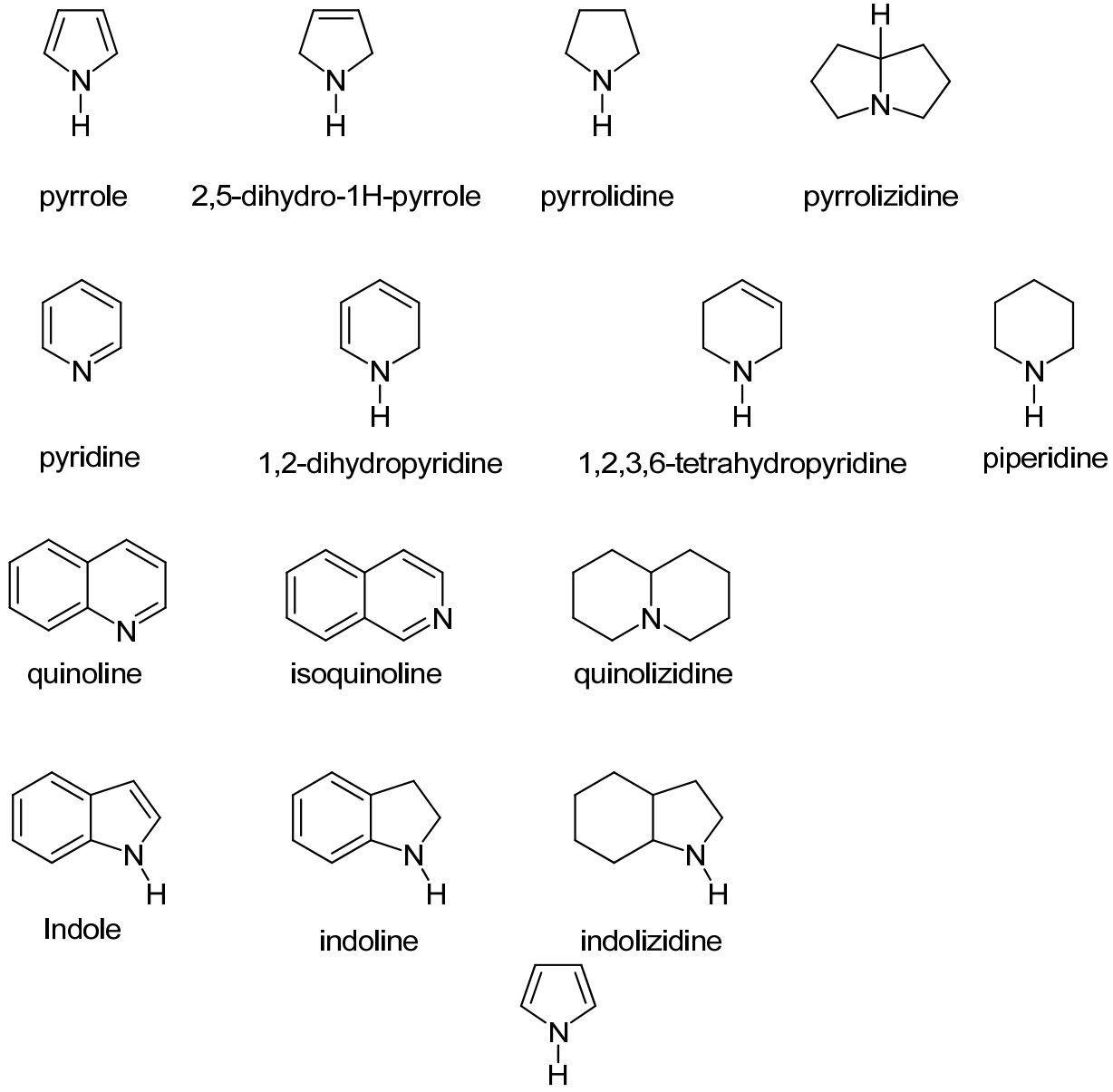
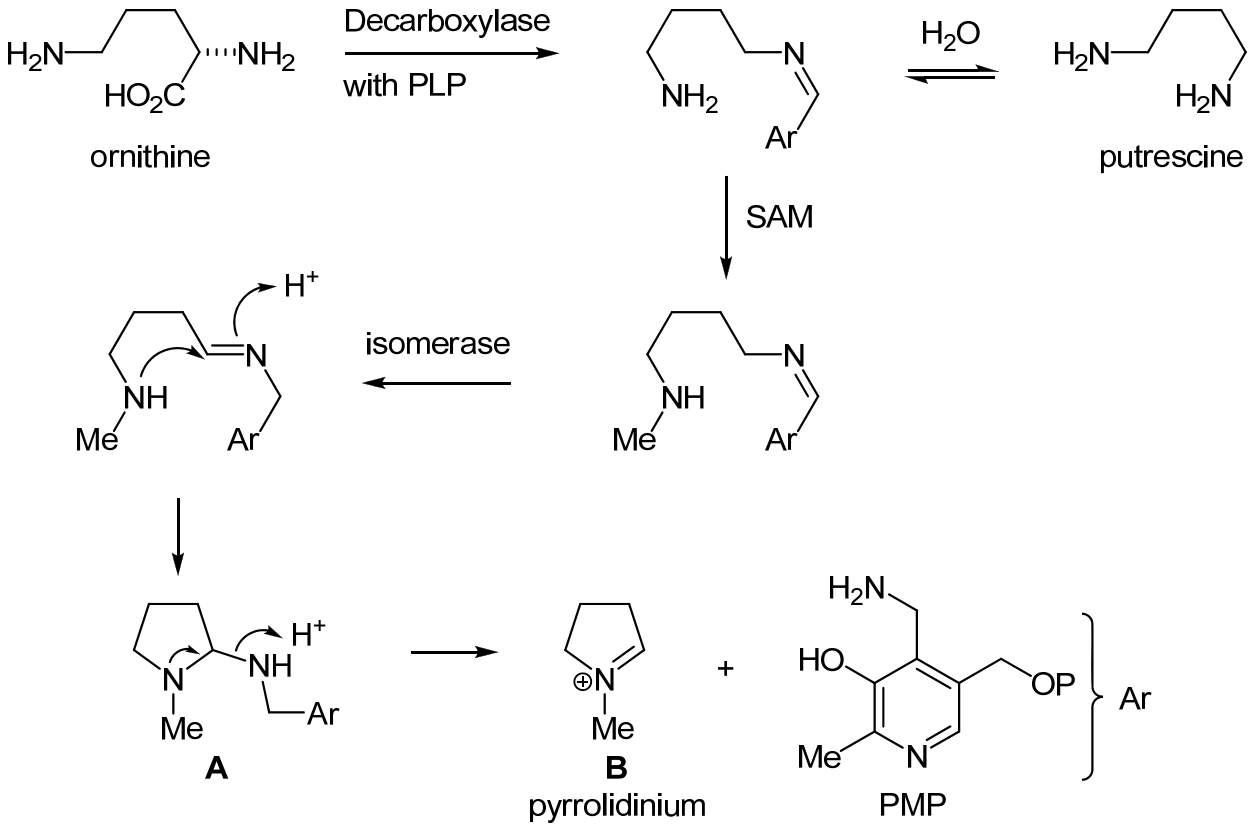


Figure V.3.2

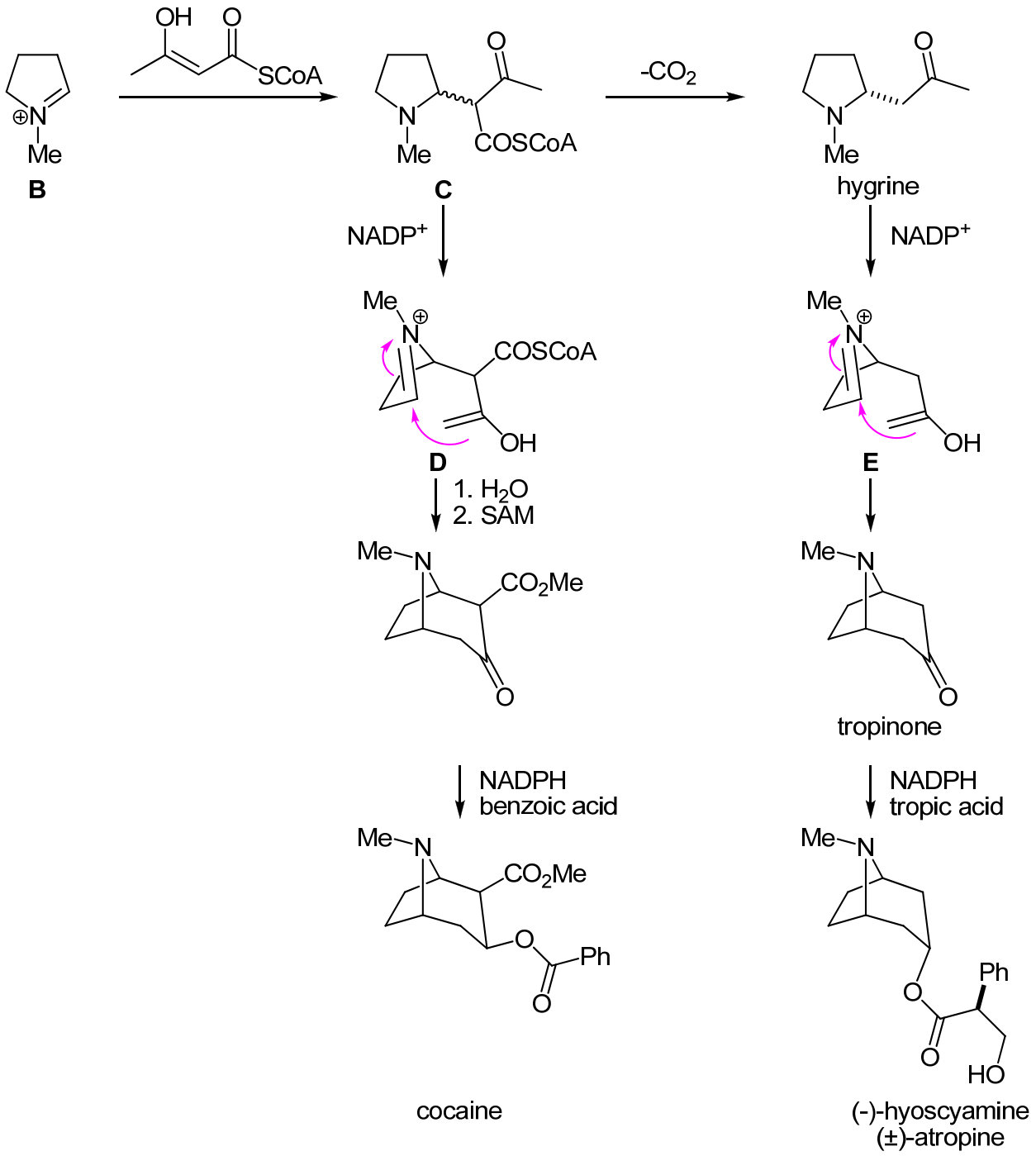
V.3.2. The Pyrrolidine Alkaloids (derived from ornithine).

Ornithine is the direct precursor of the pyrrolidine and pyrrolizidine alkaloids including the compounds containing these rings as substructures. Ornithine, in the first stages of biosynthesis of alkaloids, gets decarboxylated by the mechanism shown in Scheme V.3.1. The resulting intermediate can be hydrolyzed to the diamine putrescine which is in fact a natural product in itself (Scheme V.3.3). It is also incorporated in some alkaloids which suggested it could be an intermediate in their biosynthesis. However, studies with labelled nitrogen at position 2 or 5 of ornithine have shown that the labelled nitrogen is **not** randomly distributed in the alkaloids therefore suggesting the non-intermediacy of symmetrical putrescine. Instead, the intermediate after decarboxylation can suffer a methylation and cyclization to give the pyrrolidine ring **A**. The two nitrogens were never equivalent and are thus differentiated. This pyrrolidine ring eliminates pyridoxamine to give a cyclic ammonium salt **B** which is exceedingly electrophilic and can undergo attack by nucleophiles.



Scheme V.3.3

Such an attack occurs with acetoacetyl coenzyme A (acetate derived) from either face of the molecule (Scheme V.3.4). 2(S)-Alkaloids are obtained if the attack proceeds from the ß-face or 2(R)-alkaloids if it proceeds from the α-face. The intermediate **C** formed after the attack of acetoacetyl CoA can then suffer decarboxylation to give the natural product hygrine (isolated from coca leaves). Oxidation (removal of H^-) of intermediate **C** gives a new iminium salt **D**, while oxidation of hygrine gives iminium salt **E**. Cyclization of the methyl ketone is triggered enzymatically to give tropinone having the tropane-type skeleton. Tropane alkaloids, including cocaine, are ubiquitous in nature. Further reduction of the ketone and subsequent esterification of the resulting alcohol with tropic acid occurs leading to atropine. (-)-hyoscyamine is an insect repellant that plants use against harmful insects. Its synthetic racemic form called atropine is used in eye surgery as a local anaesthetic.

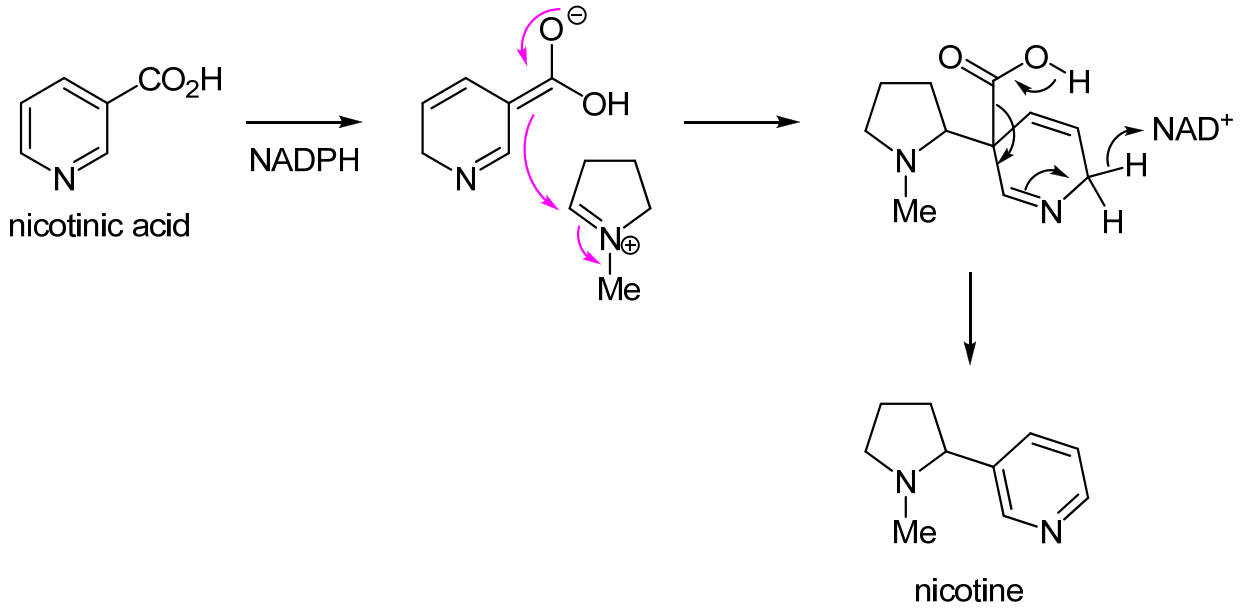


Scheme V.3.4

Oxidation to give the pyrrolidinium salt can occur before decarboxylation in compound **B** leading to pyrrolidinium salt **D**. After cyclization at the terminal methyl, the ketone is reduced and esterified with benzoic acid to give cocaine. Of the tropane alkaloids, cocaine is probably the

most interesting from a pharmacological point of view. It derives from the Coca plant, a native of the Andes, and the leaves of this plant are dried, and chewed by the local Indians. An estimated 8 million South American Indians consume large quantities of cocaine daily and they are said to derive a feeling of well-being, and alleviation of hunger pangs, with no concomitant hallucinogenic effects. However, intravenous cocaine is strongly addictive, and this compound is treated as a narcotic, though it is a useful local anaesthetic, used primarily in minor ear, nose, and throat operations.

Nucleophilic attack of the pyrrolidinium salt **A** can occur by the enolate of reduced nicotinic acid (Scheme V.3.5). A decarboxylation-oxidation sequence leads to nicotine, the addictive stimulant found in tobacco leaves.

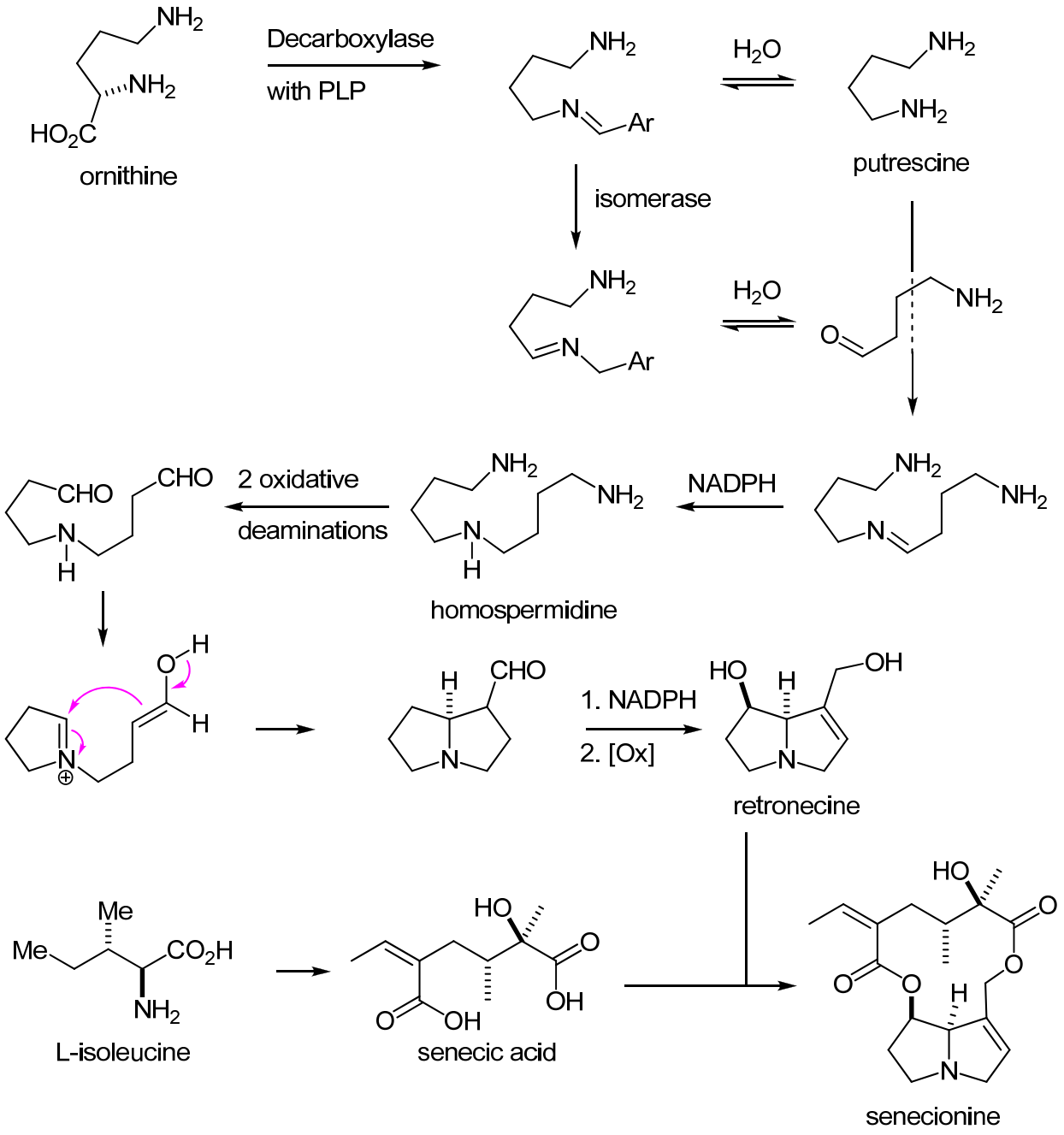


Scheme V.3.5

V.3.3. The Pyrrolizidine Alkaloids.

Two molecules of ornithine are needed to construct the pyrrolizidine alkaloids. There has not been a great deal of studies done on this class of alkaloids, except that it is known that laburnine is a metabolite of these bicyclic alkaloids and causes poisoning among cattle that eat tansy ragwort and other weeds. One molecule of ornithine is converted to putrescine as illustrated in Scheme V.3.3. The other molecule is transformed to an amino-aldehyde which can undergo condensation with putrescine to yield an imine (Scheme V.3.6). This imine is reduced with NADPH and the amino groups are oxidized and hydrolyzed to the dialdehyde. This dialdehyde suffers internal cyclization to form a cyclic imine which can in turn undergo

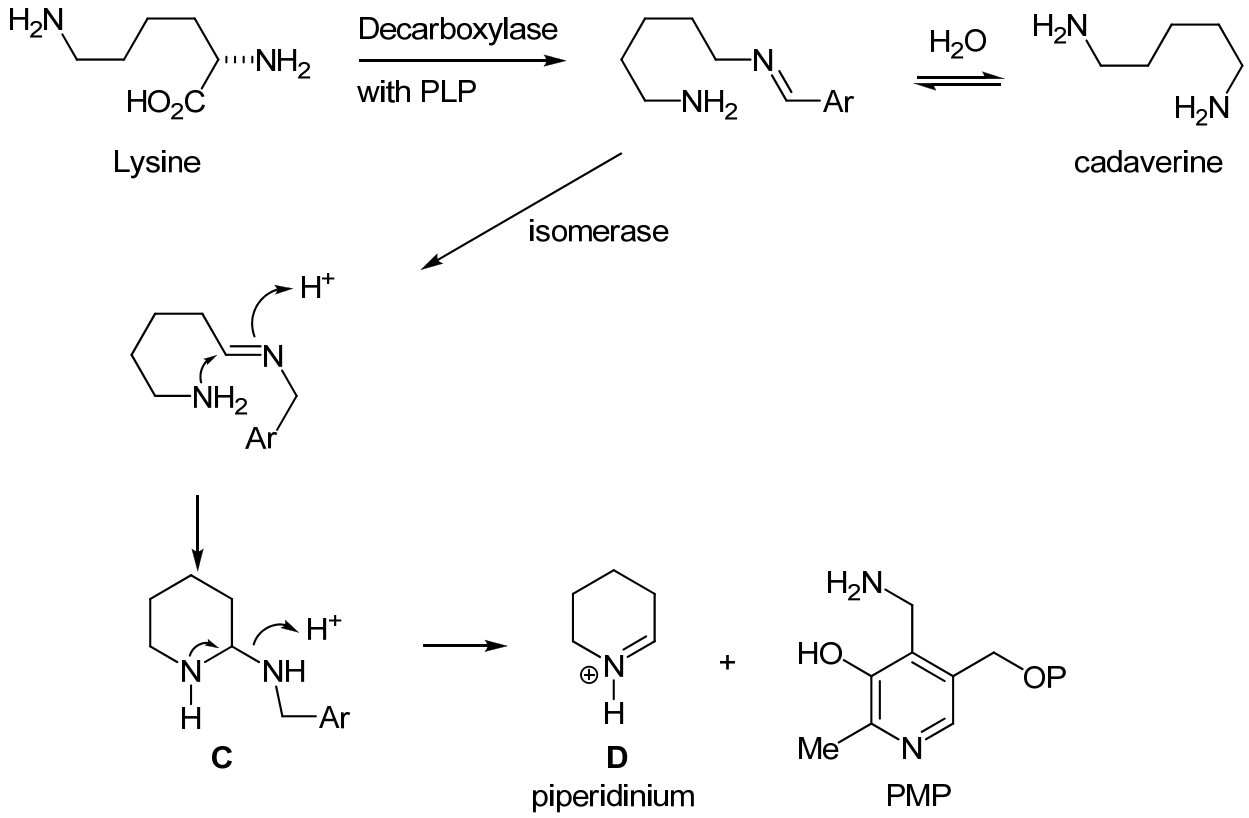
cyclization in an aldol-type reaction. The rest of the biogenesis to form alkaloids is still unknown at present and much studying remains to be done.



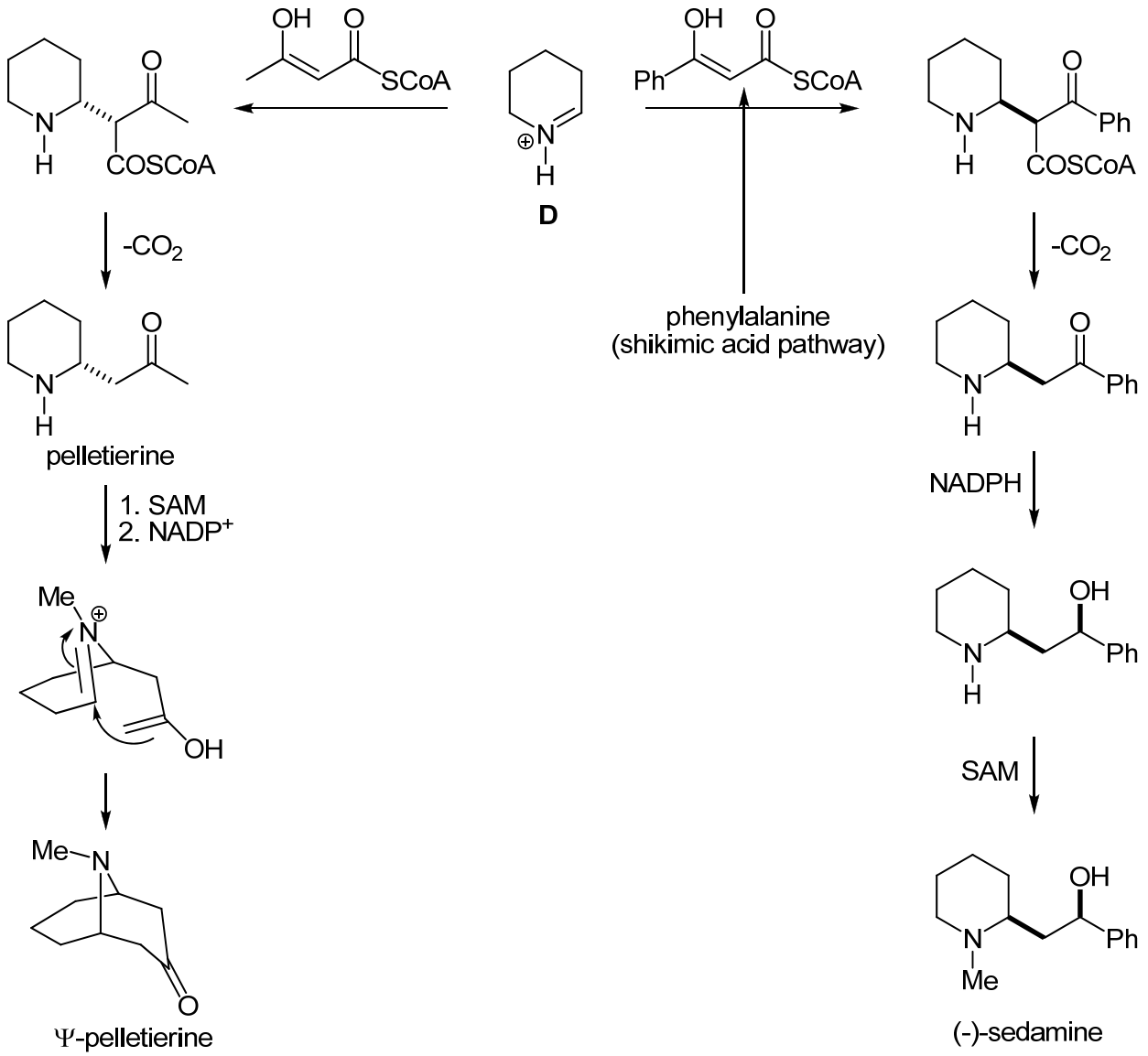
Scheme V.3.6

V.3.4 The Piperidine Alkaloids.

