Biosynthesis

Biosynthesis of Fatty Acids & Polyketides

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

Dec 2014
Format & Scope of Lectures

• **What are fatty acids?**
  – 1° metabolites: fatty acids; 2° metabolites: their derivatives
  – biosynthesis of the building blocks: acetyl CoA & malonyl CoA

• **Fatty acid synthesis by Fatty Acid Synthases (FASs)**
  – the chemistry involved
  – the FAS protein complex & the dynamics of the iterative synthesis process

• **Fatty acid secondary metabolites**
  – polyacetylenes
  – eiconasiods: prostaglandins, thromboxanes & leukotrienes
  – branched and cyclopropanated fatty acid derivatives

• **What are polyketides?**
  – definitions & variety
  – $^{13}$C labelling techniques

• **Polyketide synthesis by PolyKetide Synthases (PKSs)**
  – the chemistry involved
  – the PKS protein complexes & the dynamics of the iterative synthesis process

• **Polyketide secondary metabolites**
  – Type I modular metabolites: macrolides – e.g. erythromycin & rapamycin
  – Type I iterative metabolites: e.g. mevinolin (=lovastatin®)
  – Type II iterative metabolites: aromatic compounds and polyphenols: e.g. actinorhodin etc.
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 $\rightarrow$ citric acid cycle
  - **Biosynthetic precursors** – for elaboration to secondary metabolites

---

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

*e.g.*

- caprylic acid ($C_8, n = 2$)
- capric acid ($C_8, n = 3$)
- lauric acid ($C_{12}, n = 4$)
- myristic acid ($C_{14}, n = 5$)
- palmitic acid ($C_{16}, n = 6$)
- stearic acid ($C_{18}, n = 7$)

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

- palmitoleic acid ($C_{16}, Z - \Delta^9$)
- oleic acid ($C_{18}, Z - \Delta^9$)

* (>80% of fat in olive oil )

**POLY-UNSATURATED ACID DERIVATIVES (PUFAs)**

*e.g.*

- arachidonic acid (AA) ($C_{20}, Z - \Delta^5, Z - \Delta^8, Z - \Delta^{11}, Z - \Delta^{14}$)
- eicosapentaenoic acid (EPA) ($C_{20}, Z - \Delta^5, Z - \Delta^8, Z - \Delta^{11}, Z - \Delta^{14}, Z - \Delta^{17}$)

*(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)

![Chemical structures](image)

- **Polyacetylenes**
  - e.g. wyerone

- **Prostaglandins**
  - e.g. prostaglandin F2α (PGF2α)
  - e.g. thromboxane A2 (TXA2)

- **Thromboxanes**
  - e.g. leukotriene A4 (LTA4)

- **Leukotrienes**
  - e.g. leukotriene A4 (LTA4)

- **Eicosanoids**
Primary Metabolism - Overview

**Primary metabolism**

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

1) 'light reactions': \( hv \rightarrow \text{ATP and NADH} \)
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
  - ALKALOIDS
    - penicillins
    - cephalosporins
    - cyclic peptides
  - FATTY ACIDS & POLYKETIDES
    - polyacetylenes
    - prostaglandins
    - aromatic compounds, polyphenols
    - macrolides
  - ISOPRENOIDS
    - terpenoids
    - steroids
    - carotenoids
**Biosynthesis of Malonyl Coenzyme A**

- *Malonyl coenzyme A* is the key ‘extender unit’ for the biosynthesis of **fatty acids** (and **polyketides**):
  - is formed by the **carboxylation** of *acetyl coenzyme A* mediated by a **biotin-dependent enzyme**
  - this is the **first committed step of fatty acid/polyketide biosynthesis** (and is a rate controlling step)
Oxidative Decarboxylation of Pyruvate

