An Overview of Biosynthesis Pathways – Inspiration for Pharmaceutical and Agrochemical Discovery

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

19th Oct 2019
Lessons in Synthesis - Azadirachtin

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

  ![Chemical Structures](image)

  - *tirucallol* (cf. lanosterol)
  - *azadirachtin A* (a limanoid = tetra-nor-triterpenoid)
  - oxidative cleavage of C ring

- Intense synthetic efforts by the groups of Nicolaou, Watanabe, Ley and others since structural elucidation in 1987.
- 1st total synthesis achieved in 2007 by Ley following 22 yrs of effort
- ~40 researchers and over 100 person-years of research! – 64-step synthesis

Format & Scope of Presentation

- **Metabolism & Biosynthesis**
  - some definitions, 1° & 2° metabolites

- **Shikimate Metabolites**
  - photosynthesis & glycolysis → shikimate formation → shikimate metabolites
  - *Glyphosate* – a non-selective herbicide

- **Alkaloids**
  - acetylCoA & the citric acid cycle → α-amino acids → alkaloids
  - *Opioids* – powerful pain killers

- **Fatty Acids and Polyketides**
  - acetylCoA → malonylCoA → fatty acids, prostaglandins, polyketides, macrolide antibiotics
  - *NSAIDs* – anti-inflammatory’s

- **Isoprenoids/terpenes**
  - acetylCoA → mevalonate → isoprenoids, terpenoids, steroids, carotenoids
  - *Statins* – cholesterol-lowering agents
Metabolism and Biosynthesis
Metabolism & Natural Product Diversity

- camphor
- clavulanic acid
- caffeine
- lysergic acid
- nicotine
- patulin
- quinine
- androstenedione

\[ \text{CO}_2 \ \text{H}_2\text{O} \ \text{P}_i \ \text{N}_2 \ \text{hv} \]
Metabolism

- **Metabolism** is the term used for *in vivo* processes by which compounds are degraded, interconverted and synthesised:
  - **Catabolic** or **degradative**: primarily to release energy and provide building blocks
    - generally **oxidative** processes/sequences (glycolysis, Krebs cycle)
  - **Anabolic** or **biosynthetic**: primarily to create new cellular materials (1° & 2° metabolites)
    - generally **reductive** processes/sequences
- These two types of process are coupled – one provides the driving force for the other:

\[
\text{CO}_2 + \text{H}_2\text{O} \quad \text{hv} \quad \text{hv} \\
\text{photosynthesis} \quad \text{energy storage} \quad \text{ATP} \\
\text{NAD(P)H} \quad \text{NAD(P) + H} \quad \text{energy release} \\
\text{simple products} \quad \text{complex metabolites} \\
\]

Nutrients "\text{CO}_2\text{ fixation}" (photosynthesis) Energy & Building blocks

Simple products

Cell components & Growth "\text{nitrogen fixation}" (by diazatrophs, lightning, the Haber process)

Building Blocks
Primary Metabolism - Overview

**Primary metabolism**

\[ \text{CO}_2 + \text{H}_2\text{O} \]

1) 'light reactions': \( hv \rightarrow \text{ATP} \) and \( \text{NADPH} \)

2) 'dark reactions': \( \text{CO}_2 \rightarrow \text{sugars} \) (Calvin cycle)

**Primary metabolites**

- Oligosaccharides
- Polysaccharides
- Nucleic acids (RNA, DNA)

**Secondary metabolites**

- Shikimate metabolites
  - Cinnamic acid derivatives
  - Aromatic compounds
  - Lignans, flavinoids

- Alkaloids
  - Penicillins
  - Cephalosporins
  - Cyclic peptides

- Fatty acids & polyketides
  - Prostaglandins
  - Polyacetylenes
  - Aromatic compounds, polyphenols
  - Macrolides

- Isoprenoids
  - Terpenoids
  - Steroids
  - Carotenoids

For interesting animations' of e.g. photosynthesis see: [http://www.johnkyrk.com/index.html](http://www.johnkyrk.com/index.html)
Shikimate Metabolites
Shikimate Metabolites

- (S)-tryptophan ($\text{ArC}_9$)
- (S)-phenylalanine ($\text{ArC}_9$)
- (S)-tyrosine ($\text{ArC}_9$)
- scopoletin ($\text{ArC}_3$)
- menaquinone (vitamin $K_2$) ($\text{ArC}_7$)
- $\alpha$-tocopherol (vitamin $E$) ($\text{ArC}_7$)
- epigallocatechin (EGC) ($\text{ArC}_7$)
- podophyllotoxin ($\text{ArC}_3$)
The Shikimate Biosynthetic Pathway - Overview

- **Phosphoenol pyruvate & erythrose-4-phosphate → shikimate → chorismate → prephenate:**

  ![Chemical Diagram]

  - The detailed mechanisms of these steps have been studied intensively. Most are chemically complex and interesting. For additional details see:
    - Mann *Chemical Aspects of Biosynthesis* Oxford Chemistry Primer No. 20, 1994 (key details)
    - Haslam *Shikimic Acid – Metabolism and Metabolites* Wiley, 1993 (full details and primary Lit. citations)
    - [http://www.chem.qmul.ac.uk/iubmb/enzyme/reaction/misc/shikim.html](http://www.chem.qmul.ac.uk/iubmb/enzyme/reaction/misc/shikim.html) (interesting web-site with many biosynthetic pathways)
Rational Agrochemical Development – *Shikimate Pathway Intervention*

