# Discovery of Agrochemicals

**Taught Course**

**Contact: kenneth.ling@syngenta.com**

Classification: PUBLIC

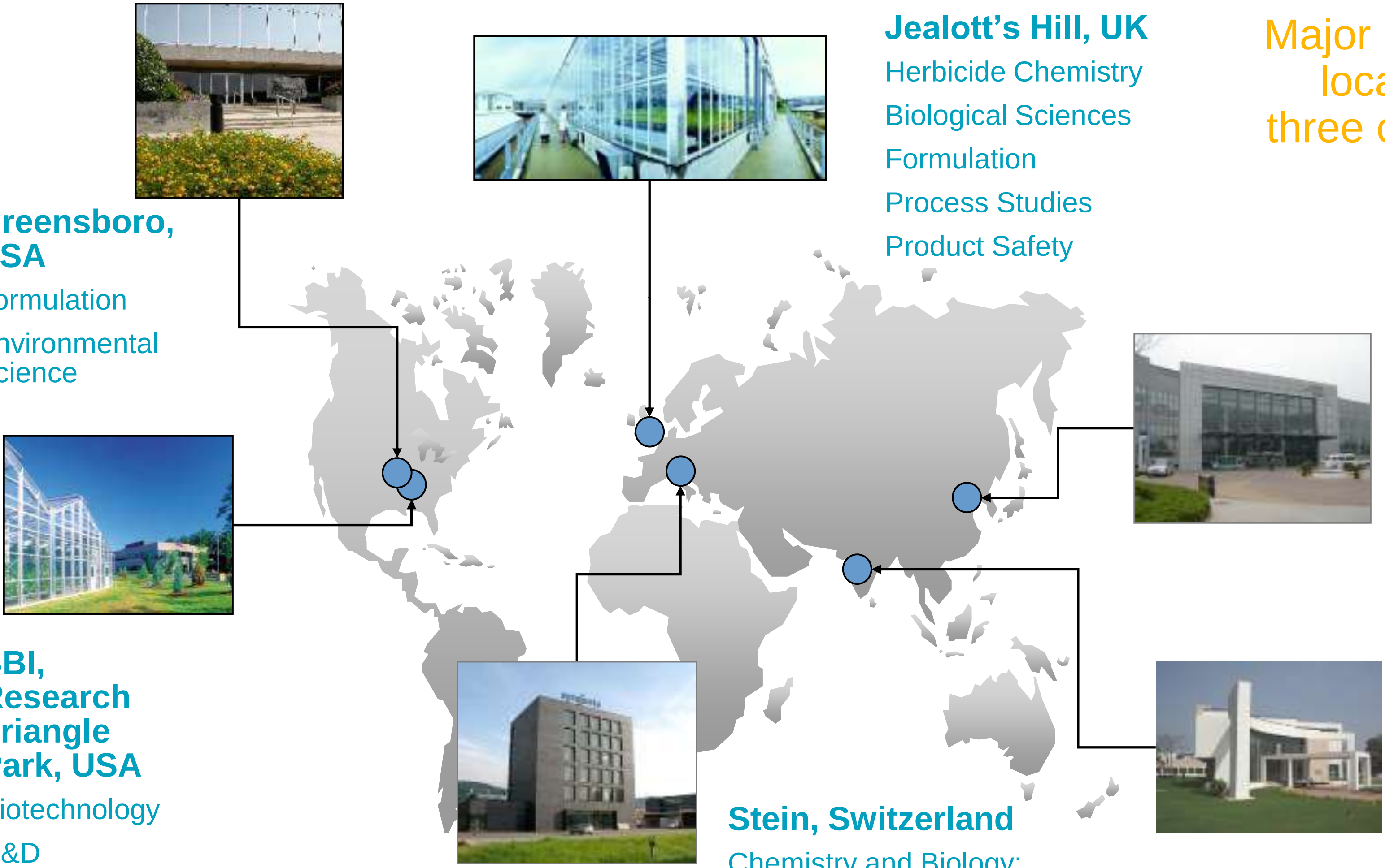
# Syngenta at a glance

$1.4 billion
R&D investment in 2014 and
more than 5000
R&D staff

Over
28000
employees
in some 90
countries

$15.1bn
sales in 2014

# Syngenta R&D locations



Major R&D sites located on three continents

**Greensboro, USA**

Formulation

Environmental Science

**SBI, Research Triangle Park, USA**

Biotechnology

R&D

**Jealott's Hill, UK**

Herbicide Chemistry

Biological Sciences

Formulation

Process Studies

Product Safety

**Stein, Switzerland**

Chemistry and Biology:

Fungicides, Insecticides & Seed Care

**Beijing, China**

Biotechnology

Natural Products Chemistry at various Chinese Universities

**Goa, India**

Research Chemistry

Kilo Lab

Process Research

# Introduction to agrochemicals and modern agronomy

**Taught Course**

Classification: PUBLIC

# Our Fundamental Challenge

By **2050**, global population will rise by about a third to **9 billion** people

but

Global calorie demand will increase by **50%**

**Oxford**

*Population ~150k*

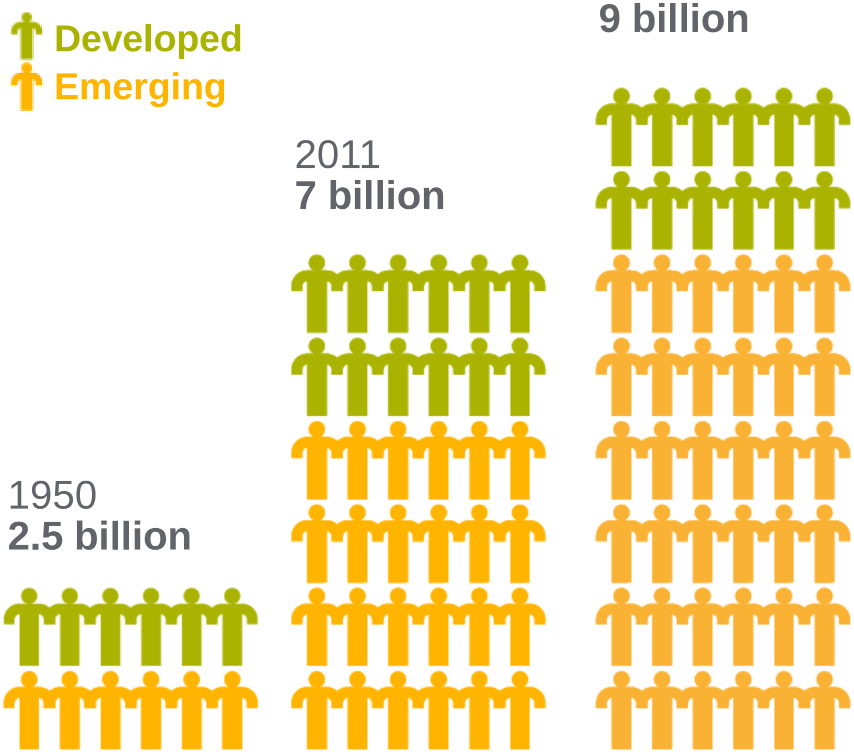
*A city the size of Oxford every 16 hours*

Source: Food and Agriculture Organisation (UN), World Bank statistics, Syngenta

# Demand for food is driven by population growth

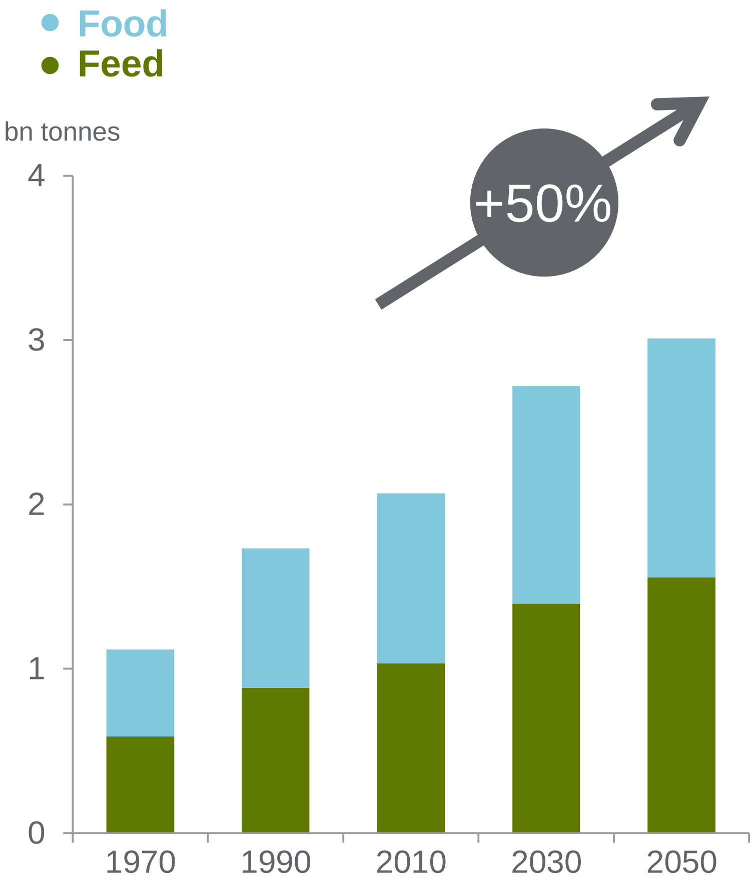
## World population

>80% of growth happens
in emerging markets



Source: FAO, Syngenta analysis

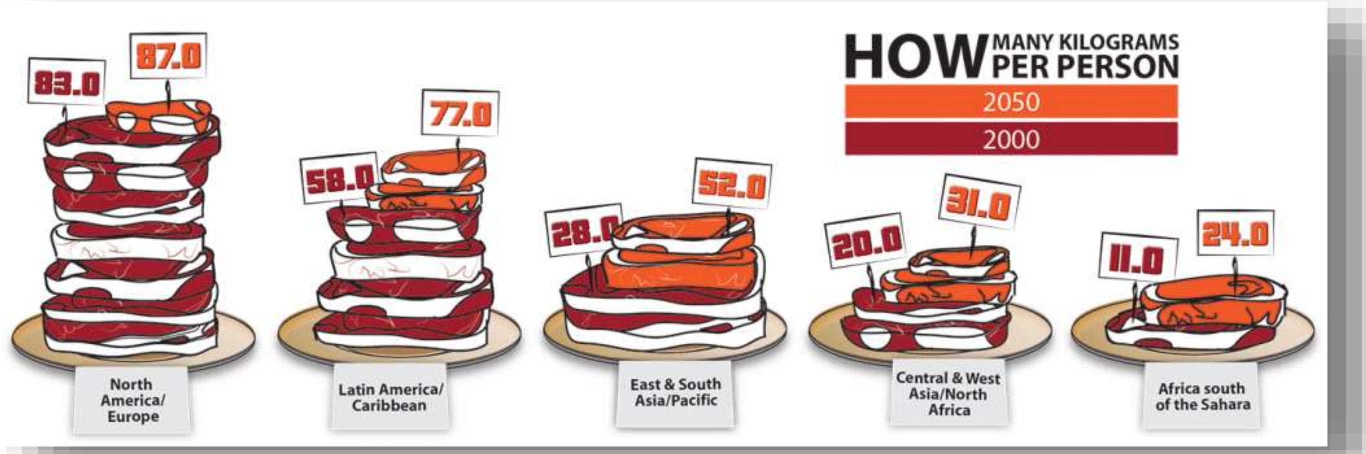
## World demand for grains* (as per FAO)



* Includes cereals, rice, corn, excluding oilseeds

# Changing diets impacts on grain demand

**Increased wealth** in developing countries leads to higher protein content and higher calorie diets



This impacts demand for agricultural feed:

Per kg of meat, production requires this much grain:

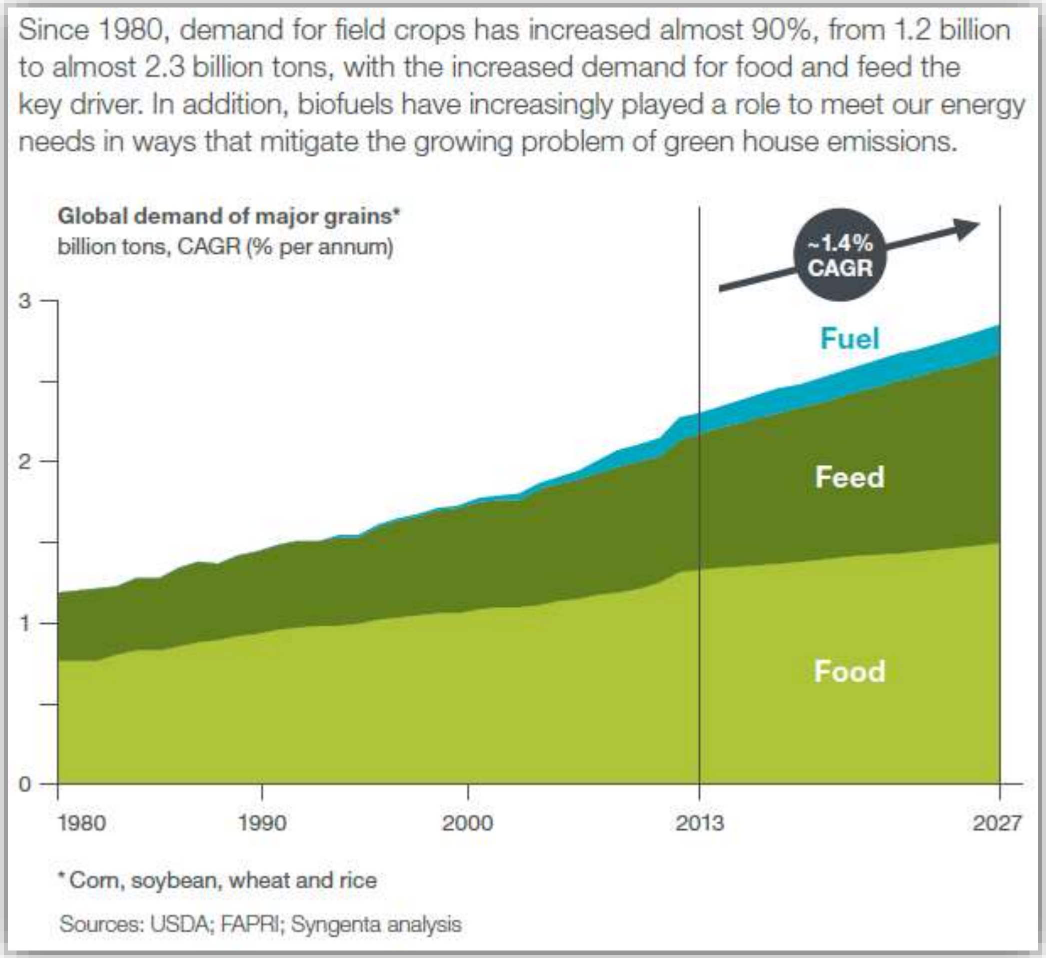
2kg

4kg

7kg

Source: International Food Policy Research Institute and FAO

# Demand



Since 1980, demand for field crops has increased almost 90%, from 1.2 billion to almost 2.3 billion tons, with the increased demand for food and feed the key driver. In addition, biofuels have increasingly played a role to meet our energy needs in ways that mitigate the growing problem of green house emissions.

**Global demand of major grains***
billion tons, CAGR (% per annum)

*Corn, soybean, wheat and rice

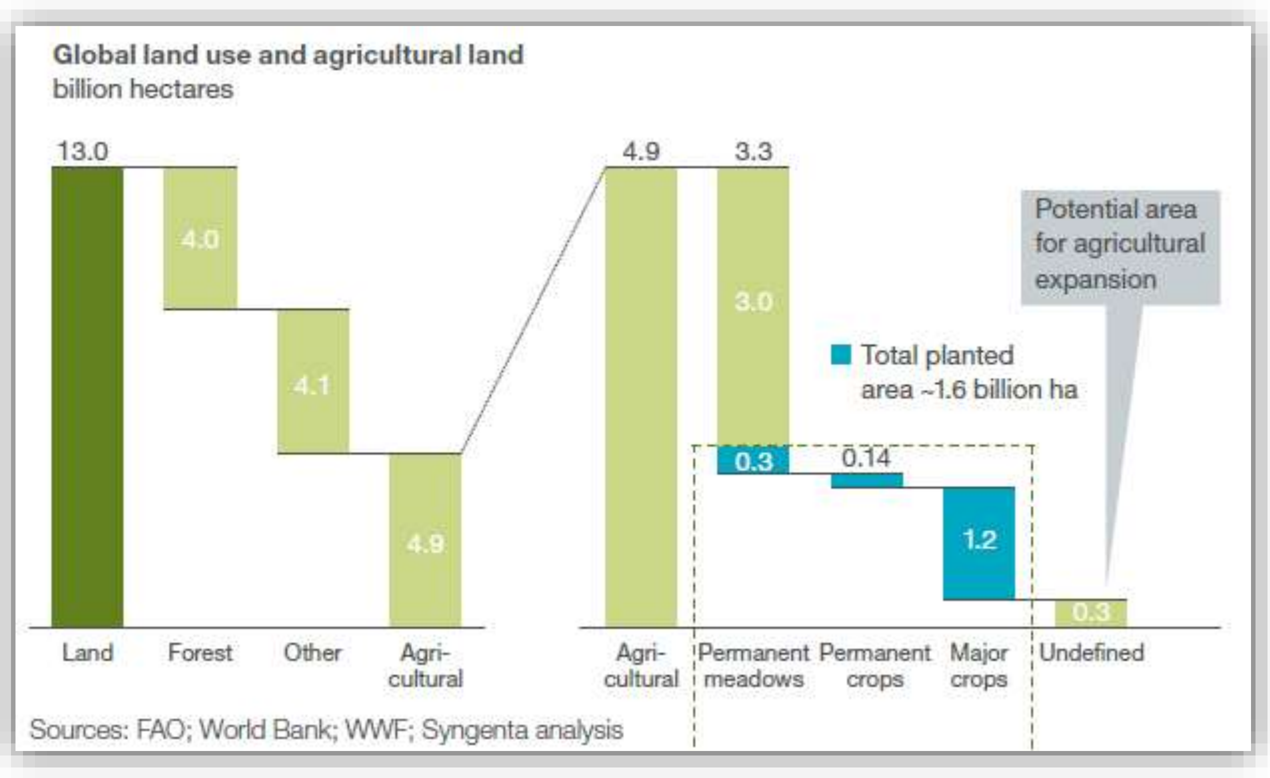
Sources: USDA; FAPRI; Syngenta analysis

# Limited land for agriculture

About a **third** of global land surface is agricultural, but only 12% is for crops

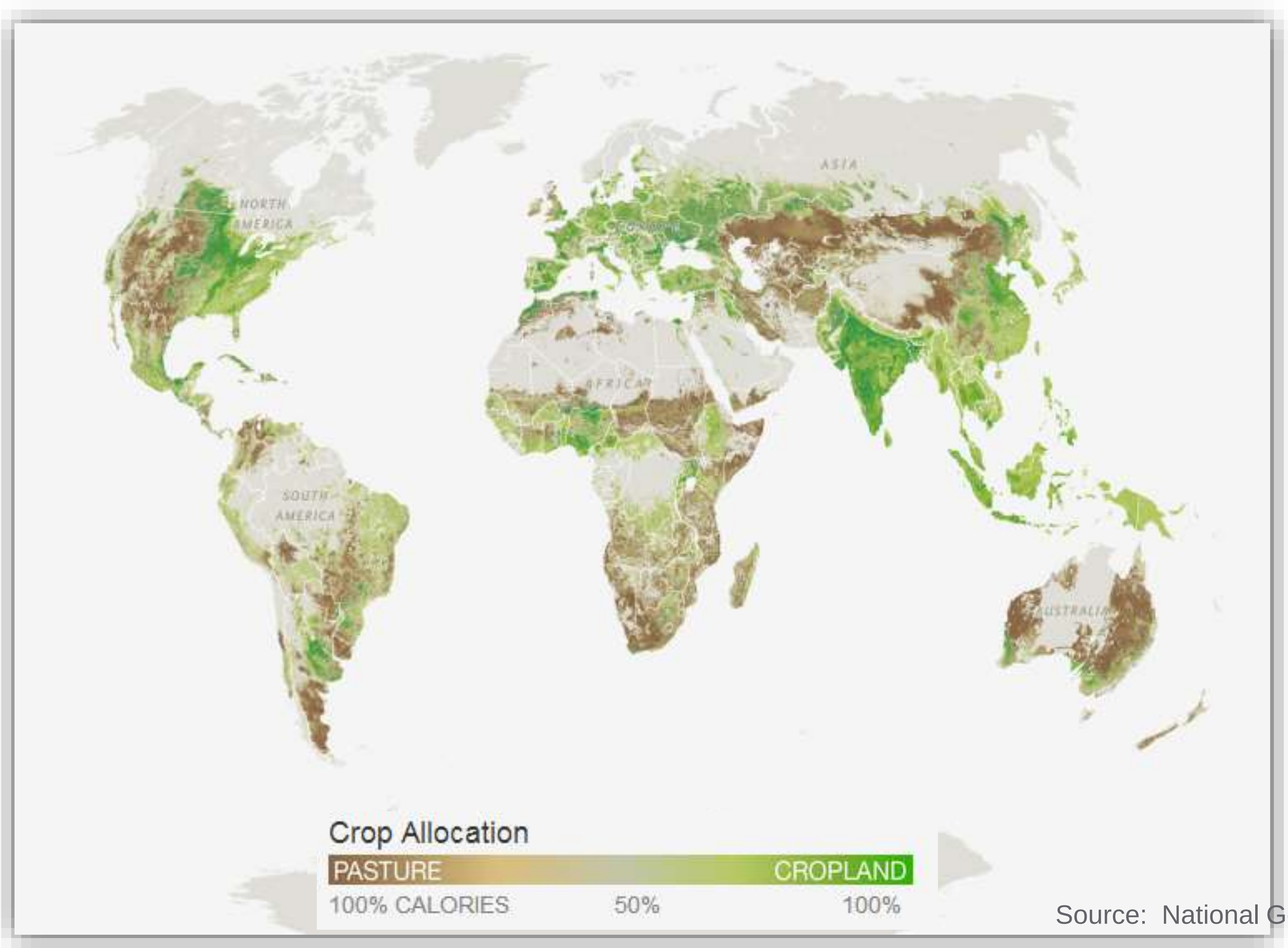
Little **land** for further expansion and **limited resources** (water, soil nutrients)

How do we meet growing demand?