- **Oxidative decarboxylation** of pyruvate is catalysed by the **Pyruvate Dehydrogenase Complex (PDC)**
  - PDC is a huge complex comprising many copies of each of 3 enzymes:
    - $24 \times E_1$ Pyruvate dehydrogenase; $24 \times E_2$ Dihydrolipoyl transferase; $12 \times E_3$ Dihydrolipoyl dehydrogenase
  - Pyruvate dehydrogenase effects the key decarboxylation using thiamine pyrophosphate as a cofactor

![Thiamine pyrophosphate (TPP)](image)

**Thiamine pyrophosphate (TPP)**
a vitamin $B_1$ derivative

![Pyruvate dehydrogenase complex](image)

**E₁ pyruvate dehydrogenase**

**E₂ dihydrolipoyl transferase**

**E₃ Dihydrolipooyl dehydrogenase**
Oxidative Decarboxylation of Pyruvate

- The Pyruvate Dehydrogenase Complex (PDC)
  - [http://www.bmsc.washington.edu/WimHol/figures/figs5/WimFigs5.html](http://www.bmsc.washington.edu/WimHol/figures/figs5/WimFigs5.html)
Biosynthesis of Malonyl Coenzyme A

- **Bicarbonate** is the source of the $CO_2$:
  - the bicarbonate is first **activated** via **phosphorylation** by **ATP**
  - then the **phosphorylated bicarbonate** carboxylates **biotin** to give **carboxybiotin**
  - then the **carboxybiotin** carboxylates the enolate of **acetyl CoA** to give **malonyl CoA**:

- the carboxylation of biotin & acetyl CoA are mediated by a **single biotin-dependent enzyme (complex)** having both **biotin carboxylase** and **transcarboxylase active sites**
- *NB.* coupling to ATP ‘hydrolysis’ provides **energy** to drive carboxylation processes
Acetyl CoA Carboxylase

- the biotin co-factor is swung between two active sites:

bicarbonate coupled to biotin

transfer to acetyl CoA
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 CO$_2$) – a *decarboxylative Claisen condensation*
  - 3-step *reduction* of the resulting ketone → methylene
- after $n = 3\text{-}8$ iterations the **C8-20 saturated fatty acid** is released from the enzyme(s):
The Decarboxylative Claisen Condensation (dCc)

- *in vitro* – the classical *Claisen condensation*:

  ![](image1)

- *in vivo* - the *decarboxylative Claisen condensation* catalysed by a *ketosynthase (KS)*

  ![](image2)

  - the energy released upon loss of CO₂ provides a driving force for the condensation
  - thioesters are also particularly reactive partners in this type of condensation...
The Claisen Condensation - Why Thioesters?

- recall the chemistry of coenzyme A (1st lecture) – properties of alkyl thioesters (cf. alkyl esters)
  - good leaving group ability of RS: (cf. RO⁻)
    - due to pKₐ (RSH) ~10 cf. pKₐ (ROH) ~16

- high acidity of protons α to the carbonyl of thioesters (cf. ester) & weak C-S bond (cf. C-O bond):
  - due to poor orbital overlap between the lone pairs on sulfur (n_S) [cf. n_O] and the carbonyl anti bonding orbital π⁺_C=O

\[
\begin{align*}
\begin{array}{c}
\text{Nu}^- \quad \text{Nu}^- \quad \text{Nu}^- \\
\text{H} \quad \text{H} \quad \text{H}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{S} \quad \text{S} \quad \text{S}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{R} \quad \text{R} \quad \text{R}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{Nu}^- \quad \text{Nu}^- \quad \text{Nu}^- \\
\text{OR} \quad \text{OR} \quad \text{OR}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{S} \quad \text{S} \quad \text{S}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{R} \quad \text{R} \quad \text{R}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{Nu}^- \quad \text{Nu}^- \quad \text{Nu}^- \\
\text{OR} \quad \text{OR} \quad \text{OR}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{S} \quad \text{S} \quad \text{S}
\end{array}
\end{align*}
\]
\[
\begin{align*}
\begin{array}{c}
\text{R} \quad \text{R} \quad \text{R}
\end{array}
\end{align*}
\]