- **The shikimate biosynthetic pathway is not found in animals/humans – only in plants**
  - selective intervention in these pathways allows development of agrochemicals with minimal human toxicity

- **Glyphosate ('Roundup')** – a Monsanto agrochemical is a potent inhibitor of the conversion of 3-phosphoshikimate (3-PS) → 5-enolpyruvylshikimate-3-phosphate (5-EPS-3P)
  - a non-selective herbicide

\[
\begin{align*}
 &\text{phosphoenol pyruvate (PEP)} \\
 &\text{erythrose-4-phosphate (E-4-P)} \\
 &\text{3-phosphoshikimate (3-PS)} \\
 &\text{5-enolpyruvylshikimate-3-phosphate (5-EPS-3P)}
\end{align*}
\]

**Glyphosate (Roundup®)** inhibits this step
Chorismate $\rightarrow$ Tryptophan, Tyrosine & Phenylalanine

- **Chorismate $\rightarrow$ anthranilate $\rightarrow$ tryptophan**

- **Chorismate $\rightarrow$ prephenate $\rightarrow$ tyrosine & phenylalanine**
  - NB. The enzyme chorismate mutase [EC 5.4.99.5] which mediates the conversion of chorismate to prephenate is the only known ‘Claisen rearrangementase’
Tyrosine/Phenylalanine → ArC₃ Metabolites

- Tyrosine & phenylalanine → cinnamate derivatives → ArC₃ metabolites
  - coumarins, lignans (stereoselective enzymatic dimerisation) & lignins (stereorandom radical polymerisation)

- phenylalanine ammonia lyase (PAL)
- cinnamate derivatives
- ferulate
- scopoletin (a coumarin) germination stimulant
- podophyllotoxin (a lignan) natural product used to treat worts
- pinoresinol (a lignan)
- fragment of lignin polymer 'woody' component of cell walls

Two ArC₃

n x ArC₃
Primary Metabolism - Overview

**Primary metabolism**

1. **'light reactions':** \(hv \rightarrow \text{ATP and NADPH}\)
2. **'dark reactions':** \(\text{CO}_2 \rightarrow \text{sugars (Calvin cycle)}\)

**Primary metabolites**

- Oligosaccharides
- Polysaccharides
- Nucleic acids (RNA, DNA)

**Secondary metabolites**

- Shikimate metabolites
- Cinnamic acid derivatives
- Aromatic compounds
- Lignans, flavonoids
- Alkaloids
- Penicillins
- Cephalosporins
- Cyclic peptides

**Other metabolites**

- Tetrapyrroles (porphyrins)
- Saturated fatty acids
- Unsaturated fatty acids
- Lipids
- Saturated fatty acids
- Unsaturated fatty acids
- Lipids
- Peptides
- Proteins
- Fatty acids & polyketides
- Prostaglandins
- Polyacetylenes
- Aromatic compounds, polyphenols
- Macrolides
- Isoprenoids
- Terpenoids
- Steroids
- Carotenoids

For interesting animations’ of e.g. photosynthesis see: [http://www.johnkyrk.com/index.html](http://www.johnkyrk.com/index.html)
Alkaloids
Alkaloids

- **Definitions:**
  - *originally* – ‘a natural product that could be extracted out of alkaline but not acidic water’ (i.e. containing a basic amine function that protonated in acid)
  - *more generally* - ‘any non-peptidic & non-nucleotide nitrogenous secondary metabolite’
The Citric Acid Cycle

- **The citric acid (Krebs) cycle** is a major catabolic pathway of 1° metabolism that provides two key building blocks for aliphatic amino acid biosynthesis - oxaloacetate & α-ketoglutarate:

![Citric Acid Cycle Diagram]

**OVERAL STOICHIOMETRY**

- 1x 'acetate'
- 1x $\text{O}_2$
- 2x $\text{CO}_2$
- 12x ATP energy!
The Biosynthesis of Lysine & Ornithine

- **Lysine & ornithine** - the two most significant, *non-aromatic* α-amino acid *precursors to alkaloids*:
  - NB. lysine (Lys) is proteinogenic whereas ornithine (Orn) is not
  - phenylalanine (Phe), tyrosine (Tyr) & tryptophan (Trp) from *shikimate* are the other important precursors
  - biosynthesis is *via* reductive amination of the appropriate α-ketoacid mediated by *pyridoxal-5’-phosphate* (PLP)

\[
\begin{align*}
\text{R} & \quad \rightarrow \quad \text{R} \\
\text{NH}_2 & \quad \rightarrow \quad \text{R} \quad \text{NH}_3 \quad \text{O} \\
\text{O} & \quad \rightarrow \quad \text{O} \\
\end{align*}
\]

- **citric acid cycle**

\[
\begin{align*}
\text{O} & \quad \rightarrow \quad \text{O} \\
\text{NH}_3 & \quad \rightarrow \quad \text{NH}_3 \\
\text{R} & \quad \rightarrow \quad \text{R} \\
\end{align*}
\]

**OVERALL: TRANSAMINATION**
PLP Chemistry – Transamination & Racemisation

- **Transamination:**

![Chemical diagram showing the process of transamination involving pyridoxamine phosphate and pyridoxal phosphate](Image)
PLP Chemistry – Decarboxylation

- **Decarboxylation:**

- **Decarboxylation of lysine & ornithine:**

- **PLP dependant decarboxylase:**

- **lysine**

- **cadaverine**

- **ornithine**

- **putrescine**

- **piperidine alkaloids**

- **pyrrolidine alkaloids**
Lysine-derived Piperidine Alkaloids – *Hemlock!*

Socrates drinking poison hemlock, 399 B.C.

