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

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

19th Oct 2018
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]

  - Intense synhtetic 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

- CO$_2$ H$_2$O P$_i$ N$_2$
  - $hv$

- Lysergic acid
- Camphor
- Caffeine
- Quinine
- Clavulanic acid
- Nicotine
- Patulin
- Androstenedione
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': \( \text{hv} \rightarrow \text{ATP} \) 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, 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](http://www.johnkyrk.com/index.html)
Shikimate Metabolites
Shikimate Metabolites

- Menaquinone (vitamin K₂) ($\text{ArC}_1$)
- α-tocopherol (vitamin E) ($\text{ArC}_3$)
- Epigallocatechin (EGC) ($\text{ArC}_3$)
- Scopoletin ($\text{ArC}_3$)
- Podophyllotoxin ($\text{ArC}_3$)
The Shikimate Biosynthetic Pathway - Overview

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

  - 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

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

\[\text{inhibits this step}\]
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 $\rightarrow$ ArC$_3$ Metabolites

- **Tyrosine & phenylalanine $\rightarrow$ cinnamate derivatives $\rightarrow$ ArC$_3$ 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) 2 x ArC$_3$
  - fragment of lignin polymer 'woody' component of cell walls

Tyrosine
Phenylalanine
Primary Metabolism - Overview

**Primary metabolism**

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

1) 'light reactions': hv -> ATP and NADPH
2) 'dark reactions': CO\(_2\) -> 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)
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:

---

**Overall Stoichiometry**

- 1x 'acetate' + 1x O₂
- 2x CO₂
- 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)

\[
\text{R} \quad \overset{\text{reductive}}{\longrightarrow} \quad \overset{\text{amination}}{\longrightarrow} \\
\text{ketocacid} \\ \longrightarrow \quad \overset{\text{amino acid}}{\longrightarrow} \\
\text{lysine (Lys)} \quad [50 \text{ ATP equivs}] \\
\text{ornithine (Orn)} \quad [<44 \text{ ATP equivs (Arg)}]
\]
PLP Chemistry – *Transamination & Racemisation*

- **Transamination:**

  [Diagram showing the mechanism of transamination with chemical structures and reactions]

  - Pyridoxamine phosphate
  - Enzyme-NH₂
  - Acceptor
  - Product
  - Imine formation
  - Transaminase
  - Oxidative deamination
  - Reductive amination
PLP Chemistry – Decarboxylation

- **Decarboxylation:**

  ![Decarboxylation diagram](image)

- **Decarboxylation of *lysine* & *ornithine*:**

  ![Decarboxylation of lysine and ornithine](image)
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-laudanosoline:

  ![Chemical diagram](image)

  - Mechanism of Pictet Spengler reaction:

  ![Chemical diagram](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\textsubscript{50} 3 nM)
  - The natural ligands for these receptors are peptides: e.g. Leu-enkephalin (Tyr–Gly–Gly–Phe–Leu) (IC\textsubscript{50} 12 nM)
Dimeric Indole Alkaloids – *Vinca extracts*

Dimeric Indole Alkaloids

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

- **vinblastine** (R = Me)
- **vincristine** (R = CHO)

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

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

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

  
  \[
  \text{tryptophan} \xrightarrow{\text{PLP}} \text{tryptamine} \xrightarrow{\text{enzymatic Pictet-Spengler reaction}} \text{strictosidine}
  \]

  
  - **Mechanism of Pictet-Spengler reaction:**
    - via *spirocyclic* intermediate then *Wagner-Meerwein* 1,2-alkyl shift:
The diversity of alkaloids derived from *strictosidine* is stunning and many pathways remain to be fully elucidated.
Primary Metabolism - Overview

**Primary metabolism**

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

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

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

- **Glycolysis**
- **Glucose** & other 4,5,6 & 7 carbon sugars
- **Phosphoenol pyruvate**
- **Erythrose-4-phosphate**

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

*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](example.png)

- PROSTAGLANDINS
  - e.g.
  - prostaglandin F₂<sub>₅</sub> (PGF₂<sub>₅</sub>)

- THROMBOXANES
  - e.g.
  - thromboxane A₂ (TXA₂)

- LEUKOTRIENES
  - e.g.
  - leukotriene A₄ (LTA₄)

- EICOSANOIDS
  - matricaria ester in chamomile tea
  - obtucallene II marine natural product
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]:

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

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/](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

O

Me

SH
FATTY ACID BIOSYNTHESIS

- AT$_1$ loads malonyl group onto ACP$_1$

Pantetheine

Malonyl-CoA

• AT$_1$ loads malonyl group onto ACP$_1$
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

O

Me

SH
FATTY ACID BIOSYNTHESIS

- **ACP₁**
- **AT₁**
- **KS₁** catalyzes Claisen condensation
- **KR₁**
- **DH₁**
- **ER₁**
- **ACP₂**
- **MT₂**

