

# Biosynthesis

## *Biosynthesis of Isoprenoids: Terpenes (Including Steroids & Carotenoids)*

Alan C. Spivey  
a.c.spivey@imperial.ac.uk

**Imperial College**  
London

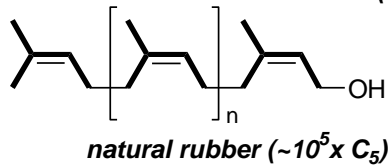
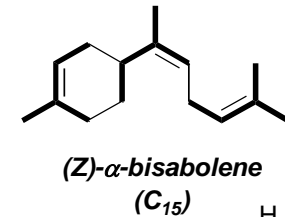
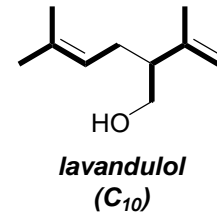
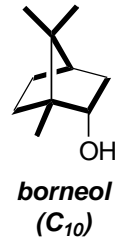
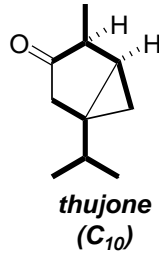
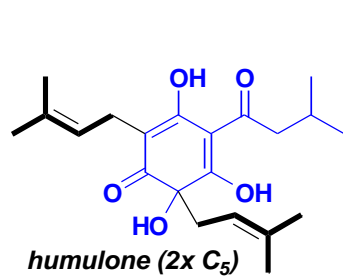
*Dec 2014*

# Format & Scope of Lectures

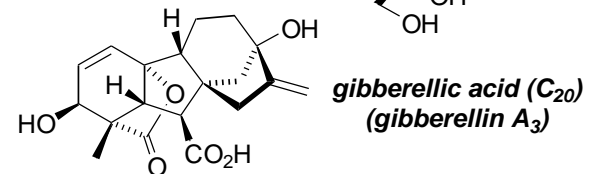
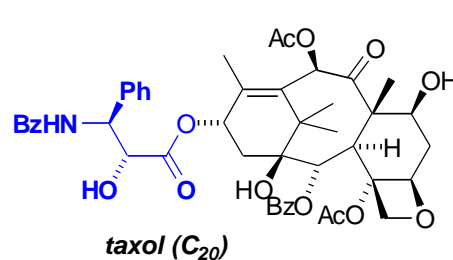
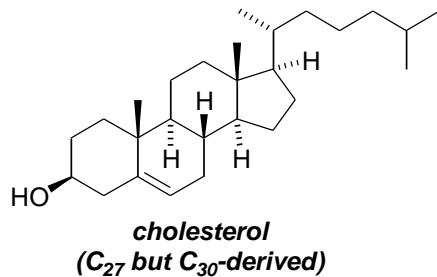
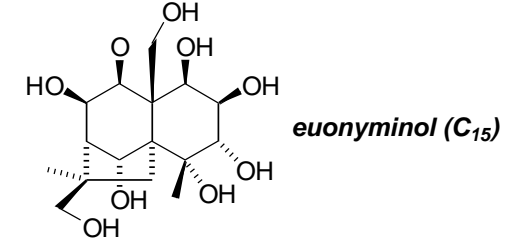
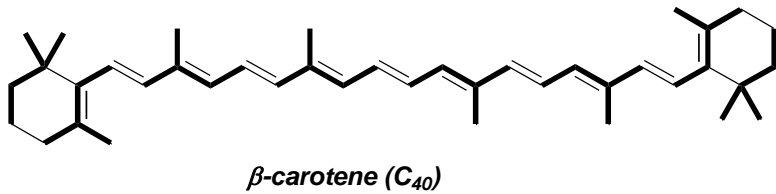
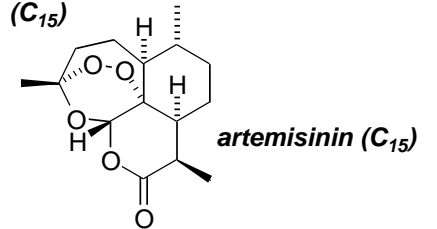
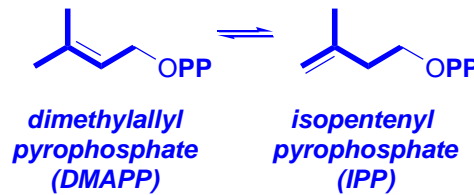
- **What are isoprenoids?**
  - $n \times C_5$  diversity: terpenes, steroids, carotenoids & natural rubber
  - ‘the isoprene rule’
  - mevalonate & 1-deoxyxylulose pathways to IPP & DMAPP
- **Monoterpenes ( $C_{10}$ )**
  - regular (‘head-to-tail’) *via* geranyl pyrophosphate
  - irregular: *incl.* iridoids (*e.g.* seco-loganin)
- **Sesquiterpenes ( $C_{15}$ )**
  - farnesyl pyrophosphate derived metabolites
  - sesquiterpene cyclases: pentalenene, aristolochene & 5-*epi*-aristolochene
- **Diterpenes ( $C_{20}$ )**
  - gibberellins & taxol
- **Triterpenes ( $C_{30}$ )**
  - hopanoids (squalene → hopene)
  - steroids (2,3-oxidosqualene → lanosterol → cholesterol → estrone)
  - ring-opened ‘steroids’: vitamin D<sub>2</sub> & azadirachtin
- **Biomimetic cationic cyclisation cascades**
- **Carotenoids ( $C_{40}$ )**
  - β-carotene, retinal & vitamin A

# Isoprenoids

- **isoprenoids** are widely distributed in the natural world
  - particularly prevalent in plants and least common in insects; >30,000 known
  - composed of integral numbers of C<sub>5</sub> 'isoprene' units:
    - **monoterpenes** (C<sub>10</sub>); **sesquiterpenes** (C<sub>15</sub>); **diterpenes** (C<sub>20</sub>); **sesterpenes** (C<sub>25</sub>, *rare*); **triterpenes** (C<sub>30</sub>); **carotenoids** (C<sub>40</sub>)



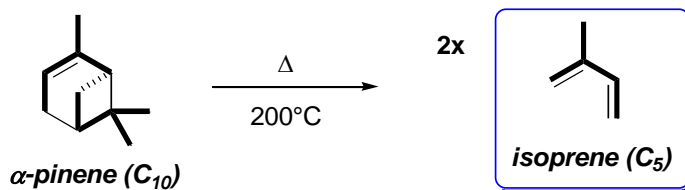
**ISOPRENOIDS**



# Historical Perspective – ‘The Isoprenoid Rule’

- **Early 1900s:**

- common **structural feature** of terpenes – **integral # of C<sub>5</sub> units**
- **pyrolysis** of many monoterpenes produced two moles of **isoprene**:



- **1940s:**

- **biogenesis** of terpenes attributed to oligomers of isoprene – ‘**the isoprene rule**’

- **1953:**

- **Ruzicka** proposes ‘**the biogenetic isoprene rule**’ to accommodate ‘irregular’ terpenoids:
  - *i.e.* that terpenes were derived from a number of **biological equivalents of isoprene** initially joined in a ‘**head-to-tail**’ manner & sometimes subsequently modified enzymatically to provide greater diversity of structure

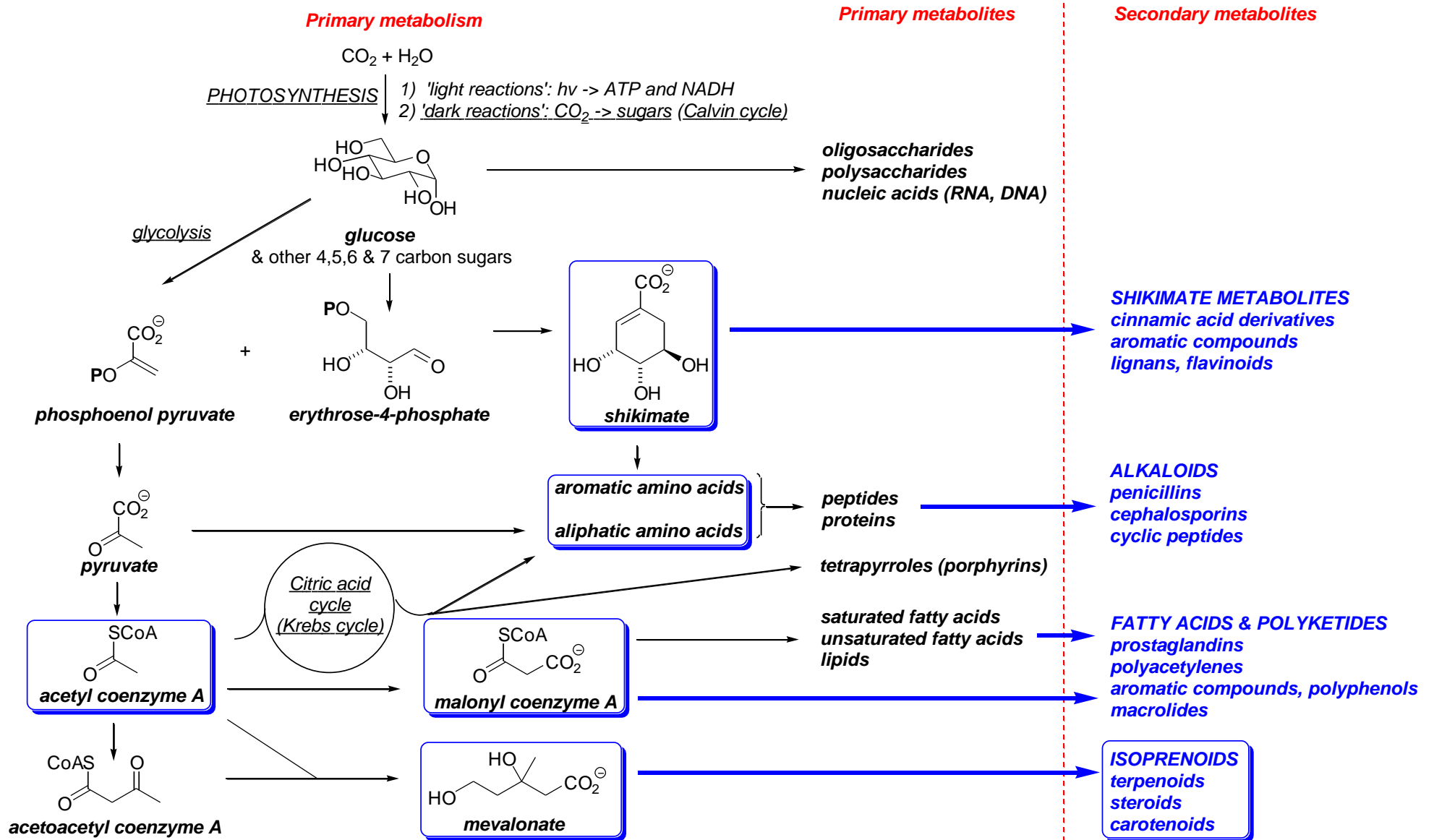
- **1964:**

- **Nobel prize** awarded to **Bloch, Cornforth & Popjak** for elucidation of biosynthetic pathway to **cholesterol** including the first steps:
  - **acetate** → **mevalonate (MVA)** → **isopentenyl pyrophosphate (IPP)** & **dimethylallyl pyrophosphate (DMAPP)**

- **1993:**

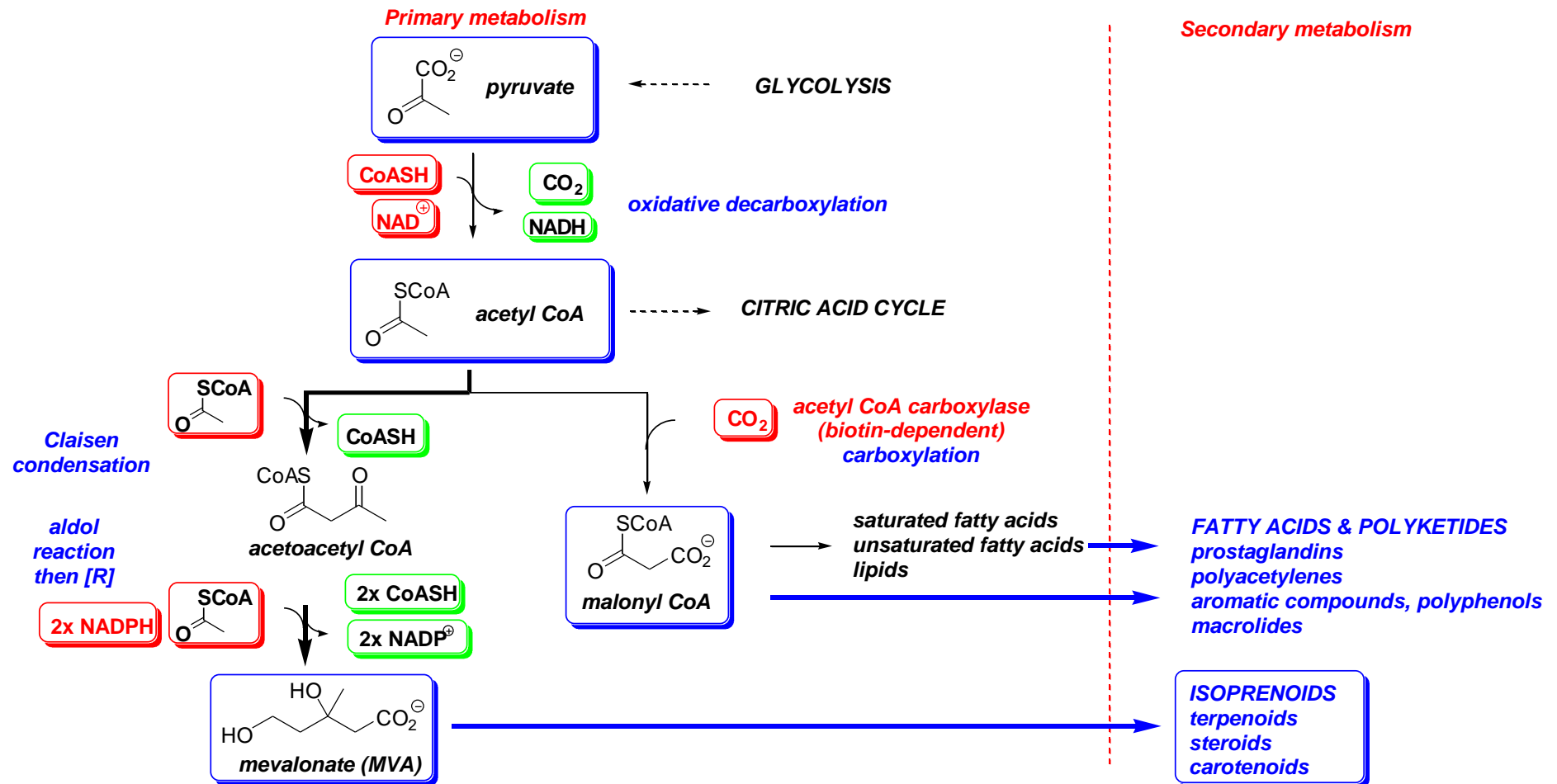
- **Rohmer, Sahn & Arigoni** elucidate an additional pathway to **IPP & DMAPP**:
  - **pyruvate + glyceraldehyde-3-phosphate** → **1-deoxyxylulose-5-phosphate** → **IPP & DMAPP**

# Primary Metabolism - Overview



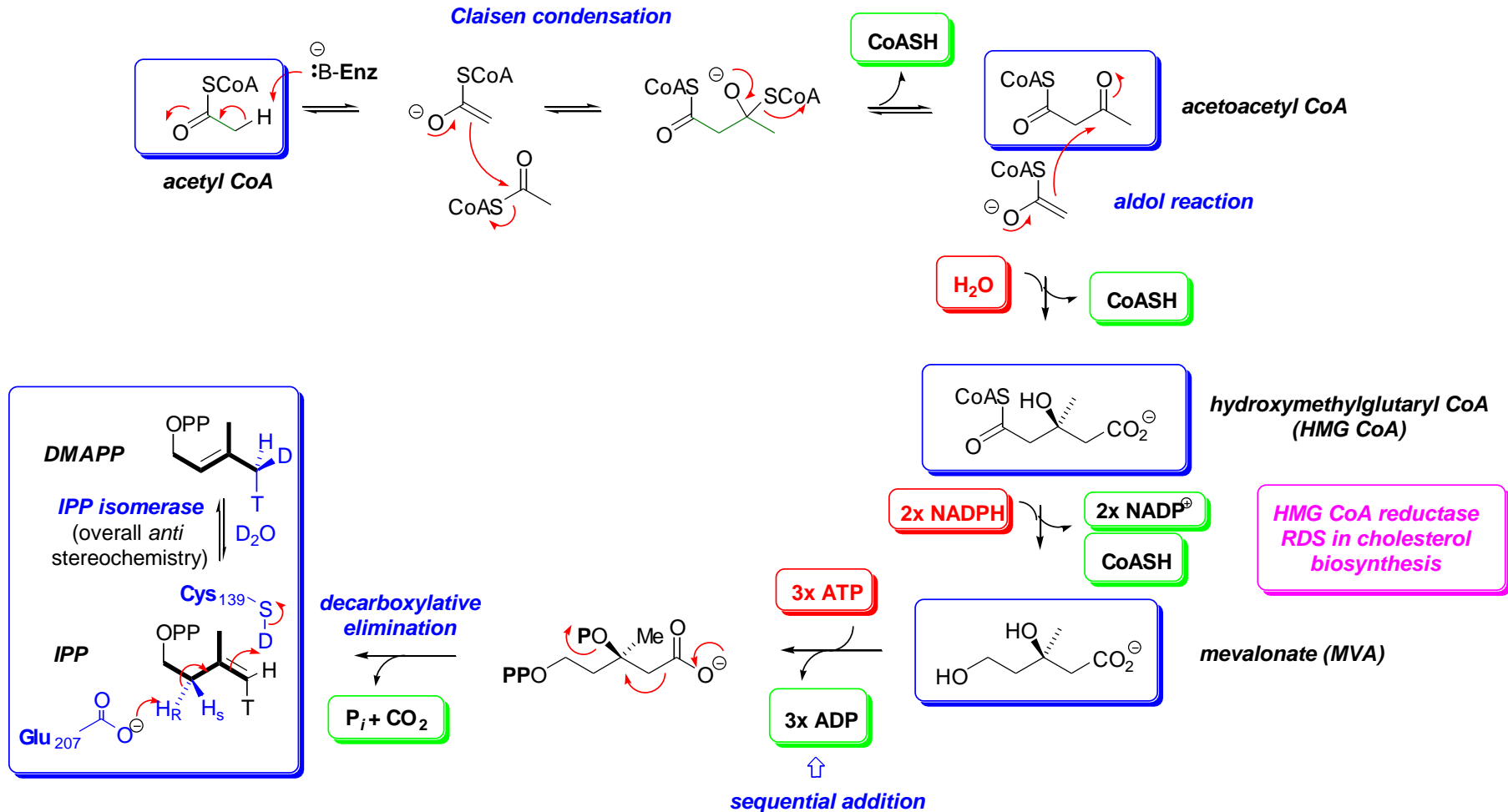
# Biosynthesis of Mevalonate

- **Mevalonate (MVA)** is the first committed step of **isoprenoid biosynthesis**
  - this key 6-carbon metabolite is formed from three molecules of **acetyl CoA** via **acetoacetyl CoA**:



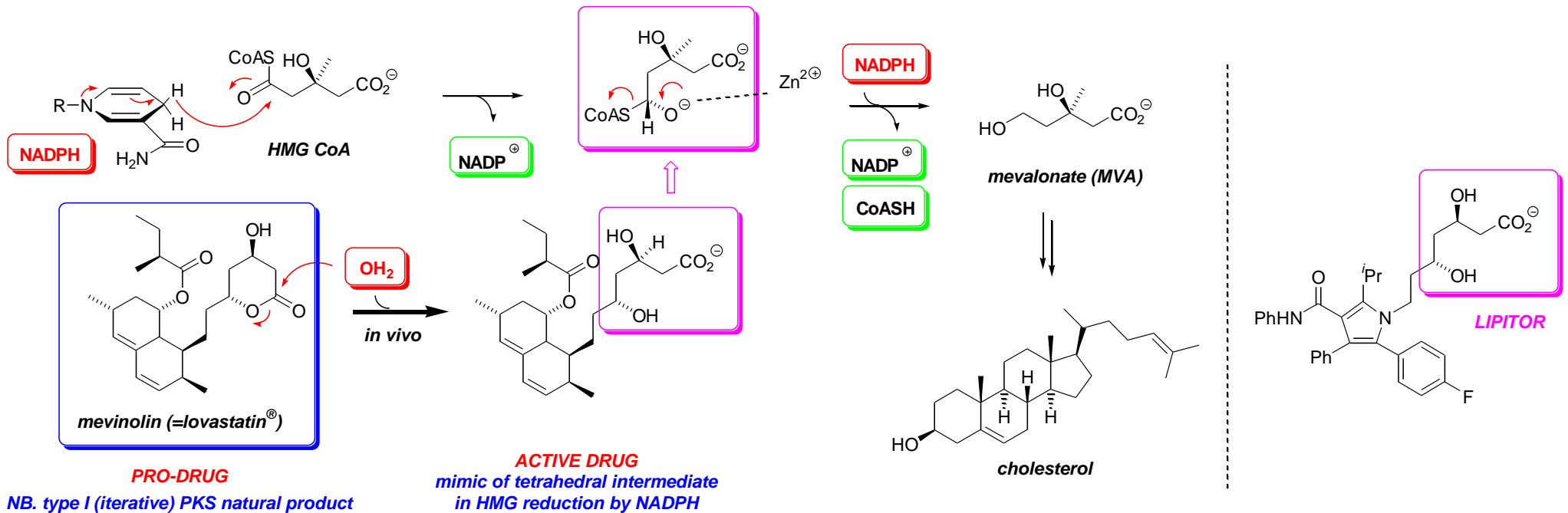
# Biosynthesis of IPP & DMAPP - via Mevalonate

- **IPP & DMAPP** are the key **C<sub>5</sub> precursors** to **all isoprenoids**
  - the **main pathway** is via: **acetyl CoA** → **acetoacetyl CoA** → **HMG CoA** → **mevalonate** → **IPP** → **DMAPP**:



# HMG CoA reductase inhibitors - *Statins*

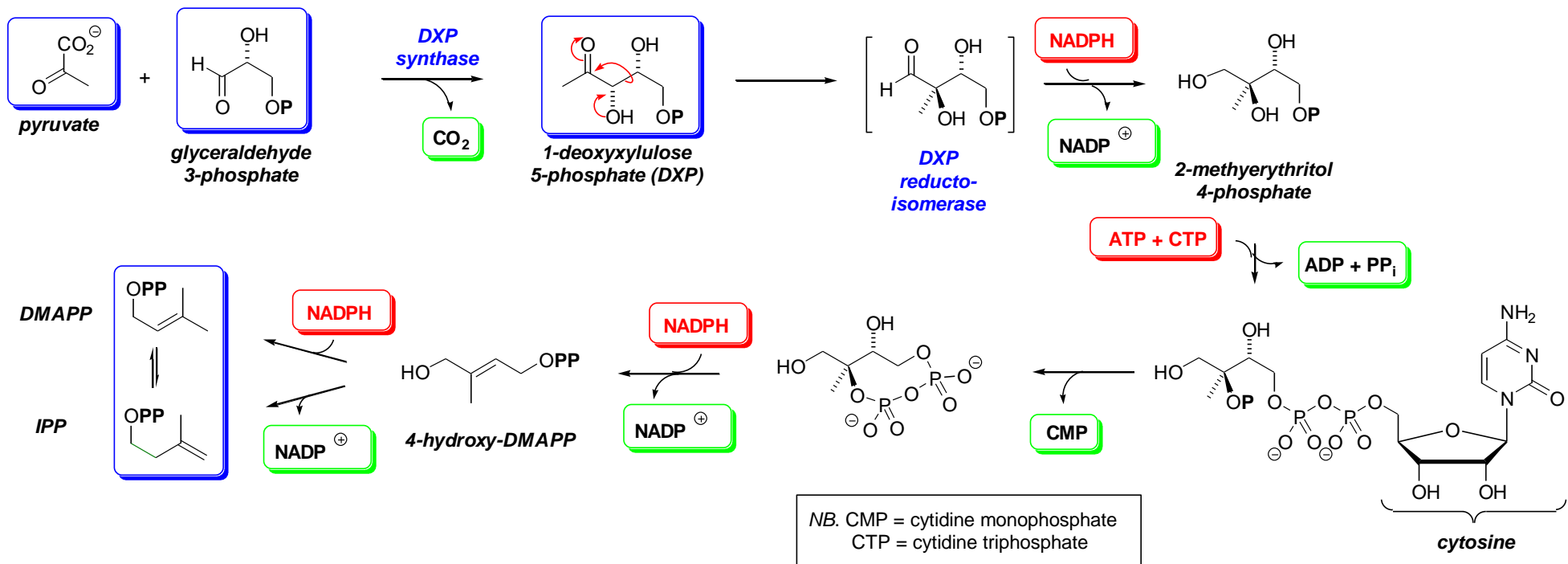
- **HMG CoA** → **MVA** is the **rate determining step** in the biosynthetic pathway to **cholesterol**
  - 33 enzyme mediated steps are required to biosynthesise cholesterol from acetyl CoA & in principle the inhibition of any one of these will serve to break the chain. In practice, control rests with HMG-CoA reductase as the result of a variety of biochemical feedback mechanisms
- ‘**Statins**’ inhibit HMG CoA reductase and are used clinically to treat **hypercholesterolemia** - a causative factor in **heart disease**
  - e.g. **mevinolin** (=lovastatin<sup>®</sup>, Merck) from *Aspergillus terreus* is a competitive inhibitor of HMG-CoA reductase





# Biosynthesis of IPP & DMAPP – via 1-Deoxyxylulose

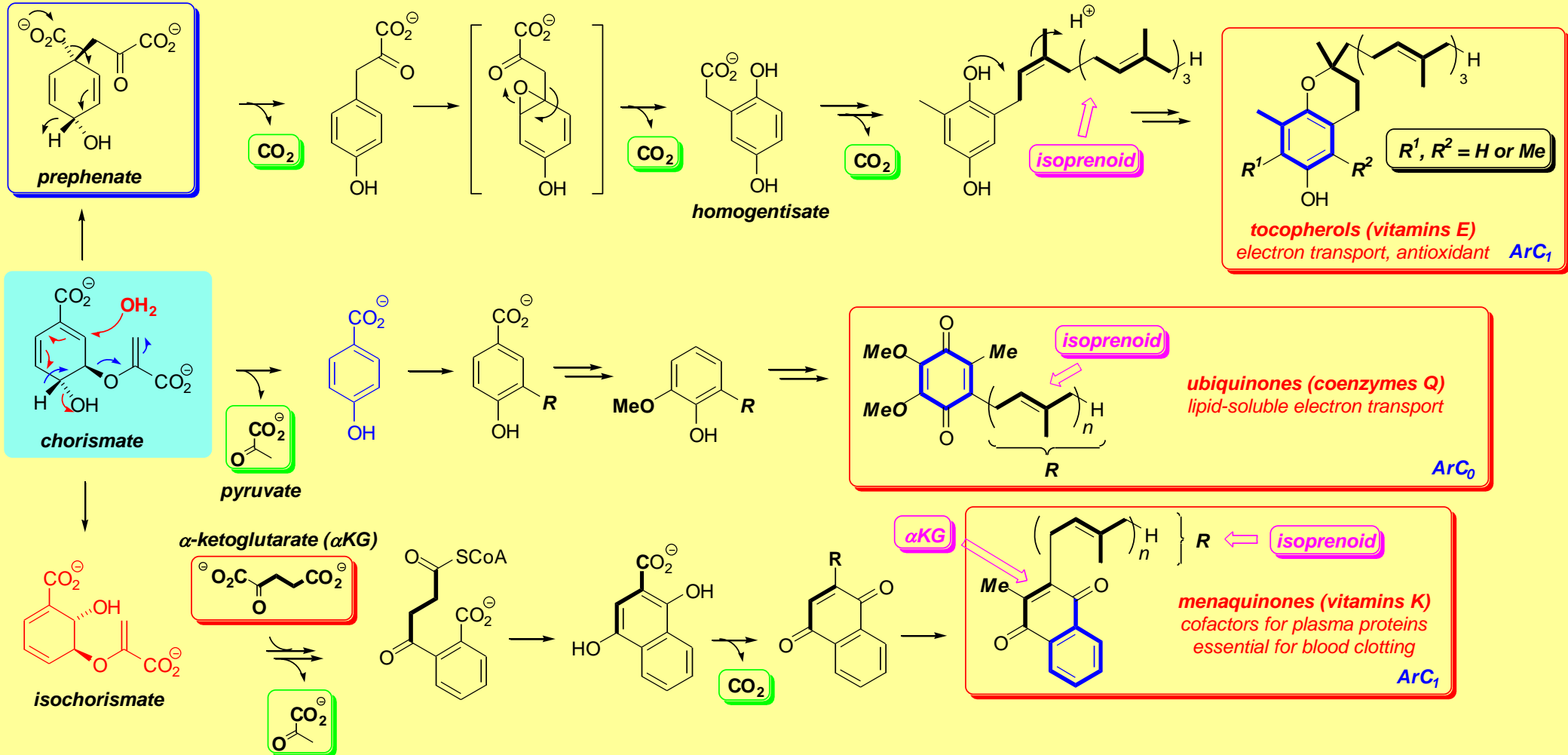
- the **mevalonate** route to **IPP** & **DMAPP** has been proven in **yeast** & **animals** & in **some plants** & for a long while was believed to be the only pathway to these key intermediates
  - However, in some bacterial labelling studies:
    - no incorporation of mevalonolactone was observed
    - the pattern of label from glucose was inconsistent with derivation *via* catabolism to acetate
- in **1993** an additional pathway to **IPP** & **DMAPP** was discovered:
  - Rohmer *et al. Biochem. J.* **1993**, 295, 517 ([DOI](#))



- The pathway is prevalent in many **pathogenic bacteria** and so its inhibition represents an exciting opportunity for anti-infective therapeutic development: Rohdich *J. Org. Chem.* **2006**, 71, 8824 ([DOI](#))

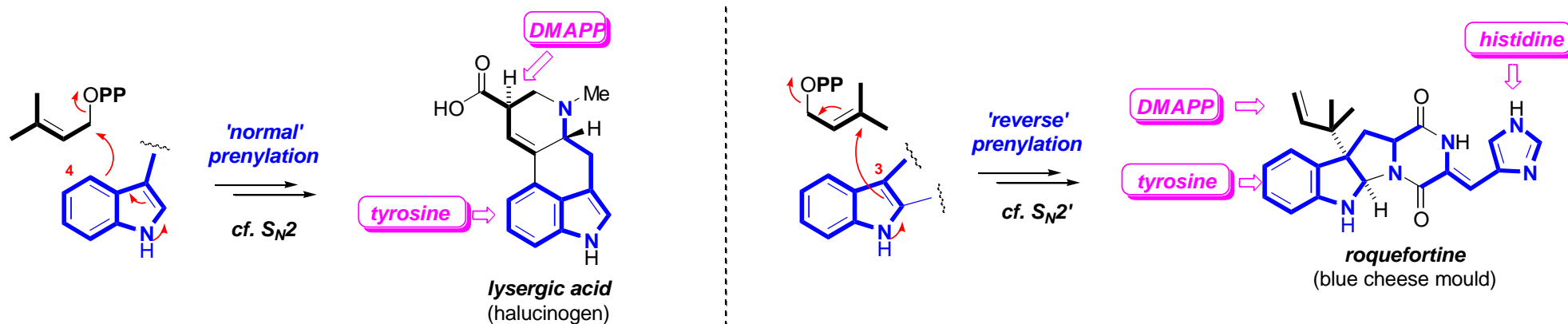
# Chorismate → Coenzymes Q & Vitamins E & K

- **Chorismate** → *p*- & *o*-hydroxybenzoic acids → **coenzymes Q & vitamins E & K**
  - NB. 'Mixed' biosynthetic origin: **shikimate/mevalonate (isoprenoid)**



# Hemi-Terpenes – ‘Prenylated Alkaloids’

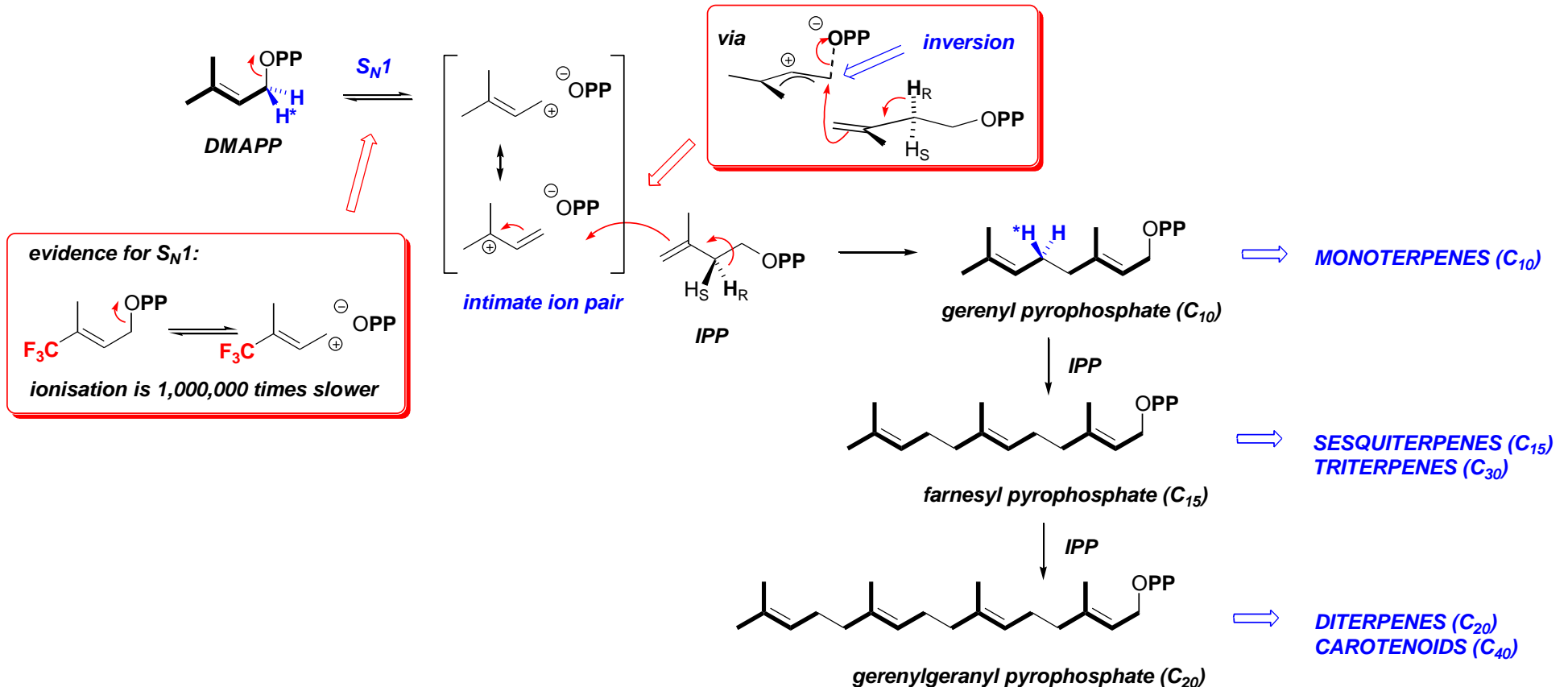
- **DMAPP** is an excellent **alkylating agent**
- **C<sub>5</sub> units** are frequently encountered as part of **alkaloids** (& **shikimate metabolites**) due to ‘late-stage’ alkylation by **DMAPP**
  - the transferred **dimethyl allyl unit** is often referred to as a ‘**prenyl group**’
  - ‘**normal prenylation**’ – ‘**S<sub>N</sub>2**’-like alkylation; ‘**reverse prenylation**’ – ‘**S<sub>N</sub>2**’-like alkylation
- e.g. **lysergic acid** (recall the **ergot alkaloids**) – a ‘normal prenylated’ alkaloid (with significant subsequent processing)
- e.g. **roquefortine** (recall **diketopiperazine alkaloids**) – a ‘reverse prenylated’ alkaloid