"The Death of Socrates" by Jacques-Louis David (1787)
Piperidine Alkaloids – Pelletierine & Coniine

- **Pelletierine**:

- **Coniine**:
  - in 399 BC Socrates was sentenced to death for impiety and executed by being forced to drink a potion made from poison hemlock. The toxic component in hemlock is coniine. Although by analogy with the above pathway, biosynthesis from lysine might be suspected, it is in fact of fatty acid origin.
Tyrosine-derived Alkaloids - *Opium Alkaloids*

Benzylisoquinoline Alkaloids

- papaverine
- morphine
Benzylisoquinoline Alkaloids – *Ring Formation*

- **Benzylisoquinoline alkaloids** constitute an extremely large and varied group of alkaloids — many, particularly the *opium alkaloids* (e.g. papaverine, morphine) are biosynthesised from two molecules of tyrosine via *nor-coclaurine* (and then *nor-laudanosoline*).

- **Mechanism of Pictet Spengler reaction:**

  \[
  \text{4-hydroxyphenylacetaldehyde} + \text{dopamine} \rightarrow \text{(S)-nor-coclaurine}
  \]
Benzylisoquinoline Alkaloids - *Papaverine*

- **Papaverine:** analgesic constituent of the opium poppy (*Papaver somniferum*):
  - biosynthesis:

  ![Chemical diagram of biosynthesis of Papaverine](image)

  - NB. The prefix *nor* means **without a methyl group**. Coclaurine, reticuline and laudanosine are the *N*-methyl compounds.
Oxidative Phenolic Coupling – Morphine & Synthetic Opioids

- **Morphine**: analgesic & sedative constituent of the *opium poppy* (*Papaver somniferum*):
  - **biosynthesis**: o-/p- oxidative phenolic coupling of *reticuline*:

  ![Chemical Structure Diagram]

  - Morphine acts by activating the **opiate receptors** in the brain (IC₅₀ 3 nM)
  - The natural ligands for these receptors are peptides: *e.g.* Leu-enkephalin (Tyr–Gly–Gly–Phe–Leu) (IC₅₀ 12 nM)
Dimeric Indole Alkaloids – *Vinca extracts*

Dimeric Indole Alkaloids

![Chemical structure of vinblastine and vincristine](image)

- **vinblastine** \((R = \text{Me})\)
- **vincristine** \((R = \text{CHO})\)

Potent *anti tumour* alkaloids used in *cancer chemotherapy*
Tryptamine + Secologanin $\rightarrow$ Strictosidine

- Most alkaloids of mixed Tryptophan/mevalonate biogenesis (>1200) are derived from strictosidine:
  - **Strictosidine** is derived from the condensation of **tryptamine** with the iridoid C$_{10}$ monoterpenec **secologanin**:

\[
\text{mevalonate (x2)} \rightarrow \text{geranyl pyrophosphate} \rightarrow \text{secologanin}
\]

- **Mechanism of Pictet-Spengler reaction:**
  - via **spirocyclic** intermediate then **Wagner-Meerwein** 1,2-alkyl shift:

\[
\text{tryptamine} + \text{secologanin} \rightarrow \text{spirocyclic intermediate} \rightarrow \text{strictosidine}
\]
Strictosidine $\rightarrow$ Vinca, Strychnos, Quinine etc.

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

1) 'light reactions': $hv \rightarrow$ ATP and NADPH

2) 'dark reactions': $CO_2 \rightarrow$ sugars (Calvin cycle)

PHOTOSYNTHESIS

glycolysis

$CO_2 + H_2O$
Fatty Acids
Fatty Acid Primary Metabolites

- **Primary metabolites:**
  - fully saturated, linear carboxylic acids & derived (poly)unsaturated derivatives:
    - constituents of essential natural waxes, seed oils, **glycerides** (fats) & phospholipids
    - structural role – **glycerides** & phospholipids are essential constituents of cell membranes
    - energy storage – **glycerides** can also be catabolised into acetate → citric acid cycle
    - biosynthetic precursors – for elaboration to secondary metabolites

**SATURATED ACIDS** \([\text{MeCH}_2(\text{CH}_2\text{CH}_3)_n\text{CH}_2\text{CO}_2\text{H} \ (n = 2-8)]\)

*E.g.*

- **Caprylic acid** (C8, n = 2)
- **Capric acid** (C8, n = 3)
- **Lauric acid** (C12, n = 4)
- **Myristic acid** (C14, n = 5)
- **Palmitic acid** (C16, n = 6)
- **Stearic acid** (C18, n = 7)

**MONO-UNSATURATED ACID DERIVATIVES (MUFAs)** *E.g.*

- **Palmitoleic acid** (C16, Z -\(\Delta^9\))
- **Oleic acid** (C18, Z -\(\Delta^9\))
  (>80% of fat in olive oil)

**POLY-UNSATURATED ACID DERIVATIVES (PUFAs)** *E.g.*

- **Arachidonic acid (AA)** (C20, Z -\(\Delta^5\), Z -\(\Delta^8\), Z -\(\Delta^{11}\), Z -\(\Delta^{14}\))
- **Eicosapentaenoic acid (EPA)** (C20, Z -\(\Delta^5\), Z -\(\Delta^8\), Z -\(\Delta^{11}\), Z -\(\Delta^{14}\), Z -\(\Delta^{17}\))
  (in cod liver oil)
Fatty Acids Derivatives – Secondary Metabolites

