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

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

Imperial College London

Oct 2016
Lessons in Synthesis - Azadirachtin

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

  - 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
- Caffeine
- Lysergic acid
- Clavulanic acid
- Nicotine
- Patulin
- Quinine
- Androstenedione

Chemical reactions:

\[ \text{CO}_2 + \text{H}_2\text{O} + \text{P}_i + \text{N}_2 \xrightarrow{hv} \]
**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:
Primary Metabolism - Overview

Primary metabolism

$\text{CO}_2 + \text{H}_2\text{O}$

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

**Photosynthesis**

- glycolysis
- glucose & other 4,5,6 & 7 carbon sugars
- phosphoenol pyruvate
- erythrose-4-phosphate
- pyruvate

**Primary metabolites**

- phosphoenol pyruvate
- erythrose-4-phosphate
- pyruvate
- $\text{CO}_2$

**Secondary metabolites**

- oligosaccharides
- polysaccharides
- nucleic acids (RNA, DNA)
- 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
  - (ArC₆)

- (S)-tyrosine
  - (ArC₃)

- (S)-phenylalanine
  - (ArC₃)

- scopoletin
  - (ArC₃)

- Menoquinone (vitamin K₂)
  - (ArC₃)

- α-tocopherol (vitamin E)
  - (ArC₃)

- Epigallocatechin (EGC)
  - (ArC₃)

- Podophyllotoxin
  - (ArC₃)
The Shikimate Biosynthetic Pathway - Overview

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

```
  Phosphoenol pyruvate (PEP)
  erythrose-4-phosphate (E-4-P)
  DAHP
  chorismate
  5-EPS-3-P
  3-phosphoshikimate (3-PS)
  3-dehydroshikimate (3-DHS)
  3-dehydroquinate (3-DHQ)

  prephenate
  tryptophan
  phenylalanine
  tyrosine

  non-concerted syn-elimination
  FMNH2-mediated non-concerted 1,4-anti-elimination
  Claisen rearrangement
```

– 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

\[
\begin{align*}
\text{phosphoenol pyruvate (PEP)} & \quad \rightarrow \quad \text{3-phosphoshikimate (3-PS)} \quad \rightarrow \quad \text{5-enolpyruvylshikimate-3-phosphate (5-EPS-3P)} \quad \rightarrow \quad \text{glyphosate (Roundup®)}
\end{align*}
\]
Chorismate → Tryptophan, Tyrosine & Phenylalanine

- **Chorismate → anthranilate → tryptophan**

- **Chorismate → prephenate → 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)

  - Ammonia lyase (PAL)
  - Coumarins:
    - Scopoletin (a coumarin) germination stimulant
  - Lignans:
    - Podophyllotoxin (a lignan) natural product used to treat worts
    - Pinoresinol (a lignan)
  - Lignin fragments:
    - 'Woody' component of cell walls

Chemical structures:
- Scopoletin (coumarin)
- Podophyllotoxin (lignan)
- Pinoresinol (lignan)
- Ferulate (cinnamate derivative)
- Phenylalanine
- ArC₃ metabolites
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': \(\text{CO}_2 \rightarrow \text{sugars (Calvin cycle)}\)

\[
\begin{align*}
glucose \\
& \rightarrow \text{erythrose-4-phosphate} \\
& \rightarrow \text{shikimate} \\
& \rightarrow \text{aromatic amino acids} \\
& \rightarrow \text{peptides proteins} \\
& \rightarrow \text{saturated fatty acids unsaturated fatty acids lipids} \\
\end{align*}
\]

Primary metabolites

Secondary metabolites

\[
\begin{align*}
\text{oligosaccharides} \\
\rightarrow \text{polysaccharides} \\
\rightarrow \text{nucleic acids (RNA, DNA)} \\
\rightarrow \text{alkaloids} \\
\rightarrow \text{penicillins} \\
\rightarrow \text{cephalosporins} \\
\rightarrow \text{cyclic peptides} \\
\rightarrow \text{saturated fatty acids} \\
\rightarrow \text{unsaturated fatty acids} \\
\rightarrow \text{lipids} \\
\rightarrow \text{fatty acids & polyketides} \\
\rightarrow \text{prostaglandins} \\
\rightarrow \text{polyacetylenes} \\
\rightarrow \text{aromatic compounds, polyphenols} \\
\rightarrow \text{macrolides} \\
\rightarrow \text{isoprenoids} \\
\rightarrow \text{terpenoids} \\
\rightarrow \text{steroids} \\
\rightarrow \text{carotenoids}
\end{align*}
\]