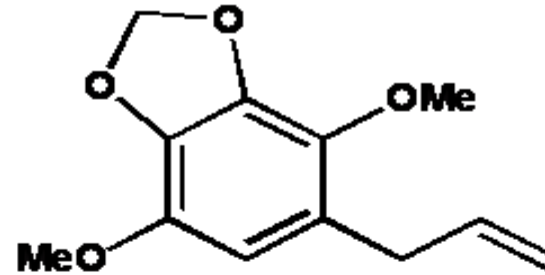
- **review:** R.M. Williams *et al.* ‘Biosynthesis of prenylated alkaloids derived from tryptophan’ *Top. Curr. Chem.* **2000**, 209, 97-173 ([DOI](#))

# Linear C<sub>5n</sub> 'head-to-tail' Pyrophosphates

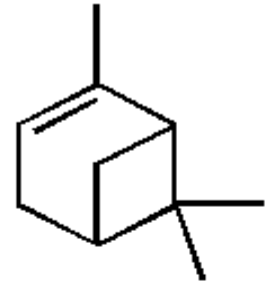
- head-to-tail **C<sub>5</sub> oligomers** are the key precursors to isoprenoids
  - geranyl** pyrophosphate (C<sub>10</sub>) is formed by **S<sub>N</sub>1 alkylation** of **DMAPP** by **IPP** → **monoterpenes**
  - farnesyl** (C<sub>15</sub>) & **geranylgeranyl** (C<sub>20</sub>) pyrophosphates are formed by **further S<sub>N</sub>1 alkylations** with **IPP**:



# Monoterpenes from Parsley & Sage

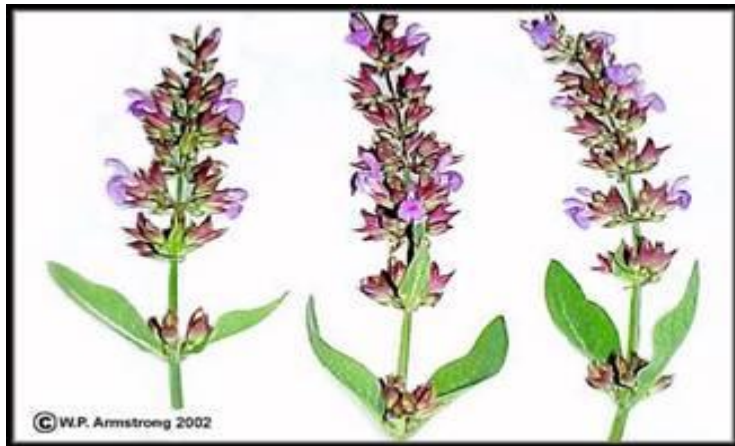


apiol

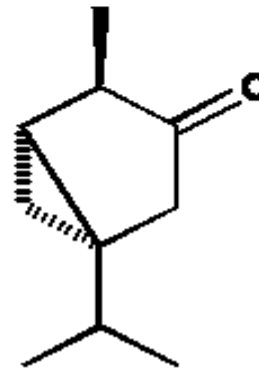


$\alpha$ -pinene

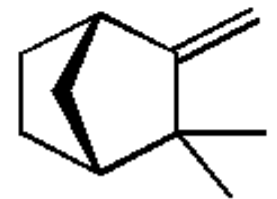
Parsley (*Petroselinum sativum*)



Sage (*Salvia officinalis*)



thujone



camphene

# Monoterpenes from Rosemary & Thyme



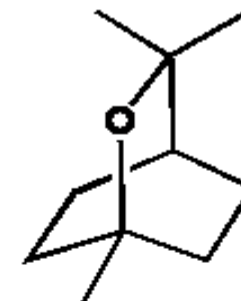
Rosemary (*Rosmarinus officinalis*)



camphor



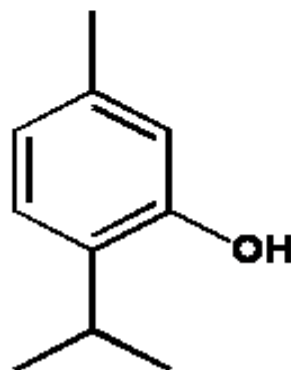
borneol



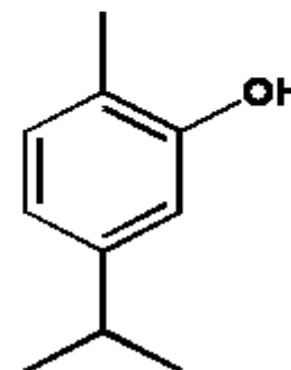
cineol



Thyme (*Thymus vulgaris*)



thymol

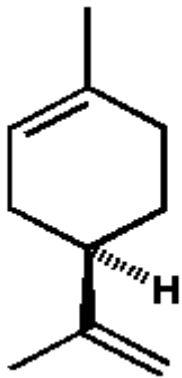


carvacrol

# Limonene & Carvone



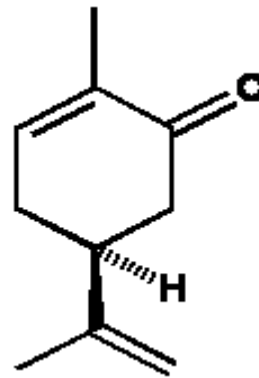
Chiroscience plc. (now Dow Inc.)



1. *S*-(-)-limonene (lemon)

2. *R*-(+)-limonene (orange)

3. *RS*-(±)-limonene (pleasant)



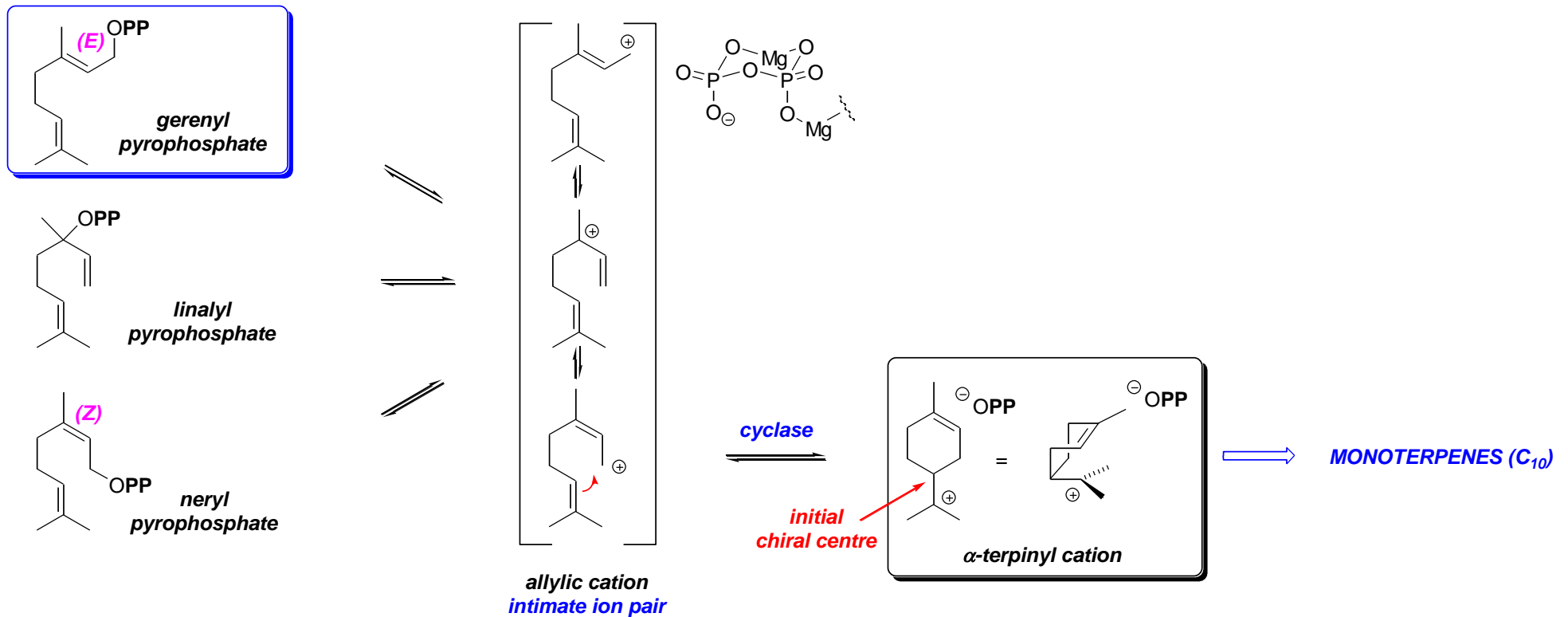
4. *R*-(-)-carvone (spearmint)

5. *S*-(+)-carvone (caraway)

6. *RS*-(±)-carvone (disgusting)

# Monoterpenes – $\alpha$ -Terpinyl Cation Formation

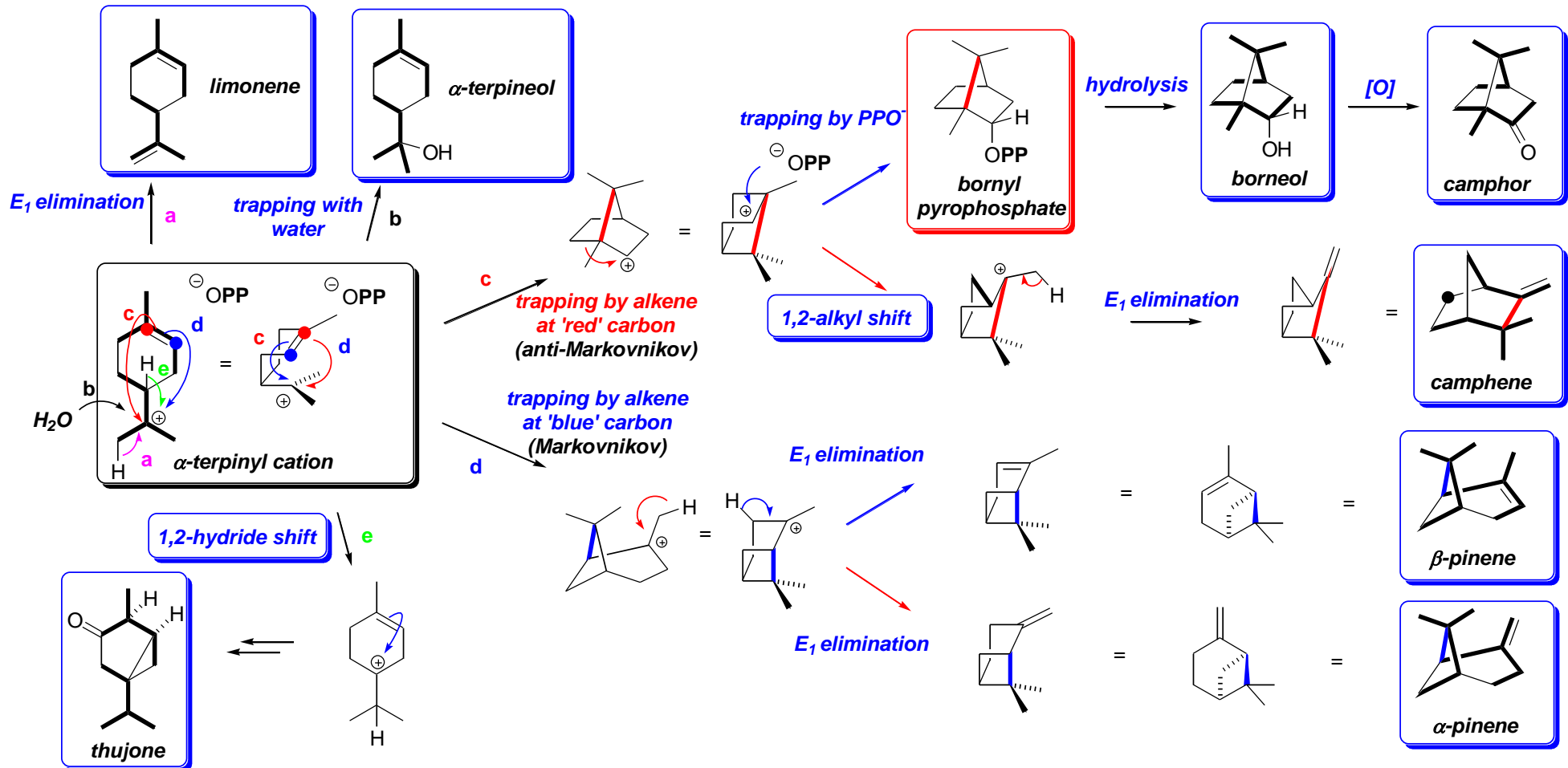
- **geranyl** pyrophosphate isomerises readily via an allylic cation to **linalyl** & **neryl** pyrophosphates
  - the leaving group ability of pyrophosphate is enhanced by coordination to  $3 \times \text{Mg}^{2+}$
  - all three pyrophosphates are substrates for **cyclases** via an  **$\alpha$ -terpinyl cation**:





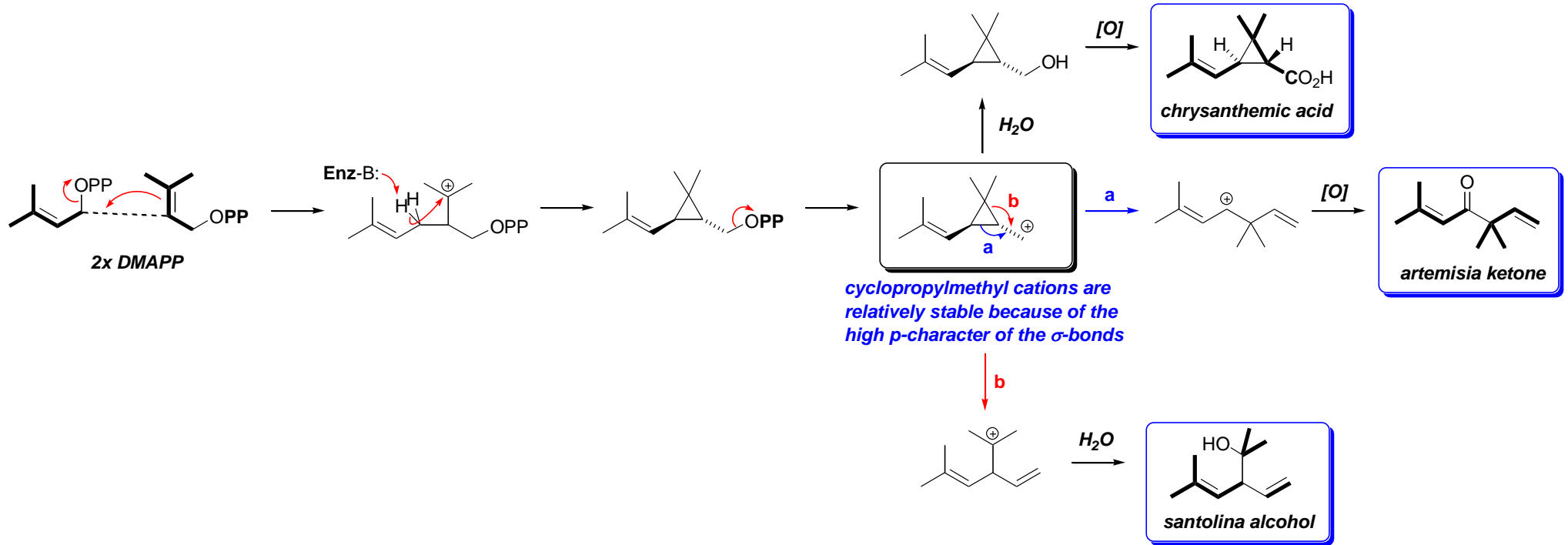
# Monoterpenes – Fate of the $\alpha$ -Terpinyl Cation

- The  $\alpha$ -terpinyl cation undergoes a rich variety of further chemistry to give a diverse array of **monoterpenes**
- Some important enzyme catalysed pathways are shown below
  - NB. intervention of **Wagner-Meerwein 1,2-hydride- & 1,2-alkyl shifts**

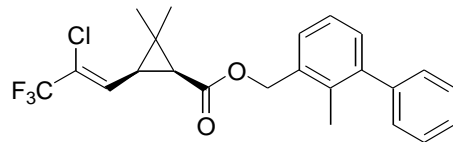


# Irregular Monoterpenes

- Non-'head-to-tail' linkage of **IPP** &/or **DMAPP** leads to '**irregular**' monoterpenes
  - e.g. **daisy** (*Compositae*) & **chrysanthemum** metabolites:



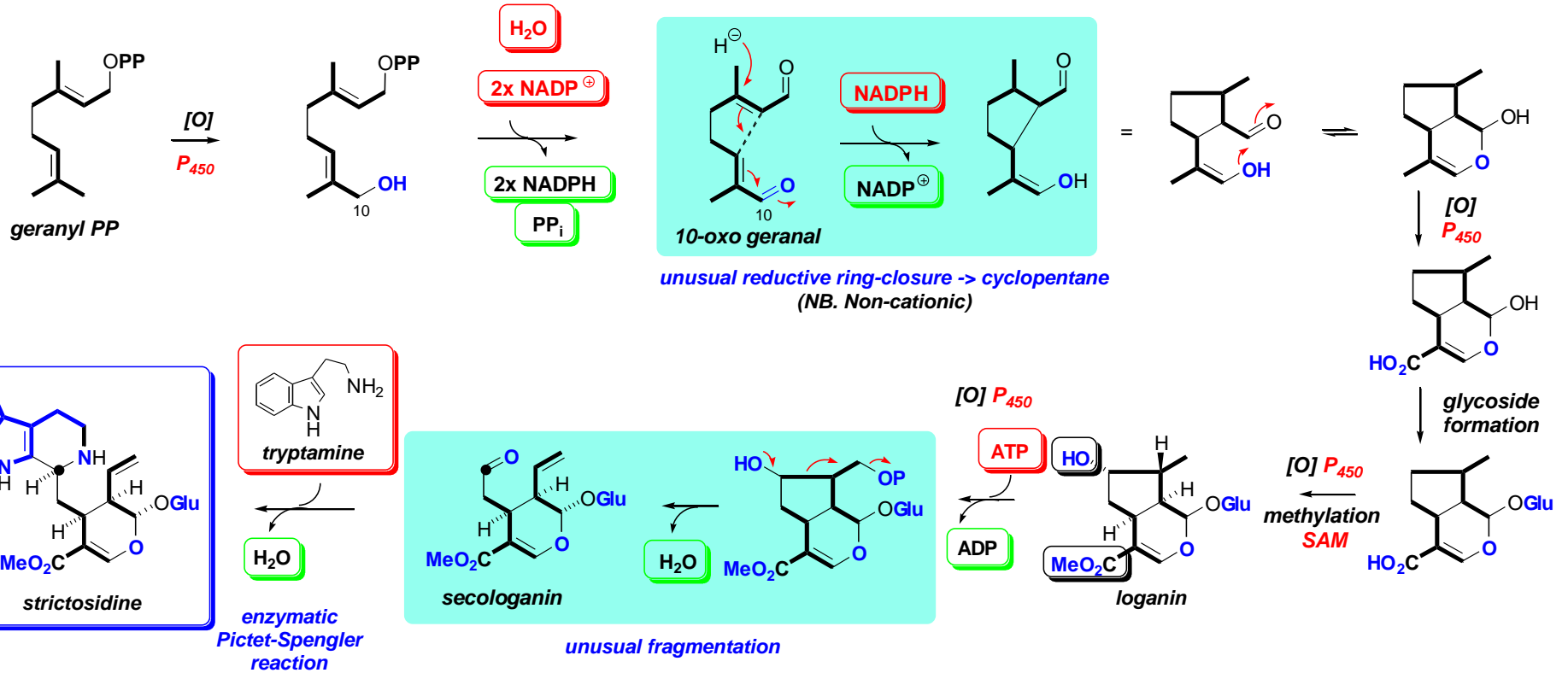
- Natural chrysanthemic acid derivatives are referred to as **pyrethrins** and are natural **insecticides**
- Synthetic analogues of chrysanthemic acid are referred to as **pyrethroids**. e.g. bifenthrin:



**Bifenthrin**  
potent insecticide (via ATPase inhibition)

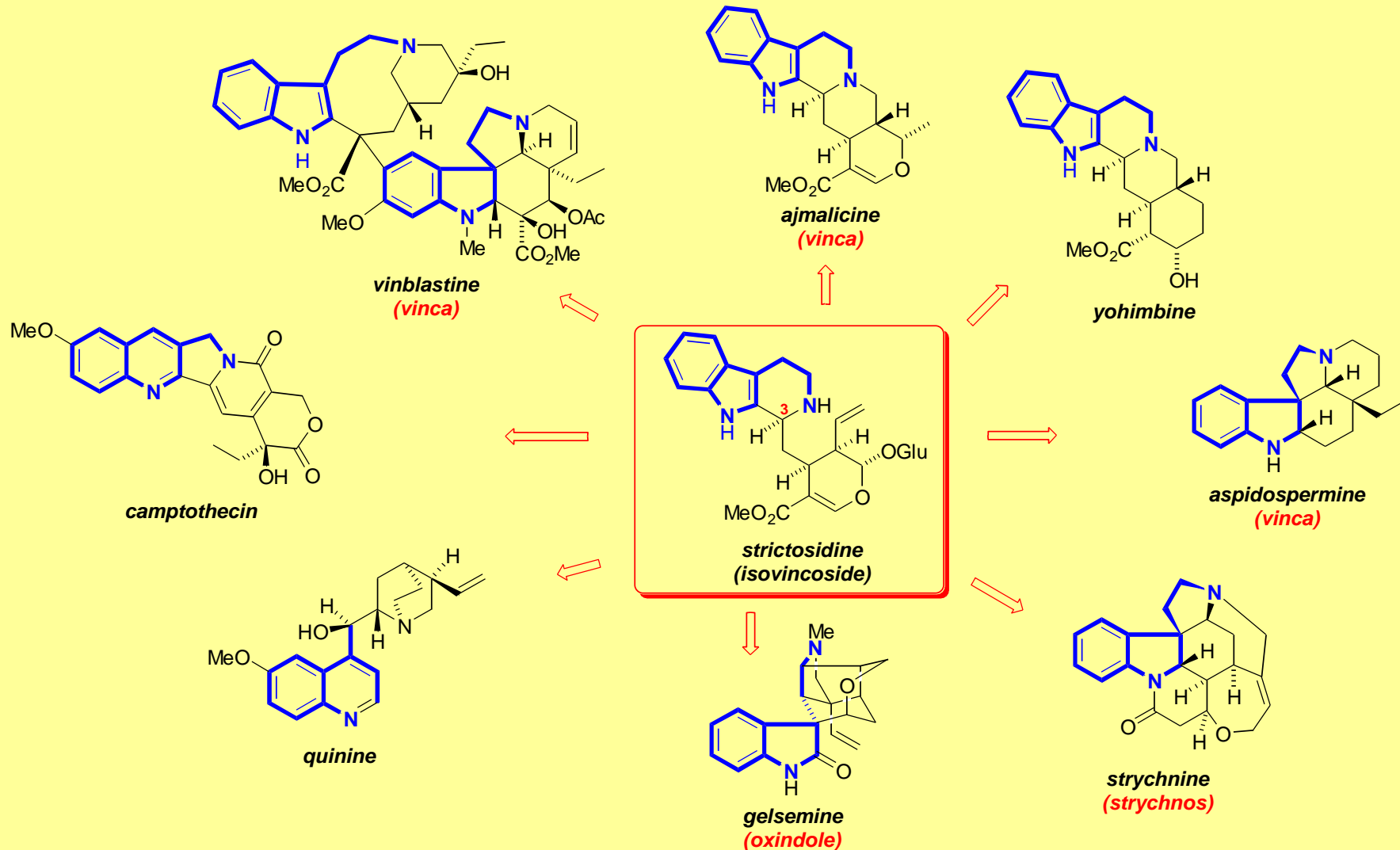
# Apparently Irregular Monoterpenes

- apparently **'irregular' monoterpenes** can also occur by **non-cationic cyclisation** of **geranyl PP** derivatives followed by **extensive rearrangement**
  - e.g. **iridoids** – named after *Iridomyrmex* ants but generally of plant origin and invariably glucosidated
    - e.g. **seco-loganin** (recall **indole alkaloids**) is a key component of **strictosidine** - precursor to numerous complex medicinally important alkaloids:



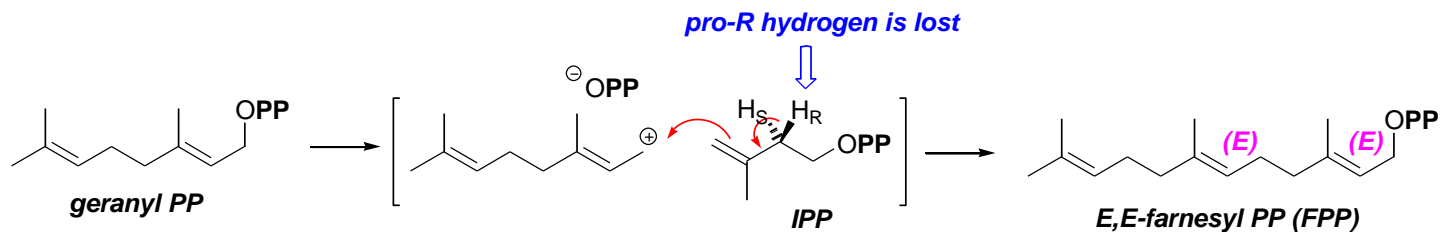
# Strictosidine → *Vinca*, *Strychnos*, *Quinine* etc.

- The diversity of alkaloids derived from **strictosidine** is stunning and many pathways remain to be fully elucidated:

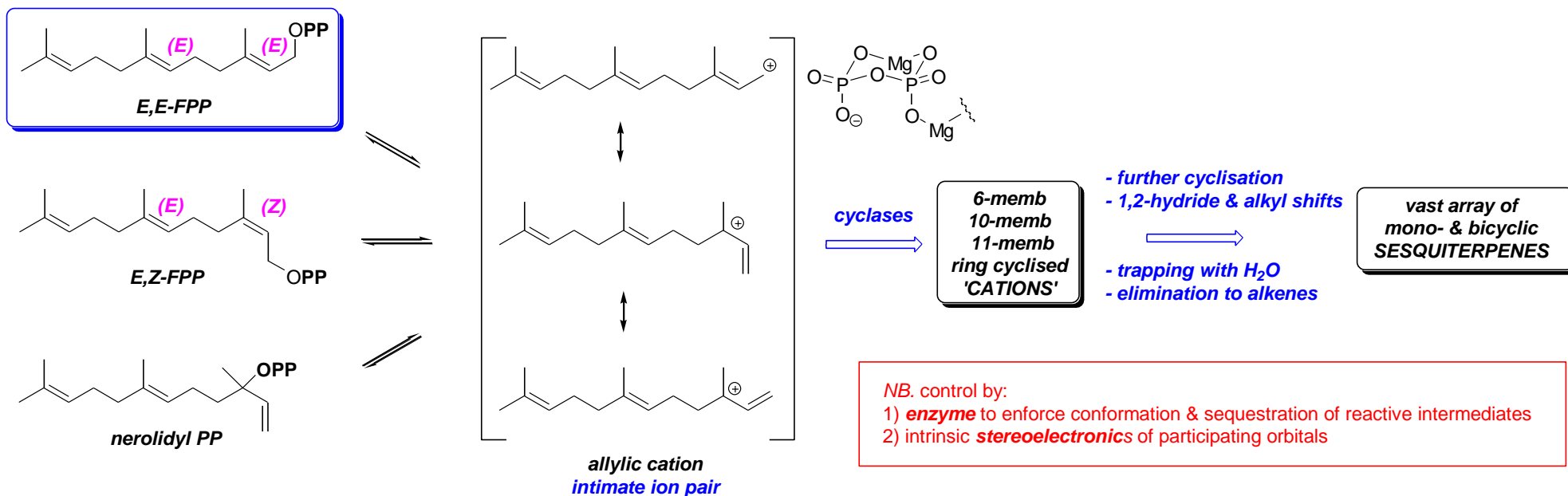


# Sesquiterpenes – *Farnesyl Pyrophosphate (FPP)*

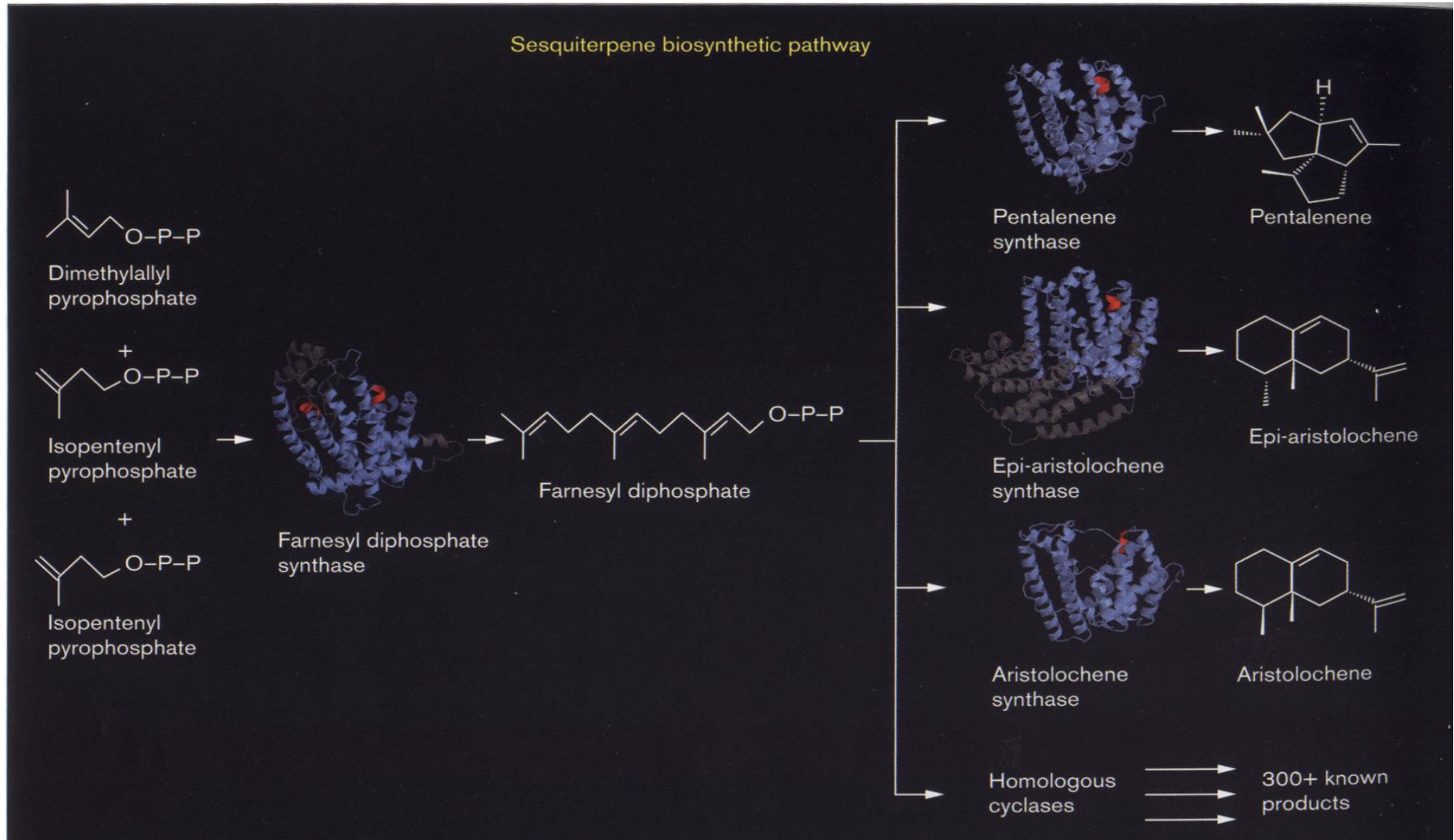
- ‘ $S_N2$ ’-like alkylation of *geranyl PP* by *IPP* gives *farnesyl PP*:



- just as *geranyl PP* readily isomerises to neryl & linalyl PPs so *farnesyl PP* readily isomerises to equivalent compounds – allowing many modes of cyclisation & bicyclisation



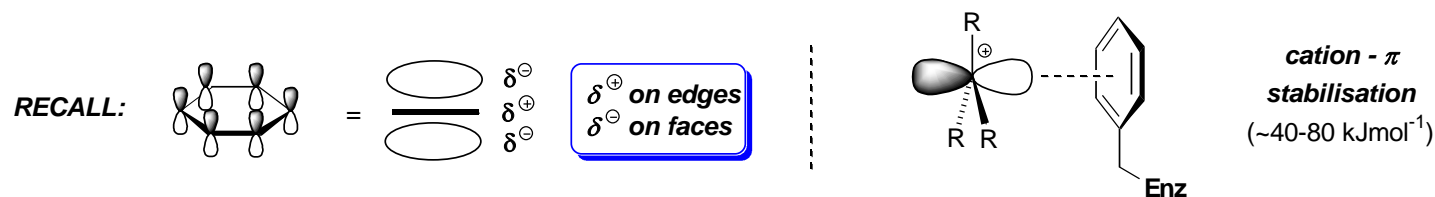
# Sesquiterpene Cyclases



# Terpene Cyclases – Control of Cyclisation

- **Functional aspects of terpenoid cyclases:**

- **Templating:** Active site provides a template for a specific conformation of the flexible linear isoprenoid starting material.
- **Triggering:** Cyclase initiates carbocation formation.
  - Metal-assisted leaving group departure (e.g. pyrophosphate ionization aided by  $Mg^{2+}$ )
  - C=C bond protonation (e.g. squalene-hopene cyclase, see later).
  - Epoxide protonation (e.g. oxidosqualene cyclase, see later).
- **Chaperoning:** Chaperones conformations of carbocationic intermediates through the reaction sequence, ordinarily leading to one specific product.
- **Sequestering:** Sequesters the carbocation intermediates by burying the substrate in a hydrophobic cavity that is generally solvent-inaccessible. Carbocations are concomitantly stabilized by the presence of aromatic residues in the active site that exert their effects *via* cation- $\pi$  interactions

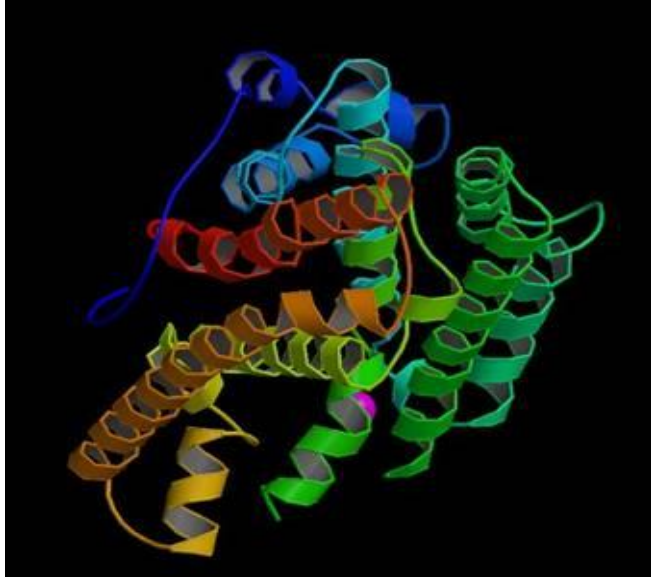


- Adapted from: Christianson *et al. Curr. Opin. Struct. Biol.* **1998**, 695 ([DOI](#))

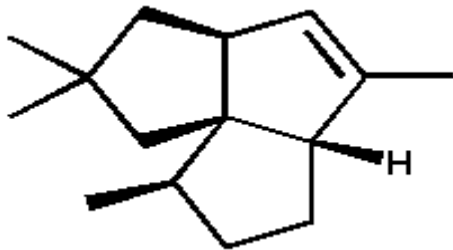
- **BUT:**

- individual terpene cyclases can give multiple products, see: Matsuda *J. Am. Chem. Soc.* **2007**, 129, 11213 ([DOI](#))