Sources: FAO; World Bank; WWF; Syngenta analysis

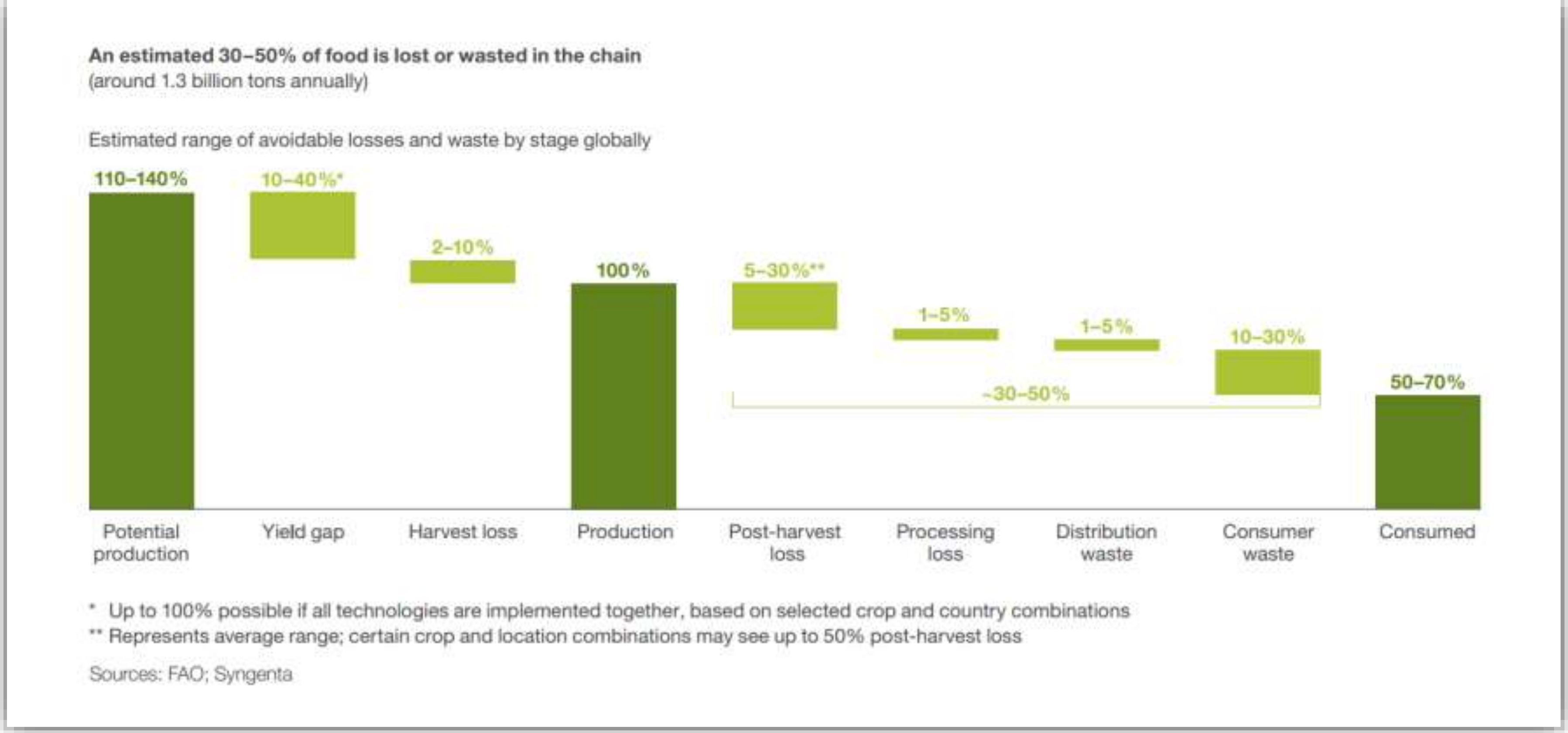
# Geographical distribution of agricultural land



Source: National Geographic

# Improving Productivity

Around **40%** of all food produced in never used



**An estimated 30–50% of food is lost or wasted in the chain**
(around 1.3 billion tons annually)

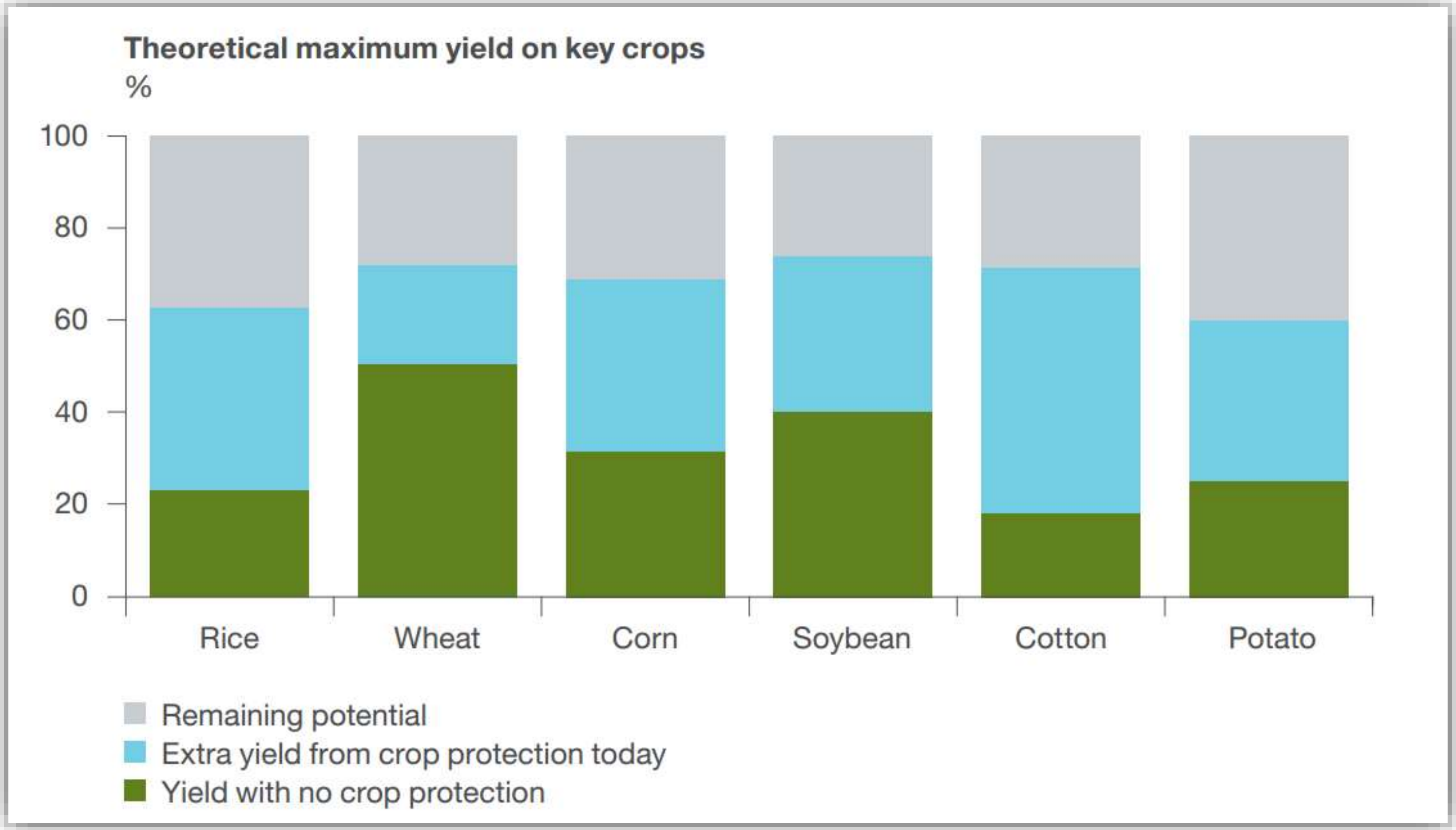
Estimated range of avoidable losses and waste by stage globally

* Up to 100% possible if all technologies are implemented together, based on selected crop and country combinations
** Represents average range; certain crop and location combinations may see up to 50% post-harvest loss

Sources: FAO; Syngenta

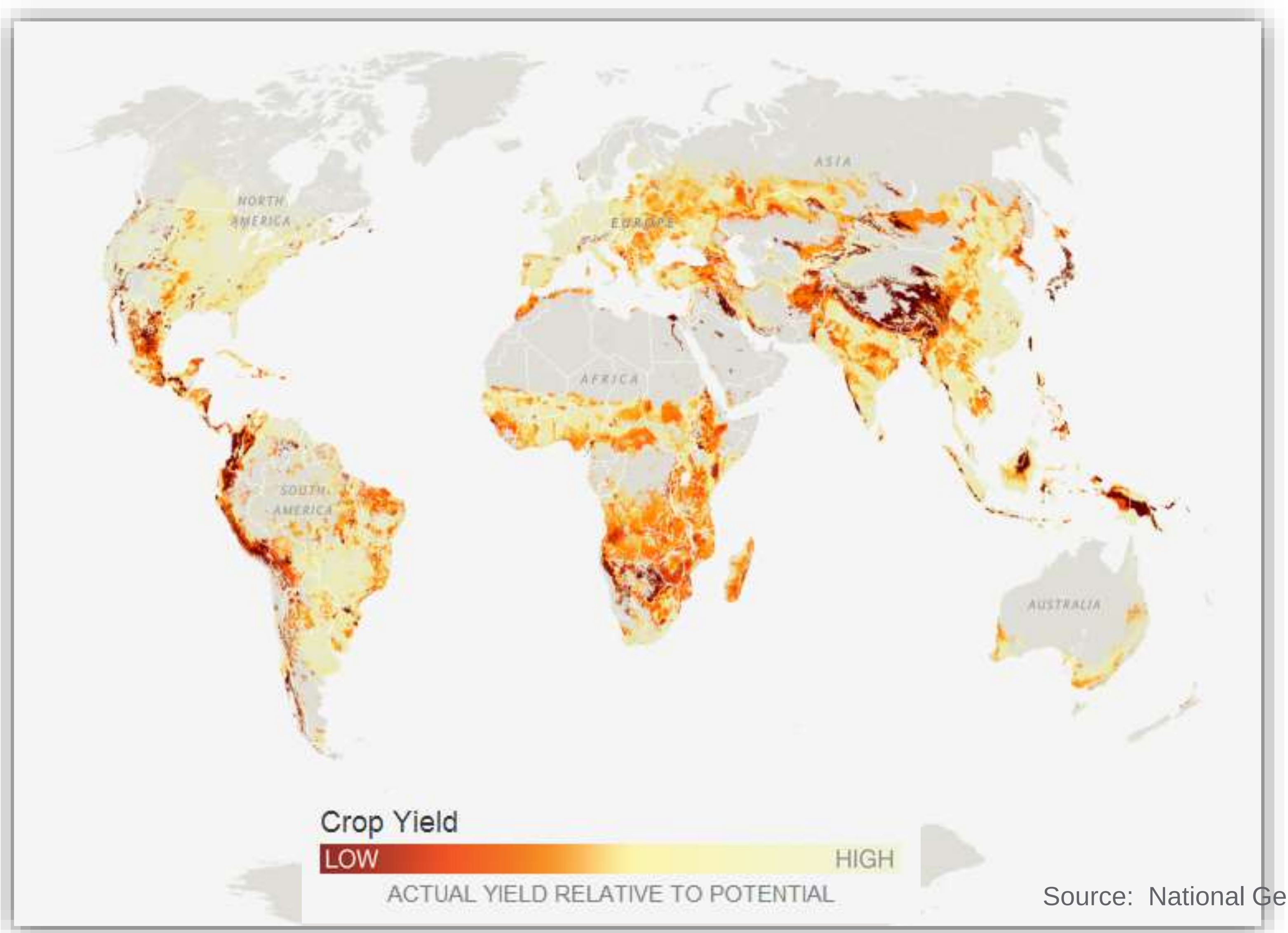
# Improving Productivity

Still significant potential to increase crop yields (but crop protection helps!)



However, there are big differences in yield globally…

# Geographical variation in crop yield



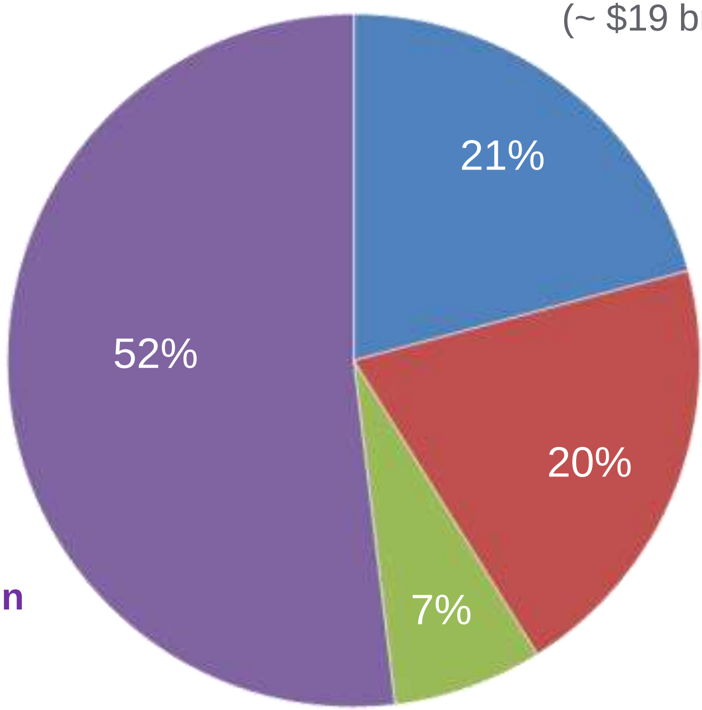
Source: National Geographic

# Technologies for Yield

- There are only 4 major technologies
  - Better Seeds
    - Traditional Breeding
    - Marker Assisted Breeding
    - GM Crops
  - Mechanisation, including irrigation
  - Synthetic fertilisers (NPK)
  - Crop Protection Chemicals
- Deployed as part of an integrated agronomic system

# Global agribusiness 2012

## Global market: ~$91 bn



**Conventional seeds**
(~ $19 bn)

**G.M. seeds**
(~ $18 bn)

**Non-crop uses**
(~ $6.4 bn)

**Crop Protection**
Herbicides
Insecticides
Fungicides
Seed treatment
(~ $47 bn)

Source: Syngenta Analysis, Philips McDougall

# What is an agrochemical?

A chemical which can safely be applied to a crop in order to give the farmer:

- Higher yields
- Better quality produce
- Reliability
- Ease of harvest

Agrochemicals include **insecticides**, **fungicides** and **herbicides** which are collectively known as agrochemicals or crop protection products.

Crops are in competition with weeds, plant diseases, insects and other organisms.