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](image)

**OVERALL STOICHIOMETRY**

- 1x [\(\text{acetate}\)]
- 1x \(\text{O}_2\)
- 2x \(\text{CO}_2\)
- 12x \(\text{ATP}\) energy!
The Biosynthesis of Lysine & Ornithine

- **Lysine & ornithine** - the two most significant, *non-aromatic* $\alpha$-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 $\alpha$-ketoacid mediated by *pyridoxal-5’-phosphate* (PLP)

---

**Diagram:**
- Reductive amination:
  - $\alpha$-ketoacid + $\mathrm{NH}_2$ \(\rightarrow\) amino acid
  - Pyridoxamine phosphate (tightly bound to enzyme)
  - Oxidative deamination:
    - glutamic acid (Glu) \(\rightarrow\) $\mathrm{NH}_3$

**Equivalents:**
- Lysine (Lys) [50 ATP equivs]
- Ornithine (Orn) [<44 ATP equivs (=Arg)]
PLP Chemistry – *Transamination* & *Racemisation*

- *Transamination* – LHS → RHS (*reductive amination*); RHS → LHS (*oxidative deamination*):
PLP Chemistry – Decarboxylation

- **Decarboxylation**:

- **Decarboxylation of lysine & ornithine**:

  - **Lysine**
    - ![Lysine](image)
    - Reaction: $\text{H}_2\text{NCH(NH}_2\text{)}\text{CH}_2\text{CO}_2\text{H}$ \rightleftharpoons $\text{RHNCH(NH}_2\text{)}\text{CH}_2\text{NH}_2$
    - Enzyme: Enz-NH$_2$
    - PLP-dependent decarboxylase
    - Product: Cadaverine
    - Reaction: $\text{H}_2\text{NCH(NH}_2\text{)}\text{CH}_2\text{NH}_2 \rightarrow \text{RHNCH(NH}_2\text{)}\text{CH}_2\text{NH}_2$
    - Further reaction: PLP-dependent decarboxylase
    - Product: Iminium salt
      - $\text{H}_2\text{NCH(NH}_2\text{)}\text{CH}_2\text{NH}_2 \rightarrow \text{RHNCH(NH}_2\text{)}\text{CH}_2\text{NH}_2$
      - Reaction: Iminium salt \rightarrow PIP

  - **Ornithine**
    - ![Ornithine](image)
    - Reaction: $\text{H}_2\text{NCH(NH}_2\text{)}\text{CH}_2\text{CO}_2\text{H}$ \rightleftharpoons $\text{RHNCH(NH}_2\text{)}\text{CH}_2\text{NH}_2$
    - Enzyme: Enz-NH$_2$
    - PLP-dependent decarboxylase
    - Product: Putrescine
    - Reaction: $\text{H}_2\text{NCH(NH}_2\text{)}\text{CH}_2\text{NH}_2 \rightarrow \text{RHNCH(NH}_2\text{)}\text{CH}_2\text{NH}_2$
    - Further reaction: PLP-dependent decarboxylase
    - Product: Iminium salt
      - $\text{H}_2\text{NCH(NH}_2\text{)}\text{CH}_2\text{NH}_2 \rightarrow \text{RHNCH(NH}_2\text{)}\text{CH}_2\text{NH}_2$
      - Reaction: Iminium salt \rightarrow PIP
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**:

  - 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-laudanosoline**:

  ![Mechanism of Pictet Spengler reaction](image)

  - **Mechanism of Pictet-Spengler reaction:**

    ![DHPP + dopamine](image)
Benzylisoquinoline Alkaloids - Papaverine

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

- *NB*. The prefix *nor* means *without a methyl group*. Laudanosoline, 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*:

Morphine acts by activating the **opiate receptors** in the brain (IC\(_{50} 3\) nM)
- The natural ligands for these receptors are peptides: e.g. Leu-enkephalin (Tyr–Gly–Gly–Phe–Leu) (IC\(_{50} 12\) nM)
Dimeric Indole Alkaloids – *Vinca extracts*

Dimeric Indole Alkaloids

![Chemical structure of dimeric indole alkaloids](image)

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

Potent *anti tumour* alkaloids used in *cancer chemotherapy*
Tryptamine + Secologanin → 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} monoterpenic secologanin:

  \[
  \text{tryptophan} + \text{secologanin} \rightarrow \text{strictosidine}
  \]

- Mechanism of Pictet-Spengler reaction:
  - via spirocyclic intermediate then Wagner-Meerwein 1,2-alkyl shift:
Strictosidine → Vinca, Strychnos, Quinine etc.