Lysine gives rise to the piperidine alkaloids and related compounds. Lysine suffers very similar biosynthetic transformations as ornithine in the first stages of biogenesis (Scheme V.3.7). Cadaverine is the symmetrical diamine related to putrescine. It seems that N-methylation in the piperidine alkaloids comes at a later stage in the biogenesis. For that reason the piperidinium salt **A** is not already methylated as was the case for the pyrrolidinium salt. Nucleophilic attack of this salt by acetoacetyl CoA followed by a similar sequence of reaction to the pyrrolidine alkaloids lead to pelletierine and ψ-pelletierine both related to hygrine and tropine respectively (Scheme V.3.8). On the other hand nucleophilic attack by benzoylacetyl CoA, derived from phenylalanine, leads to sedamine, after a series of decarboxylation, reduction and methylation.

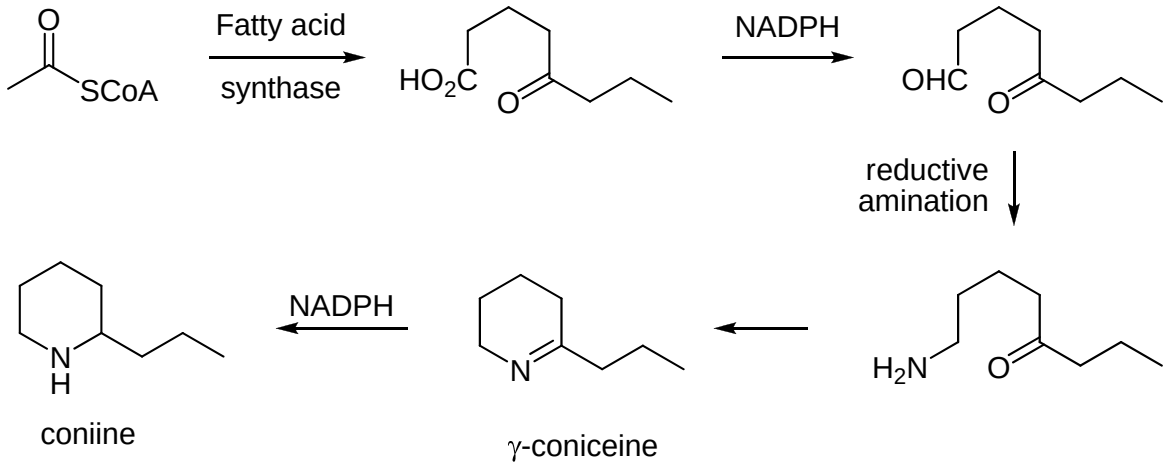


Scheme V.3.7



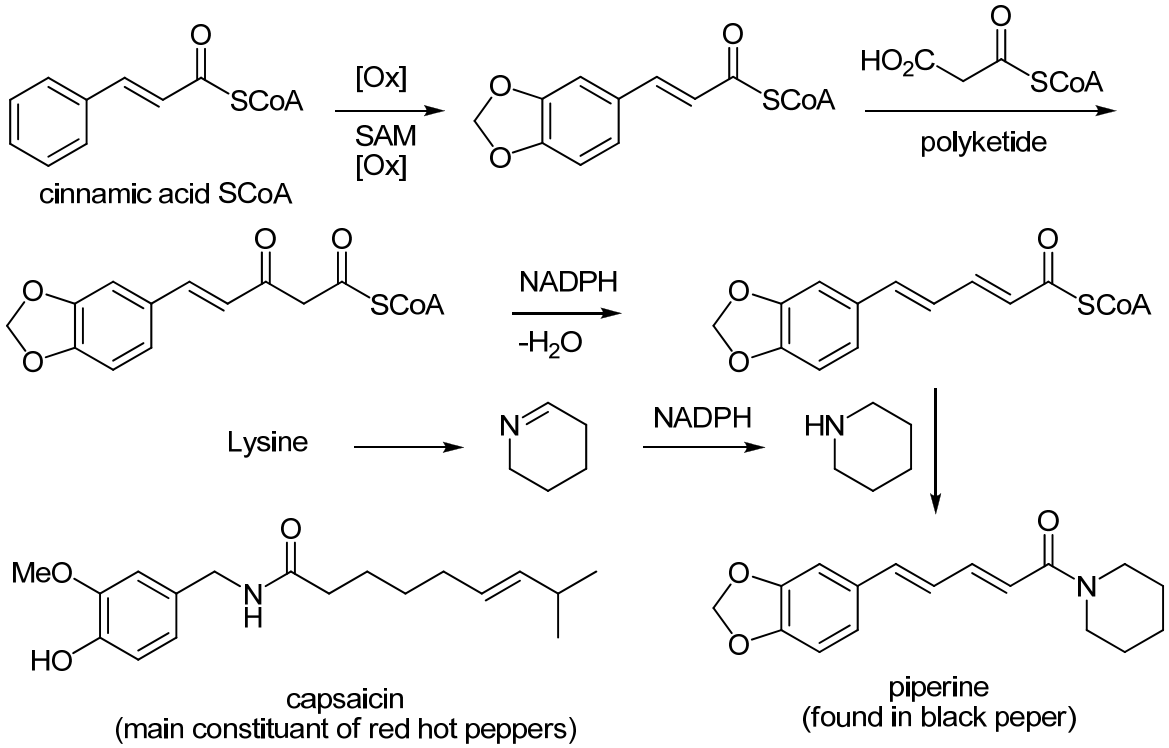
Scheme V.3.8

Not all alkaloids are derived from amino acids as we have already mentioned. In fact some piperidine alkaloids structurally similar to pelletierine and hygrine have a totally different biogenetic origin. Indeed, coniceine and coniine are two piperidine alkaloids that derive from the acetogenin pathway (Scheme V.3.9). This was proven by labelling studies of course, and it was shown that the nitrogen of the piperidine ring came from ammonia. This phenomenon demonstrates nature's resourcefulness at being able to make structurally similar compounds in totally different ways. Coniine is an acute poison that was used by ancient Greeks to execute state prisoners of which Socrates was certainly the most famous.



Scheme V.3.9

Other types of piperidine alkaloids are shown in Scheme V.3.10. Piperine is a bitter constituent of black pepper, an irritant. It is biosynthesized from phenyalanine (via cinnamic acid), acetate (via malonate) and piperidine (from lysine). Peppers comprise many alkaloids that are known as the pepper alkaloids some of which, like piperine, are piperidine alkaloids. Red and green peppers contain mostly capsaicin. The concentrated alkaloids are responsible for the "hot" taste of pepper when dried.



Scheme V.3.10

APP.V.1

Proposez une biosynthèse de la capsaicine sachant que la chaîne amide de droite inclut comme 'starter unit' une molécule d'isobutyryl-CoA.

V.3.5. The Quinolizidine Alkaloids.

These alkaloids are sometimes known as the Lupin alkaloids since they occur mostly in plants of the species *Lupinus*. Their biogenesis from lysine is perhaps analogous to that of the pyrrolizidine alkaloids. Some examples of simple and complex quinolizidine alkaloids are included in Figure V.3.4. The most common quinolizidine alkaloid is sparteine. Cytisine is a toxic principle of laburnum and has a similar effect to strychnine, the potent poison used as rodenticide and vermin killer (also killed many beloved pets). It causes nausea, convulsions, respiratory failure, etc.

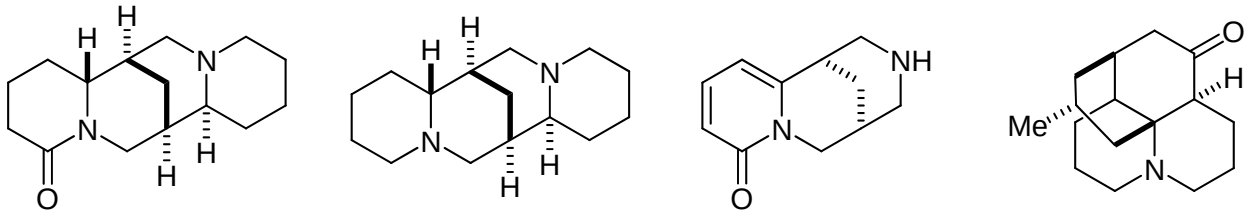
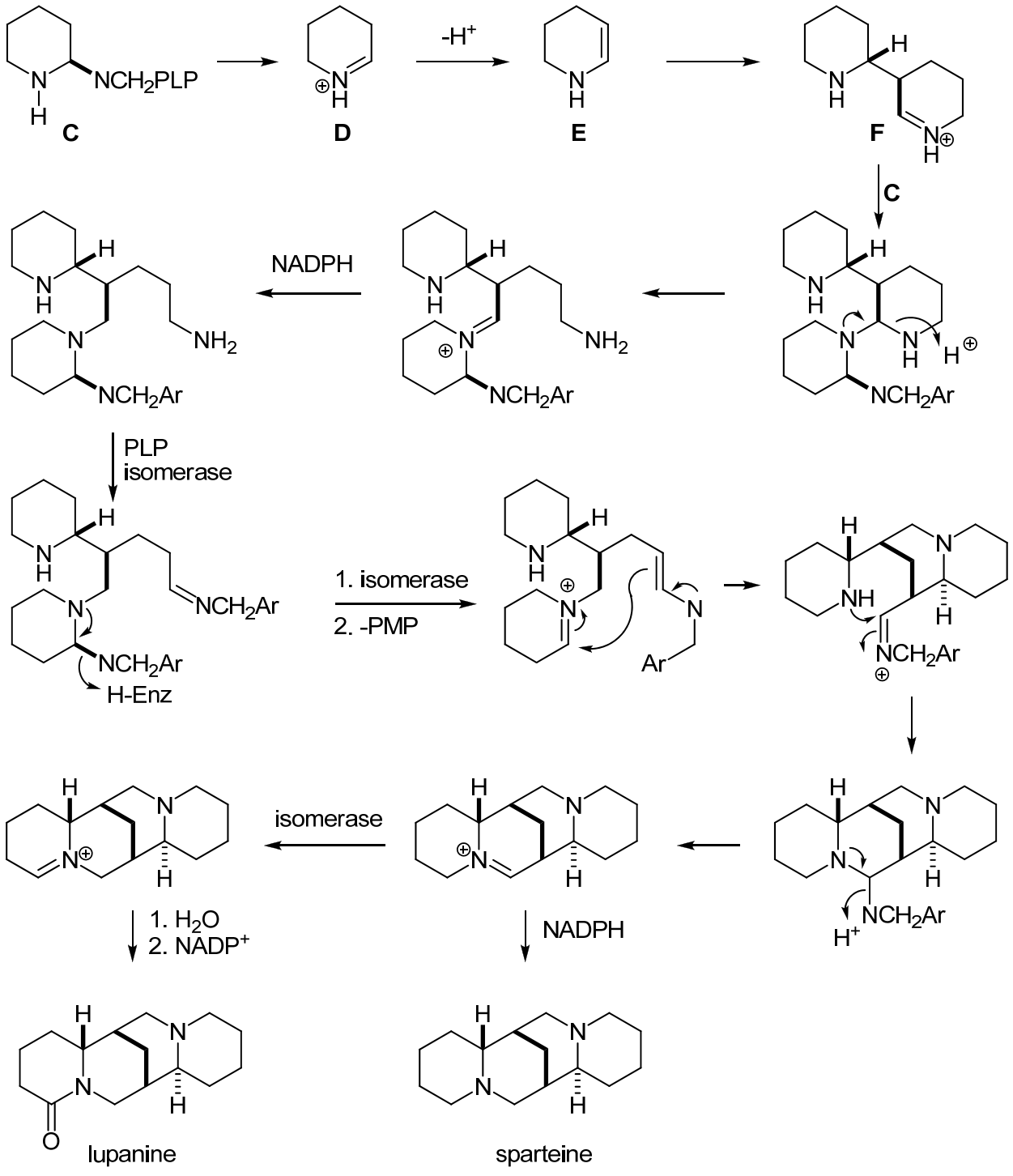


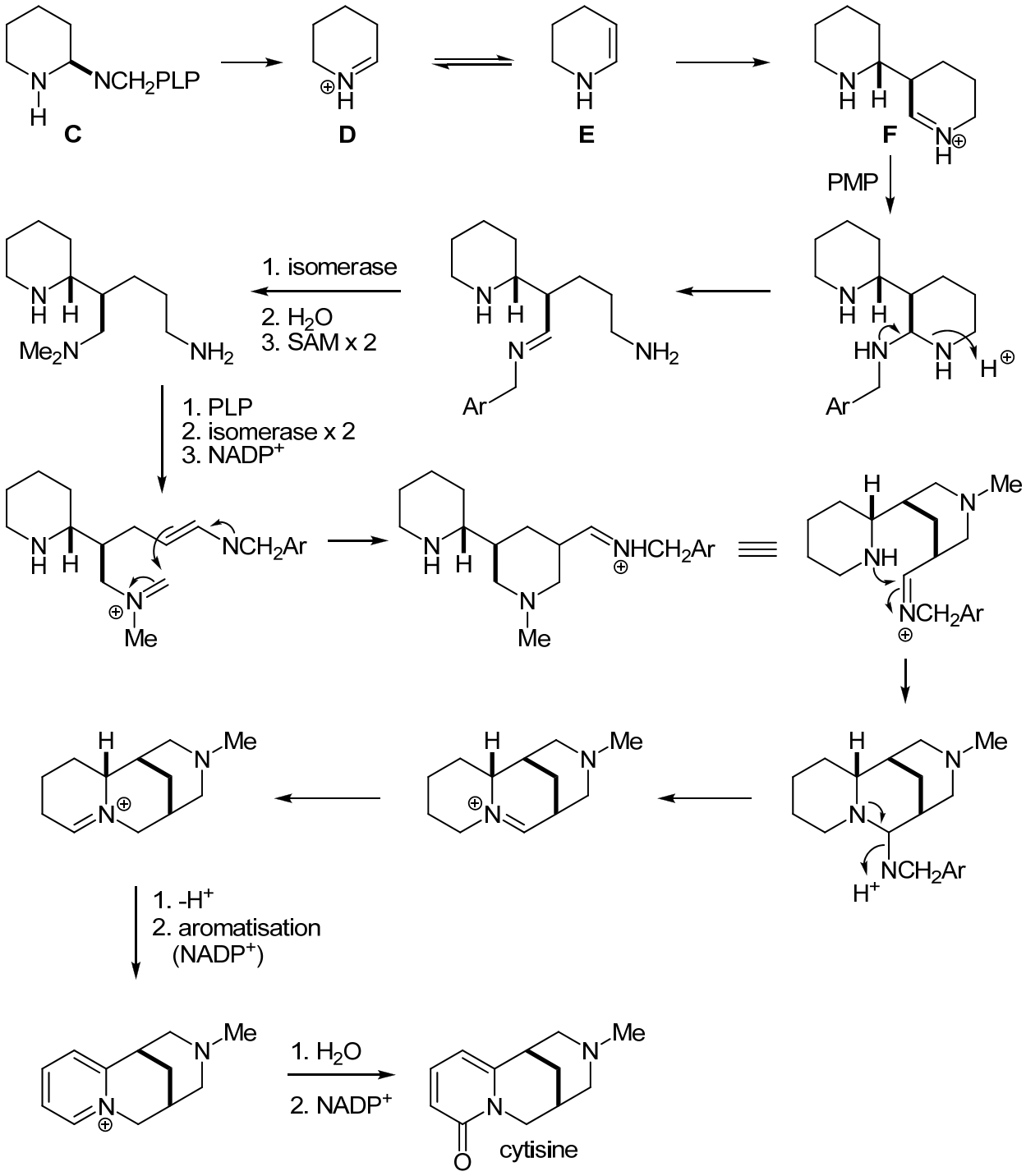
Figure V.3.4

The biosynthesis of lupanine and sparteine involve three molecules of lysine and is shown in Scheme V.3.11. It starts with piperidinium **D**, which can be lose a proton to produce enamine **E**. Piperidinium **D** and enamine **E** are coupled to give a new iminium, which can add piperidine **C**. The tris-piperidine compound is reduced by NADPH. Further transformation involving another molecule of PLP and an oxidative amination ends in a bis-cyclization to give the sparteine carbon skeleton. From here to the natural alkaloids is a simple matter of reduction or oxidation.

The biogenesis of cytosine is slightly different (Scheme V.3.12). It only involves two molecules of lysine. Formation of iminium **F** is identical but the amine of PMP is adding to it rather then piperidine **C**. One primary amine is dimethylated and after generation of the enamine (much the same way as for the synthesis of sparteine), the latter adds to the iminium formed from the oxidation of one of the methyl groups. Cyclization also follows the sparteine/lupanin pathway but here aromatization occurs by way of $NADP^+$ oxidation. The pyridinium suffers an attack by water at the only available ortho position and a last oxidation produces cytisine.


NCH2PLP
C
D
-H+
E
F
C
NADPH
NH2
NCH2Ar
PLP
isomerase
NCH2Ar
H-Enz
1. isomerase
2. -PMP
Ar
NH
NCH2Ar
H+
isomerase
1. H2O
2. NADP+
NADPH
lupanine
sparteine

Scheme V.3.11


NCH2PLP
C
D
E
F
PMP
1. isomerase
2. H2O
3. SAM x 2
Me2N
NH2
1. PLP
2. isomerase x 2
3. NADP+
NCH2Ar
NHCH2Ar
Me
Ar
HN
NH
1. -H+
2. aromatisation
(NADP+)
1. H2O
2. NADP+
cytisine
O

Scheme V.3.12

V.4. Alkaloids Derived from Aromatic Amino Acids.

The alkaloids derived from the amino acids phenylalanine, tyrosine and tryptophan are of an extremely diverse nature. They range from quite simple amines such as mescaline, through simple tetrahydroisoquinolines such as pellotine to complex monomeric alkaloids such as morphine (Figure V.4.1). There are also examples where the tyrosine-derived portion of the molecule is less discernible e.g. the mycotoxin alkaloid gliotoxin. The latter induces apoptosis in immune cells, including macrophages, and is thus immunosuppressant. With so many varied structural types, it is not surprising to find that this group of alkaloids covers the most part of pharmacological responses, and many of these are of considerable therapeutic significance. One of the special features of this group of compounds is their distribution, which stretches from the bacterial and fungal world into marine organisms, plants, lower animals, and all the way to humans. The following sections deal with some important tyrosine, phenylalanine, and tryptophan derived alkaloids in order of increasing complexity.

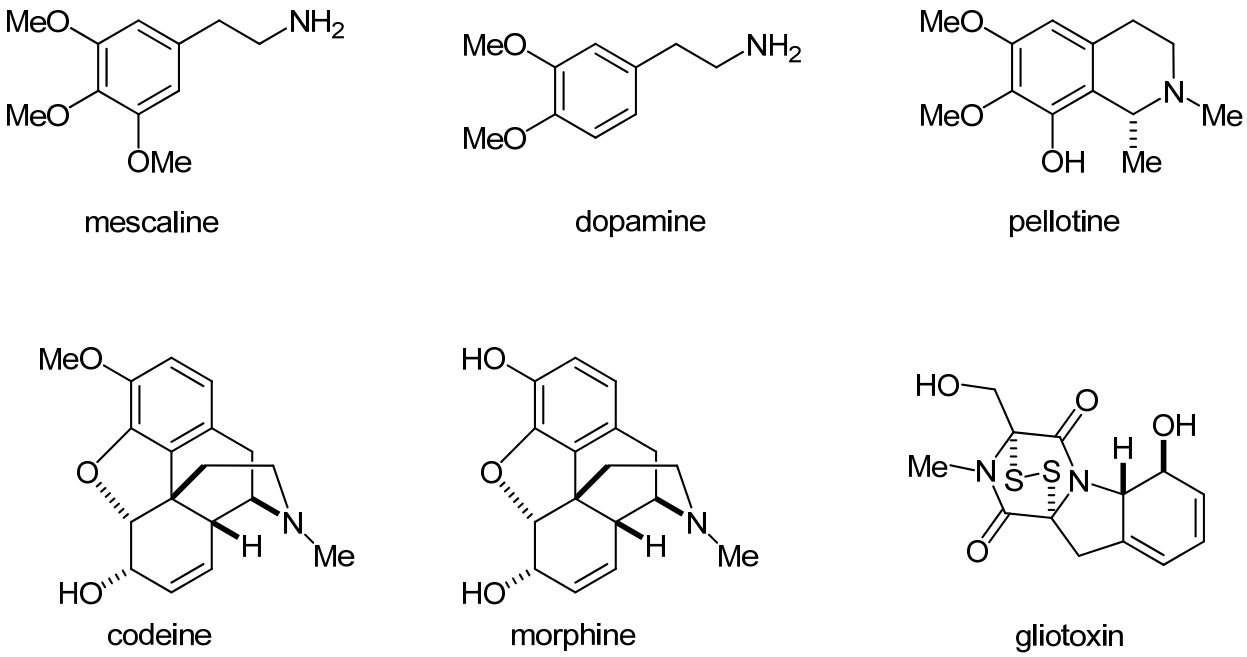
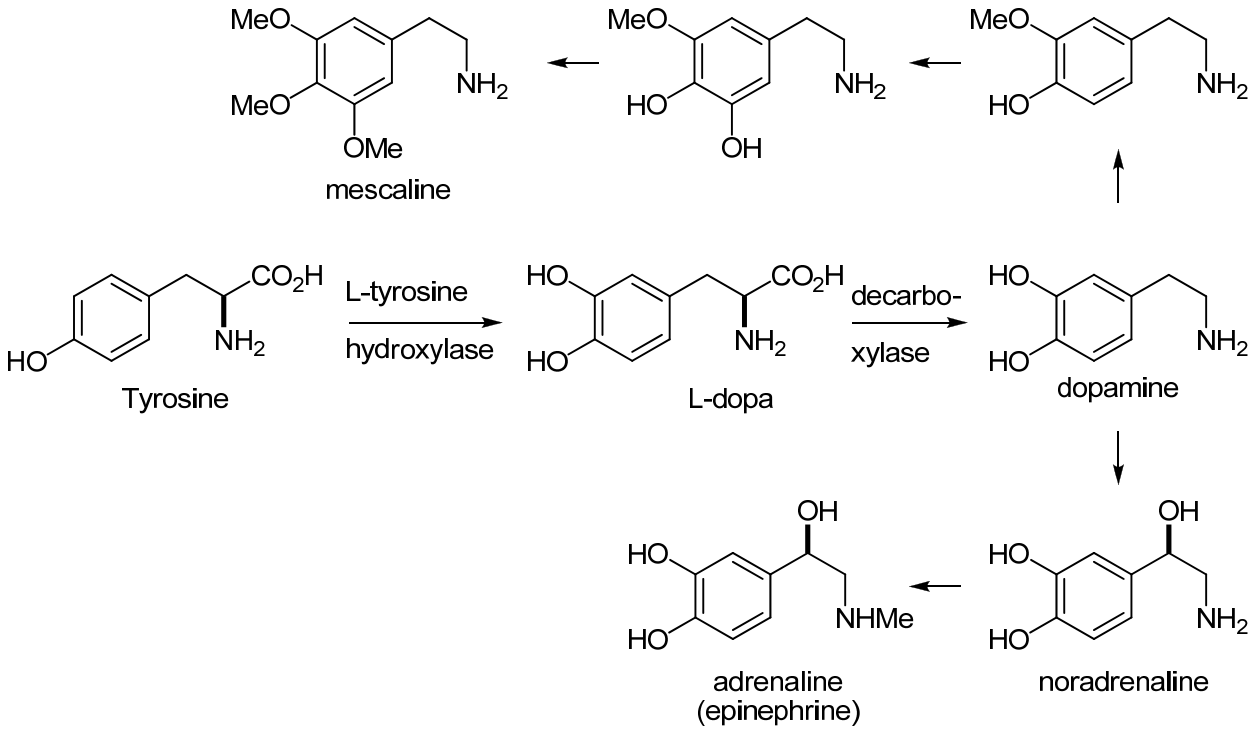


Figure V.4.1

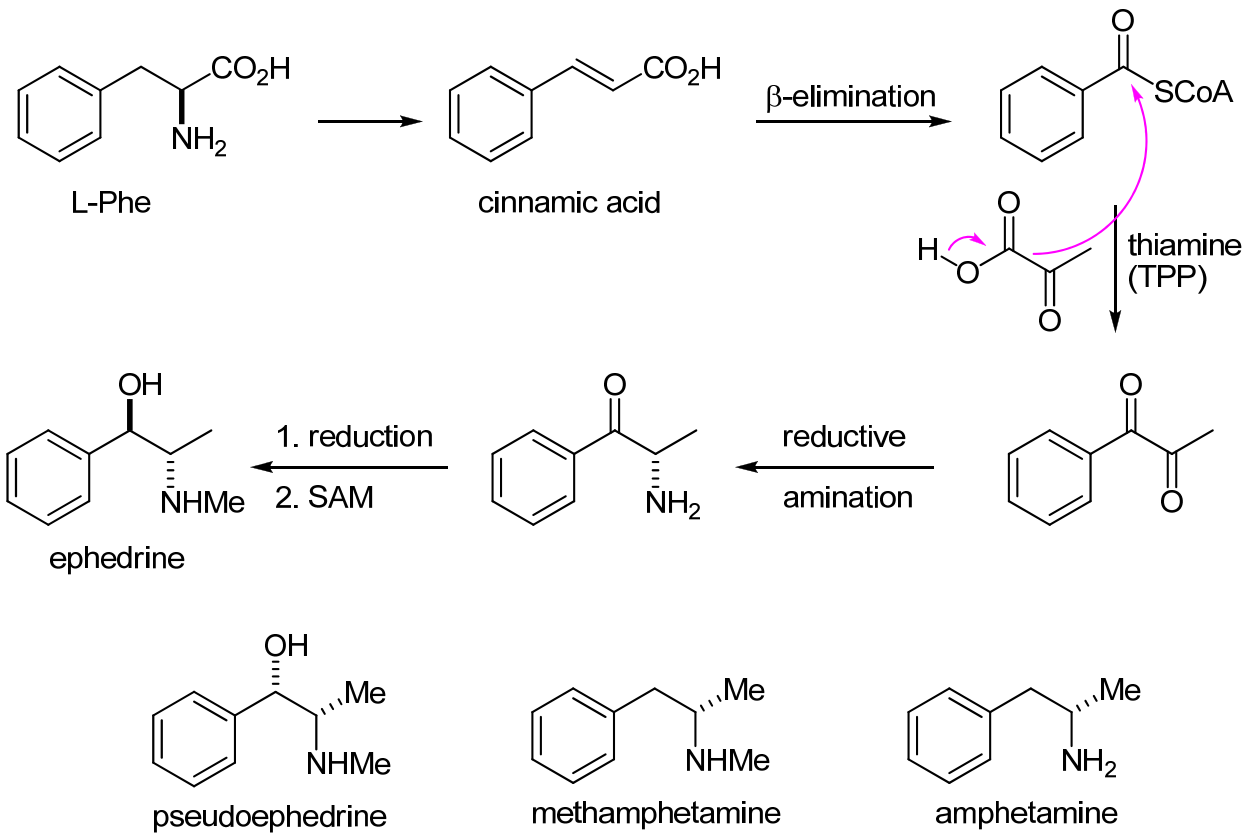
V.4.1. Simple Monocyclic Tyrosine Alkaloids.

