Biosynthesis of Natural Products

Introduction to Secondary Metabolism & the Shikimate Pathway

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

Imperial College
London

Oct 2019
Lessons in Synthesis - Azadiractin

- **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

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) $\rightarrow$ 5-enolpyruvylshikimate-3-phosphate (5-EPS-3P)
  - a non-selective herbicide

\[
\text{phosphoenol pyruvate (PEP)} \rightarrow \text{3-phosphoshikimate (3-PS)} \rightarrow \text{5-enolpyruvylshikimate-3-phosphate (5-EPS-3P)}
\]

- **glyphosate (Roundup®)** inhibits this step
Inspiration for Med Chem - 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**
  - *e.g. Lipitor* (Atorvastatin calcium, Pfizer) is 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)]
Format & Scope of Lecture

- **What is biosynthesis?**
  - some definitions – phototrophs, chemotrophs; metabolism (catabolism/anabolism), 1° & 2° metabolites

- **Overview of primary metabolism → secondary metabolites**

- **Biological/biosynthetic reactions – enzyme & cofactor chemistry**
  - free energy source – ATP
  - C-C & C-O bond formation – CoASH, SAM, DMAPP, biotin
  - oxidation – NAD⁺, FAD/FMN, haem iron oxo monooxygenases
  - reduction – NADPH
  - C-N bond formation – pyridoxal

- **The Shikimate Pathway**
  - phenylalanine, tyrosine, tryptophan
  - coumarins, lignans & lignins
Metabolism & Natural Product Diversity

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

\[ \text{CO}_2 \text{ H}_2\text{O} \text{ P}_i \text{ N}_2 \text{ hv} \]
Phototrophs & Chemotrophs

- **Living organisms** are not at equilibrium. They **require** a continuous influx of **free energy** to perform mechanical work & for cellular growth/repair:

  - **Phototrophs** (e.g. green plants, algae & photosynthetic bacteria): derive free energy from the sun via **photosynthesis** (‘CO₂ fixation’):
    - 10¹⁵ kg/year by green plants, which constitute 99% of Earths biomass (*i.e.* 10¹² tons of dry matter)
    - 1g of carbon processed = >6250 litres of air

  ![Photosynthesis](image)

  \[ \text{CO}_2 + \text{H}_2\text{O} \xrightarrow{hv} (\text{CHO}) + \text{O}_2 \]

  - **Chemotrophs** (e.g. animals, fungi, most bacteria): derive free energy by **oxidising nutrients** (e.g. carbohydrates, lipids, proteins) obtained from other organisms, ultimately phototrophs
    - some bacteria & fungi require just D-glucose
    - mammals require sugars, essential amino acids & certain vitamins (enzyme co-factors or precursors)

  - Degradation of the nutrients is coupled to the stoichiometric production of ‘high energy’ phosphate compounds, particularly adenosine triphosphate (ATP, see later). All metabolic function is underpinned by ATP energetic coupling

  - By analogy with a money-based economy, the metabolic cost of production of a given metabolite from another can be quantified in terms of ‘ATP equivalents’ defined as the # of moles of ATP consumed/produced per mole of substrate converted in the reaction or sequence
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

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

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

**glycolysis**

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

**phosphoenol pyruvate**

**erythrose-4-phosphate**

**pyruvate**

**acetyl coenzyme A**

**citric acid cycle (Krebs cycle)**

**malonyl coenzyme A**

**mevalonate**

**shikimate**

- **aromatic amino acids**
- **aliphatic amino acids**

**peptides proteins tetrahydropyrroles (porphyrins)**

**saturated fatty acids unsaturated fatty acids lipids**

**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)
Biological/Biosynthetic Reactions – Enzyme Catalysis & Cofactors

• Most biosynthetic steps are catalysed by specific, individual enzymes. They generally perform familiar processes such as oxidation, reduction, alkylation, hydrolysis, acylation, hydroxylation, elimination etc.

• Different enzymes carrying out related reactions often employ common co-factors: small organic functional fragments and/or metal ions. e.g.