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

PHOTOSYNTHESIS

1) 'light reactions': hv -> ATP and NADPH
2) 'dark reactions': CO₂ -> sugars (Calvin cycle)

glycolysis

& other 4,5,6 & 7 carbon sugars

phosphoenol pyruvate

erythrose-4-phosphate

shikimate

\( CO_2 + H_2O \)

CO₂ + H₂O

For interesting animations' of e.g. photosynthesis see: [http://www.johnkyrk.com/index.html](http://www.johnkyrk.com/index.html)
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** (fats) 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{)}_n\text{CH}_2\text{CO}_2\text{H}} (n = 2-8)\] e.g.

8 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] caprylic acid (C8, n = 2)

10 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] capric acid (C8, n = 3)

12 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] lauric acid (C12, n = 4)

14 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] myristic acid (C14, n = 5)

16 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] palmitic acid (C16, n = 6)

18 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] stearic acid (C18, n = 7)

MONO-UNSATURATED ACID DERIVATIVES (MUFAs) e.g.

16 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] palmitoleic acid (C16, Z-\(\Delta^9\))

18 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] oleic acid (C18, Z-\(\Delta^9\))

(P>80% of fat in olive oil)

POLY-UNSATURATED ACID DERIVATIVES (PUFAs) e.g.

11 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] arachidonic acid (AA)

(C20, Z-\(\Delta^5\), Z-\(\Delta^8\), Z-\(\Delta^{11}\), Z-\(\Delta^{14}\))

11 \[\begin{array}{c}
\text{CO}_2\text{H} \\
1 \\
\end{array}\] 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**
- *wyerone* anti-fungal

**Prostaglandins**
- *prostaglandin F₂₀ (PGF₂₀)*

**Thromboxanes**
- *thromboxane A₂ (TXA₂)*

**Leukotrienes**
- *leukotriene A₄ (LTA₄)*

**Eicosanoids**
- *obtucallene II* marine natural product

Further elaborated derivatives of PUFAs include polyacetylenes and 'eicosanoids'. These compounds are biologically active substances produced by the body and are involved in various physiological processes. They include prostaglandins, thromboxanes, and leukotrienes, each with specific roles and functions in the body.
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]:

\[
\text{CoAS} \xrightarrow{\text{AT}} \text{ACP} \xrightarrow{\text{KS}} \text{ACP} \xrightarrow{\text{MT}} \text{ACP} \xrightarrow{\text{KR}} \text{ACP} \xrightarrow{\text{DH}} \text{ACP} \xrightarrow{\text{ER}} \text{ACP} \xrightarrow{\text{TE}} \text{ACP}
\]

\[
\text{CoAS} \xrightarrow{\text{AT}} \text{ACP} \xrightarrow{\text{KS}} \text{ACP} \xrightarrow{\text{MT}} \text{ACP} \xrightarrow{\text{KR}} \text{ACP} \xrightarrow{\text{DH}} \text{ACP} \xrightarrow{\text{ER}} \text{ACP} \xrightarrow{\text{TE}} \text{ACP}
\]

\[
\text{CoAS} \xrightarrow{\text{AT}} \text{ACP} \xrightarrow{\text{KS}} \text{ACP} \xrightarrow{\text{MT}} \text{ACP} \xrightarrow{\text{KR}} \text{ACP} \xrightarrow{\text{DH}} \text{ACP} \xrightarrow{\text{ER}} \text{ACP} \xrightarrow{\text{TE}} \text{ACP}
\]

\[
\text{CoAS} \xrightarrow{\text{AT}} \text{ACP} \xrightarrow{\text{KS}} \text{ACP} \xrightarrow{\text{MT}} \text{ACP} \xrightarrow{\text{KR}} \text{ACP} \xrightarrow{\text{DH}} \text{ACP} \xrightarrow{\text{ER}} \text{ACP} \xrightarrow{\text{TE}} \text{ACP}
\]

\[
\text{CoAS} \xrightarrow{\text{AT}} \text{ACP} \xrightarrow{\text{KS}} \text{ACP} \xrightarrow{\text{MT}} \text{ACP} \xrightarrow{\text{KR}} \text{ACP} \xrightarrow{\text{DH}} \text{ACP} \xrightarrow{\text{ER}} \text{ACP} \xrightarrow{\text{TE}} \text{ACP}
\]