Tyrosine is obtained by hydroxylation of prephenylalanine as discussed in section IV.1. Dopamine, adrenaline and noradrenaline are widely distributed alkaloids in the animal kingdom

and humans where they serve as a neurotransmitters i.e. chemical messengers between neuron cells. The steps involving their formation is shown in Scheme V.4.1 (top). Tyrosine is first hydroxylated ortho to the hydroxy group and a decarboxylation provides dopamine. An aliphatic hydroxylation gives noradrenaline and methylation of the latter affords adrenaline. Methamphetamine, amphetamine and ephedrine cause the liberation of noradrenaline from storage sites in the sympathetic nerve ending and thus stimulate the central nervous system. The first two are prone to abuse since they excite the 'reward' psychological system by increasing the levels of serotonine, dopamine and norephedrine. The biosynthesis of ephedrine and related alkaloids are shown in Scheme V.4.2.



Scheme V.4.1



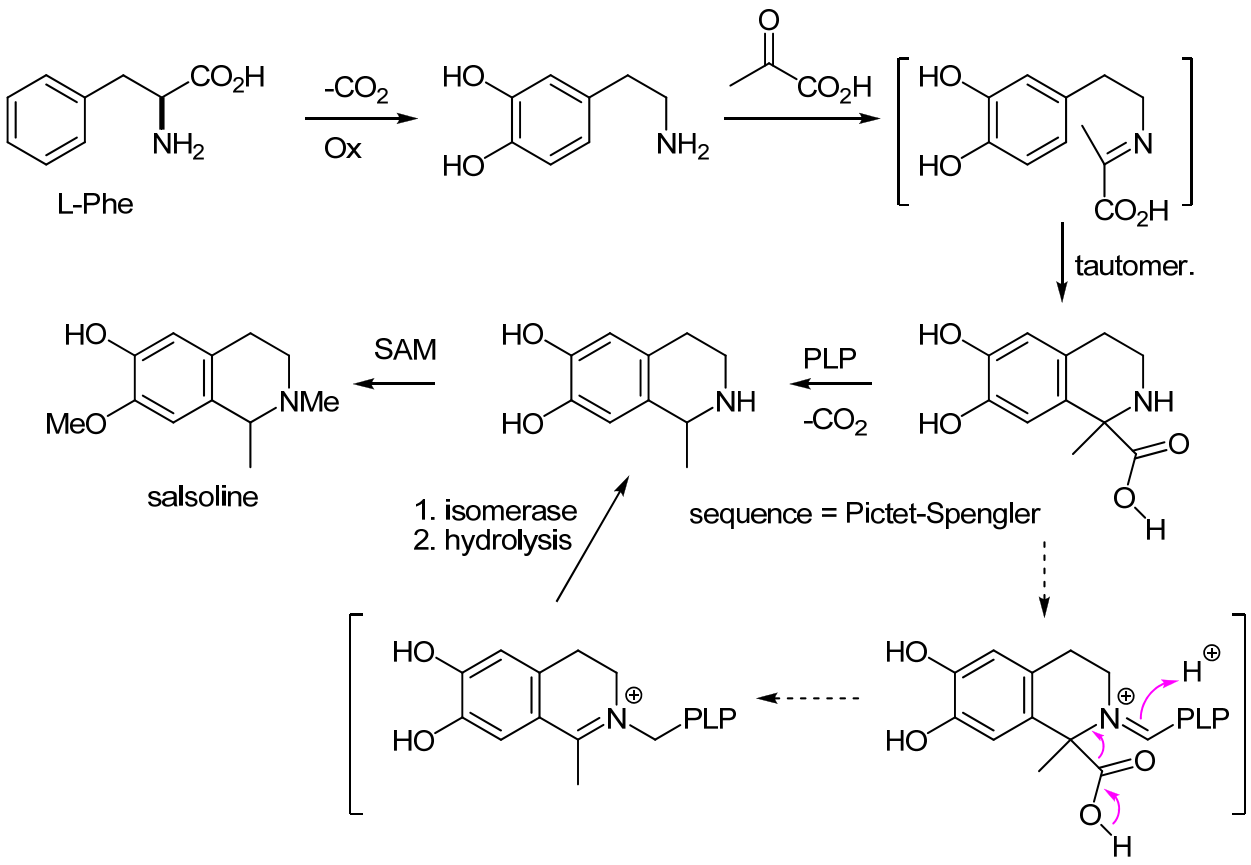
Scheme V.4.2

Peyote is one of the very old hallucinogenic drugs of the New World. The concoction is made from extracts of the peyote cactus found in the Chihuahuan Desert of Texas and Mexico (Figure V.4.2). Native Indians are still using it today for religious ceremonies in the U.S.A. but all other uses are federally controlled. Chemically, peyote is composed of many alkaloids (over 50) of which mescaline is the principal hallucinogen. Its biosynthesis is shown in Scheme V.4.1 (top). The effect of mescaline in humans has been known since the late 16th century. Some described its effect as good, charming, enhancing the beauty of life. Others, however, have described it as frightful and unpleasant. The initial physical reactions are nausea and vomiting after about 30 minutes. These then subside before psychic effects begin. Daily use of mescaline produces tolerance and a cross tolerance to lysergic acid diethylamide (LSD) but does not produce addiction. Thus mescaline cannot be regarded as a narcotic drug.

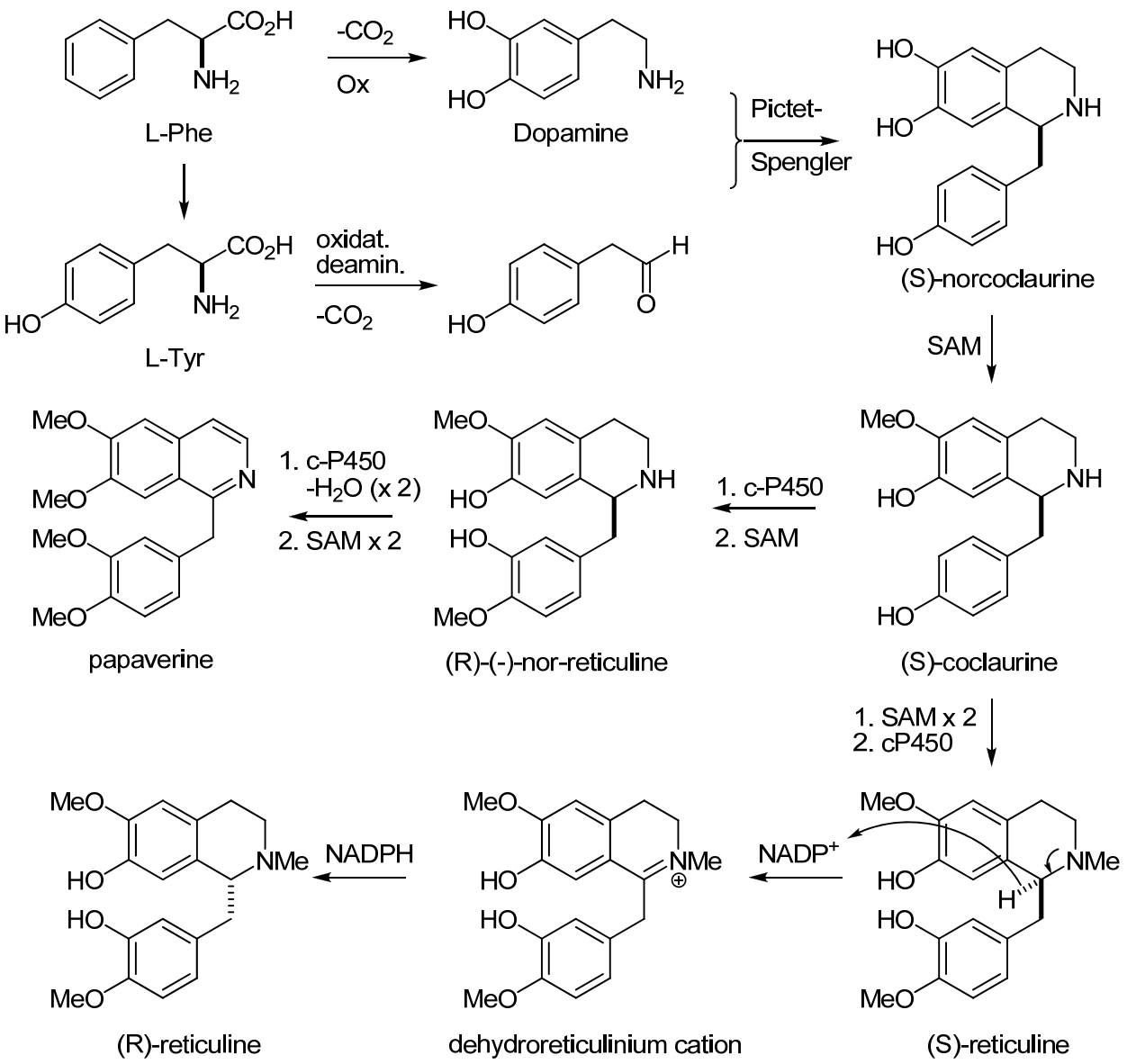
Figure V.4.2

V.4.2. Isoquinoline and Benzylisoquinoline Alkaloids.

This family of alkaloids is probably the largest of any group. The isoquinoline nucleus occurs in a vast array of structure types and all are derived from phenylalanine or tyrosine. The simple isoquinoline structure and biosynthesis of salsoline is depicted in Scheme V.4.3 and it is formed from phenylalanine or tyrosine and pyruvic acid via the Pictet-Spengler cyclization. Decarboxylation occurs during the Pictet-Spengler via PLP, as shown, when small molecules are involved (pyruvic acid, glyoxylic acid). When larger molecules are involved, decarboxylation has occurred before the Pictet-Spengler, which becomes a simpler transformation. The benzylisoquinoline alkaloids comprise a large array of structures and are derived from phenylalanine or tyrosine and β-phenylpyruvic acid. The benzylisoquinoline nucleus is depicted in Scheme V.4.4 along with the proposed biosynthesis of its most important members (pharmacologically speaking) (*S*)-reticuline, nor-reticuline and papaverine. The latter was isolated from opium in 1848 (0.8-1.0% in weight) and is a smooth muscle relaxant and a cerebral vasodilator. It is used in treatment of angina where it releases the pressure in the artery by vasodilation. It is not a narcotic as it creates no physiological dependence. Note that many of the opium alkaloids are made from (*R*)-reticuline and apparently it itself comes from (*S*)-reticuline via an $NADP^+$ oxidation and reduction.

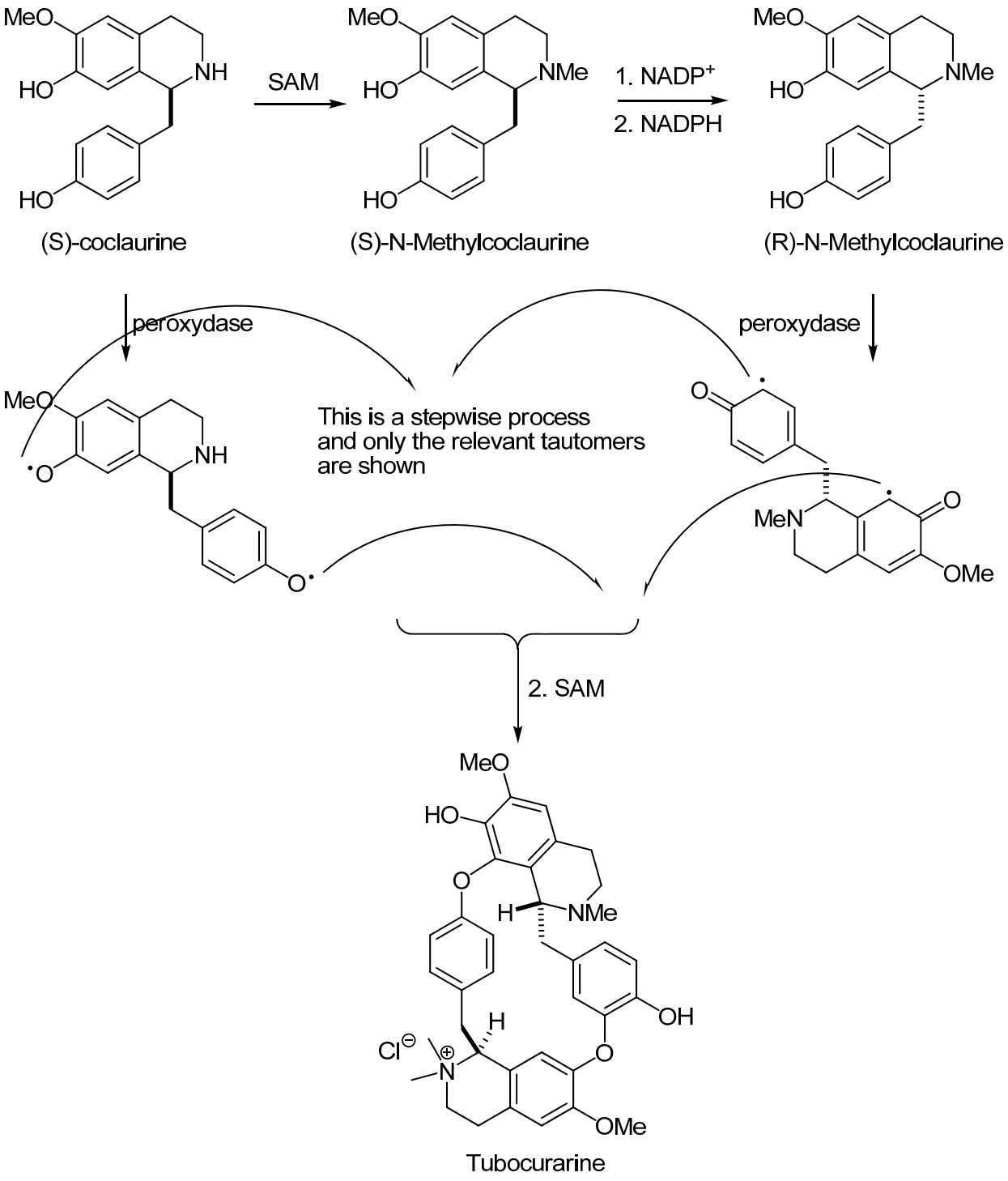

L-Phe
-CO2
Ox
CO2H
tautomer.
PLP
-CO2
SAM
salsoline
1. isomerase
2. hydrolysis
sequence = Pictet-Spengler
PLP

Scheme V.4.3



Scheme V.4.4

The more important aspect of the benzylisoquinoline alkaloids is that they are the immediate precursors of the morphinan alkaloids. Several types of oxidative phenolic coupling give rise to different types of modified benzylisoquinoline alkaloids, of which the morphinans is one. Figure V.4.3 lists a few alkaloid skeletons that are biogenetic derivatives of the benzylisoquinolines. We will discuss examples of biosyntheses from each family. They mainly differ by the position of the coupling (para-ortho, para-para etc.). Even coupling between carbon and oxygen do occur as is evident in the curare alkaloids. The biosynthesis of important members of this first class of modified benzylisoquinoline alkaloids is depicted in Scheme V.4.5.


MeO
HO
NH
HO
SAM
MeO
HO
NMe
HO
1. NADP+
2. NADPH
MeO
HO
NMe
HO
(S)-coclaurine
(S)-N-Methylcoclaurine
(R)-N-Methylcoclaurine
peroxydase
peroxydase
This is a stepwise process
and only the relevant tautomers
are shown
MeO
NH
O
O
O
MeN
O
OMe
2. SAM
MeO
HO
O
H
NMe
H
OH
Cl⊖
N⊕
O
OMe
Tubocurarine

Scheme V.4.5

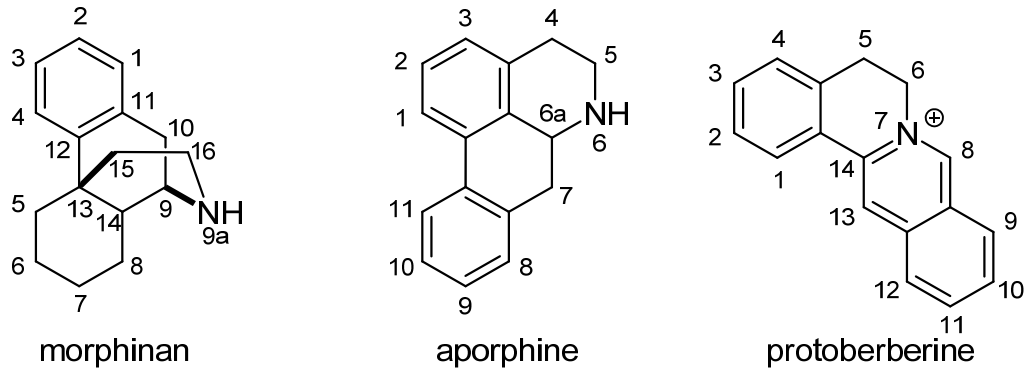
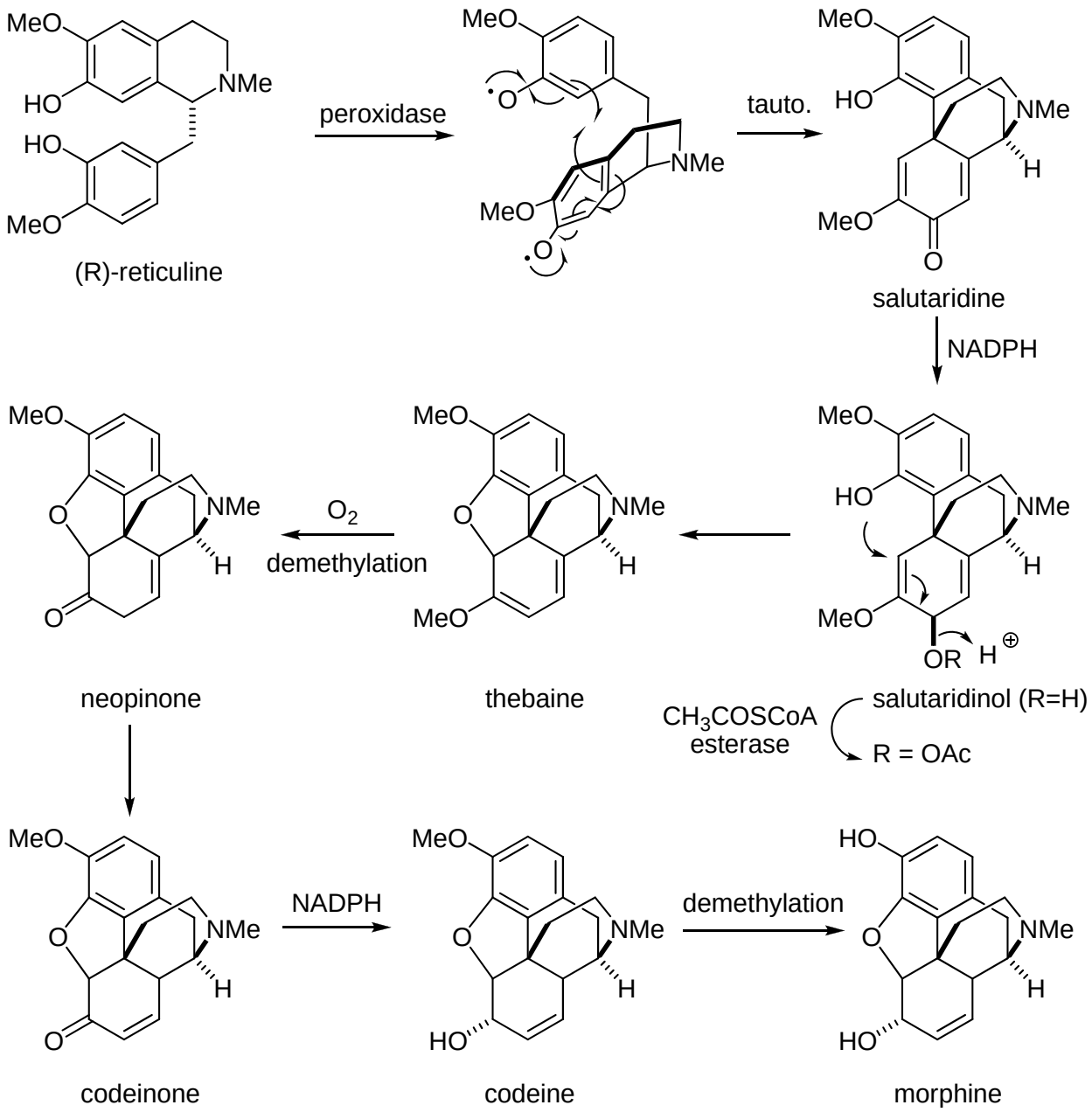


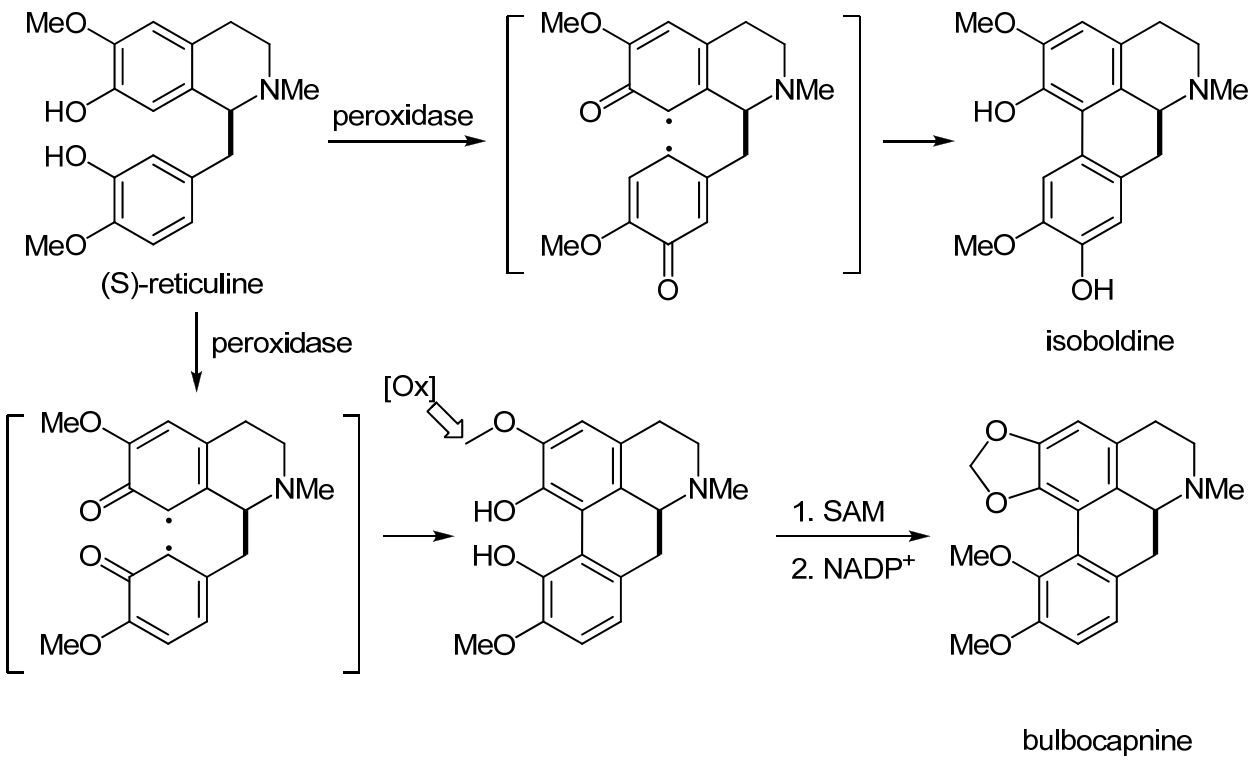
Figure V.4.3

Thus (*R*)-reticulene (made from (*S*)-reticulene, see Scheme V.4.4) leads to morphine and codeine by oxidative phenolic coupling (see section III.4.4). Scheme V.4.6 illustrates the biosynthesis of these two morphine alkaloids from reticuline. The phenolic coupling takes place ortho to the hydroxy group in the benzyl ring and para to the hydroxy group in the isoquinoline ring as shown. Historically, the most important morphine alkaloids are those of the ancient drug opium. Opium is the air-dried milky extract from incised, unripe capsules of *Papaver somniferum*. It was well known in ancient Greece but the smoking of opium was first noted to be extensive in China and the Far East in the latter part of the eighteenth century. Addiction to the drug subsequently led to its illegalisation and traffic proved to be a large source of revenue. Over 40 alkaloids are known and several are of major significance. These are morphine, codeine, thebaine, noscapine (previously known as narcotine), and papaverine. The morphine content of opium is in the range of 4-21%, noscapine 4-8%, and all the other previously mentioned alkaloids 0.5-2.5%. These alkaloids occur, at least in part, bound to meconic acid.