# Sesquiterpene Cyclase Crystal structures



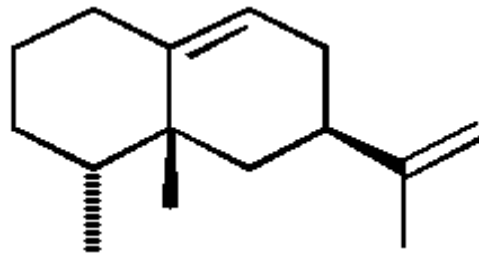
**pentalenene synthase**



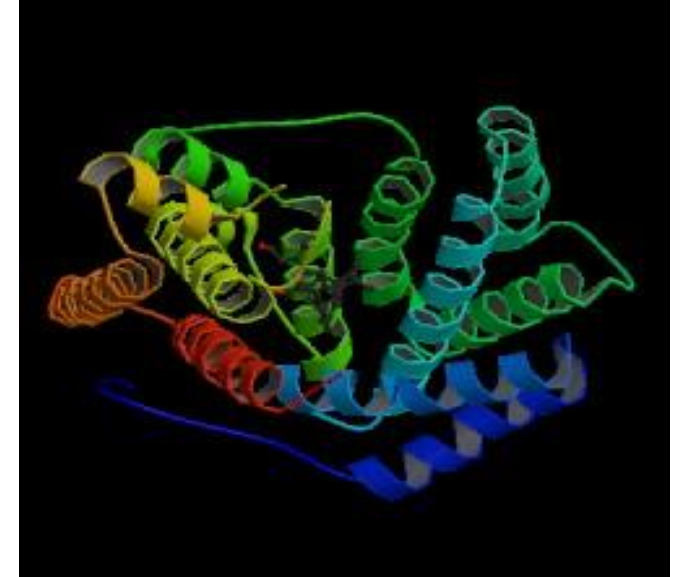
→ pentalenolactone (antibiotic)



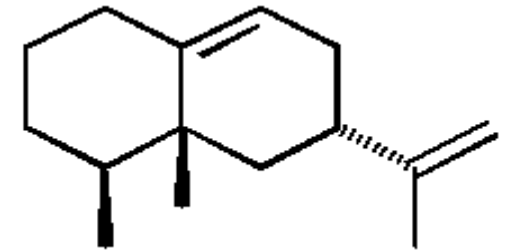
**5-*epi*-aristolochene synthase**



→ fungal (myco)toxins (*e.g.* bipolaroxin, PR-toxin)



**aristolochene synthase**

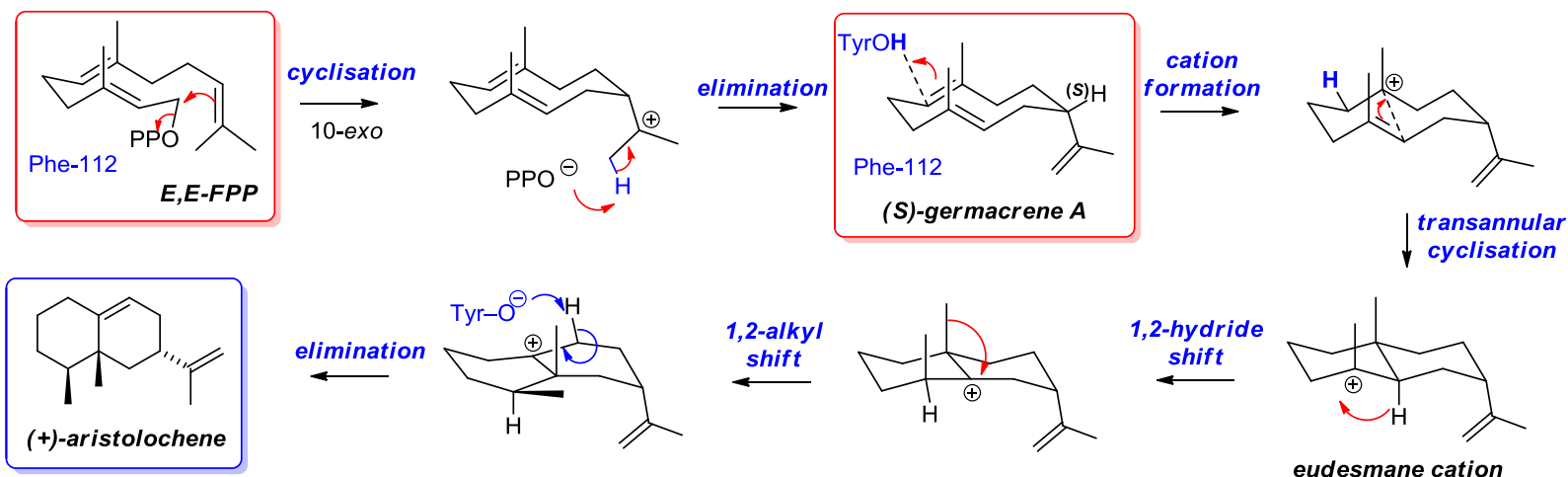




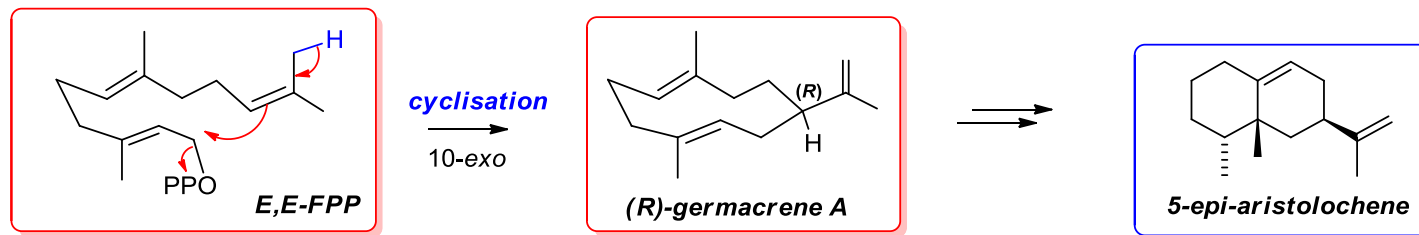
# Aristolochene & 5-*Epi*-Aristolochene Synthases

- molecular modelling studies indicate that the **shape of the active sites** determines the **conformation of FPP** and thus the **stereochemistry** of the final product

– *Penicillium roqueforti* **aristolochene synthase** – Felicetti & Cane *J. Am. Chem. Soc.* **2004**, 126, 7212 ([DOI](#))

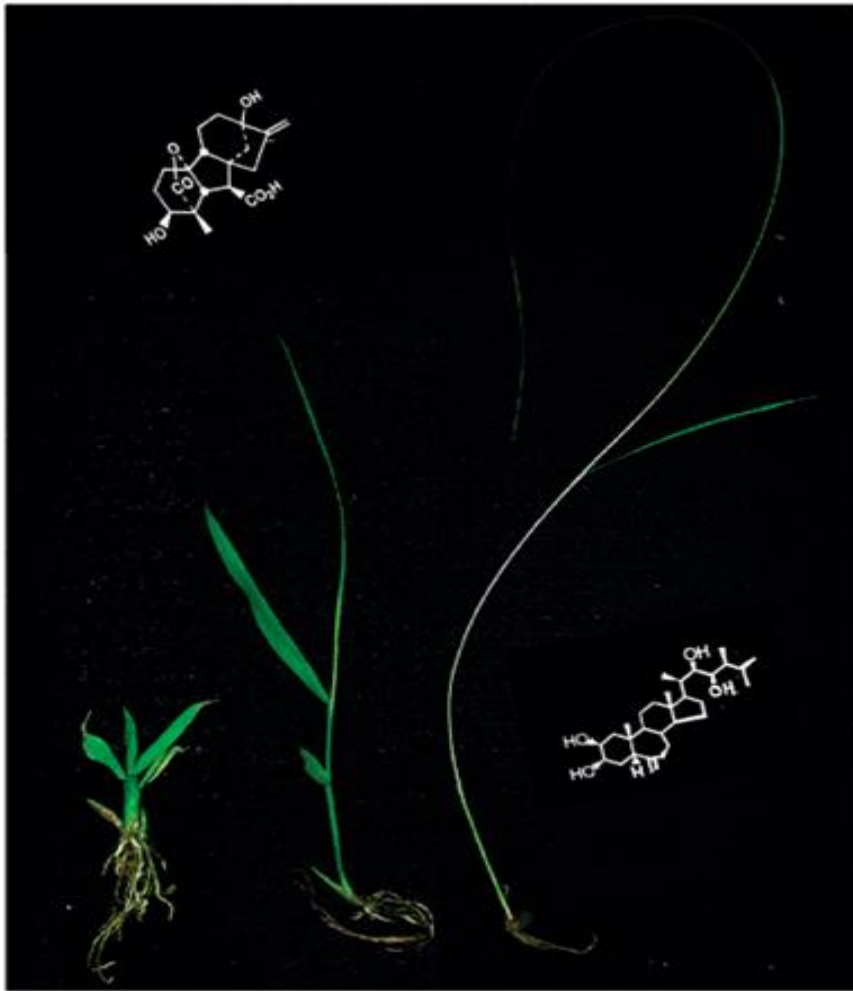


– tobacco **5-*epi*-aristolochene synthase** – Starks, Back, Chappell & Noel *Science* **1997**, 277, 1815 ([DOI](#))



*held by enzyme active site in different conformations from above*

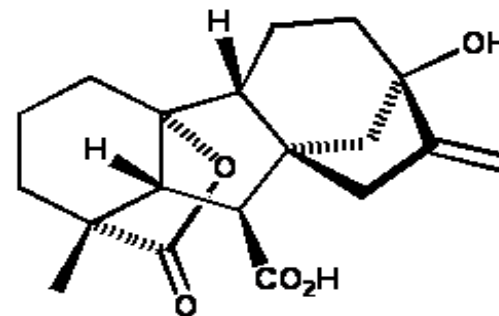
# Diterpenes - *Gibberellins*



effects of gibberellin A<sub>1</sub> and brassinolide on rice seedlings



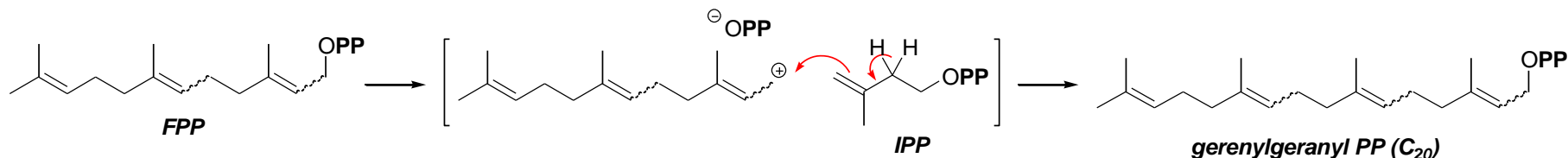
Dwarf rice seedlings (on the left) have a defect in the gibberellin-dependent signalling mechanism



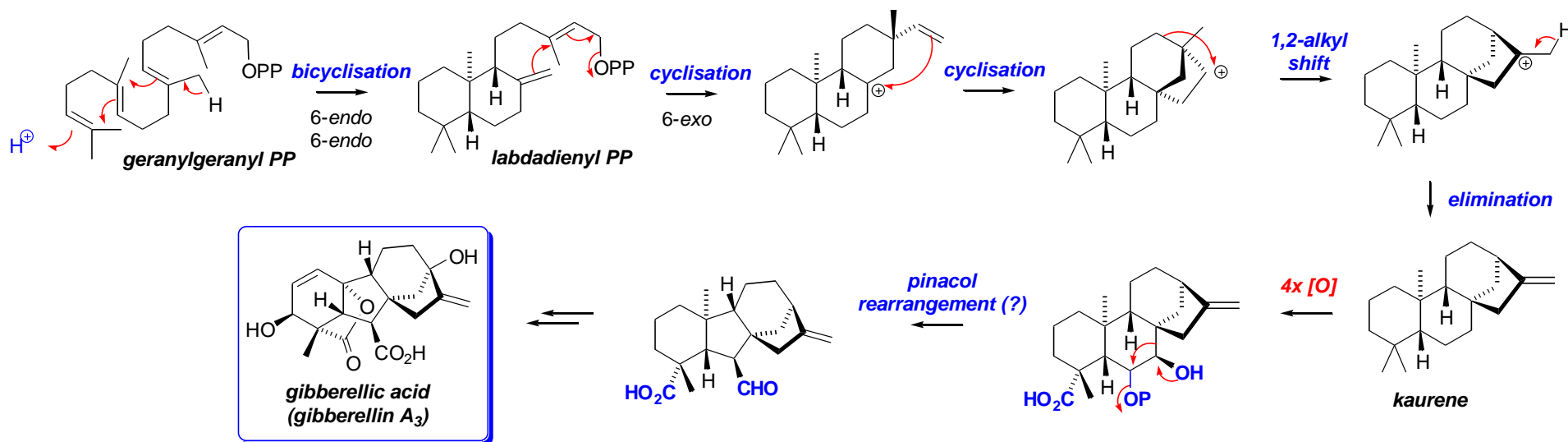
gibberellin A<sub>20</sub>

# Diterpenes – Geranylgeranyl Pyrophosphate

- $S_N2$  alkylation of farnesyl PP by IPP gives geranylgeranyl PP:

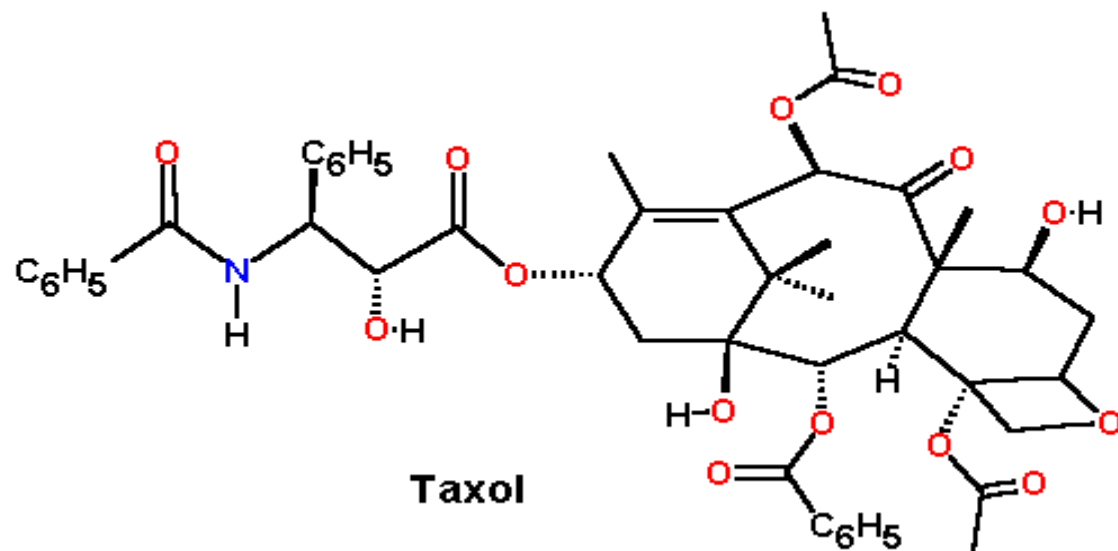
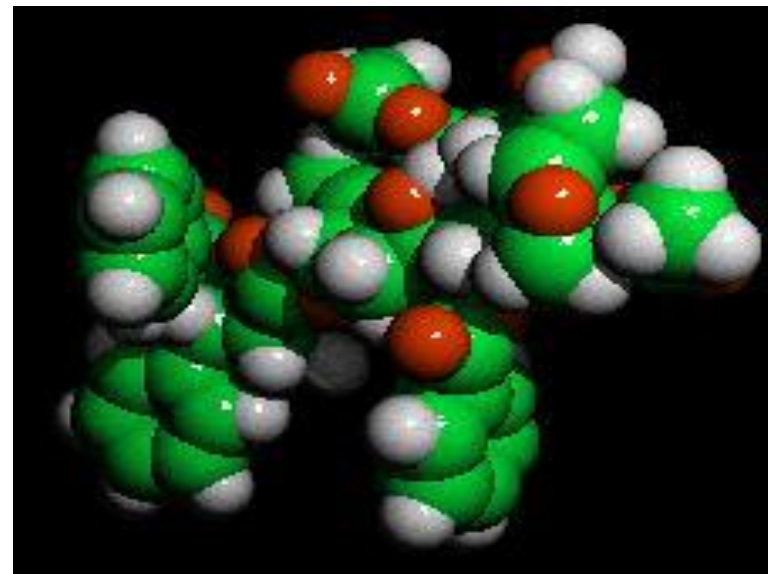


- geranylgeranyl PP readily cyclises to give numerous multicyclic diterpenes
  - e.g. gibberellins – plant growth hormones
    - NB. cyclisation initiated by alkene protonation NOT loss of PPO<sup>-</sup>



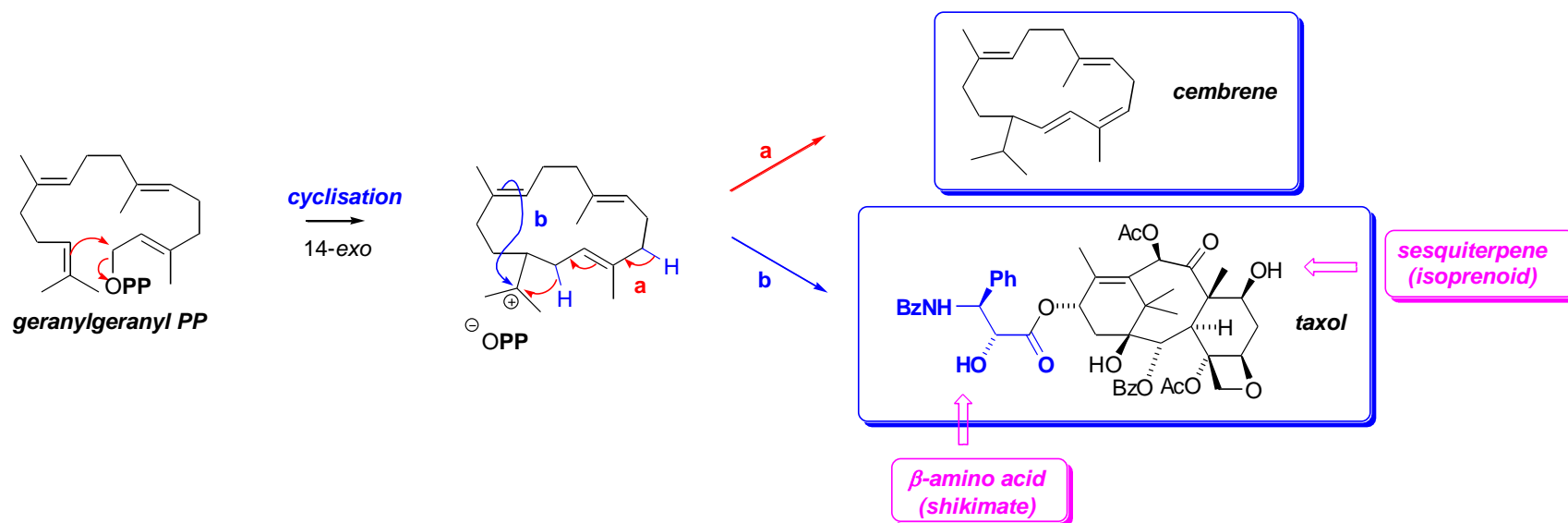
- **review:** L.N. Mander 'Twenty years of gibberellin research' *Nat. Prod. Rep.* **2003**, 20, 49-69 ([DOI](#))