- **Secondary metabolites**
  - further *elaborated* derivatives of **polyunsaturated fatty acids (PUFAs)**
  - e.g. polyacetylenes & ‘eicosanoids’ (prostaglandins, thromboxanes & leukotrienes)

**Polyacetylenes**
e.g. wyerone
anti-fungal

**Prostaglandins**
e.g. prostaglandin F$_{2\alpha}$ (PGF$_{2\alpha}$)

**Thromboxanes**
e.g. thromboxane A$_2$ (TXA$_2$)

**Leukotrienes**
e.g. leukotriene A$_4$ (LTA$_4$)

**Eicosanoids**

**Obtusallene II**
marine natural product

**Matricaria ester**
in chamomile tea
Biosynthesis of Fatty Acids – *Iterative Oligomerisation*

- **fatty acids** are biosynthesised from **acetyl CoA** as a **starter unit** by **iterative** ‘head-to-tail’ **oligomerisation** involving:
  - condensation with **malonyl CoA** as an **extender unit** (with loss of **CO₂**) – a **decarboxylative Claisen condensation**
  - 3-step **reduction** of the resulting **ketone** → **methylene**
- after **n = 2-8 iterations** the **C8-20 saturated fatty acid** is released from the enzyme(s):
Biosynthesis of Fatty Acids – Overview of FAS

- The *in vivo* process by which all this takes place involves a ‘molecular machine’ - **Fatty Acid Synthase (FAS)**
  - **Type I FAS**: *single multifunctional protein complex* (e.g. in mammals incl. humans)
  - **Type II FAS**: *set of discrete, dissociable single-function proteins* (e.g. in bacteria)
  - **All FASs** comprise **8 components** (ACP & 7× catalytic activities): ACP, KS, AT, MT, KR, DH, ER & [TE]:

\[
\text{ACP, KS, AT, MT, KR, DH, ER & [TE] :}
\]

\[
\text{CoAS \rightarrow AT \rightarrow KS\_ACP \rightarrow MT \rightarrow ACP \rightarrow translocation}
\]

KS = keto synthase (also known as CE = condensing enzyme); AT = acetyl transferase; MT = malonyl transferase; KR = keto reductase; DH = dehydratase; ER = enoyl reductase; TE = thioesterase; ACP = acyl carrier protein
Human Fatty Acid Synthase (FAS)

- the first three-dimensional structure of human fatty acid synthase (272 kDa) at 4.5 Å resolution by X-ray crystallography:
  - Maier, Jenni & Ban *Science* 2006, 311, 1258 [DOI](https://doi.org/10.1126/science.1124798); also Fungal FAS @ 3.1 Å resolution see: Jenni *et al.* *Science* 2007, 316, 254 & 288

---

**Structural overview.**
(A) Front view: FAS consists of a lower part comprising the KS (lower body) and MAT domains (legs) connected at the waist with an upper part formed by the DH, ER (upper body), and KR domains (arms). (B) Top view of FAS with the ER and KR domains resting on the DH domains. (C) Bottom view showing the arrangement of the KS and MAT domains and the continuous electron density between the KS and MAT domains.
FATTY ACID BIOSYNTHESIS (type II FAS)

NB. the following sequence of slides have been adapted from: http://www.courses.fas.harvard.edu/%7echem27/
FATTY ACID BIOSYNTHESIS

- $\text{AT}_1$ loads acetyl group onto $\text{KS}_1$

Diagram:

- ACP$_1$ → AT$_1$ → KS$_1$ → KR$_1$ → DH$_1$ → ER$_1$ → ACP$_2$ → MT$_2$
- Pantetheine
- Acetyl-CoA
- SH
FATTY ACID BIOSYNTHESIS

ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → MT₂

Pantetheine

Cys

O

Me

SH
FATTY ACID BIOSYNTHESIS

- AT_1 loads malonyl group onto ACP_1

- Pantetheine

- Malonyl-CoA

- SH
FATTY ACID BIOSYNTHESIS

ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → MT₂

Pantetheine

Cys

O\("Me"

SH
FATTY ACID BIOSYNTHESIS

ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → MT₂

• KS₁ catalyzes Claisen condensation
FATTY ACID BIOSYNTHESIS

ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → MT₂

Pantetheine

Cys

SH

S

O=S

O=Me
FATTY ACID BIOSYNTHESIS

• KR$_1$ catalyzes reduction of ketone
FATTY ACID BIOSYNTHESIS

ACP1 → AT1 → KS1 → KR1 → DH1 → ER1 → ACP2 → MT2

Pantetheine

Cys → SH → S

O

Me

SH
• DH₁ catalyzes dehydration of alcohol
FATTY ACID BIOSYNTHESIS

ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → MT₂

Pantetheine

Cys

S

Me

SH

SH
FATTY ACID BIOSYNTHESIS

- ER$_1$ catalyzes reduction of alkene
FATTY ACID BIOSYNTHESIS

ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → MT₂

Pantetheine

Cys

SH

S

Me

SH
FATTY ACID BIOSYNTHESIS

- KS$_2$ catalyzes translocation to module 2
FATTY ACID BIOSYNTHESIS

H₁ → ER₁ → ACP₂ → MT₂ → KS₂ → KR₂ → DH₂ → ER₂ → TE

Pantetheine
FATTY ACID BIOSYNTHESIS

- MT$_2$ loads malonyl group onto ACP$_2$

Malonyl-CoA

Pantetheine

H$_1$ ER$_1$ ACP$_2$ MT$_2$ KS$_2$ KR$_2$ DH$_2$ ER$_2$ TE

Ser OH

O=C=O S—Co

O=C=O S—Co
FATTY ACID BIOSYNTHESIS

H1 -> ER1 -> ACP2 -> MT2 -> KS2 -> KR2 -> DH2 -> ER2 -> TE

Pantetheine
FATTY ACID BIOSYNTHESIS

- KS\textsubscript{2} catalyzes Claisen condensation
FATTY ACID BIOSYNTHESIS
FATTY ACID BIOSYNTHESIS