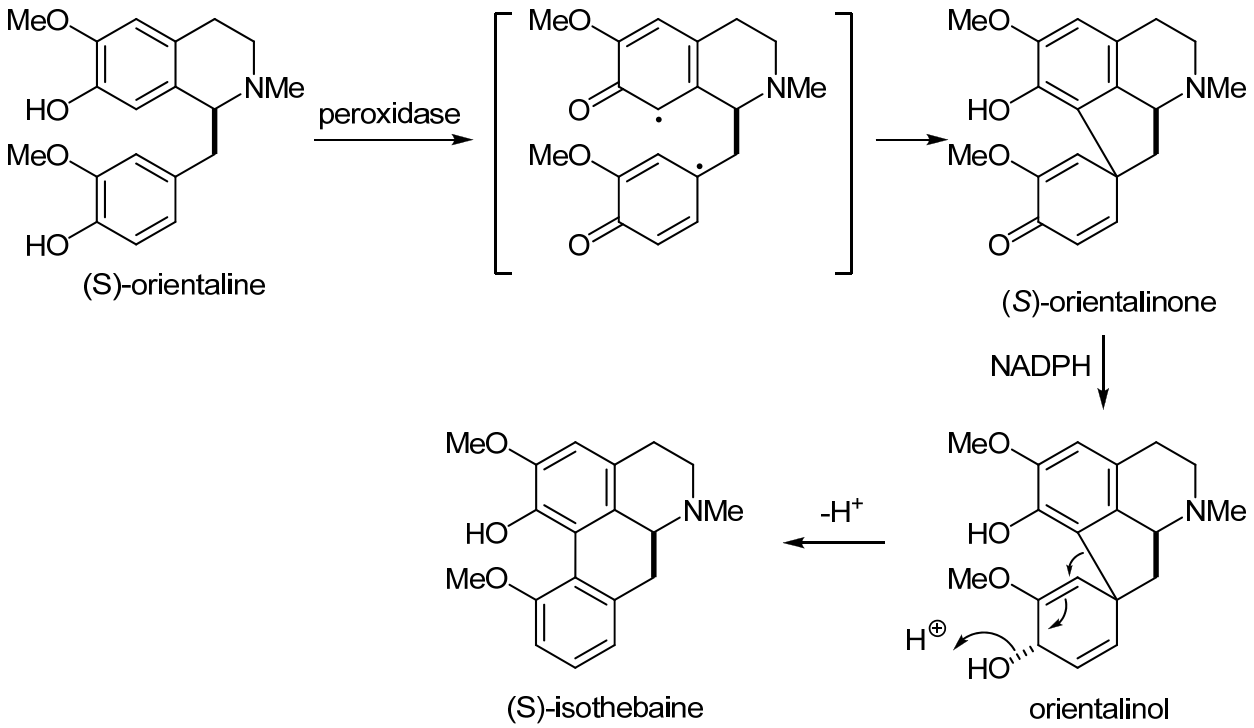
Scheme V.4.6

Other types of phenolic coupling lead to the formation of such alkaloid as isoboldine, stemming from a coupling ortho to the hydroxy group in the isoquinoline ring and para to the hydroxy group in the benzyl ring (Scheme V.4.7). Bulbocapnine represents a case of ortho-ortho coupling.

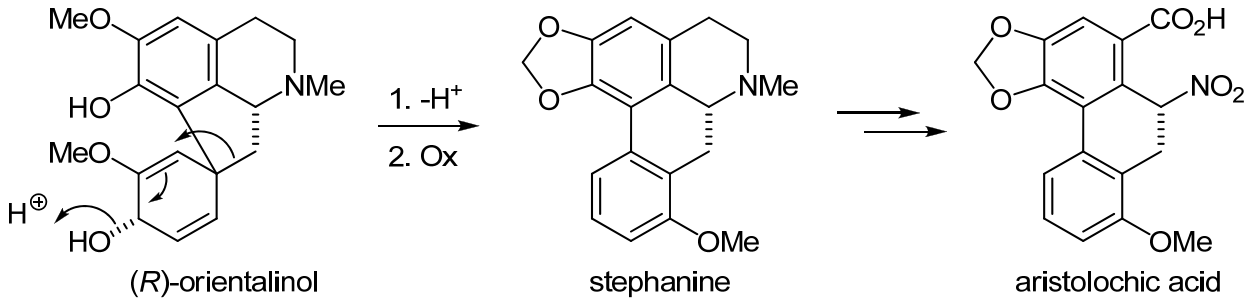


Scheme V.4.7

Alkaloids that are seemingly resulting from a ortho (isoquinoline)-meta (benzyl) coupling are in fact derived from a ortho (isoquinoline)-para (benzyl) coupling. A diradical coupling (*S*)-orientaline, for example, give the intermediate (*S*)-orientalinone (Scheme V.4.8). After reduction to orientalinol, departure of water spurs a rearrangement of the skeleton driven by the aromaticity of the product (*S*)-isothebaine. The alkaloid stephanine isolated from *stephania* species is made analogously but starting from (*R*)-orientaline and involving the migration of the methylene rather than the aryl group (Scheme V.4.9). The aporphine aristolochic acid is an oxidation product of stephanine.



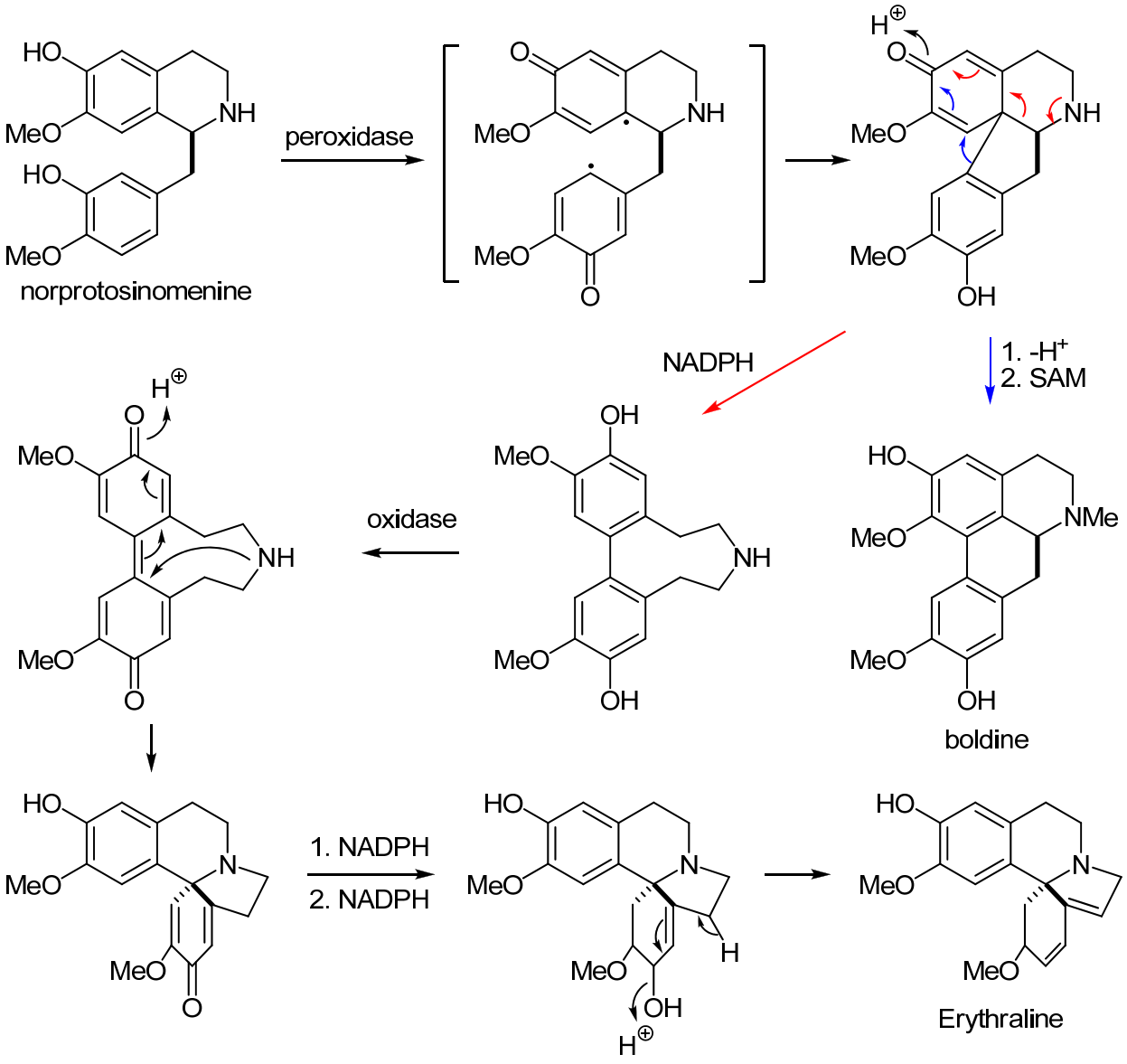
Scheme V.4.8



Scheme V.4.9

Para-para coupling also occur and two biosynthetic routes are accessible from the intermediate of that series. Boldine (compare to isoboldine in Scheme V.4.7) is obtained when the enzyme catalyzes the migration of the aryl bond to the meta position (blue arrows in Scheme V.4.10). Norprotoseinomenine was thought to be the principal precursor of the erythrina alkaloids via the cleavage of the C-N bond (red arrows in Scheme V.4.10) to give an iminium ion that is then reduced with NADPH. Recently, however, norreticuline (see Scheme V.4.4) was shown to be the precursor of the erythrina alkaloids, which would imply methylation and demethylation steps.[2] Regardless, the 9-membered cyclic amine cyclizes, once the bisphenol ring

system has been oxidized to the bisquinone system. This cyclization affords the erythraline skeleton and standard reduction and elimination steps leads to erythraline. Erythrinas constitue a large family of trees that bear bright red flowers (mostly). Pictures of *Erythrina crista-galli* and of *Erythrina xsykesii* (found in New-Zealand) are shown in Figure V.4.4

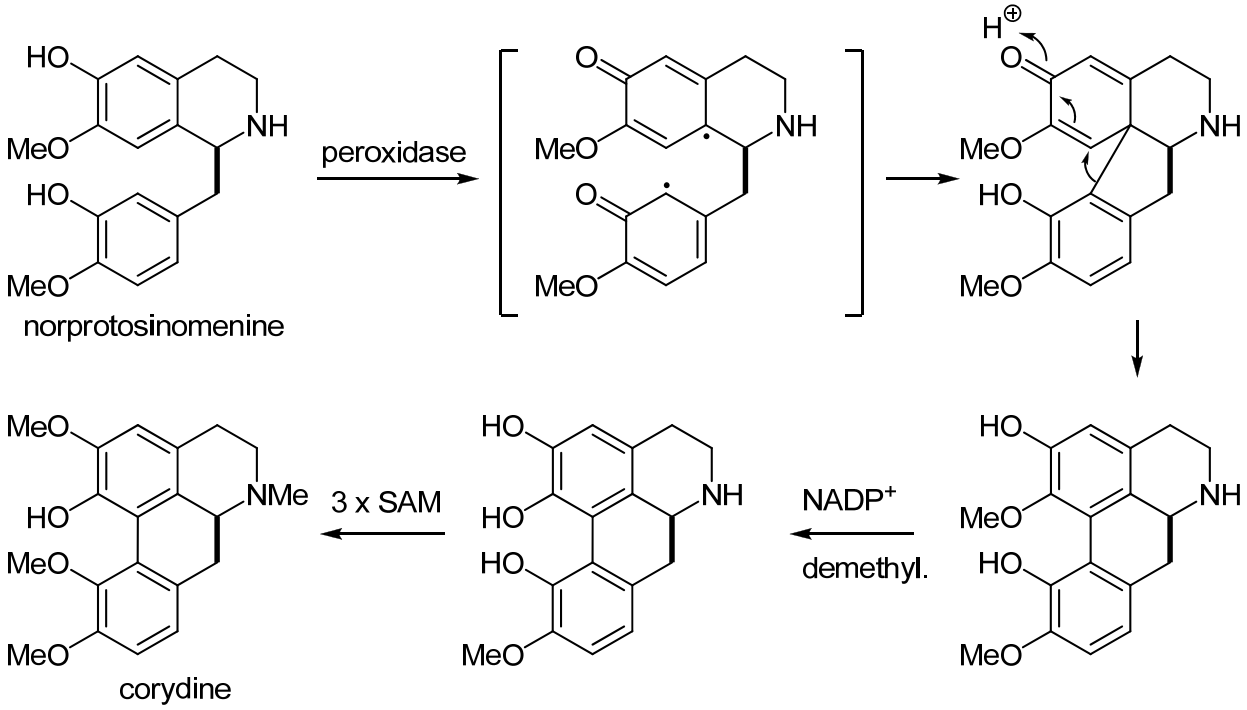


Scheme V.4.10

Erythrina crista-galli *Erythrina xsykesii*

Figure V.4.4

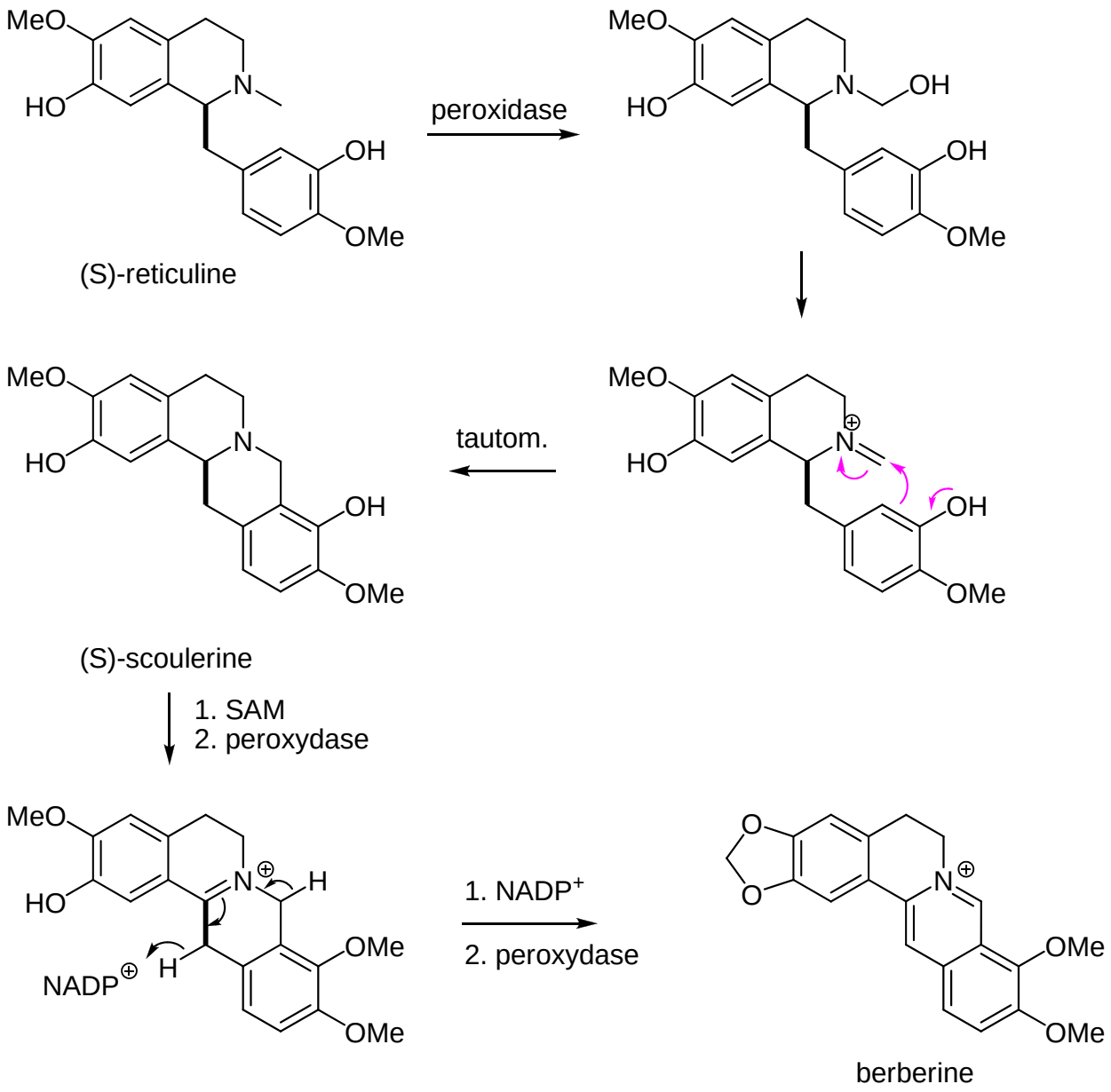
Finally, a last phenolic coupling type is pictured in Scheme V.4.11. It involves coupling at the para position of the isoquinoline fragment with that of the ortho position of the benzyl fragment. After demethylation and methylation of two hydroxyls and the amine, the alkaloid corydine is produced.



Scheme V.4.11

Reticuline is precursor to yet another family of alkaloids named protoberberine. Berberine is the prime example and its biosynthesis is shown in Scheme V.4.12. The so-called ‘berberine

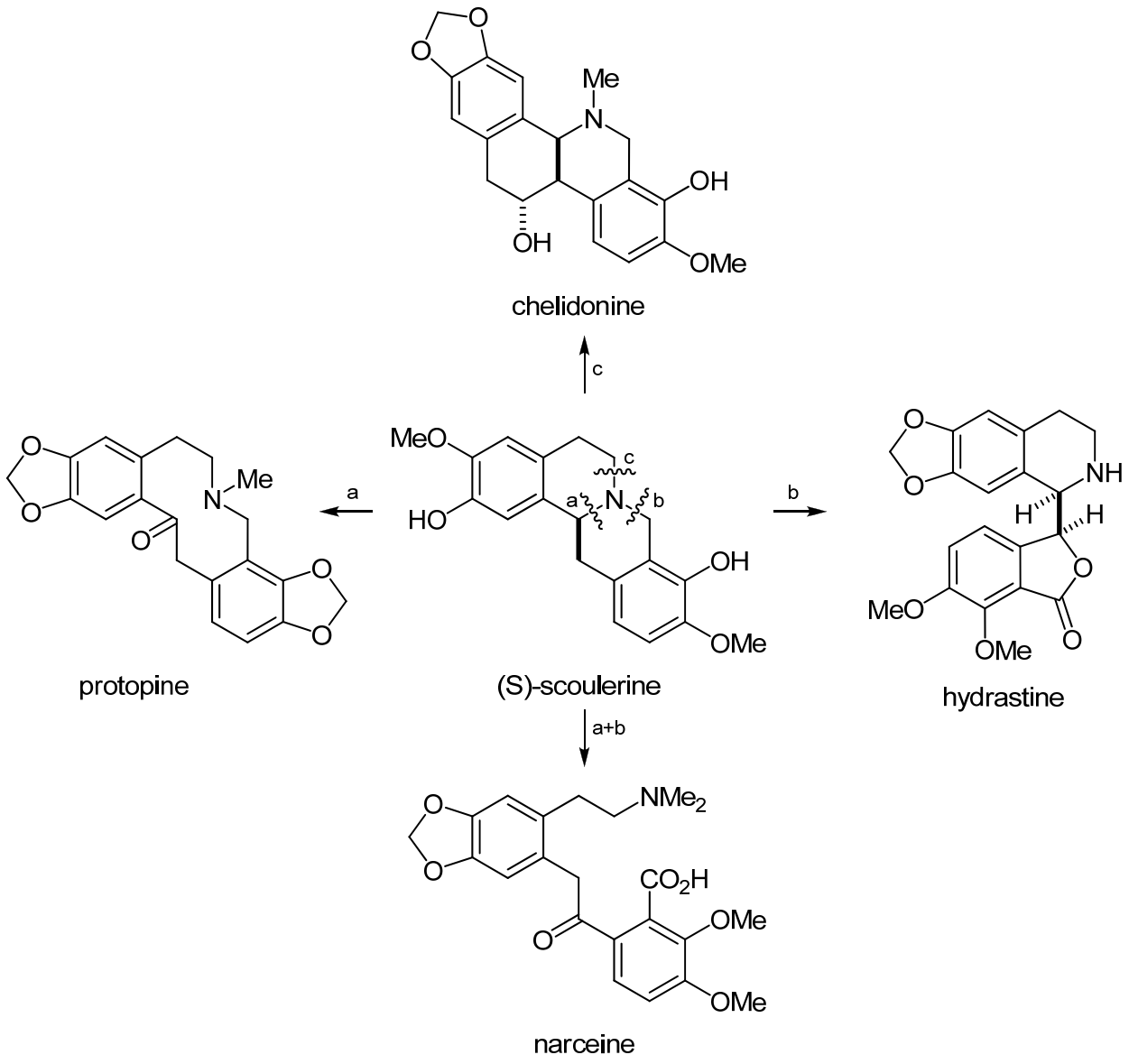
bridge' that effectively bridges the nitrogen's methyl group with the benzyl aromatic nucleus comes from an iminium ion. The latter is generated by a peroxidase oxidation of the methyl group (see the annexe section VII.1.5). Methylation and further oxidation leads to a fully aromatic system.



Scheme V.4.12

The intermediate (*S*)-scoulerine is a precursor of a whole sub-family of benzylisoquinoline alkaloids. As shown in Scheme V.4.13, enzymatic cleavage of one or two C-N bonds lead to

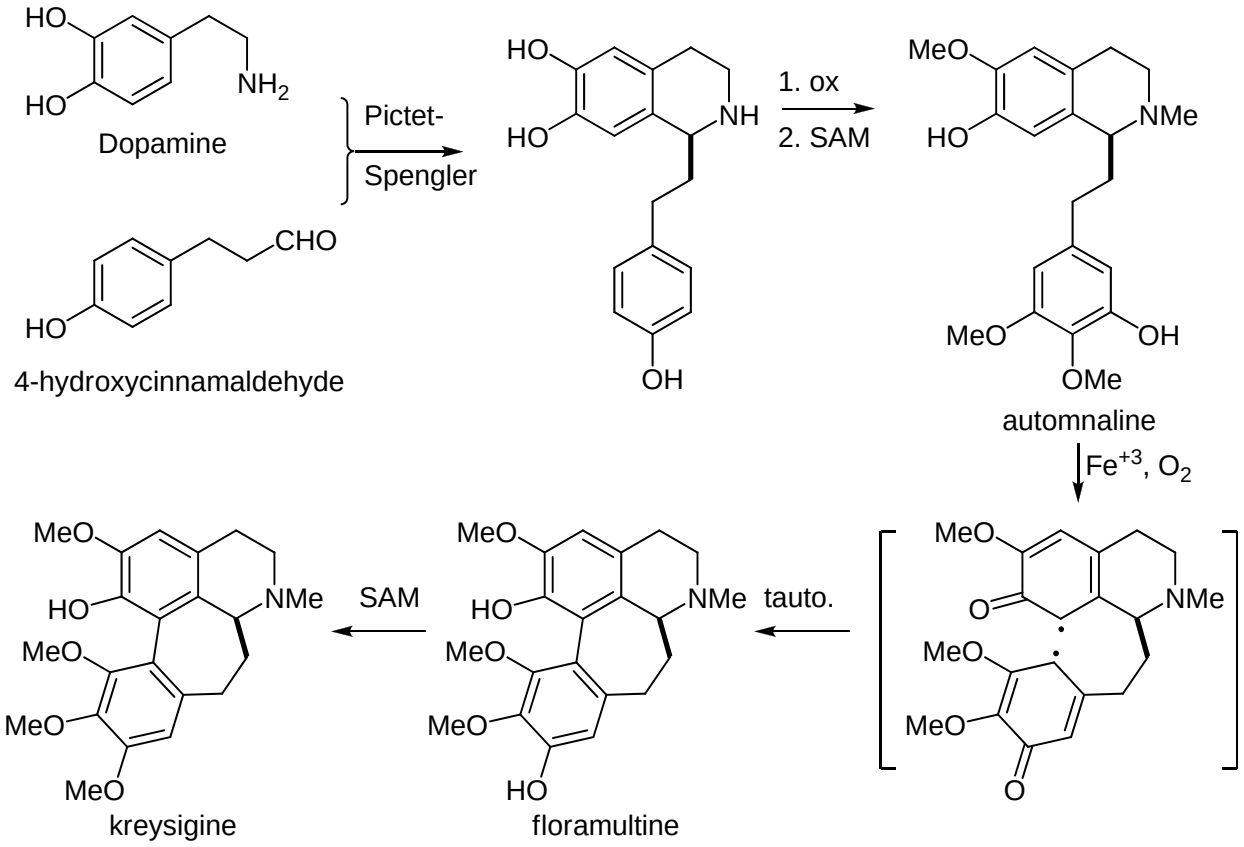
several alkaloids. Cleavage of the internal C-N bond labeled ‘a’ in Scheme V.4.10 gives protopine, after several standard biosynthetic steps. Cleavage of bond ‘b’ leads to hydrastine belonging to the phthalideisoquinoline family of alkaloids. Cleavage of bond ‘c’ affords chelidonine of the benzophenanthridine alkaloid clan. Finaly, if bonds ‘a’ and ‘b’ are cleaved, then narceine is produced, an opium alkaloid from *papaver somniferum*.



Scheme V.4.13

V.4.3. Phenethylisoquinoline Alkaloids.

Several alkaloids are derived from dopamine and dihydrocinnamaldehyde in a biosynthetic sequence very alike that of the benzylisoquinoline. While tyrosine was the ultimate source of the two fragments of benzylisoquinoline alkaloids (the isoquinoline and benzyl fragment), dopamine is derived from tyrosine but dihydrocinnamaldehyde is derived from phenylalanine. The biosynthesis of kreysigine is shown in Scheme V.4.14 and needs little comments beyond what has been said for the benzylisoquinoline alkaloids.

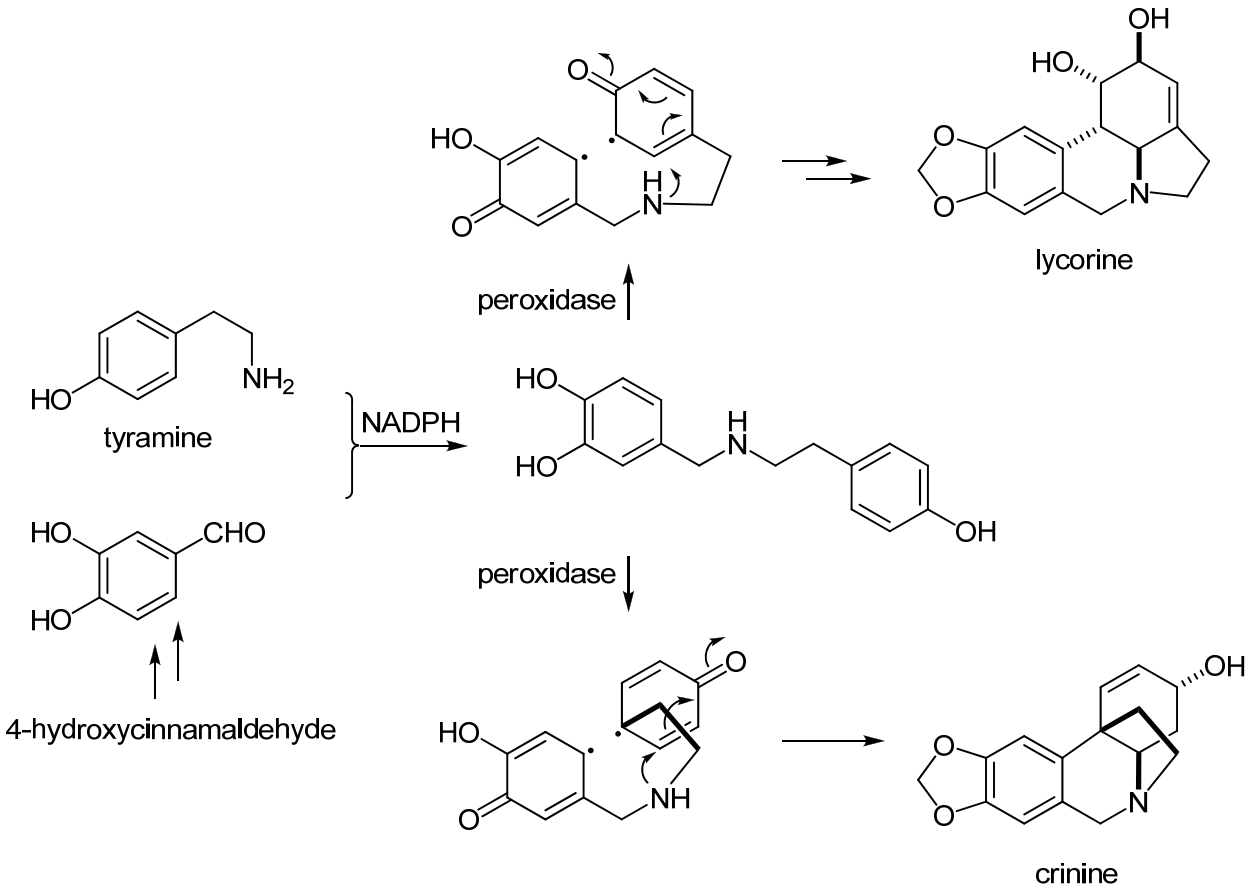


Scheme V.4.14

V.4.4. Amaryllidacea Alkaloids.