- less effective
  - pKₐ = 20
  - acyl-sulfur bond weak

- more effective
  - pKₐ = 25
  - acyl-oxygen bond strong
Ketone → Methylene - Reduction

- ketone → methylene reduction is achieved via a 3-step process:
  1. NADPH-mediated ketone → alcohol reduction catalysed by a keto reductase (KR)
  2. syn-elimination of water catalysed by a dehydratase (DH)
  3. NADPH-mediated hydrogenation of the double bond catalysed by an enoyl reductase (ER)

- all steps are generally stereospecific but stereospecificity varies from organism to organism
  - indicated specificities are for human FAS
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]**:

*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
The Acyl Carrier Protein (ACP)

- The **Acyl Carrier Protein (ACP)** is the key protein that allows the growing oligomer to access the appropriate active sites.
- The ACP is first **primed** by the post-translational modification of one of its serine hydroxyl groups:
  - the introduction of a *phosphopantetheine ‘swinging-arm’* by reaction with *acetyl coenzyme A*:
    - this swinging-arm provides **flexibility** for module-module acyl transfer & provides **binding energy** for catalysis
    - the ACP is inactive prior to priming
Human Fatty Acid Synthase (FAS)

- Human FAS (EC 2-3-185) is a type I FAS – a homodimer of a multifunctional protein (272 kDa)
  - each monomer is 'barrel' shaped with diameter ~210 Å & length ~250 Å
  - each subunit protein contains seven catalytic activities plus the acyl carrier protein (ACP)

- NB. keto synthases (KS) are also sometimes referred to as condensing enzymes (CE)
Human Fatty Acid Synthase (FAS)

• the first three-dimensional structure of human fatty acid synthase at 4.5 Å resolution by X-ray crystallography:

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)

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

- Pantetheine

NB. the following sequence of slides have been adapted from: [http://www.courses.fas.harvard.edu/~7echem27/](http://www.courses.fas.harvard.edu/~7echem27/)

- Cys
- SH
- SH
- SH
• AT₁ loads acetyl group onto KS₁
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

O

Me

SH
• AT<sub>1</sub> loads malonyl group onto ACP<sub>1</sub>
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

O

Me

SH
FATTY ACID BIOSYNTHESIS

• KS₁ catalyzes Claisen condensation
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

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

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

Pantetheine

Cys

SH

O

OH

Me
FATTY ACID BIOSYNTHESIS

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

- DH₁ catalyzes dehydration of alcohol
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

SH

S

Me
• $ER_1$ catalyzes reduction of alkene

FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys $\leftrightarrow$ SH

S $\rightarrow$ O $\rightarrow$ Me
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys

SH

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

SH

Cys

Me

OH

Ser
FATTY ACID BIOSYNTHESIS

- \( MT_2 \) loads malonyl group onto \( ACP_2 \)
FATTY ACID BIOSYNTHESIS

The process involves the following enzymes:

1. $H_1$ → $ER_1$ → $ACP_2$ → $MT_2$ → $KS_2$ → $KR_2$ → $DH_2$ → $ER_2$ → $TE$

Each enzyme plays a specific role in the biosynthesis pathway. For example, $ACP_2$ is involved in the activation of fatty acids, $MT_2$ in their elongation, $KS_2$ in the desaturation reaction, $KR_2$ in the reductive condensation, $DH_2$ in the dehydration, and $ER_2$ in the elongation. The final product is coupled with a Ser residue through an ester linkage.
• KS$_2$ catalyzes Claisen condensation
FATTY ACID BIOSYNTHESIS