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) ; 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

- AT$_1$ loads acetyl group onto KS$_1$
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

SH
FATTY ACID BIOSYNTHESIS

- AT₁ loads malonyl group onto ACP₁

Malonyl-CoA

Pantetheine

SH

ACP₁  AT₁  KS₁  KR₁  DH₁  ER₁  ACP₂  MT₂
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

O

O

Me

SH
FATTY ACID BIOSYNTHESIS

• **KS<sub>1</sub>** catalyzes Claisen condensation
FATTY ACID BIOSYNTHESIS

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

Cys → SH

Pantetheine

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
FATTY ACID BIOSYNTHESIS

- DH₁ catalyzes dehydration of alcohol
FATTY ACID BIOSYNTHESIS

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

Pantetheine

S

Cys

SH

O

Me
FATTY ACID BIOSYNTHESIS

- ER\textsubscript{1} catalyzes reduction of alkene
FATTY ACID BIOSYNTHESIS

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

Cys SH

Pantetheine

S
O
Me

SH
FATTY ACID BIOSYNTHESIS

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

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

![Diagram of fatty acid biosynthesis pathway](image-url)
FATTY ACID BIOSYNTHESIS

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

Pantetheine
FATTY ACID BIOSYNTHESIS

- $\text{KS}_2$ catalyzes Claisen condensation
FATTY ACID BIOSYNTHESIS

H₁ \( \rightarrow \) ER₁ \( \rightarrow \) ACP₂ \( \rightarrow \) MT₂ \( \rightarrow \) KS₂ \( \rightarrow \) KR₂ \( \rightarrow \) DH₂ \( \rightarrow \) ER₂ \( \rightarrow \) TE

Pantetheine

Cys \( \text{SH} \)

\( \text{Me} \)

\( \text{O} = \text{S} \)

\( \text{O} = \text{O} \)
• KR$_2$ catalyzes reduction of ketone
FATTY ACID BIOSYNTHESIS

\[ H_1 \rightarrow ER_1 \rightarrow ACP_2 \rightarrow MT_2 \rightarrow KS_2 \rightarrow KR_2 \rightarrow DH_2 \rightarrow ER_2 \rightarrow TE \]

Pantetheine
FATTY ACID BIOSYNTHESIS

- DH$_2$ catalyzes dehydration of alcohol
FATTY ACID BIOSYNTHESIS

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

Pantetheine
FATTY ACID BIOSYNTHESIS

- $ER_2$ catalyzes reduction of alkene
FATTY ACID BIOSYNTHESIS

Diagram showing the fatty acid biosynthesis pathway with enzymes ACP, MT, KS, KR, DH, and ER. The pathway also involves Pantetheine, a compound derived from cysteine (Cys) and serine (Ser).
FATTY ACID BIOSYNTHESIS

- TE catalyzes transesterification
FATTY ACID BIOSYNTHESIS

- $H_1$ → $ER_1$ → $ACP_2$ → $MT_2$ → $KS_2$ → $KR_2$ → $DH_2$ → $ER_2$ → TE

- Pantetheine

- Cys $\text{SH}$

- O → Ser

- Me $\text{SH}$
FATTY ACID BIOSYNTHESIS

- TE catalyzes hydrolysis
FATTY ACID BIOSYNTHESIS

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

Cys \rightarrow \text{SH}

\text{SH}

O\rightarrow\text{OH}

Me
Biosynthesis of Unsaturated Fatty Acids

- **two mechanisms** are known for the introduction of double bonds into fatty acids:
  - in **BACTERIA**: *anaerobic [O]* → monounsaturated FAs (**MUFAs**)
  - in **MAMMALS, INSECTS & PLANTS**: *aerobic [O]* → **MUFAs** & polyunsaturated FAs (**PUFAs**)

**ANAEROBIC ROUTE** (bacteria)  
(dehydrogenation occurs during chain elongation)  
mainly MUFAs but some PUFAs

**AEROBIC ROUTE** (mammals, insects & plants)  
(dehydrogenation occurs after chain elongation)  
MUFAs & PUFAs

NB. **in both cases cis - alkenes are produced**
Rational Anti-inflammatory Development – Prostaglandin & Thromboxane Pathway Intervention

- **prostaglandins** & **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 biosynthethetic steps in these pathways