So far, we have seen families of alkaloids that are derived from tyrosine and either a Ar-C_2 (section V.3.2) or Ar-C_3 (section V.3.3) shikimate. This section dicusses a small family of alkaloids that is derived from tyrosine coupled to an Ar-C_1 shikimic unit. The latter is derived from phenylalanine and undergoes a reductive coupling with tyramine, the amine derived from

tyrosine by decarboxylation (Scheme V.4.15). Depending on the type of phenolic coupling that the pair of aromatic rings undergo, alkaloids like lycorine (found in dafodill flowers) and crinine (found in pancratium maritimum or sea daffodils) are produced.



Scheme V.4.15

V.4.5. Tryptophan Derived Alkaloids : The Indole Alkaloids.

Apart from some simple derivatives of tryptophan, e.g. indolylalkylamines, physostigmines, and ß-carbolines, the overwhelming majority of the indole alkaloids have as many carbons coming from other sources as it has from tryptophan. The indole moiety is invariably present and rarely modified. Both their significant neurophysiological action and the synthetic challenge they represent have captured the interest of chemists and biochemists.

In a few instances the indole nucleus is modified to an isoquinoline nucleus, for example in the cinchonine alkaloids. Figure V.4.5 shows the indole nucleus and some examples of indole

alkaloids. In many cases the aliphatic part of complex indole alkaloids is isoprenoid in nature and proof of that came with incorporation studies of mevalonate and geranyl pyrophosphate.

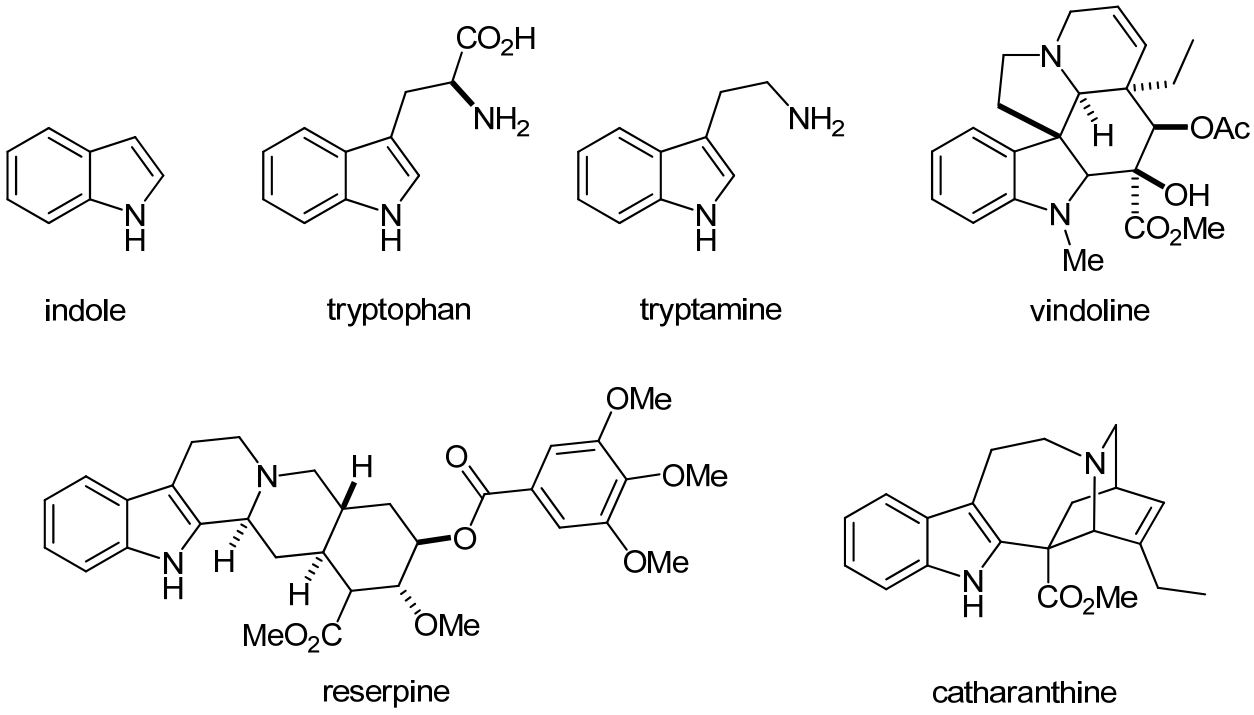
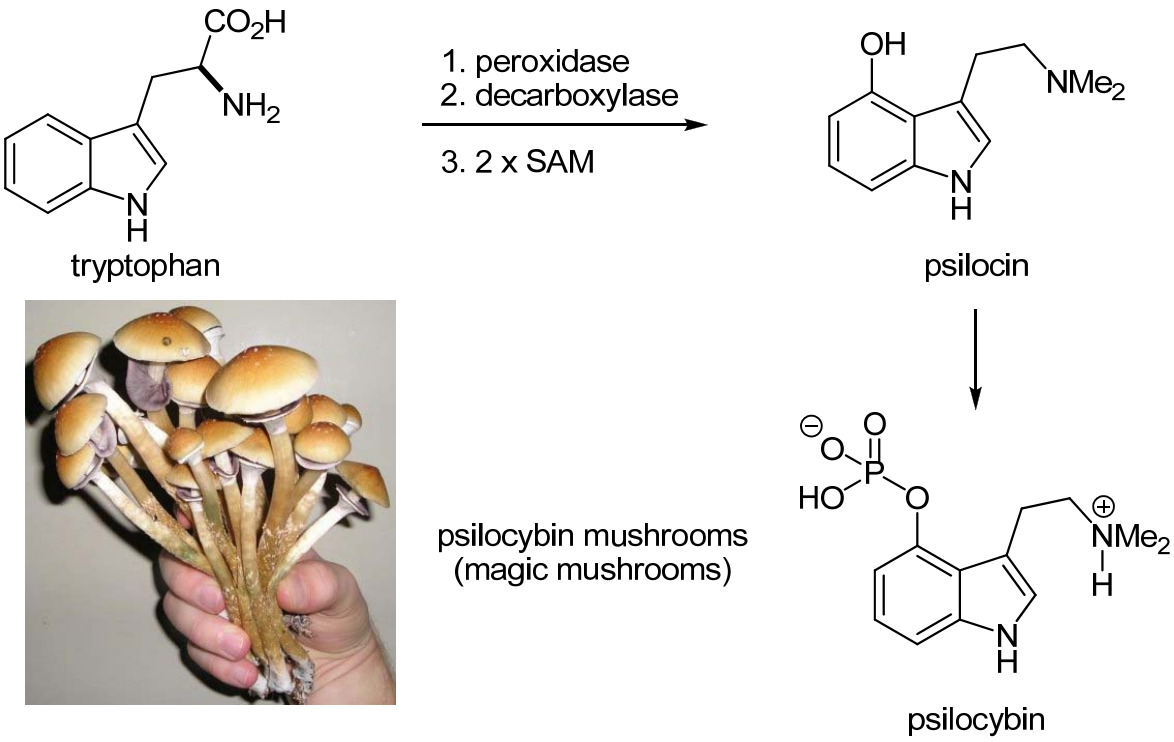


Figure V.4.5

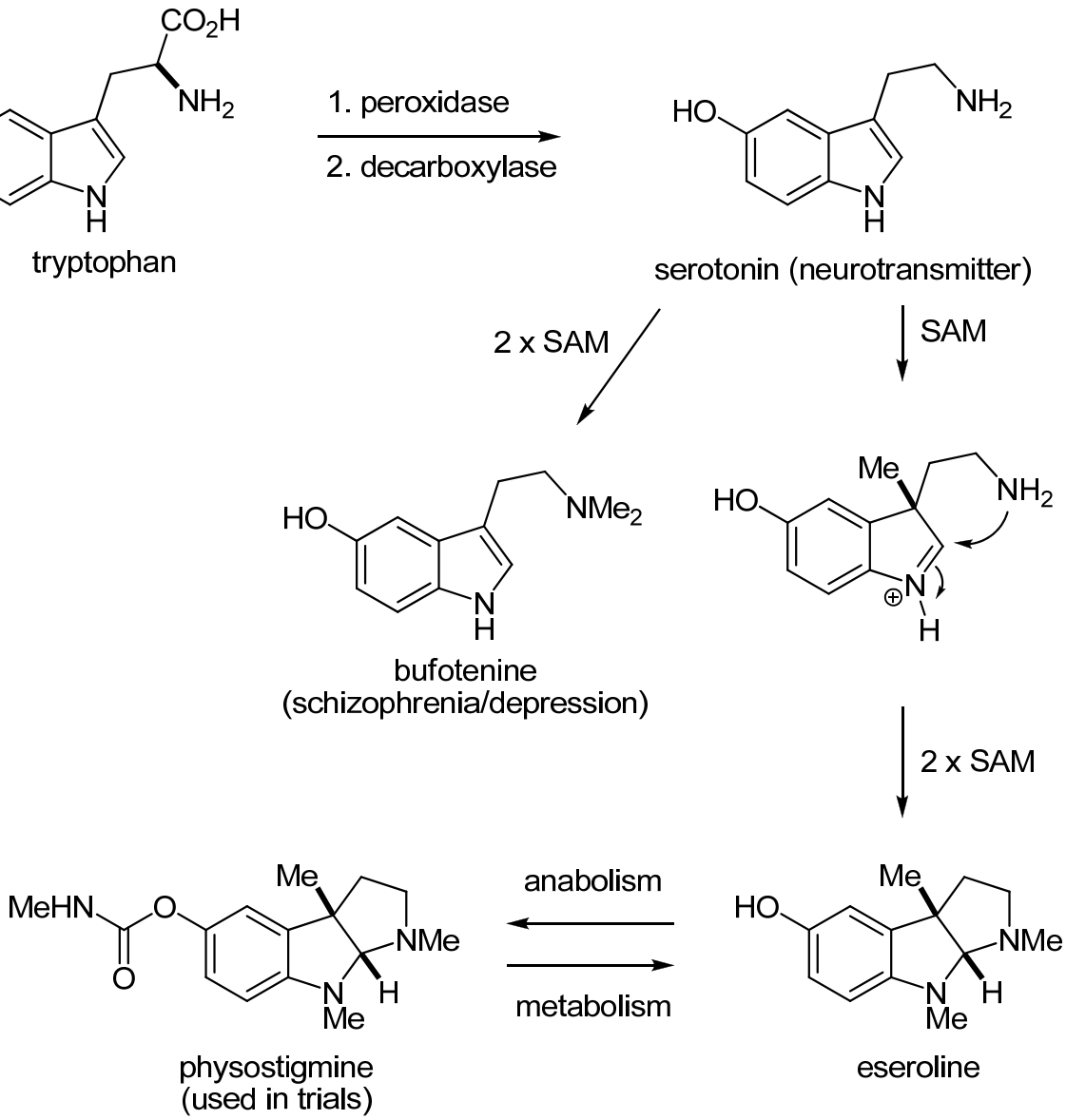
Schemes V.3.16 and V.3.17 depicts the biosynthesis of some simple tryptophan derived alkaloids which incorporate carbons from simple sources. Psilocin and psilocybin are the main hallucinogenic agents contained in "magic mushrooms" found in Mexico and South America (Scheme V.4.16). It was used by the ancient Maya's upper priesthood to gain spiritual contact with their gods.



Scheme V.4.16

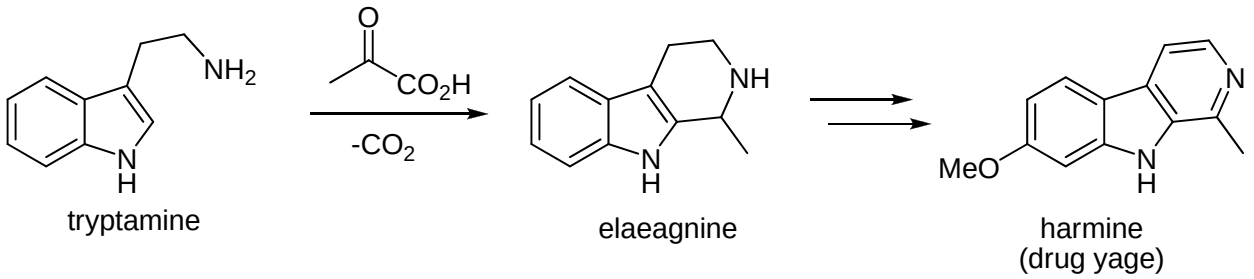
Bufotenine is formed in much the same way as psilocin except that oxidation takes place at C_5 (Scheme V.4.16). The former is found in the skin of many poison frogs along with other alkaloids (c.f. Figure V.6.3). Physostigmine (also called eserine) is a poison extracted from the African calabar bean (Figure V.4.6). The extracts from this poisonous bean was fed to defendants in Africa and in some ancient maritime trials (Figure V.4.6). It was believed to be a potent 'truth serum'.[3] Innocent people, confident of the verdict, would consume the beverage rapidly and the natural product would induce instant vomiting therefore ensuring the survival of the defendant. Survivors were therefore pronounced innocent. The guilty defendant, however, afraid of the verdict, would reluctantly and slowly "sip" the concoction which caused the active compound to be absorbed in the absence of vomiting. The defendant would quickly die of poisoning and be pronounced guilty.

Figure V.4.6



Scheme V.4.17

The formal acetaldehyde molecule contained in simple β-carboline alkaloids in reality comes from pyruvic acid, which after condensation (Pictet-Spengler) suffers a decarboxylation to give for example in elaeagnine and harmine (Scheme V.4.18). Other alkaloids may include more complex aldehydes and will be the subject of following sections. Further oxidation of the piperidine and aromatic ring gives harmine, the main constituent of the drug *yage*, a hallucinogen used by some Amazonian tribes. Burroughs and Ginsberg's voyage to the amazone have been chronicled in a book called the 'Yage Letters' first published in 1963. The drug was said to lend telepathic abilities to the user.



SchemeV.3.18

Having determined that the non-tryptamine part of numerous indole alkaloids was isoprenoid in origin, the exact intermediate by which the isoprenoid part was brought in was still unclear for many years. Eventually, it was found that loganin (a monoterpene belonging to the iridane family, see section II.3.1) was the precursor of the isoprenoid chain and that after oxidation of this compound into secologanin (Figure V.4.7) the latter is directly condensed by a tryptamine of the appropriate alkaloids.

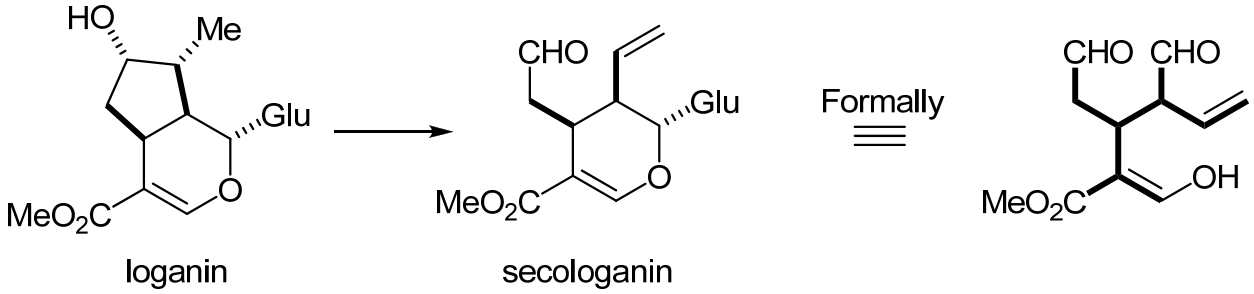


Figure V.4.7

Depending on whether this secologanin moiety suffers rearrangement or not, different types of alkaloids ensue (Figure V.4.8). There are three general types of ‘loganin skeleton’, namely the corynanthe type, which has not suffer any rearrangement from secologanin, the aspidosperma and the iboga types, which have underwent rearrangement has shown in Figure V.4.6. Each type of skeleton is flanked with one or two examples of natural akaloids possessing this particular type of skeleton.

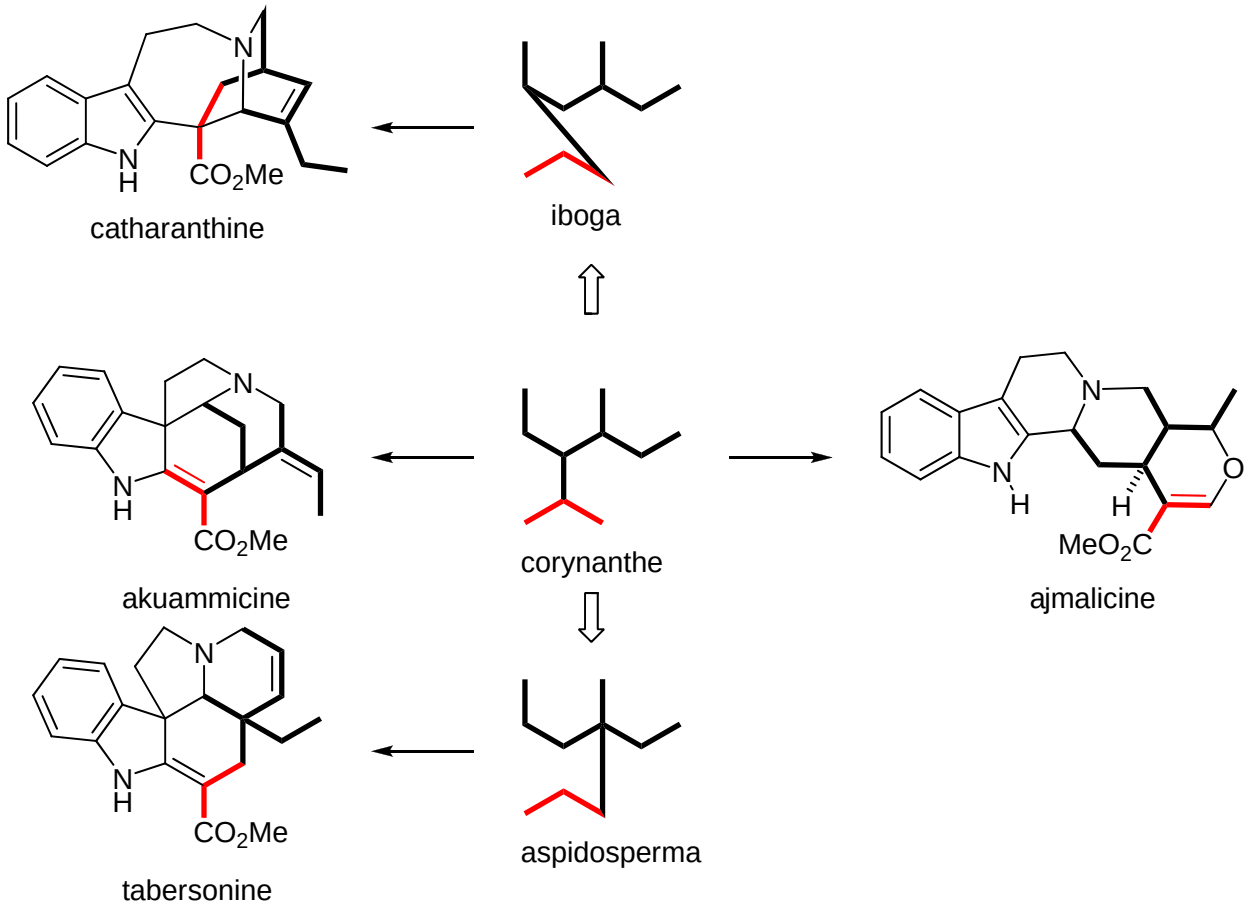
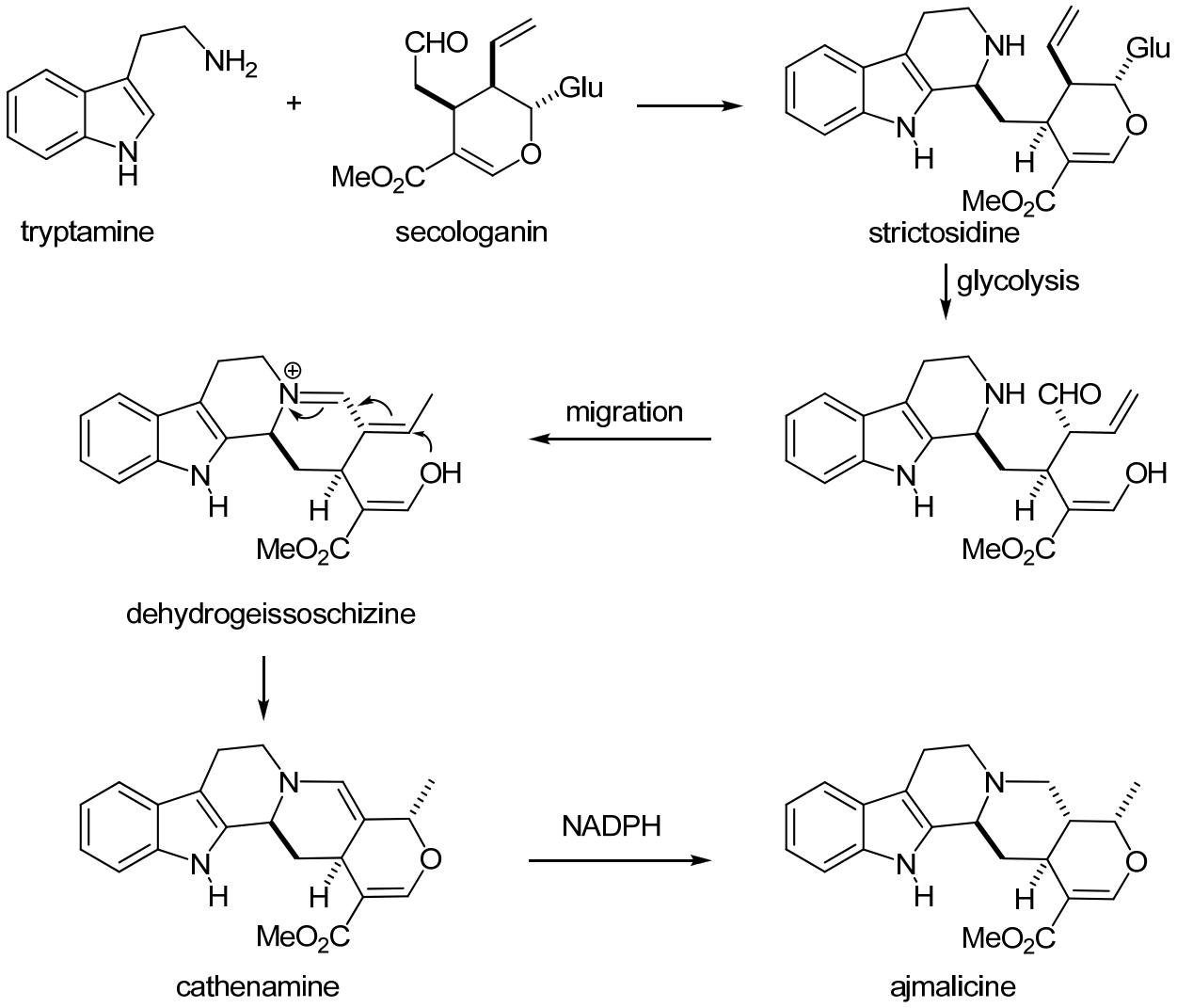


Figure V.4.8

Loganin is condensed and cyclized to give the now familiar six-membered heterocyclic ring (Pictet-Spengler). Then the tryptamine-loganin condensed intermediate (also a natural product known by the name **strictosidine** or also **isovincoside**) suffers a number of interesting reactions (Scheme V.4.19). Firstly the glucose moiety attached to loganin is hydrolyzed (glycolysis) and the double bond migrates to the exocyclic position. The aldehyde formed in the glycolysis condenses with the amine of the heterocyclic six-membered ring to give an immonium ion. This cation serves as electrophile to the enol of the ß-aldehydo-ester and cyclization to form **cathenamine** thus occurs. Further reduction of the double bond gives **ajmalicine**. The first reactions in this scheme i.e. the condensation of loganin with tryptamine and further cyclizations occur quite often in the biosynthesis of indole alkaloids.



Scheme V.4.19

Figure V.4.9 contains examples of alkaloids made up of tryptophan and seco-loganin. The loganin part of the molecule is highlighted to indicate the origin of the carbons. The biogenesis of each of these alkaloids can be derived from different rearrangements of the seco-loganin portion of the molecule. Yohimbine is an adrenergic blocking agent used in treatments of arteriosclerosis. It is also aphrodisiac. Vindoline does not have a known physiological activity. Only strychnine lost one carbon from the 10 original seco-loganin carbons. Strychnine is one potent poison that blocks inhibitory controls of motor neurons. It creates convulsion, paralysis, and ultimately death.