# Diterpenes - *Taxol*



# Diterpenes – Geranylgeranyl PP → Taxol

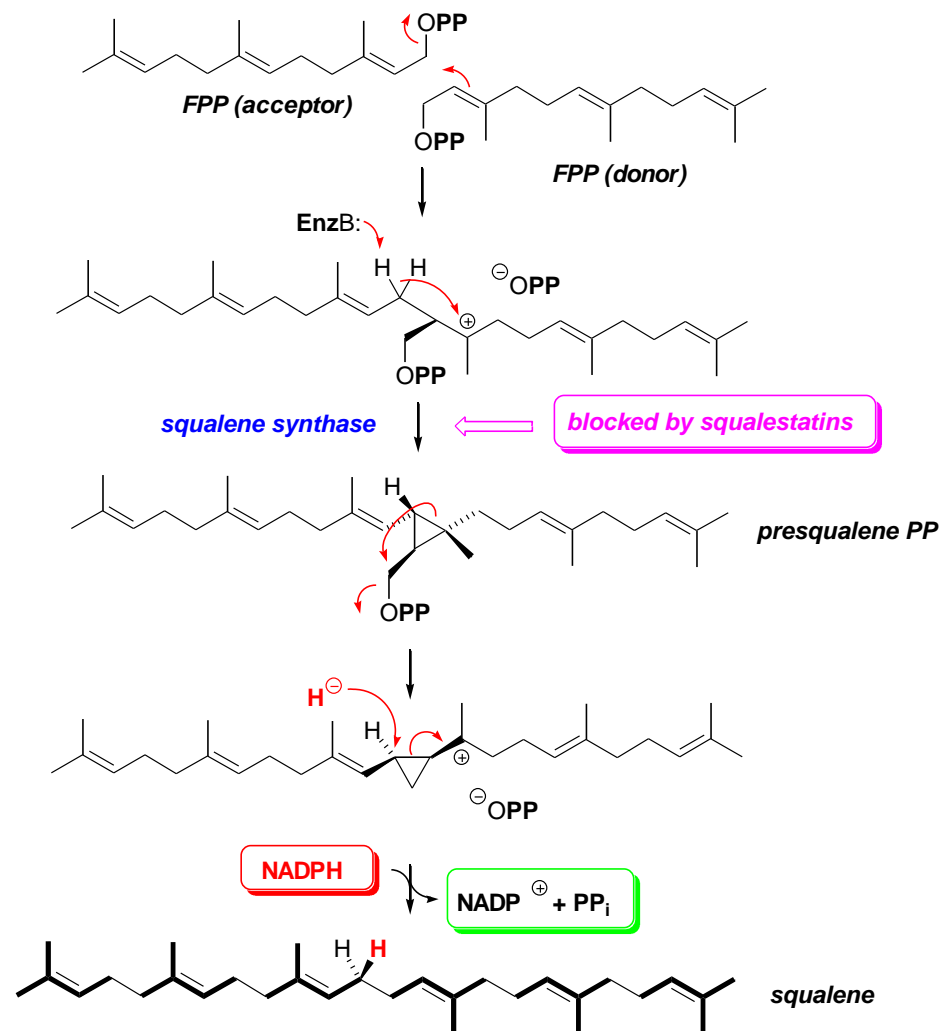
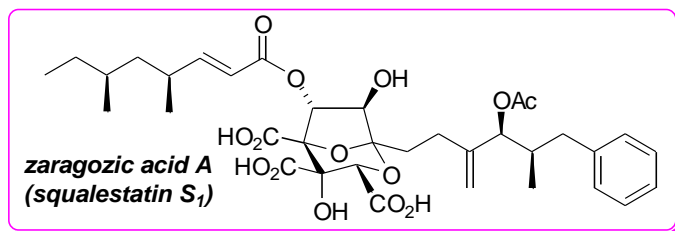
- **Taxol** is a potent **anti-cancer agent** used in the treatment of **breast & ovarian cancers**
  - comes from the bark of the **pacific yew** (*Taxus brevifolia*)
  - binds to tubulin and interferes with the assembly of microtubules
- biosynthesis is from **geranylgeranyl PP**:



- for details see: <http://www.chem.qmul.ac.uk/iubmb/enzyme/reaction/terp/taxadiene.html>
- home page is: <http://www.chem.qmul.ac.uk/iubmb/enzyme/>
  - recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzyme-Catalysed Reactions
  - based at Department of Chemistry, Queen Mary University of London

# Triterpenes – *FPP* → *Squalene*

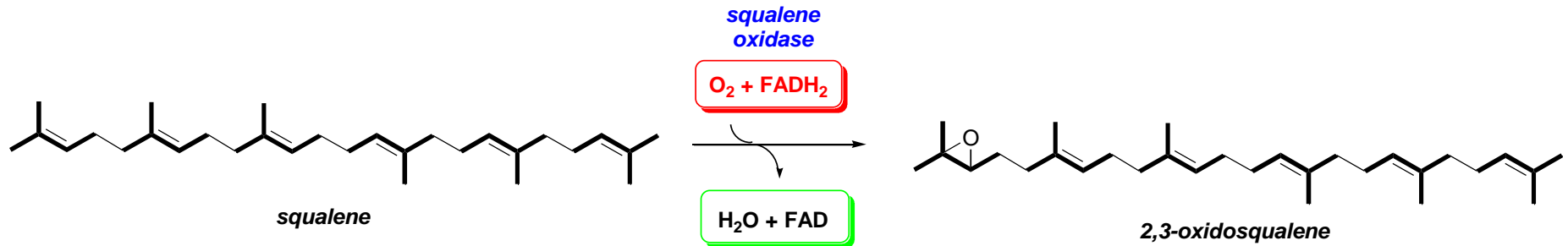
- **triterpenes** ( $C_{30}$ ) arise from the ‘**head to head**’ coupling of **two farnesyl PP** units to give **squalene** catalysed by **squalene synthase**:
  - squalene was first identified as a steroid precursor from **shark liver oil**
  - the dimerisation proceeds *via* an unusual mechanism involving electrophilic cyclopropane formation - rearrangement to a tertiary cyclopropylmethyl cation and reductive cyclopropane ring-opening by NADPH (NB. exact mechanism disputed)
  - **Zaragozic acids (squalestatins)** mimic a rearrangement intermediate and inhibit squalene synthase. They constitute interesting leads for development of new treatments for **hypercholesterolemia & heart disease** (*cf.* statins)



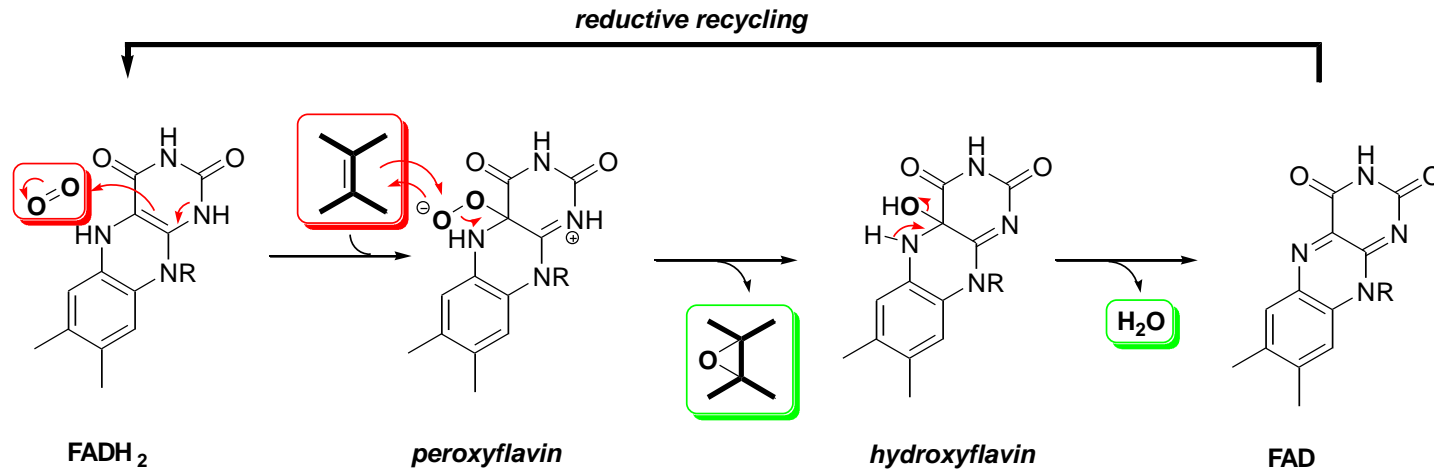
For an interesting account of the elucidation of this pathway see:  
Poulter *J. Org. Chem.* **2009**, *74*, 2631 ([DOI](#))

# Triterpenes – Squalene → 2,3-Oxidosqualene

- squalene* is oxidised to *2,3-oxidosqualene* by *squalene oxidase* – which is an  $O_2/FADH_2$ -dependent enzyme:

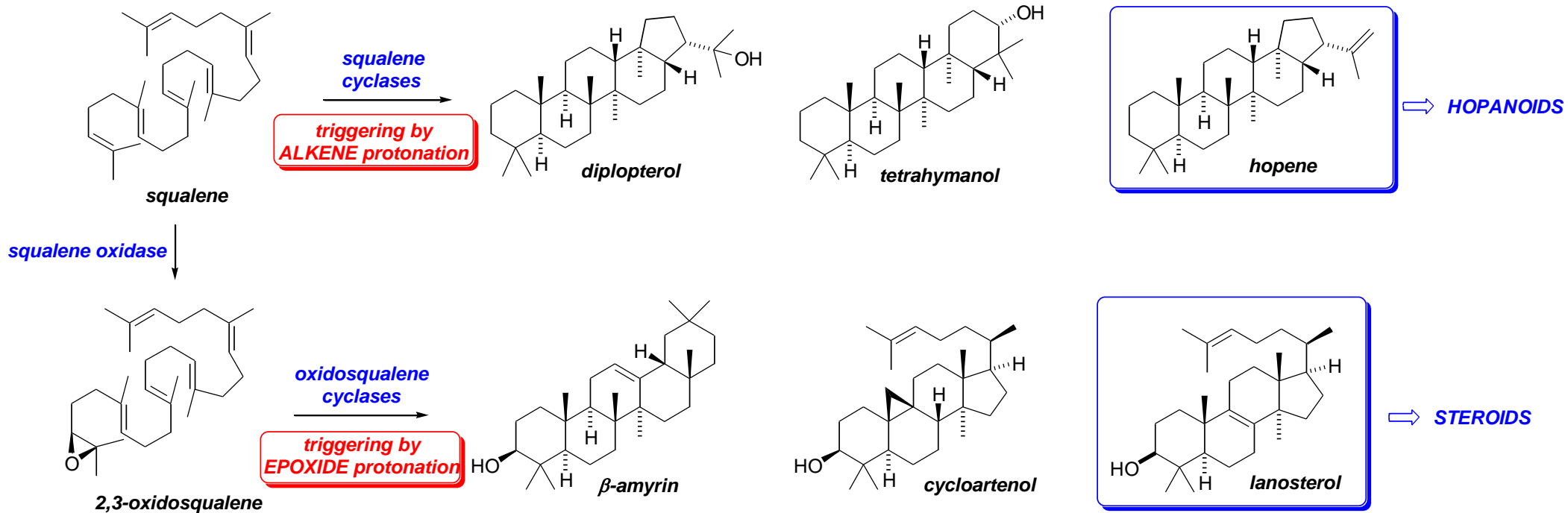


- the key oxidant is therefore a *peroxyflavin*:



# Modes of Cyclisation of Squalene


- all triterpenes (steroids, hopanoids etc.) are formed by the action of **cyclase enzymes** on either **squalene** or **2,3-oxidosqualene**
  - i.e. different methods of **'triggering'** cyclisation




## STEROID/TRITERPENE NOMENCLATURE:

$\beta$ -face = top face (as drawn here)

$\alpha$ -face = bottom face (as drawn here)

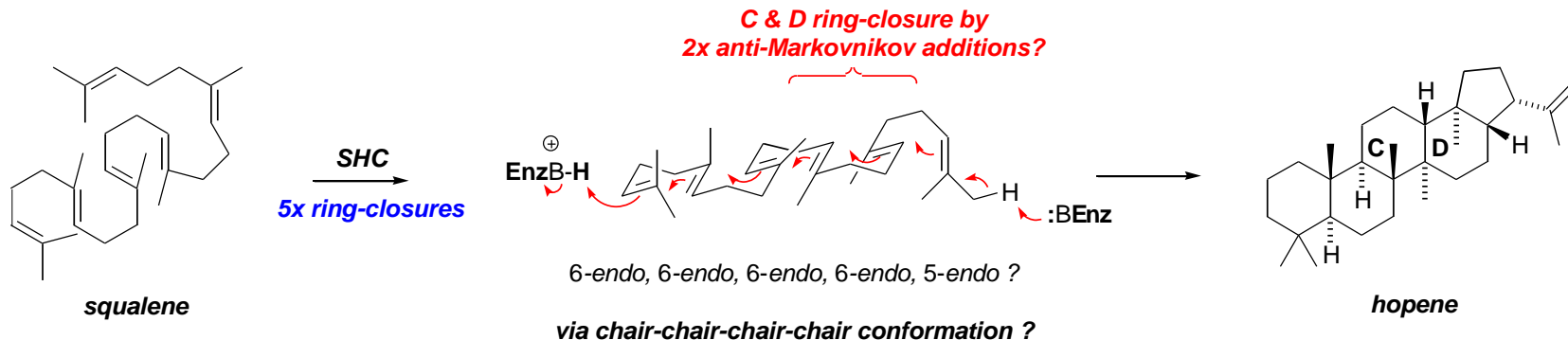
 = hydrogen up (on  $\beta$ -face)

 = hydrogen down (on  $\alpha$ -face)

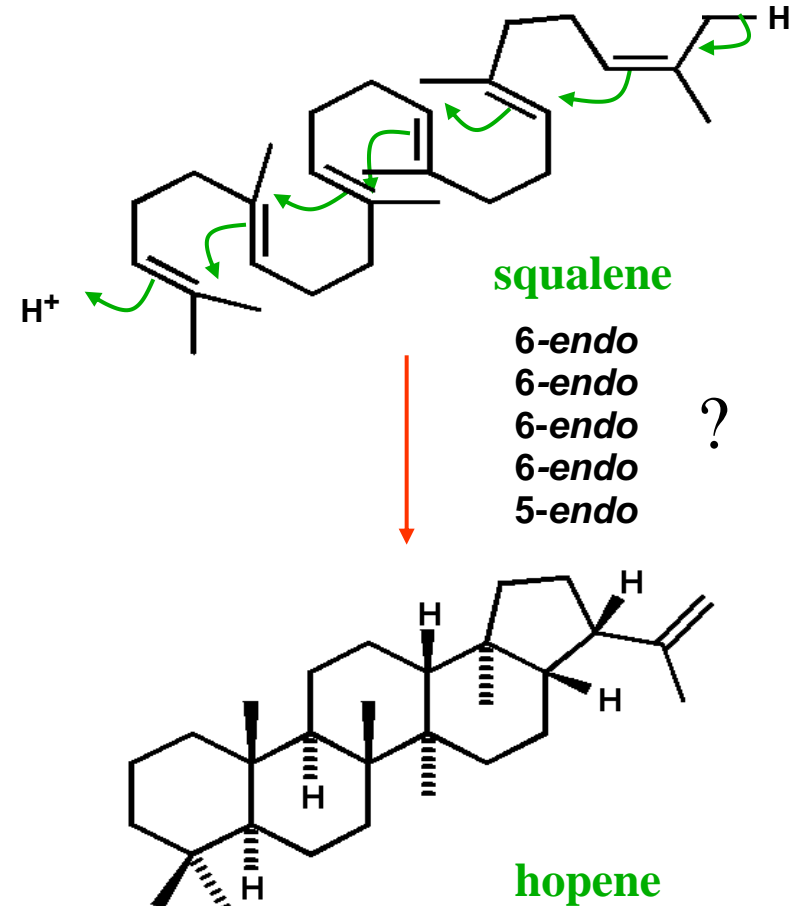
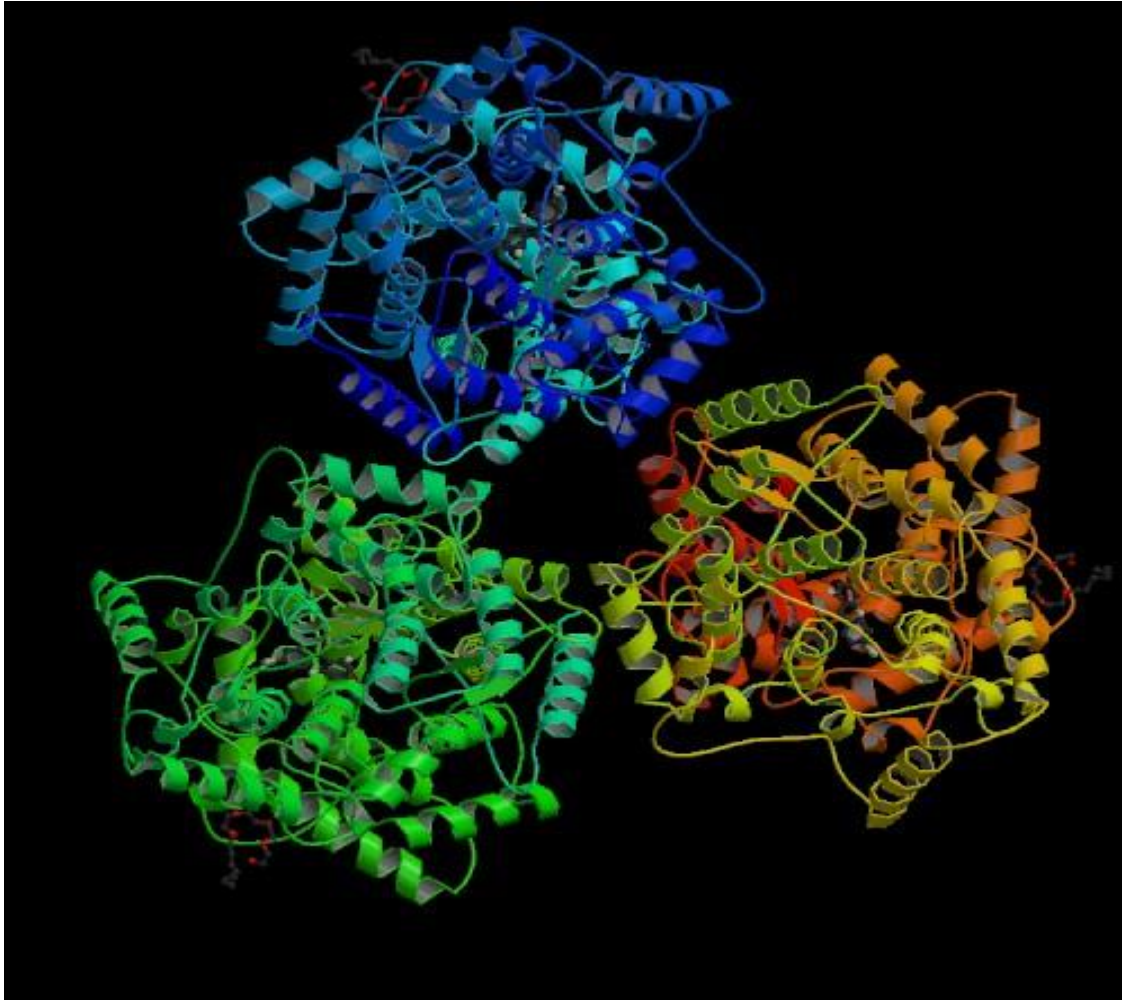