  - About 10,000 insect species are classified as pests
  - At least 600 species of plants are classified as weeds
  - Some 1,500 different fungi cause plant diseases

# History of Agrochemicals

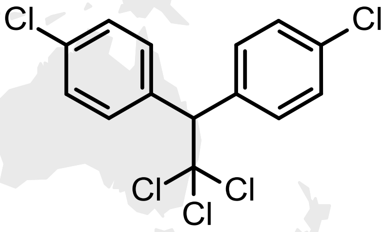
Agriculture begins in the Fertile Crescent of Mesopotamia

$NaClO_3$, $FeSO_4$, $CuSO_4$, $H_2SO_4$, $As_2O_3$, $Pb_3(AsO_4)_2$, $HgCl_2$, PhHgOAc, $Cu(OH)_2/CaSO_4$

Inorganic pest control

~8000BC ~2500BC 0 1800 1900 1940

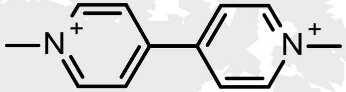
Sumerians use sulphur to control insects and mites

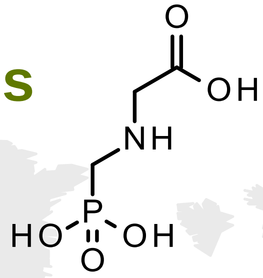


DDT
First synthetic insecticide

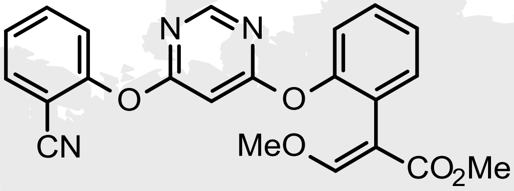
# History of Agrochemicals

**Paraquat**
**(Herbicide)**

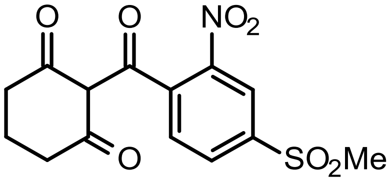




**Glyphosate**
**(Herbicide)**

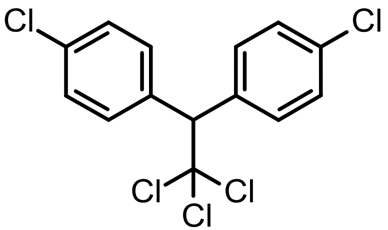


**Azoxystrobin**
**(Fungicide)**

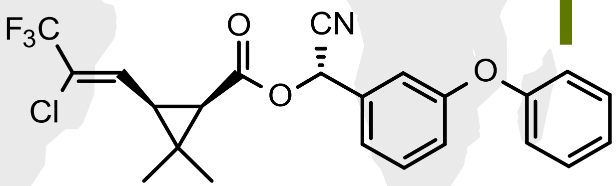


**Mesotrione**
**(Herbicide)**

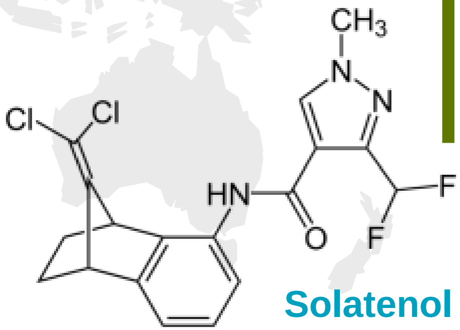
**1940** **1960** **1970** **1980** **1990** **2000**



**DDT (Insecticide)**



**Lambda-cyhalothrin**
**(Insecticide)**



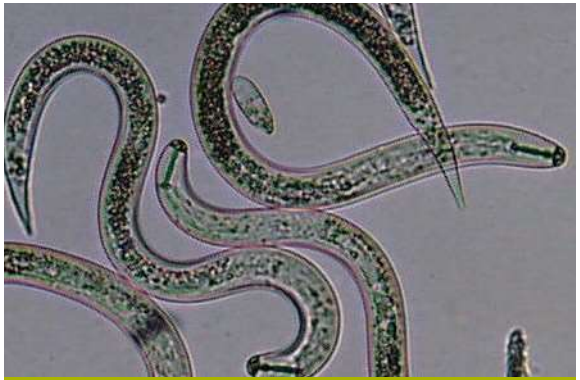
**Solatenol**
**(Fungicide)**

# Crop Protection: 6 key areas of grower need

Weed control

Fungal control

Insect control



Other biotic stress

Abiotic stress

Yield and quality

# Crop Protection: Herbicides

- 'Weed' is a term which describes any undesired vegetation and it consequently covers a very large spectrum of plants! Weeds are generally categorised as grasses or broad leaf weeds

- The presence of weeds in a field plot can dramatically reduce yield

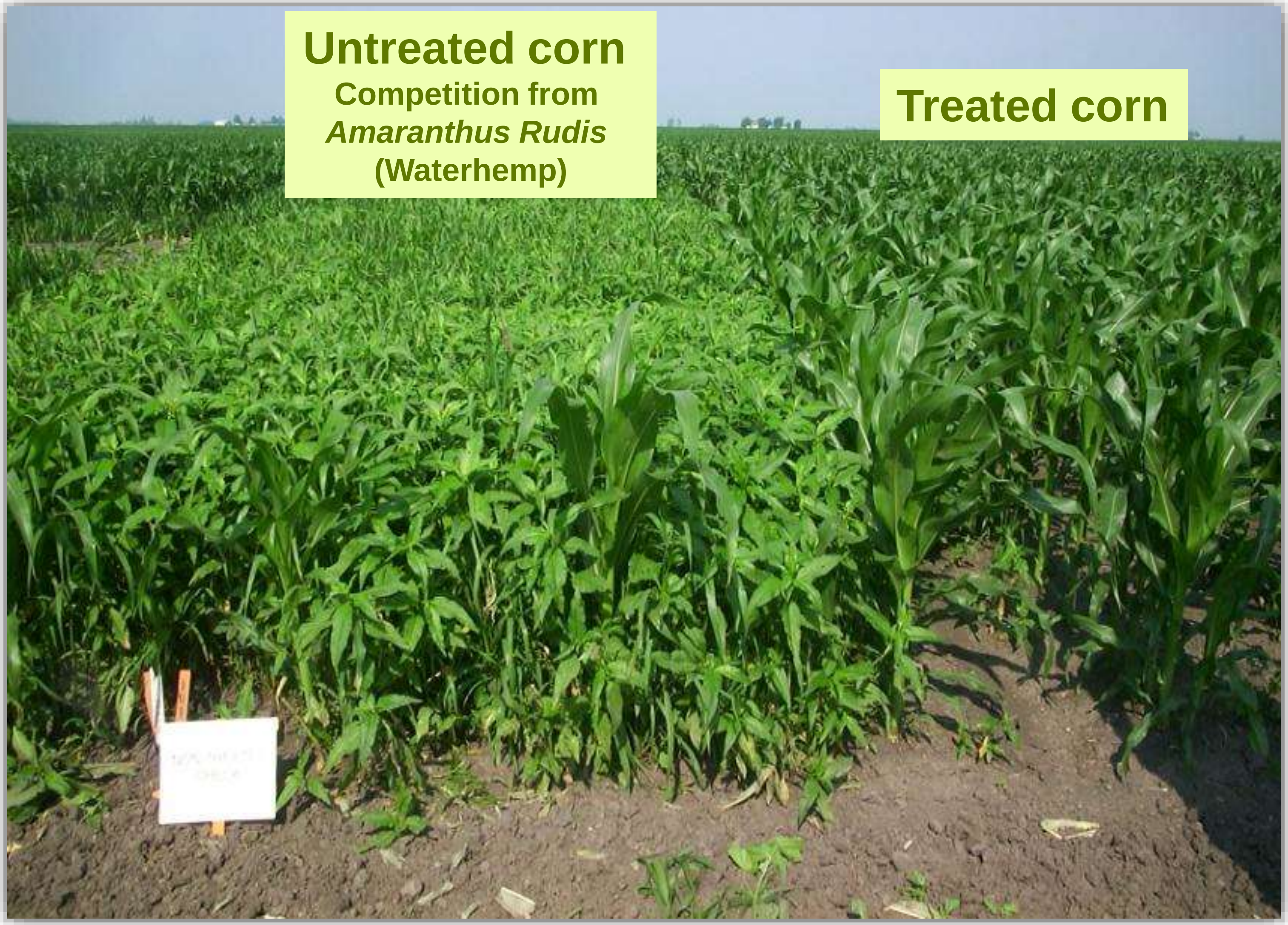
- **Herbicides** control weeds that compete with crops for light and nutrients

- Herbicides can also prevent soil erosion and water loss by replacing or reducing the need for cultivation

**Grass weed example:**
***Echinochloa crus-galli***
**"Barnyard grass"**

**Broad leaf weed example:**
***Abutilon theo.***
**"Velvetleaf"**

# **Herbicides:** Protect Crops from Competition by Weeds



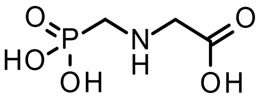
# Crop Protection: Herbicides

- Herbicides can be:
  - **Selective:** Meaning they will only affect a certain type of plant and not another. This allows the spraying of a crop, leaving it unaffected, whilst controlling weed growth
  - **Non-selective:** Meaning that all vegetation is controlled within the sprayed area

- Herbicides also have non-crop uses, for example non-selective herbicides are used to keep train tracks clear and selective herbicides are used in gardens

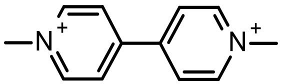
# Herbicides

## Non-selective



**Glyphosate**

Touchdown®

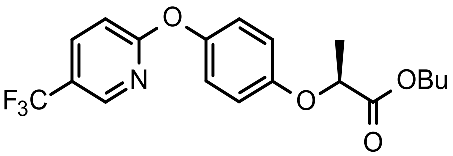


**Paraquat**

Gramoxone®

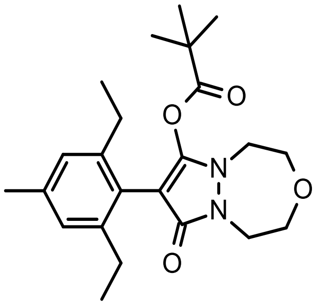
**Treated with Mesotrione (Callisto)**

## Selective



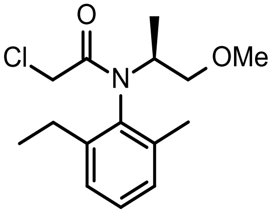
**Fluazifop-P butyl**

Fusilade®



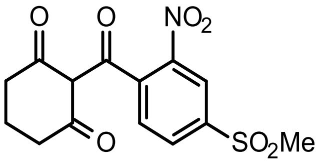
**Pinoxaden**

Axial®



**S-Metolachlor**

DualGold®



**Mesotrione**

Callisto®

# Crop Protection: Fungicides

- **Fungicides** play a key role in keeping a crop healthy from fungal disease which can have severe adverse effects on crop yield and quality

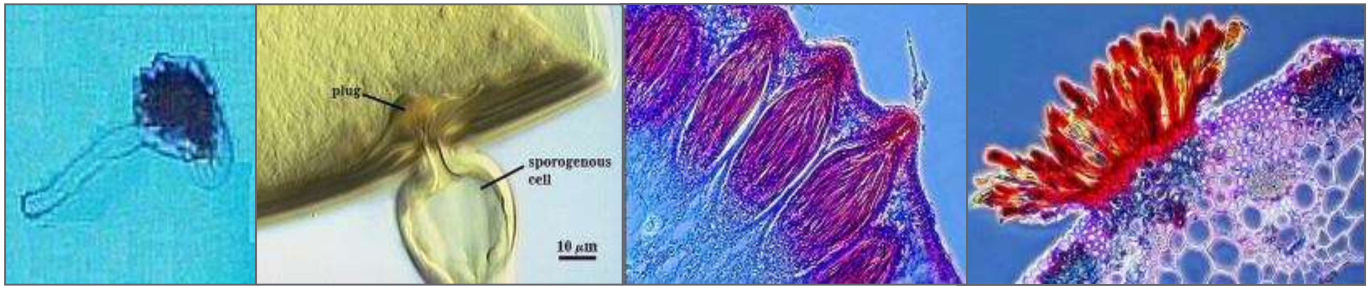
- During the Irish potato famine in the 1840s, 1 million people died and another 1 million people left Ireland. The famine was caused by *Phytophthora infestans* (late potato blight)

- Some fungal species (e.g. *Fusarium graminearium*) produce mycotoxins which are known carcinogens. Poisoning caused by ergot alkaloids from fungi is also potentially serious

- Fungicides are described as **broad spectrum** (e.g. effective on a wide range on fungi across the taxonomical groups) or **specific** (e.g. mildew-specific or oomycete-specific fungicides)

**Potato late blight**

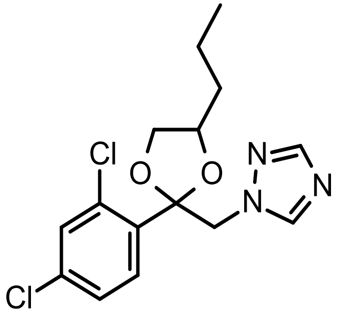
# Crop Protection: Fungicides



| SPORE GERMINATION | INFECTION | VISIBLE SYMPTOMS | SPORULATION |
|---|---|---|---|
| **Preventative** | **Curative** | **Eradicant** | **Antisporulant** |

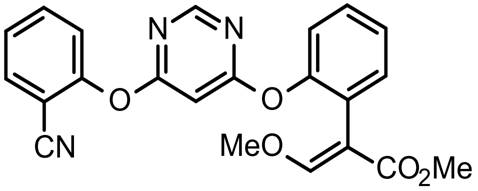
- A fungicide can be:
    - **Preventative**: Prevents the establishment of infection
    - **Curative**: Inhibits the development of an established infection which is **not** showing visible symptoms of disease
    - An **eradicant**: Inhibits the development of an established infection which **is** showing visible symptoms
    - An **antisporulant**: Prevents or reduces sporulation without necessarily stopping vegetative growth

# Fungicides



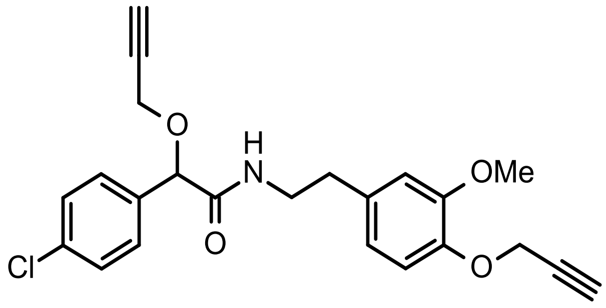
**Propiconazole**

Tilt®



**Azoxystrobin**

Amistar®



**Mandipropamid**

Revus®

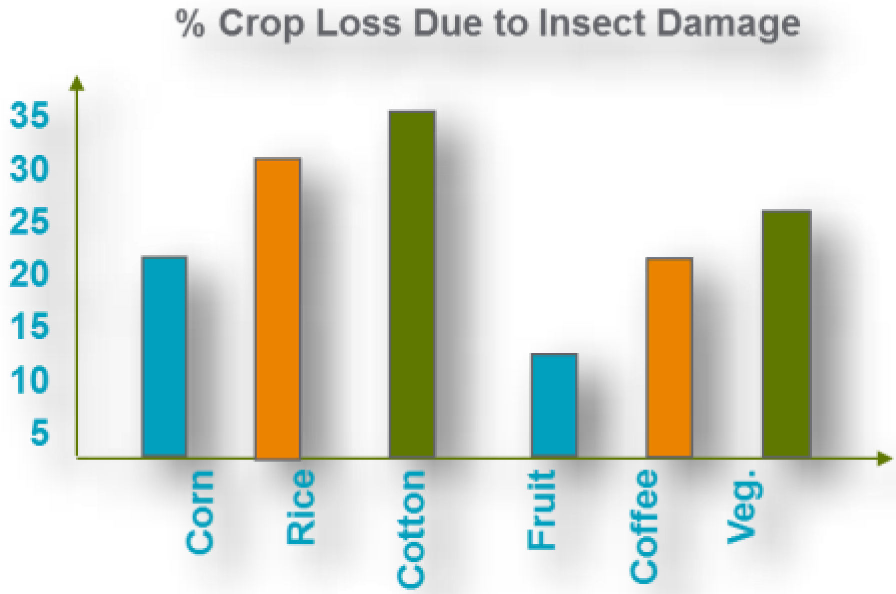
Root and crown rot *Rhizoctonia solani*

**Untreated**

**Treated with Azoxystrobin (Amistar)**

# Crop Protection: Insecticides

- **Insecticides** protect crops before and after harvest from potentially devastating pests that threaten yield and quality.
- Some insects are **chewing pests** (e.g. caterpillars & beetles) which feed on plant material itself (leaves, fruits, roots, etc.). This leads to a reduction of the photosynthetic area and yield. The damage caused can seriously weaken the plant. They can also cause serious loss during material storage.
- Others are **sucking pests** which feed on sap (e.g. Aphids). This method of feeding can lead to serious viral transmission that can destroy a crop.



% Crop Loss Due to Insect Damage

# Crop Protection: Insecticides

**Lygus bugs**

**Aphids, hoppers & whitefly**

**Tobacco budworm**

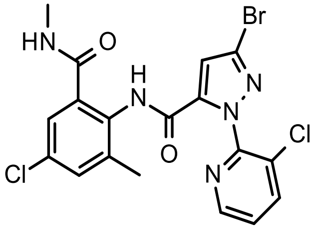
**Beetles**

- Insecticides control insect pests whilst minimising impact on beneficial insects such as:

**Bees**

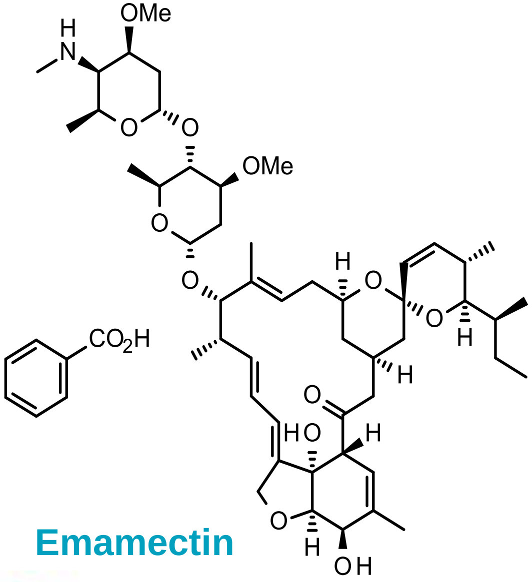
**Ladybirds**

# Insecticides



**Chlorantraniliprole**

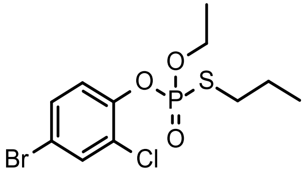
Durivo®



**Emamectin**

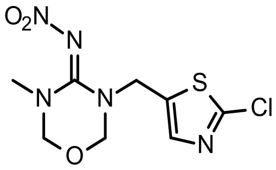
Proclaim®

**Cotton Alabama Larva & leaf damage**



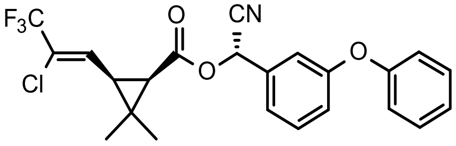
**Profenaphos**

Curacron®



**Thiamethoxam**

Actara®



**Lambda-cyhalothrin**

KarateZeon®

# Non-crops uses of agrochemicals

- Chemicals developed for crop protection also find uses in a variety of other areas:

## Lawn & Garden

**Turf**

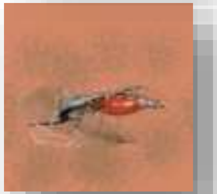
**Ornamentals**

**Garden**

## Home Care & Public Health

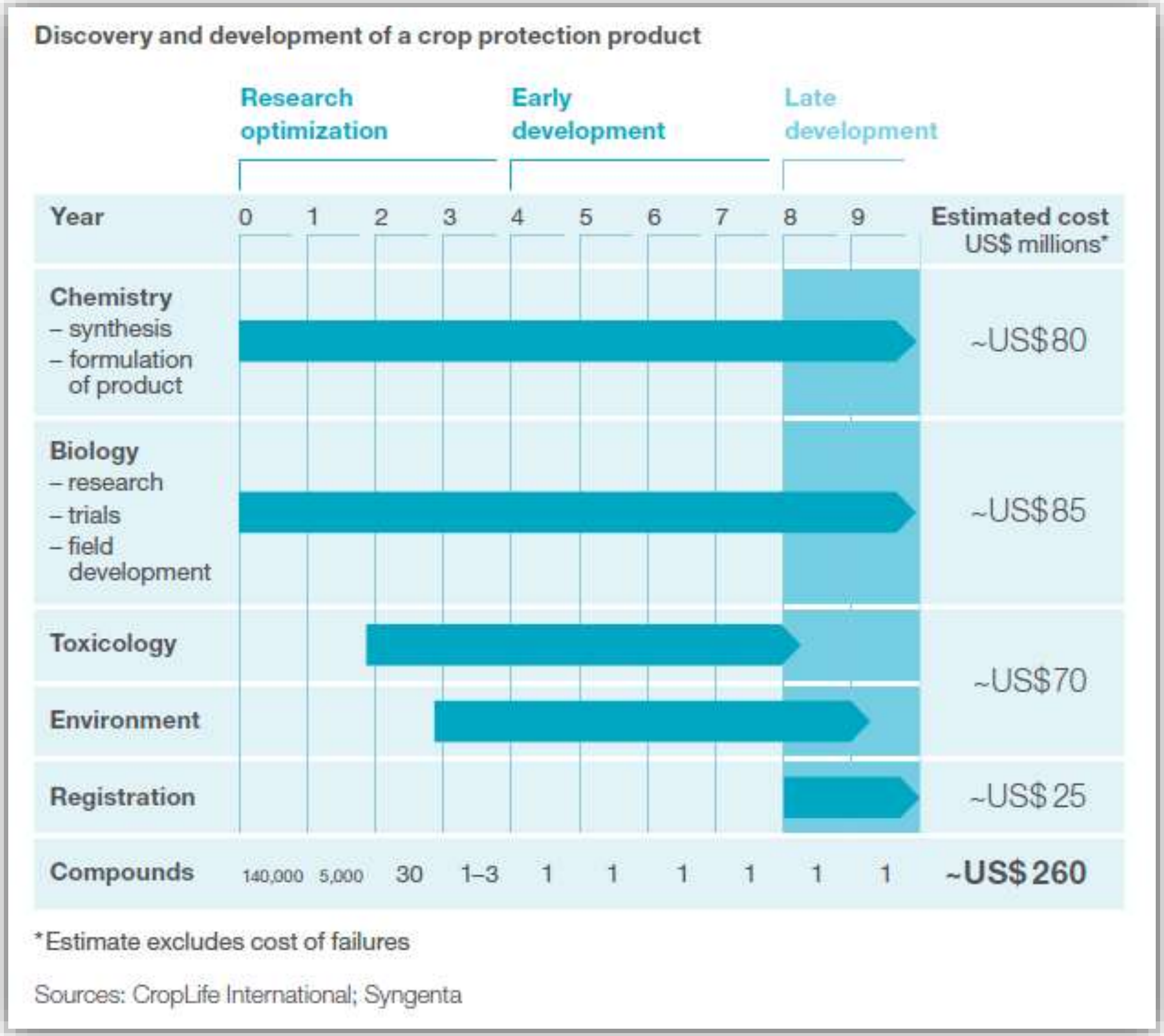
**Materials Protection**

**Household Pest Management**



**Vector Control**

# The Route to Market



**Discovery and development of a crop protection product**

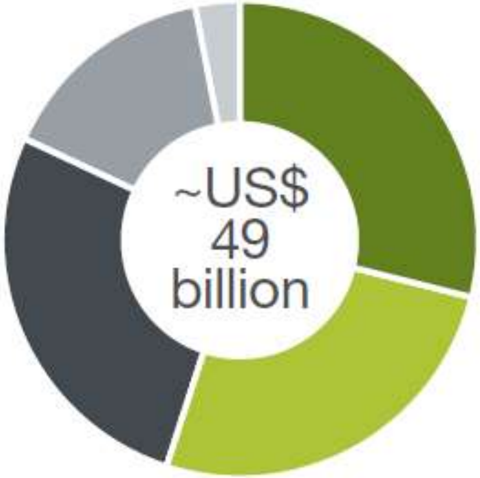
| Year | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Estimated cost US$ millions* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Chemistry**<br>– synthesis<br>– formulation of product | | | | | | | | | | | ~US$80 |
| **Biology**<br>– research<br>– trials<br>– field development | | | | | | | | | | | ~US$85 |
| **Toxicology** | | | | | | | | | | | ~US$70 |
| **Environment** | | | | | | | | | | | |
| **Registration** | | | | | | | | | | | ~US$ 25 |
| **Compounds** | 140,000 | 5,000 | 30 | 1–3 | 1 | 1 | 1 | 1 | 1 | 1 | **~US$ 260** |

*Estimate excludes cost of failures

Sources: CropLife International; Syngenta

# The Market

**2012 global crop protection market by product**

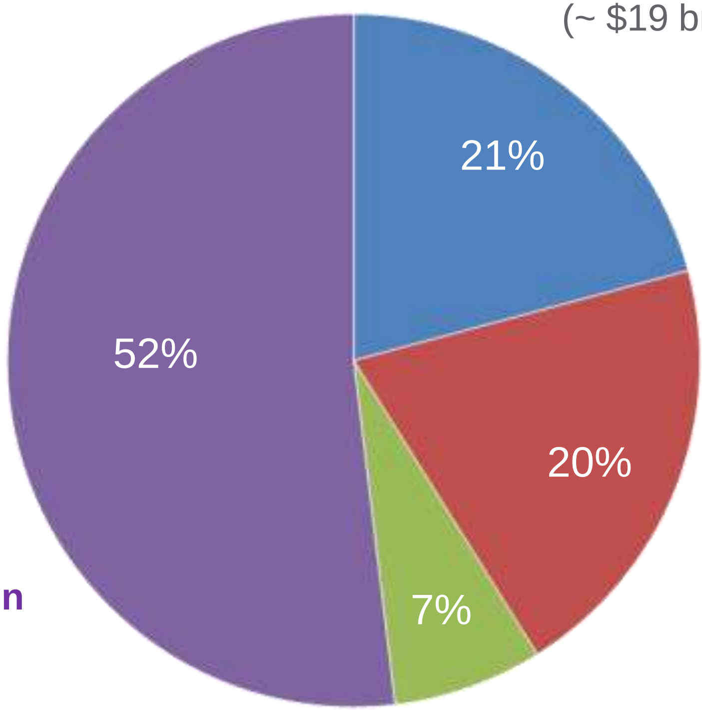


| | | |
|---|---|---|
| ■ | Selective herbicides | 29% |
| ■ | Fungicides | 26% |
| ■ | Insecticides | 27% |
| ■ | Non-selective herbicides | 15% |
| ■ | Seed treatment | 3% |

Source: Syngenta

# Global agribusiness 2012

## Global market: ~$91 bn



**Conventional seeds**
(~ $19 bn)

**G.M. seeds**
(~ $18 bn)

**Non-crop uses**
(~ $6.4 bn)

**Crop Protection**
Herbicides
Insecticides
Fungicides
Seed treatment
(~ $47 bn)

Source: Syngenta Analysis, Philips McDougall

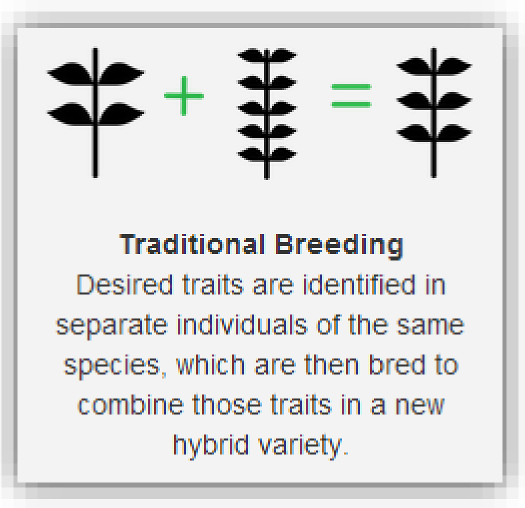
# Seeds: history of plant breeding

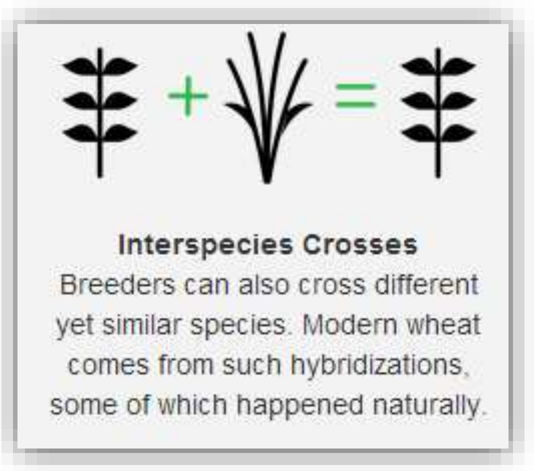
- Since the practice of agriculture began (~10,000yrs ago), farmers have been conducting ‘plant breeding’

- The technology has evolved dramatically through an understanding of Mendelian genetics and modern molecular biology techniques

- Typically, breeders have tried to incorporate the following traits
  - Increased yield and quality of the crop
  - Increased tolerance to environmental stress (e.g. drought)
  - Resistance to viruses, fungi and bacteria
  - Increased tolerance to insects
  - Increased tolerance to herbicides

# Seeds: plant breeding technology

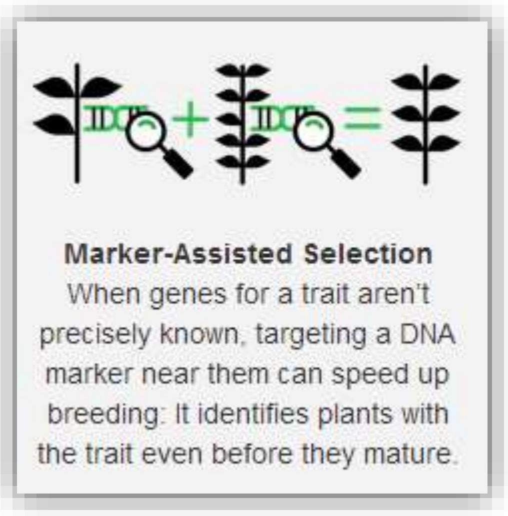

**Traditional Breeding**
Desired traits are identified in separate individuals of the same species, which are then bred to combine those traits in a new hybrid variety.


**Interspecies Crosses**
Breeders can also cross different yet similar species. Modern wheat comes from such hybridizations, some of which happened naturally.

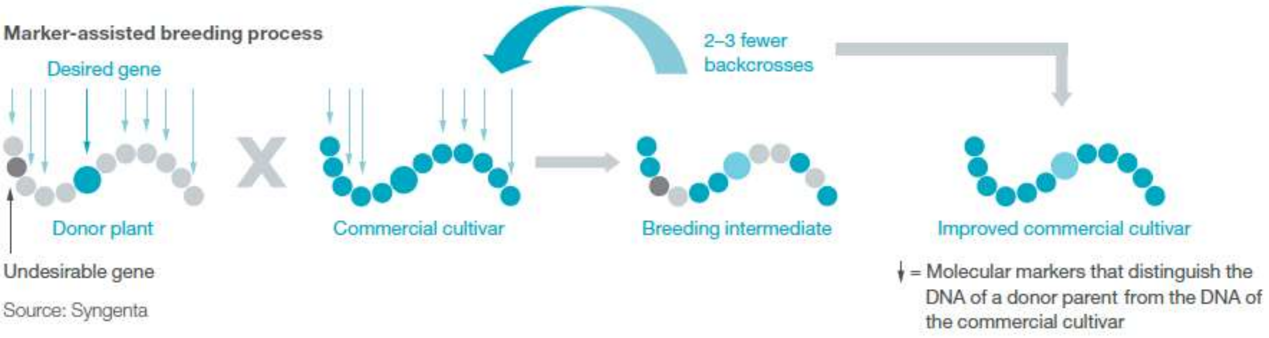
**Loganberry**
(raspberry x blackberry)

**Grapefruit**
(pomelo x Jamaican sweet orange)

# Seeds: plant breeding technology

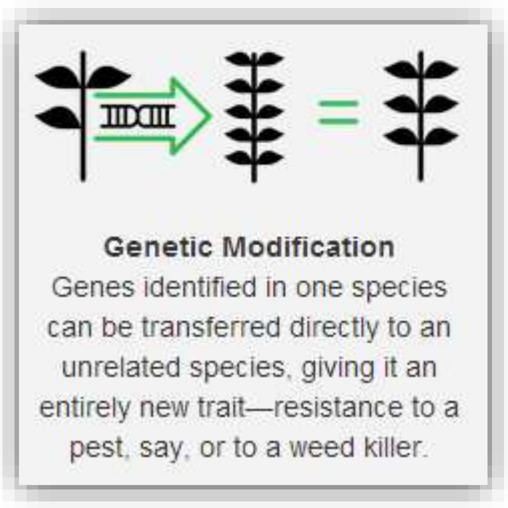


**Marker-Assisted Selection**
When genes for a trait aren't precisely known, targeting a DNA marker near them can speed up breeding: It identifies plants with the trait even before they mature.



Source: Syngenta

↓ = Molecular markers that distinguish the DNA of a donor parent from the DNA of the commercial cultivar



**Genetic Modification**
Genes identified in one species can be transferred directly to an unrelated species, giving it an entirely new trait—resistance to a pest, say, or to a weed killer.

**Golden Rice** – engineered to contain β-carotene, a direct precursor of Vitamin A whose dietary deficiency is responsible for 1-2 million deaths per annum.

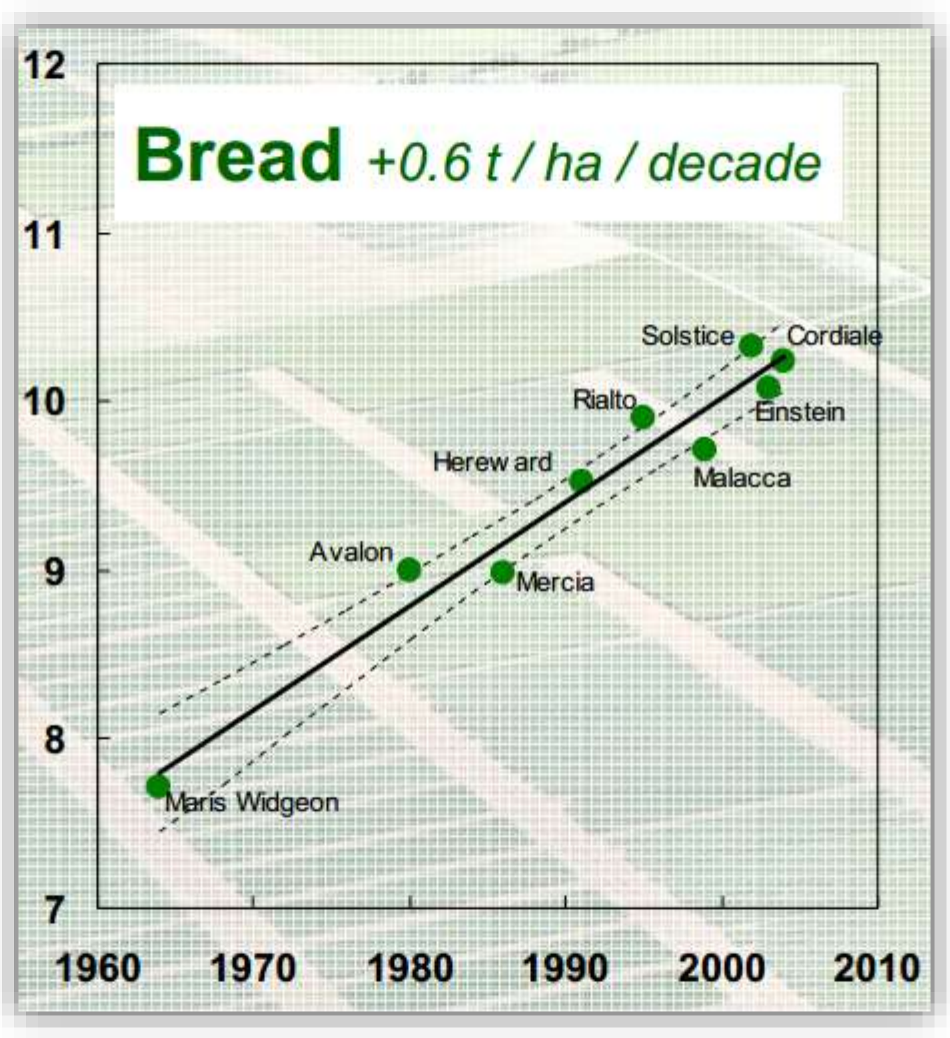
# Wheat: thousands of years in the making

*From left to right:*

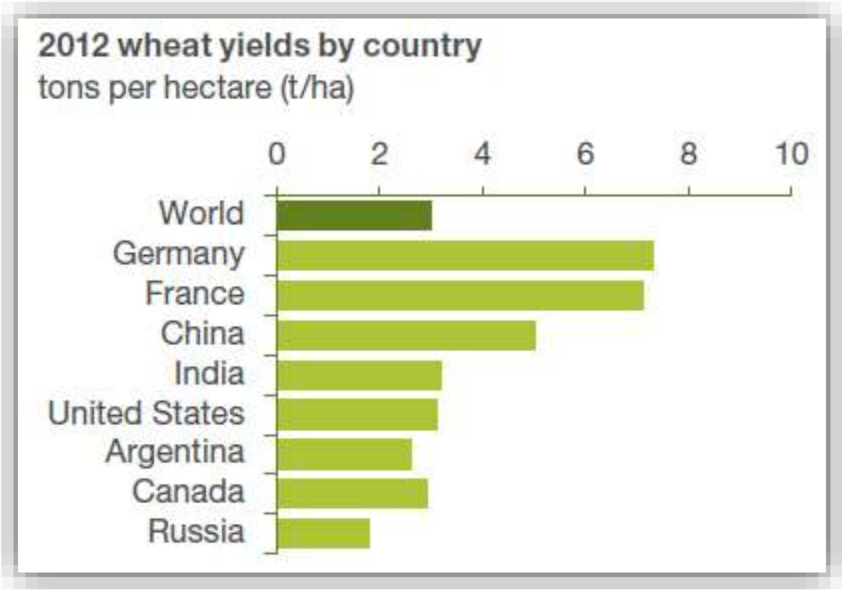
1. **Ancient wild wheat**
2. **Einkorn wheat** – earliest cultivated form (~7500BC)
3. **Durum wheat** – hybrid variety used primarily for pasta (einkhorn x *Aegilops speltoides*)
4. **Modern (common) wheat** – hybrid variety used primarily for bread (durum x *Aegilops tauschii*)
5. New variety with shorter stalks and bigger seeds

Source: National Geographic

# Wheat: breeding increases yields

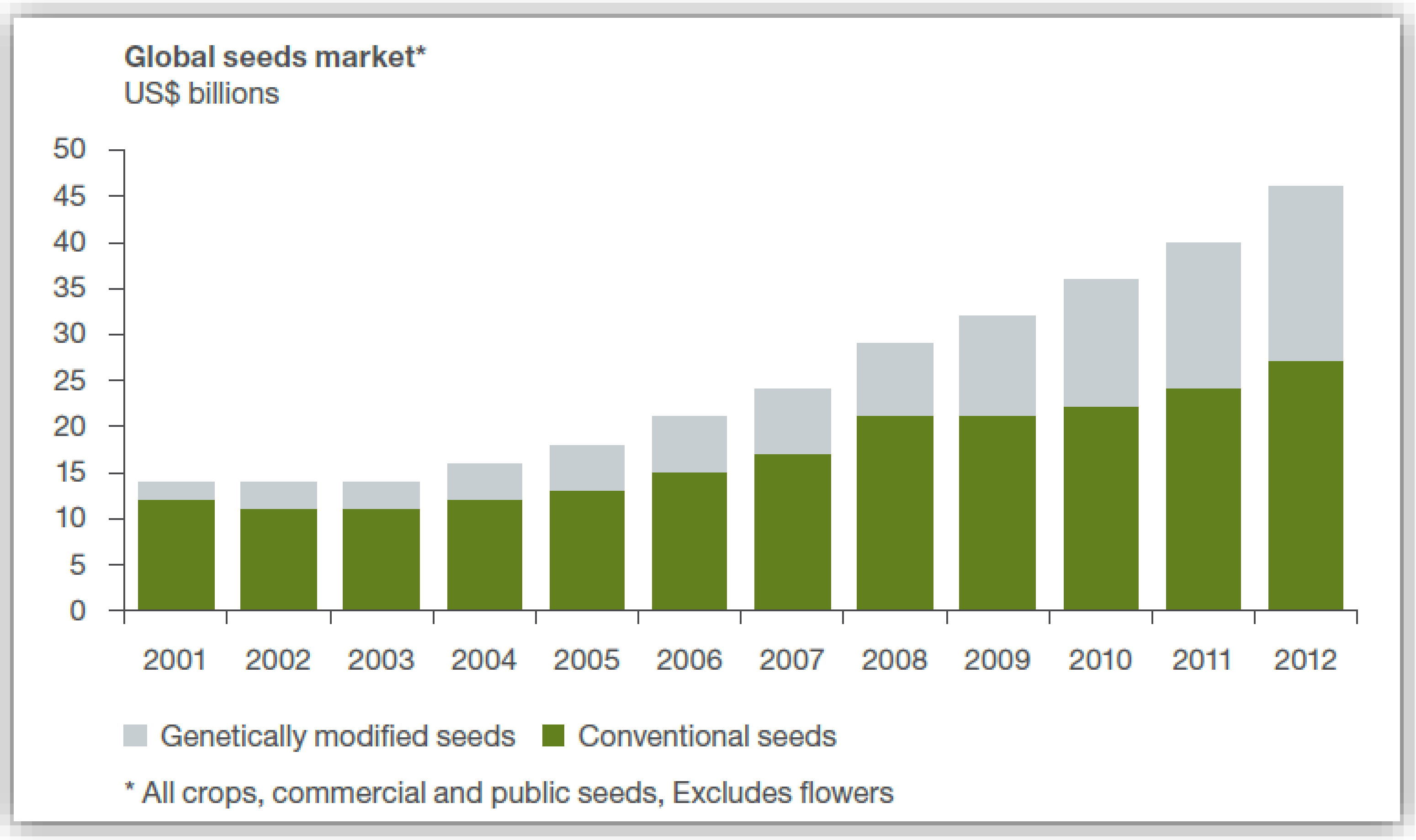


- Yields have increased dramatically in the last few decades
- **Current UK record is 14.5t/ha**
- Global average (2012) is about 3t/ha