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

Pantetheine

Cys SH

Me

O=S

O=S

Ser OH
• KR\textsubscript{2} catalyzes reduction of ketone
FATTY ACID BIOSYNTHESIS

\[ \begin{align*}
\text{H}_1 & \rightarrow \text{ER}_1 \\
& \rightarrow \text{ACP}_2 \rightarrow \text{MT}_2 \rightarrow \text{KS}_2 \rightarrow \text{KR}_2 \rightarrow \text{DH}_2 \rightarrow \text{ER}_2 \rightarrow \text{TE}
\end{align*} \]

- Pantetheine
- Cys
- Me
- Ser
- OH

\[ \text{S} \quad \text{O} \quad \text{OH} \]

\[ \text{S} \quad \text{O} \quad \text{OH} \]
FATTY ACID BIOSYNTHESIS

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

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

Pantetheine

\[ \text{Cys} \rightarrow \text{SH} \rightarrow \text{S} \rightarrow \text{C=O} \rightarrow \text{Me} \]

\[ \text{Ser} \rightarrow \text{OH} \]
FATTY ACID BIOSYNTHESIS

- ER₂ catalyzes reduction of alkene
FATTY ACID BIOSYNTHESIS
• TE catalyzes transesterification
FATTY ACID BIOSYNTHESIS
FATTY ACID BIOSYNTHESIS

• TE catalyzes hydrolysis
FATTY ACID BIOSYNTHESIS
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)

**ANAOBERIC 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
Biosynthesis of Polyacetylenes

- A family of over 1000 natural products!
  - review: Tykwinski Angew. Chem. Int. Ed. 2006, 45, 1034 (DOI)
- Few detailed pathways have been established but generally involve sequential dehydrogenations:
  - e.g. biosynthesis of matricaria ester (Matricaria chamomilla):
    - component of chamomile tea

```
\begin{align*}
\text{stearate} & \xrightarrow{\text{Enz}} \text{oleate} \\
\text{oleate} & \xrightarrow{\text{Enz}} \text{linoleate} \\
\text{linoleate} & \xrightarrow{\text{Enz}} \text{crepenynate} \\
\text{crepenynate} & \xrightarrow{\text{Enz}} \text{matricaria ester}
\end{align*}
```

Matricaria chamomilla
Biosynthesis of Prostaglandins & Thromboxanes

- **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
Biomimetic Synthesis of Prostaglandins

- In 1984 Corey published a classic biomimetic total synthesis of prostaglandins

\[
\text{HO} \quad \text{O} \quad \text{O}
\]

\[
\text{C}_5\text{H}_{11}
\]

\[
\text{OH}
\]

\[
\text{prostacyclin (PGI}_2\text{)}
\]

\[
\text{HO} \quad \text{O}
\]

\[
\text{O}
\]

\[
\text{C}_5\text{H}_{11}
\]

\[
\text{OMe}
\]

\[
\text{ClHg}
\]

\[
\text{Bu}_3\text{SnH}
\]

\[
\text{O}_2
\]

\[
\text{disrotatory (only syn products)}
\]

**Biosynthesis of Leukotrienes**

- **leukotrienes** are the other main class of 2° metabolites derived from *arachidonic acid*
  - they are potent (<nM) *inflammatory substances* released during allergic reactions

\[
\begin{align*}
\text{leukotriene } A_4 (LTA_4) & \quad \text{LTC}_4 \text{ synthase} \\
\text{leukotriene } B_4 (LTB_4) & \quad \text{LTA}_4 \text{ hydrolase}
\end{align*}
\]
Branched & Cyclopropanated Fatty Acids

- fatty acid metabolites occasionally contain 'extra' methyl groups:
  - there are **two methods** by which these are added:
    - by use of a different extender unit – *methyl malonyl CoA*:
    
      \[
      \text{Met, Val or Ile} \rightarrow \text{CoA} \xrightarrow{\text{biotin-dependent carboxylase}} \text{propionyl CoA} \xrightarrow{\text{FAS}} \text{methyl malonyl CoA}
      \]
    