- Pantetheine
- Cys
- O
- Me
- CO₂
- SH
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

O

O

S

Me

SH

SH
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 − OH

Me
FATTY ACID BIOSYNTHESIS

- DH$_1$ catalyzes dehydration of alcohol
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

SH

S

Me

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 BIOSYNTHESES

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

Pantetheine

Cys

SH

OH

Ser
FATTY ACID BIOSYNTHESIS

- MT$_2$ loads malonyl group onto ACP$_2$
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys
Me
OH
Ser
FATTY ACID BIOSYNTHESIS

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

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

Pantetheine

Cys SH

Me

Ser OH
FATTY ACID BIOSYNTHESIS

- \( \text{KR}_2 \) catalyzes reduction of ketone
FATTY ACID BIOSYNTHESIS
FATTY ACID BIOSYNTHESIS

- DH$_2$ catalyzes dehydration of alcohol

Pantetheine
FATTY ACID BIOSYNTHESIS
FATTY ACID BIOSYNTHESIS

- \( \text{ER}_2 \) catalyzes reduction of alkene
FATTY ACID BIOSYNTHESIS

The diagram illustrates the fatty acid biosynthesis process, highlighting key enzymes and their interactions. The process begins with the activation of acyl-CoA by H₁ and ER₁, followed by a series of enzymatic steps involving ACP₂, MT₂, KS₂, KR₂, DH₂, ER₂, and TE. Specific molecules such as Cys and Pantetheine play crucial roles in the synthesis pathway. The final product is a complex lipid, possibly a fatty acid, with functional groups like OH and Me.
FATTY ACID BIOSYNTHESIS

• TE catalyzes transesterification
FATTY ACID BIOSYNTHESIS
FATTY ACID BIOSYNTHESIS

• TE catalyzes hydrolysis
FATTY ACID BIOSYNTHESIS

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

Cys \( \text{SH} \)

OH

Ser

Me

O = OH
Biosynthesis of Unsaturated Fatty Acids

• **two mechanisms** are known for the introduction of double bonds into fatty acids:
  - in **BACTERIA**: *anaerobic* $[O] \rightarrow$ monounsaturated FAs (*MUFAs*)
  - in **MAMMALS, INSECTS & PLANTS**: *aerobic* $[O] \rightarrow$ *MUFAs* & polyunsaturated FAs (*PUFAs*)
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 biosynthetic steps in these pathways

![Chemical Structures](image)
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

- **linear & cyclised polyketides**
  - 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)
  - **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)

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

ACP₀  AT₀  ACP₁  AT₁  KS₁  KR₁  DH₁  ER₁  ACP₂  A

Pantetheine

Pantetheine

Pantetheine

SH

Cys

SH

SH

NB. the following sequence of slides has also been adapted from: [http://www.courses.fas.harvard.edu/%7echem27/](http://www.courses.fas.harvard.edu/%7echem27/)
• AT₀ loads starting group (propionyl) onto ACP₀
POLYKETIDE BIOSYNTHESIS

ACP<sub>0</sub> - AT<sub>0</sub> - ACP<sub>1</sub> - AT<sub>1</sub> - KS<sub>1</sub> - KR<sub>1</sub> - DH<sub>1</sub> - ER<sub>1</sub> - ACP<sub>2</sub> - A

Pantetheine

O=S

Me

SH

SH

Cys

SH
POLYKETIDE BIOSYNTHESIS

- $KS_1$ catalyzes translocation to module 1
POLYKETIDE BIOSYNTHESIS

ACP₀  AT₀  ACP₁  AT₁  KS₁  KR₁  DH₁  ER₁  ACP₂
POLYKETIDE BIOSYNTHESIS

- $AT_1$ loads methylmalonyl group onto $ACP_1$

**Diagram**: Polyketide biosynthesis pathway with components labeled as $P_0$, $AT_0$, $ACP_1$, $AT_1$, $KS_1$, $KR_1$, $DH_1$, $ER_1$, $ACP_2$, and $AT_2$. Additional components include Pantetheine and Methylmalonyl-CoA.
POLYKETIDE BIOSYNTHESIS
• KS$_1$ catalyzes Claisen condensation
POLYKETIDE BIOSYNTHESIS

P0 \rightarrow AT0 \rightarrow ACP_1 \rightarrow AT_1 \rightarrow KS_1 \rightarrow KR_1 \rightarrow DH_1 \rightarrow ER_1 \rightarrow ACP_2 \rightarrow AT_2