- KR$_2$ catalyzes reduction of ketone
FATTY ACID BIOSYNTHESIS

H\textsubscript{1} → ER\textsubscript{1} → ACP\textsubscript{2} → MT\textsubscript{2} → KS\textsubscript{2} → KR\textsubscript{2} → DH\textsubscript{2} → ER\textsubscript{2} → TE

Pantetheine

Cys

SH

\begin{align*}
\text{Me} & \\
\text{O} & \\
\text{OH} & 
\end{align*}
• DH$_2$ catalyzes dehydration of alcohol
FATTY ACID BIOSYNTHESIS

H₁ → ER₁ → ACP₂ → MT₂ → KS₂ → KR₂ → DH₂ → ER₂ → TE

Pantetheine

Cys SH

S
O
Me
OH
Ser
• ER₂ catalyzes reduction of alkene
FATTY ACID BIOSYNTHESIS

Pantetheine
• TE catalyzes transesterification
FATTY ACID BIOSYNTHESIS

\[ \text{H}_1 \rightarrow \text{ER}_1 \rightarrow \text{ACP}_2 \rightarrow \text{MT}_2 \rightarrow \text{KS}_2 \rightarrow \text{KR}_2 \rightarrow \text{DH}_2 \rightarrow \text{ER}_2 \rightarrow \text{TE} \]

- Pantetheine
- SH
- Cys
- Me
- Ser
FATTY ACID BIOSYNTHESIS

- TE catalyzes hydrolysis

- Pantetheine

- Cys SH

- Ser OH

- Me
FATTY ACID BIOSYNTHESIS

\[ \text{ACP}_2 \rightarrow \text{MT}_2 \rightarrow \text{KS}_2 \rightarrow \text{KR}_2 \rightarrow \text{DH}_2 \rightarrow \text{ER}_2 \rightarrow \text{TE} \]

Pantetheine

Cys

\( \text{SH} \)

OH

O

Me

SH
Biosynthesis of Unsaturated Fatty Acids

- **two mechanisms** are known for the introduction of double bonds into fatty acids:
  - in **BACTERIA**: anaerobic \([O]\) → monounsaturated FAs (MUFA\(s\))
  - in **MAMMALS, INSECTS & PLANTS**: aerobic \([O]\) → MUFA\(s\) & polyunsaturated FAs (PUFA\(s\))
Rational Anti-inflammatory Development – Prostaglandin & Thromboxane Pathway Intervention

- **prostaglandins** and **thromboxanes** are derived from further oxidative processing of arachidonic acid
- both are important **hormones** which control e.g. smooth **muscle contractility** (blood pressure), **gastric secretion**, **platelet aggregation** & **inflammation** (<nM activity)
  - various pharmaceuticals including **corticosteroids** & **aspirin** inhibit biosynthetic steps in these pathways

![Chemical reactions and structures](image-url)
Polyketides
Polyketides

- the structural variety of **polyketide secondary metabolites** is very wide:
  - NB. starter units marked in red; extender units in bold black; post oligomerisation appended groups in blue

- 6-methylsalicylic acid (antibiotic)
- orsellinic acid
- citrinin (kidney toxin, ‘yellow rice disease’)
- Griseofulvin (treatment for ring worm infections)
- actinorhodin (antibiotic)
- aflatoxin B1 (mycotoxic carcinogen)
- rapamycin (immunosuppressant)
- 6-deoxyerythronolide B (NB. a polypropionate)
- erythromycin A (antibiotic)
- mevinolin (lovastatin®) (anti-cholesterol)
Biosynthesis of Polyketides – Oligomerisation Steps

- **polyketides** are biosynthesised by a process very similar to that for **fatty acids**
  - the key **differences** are:
    - greater variety of starter units, extender units & termination processes
    - absent or incomplete reduction of the iteratively introduced β-carbonyl groups: ie. each cycle may differ in terms of KR, DH & ER modules & stereochemistry

- this leads to enormous diversity...
Biosynthesis of Polyketides – Overview of PKS

- The in vivo process of polyketide synthesis involves PolyKetide Synthases (PKSs):
  - PKSs (except Type II, see later) comprise the same 8 components as FASs, i.e. (ACP & 7x catalytic activities): ACP, KS, AT, MT, [KR, DH, ER & TE]
  - Type I PKSs: single (or small set of) multifunctional protein complex(es)
    - modular (microbial) - each ‘step’ has a dedicated catalytic site (→ macrolides)
    - iterative (fungal) – single set of catalytic sites, each of which may operate in each iteration (cf. FASs) (→ aromatics/polyphenols - generally)
  - Type II PKSs: single set of discrete, dissociable single-function proteins
    - iterative (microbial) - each catalytic module may operate in each iteration (cf. FASs) (→ aromatics/polyphenols)

\[ \text{KS} = \text{keto synthase; AT} = \text{acetyl transferase; MT} = \text{malonyl transferase;}
\]
\[ \text{KR} = \text{keto reductase; DH} = \text{dehydratase; ER} = \text{enoyl reductase; TE} = \text{thioesterase; ACP} = \text{acyl carrier protein} \]
POLYKETIDE BIOSYNTHESIS [Type I – (modular)]