    - by SAM-mediated *methylation/cyclopropanation* process:

      \[
      \text{Me} \xrightarrow{\text{SAM}} \text{CoA} \xrightarrow{\text{Me}} \text{CoA} \xrightarrow{\text{NADPH}} \text{Me}
      \]

      *tuberculostearic acid*

      *dihydrosterculic acid*
The Polyketide Pathway

- **Polyketides** are also sometimes known as **acetogenins**
- **Acetyl CoA** is also the starting point for the biosynthesis of **polyketide** secondary metabolites
- These metabolites are topologically very different to the fatty acid metabolites but are in fact synthesised in a very similar fashion. The significant difference is that during the iterative cycle of chain extension **the β-keto group is generally not completely reduced out**. This gives rise to huge structural diversity based around a 1,3-oxygenation pattern & cyclisation to give aromatic compounds

\[ \text{FATTY ACIDS (FAs)} \]
\[ \text{POLYKETIDES (PKs)} \]

- **NB.** unlike fatty acids, polyketides are NOT biosynthesised by humans – only microorganisms (bacteria) & fungi
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

![Chemical structures](image-url)
Historical Perspective – ‘The Acetate Hypothesis’

**1907: James Collie** (University of London) converts **dehydroacetic acid** to **orcinol** by boiling with \( \text{Ba(OH)}_2 \) (while trying to deduce the structure of the former):

Collie perceptively postulated the **triketone** as an intermediate & suggested that this might also be an **intermediate** in the **biosynthesis** of **orcinol** (the ‘polyketide hypothesis’)

**1955: Arthur Birch** used \(^{14}\text{C} \) labelled acetate to show that 6-methylsalicylic acid (ex. *Penicillium patulum*) was biosynthesised by head-to-tail oligomerisation of **4 × acetate units** and proposed the following biogenesis – proceeding **via a tetraketide intermediate** (cf. Collie!):
**Isotopic Labelling Studies – Use of $^{13}$C**

- **1970s:** Commercial availability of $^{13}$C & $^2$H labelled precursors & NMR instruments allowed rapid determination of labelling patterns of polyketides (cf. radiolabelling/degradation)

- **single labelled acetate** $\{[1-^{13}\text{C}]$- or $[2-^{13}\text{C}]$-acetate, cf. $^{14}$C label used by Birch\}
  1. feed ~99.9% $^{13}$C enriched acetate
  2. verify uniform incorporation along backbone (ideally obtain incorporation to give ~ doubling of signal size)
  3. assign positions of labelled carbons by reference to standard $^{13}$C spectrum

- **double labelled acetate** $\{[1,2-di-^{13}\text{C}]$-acetate\}
  1. feed >90% 2× $^{13}$C enriched acetate
  2. observe pairs of $^{13}$C-$^{13}$C coupled doublets (NB. again incorporation to a level representing ~ doubling of signal sizes is standard; since natural abundance is ~1% this amounts to ~1% incorporation...this ensures that statistically very few (1 in $10^4$) labelled acetates will be sequentially incorporated & result in inter-unit coupling patterns)
  3. assign positions of labelled carbons by reference to standard $^{13}$C spectrum

- **e.g.**

```
\begin{align*}
\text{4x} & \quad \text{ACOH} \quad \text{PKS} \\
\text{[1,2-$^{13}$C]-acetate} & \quad 3 \times \text{H}_2\text{O} \\
\text{tetraketide} & \quad \text{orsellinic acid} \\
\text{J} = 44\text{Hz} & \quad \text{penicillic acid} \quad \text{J} = 77\text{Hz}
\end{align*}
```

---

The diagrams represent the chemical transformations involved in the synthesis of orsellinic acid and penicillic acid, illustrating the incorporation of labelled carbons through stepwise reactions and the observation of coupling patterns in NMR spectra.
Isotopic Labelling Studies – Use of $^{13}\text{C}/^{2}\text{H}$

- the ‘$\alpha$-shift’ technique for following the fate of hydrogens:
  - allows # of hydrogens (deuteriums) retained at the labelled carbon following biosynthesis → evidence of redox processing etc.
  - useful for identifying starter units e.g.