Sources: Yield Enhancement Network, COCERAL

# The Seeds Market



**Global seeds market***
US$ billions

* All crops, commercial and public seeds, Excludes flowers

# Modern Agriculture

- On top of day-to-day farming, farmers have to consider a wide range of external factors such as climate, market trends, the political environment, agricultural regulations
- In addition, farmers are responsible for procurement, finance, distribution and marketing of their products

# Modern Agriculture: precision farming

- Farmers are increasingly looking to technology to refine agricultural practices
- Advanced warning sensors, automatic yield mapping, GPS positioning, drones and many more are all becoming accepted tools for farmers

# Lead Generation

**Taught Course**

Classification: PUBLIC

# The Route to Market

Lead Generation

Optimisation

Development

Registration

Commercial Pesticide

Number of compounds

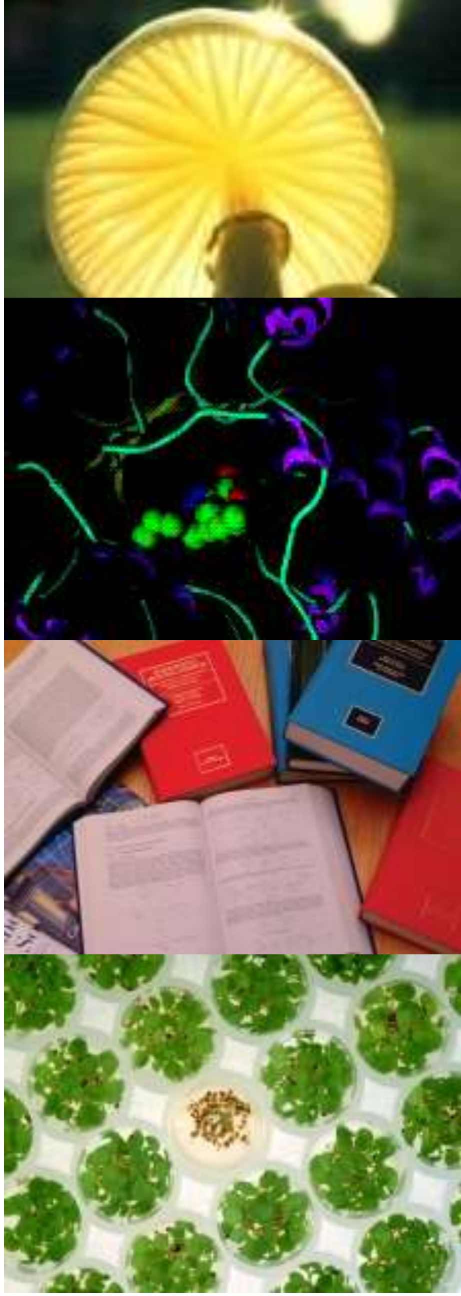
140,000* → 5000 → 30 → 1-3 → 1 → 1 → 1

*average number of compounds synthesised to deliver one new market introduction in 2005

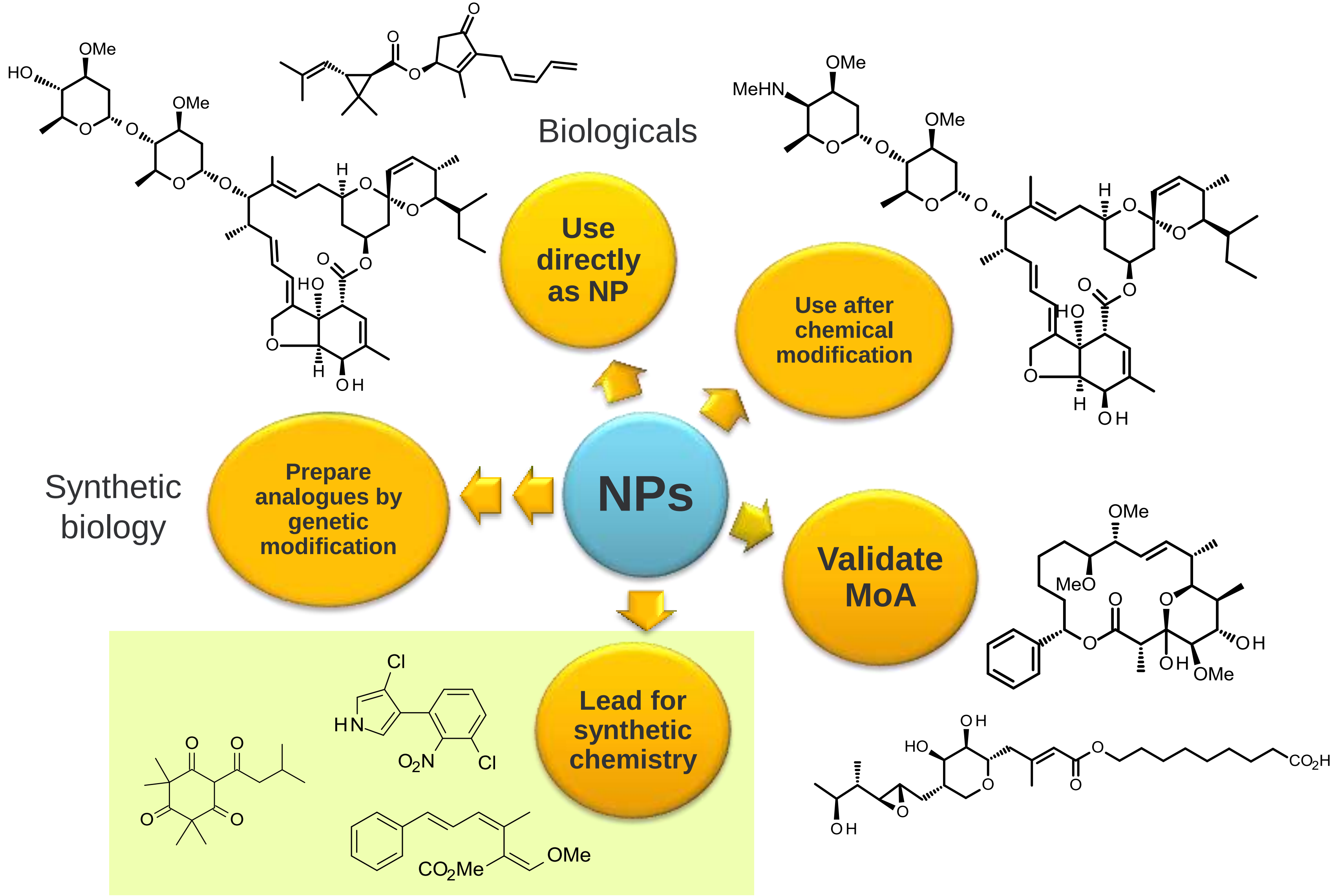
# Lead Generation

## Where do New Leads Come From?

- Natural products
- Mechanism & structure-based design
- Ideas from the literature and other companies’ patents
- High-throughput screening (HTS)



# Natural Products – uses in the crop protection industry



Biologicals

**Use directly as NP**

**Use after chemical modification**

Synthetic biology

**Prepare analogues by genetic modification**

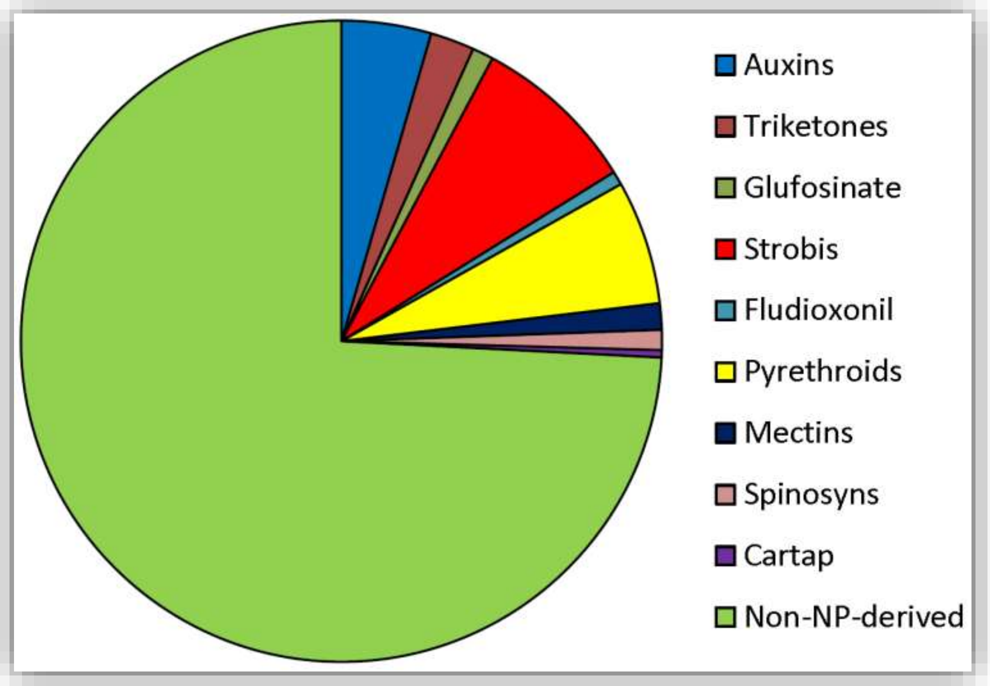
**NPs**

**Validate MoA**

**Lead for synthetic chemistry**

# Natural products – a proven source of innovation

- Analysis based on leading commercial crop protection chemicals with annual sales >$100 million (2011)
- About a quarter of these leading products have their origins in NP leads, with total sales worth ~$10 billion p.a.
- Commercial products based on NPs usually have novel, resistance-breaking modes of action



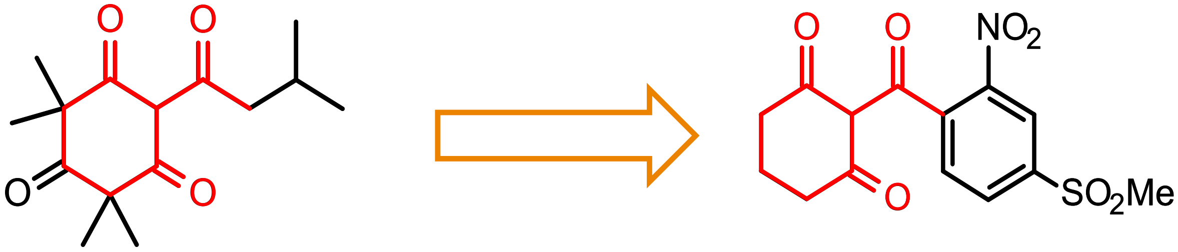
*Market for crop protection chemicals, showing relative values of NP-derived classes (2011)*

*Total value 2011 = $44 billion*

*Source: Phillips McDougall AgriService 2013 v2.0, Product Directory*

# Natural Products

- Often have interesting biological activity and a high degree of structural variation
- Successes are infrequent, but rewards can be huge
- For example, Leptospermone is produced by the bottle brush plant (*Callistemon citrinus*) as a natural chemical defence against competing plant species
- Optimisation led to the commercial herbicide Mesotrione



***Leptospermone***

***Mesotrione***

Callisto®

***Callistemon citrinus***

# Natural Products

- **Sources of natural products:**

    - Literature reports
    - HTS of crude extracts (plants, fungi, bacteria, marine organisms, microbial broths)
    - Commercial collections

- **Strategy:**

    - Confirm the structure from an authentic sample (obtained from the author or *via* targeted isolation or synthesis)
    - Confirm reported biology in our own assays
    - Optimisation

# Mechanism & Structure-Based Design

- **Mechanism-Based Design:**
  - Involves designing an inhibitor of an essential biochemical process in the pest organism
  - This is usually based on the natural inhibitor
  - The new lead is designed to block the biochemical process irreversibly

- **Structure-Based Design:**
  - New lead compounds are modelled to fit within the binding pocket of a protein involved in the targeted mode of action
  - This may be based on actual structures of binding proteins or by modelling related proteins (‘homology models’)
  - Allows the rational design of leads or libraries of compounds

- **Fragment-Based Design:**
  - Small ‘lead-like’ fragments are docked initially into a model binding site and the structures are then built up in an iterative fashion

# Chemical literature and patents

- Many ideas are based on articles appearing in the chemical literature:
    - Published papers describing novel structures with interesting activity (e.g. natural products)
    - Patents published by competitors describing their lead areas
- New projects can be created by exploring gaps in patent claims

| Advantages | Disadvantages |
|---|---|
| Probably the most reliable way to achieve biological activity | Competitors have a head start and will have covered chemical space in filed patents that have not yet been published. |
| Can start lead optimisation process quickly | Several companies working in the same area – similar IP issues as above. |
| | Possibility of insufficient differentiation of products in the market place; products will have the same mode of action |

# *in vivo* High Throughput Screening (HTS)

- In the last decade, biological screens have been miniaturised and automated; increasing potential throughput to >100,000 compounds per year
- Compounds from in-house chemistry projects, combinatorial libraries, collaborations & commercial suppliers are evaluated against business critical targets
- 96-well plates containing live weeds, insects or fungi are cultured in growth cabinets and used to test for activity ***in vivo***
- Activity is assessed and favourable compounds are further investigated in profiling screens in the greenhouse using whole plants and more pest and crop species to get information about potency and selectivity.

**Compound Store**

↓

**HTS**

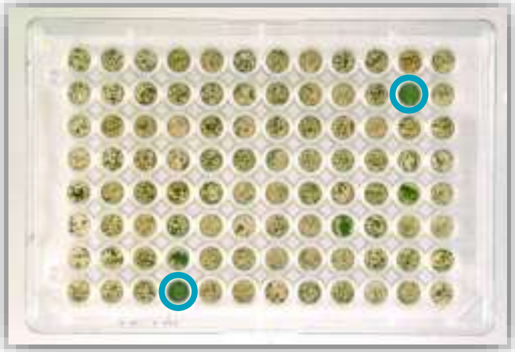
↓

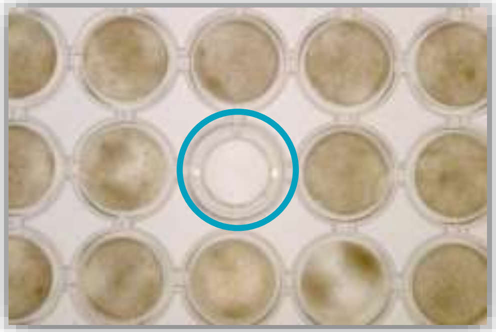
**Assay Analysis**

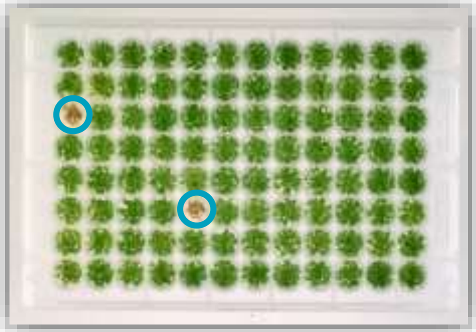
↓

**Profiling screens**

# Screening cascade


**Insecticide HTS**


**Fungicide HTS**


**Herbicide HTS**

- Promising hits from HTS are assessed in profiling screens
- For example, in herbicide projects different weed varieties, rates of application, and pre- or post-emergence application modes are tested


**Screen Preparation**


**Profiling Screens**


**Field Trials**

- Key optimisation compounds are then tested in the field

# Agrochemicals: from application to effect

**Taught Course**

Classification: PUBLIC

# Agrochemicals: From application to effect

How do we go from this….

to this…?

# Agrochemicals: From application to effect

1. **Application**: application of agrochemical to field.
2. **Transfer Processes**: agrochemical movement to the point of absorption.
3. **Absorption**: agrochemical enters the organism.
4. **Translocation**: agrochemical moves within the organism to the active site.
5. **Biological effect**: agrochemical reaches the active site and triggers the desired biological effect.