NB. the following sequence of slides has also been adapted from: http://www.courses.fas.harvard.edu/%7echem27/
• AT₀ loads starting group (propionyl) onto ACP₀
POLYKETIDE BIOSYNTHESIS
• $KS_1$ catalyzes translocation to module 1
POLYKETIDE BIOSYNTHESIS

ACP_0 → AT_0 → ACP_1 → AT_1 → KS_1 → KR_1 → DH_1 → ER_1 → ACP_2 → A

Pantetheine

SH
- **AT\textsubscript{1}** loads methylmalonyl group onto ACP\textsubscript{1}
POLYKETIDE BIOSYNTHESIS

P0 AT0 ACP1 AT1 KS1 KR1 DH1 ER1 ACP2 AT2

Pantetheine

Cys

O

O

O

Me

SH
• \( KS_1 \) catalyzes Claisen condensation
POLYKETIDE BIOSYNTHESIS

P_0 \rightarrow \text{AT}_0 \rightarrow \text{ACP}_1 \rightarrow \text{AT}_1 \rightarrow \text{KS}_1 \rightarrow \text{KR}_1 \rightarrow \text{DH}_1 \rightarrow \text{ER}_1 \rightarrow \text{ACP}_2 \rightarrow \text{AT}_2

Pantetheine

Stereocenter

SH
• **KR$_1$** catalyzes reduction of ketone
POLYKETIDE BIOSYNTHESIS

ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → AT₂

Pantetheine

Cys

Stereocenter

SH
POLYKETIDE BIOSYNTHESIS

- no DH$_1$ activity
POLYKETIDE BIOSYNTHESIS

• no ER$_1$ activity
• KS$_2$ catalyzes translocation to module 2
Biosynthesis of Erythromycin – Type I(modular) PKS

- 6-deoxyerthronolide is a precursor to **erythromycin A** – **bacterial** antibiotic (*Streptomyces erythreus*):
  - propionate based **heptaketide**; 3 multifunctional polypeptides (DEBS1, DEBS2 & DEBS3, all ~350 kDa)

![Diagram of the biosynthesis of erythromycin A](image)

**Key enzymes:**
- **AT** = Acyl Carrier Protein
- **KS** = Ketosynthase
- **KR** = Ketoreductase
- **DH** = Dehydratase
- **ER** = Epoxydase
- **TE** = Thioesterase

**Release:** erythronolide B

**Loading:**

- 6-deoxyerthronolide
Type II PKSs – Enzyme Clusters (Microbial)

- **Type II PKSs**: single set of discrete, dissociable single-function proteins (ACP & 6× catalytic functions): \( ACP, KS_\alpha, KS_\beta, [KR, DH, ER, & TE] \) \([NB. \ NO\ acetyl\ or\ malonyl\ transferases\ (AT,\ MT)]\)
  - **iterative** - each catalytic module *may* operate in each iteration (cf. FASs) \(\rightarrow\) aromatics/polyphenols
- these clusters (generally) use *malonate* as BOTH *starter* & *extender* unit
- their *ACP proteins* are able to load malonate direct from malonyl CoA (no MT required)
  - the *starter malonate* is *decarboxylated* by ‘keto synthase’ \(\beta (KS_\beta)\) to give S-acetyl-ACP
  - the *extender malonates* undergo *decarboxylative Claisen condensations* by keto synthase \(\alpha (KS_\alpha)\)
- these clusters rarely utilise *KR, DH* or *ER* activities and produce ‘true’ polyketides:

\[
\begin{align*}
\text{CoAS} & \quad \text{CONH}_2 \\
\text{ACP} & \quad \text{SH} \\
\text{KS}_\beta \quad & \quad \text{ACP} \\
\text{KS}_\alpha \quad & \quad \text{ACP} \\
\text{KS}_\alpha \quad & \quad \text{ACP} \\
\text{KS}_\alpha \quad & \quad \text{ACP} \\
\text{KS}_\beta \quad & \quad \text{ACP} \\
\text{SH} & \quad \text{SH} \\
\text{SH} & \quad \text{SH} \\
\text{SH} & \quad \text{SH} \\
\text{SH} & \quad \text{SH} \\
\end{align*}
\]

\( KS_\beta \) = ‘keto synthase \(\beta\)’ (=decarboxylase!); \( KS_\alpha \) = ‘keto synthase \(\alpha\)’ (=ketosynthase!);
\( KR \) = keto reductase;
\( DH \) = dehydratase;
\( ER \) = enoyl reductase;
\( TE \) = thioesterase;
\( ACP \) = acyl carrier protein
Biosynthesis of Actinorhodin – *Type II PKS*

- **actinorhodin** – octaketide *bacterial antibiotic* (Streptomyces coelicolor)
  - Hopwood Chem. Rev. 1997, 97, 2465 (DOI)

- Timing of 1st cyclisation and mechanism of control of chain length uncertain
  - octaketide synthesis then cyclisation? (as shown above)
  - hexaketide synthesis then cyclisation then two further rounds of extension?
- Indications can sometimes be gleaned from *biomimetic syntheses.*
Scope of Structures - Type II PKS