- $^{13}\text{C}$ & $3\times^{2}\text{H}$ labelled acetate $\{[2^{13}\text{C},^{2}\text{H}_3]\text{-acetate}\}$
  1. feed multiply labelled acetate
  2. observe shifts of labelled carbons
  3. assign positions of labelled carbons & determine fate of attached hydrogens

\[ \text{[2-}^{13}\text{C,}^{2}\text{H}_3\text{-acetate]} \]

- $5x$ D, D, D, D, D

- PKS

- $4x$ H$_2$O

- pentaketide

- terrein
**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

  
  \[
  \text{CoA}S\xrightarrow{\text{Enz}} \text{R} \xrightarrow{\text{S}_{\text{Enz}}} \text{R} \xrightarrow{\text{CO}_2} \text{R} \xrightarrow{\text{Enz}} \text{R'} \xrightarrow{\text{OH}} \text{R'} \xrightarrow{\text{[funct]}^\#1} \text{R} \xrightarrow{\text{[funct]}^\#2} \text{R} \xrightarrow{\text{[funct]}^\#3} \text{R} \xrightarrow{\text{dCc}^\#1} \text{R} \xrightarrow{\text{dCc}^\#2} \text{R} \xrightarrow{\text{dCc}^\#3} \text{R} \]

  
  - **linear & cyclised polyketides**
  - this leads to **enormous diversity**...
Polyketide Diversity

- **starter units:**
  - R = Me
  - R = Et
  - R = 'Pr & 'Pr
  - CoAS
  - acetyl CoA
  - propionyl CoA
  - butyryl CoA
  - isobutyril CoA

- **extender units:**
  - CoAS
  - malonyl CoA (C₂)
  - Me-malonyl CoA (C₃)
  - Et-malonyl CoA (C₄)

- **non-functional or missing KR, DH, ER:**
  - no KR
  - no ER
  - no DH
  - none missing

- **stereochemistry:**
  1) side chain stereochemistry (determined by KSₙ)
  2) OH stereochemistry (determined by KRₙ₊₁)
  3) alkene stereochemistry (determined by DHₙ₊₁)

- **termination step:**
  - depends on nucleophile that releases product at TE stage:
    - Nu = H₂O
    - Nu = OH
    - Nu = CoA
    - NADH
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 ‘iteration’ has a dedicated set of catalytic sites (→ 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 (see later)
    - **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 → SH → SH

Cys → 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

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

Pantetheine

Cys

SH

SH

O

S

Me
• KS₁ catalyzes translocation to module 1
• **AT\textsubscript{1}** loads methylmalonyl group onto ACP\textsubscript{1}
POLYKETIDE BIOSYNTHESIS

Pantetheine

SH

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> - AT<sub>2</sub>
**POLYKETIDE BIOSYNTHESIS**

- **KS₁** catalyzes Claisen condensation
POLYKETIDE BIOSYNTHESIS

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

Pantetheine

Stereocenter
• KR₁ catalyzes reduction of ketone
POLYKETIDE BIOSYNTHESIS

Pantetheine

SH

Stereocenter
POLYKETIDE BIOSYNTHESIS

- no DH$_1$ activity
• no ER₁ activity
**POLYKETIDE BIOSYNTHESIS**

- KS₂ catalyzes translocation to module 2
• **Electron cryo-microscopy has recently thrown additional light on how this process works for the Type 1 PKS that synthesises pikromycin in Streptomyces venezuelae.**
  – For a video of the process see: [http://cen.acs.org/articles/92/i25/Polyketide-Synthase-Secrets-Revealed.html](http://cen.acs.org/articles/92/i25/Polyketide-Synthase-Secrets-Revealed.html)
‘Deconvolution’ of Type I(modular) PKSs