  - **FREE ENERGY RELEASING COUPLE:** Adenosine triphosphate (ATP)

  - **C-C & C-O BOND FORMATION:** Coenzyme A (CoASH); S-adenosyl methionine (SAM); dimethylallylpyrophosphate (DMAPP); biotin

  - **OXIDATION:** NAD(P)+; FAD/FMN; Haem iron oxo species (e.g. P₄₅₀)

  - **REDUCTION:** NAD(P)H; (FADH₂/FMNH₂)

  - **C-N BOND FORMATION:** Pyridoxal
**Free Energy Releasing Couple - ATP**

- **Adenosine triphosphate (ATP)**
  - phosphorylation of an alcohol by adenosine diphosphate (ADP) is highly *exothermic* (i.e. liberates energy):
    
    $$\Delta G^\circ = -31 \text{ kJmol}^{-1}$$

    - The phosphorylated alcohol ($\text{ROP}$) is then activated towards nucleophilic displacement:
      
      $$\text{Nu}^\ominus + \text{ROP} \rightarrow \text{R-Nu} + \text{OP}$$

    - So, overall the *endothermic* process $\text{ROH} + \text{Y}^- \rightarrow \text{RY} + \text{OH}^-$ has been achieved by ‘coupling’ the process to the ‘hydrolysis of ATP’

    - The situation is analogous to the use of tosylate activation to achieve nucleophilic displacement of an alcohol

    - In general, the exothermicity associated with phosphorylation shifts the equilibria of ‘coupled’ process by a *factor of $\sim 10^8$*
Acyl & C-C Bond Formation α to C=O – CoASH

- **Coenzyme A (CoASH)**
  - Coenzyme A acts as an acyl transfer/α-carbon activation reagent by forming reactive acyl thioesters:
    - Acyl CoA derivatives can act as nucleophiles or electrophiles depending on the circumstances
    - These modes of reactivity reflect inherent properties of alkyl thioesters:
      - The *good leaving group ability* of RS⁻ (cf. RO⁻) reflects: $pK_a (RSH) \approx 10$ cf. $pK_a (ROH) \approx 16$
      - The *high electrophilic character of a thioester carbonyl carbon* (cf. normal esters) reflects the poor orbital overlap between the lone pairs on sulfur ($n_S$) [cf. $n_O$] and the carbonyl anti bonding molecular orbital $\pi^*_{\text{C}=\text{O}}$
      - The *enhanced acidity of protons α to the carbonyl of thioesters* (cf. normal esters) reflects the same poor $n_S \leftrightarrow \pi^*_{\text{C}=\text{O}}$ resonance:

  - Sulfur is in the 2nd period so its lone pair has poor size/energy match with the $\pi^*_{\text{C}=\text{O}}$ orbital
  - Hence: $pK_a (RCH_2COSR') \approx 20$ cf. $RCH_2COOR' \approx 25$
  - *i.e. α to a thioester is similar to α to a ketone*
Methylation/Dimethylallylation – SAM & DMAPP

- **S-Adenosyl methionine (SAM)**
  - SAM acts as a versatile O-, C-, N- & S- methylating reagent *in vivo*
  - Equivalent to performing an $S_N2$ methylation using MeI in the laboratory

- **Dimethylallyl pyrophosphate (DMAPP)**
  - DMAPP acts a dimethylallylating reagent – the pyrophosphate (+ Mg$^{2+}$/Mn$^{2+}$) is an excellent leaving group
  - Equivalent to performing an $S_N2$ allylation using allyl bromide in the laboratory
Carboxylation – *Biotin*

- **Biotin**
  - Biotin in the presence of bicarbonate, ATP and Mg$^{2+}$ enables nucleophile carboxylation *in vivo*:
Oxidation – \( \text{NAD}^+ \)

- **Nicotinamide-adenine dinucleotide (NAD\(^+\))** [and its phosphorylated analogue (NADP\(^+\))] are mediators of **biological oxidation** (e.g. alcohol to ketone oxidation)
  - In general, the couple NAD\(^+\)/NADH is used by enzymes in **catabolic oxidation** (degradation)
  - The reagent is a *stereospecific* **hydride acceptor**:

  - Different enzymes show different absolute specificities but are generally specific for the *pro-R* or *pro-S* hydrogens both for removal and delivery
Oxidation – Flavins (FAD & FMN)