- *microbial polyphenolic* metabolites:
  - **pentaketides (5x C₂)**
    - Chemical structure: ![Structure of pentaketide](image)
    - Example: eugenone
  - **hexaketides (6x C₂)**
    - Chemical structure: ![Structure of hexaketide](image)
    - Example: plumbagin
  - **heptaketides (7x C₂)**
    - Chemical structure: ![Structure of heptaketide](image)
    - Example: rubrofusarin
  - **octaketides (8x C₂)**
    - Chemical structure: ![Structure of octaketide](image)
    - Example: emodin
  - **nonaketides (9x C₂)**
    - Chemical structure: ![Structure of nonaketide](image)
    - Example: tetracycline
  - **decaketides (10x C₂)**
    - Chemical structure: ![Structure of decaketide](image)
    - Example: rabelomycin

- many display interesting biological activities...
Primary Metabolism - Overview

**Primary metabolism**

\[ \text{CO}_2 + \text{H}_2\text{O} \]

1. **'light reactions':** \( \text{hv} \rightarrow \text{ATP and NADPH} \)
2. **'dark reactions':** \( \text{CO}_2 \rightarrow \text{sugars (Calvin cycle)} \)

**Primary metabolites**

- oligosaccharides
- polysaccharides
- nucleic acids (RNA, DNA)

**Secondary metabolites**

- SHIKIMATE METABOLITES
  - cinnamic acid derivatives
  - aromatic compounds
  - lignans, flavinoids

- ALKALOIDS
  - penicillins
  - cephalosporins
  - cyclic peptides

- FATTY ACIDS & POLYKETIDES
  - prostaglandins
  - polyacetylenes
  - aromatic compounds, polyphenols
  - macrolides

- ISOPRENOIDS
  - terpenoids
  - steroids
  - carotenoids

For interesting animations' of e.g. photosynthesis see: [http://www.johnkyrk.com/index.html](http://www.johnkyrk.com/index.html)
Isoprenoids
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 C₅ ‘isoprene’ units:
    - *monoterpenes* (C₁₀); *sesquiterpenes* (C₁₅); *diterpenes* (C₂₀); *sesterpenes* (C₂₅, rare); *triterpenes* (C₃₀); *carotenoids* (C₄₀)

![Chemical structures and reactions]

- dimethylallyl pyrophosphate (DMAPP)
- isopentenyl pyrophosphate (IPP)
- cholesterol (C₂₇ but C₃₀-derived)
- humulone (2x C₅)
- natural rubber (~10⁵x C₅)
- thujone (C₁₀)
- borneol (C₁₀)
- lavandulol (C₁₀)
- (Z)-α-bisabolene (C₁₅)
- artemisinin (C₁₉)
- β-carotene (C₄₀)
- euonyminol (C₁₅)
- gibberellic acid (C₂₀) (gibberellin A₃)
- taxol (C₂₀)
Biosynthesis of IPP & DMAPP - via Mevalonate

- **IPP & DMAPP** are the key $C_5$ precursors to all isoprenoids
  - the main pathway is via: acetyl CoA $\rightarrow$ acetoacetyl CoA $\rightarrow$ HMG CoA $\rightarrow$ mevalonate $\rightarrow$ IPP $\rightarrow$ DMAPP:

\[\begin{align*}
&\text{acetyl CoA} \\
\rightarrow &\text{acetoacetyl CoA} \\
\rightarrow &\text{HMG CoA} \\
\rightarrow &\text{mevalonate} \\
\rightarrow &\text{IPP} \\
\rightarrow &\text{DMAPP}
\end{align*}\]
Rational Anti-cholesterol Development - Statins

- **HMG CoA → MVA** is the *rate determining step* in the biosynthetic pathway to **cholesterol**
- ‘**Statins**’ inhibit HMG CoA reductase and are used clinically to treat **hypercholesterolaemia** - a causative factor in **heart disease**, see: Wu et al. *Tetrahedron* 2015, 71, 8487 (DOI)
  - *e.g. mevinolin* (=lovastatin®, Merck) from *Aspergillus terreus* is a competitive inhibitor of HMG-CoA reductase
  - *e.g. lipitor* (Atorvastatin calcium, Pfizer) is also a competitive inhibitor of HMG-CoA reductase and the world's biggest selling drug [first drug to reach $10 billion sales (2004: $10.8 bn)]
Linear C$_{5n}$ ‘head-to-tail’ Pyrophosphates

- head-to-tail C$_5$ oligomers are the key precursors to isoprenoids
  - geranyl pyrophosphate (C$_{10}$) is formed by $S_{N1}$ alkylation of DMAPP by IPP → monoterpenes
  - farnesyl (C$_{15}$) & geranylgeranyl (C$_{20}$) pyrophosphates are formed by further $S_{N1}$ alkylations with IPP:

\[
\text{DMAPP} \xrightarrow{S_{N1}} \text{IPP} \xrightarrow{\text{farnesyl}} \text{IPP} \xrightarrow{\text{geranylgeranyl}}
\]