• **deduce the module structure for the type I modular PKS responsible for the synthesis of this hexaketide:**

1. identify the **last building block**:  
   - i.e. \[\text{O}\] \[\begin{array}{c}
   \text{\text{O}} \\
   \text{\text{HO}} \\
   \text{\text{or}} \\
   \text{\text{O}} \\
   \text{\text{HO}}
   \end{array}\]

2. identify **each extender unit** (working back from the last one):  
   - 2C in the **backbone**  
   - + 0, 1, 2 (or more) C in the **sidechain**

3. identify the **starter unit**:  
   - the module that appended this unit is designated **module 0**

4. deduce what happens to each ketone:  
   
   **NB.** module \(n\) modifies the ketone of the building block added by module \(n-1\)

<table>
<thead>
<tr>
<th>module#</th>
<th>#5</th>
<th>#4</th>
<th>#3</th>
<th>#2</th>
<th>#1</th>
<th>#0</th>
</tr>
</thead>
<tbody>
<tr>
<td>starting material</td>
<td>Mal-CoA</td>
<td>Et-Mal-CoA</td>
<td>Mal-CoA</td>
<td>Me-Mal-CoA</td>
<td>Mal-CoA</td>
<td>Ac-CoA</td>
</tr>
<tr>
<td>ACP</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>AT</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>KS</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>KR</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>DH</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>ER</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>TE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>
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)

---

**Diagram**: 
- Loading
- DEBS1 (modules 1-2)
- DEBS2 (modules 3-4)
- DEBS3 (modules 5-6)
- Release
- **erythronolide B**

- **AT** = Acyl Carrier Protein
- **KR** = Ketosynthase
- **KS** = Ketolactonase
- **TE** = Thioesterase
**Biosynthesis of Rapamycin – Type I(modular) PKS**

- **rapamycin** – *bacterial* immunosuppressant used in organ transplant surgery:
  - mixed polyketide (*acetate* & *propionate*)/peptide with novel cyclohexyl carboxamide starter unit
  - 3 multifunctional polypeptides with 70 catalytic functions!
  - RAPS1 (~900 kDa, 4 modules), RAPS2 (1.07 MDa, 6 modules), RAPS3 (660 kDa, 4 modules)
Biomimetic Decarboxylative Thioester Aldol

- recall the key C-C bond forming process in both FAS and PKS chain extension is a decarboxylative Claisen condensation of enzyme thioester-bound acetyl and malonyl residues:

  - Shair has developed an exceptionally mild aldol reaction of malonic acid half thioesters (MAHTs) inspired by this process:
    - Shair et al. J. Am. Chem. Soc. 2003, 125, 2852 (DOI)

  {\[
  \begin{align*}
  & \text{aldehyde} \\
  & \text{outcome} \\
  & \text{Ph} & 3.5 \text{h}, 82\% \\
  & \text{3.5h, 85\%} \\
  & \text{2.5h, 81\%} \\
  & \text{8h, 82\%}
  \end{align*}
  \]}

  - decarboxylative Claisen condensation
  - decarboxylative thioester aldol

  \[
  \begin{align*}
  \text{CO}_2 & \\
  \text{KS} & = \text{keto synthase} \\
  \text{ACP} & = \text{acyl carrier protein}
  \end{align*}
  \]
Biomimetic Iterative Claisen-Like Condensations

- Harrison has developed a glycoluril ‘template’ to mimic the proximal ketosynthase (KS) & acyl carrier protein (ACP) units in FAS and PKS and achieved iterative chain extension of up to eight carbons:

\[
\begin{align*}
\text{glycoluril} & \quad \text{Ac}_2\text{O}, \Delta \text{20h} (=A) \quad \text{[89%]} \\
\text{\textsuperscript{6}BuLi, THF, 1h} \text{then AcCl} (=B) & \quad \text{[78%]} \\
\text{\textsuperscript{t}BuOLi, THF, 0°C 20min} (=C) & \quad \text{[93%]} \\
\text{NaBH}_4, \text{MeOH, 10min then AcOH} (=D) & \quad \text{[93%]}
\end{align*}
\]

\[
\begin{align*}
\text{[89%]} & \quad \text{[78%]} & \quad \text{[93%]} & \quad \text{[72%]} & \quad \text{[83%]} \\
\text{[44%]} & \quad \text{[60%]} & \quad \text{[55%]} & \quad \text{[86%]} & \quad \text{[86%]}
\end{align*}
\]