Application

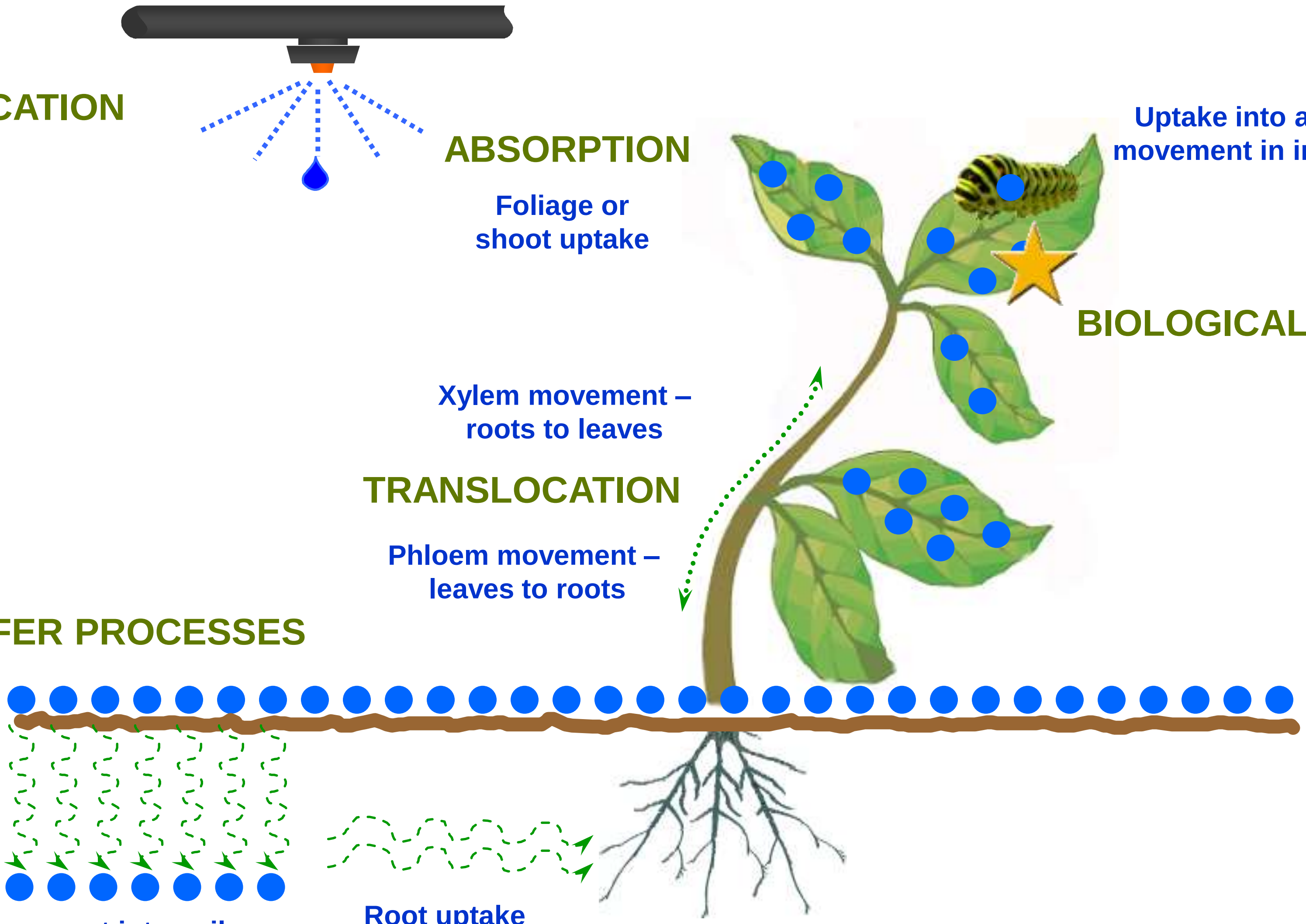
Transfer Processes

Absorption

Translocation

Biological Effect

# From application to effect – the ideal situation



APPLICATION

ABSORPTION

Foliage or
shoot uptake

Uptake into and
movement in insect

BIOLOGICAL EFFECT

Xylem movement –
roots to leaves

TRANSLOCATION

Phloem movement –
leaves to roots

TRANSFER PROCESSES

Movement into soil

Root uptake

ABSORPTION

# Application of herbicides

## Pre-emergent

Applied to soil before germination, before the shoots start emerging.

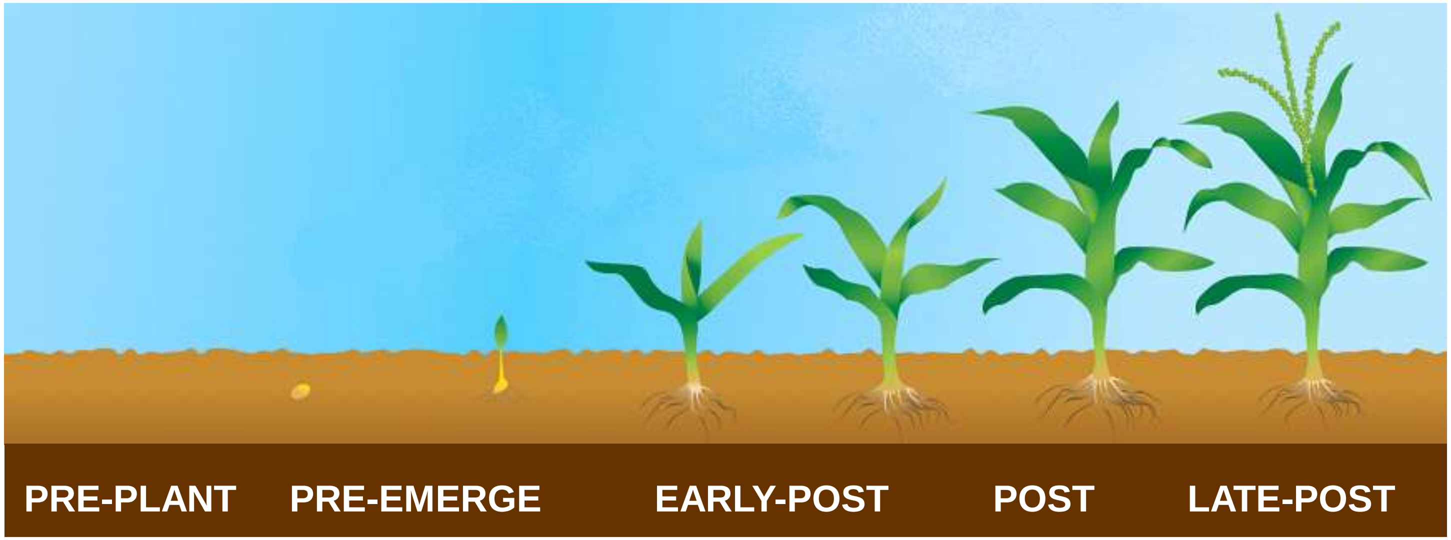
Root or shoot is destroyed during germination.

Pre-emergent herbicides have little effect on established plants.

## Post-emergent (early or late)

Applied to foliage after germination.

The established plant is controlled after herbicide application.



# Application of fungicides and insecticides

- **Post-emergent** application most common
- May be preventative or curative

- **Seed treatment** is an alternative
  - Involves the chemical treatment of the seeds/seedlings (usually with fungicide or insecticide, rarely used as a herbicide treatment) to protect it from a range of pathogenic organisms in the environment such as:
    - Soil-borne diseases
    - Early foliar diseases
    - Insect pests
    - Nematodes

# Spray application

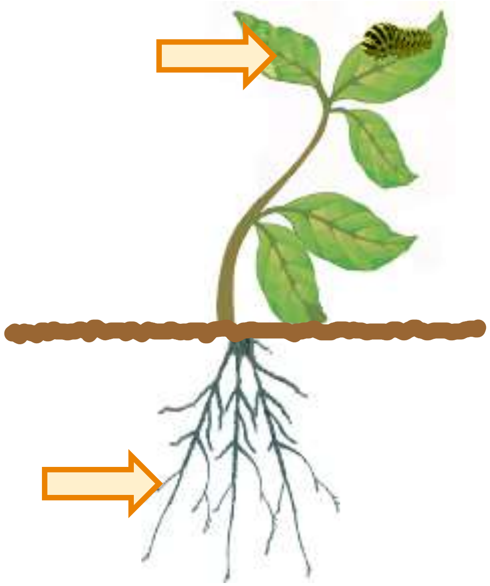
- Most pesticides are applied as dilute aqueous solutions (or suspensions/emulsions)
- The exact **formulation** can have a dramatic effect on activity
- Type of spraying used depends upon a number of factors

# Absorption of agrochemicals into plants

- **Absorption** is the uptake of agrochemicals into organisms and is essential for them to be effective

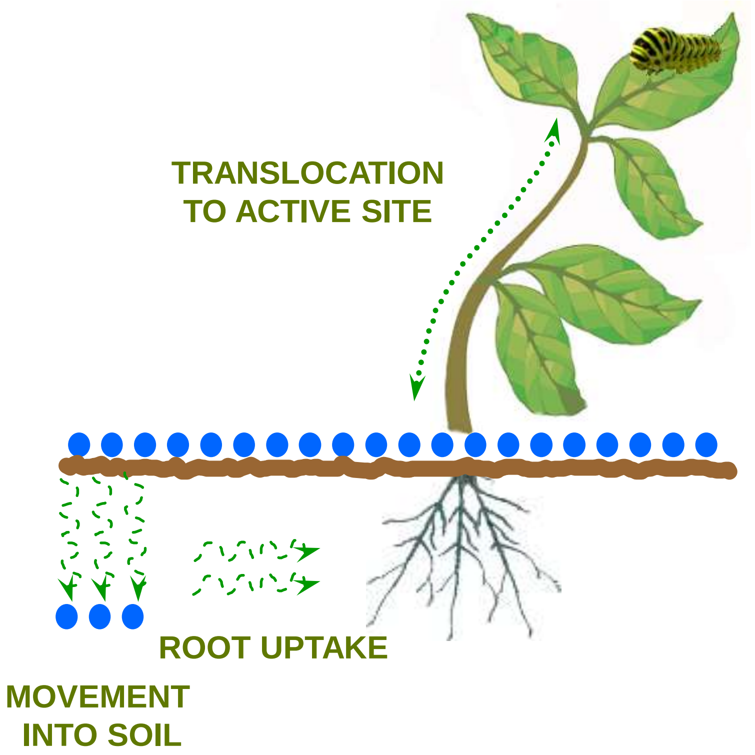
- The common methods in which an agrochemical may enter a plant are:
  - Shoot or foliar uptake (post-emergence)
  - Root uptake (pre-emergence)

- Both physical and chemical properties of an agrochemical will affect how readily it is absorbed



# Absorption of pre-emergence agrochemicals

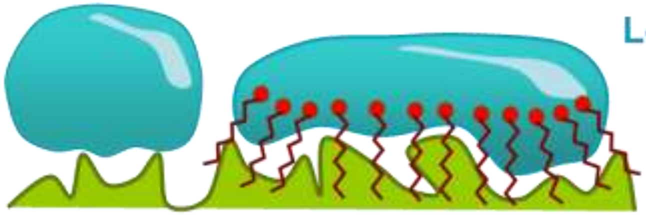
- Active ingredient must be **transferred** successfully through the soil to the plant or organism for absorption to take place
- Agrochemicals that strongly adsorb to the soil are less likely to be taken up by the roots/shoots of the plant
- Less strongly adsorbed agrochemicals that leach or form deposits on the top layer of soil have less chance of reaching the plant target
- A balance of adsorption characteristics is required (see later)



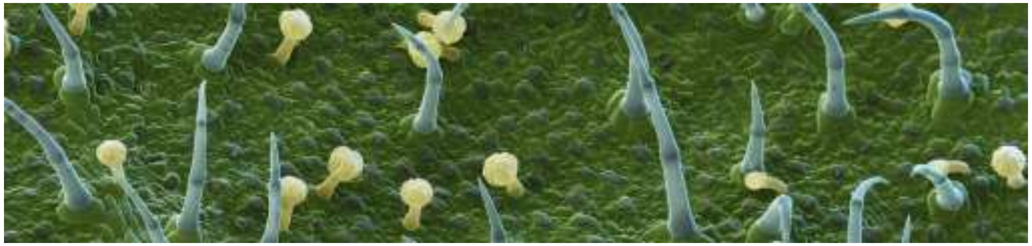
# Absorption of post-emergence agrochemicals

- Absorption is affected by:

1. The **surface tension** of the spray solution.



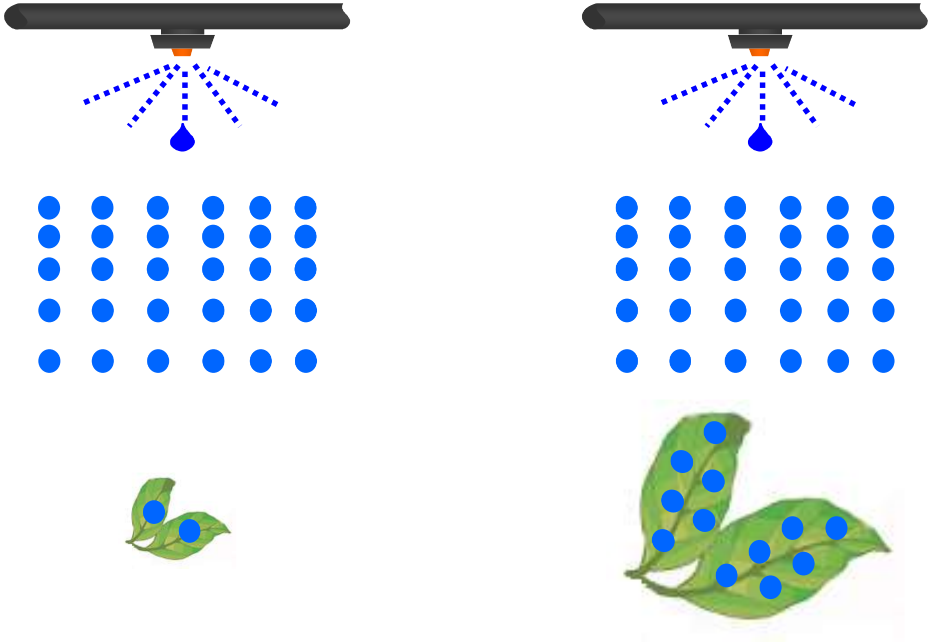
2. The inherent **structure** of the leaf surface.
   - The amount of cuticular wax and physical structure of the wax.
   - The hairiness (number of trichomes) on the leaf surface.



# Physical factors affecting absorption

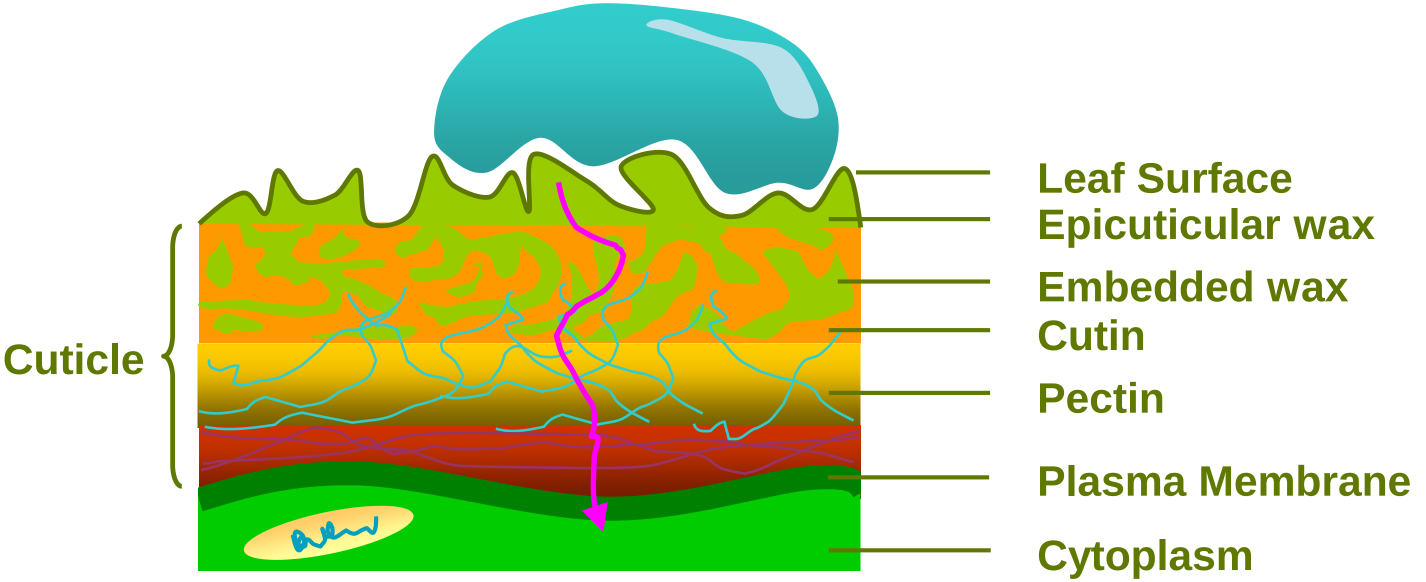
3. Leaf **orientation** with respect to incoming spray droplets.

4. The total **leaf area** per plant (probability of intercepting a spray droplet).



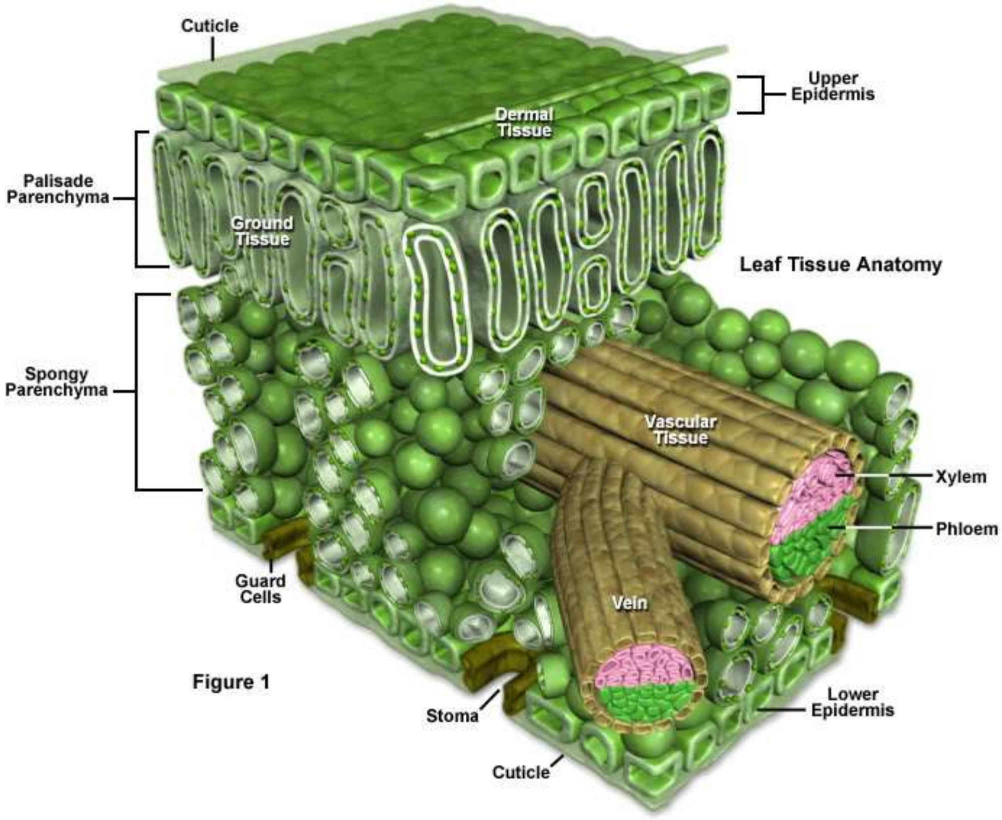
# Chemical factors affecting absorption

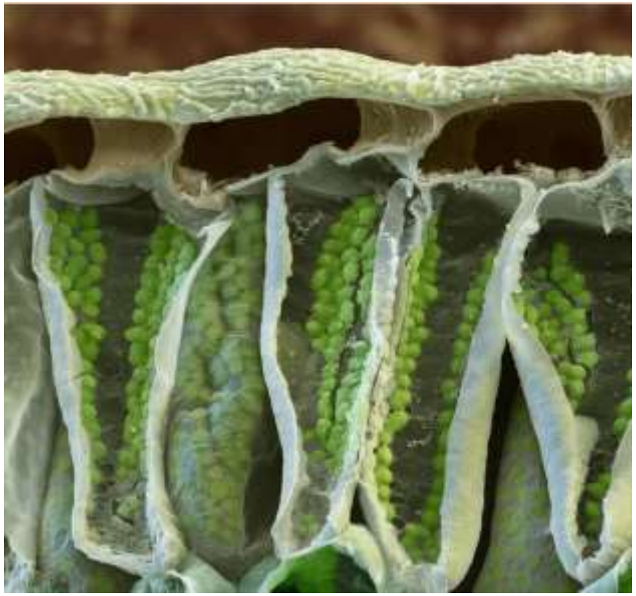
- The cuticle is the primary barrier for absorption of pesticides and consists of waxes (lipophilic), cutin (less lipophilic) and pectin (hydrophilic).