**MONOTERPENES (C$_{10}$)**

**SESQUITERPENES (C$_{15}$)**

**TRITERPENES (C$_{30}$)**

**DITERPENES (C$_{20}$)**

**CAROTENOIDS (C$_{40}$)**
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 Mg$^{2+}$ ions
  - all three pyrophosphates are substrates for *cyclases* via an $\alpha$-terpinyl cation:

\[
\begin{align*}
\text{gerenyl} & \quad \text{pyrophosphate} \\
\text{linalyl} & \quad \text{pyrophosphate} \\
\text{neryl} & \quad \text{pyrophosphate}
\end{align*}
\]

\[
\begin{align*}
\text{allylic cation} \\
\text{intimate ion pair}
\end{align*}
\]

\[
\begin{align*}
\text{cyclase} & \quad \rightarrow \\
\alpha\text{-terpinyl cation}
\end{align*}
\]

\[
\begin{align*}
\text{MONOTERPENES (C}_{10}\text{)}
\end{align*}
\]
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

![Chemical structures and reaction schemes involving $\alpha$-terpineol, limonene, $\alpha$-pinene, thujone, and other monoterpenes showing various reaction mechanisms such as $E_1$ elimination, trapping by PPO, hydrolysis, and 1,2-hydride and 1,2-alkyl shifts.](image)
Sesquiterpenes – *Farnesyl Pyrophosphate (FPP)*

- ‘$S_N 2$’-like alkylation of geranyl PP by IPP gives farnesyl PP:

\[
\text{geranyl PP} \rightarrow \text{farnesyl PP} (FPP)
\]

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

**NB. control by:**
1) *enzyme* to enforce conformation & sequestration of reactive intermediates
2) intrinsic *stereoelectronics* of participating orbitals
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](http://www.chem.qmul.ac.uk/iubmb/enzyme/reaction/terp/taxadiene.html)
  - home page is: [http://www.chem.qmul.ac.uk/iubmb/enzyme/](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 – \( FPP \rightarrow Squalene \)

- **triterpenes** \((C_{30})\) arise from the ‘**head to head**’ coupling of two 'fanesyl 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 **hypercholesterolaemia & heart disease** (cf. statins)

\[
\begin{align*}
\text{FPP (donor)} & \quad \text{EnzB:} \\
\text{FPP (acceptor)} & \quad \text{squalene synthase} \\
\text{FPP (donor)} & \quad \text{blocked by squalestatins} \\
\text{presqualene PP} & \\
\text{NADPH} & \quad \text{NADP}^+ + \text{PP}_i \\
\text{squalene} & \\
\end{align*}
\]
Oxidosqualene-Lanosterol Cyclase – *Mechanism*

- **oxidosqualene-lanosterol cyclase** catalyses the formation of *lanosterol* from 2,3-oxidosqualene:
  - this cascade establishes the characteristic ring system of *ALL steroids*
  - ring-expansion sequence to establish the C ring
  - the process is *NOT concerted*, discrete *cationic intermediates* are involved & *stereoelectronics dictate* the *regio- & stereoselectivity* although the enzyme undoubtedly lays a role in pre-organising the ~chair-boat-chair conformation

  “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”

Lanosterol → Cholesterol – Oxidative Demethylation

Several steps are required for conversion of lanosterol to cholesterol:

1) $\Delta^{24}$ hydrogenation
   - NADPH
   - NADP$^+$

2) 14$\alpha$ DEMETHYLATION
   - $2x\ O_2$
   - $P_{450}$
   - $2x\ H_2O$

3) 4$\alpha$ & 4$\beta$ DEMETHYLATION
   - $4x\ O_2$
   - $P_{450}$
   - $4x\ H_2O$
   - $CO_2$

4) $\Delta^8$ to $\Delta^\ast$ rearrangement
   - Isomerase
   - NAD$^+$
   - NADH
   - NADP$^+$

Flat, rigid structure
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 *P₄₅₀* enzymes
  - *estrone* is produced from *androstendione* by oxidative demethylation with concomitant aromatisation:

\[
\text{cholesterol} \xrightarrow{2 \times \text{H}_2\text{O}} \text{estrone (oestrone)} \xrightarrow{\text{NAD}^+} \text{androstendione (X = O)} \xrightarrow{2 \times \text{H}_2\text{O}} \text{testosterone (X = H, } \beta\text{-OH)}
\]

**DEMETHYLATIVE aromatisation by 'aromatase' enzyme**
Steroid Ring Cleavage - Vitamin D & Azadirachtin

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

```
\[ \text{cholesterol} \xrightarrow{\text{UV light (hv)}} \text{vitamin D}_2 \]
```

- 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:

```
\[ \text{tirucallol (cf. lanosterol)} \xrightarrow{\text{oxidative cleavage of C ring}} \text{azadirachtin A (a limanoid = tetra-nor-triterpenoid)} \xrightarrow{[O]} \text{azadirachtin} \]
```

- highly hindered C-C bond for synthesis!
Summary of Presentation

- **Metabolism & Biosynthesis**
  - some definitions, 1° & 2° metabolites
- **Shikimate Metabolites**
  - photosynthesis & glycolysis → shikimate formation → shikimate metabolites
  - *Glyphosate* – a non-selective herbicide
- **Alkaloids**
  - acetylCoA & the citric acid cycle → α-amino acids → alkaloids
  - *Opioids* – powerful pain killers
- **Fatty Acids and Polyketides**
  - acetylCoA → malonylCoA → fatty acids, prostaglandins, polyketides, macrolide antibiotics
  - *NSAIDs* – anti-inflammatory’s
- **Isoprenoids/terpenes**
  - acetylCoA → mevalonate → isoprenoids, terpenoids, steroids, carotenoids
  - *Statins* – cholesterol-lowering agents