'tetraketide' cleavage from template
Biosynthesis of Mevinolin – *Type I (iterative)* PKS

- **mevinolin (=lovastatin®)** – cholesterol lowering metabolite of filamentous *fungus* *Aspergillus terreus*
  - inhibits HMG-CoA → mevalonate (see next lecture) – rate-limiting step in biosynthesis of *cholesterol*
  - acetate based polyketide composed of a diketide and nonaketide linked by an ester
  - 2 × Type I (iterative) PKSs: LNKS and LDKS...both contain *MeT* (*methyl transferase*) activities
  - Hutchinson *et al.* *Science* 1999, 284, 1368 ([DOI](https://doi.org/10.1126/science.284.5413.1368))

\[
\begin{align*}
\text{Acetate} & \rightarrow \text{Mevinolin (=lovastatin®)} \\
\end{align*}
\]
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’ β* (KS$_{\beta}$) to give S-acetyl-ACP
  - the *extender malonates* undergo *decarboxylative Claisen condensations* by *ketosynthase α* (KS$_{\alpha}$)
- these clusters rarely utilise KR, DH or ER activities and produce ‘true’ polyketides:

KS$_{\beta}$ = ‘ketosynthase β’ (=decarboxylase!); KS$_{\alpha}$ = ‘ketosynthase α’ (=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*...
Biomimetic Synthesis of Quinone Antibiotics

- Pioneered by Harris. *e.g.* classic biomimetic synthesis of *chrysophanol*:
  - position of ‘reduced’ ketone dictates cyclisation site

- Abell & Staunton’s biomimetic syntheses of *rubrofusarin* & *alternariol*:
  - timing of pyrone ring formation dictates subsequent cyclisation-aromatisation pathway

Biosynthesis of Citrinin - *Type II PKS*

- **Citrinin** is a liver toxic metabolite of the mould *Penicillium citrinum*

  - **Note:**
    - role of **SAM** for introduction of methyl groups
    - *P₄₅₀* then **NADP⁺** for *Me* → *CO₂H* oxidation...
Methylation by SAM

- Methylation at carbon by SAM takes place at the 2-position of 1,3-dicarboxyls (or at the 2-position of phenols):

![Methylation Diagram](image-url)
Oxidation by P₄₅₀ Enzymes

- Hydroxylation at unactivated CH positions is achieved by the haem co-factor in P₄₅₀ enzymes:
Griseofulvin Biosynthesis - Type II PKS

- **Griseofulvin** is a mould metabolite (*Penicillium griseofulvum, Penicillium janczewskii*) used to treat worm infections in animals and humans
  - Birch delineated the basic biogenesis in the 1950s & in 1959 griseofulvin was marketed as an oral fungicide by ICI & Glaxo (as Fulcin® & Grisovin®, respectively)
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**

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

**PHOTOSYNTHESIS**

1) 'light reactions': \( hv \to \text{ATP and NADH} \)
2) 'dark reactions': \( \text{CO}_2 \to \text{sugars (Calvin cycle)} \)

**glycolysis**

- glucose
- & other 4,5,6 & 7 carbon sugars

**phosphoenol pyruvate**

**erythrose-4-phosphate**

\[ \text{shikimate} \]

**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

**Citric acid cycle (Krebs cycle)**

- pyruvate
- acetyl coenzyme A
- malonyl coenzyme A
- mevalonate