- A balance between hydrophilicity and lipophilicity is required so that absorption and movement through the cuticle can occur – **see later**

# Absorption of post-emergence agrochemicals: Leaf structure

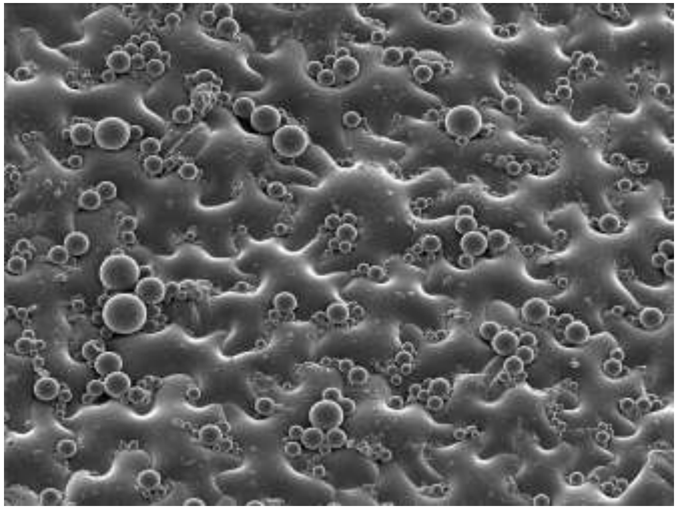




# Absorption of insecticides

- **Contact:**
  - Insecticide residues remain on the surface of the plant. The insect comes in contact with the material as it walks across the treated surface
  - If the insect is present at the time of application, the spray may also cover the insect and penetrate its body directly
  - e.g. Lambda-cyhalothrin

Karate® with Zeon technology



*Microcapsules on a Bean Leaf*

*Microcapsules on an Ant*

1-20 microns

# Absorption of insecticides

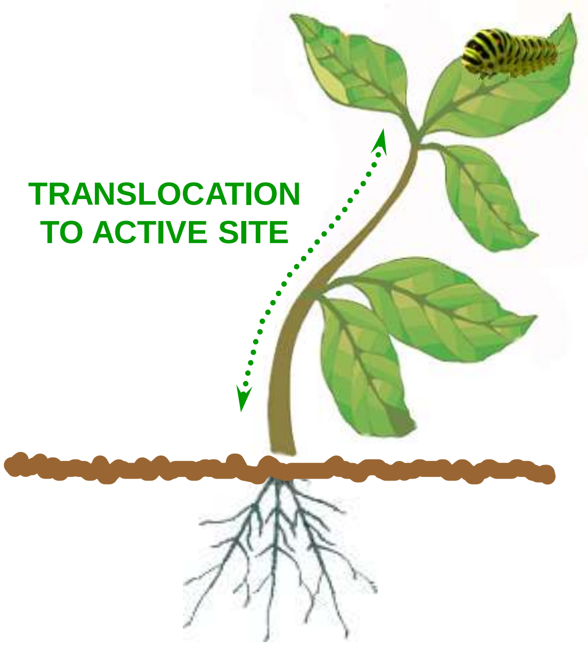
- **Inhalation:**
  - Insecticide vapour enters the insect's breathing apparatus.
  - Useful in enclosed areas where the vapours can remain concentrated.
- **Ingestion:**
  - The insect ingests the insecticide and absorbs it through the stomach lining
  - Ingestion can be more toxic to the insect than direct contact
  - Different pests feed in different ways e.g. chewing pests vs sucking pests

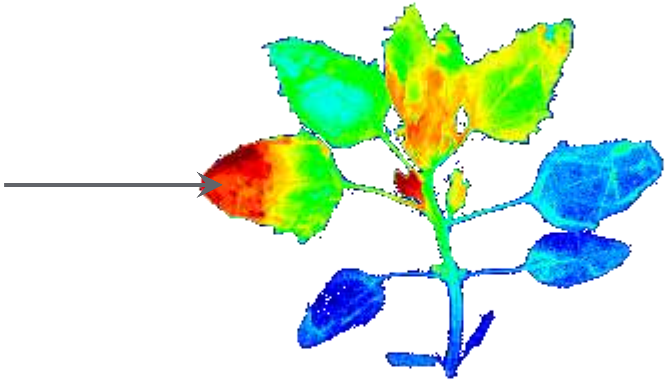
# Absorption of fungicides

- **Contact:**
  - Remain on the surface of the plant.
  - Inhibit spore germination but do not work if a plant is already infected.
- **Systemic:**
  - Plant absorbs the fungicide, which then translocates to other parts of the plant which is infected by the fungus.
  - Systemic fungicides work best by inhibiting growth of fungi after infection.

# Translocation in plants

- Once absorbed, the agrochemical must pass through various cells and eventually enter the plant's vascular system in order to be effective.
- **Translocation** is the movement of soluble materials away from the site of absorption *via* the xylem and phloem.
- Both physical and chemical properties of an agrochemical will affect translocation and hence, its effectiveness.

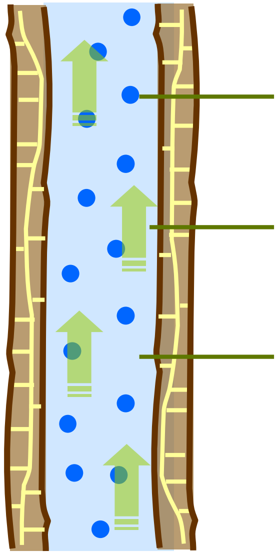




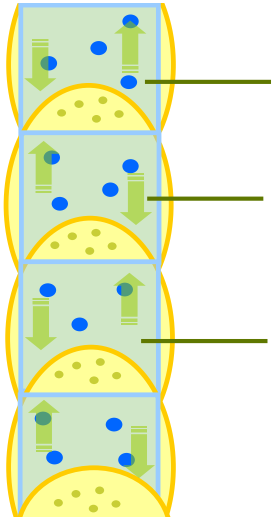
# Methods of translocation

- The **xylem** and **phloem** are the two main transport systems that move nutrients and agrochemicals throughout the plant.

- In both systems, the agrochemical is dissolved in water and moves along with the mass flow of water.



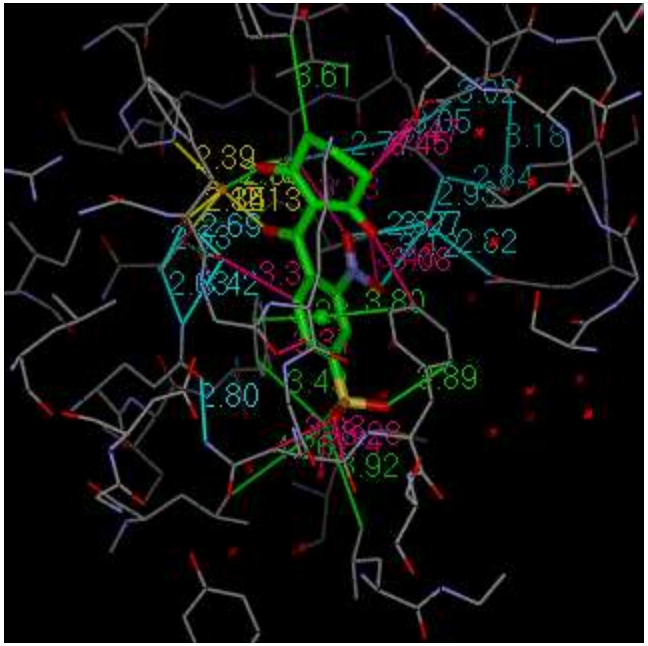
Xylem vessel

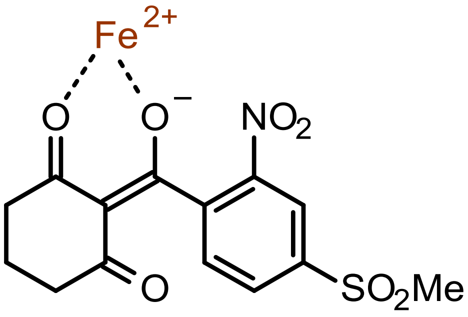


Phloem vessel

# Biological effect

- After reaching the target site, the agrochemical has its desired **biological effect** by disrupting or interfering with vital physiological functions
- This is achieved by binding to active sites (*via* covalent bonding, metal complexation, ionic bonding, H-bonding, hydrophobic interactions, or charge transfer interactions)
- The method in which the specific physiological effect is disrupted is called the **mode of action** (MoA)
- E.g. the herbicide mesotrione inhibits an iron-containing enzyme that is involved in photosynthesis





**Mesotrione**

# Molecular interactions

- Intrinsic binding is comprised of **enthalpic** and **entropic** contributions

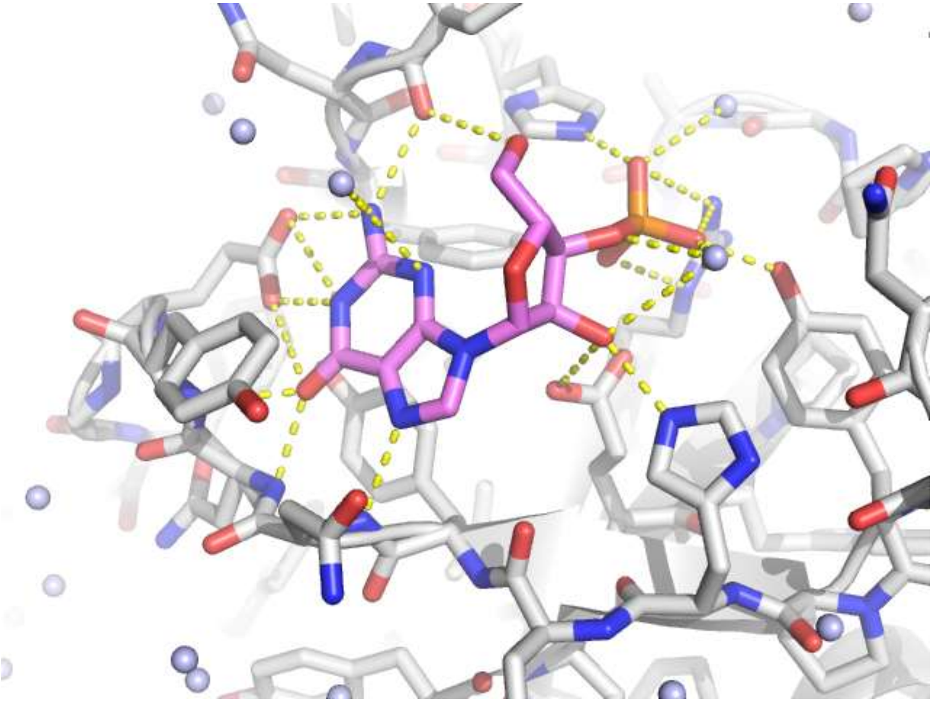
$$\Delta G^o = \Delta H^o - T\Delta S^o$$

- Interactions may be **specific** or **non-specific**
- A large contribution to binding often comes from non-specific lipophilic interactions (e.g. van der Waals interactions)
- It is important to also consider the effect of **desolvation** (ligand and protein)
- Specific interactions include hydrogen-bonding, π-stacking, halogen-bonding, cation-π interactions…and many more
- See *J. Med. Chem.*, **2012**, *53*, 5061-5084 *“A Medicinal Chemist’s Guide to Molecular Interactions”*

# Molecular interactions – hydrogen-bonding

- Hydrogen-bonding interactions are by far the most frequently observed specific interactions in biological systems

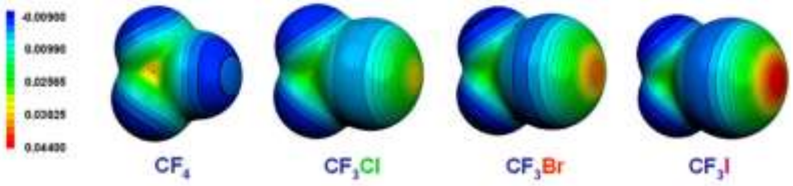
- Hydrogen-bond strength is a function of the H-bond acidity (α) of the donor, the H-bond basicity (β) of the acceptor and the geometry of the H-bond



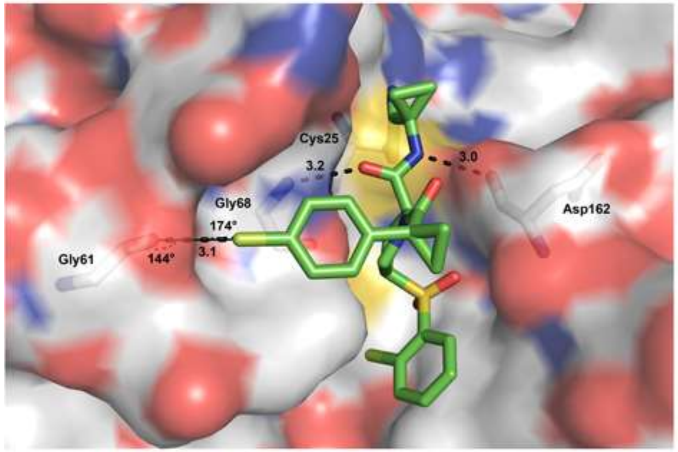
# Molecular interactions – halogen-bonding

- Halogen-bonding is much less common but can be important: see *J. Med. Chem.*, **2013**, *56*, 1363-1388
  *“Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology”*

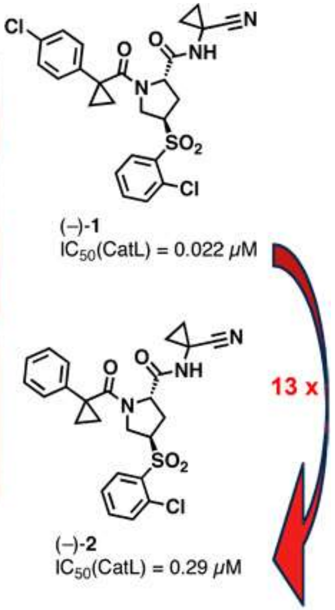
- Anisotropic effects around halogen nuclei result in areas of positive charge in the plane of the C-X bond (the σ-hole)



- The effect is most pronounced for iodine but significant effects can be observed for other halogens e.g.



Cathepsin L inhibitors – *Francois Diedrich et al.*

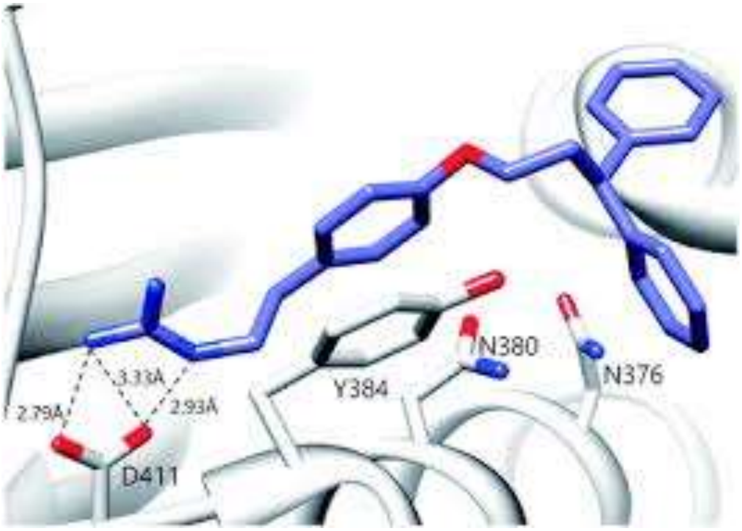


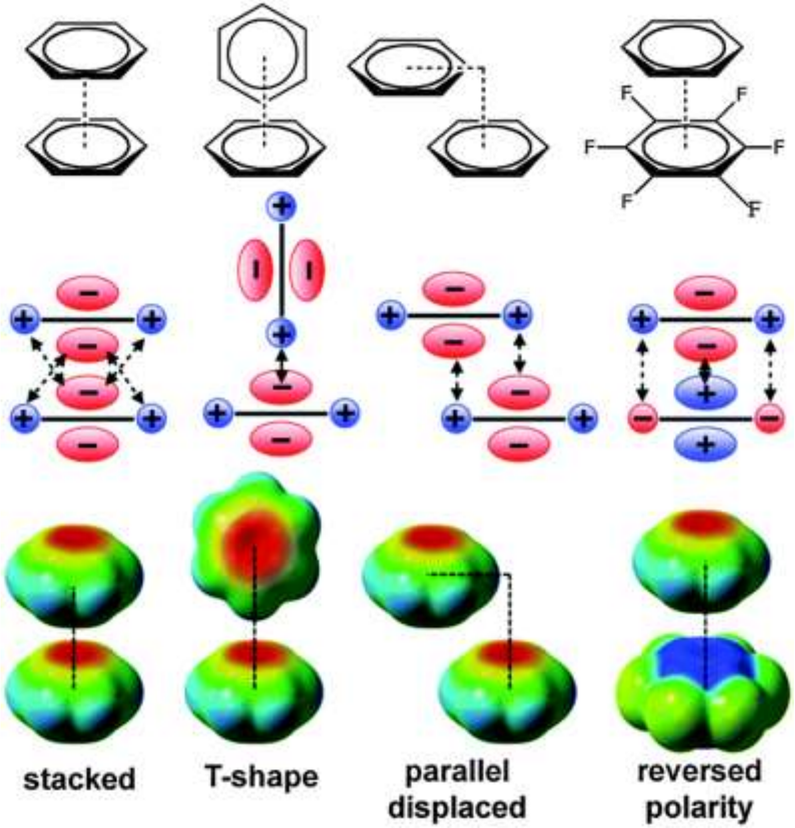
(–)-**1**
$IC_{50}$(CatL) = 0.022 μM

(–)-**2**
$IC_{50}$(CatL) = 0.29 μM

# Molecular interactions – π-stacking

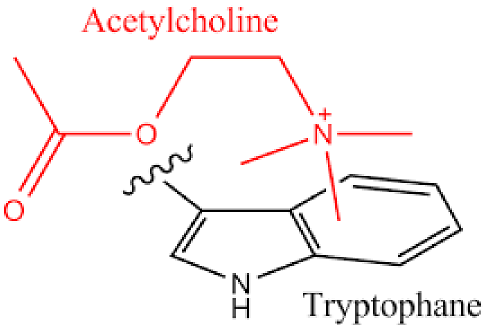
- Interactions between π-systems are commonly observed in biological systems
- Strength and orientation of the interaction depends on the electronics of the π-systems involved

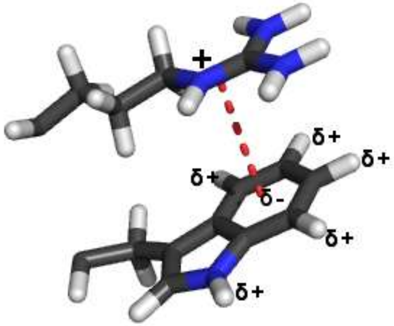


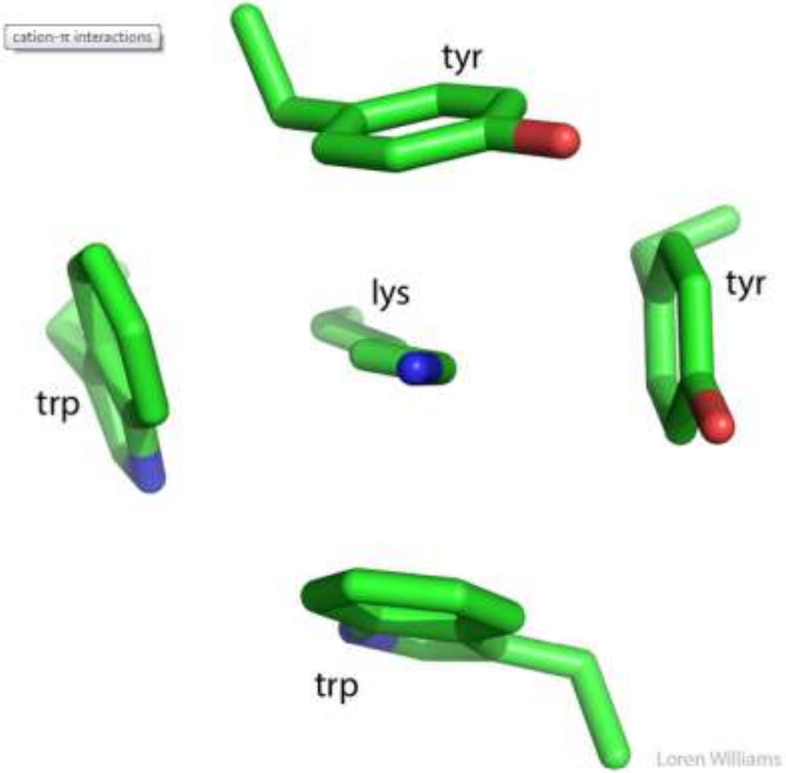


# Molecular interactions – cation-π interactions

- Cation-π interactions are often observed between e-rich aromatic systems (e.g. indoles, phenols) and cations (e.g. ammonium, guanidinium)
- For a review, see *Chem. Rev.*, **1997**, *97*, 1303-1324