![Chemical Diagram]

**Arachidonic Acid**

**Cyclooxygenase (COX)**

**Prostaglandins & Thromboxanes**

- **Thromboxane A2 (TXA2)** induces platelet aggregation
- **Prostacyclin (PGI2)** inhibits platelet aggregation
- **Thromboxane B2 (TXB2)**
- **Prostaglandin F2α (PGF2α)** used to induce labour

**Rofecoxib (Vioxx)**

Release from cellular stores inhibited by corticosteroids

Inhibited by aspirin, celebrex, vioxx etc.
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)

NB. a mixed polypropionate/acetate

- 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 & 7× 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)

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

ACP₀ → AT₀ → ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → A

Pantetheine

SH

Cys

SH

SH

NB. the following sequence of slides has also been adapted from: [http://www.courses.fas.harvard.edu/~echem27/](http://www.courses.fas.harvard.edu/~echem27/)
• \( \text{AT}_0 \) loads starting group (propionyl) onto \( \text{ACP}_0 \)
POLYKETIDE BIOSYNTHESIS
• $KS_1$ catalyzes translocation to module 1
POLYKETIDE BIOSYNTHESIS

ACP₀ → AT₀ → ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → A
• $\text{AT}_1$ loads methylmalonyl group onto $\text{ACP}_1$
POLYKETIDE BIOSYNTHESIS

Pantetheine

SH
\[ \text{KS}_1 \text{ catalyzes Claisen condensation} \]
POLYKETIDE BIOSYNTHESIS

P0 \xrightarrow{AT0} ACP1 \xrightarrow{AT1} KS1 \xrightarrow{KR1} DH1 \xrightarrow{ER1} ACP2 \xrightarrow{AT2}

Pantetheine

Stereocenter
• KR$_1$ catalyzes reduction of ketone
POLYKETIDE BIOSYNTHESIS
POLYKETIDE BIOSYNTHESIS

• no DH₁ activity
POLYKETIDE BIOSYNTHESIS

- no ER$_1$ activity
• KS$_2$ catalyzes translocation to module 2
POLYKETIDE BIOSYNTHESIS

Pantetheine

\[ \text{ACP}_2 \rightarrow \text{AT}_2 \rightarrow \text{KS}_2 \rightarrow \text{KR}_2 \rightarrow \text{DH}_2 \rightarrow \text{ER}_2 \rightarrow \text{TE} \]
Biosynthesis of Erythromycin – Type I(modular) PKS

- **6-deoxyerythronolide** is a precursor to **erythromycin A** – bacterial antibiotic (*Streptomyces erythreus*):
  - propionate based **heptaketide**; 3 multifunctional polypeptides (DEBS1, DEBS2 & DEBS3, all ~350 kDa)
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) (→ 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 ‘ketosynthase’ \(\beta\) (KS\(\beta\)) to give S-acetyl-ACP
  - the extender malonates undergo decarboxylative Claisen condensations by ketosynthase \(\alpha\) (KS\(\alpha\))
- these clusters rarely utilise KR, DH or ER activities and produce ‘true’ polyketides:

\[
\begin{align*}
\text{KS}_\beta &= \text{‘keto synthase } \beta\text{’ (=decarboxylase!); } \\
\text{KS}_\alpha &= \text{‘keto synthase } \alpha\text{’ (=ketosynthase!); } \\
\text{KR} &= \text{ketoreductase; } \\
\text{DH} &= \text{dehydratase; } \\
\text{ER} &= \text{enoyl reductase; } \\
\text{TE} &= \text{thioesterase; } \\
\text{ACP} &= \text{acyl carrier protein}
\end{align*}
\]
Biosynthesis of Actinorhodin – *Type II PKS*

- **actinorhodin** – octaketide *bacterial antibiotic* (*Streptomyces coelicolor*)
  - Hopwood *Chem. Rev.* 1997, 97, 2465 ([DOI](https://doi.org/10.1021/cr960046l))

  ![Chemical Diagram]

  - **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)
    - eugenone
  - hexaketides (6x C₂)
    - plumbagin
  - heptaketides (7x C₂)
    - rubrofusarin
  - octaketides (8x C₂)
    - emodin
  - nonaketides (9x C₂)
    - tetracycline
  - decaketides (10x C₂)
    - rabelomycine

- many display interesting biological activities...
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)

PHOTOSYNTHESIS

glycolysis

\( \text{CO}_2 \rightarrow \text{glyceraldehyde-3-phosphate} \)