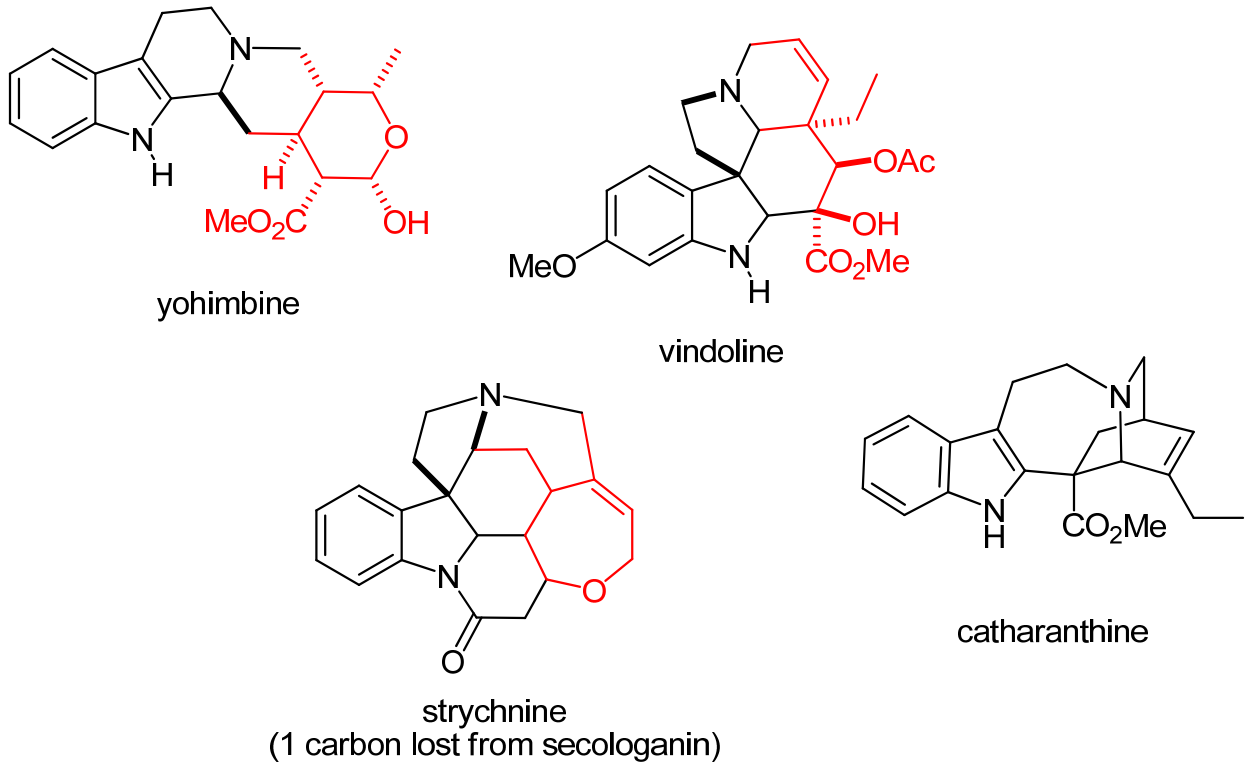
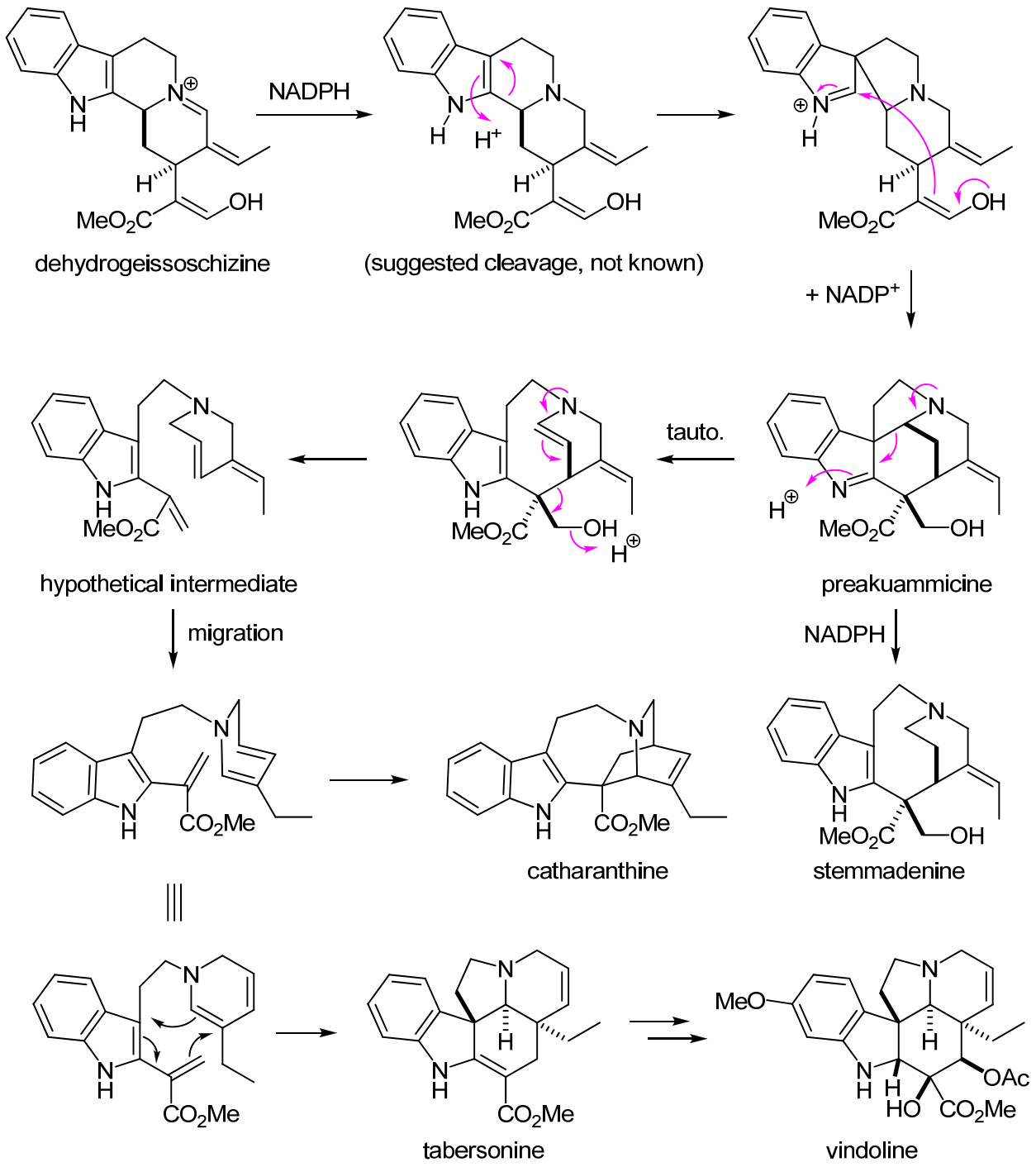


Figure V.4.9

From the same intermediate dehydrogeissoschizine, another biosynthetic route leads to catharanthine, tabersonine, and vindoline (Scheme V.4.20). There are three key carbon-carbon bond cleavage, the first one is not well known and we suggest an oxidation-assisted migration of the ring from position 2 to position 3 of the indole. After the enol cyclized on the indole-iminium ion as shown, there is the second cleavage assisted by the lone pair on nitrogen followed by a tautomerisation to the enamine. Then, the resulting enamine causes the third carbon-carbon bond cleavage promoted by the departure of water.

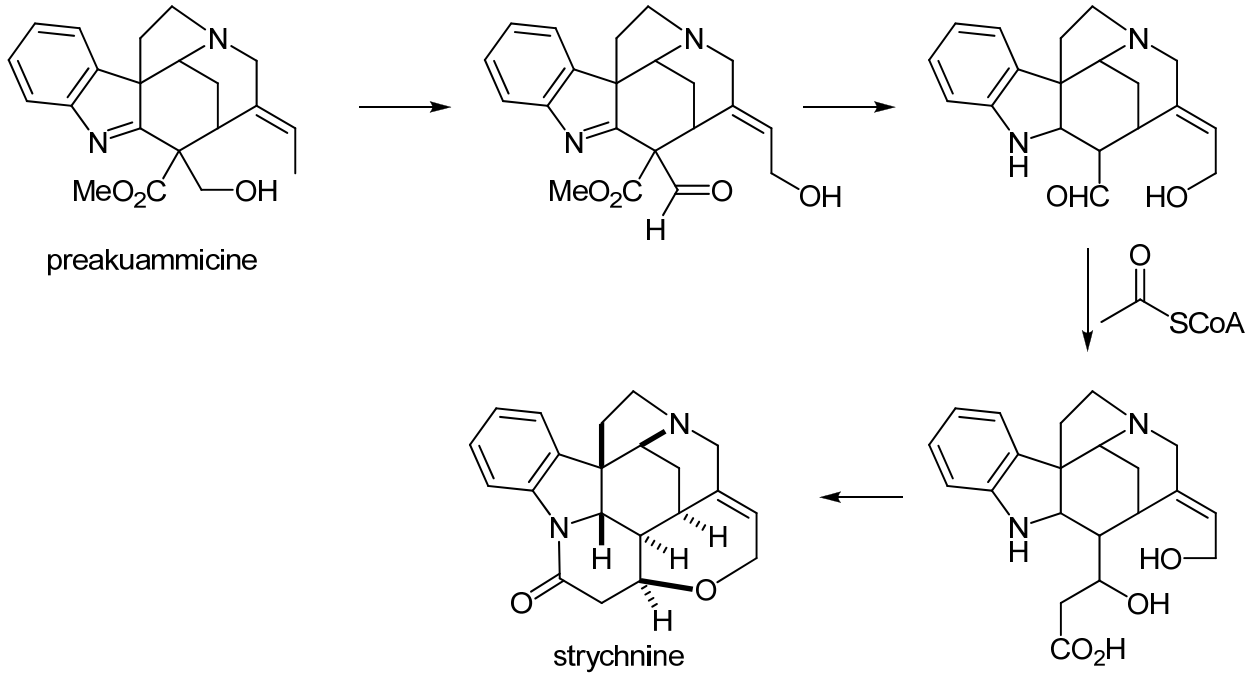
From there, two different formal 'Diels-Alder' cycloaddition may take place. Note that no 'Diels-Alderase' enzyme has been unambiguously identified therefore this proposal remains a formalism for the moment. Nonetheless, one of these formal cycloadditions leads to catharanthine (iboga type skeleton) while the other leads to tabersonine and vindoline (aspidosperma type)



Scheme V.4.20

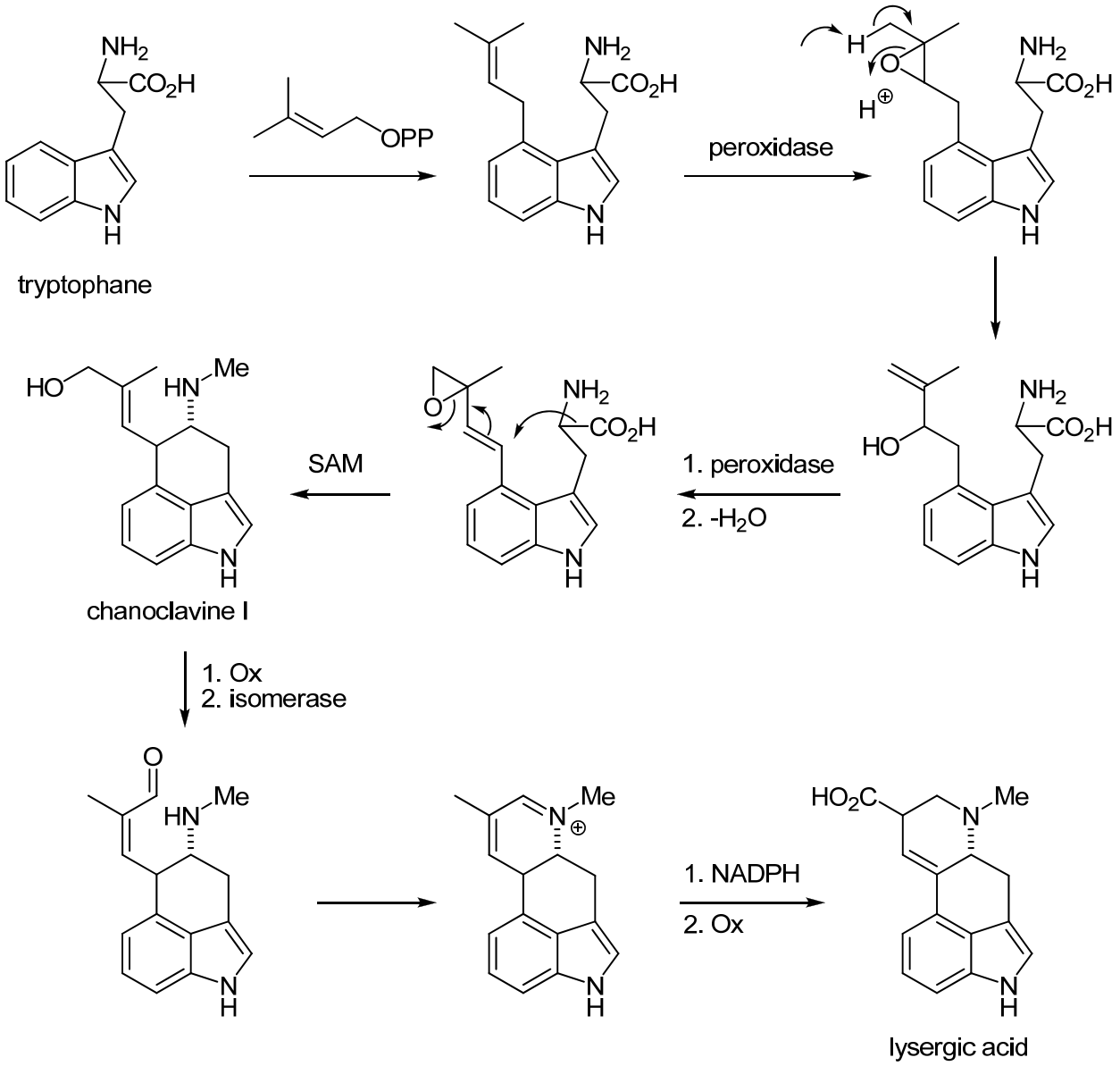
Strychnine is a highly toxic alkaloid from the strychnine family and it possesses the corynanthe carbon skeleton. Preakuammicine is its precursor, the same intermediate seen in Scheme V.4.20 in the synthesis of catharantine and tubersonine. However, in the case of

strychnine, one carbon has been lost in the form of carbon dioxide while two carbons were added in the form of acetyl CoA (Scheme V.4.21).



Scheme V.4.21

Lysergic acid is a tryptophan derived alkaloid plus a C_5 terpenoid unit. It is a member of the so-called ergot alkaloids (ergot = fungus on rye grain) which caused generalized gangrene (vasoconstriction), convulsions, and neural illness for those people who ate infected rye bread. The illness known as St. Anthony's fire which gave burning sensations and blackened limbs (gangrene) was caused by the same ergot alkaloids. The biosynthesis of lysergic acid is shown in Scheme V.4.22. A dimethylallyl pyrophosphate unit is brought in and probably undergoes a biogenetic equivalent of the Friedel-Craft reaction with the aromatic ring. Further oxidation of the isoprenyl chain gives the alkene-epoxide, which undergoes cyclization with concomitant loss of CO_2 to afford chanoclavine I, an ergot alkaloid. Further oxidation of the isoprenyl chain and cyclization give the third six-membered ring and leads to lysergic acid as shown. The diethylamide derivative of lysergic acid gives the well known hallucinogen lysergic acid diethylamide (LSD).



Scheme V.4.22

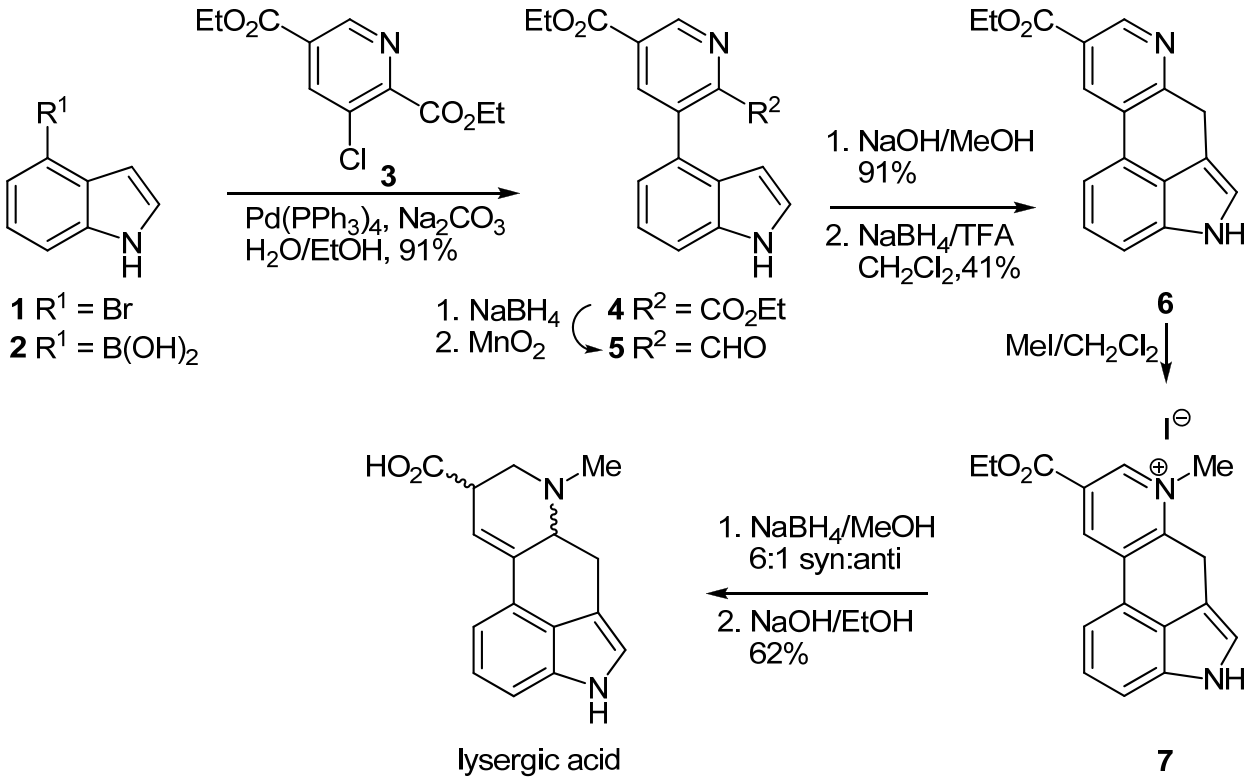
V.5. Chemistry of alkaloids.

V.5.1. Synthesis of Lysergic Acid.

Alkaloids are sometimes abundant in plants and can be isolated in large quantities. For example, morphine is easily isolated from opium where it is found in up to 21% in dry weight in the seeds of *Papaver somniferum*. However, other alkaloids are found in trace quantities only

and thus total synthesis becomes the only way to obtain sufficient quantities for pharmacological testing. But the prime reason for organic chemists to embark in the total synthesis of alkaloids is no doubt the immense challenge their complex structure has created. Many new reactions and strategies were developed in total syntheses of natural or unnatural alkaloids. We therefore shall study a few of these syntheses in some detail.

One recent synthesis of lysergic acid was performed by Hendrickson and Wang in 2004.[4] Scheme V.5.1 depicts their approach. In a key step, the 4-boronic acid indole **1** was coupled with the chloride **3**. Reduction/oxidation of the more reactive *o*-ester (because of coordination of the reagent to the pyridine nitrogen) give the aldehyde **5**, which was then condensed in basic methanol to give the product of a Mannich reaction. The alcohol was dehydrated in acid and reduced *in situ* to **6**. Methylation give a very electrophilic nitrogen, prone to reduction with excess $NaBH_4$. Hydrolysis affords lysergic acid.

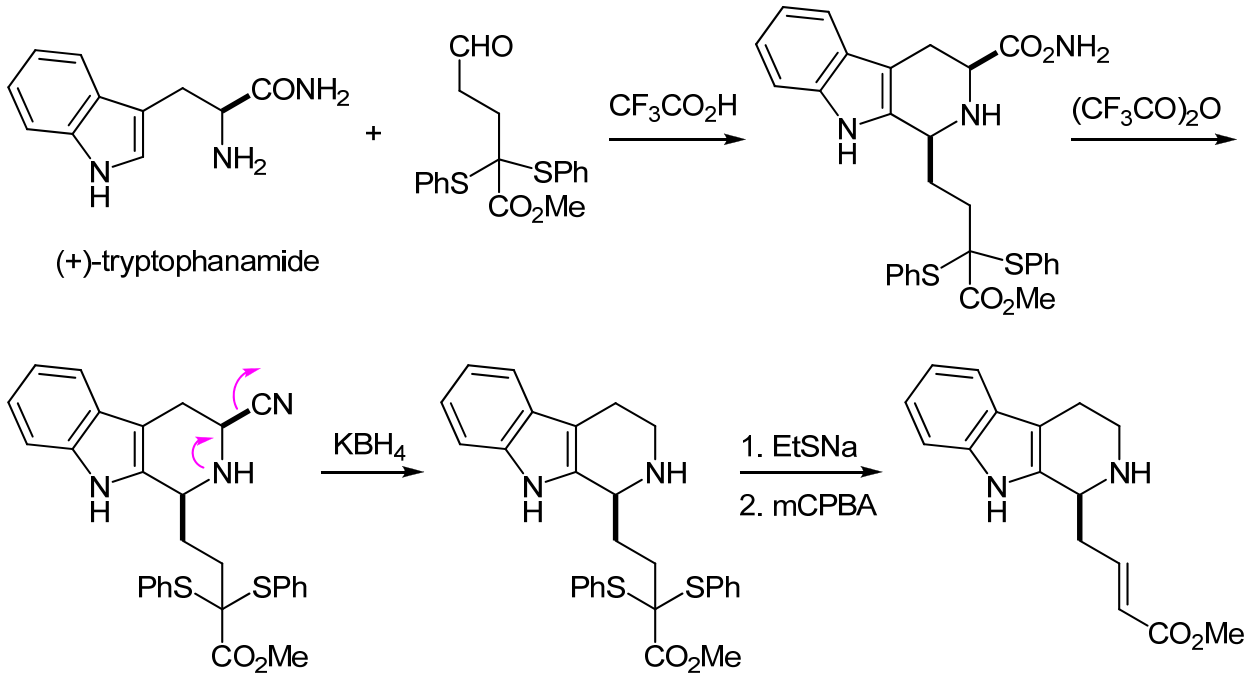


Scheme V.5.1

V.5.2. Synthesis of Ajmalicine.

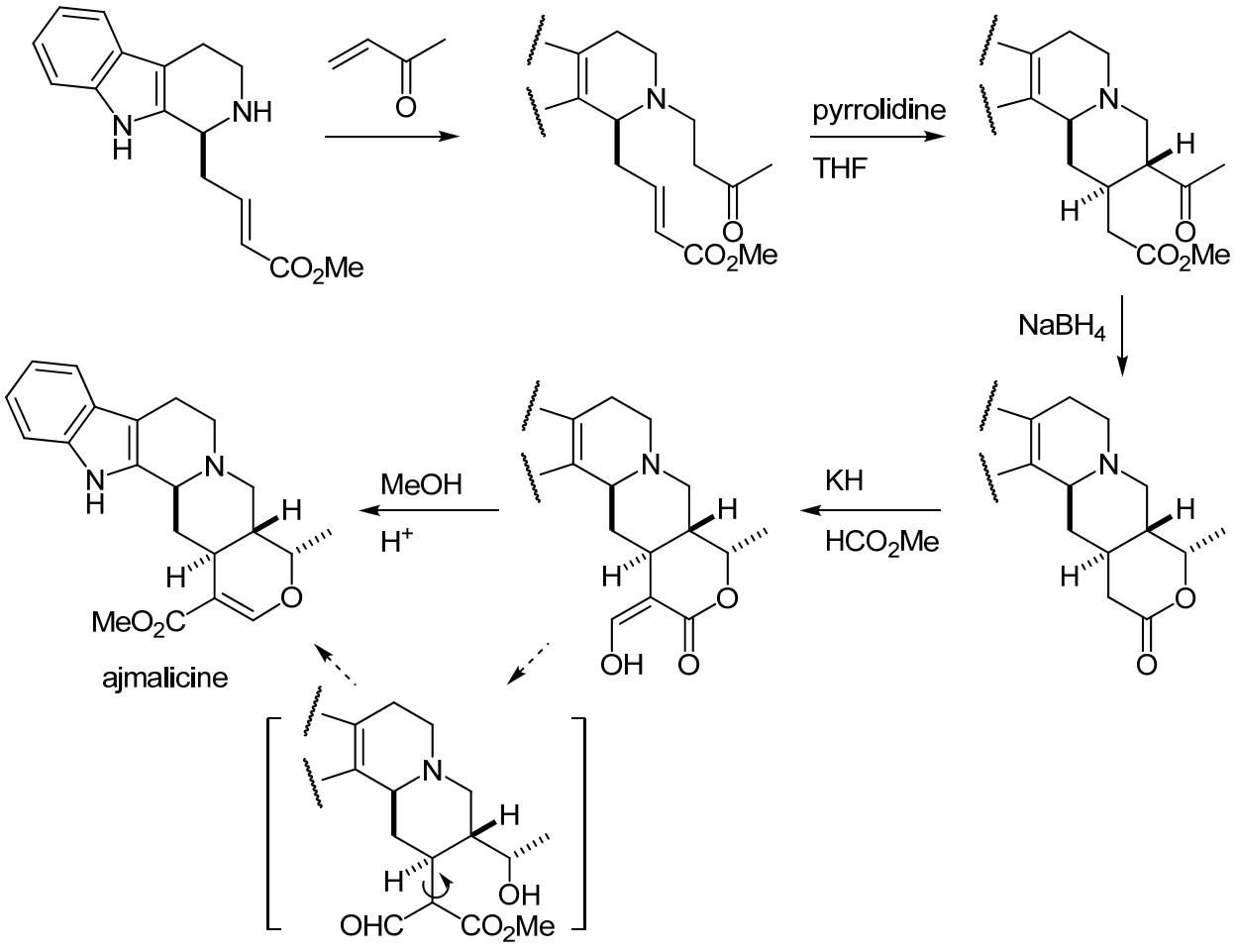
Schemes V.5.2 and V.5.3 show the total synthesis of ajmalicine. The starting material is derived from tryptophan where the acid has been converted to the primary amide. Amides are much less nucleophilic than amines because the free pair of electrons on the amide are in resonance with the carbonyl thereby putting a partial positive charge on the nitrogen. Thus the amine is condensed preferentially with an aldehyde to give the expected imine. Trifluoroacetic acid, a very strong organic acid, is used to trigger cyclization to form the *N*-heterocycle 6-membered ring. Note that this overall transformation is called the Pictet-Spengler reaction sequence which is the same as we saw in the biosynthesis of alkaloids. The cyclization proceeded to give the thermodynamically more stable isomer in which both groups can be pseudo-equatorial. Now that the amide has served in directing the cyclization stereochemistry, it is removed by first dehydrating the amide to a cyano group with trifluoroacetic anhydride followed by reductive elimination with potassium borohydride. The mechanism for this reductive elimination proceeds via the iminium ion because cyanide may be expelled by the lone pair on the nitrogen. The iminium is immediately reduced by the hydride molecule.

To begin the formation of the other two rings of ajmalicine, the side chain is functionalized appropriately. Thus one of the sulfide groups is removed using sodium ethanethiol and base. The sulfur anion attacks one of the sulfide at the sulfur atom to form the disulfide PhS-SEt and the anion next to the ester group on the chain. This latter anion being more stable than EtS^-, the equilibrium is shifted to the right. Then the remaining sulfide is oxidized and undergoes a *cis*-elimination upon heating. The elimination is *cis* but the two groups on the new double bond end up in the more stable *trans* geometry.



Scheme V.5.2

Finally the other side chain is brought in using a Michael addition (1,4-) of the piperidine nitrogen and methyl vinyl ketone (Scheme V.5.3). The resulting molecule is cyclized in yet another Michael reaction. Note that the Claisen condensation would afford an 8-membered ring instead of the 6-membered ring that the Michael addition reaction yields. Because the latter is more easily formed, the Claisen condensation does not compete. After reducing the ketone chemoselectively (sodium borohydride is incapable of reducing esters), the alcoholate intermediate internally cyclizes to give the lactone. Next, the formate group is introduced in a Claisen condensation. This formate group is stabilized via internal hydrogen-bonding. Acid in hot methanol is capable of opening the lactone ring. The resulting ß-formyl ester can rotate and the alcohol recyclizes with the aldehyde under these acidic conditions. All these reactions being in equilibrium, the cyclization with the aldehyde is favored because the resulting lactol eliminates water and thus forms the desired product ajmalicine. The latter is not rehydrated under the reaction conditions and is thus stable and accumulates.