- **Flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN)** are also mediators of biological oxidations (e.g. dehydrogenations – alkane to alkene)
  - Unlike NAD\(^+\), which readily diffuses from enzyme to enzyme, FAD/FMN is usually tightly bound to a given enzyme, sometimes covalently

- Re-oxidation of the FADH\(_2\) back to FAD is generally by molecular oxygen. The intermediate peroxylavin can also mediate hydroxylation, epoxidation & other oxygen transfer reactions (see next slide):
Oxidation Reactions Mediated by Flavins

- **Dehydrogenation by flavins** – *e.g.* dehydrogenation of succinate → fumarate:

  \[
  \text{succinate} \quad \xrightarrow{\text{dehydrogenation}} \quad \text{fumarate}
  \]

  - *ant* conformation imposed by enzyme via arginine salt bridges

- **Baeyer-Villiger-type oxidation by peroxyflavins** – *e.g.* ketone monooxygenase:

  \[
  \text{peroxyflavin} \quad \xrightarrow{\text{Baeyer-Villiger oxidation}} \quad \text{product}
  \]

  - Reaction catalyzed by peroxyflavin with subsequent reduction by NADPH
Oxidation – *Haem Iron oxo Species* \((P_{450})\)

- *Haem iron oxo species* e.g. in cytochrome \(P_{450}\) (a ubiquitous *heam monooxygenase*) are also mediators of *biological oxidation* (e.g. phenolic coupling, epoxidation, *hydroxylation*):

  - The porphyrin ring acts as a tetradeutate ligand for the octahedral iron. The two axial positions are occupied by an enzyme amino acid ligand (typically a histidine nitrogen) and hydroxy/hydroperoxy residue respectively.
Reduction - NADPH

- **Dihydro-nicotinamide-adenine dinucleotide phosphate (NADPH)** [and its de-phosphorylated analogue (NADH)] are mediators of **biological reduction** (e.g. ketone to alcohol reduction)
  - In general, the couple NAPH/NADP$^+$ is used by enzymes in **anabolic reduction** (biosynthesis)
  - The reagent is a stereospecific **hydride donor**:

  - As for the reverse process, different enzymes show different absolute specificities but are generally specific for the **pro-R** or **pro-S** hydrogens both for removal and delivery
Transamination - PLP

- **Pyridoxine (vitamin B₆) → pyridoxal-5’-phosphate (PLP)**
  - PLP forms **amines** (Schiffs bases) with **primary amines**. This forms the basis of **in vivo transamination** of **α-ketoacids** to give **α-amino acids** (& also **racemisation/decarboxylation** processes, see ‘alkaloids’)

- The α-carbon protonation is stereospecific and generally gives the (S) configured chiral centre
Primary Metabolism - Overview

**Primary metabolism**

1) *light reactions*: hv $\rightarrow$ ATP and NADPH
2) *dark reactions*: CO$_2$ $\rightarrow$ 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

- (S)-tryptophan ($ArC_2$)
- (S)-phenylalanine ($ArC_2$)
- (S)-tyrosine ($ArC_3$)
- Menaquinone (vitamin K$_2$) ($ArC_1$)
- $\alpha$-tocopherol (vitamin E) ($ArC_1$)
- Epigallocatechin (EGC) ($ArC_2$)
- Scopoletin ($ArC_3$)
- Podophyllotoxin ($ArC_3$)
**The Shikimate Biosynthetic Pathway - Overview**

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

  ![Chemical diagram of the shikimate biosynthetic pathway](image)

  - 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)
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)

![Diagram showing the metabolic pathways and chemical structures](image)
Primary Metabolism - Overview

**Primary metabolism**

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

**PHOTOSYNTHESIS**

1) *light reactions*: hv → ATP and NADPH
2) *dark reactions*: CO₂ → sugars (Calvin cycle)

**glycolysis**

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

**phosphoenol pyruvate**

**erythrose-4-phosphate**

**shikimate**

**aromatic amino acids**

**peptides**

**proteins**

**tetrapyroles (porphyrins)**

**saturated fatty acids**

**unsaturated fatty acids**

**lips**

**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)