& other 4,5,6 & 7 carbon sugars

\( \text{CO}_2 \rightarrow \text{glucose} \)

Primary metabolites

oligosaccharides
polysaccharides
nucleic acids (RNA, DNA)

Secondary metabolites

SHIKIMATE METABOLITES

cinnamic acid derivatives
aromatic compounds
lignans, flavonoids

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
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 C$_5$ ‘isoprene’ units:
    - **monoterpenes** (C$_{10}$); **sesquiterpenes** (C$_{15}$); **diterpenes** (C$_{20}$); **sesterpenes** (C$_{25}$, rare); **triterpenes** (C$_{30}$); **carotenoids** (C$_{40}$)
Biosynthesis of IPP & DMAPP - via Mevalonate

- **IPP & DMAPP** are the key \( C_5 \) precursors to all isoprenoids
  - the main pathway is via: acetyl CoA → acetoacetyl CoA → HMG CoA → mevalonate → IPP → DMAPP:
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](https://doi.org/10.1016/j.tet.2015.07.028))
  - *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 worlds biggest selling drug [first drug to reach $10 billion sales (2004: $10.8 bn)]

---

**Lovastatin**

- **PRO-DRUG**
- *NB.* type I (iterative) PKS natural product

**Lipitor**

- **ACTIVE DRUG**
  - mimic of tetrahedral intermediate in HMG reduction by NADPH

---

**Diagram: Statins inhibit this step**
Linear C₅ᵣ ‘head-to-tail’ Pyrophosphates

- head-to-tail C₅ oligomers are the key precursors to isoprenoids
  - geranyl pyrophosphate (C₁₀) is formed by Sₙ₁ alkylation of DMAPP by IPP → monoterpenes
  - farnesyl (C₁₅) & geranylgeranyl (C₂₀) pyrophosphates are formed by further Sₙ₁ alkylations with IPP:

![Chemical diagram showing the synthesis of terpenes from geranyl pyrophosphate, farnesyl pyrophosphate, and geranylgeranyl pyrophosphate.](attachment:terpene_diagram.png)
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**:

\[
\begin{align*}
\text{gerenyl pyrophosphate} & \quad \text{linalyl pyrophosphate} \\
\text{linalyl pyrophosphate} & \quad \text{neryl pyrophosphate}
\end{align*}
\]
Monoterpenes – *Fate of the α-Terpinyl Cation*

- The *α-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.

![Diagram of monoterpenes](image)
Sesquiterpenes – *Farnesyl Pyrophosphate (FPP)*

- ‘\(S_N^2\)-like alkylation of geranyl PP by IPP gives farnesyl PP:

\[
\text{geranyl PP} \quad \xrightarrow{\text{IPP}} \quad \text{E,E-farnesyl PP (FPP)}
\]

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

\[
\begin{align*}
\text{geranyl PP} & \quad \rightarrow \quad \text{neryl PP} \\
\text{geranyl PP} & \quad \rightarrow \quad \text{linalyl PP} \\
\text{farnesyl PP} & \quad \rightarrow \quad \text{equivalent compounds}
\end{align*}
\]

\[\text{allylic cation intimate ion pair}\]

\[\text{cyclases}\]

- further cyclisation
- 1,2-hydride & alkyl shifts
- trapping with \(H_2O\)
- elimination to alkenes

\[\text{vast array of mono- & bicyclic SESQUITERPENES}\]

\[\text{NB. control by:}\]
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:**

  ![Chemical Structures]

  - 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 – \( \text{FPP} \rightarrow \text{Squalene} \)

- **triterpenes** (C\textsubscript{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 
    **hypercholesterolaemia & heart disease** (cf. statins)
Oxidosqualene-Lanosterol Cyclase – Mechanism

- oxidosqualene-lanosterol cyclase catalyses the formation of lanosterol from 2,3-oxidósqualene:
  - 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 $\rightarrow$ 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 $\Delta^7$ 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**:

\[
\text{cholesterol} \xrightarrow{2x \text{ H}_2\text{O}} \text{estrone} \quad \text{androstendione (X = O)} \quad \text{testosterone (X = H, } \beta\text{-OH)}
\]

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

- **Vitamin D₂** is biosynthesised by the **photolytic cleavage** of **Δ⁷-dehydrocholesterol** by UV light:
  - a classic example of **photo-allowed, conrotatory electrocyclic ring-opening**:

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