Scheme V.5.3

V.5.3. Synthesis and Chemistry of Pyridine, Quinoline and Isoquinoline Alkaloids.

As we have mentioned earlier, the 6-membered ring *N*-heteroaromatic compounds are all basic in nature because they possess a lone pair of electrons that is not involved in the aromaticity. In addition, the *N*-heteroaromatic rings are electron deficient compared to their carbon analogues benzene and naphthalene. This is due to the electron induction from the nitrogen atom which can be clearly demonstrated by resonance forms (Figure V.5.1). This phenomenon is seen in the ^{1}H and ^{13}C nmr of pyridines, quinolines, and isoquinolines. For example, nicotine has the ^{1}H nmr shown in Figure V.5.1. Note that the protons in the α-positions are shifted downfield relative to benzene protons. This reflects their partial positive charge or lower electron density. The γ-proton is also shifted upfield, but less since it is further away from the nitrogen. Finally the β-hydrogen is more or less at the expected position for benzene protons. Indeed it does not bear a partial positive charge like the other hydrogens. In general the α-

hydrogens of pyridine and analogous rings will appear in the 8.5 ppm range, γ-hydrogens in the 7.3-7.75 ppm range while the β-protons will be in the usual 7-7.5 ppm range.

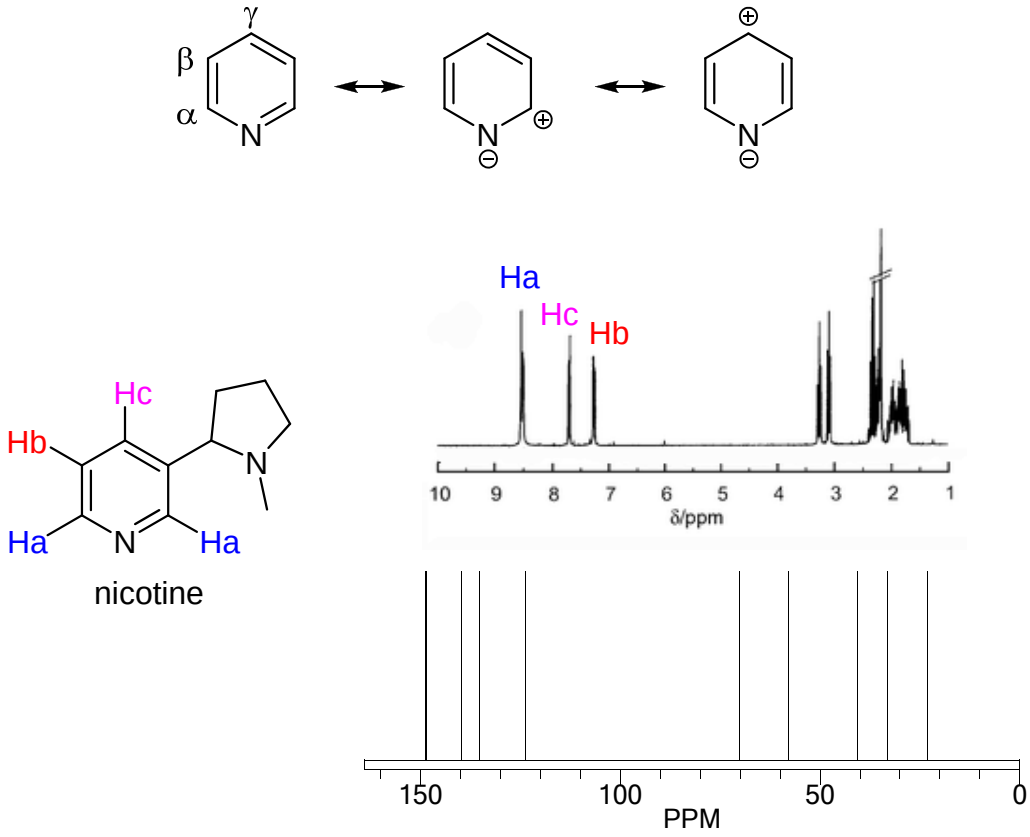
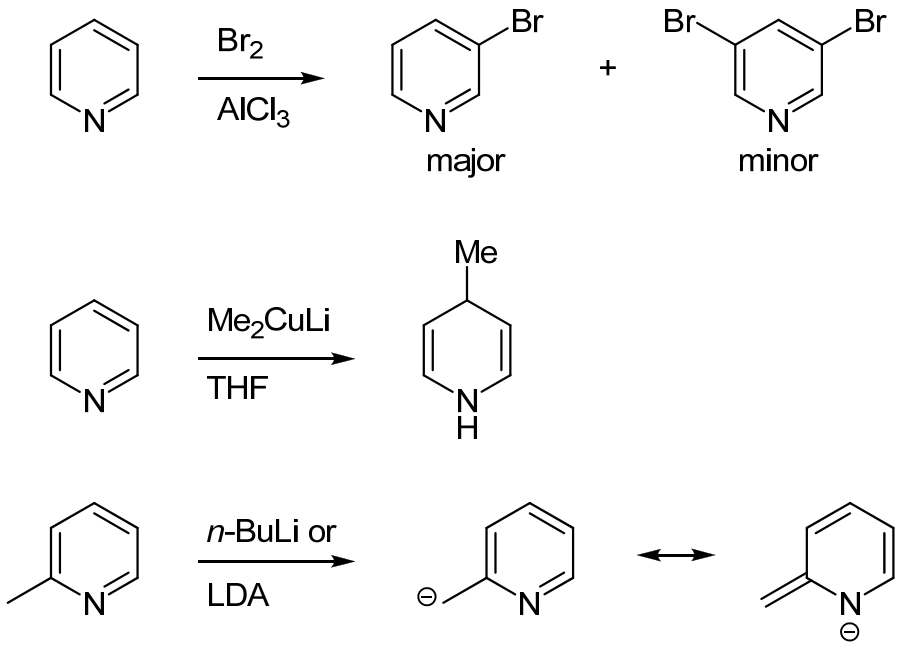


Figure V.5.1

By analogy, the carbon resonance of the α-carbons will be most deshielded in the 150 ppm range (Figure V.5.1). The γ-carbon will be found at 138 ppm and the ß-carbons in the 120-125 ppm range. Note that the ß-carbon of nicotine bearing the pyrrolidine ring is found to be more deshielded because of that group (~ 135 ppm).

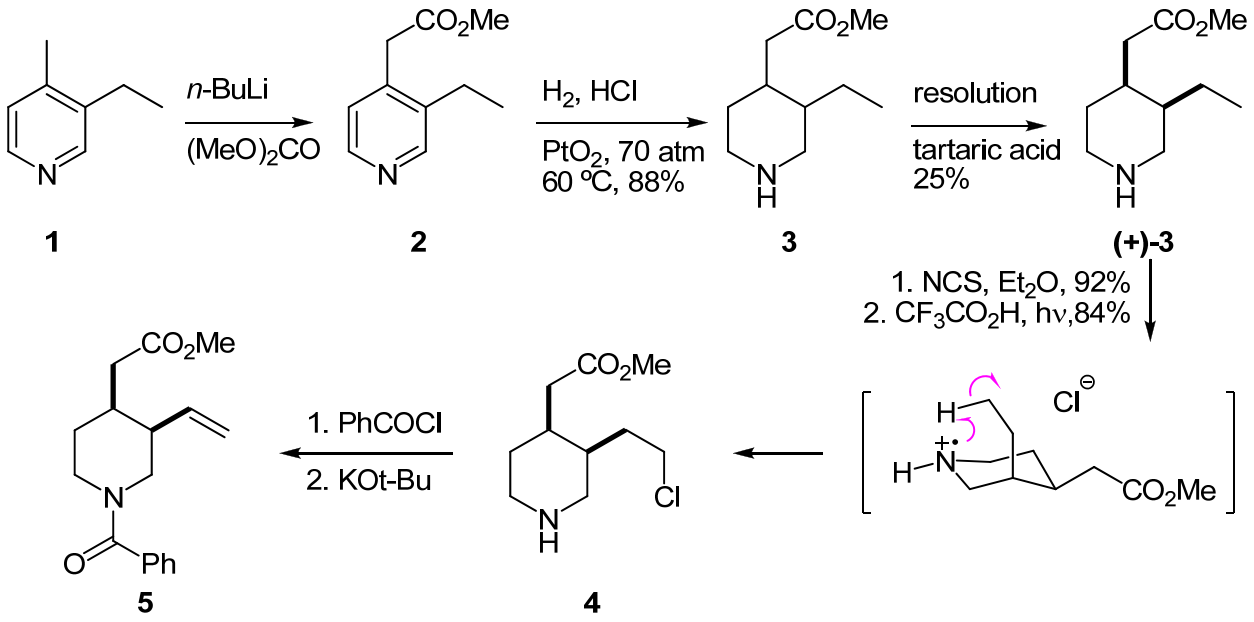
Because the pyridine ring is more electron deficient than benzene it will be more difficult to perform an electrophilic attack on it. For example bromination is slower and proceeds only to the mono-bromide (Scheme V.5.4). In contrast, phenols are electron-rich and are tribrominated quite easily. On the other hand, the pyridine ring, being deficient in electrons, is susceptible to nucleophilic attack. For example alkyl cuprates add to the pyridine ring at the 4-position to give dihydropyridines. Also, protons at benzylic positions on pyridine rings can be removed by a strong base and alkylated or acylated. The anion is 1800 times more stable at the 4-position than

the 2-position which is itself 130 times more stable than the 3-position. This can be understood from the resonance structures of such anions.



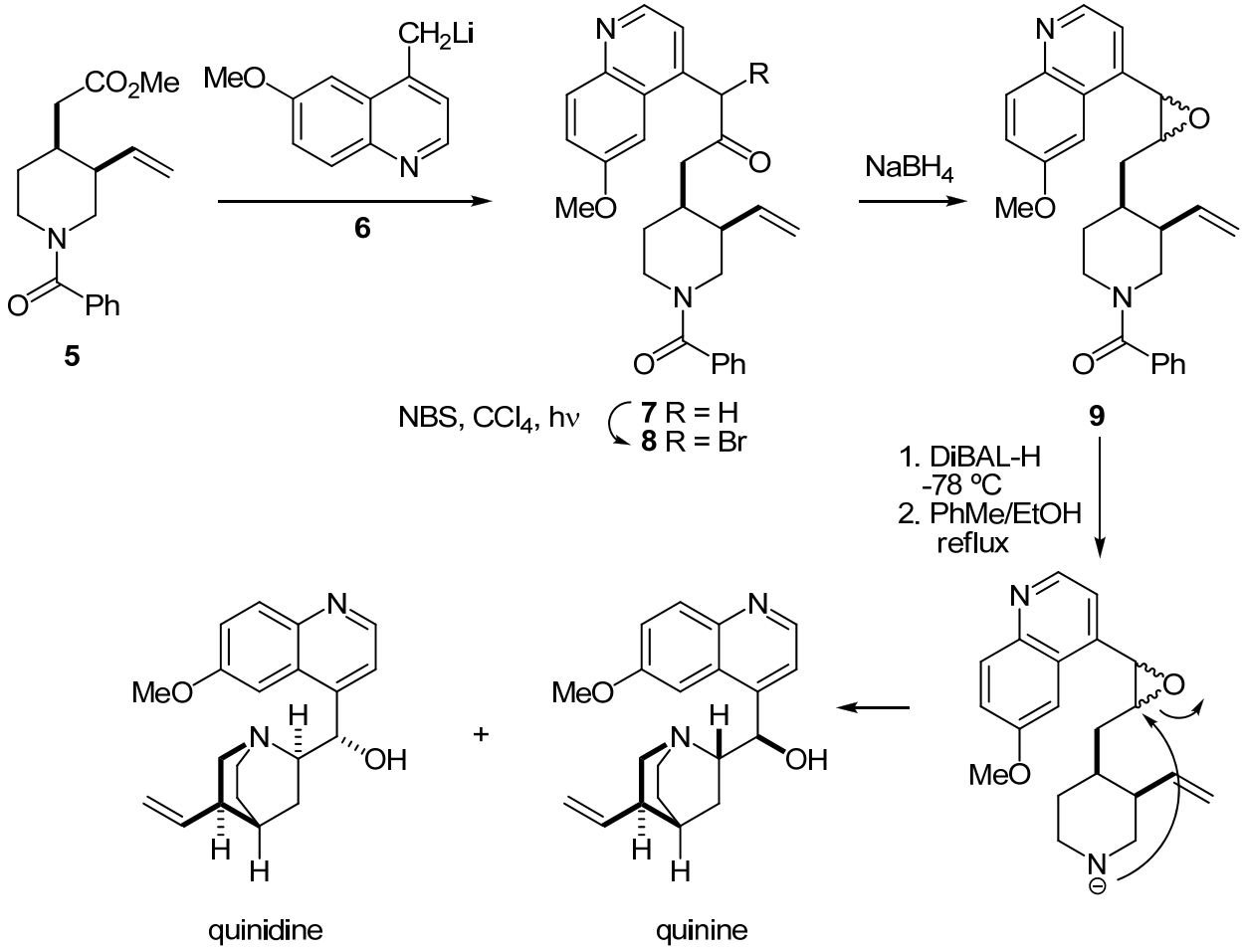
Scheme V.5.4

One synthesis of quinine, an isoquinoline alkaloid, demonstrates some of that chemistry (Scheme V.5.5). This synthesis was achieved at Hoffmann-Laroche Pharmaceutical Company and can be found in *J. Am. Chem. Soc.* **1970**,*92*, 203 and **1971**, *93*, 5902. The synthesis starts with 3-ethyl-4-methylpyridine **1**. The anion of the methyl at the 4-position is preferentially formed and reacted with diethyl carbonate to give the ester **2**. The ring is then reduced completely to the piperidine ring with hydrogen and a platinum catalyst to give **3**. The molecule is of course racemic at this point and is resolved to an optically pure compound. *N*-Chlorosuccinimide is used to chlorinate the nitrogen which undergoes a photochemical "Hofmann-Löffler-Freytag" rearrangement. The reactive conformation depicted in Scheme V.5.4 shows that hydrogen abstraction is poised to occur at the terminal methyl group. The amine is protected again as its benzoyl amide. Elimination of chlorine with potassium *t*-butoxide to give molecule **5** which will form the bicyclo upper right portion of quinine.



Scheme V.5.5

The lower portion of quinine is quite easily accessible from the commercial 7-methoxy-4-methylisoquinoline **6** (Scheme V.5.6). Lithium diisopropylamide deprotonates the methyl and the resulting anion can be condensed with the ester in molecule **5**. The resulting ketone **7** is photochemically brominated with NBS to **8** and reduced with diisobutylaluminium hydride to give a 1:1 mixture of epoxides. Then it is heated in a mixture of toluene and ethanol causing an internal cyclization of the nitrogen onto the epoxide. The two isomeric alcohols are quinine and quinidine, respectively and were separated.



Scheme V.5.6

V.6. Naturally Occurring Toxins.

We will conclude our studies of natural products by looking at some **natural toxins**, both alkaloidal and others. Table V.5.1 gives some examples of natural alkaloidal poisons together with their biological activities and uses.

Table V.5.1. Toxicity, biological activity and uses of alkaloids.

Amphetamine:	Stimulates the adrenergic response.
Atropine:	Natural insect repellant. Used in eye surgery as anaesthetic
Cocaine:	Isolated from coca leaves from the Andes. Chewed by ~8 million South American Indians to alleviate digestion problems and other ailments. Not addictive when chewed, however intravenous pure cocaine is addictive. It is a narcotic but used as local anaesthetic in ear, nose, throat, and mouth surgery.
Coniine:	Acute poison. Used by ancient Greeks for State executions. Socrates is believed to have been killed by an injection of extracts of Hemlock containing coniine. Acetate derived metabolite, not from amino acid.
Cytisine:	Toxic substance much like strychnine, used as rat poison.
Harmine:	Principal constituent of the drug Yage (Native Indians)
Lycopodine:	First alkaloid to ever be isolated and characterized. Its structure was not elucidated before the mid-20th century.
Mescaline:	Hallucinogenic component of the cactus extract *Peyote* used by North American Indian (Southern US) in religious ceremonies. Mescaline is not addictive and thus not a narcotic. Humans develop a tolerance to it as well as a cross tolerance for lysergic acid diethylamide (LSD).
Morphine:	Main constituent of Opium (see below). Hallucinogenic, narcotic, and pain killer. Used in cocktails for terminally ill cancer patients.
Nicotine:	Addictive stimulant found in tobbaco leaves. Increases heart rate and blood pressure.
Opium:	Extract from the outer shell of unripe *Papaver somniferum* seeds. Contains over 40 alkaloids with morphine (21%), noscapine (8%), papaverine (1%), codeine (1%), and thebaine (0.5%) being the most important and active. Often used in Ancient China for sensation of well-being.
Papaverine:	Found in opium. It is a smooth muscle relaxant as well as a cerebral vasodilator. Used in treatment of angina. Not a narcotic by itself.
Pelletierine:	Anthelmintic used to treat tapeworm and other parasites.
Physostigmine:	Acute poison used in ancient trials. Rapid ingestion causes vomiting and survival of defendant who would be pronounced innocent. Slow ingestion causes absorption and death of the defendant.
Piperine:	Bitter and irritant constituent of black pepper. Gives impressions of burn and is responsible for the "hot" taste of pepper.
Psilocine:	Main hallucinogenic constituent of the "magic mushrooms" used by ancient Mayas to reach spiritual contact with their gods.

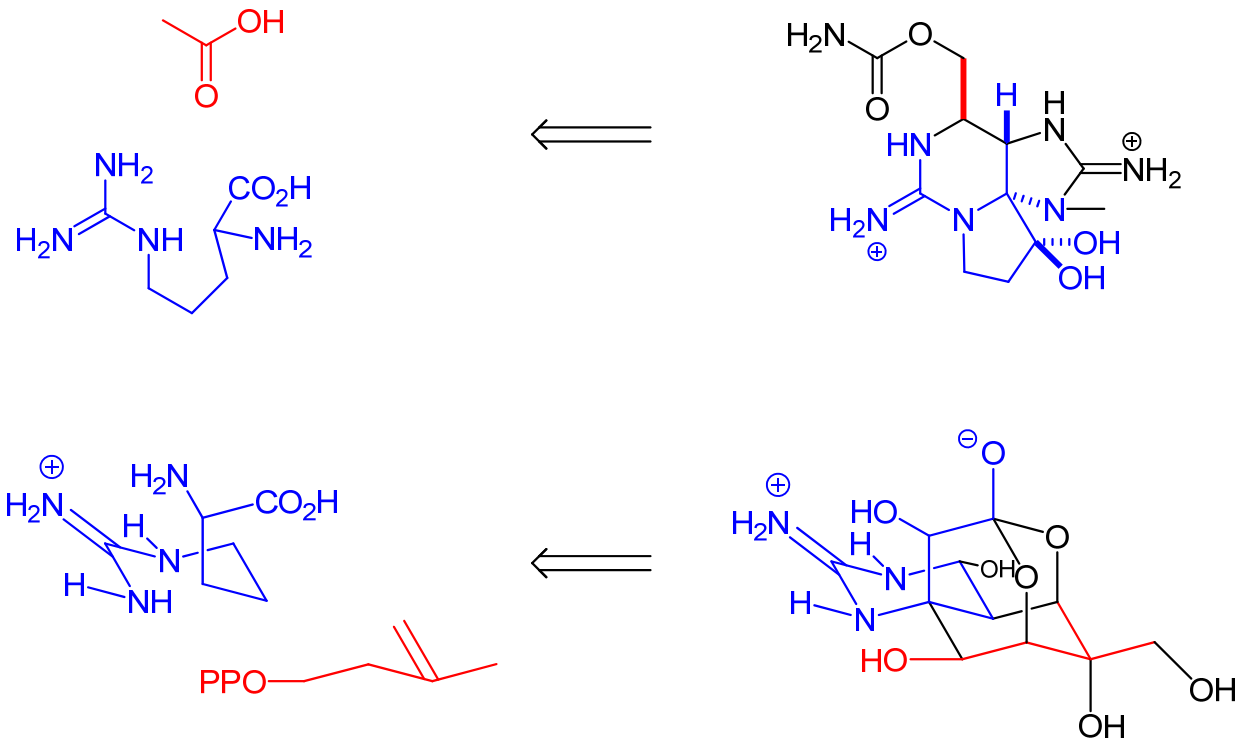
V.6.1. Fungi and algae toxins.

Figure V.6.1 gives the structures of the various toxins found in some terrestrial fungi and some marine algae. Aflatoxin is a polyketide derived metabolite that has suffered extensive rearrangements. It is found in a fungus *Aspergillus flavus* and is extremely toxic and carcinogenic. The toxicity may be due to an inhibition of DNA synthesis either by DNA-synthetase enzyme inhibition or toxin-DNA template reaction where the toxins irreversibly and chemically react with DNA molecules.

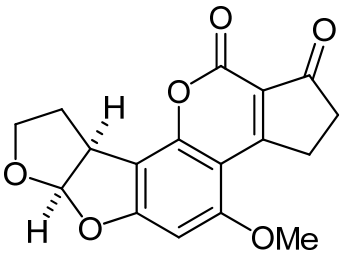
Aflatoxin is acetogenic in origin, though extensive rearrangement has occurred. It is a toxic and carcinogenic substance produced by the fungi *Aspergillus flavus*. It works by inhibiting DNA synthesis. The acetogenin **saxitoxin** is one of the toxin responsible for the "red tide" in certain region of the Atlantic and Pacific Ocean near the North American Coastal lines (Figure V.6.1). The toxin is produced by algae but is rapidly concentrated in certain shellfish e.g. mussels, oysters or others. The symptoms produced by ingestion of the contaminated fish are numbness of the extremities, a lack of muscular coordination which leads to respiratory problems and ultimately death if not treated.

A similar toxin, **tetrodotoxin**, is found in the puffer fish or "Fugu" in Japanese. This fish is considered a delicacy in Japan and the preparation of the fish is absolutely crucial if the customers are to survive their meal! Indeed, this deadly poison resides in the ovaries, the liver, and the intestines of the fish and if by mistake these are cut open, the toxin may leak into the flesh of the fish. Cooking will not destroy the toxin and the customer ingesting the fish would almost certainly die of poisoning. Restaurants in Japan take great pride in serving this very expensive meal and the reputation of the chef is, needless to say, quite important...

The biosyntheses of saxitoxin and tetrodotoxin start with arginine (Scheme V.6.1). In the former case, a molecule of acetate is added and after extensive oxidation, leads to the natural product. Tetrodotoxin couples arginine with a molecule of IPP (the basic terpene unit). The carbon skeleton of the natual toxin also suffered extensive oxidation.

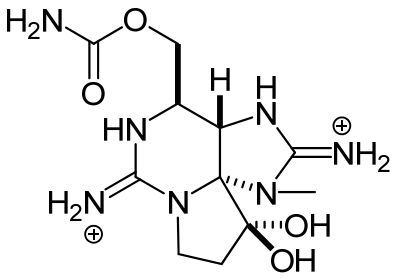

OH
O
NH2
H2N
NH
CO2H
NH2
H2N
O
O
HN
H
H
N
NH2
H2N
N
N
OH
OH
H2N
CO2H
H
N
H
NH
PPO
O
HO
H2N
H
N
OH
O
O
H
N
HO
OH
OH

Scheme V.6.1



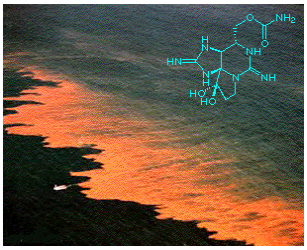
Aflatoxin

Aspergillus flavus (fungi)
Toxic and carcinogenic
LD_{50} 9.5 mg/kg
Acetate in origin with extensive rearrangement

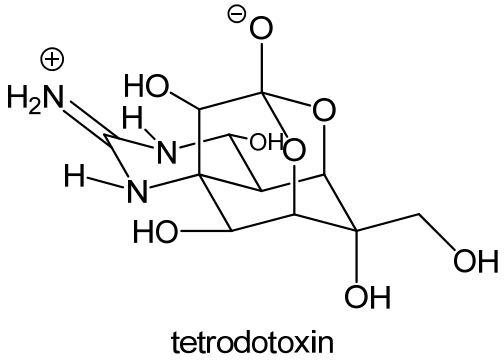


saxitoxin

Dinoflagellates
Toxins of **red tide**



dinoflagellate alexandrium



tetrodotoxin

Puffer fish or 'Fugu' in japanese
LD_{50} 10 μg / kg

Figure V.6.1

Certain species of mushrooms are not comestible because they contain potent toxins. Gyromitrin and muscarine are two examples (Figure V.6.2). The latter is hypotensive agent and acts at the postganglionic parasympathetic effector sites in the autonomous nervous system. More potent poisons are the polypeptides phalloidin and amanitin (100 g of mushroom tissue contains 10 mg phalloidin, 08 mg α-amanitin, and 05 mg β-amanitin). They are believed to be responsible for over 95% of all mushroom poisoning. They immediately cause violent emesis and diarrhea. They irreversibly destroy liver cells at a very low dosage (LD_{50} is 0.1 mg/kg).

Interestingly, *A. phalloides* contain an antitoxin, antanamide, that completely protects mice against the poison even at LD_{100} i.e. dosage of 0.5 mg/kg. The mechanisms of action are different for these toxins, phalloidin affecting the endoplasmic reticulum of the cells and amanitin affecting the cell nucleus. The latter has been demonstrated to inhibit RNA polymerase from nucleoplasm in vitro in liver cells. In vivo studies showed that it also inhibits the synthesis of ribosomal DNA.