Pantetheine

Cys

SH

Stereocenter

\text{Me}
• KR$_1$ catalyzes reduction of ketone
POLYKETIDE BIOSYNTHESIS

P₀ → AT₀ → ACP₁ → AT₁ → KS₁ → KR₁ → DH₁ → ER₁ → ACP₂ → AT₂

Pantetheine

Cys

S

Me

OH

Me

Stereocenter
POLYKETIDE BIOSYNTHESIS

- no DH$_1$ activity
POLYKETIDE BIOSYNTHESIS

- no ER₁ activity
• KS₂ catalyzes translocation to module 2
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 ‘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:

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

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

- 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)**
  - **hexaketides (6x C₂)**
  - **heptaketides (7x C₂)**
  - **octaketides (8x C₂)**
  - **nonaketides (9x C₂)**
  - **decaketides (10x C₂)**

- 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 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 & polycyclic compounds: saturated fatty acids, unsaturated fatty acids, lipids, 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)
Biosynthesis of IPP & DMAPP - via Mevalonate

- IPP & DMAPP are the key C5 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.2015.07.069))
  - *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)]

---

**Lovastatin PRO-DRUG**

*NB.* type I (iterative) PKS natural product

**ACTIVE DRUG**

mimic of tetrahedral intermediate in HMG reduction by NADPH

**Lipitor**

- NADPH
- NADP⁻
- CoASH
- HMG CoA
- mevalonate (MVA)
- cholesterol
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 $S_N 1$ alkylation of DMAPP by IPP → monoterpenes
  - farnesyl (C$_{15}$) & geranylgeranyl (C$_{20}$) pyrophosphates are formed by further $S_N 1$ alkylations with IPP:
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**:

---

\[ \text{gerenyl pyrophosphate} \quad \text{linalyl pyrophosphate} \quad \text{neryl pyrophosphate} \]

\[ \text{allylic cation \; intimate ion pair} \quad \text{cyclase} \quad \text{MONOTERPENES (C}_{10} \text{)} \]
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**

```latex
\begin{align*}
\text{limonene} & \xrightarrow{E_1 \text{ elimination}} \text{α-terpineol} \\
\text{α-terpinyl cation} & \xrightarrow{1,2-hydride \text{ shift}} \text{thujone} \\
\text{bornyl pyrophosphate} & \xrightarrow{\text{trapping by PPO}} \text{borneol} \\
\text{H}_2\text{O} & \xrightarrow{\text{trapping with water}} \text{camphor} \\
\text{α-terpinyl cation} & \xrightarrow{\text{trapping by alkene at 'red' carbon (anti-Markovnikov)}} \text{camphene} \\
\text{E}_1 \text{ elimination} & \xrightarrow{\text{trapping by alkene at 'blue' carbon (Markovnikov)}} \text{β-pinene} \\
\text{E}_1 \text{ elimination} & \xrightarrow{\text{E}_1 \text{ elimination}} \text{α-pinene}
\end{align*}
```
Sesquiterpenes – **Farnesyl Pyrophosphate (FPP)**

- ‘\( S_{N2} \)'-like alkylation of *geranyl PP* by *IPP* gives *farnesyl PP*:

\[
\text{geranyl PP} \xrightarrow{\text{pro-R hydrogen is lost}} \text{farnesyl PP (FPP)}
\]

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

\[ \text{E,E-farnesyl PP (FPP)} \]

\[ \text{E,Z-farnesyl PP} \]

\[ \text{nerolidyl PP} \]

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

**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 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)
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
  2. $14\alpha$ DEMETHYLATION
  3. $4\alpha$ & $4\beta$ DEMETHYLATION
  4. $\Delta^8$ to $\Delta^7$ rearrangement

Chemical reactions:
1. $\Delta^{24}$ hydrogenation
   - NADPH to NADP$^+$
   - $24$ hydrogenation

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

3. $4\alpha$ & $4\beta$ DEMETHYLATION
   - $4x\ O_2$ to $P_{450}$
   - $4x\ H_2O$
   - CO$_2$

4. $\Delta^8$ to $\Delta^7$ rearrangement
   - Isomerase
   - NAD$^+$ to NADH
   - NADH to NAD$^+$

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 $P_{450}$ enzymes
  - **estrone** is produced from **androstendione** by oxidative demethylation with concomitant aromatisation:

\[
\text{cholesterol} \xrightarrow{2x \text{O}_2, \text{P}_{450}} \text{2x H}_2\text{O} \xrightarrow{2x \text{O}_2, \text{P}_{450}} \text{progesterone} \xrightarrow{\text{NAD}^{\ominus}} \text{estrone (œstrone)} \xrightarrow{\text{HCO}_2\text{H}} \xrightarrow{2x \text{O}_2, \text{P}_{450}} \text{2x H}_2\text{O} \xrightarrow{\text{androstendione (X = O)}} \text{testosterone (X = H, } \beta\text{-OH)}
\]

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

- **vitamin D$_2$** 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; defficiency** 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:
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