# Squalene-Hopene Cyclase (SHC)

- **Squalene-Hopene Cyclase (SHC)** catalyses the formation of **hopene** from **squalene**:
  - what does the enzyme have to do to achieve such exquisite **regio-** & **stereoselectivity** over the formation of 9 new stereogenic centres?
    - **enforce** an appropriate **conformation** of squalene
    - **activate** the **C2 alkene** by **protonation**
    - **shield** reactive **cations** from nucleophiles (e.g. H<sub>2</sub>O) using aromatic residues (**cation- $\pi$** )
    - **position** a **general base** precisely to facilitate the **terminal elimination**
  - until recently, the ‘appropriate’ conformation was believed to be the formally appealing ‘**all-chair**’ conformation allowing for a **concerted cationic ring-closure cascade** (shown below)
  - remaining stereocontrol would then be taken care of by **intrinsic stereoelectronics**
    - *i.e. correct orbital overlap*
  - however, this requires **anti-Markovnikov regioselectivity** for the **C & D ring-closures**...



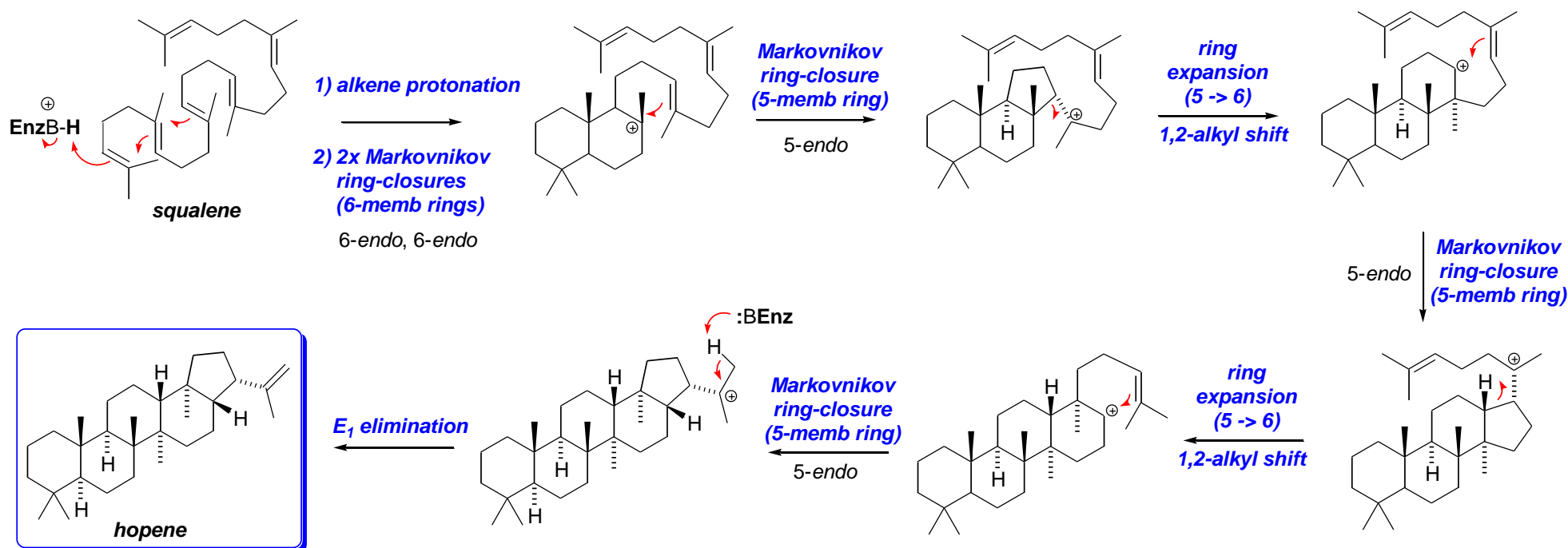
# Squalene-Hopene Synthase (SHC)



X-ray crystal structure: Wendt, Poralla & Schulz *Science*, 1997, 277, 1811 ([DOI](#))

# Squalene-Hopene Synthase - Mechanism

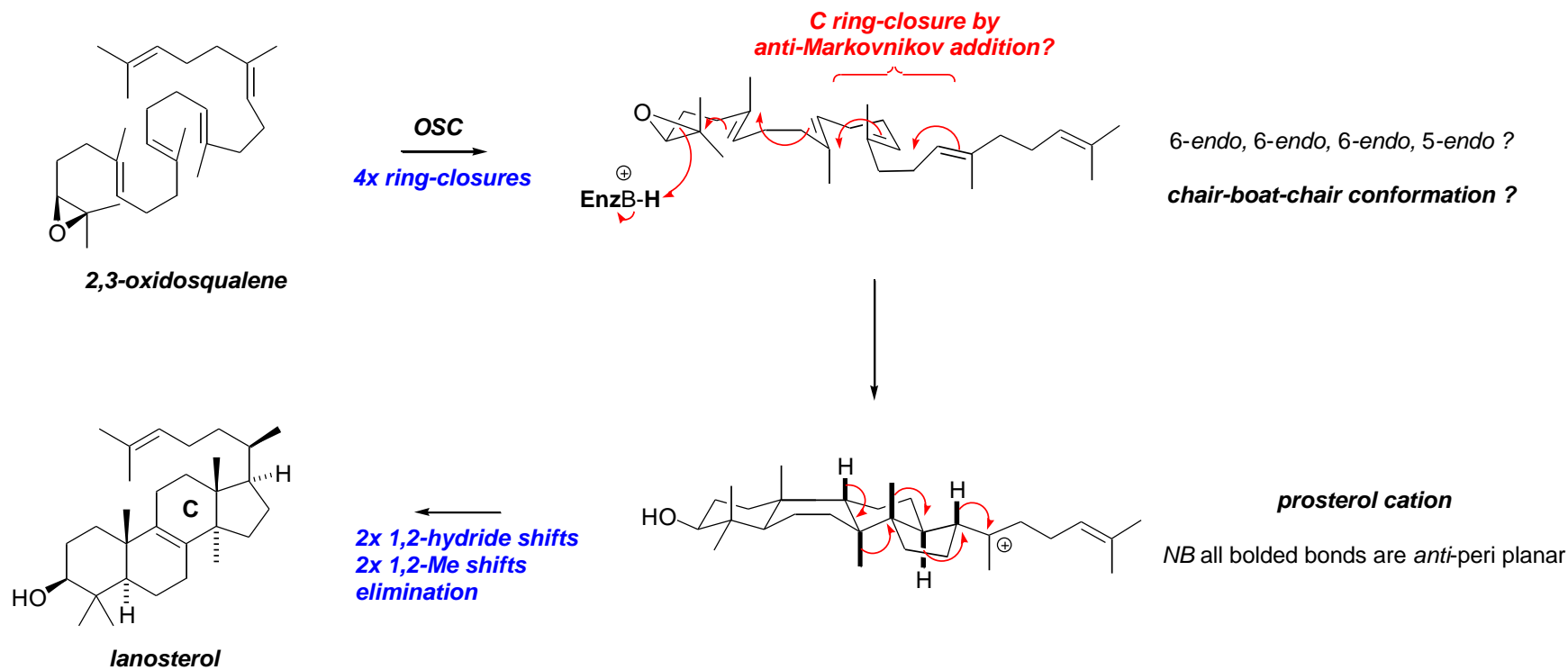
- The process is apparently more complex & has more recently been shown to involve:
  - **2x Markovnikov ring-closure/1,2-alkyl-shift ring-expansion sequences** to establish the **C & D rings**
  - lessons?
    - the **conformation** enforced by the enzyme is **NOT** strictly an **all-chair** one! (although probably very close)
    - the process is **NOT concerted**, discrete **cationic intermediates** are involved
    - **stereoelectronics dictate** the **regio- & stereoselectivity**



- **review:** Wendt *et al. Angew. Chem. Int. Ed.* **2000**, 39, 2812 ([DOI](#)) & Wendt *ibid* **2005**, 44, 3966 ([DOI](#))

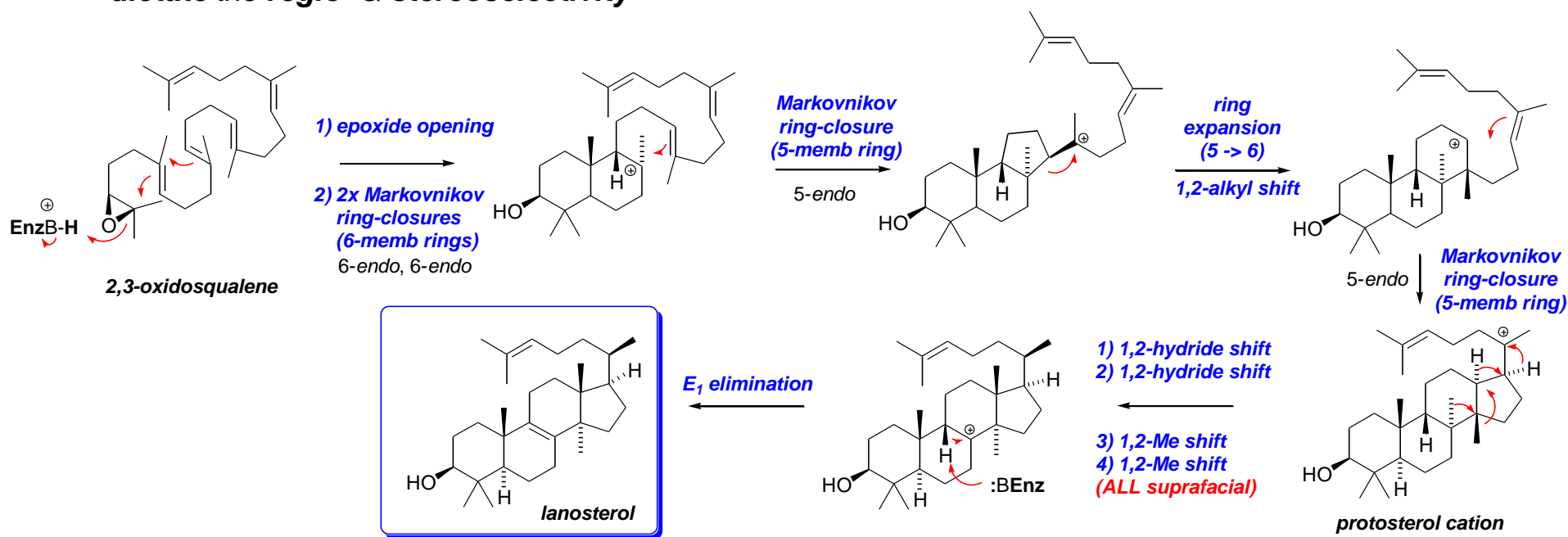
# Oxidosqualene-Lanosterol Cyclase (OSC)

- **oxidosqualene-lanosterol cyclase** catalyses the formation of **lanosterol** from **2,3-oxidosqualene**:
  - this cascade establishes the characteristic ring system of **ALL steroids**
  - until recently, as for SHC, the enzyme was believed to enforce a **chair-boat-chair conformation** to allow a **concerted cationic ring-closure cascade** followed by a series of **suprafacial 1,2-shifts** (shown below)
  - however, this also requires **anti-Markovnikov regioselectivity** for the **C ring-closure**...



# Oxidosqualene-Lanosterol Cyclase – Mechanism

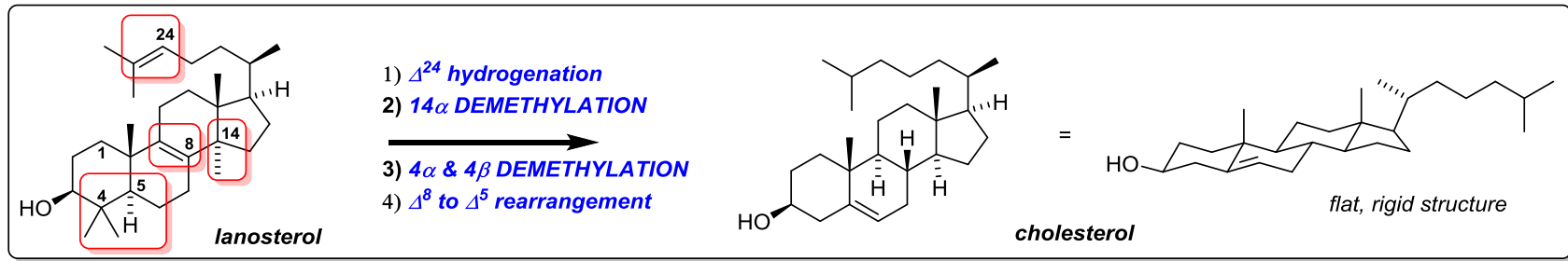
- This process has also been shown to involve a **Markovnikov ring-closure/1,2-alkyl-shift ring-expansion sequence** to establish the **C ring**
  - again, the **conformation** enforced by the enzyme is **NOT strictly a chair-boat-chair** one (although probably close), the process is **NOT concerted**, discrete **cationic intermediates** are involved & **stereoelectronics dictate** the **regio- & stereoselectivity**



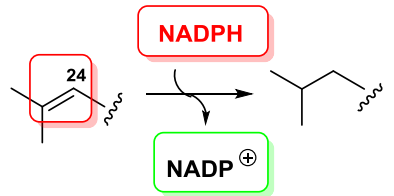
- “The enzyme’s role is most likely to shield intermediate carbocations... thereby allowing the hydride and methyl group migrations to proceed down a thermodynamically favorable and kinetically facile cascade”
  - Wendt *et al.* *Angew. Chem. Int. Ed.* **2000**, 39, 2812 ([DOI](#)) & Wendt *ibid* **2005**, 44, 3966 ([DOI](#))

# Lanosterol $\rightarrow$ Cholesterol – Oxidative Demethylation

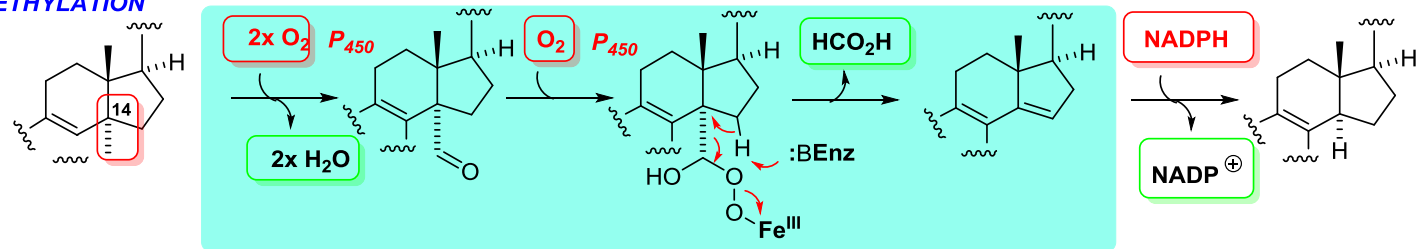
- Several steps are required for conversion of *lanosterol* to *cholesterol*:



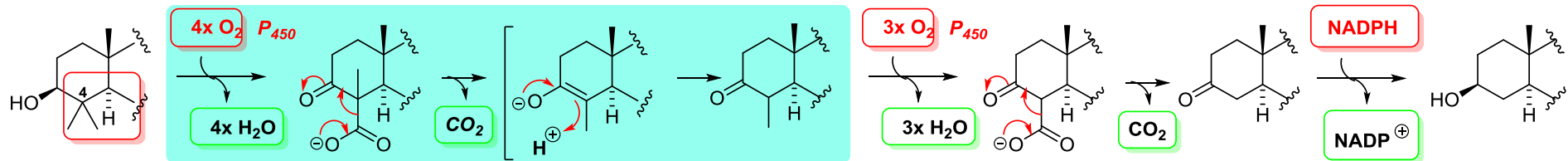
## 1) $\Delta^{24}$ hydrogenation



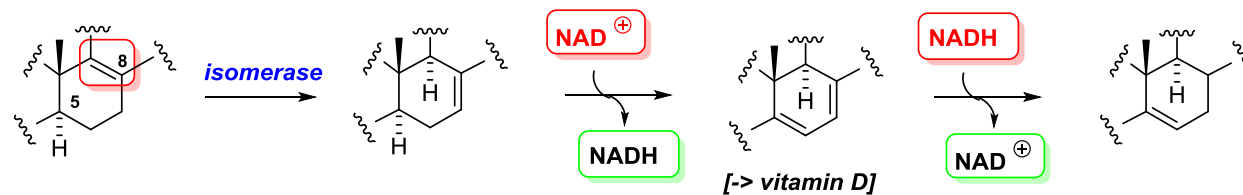
## 2) $14\alpha$ DEMETHYLATION



## 3) $4\alpha$ & $4\beta$ DEMETHYLATION

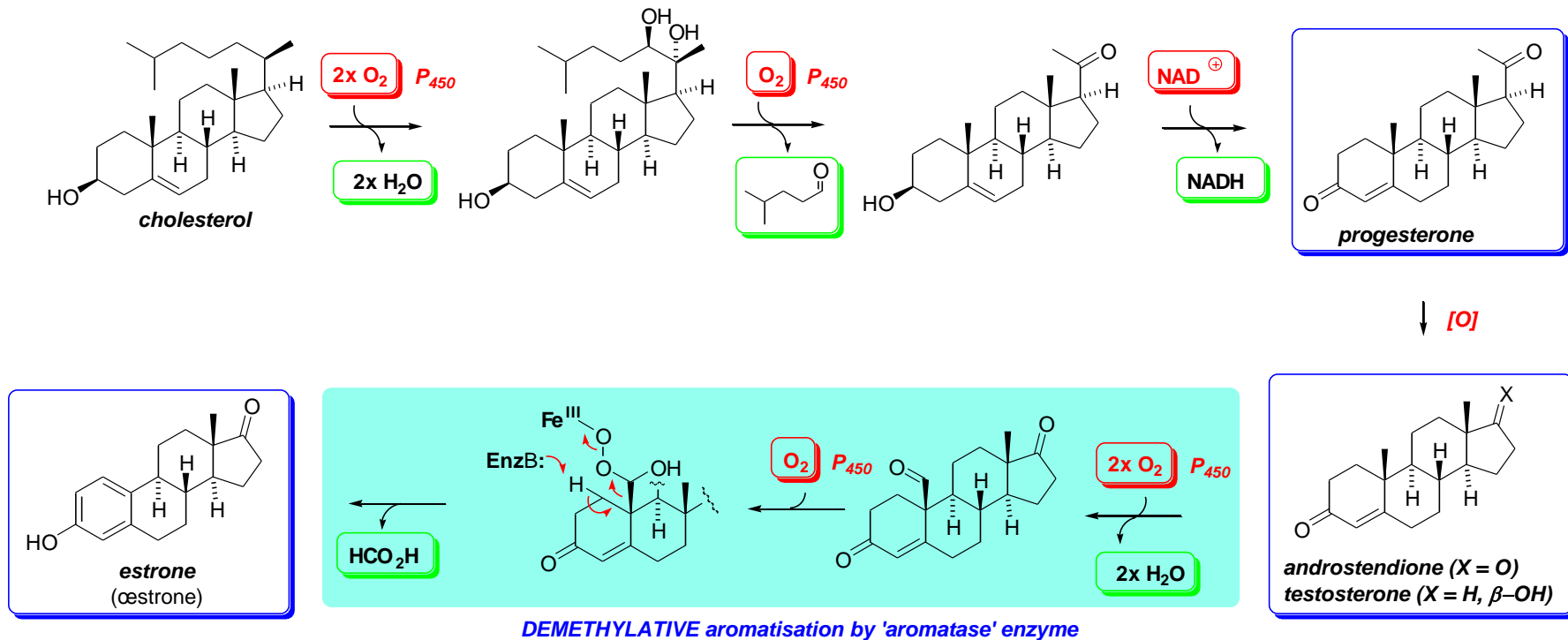


## 4) $\Delta^8$ to $\Delta^5$ rearrangement



# Cholesterol → Human Sex Hormones

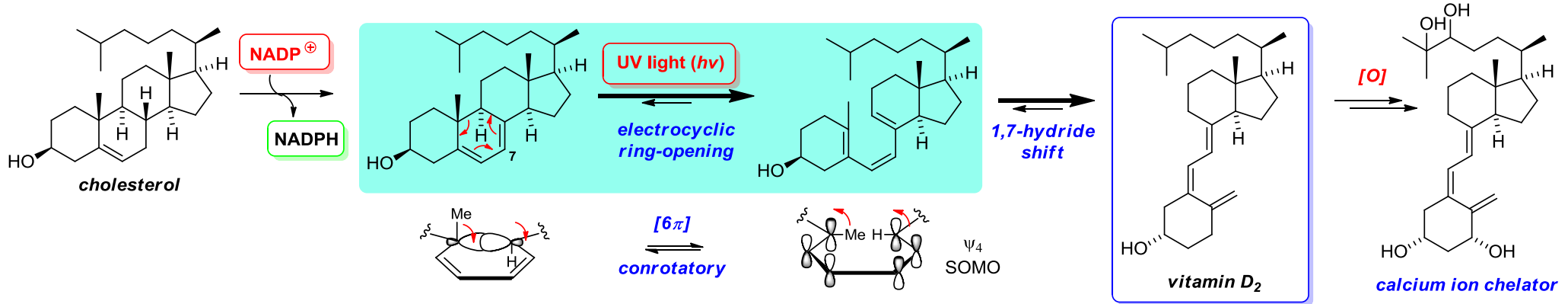
- **cholesterol** is the precursor to the human sex hormones – **progesterone**, **testosterone** & **estrone**
  - the pathway is characterised by **extensive oxidative processing** by  $P_{450}$  enzymes
  - **estrone** is produced from **androstendione** by **oxidative demethylation** with **concomitant aromatisation**:



*NB.* The involvement of a peroxyacetal during aromatase demethylation has recently been disputed, see: Guengerich *J. Am. Chem. Soc.* **2014**, 136, 15036 ([DOI](#)).

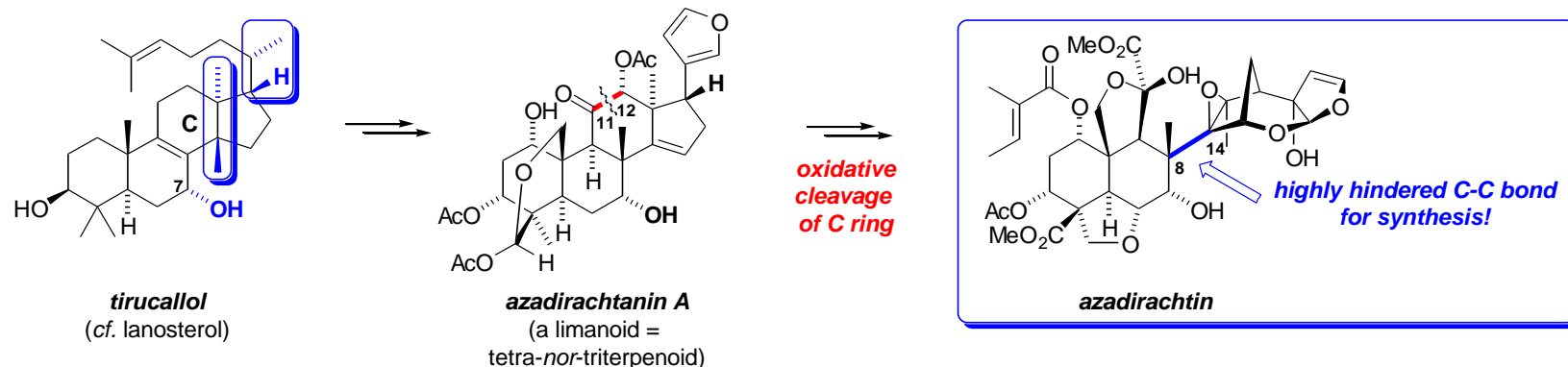
# Steroid Ring Cleavage - Vitamin D & Azadirachtin

- **vitamin D<sub>2</sub>** is biosynthesised by the **photolytic cleavage** of  **$\Delta^7$ -dehydrocholesterol** by UV light:
  - a classic example of **photo-allowed, conrotatory electrocyclic ring-opening**:



- D vitamins are involved in **calcium absorption**; **deficiency** leads to **rickets** (brittle/deformed bones)

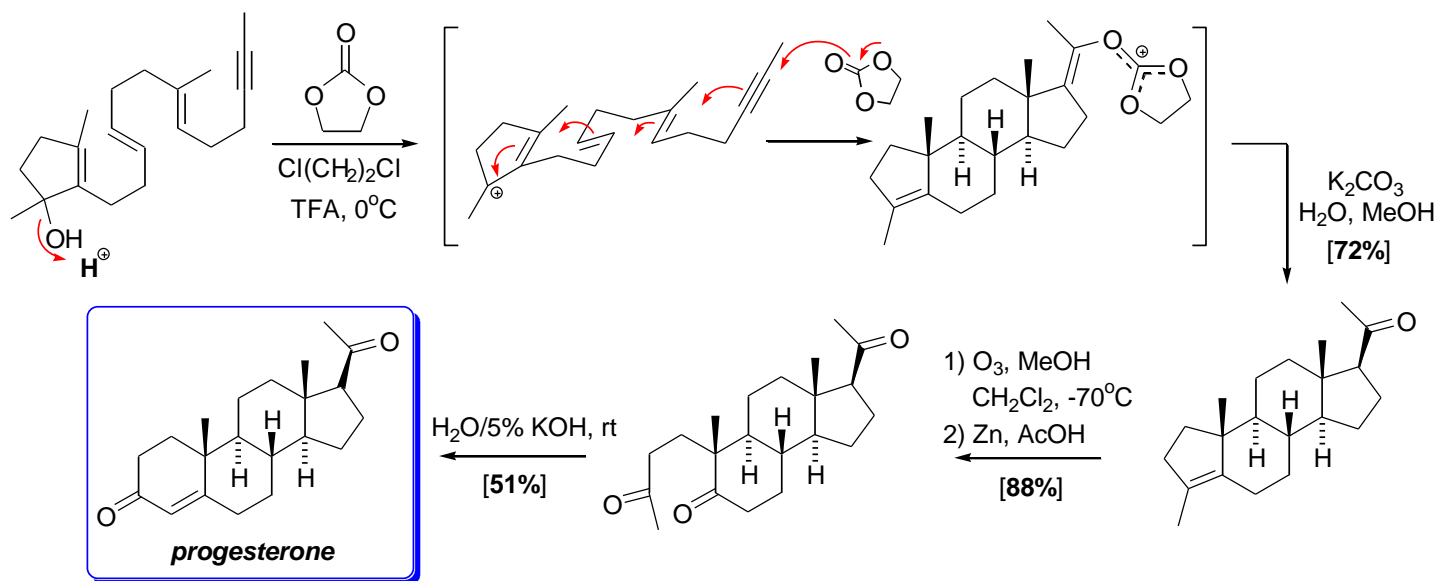
- **Azadirachtin** is a potent **insect anti-feedant** from the Indian **neem tree**:
  - exact biogenesis unknown but certainly **via steroid modification**:





# Biomimetic Cationic Cyclisations - *Progesterone*

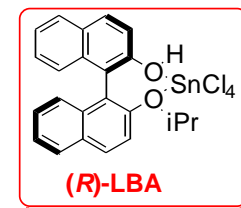
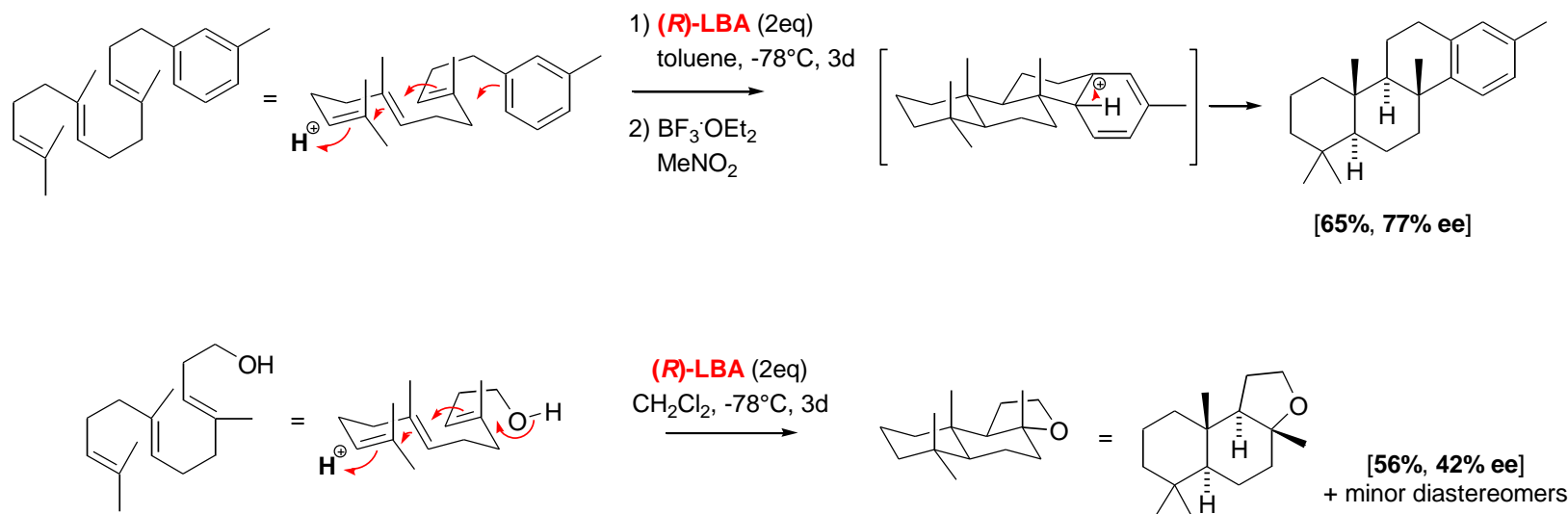
- in **1971**, **W.S. Johnson** utilized a biomimetic polyolefin cyclization in a pioneering & elegant total synthesis of the hormone **progesterone**
  - the substrate's preference for the '**chair-chair-chair**' **conformation** provided the progesterone core with impressive stereoselectivity
  - the cascade was **initiated** by **protonation** of a **tert-alcohol**



- Johnson, Gravestock & McCarry *J. Am. Chem. Soc.* **1971**, 93, 4332 ([DOI](#))

# Biomimetic Cationic Cyclisations – *Enantioselective*

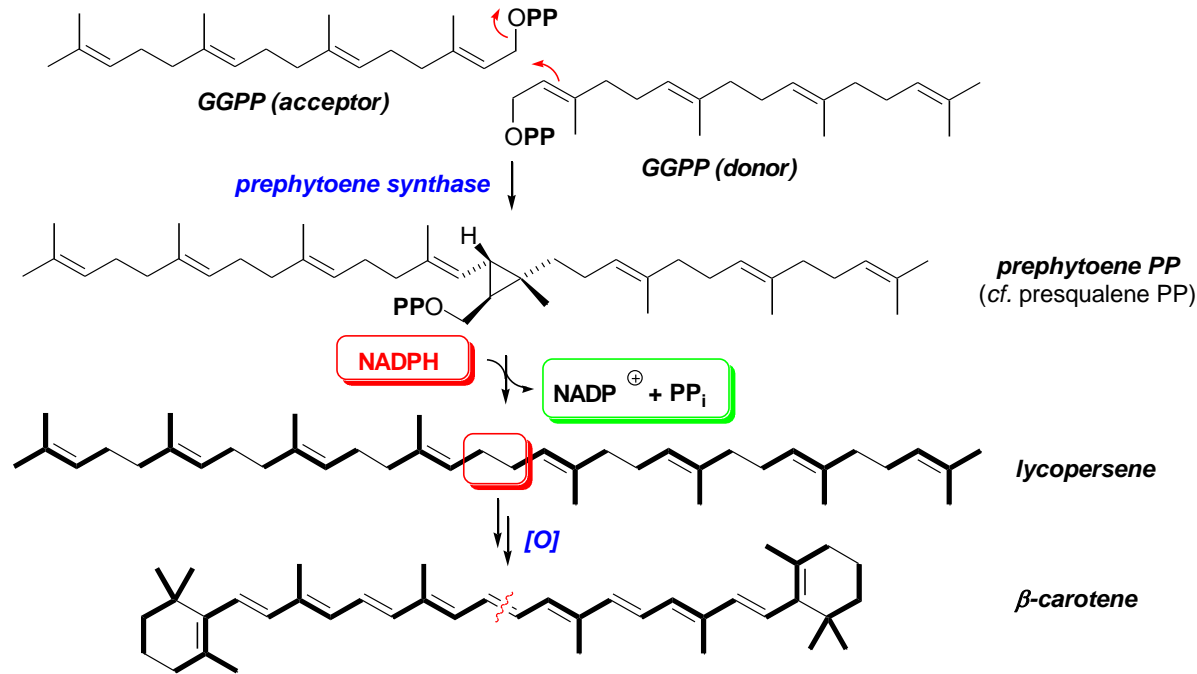
- **Yamamoto** has achieved several **enantioselective** cationic cascade cyclisations using a chiral 'Lewis acid assisted Brønsted acid' (LBA) prepared by mixing **binol** & **SnCl<sub>4</sub>**:



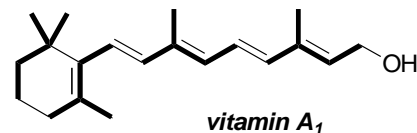
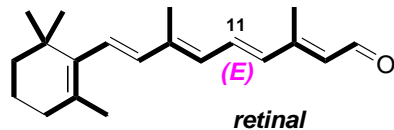
- Yamamoto *et al.* *J. Am. Chem. Soc.*, **1999**, 121, 4906 ([DOI](#))
- Yamamoto *et al.* *J. Am. Chem. Soc.* **2001**, 123, 1505 ([DOI](#))

# Carotenoids – $\beta$ -Carotene & vitamin A<sub>1</sub>

- **Carotenoids** (C<sub>40</sub>) are coloured pigments made by photosynthetic plants & certain algae, bacteria & fungi. Dietary ingestion by birds and further processing gives rise to bright feather pigments etc.
  - biosynthesised by **head-to-head** coupling of two **geranylgeranyl PP units** to give **lycopersene**:



- subsequent oxidative degradation (cf. ozonolysis!) gives **retinal** (mediator of vision) & **vitamin A<sub>1</sub>**:



# Primary Metabolism - Overview