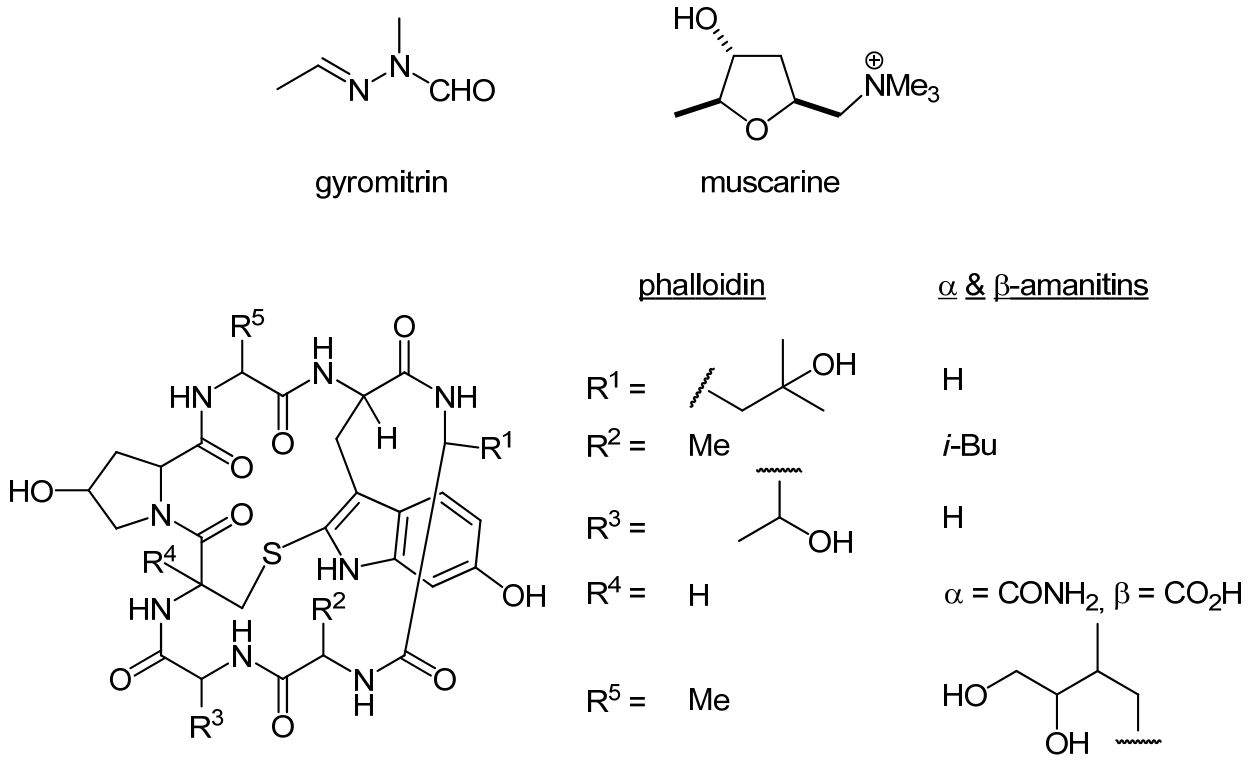


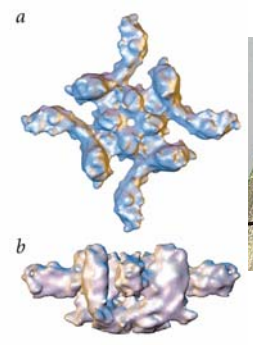
Figure V.6.2

V.6.2. Venoms.

There are two functional types of venoms: (1) those that bring about blood coagulation whether by direct action of fibrinogen or by converting prothrombin (a prostaglandin) to thrombin (a prostaglandin coagulant); (2) those that are neurotoxins and act on the central nervous system e.g. by inactivation of acetylcholine. Rattlesnake and moccasin venom are examples of (1) and cobra venom is an example of (2). Venomous is different than poisonous. A venomous organism may use its poison externally wheras a poisonous organism only contains a poisonous substance and may be harmless until it is ingested.

It is the enzymes, therefore proteins, of snake venoms that are thought to be the active toxic principles. In front-fanged snakes these include L-amino acid oxidase, protease, cholinesterase (responsible for the synthesis of acetylcholine neurotransmitter). Back-fanged snake venoms comprise the same types of enzymes and other types including fibrinolytic and bradykininogen enzymes. The latter causes tissue necrosis, respiratory distress and internal hemorrhaging. Black widow spiders have a potent neurotoxin as well, which causes intense pain and cramping of back, legs, abdomen, and chest. Muscles may convulse and cramp causing fatality in about 4% of cases. However, most venoms have antidotes available nowadays.

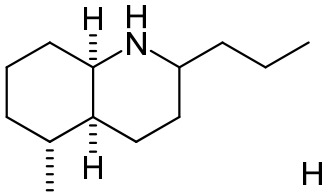
Dendrobitidae frogs (South American Frogs) secrete venoms from their skin to repel predators. These include pumiliotoxin C, histrionicotoxin, and gephyrotoxin (Figure V.6.3). The specific activity of these toxins is inhibition of neuromuscular cholinergic receptors. The Frog do not produce these alkaloids themselves but rely on their diet intake (ants and termites mostly) to stock-up on them.



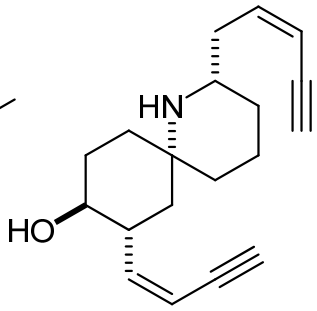
latratoxin

blackwidow

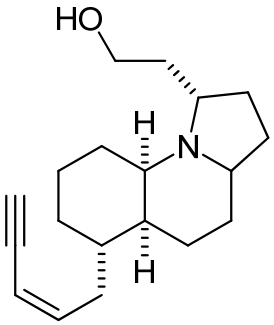
cobra : lecithinase



pumiliotoxin C



histrionicotoxin



gephyrotoxin

Figure V.6.3

References and notes

2. U. H. Maier and M. H. Zenk *Chem. Commun.*, **1997**, 2313-2314.
3. (a) J. E. Dowling In *Neurons and Networks: An Introduction to Behavioral Neuroscience*, 2nd edi., **2001**, Harvard University Press, Harvard, USA. (b) S. Funayama, G. A. Cordell In *Alkaloids: A Treasury of Poisons and Medicines*, Academic Press, Elsevier, London, U.K., **2015**.
4. J. B. Hendrickson and J. Wang *Org. Lett.* **2004**, *6*, 3-5.

VI. Natural Products of Marine Origin

(under construction, sorry for the inconvenience)

Marine natural metabolites are biosynthesized in much the same way as terrestrial secondary metabolites. There are, however, several structural features that are particular to natural products of marine origine that may be a consequence of the environment specific to marine organisms (such as halogens content) or a difference in interrelationship between species (such as the rarity of alkaloids). This chapter does not re-discuss the biosynthetic pathways, though it is not impossible that new pathways will be discovered as this relatively new field of study matures, but instead it lists several metabolites according to the terrestrial taxonomy (e.g. steroids, acetogenins, neurotoxins, or polyether antibiotics). Excellent reviews were published in a 1993 issue of Chemical Reviews in which the whole issue is dedicated to marine natural products.[5]

VI.1. Terpenoids

VI.1.1. Monoterpenes

Le pentachlorooctatriene a été isolé du tunicier clavelina lepadiformis mais ce composé est typique des algues. Des microalgues du type Prochloron (cyanobactéries) ont été identifiées en tant que symbions aux tuniciers.

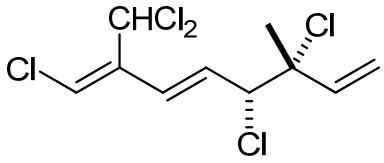


Figure VI.1.2

VI.1.2. Sesquiterpenes

From …

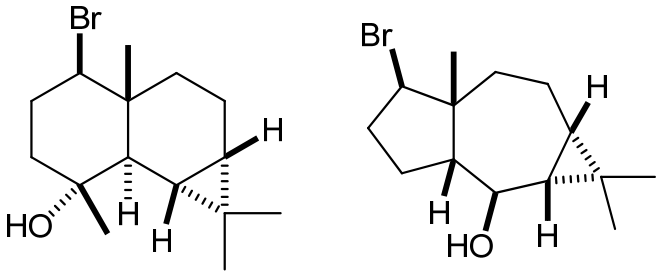
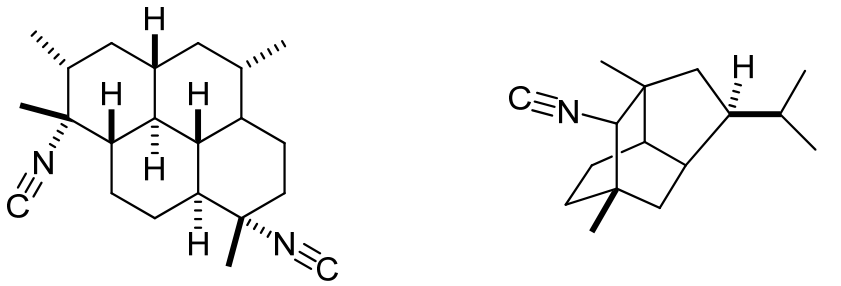


Figure VI.1.1

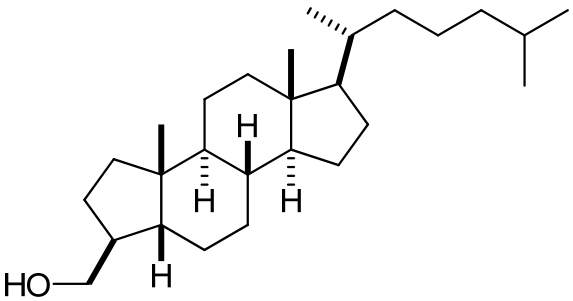
Terpenes from marine sponges.



Diisocyanoadvérane 2-Isocyanopupukéanane

Figure VI.1.3

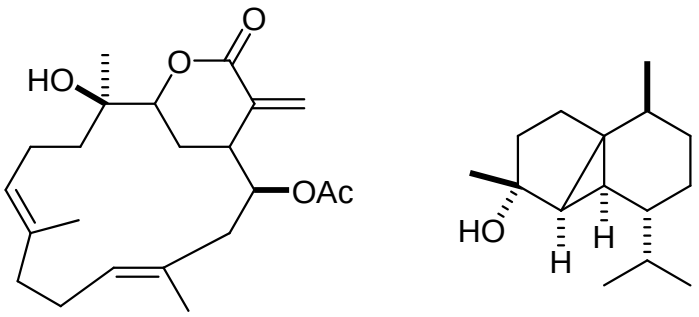
Steroids from marine sponges. Biosynthesis from cholesterol.



A-Norstanol

Figure VI.1.4

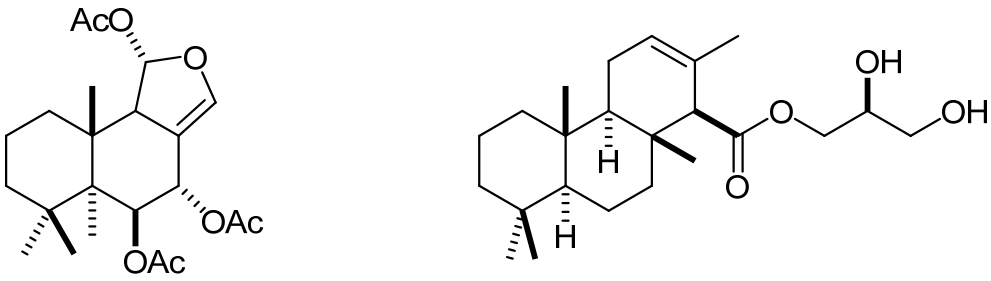
Terpenoids from corals.



crassin acetate cubétol

Figure VI.1.5

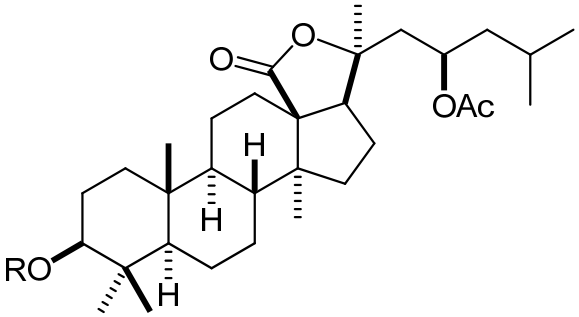
Terpenes from mollusks



Sesquiterpène Diterpènes

Figure VI.1.6

Echinoderms.



Stichopogénine A_4, R = diglycoside

Figure VI.1.7

VI.2. Polyacetates

Several C16 aromatic acids of acetogenic origin are produced by marine bacteria… They are bronchodilators…. The cyclic C12 hydrocarbon comes from macroscopic algae.

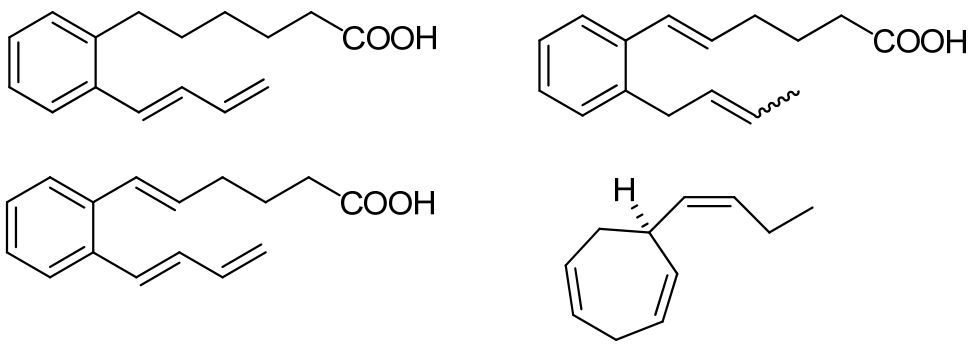
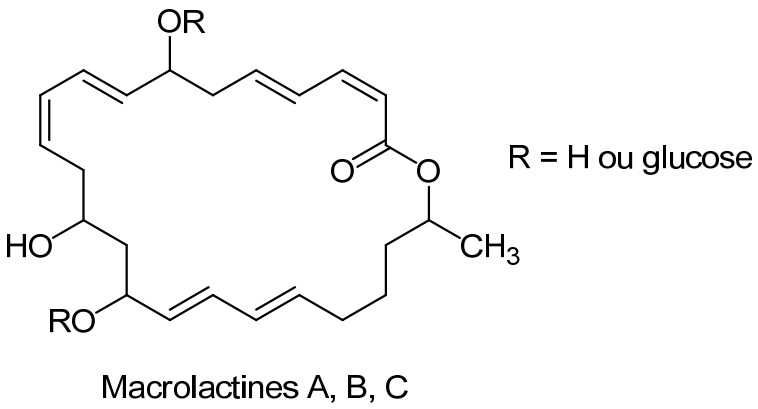


Figure VI.2.1

Antiviral and cytotoxic macrolides found in bacteria inside bottom sediments.



Macrolactines A, B, C

Figure VI.2.2

Cytotoxic acetogenins in bacteria on the water surface

Figure VI.2.3

Anticancer alkaloids in bacteria on the water surface

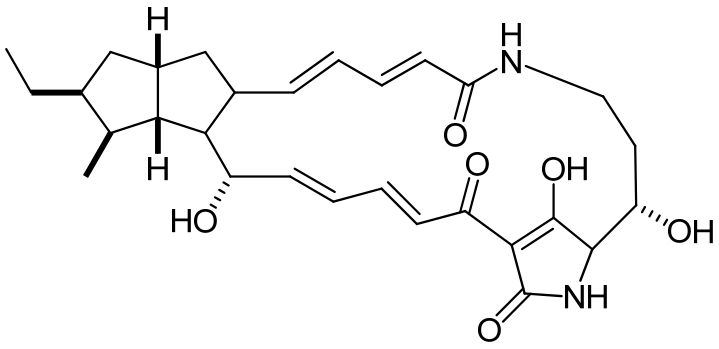
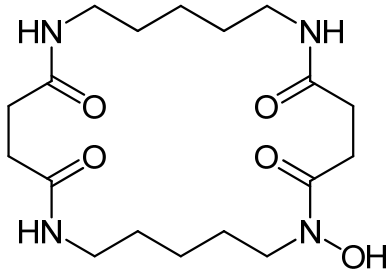


Figure VI.2.4

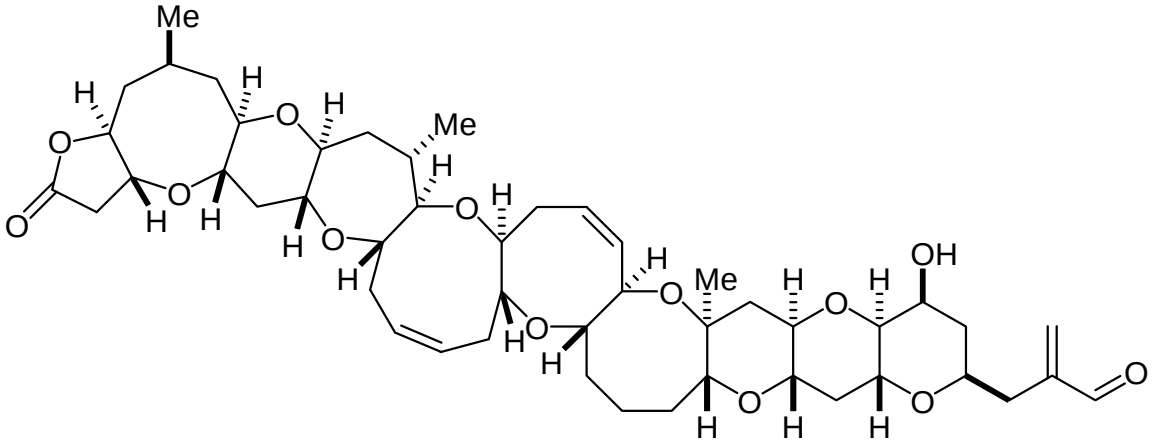
Bisacaberine renders tumoral cells sensitive to cytolytic action of certain macrophages. It is a marine bacterial toxin.



Bisacabérine

Figure VI.2.5

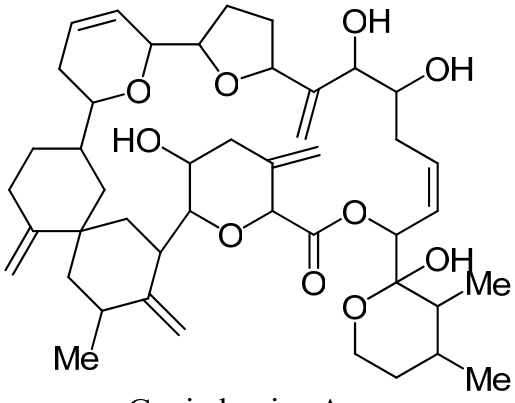
Brevitoxines is a member of the polyethers family coming from microscopic dinoflagellates algae.



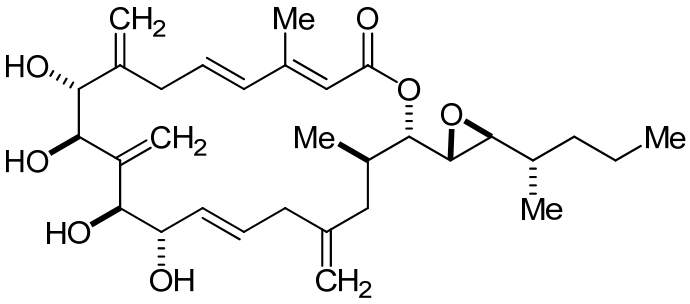
Brévitoxine A

Figuve VI.2.6

Macrolides also from microscopic dinoflagellates algae



Goniodomine A



Amphidinolide A

Figuve VI.2.7

Tolytoxine (macrolide) from blue or green algae

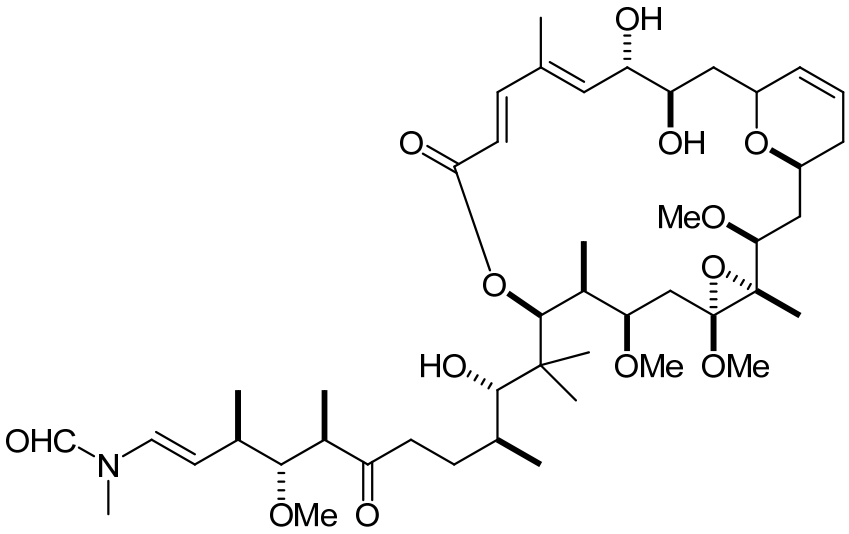
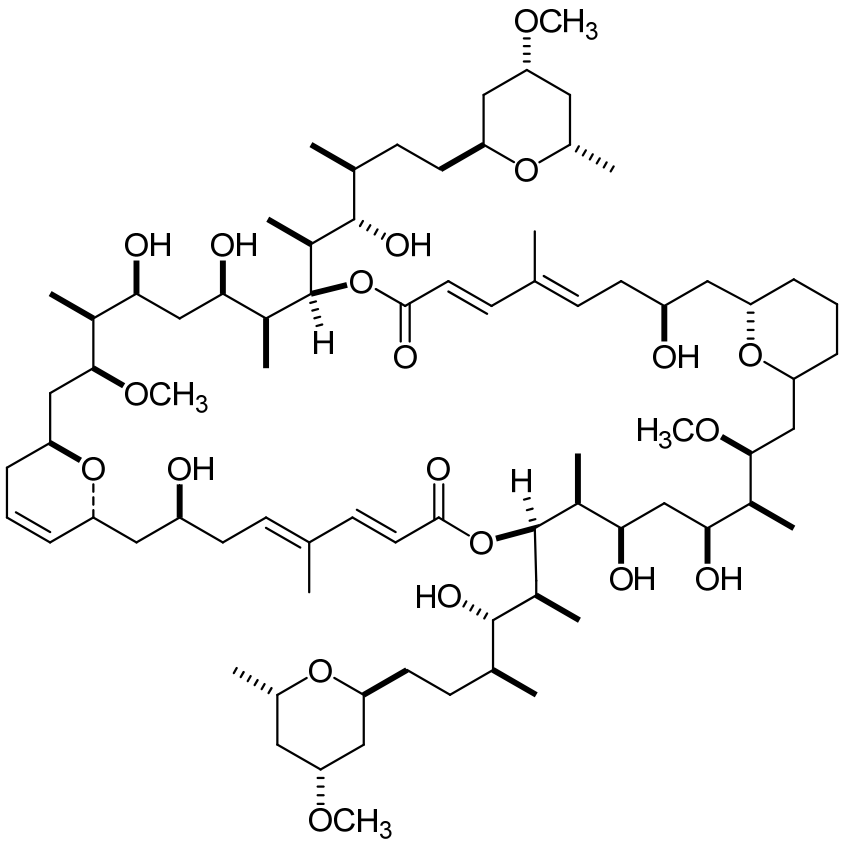


Figure VI.2.8

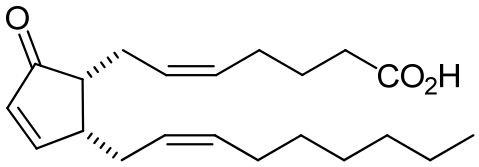
Macrolides from marine sponges.



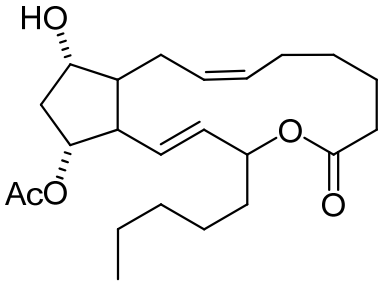
Swinholide

Figure VI.2.9

Prostaglandines from corals or mollusks



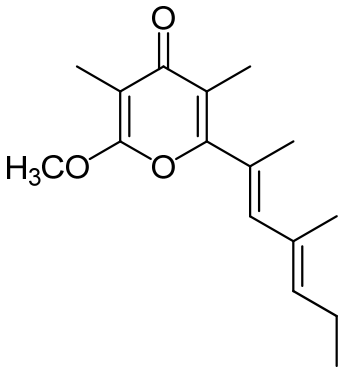
Precuvalone A
(from coral)



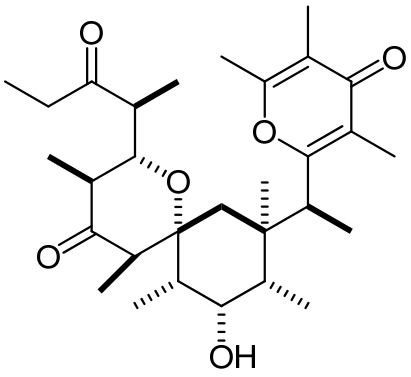
PGF2a 1,15-lactone
(from mollusks)

Figure VI.2.10

Polyacetate from mollusks



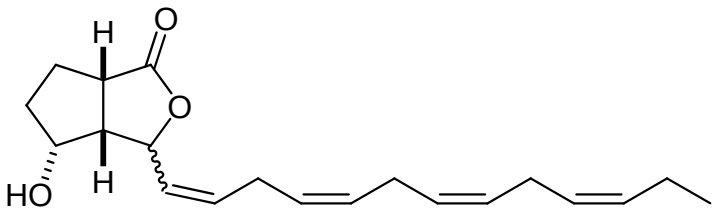
cyercène



siphonarine A

Figure VI.2.11

Bacillariolides contain de structurally unique oxylipin skeleton. They are isolated from the marine diatom *Pseudonitzschia multiseries*, a causative diatom of so-called amnesic shellfish poisoning.



bacillariolide I = β
bacillariolide II = α

Figure VI.2.11

VI.3. Alkaloids

VI.3.1.

Tunichromes (tripeptides derived from tyrosine and phenylalanine) (pigments du sang)

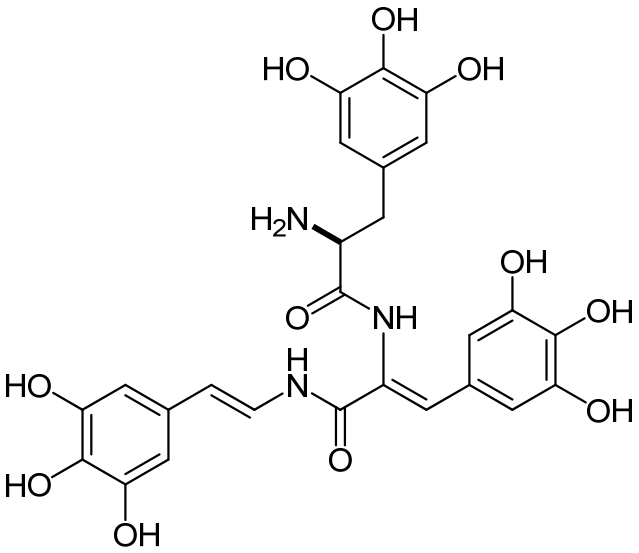
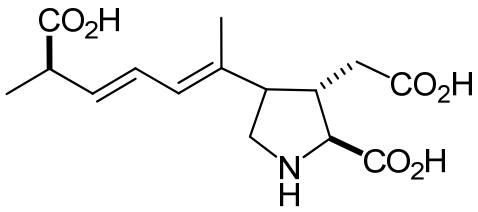


Figure VI.3.1

Domoic acid (biosynthesis : GPP + glutamic acid) from diatoms (microscopic algae)



domoic acid

Figure VI.3.2

---

5. *Chem. Rev.* **1993**, *93*, 1673…1937