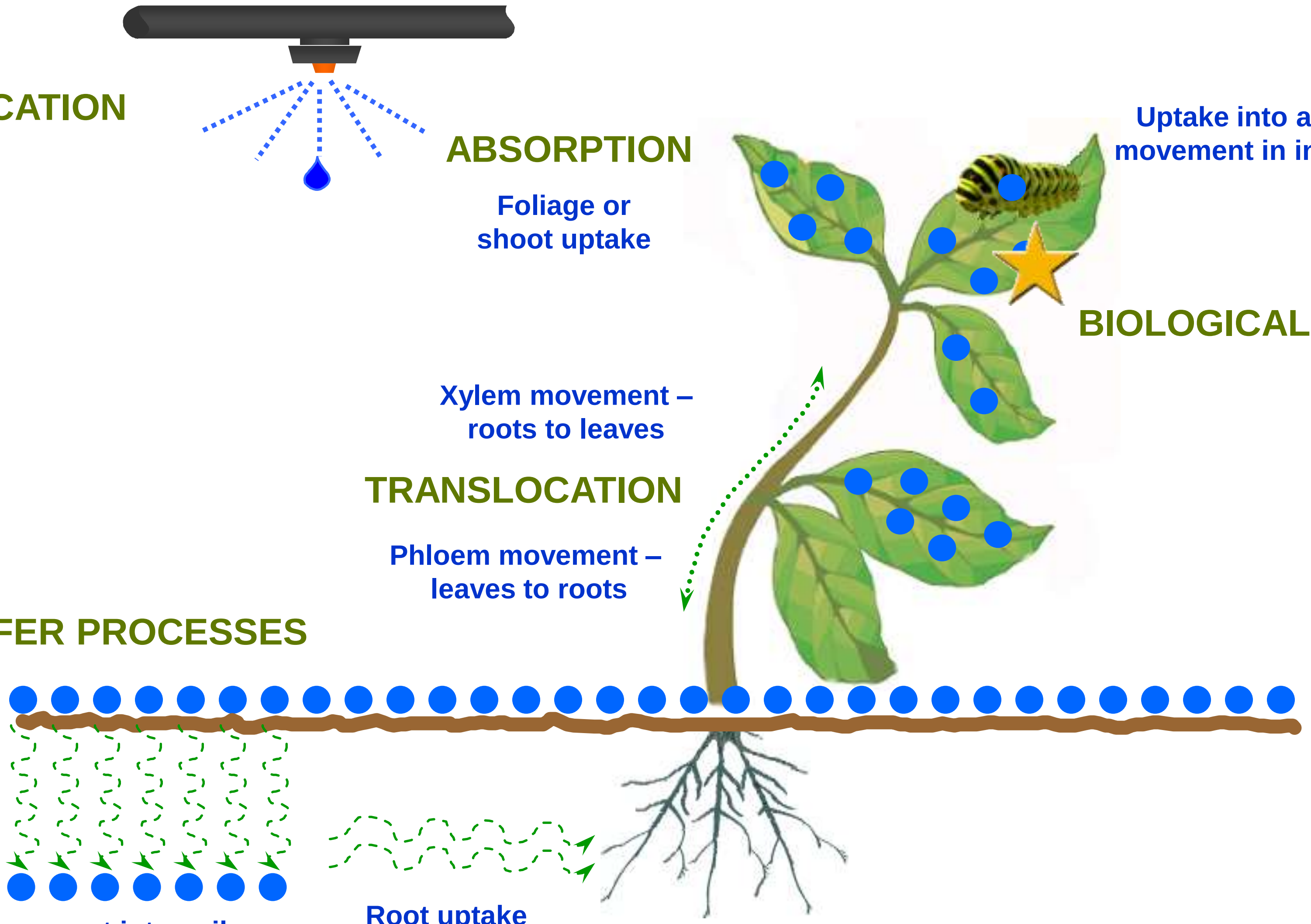






# Application to effect – the reality

# From application to effect – the ideal situation



APPLICATION

ABSORPTION

Foliage or
shoot uptake

Uptake into and
movement in insect

BIOLOGICAL EFFECT

Xylem movement –
roots to leaves

TRANSLOCATION

Phloem movement –
leaves to roots

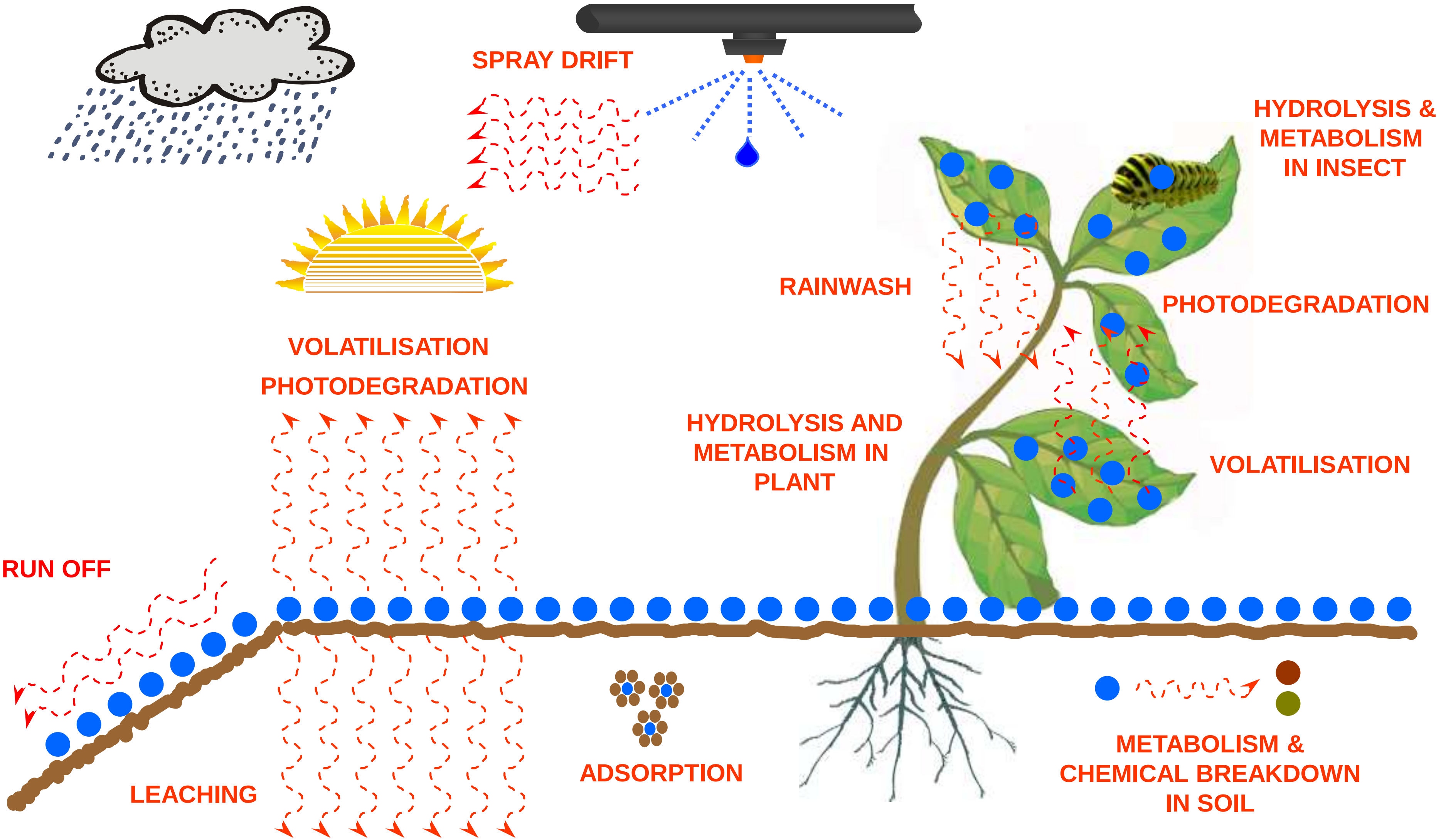
TRANSFER PROCESSES

Movement into soil

Root uptake

ABSORPTION

# Application to effect – the reality



SPRAY DRIFT

HYDROLYSIS & METABOLISM IN INSECT

RAINWASH

PHOTODEGRADATION

VOLATILISATION

PHOTODEGRADATION

HYDROLYSIS AND METABOLISM IN PLANT

VOLATILISATION

RUN OFF

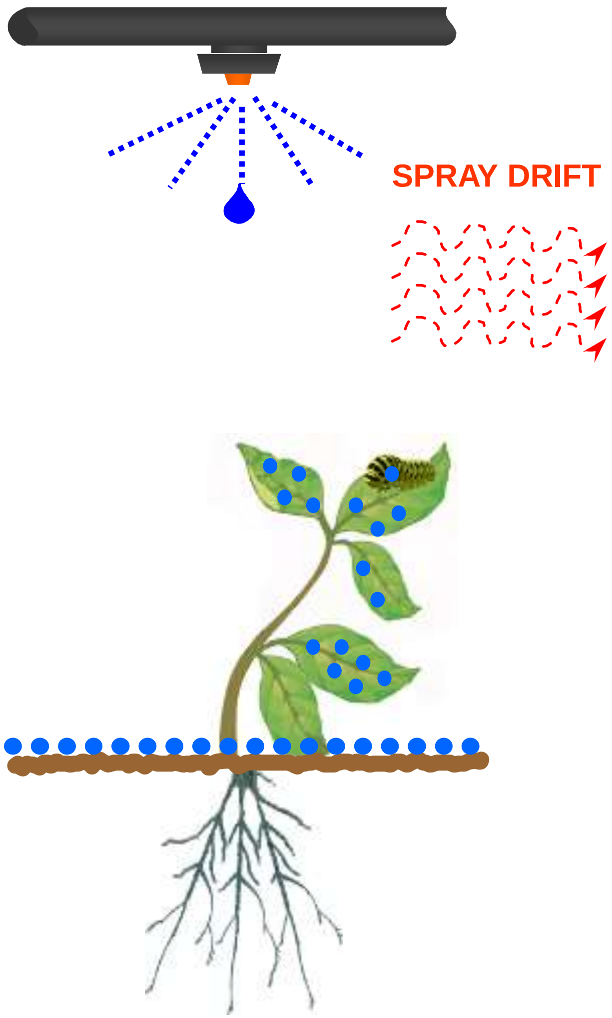
LEACHING

ADSORPTION

METABOLISM & CHEMICAL BREAKDOWN IN SOIL

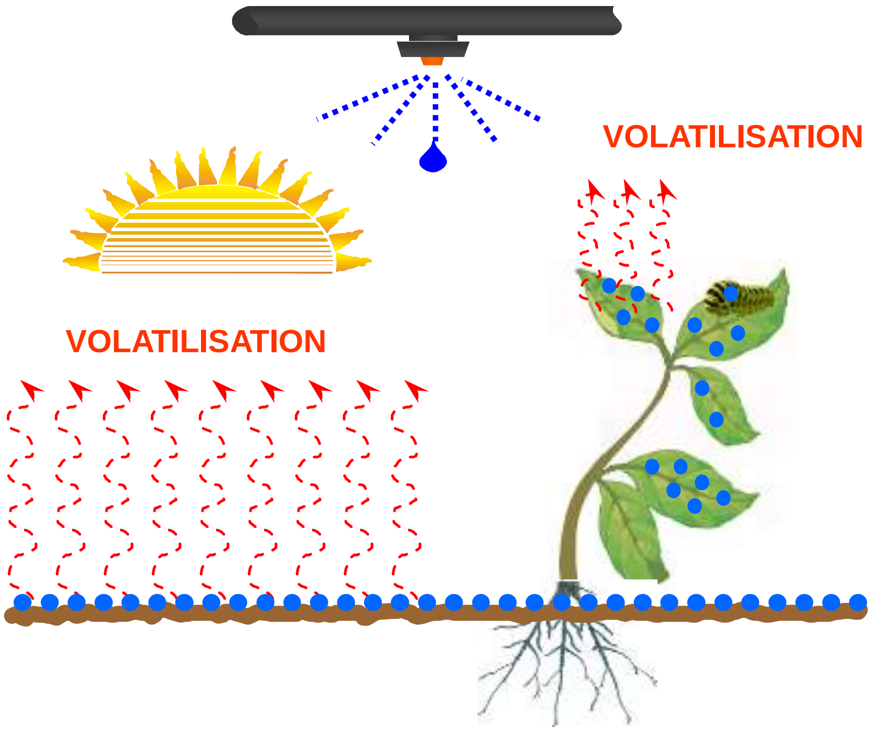
# Spray Drift

- **Spray drift** is the airborne movement of spray droplets away from treatment site during application
- Can damage nearby sensitive crops or can contaminate crops ready to harvest
- Can be successfully managed by using specially designed spray nozzles and monitoring of local weather changes.



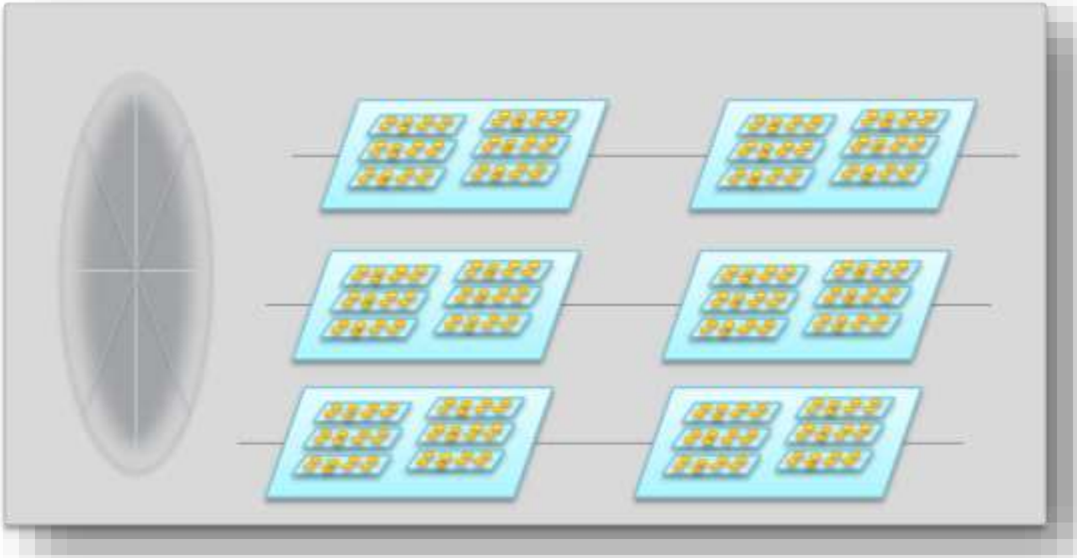
# Volatilisation

- Process of solids or liquids converting into a gas, thus moving away from the initial site of application. This movement is called **vapour drift**
- Hot, dry, or windy weather increase volatilisation
- Active ingredients with a high boiling point and molecular weight tend to be less volatile



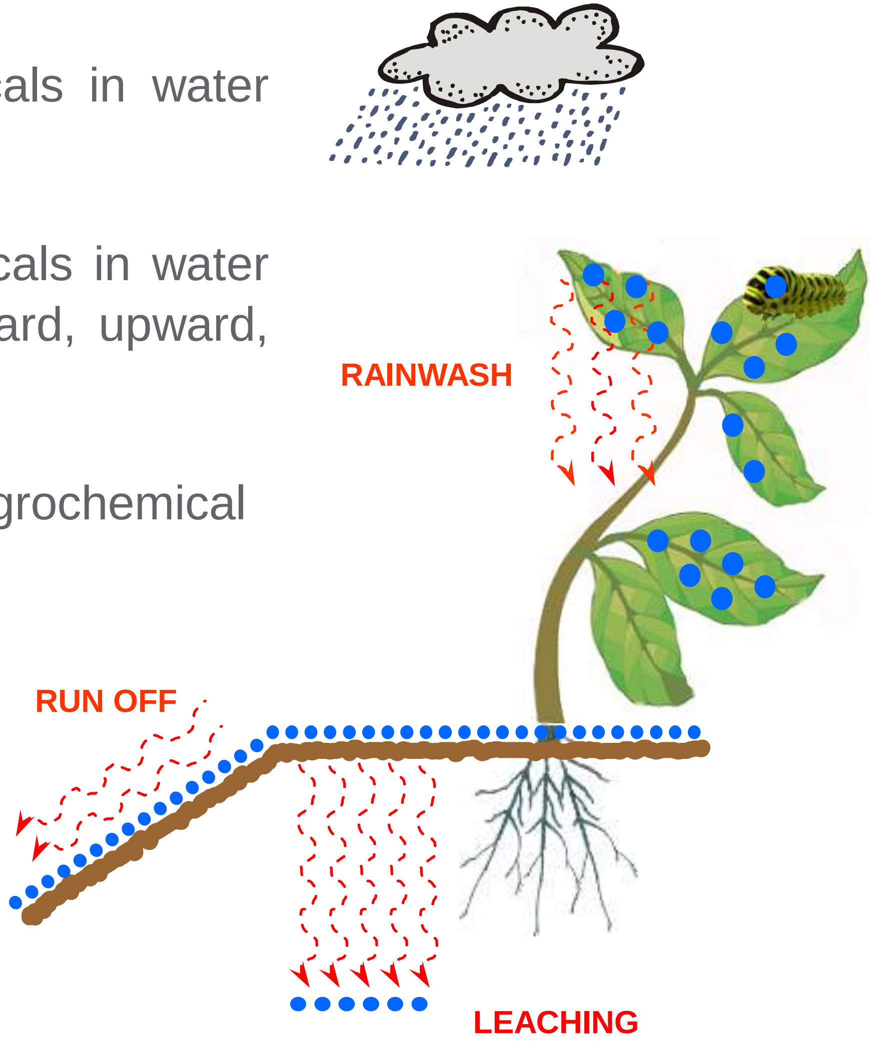
# Volatility and vapour pressure

- Volatility can be measured by recording loss of a compound over time in a wind tunnel
- A $T_{50}$ (time for 50% loss) can be measured and can vary from minutes to days



# Run Off & Leaching

- **Run off** is the movement of agrochemicals in water over a sloping surface
- **Leaching** is the movement of agrochemicals in water through the soil. Leaching occurs downward, upward, or sideways
- Heavily influenced by the solubility of the agrochemical
- As well as decreasing the effectiveness of a treatment, run off and leaching can cause water contamination and affect livestock and crops downstream



# Run Off & Leaching

- **Low solubility** (<1 ppm):
  - Lower chance of leaching and moving with surface run off
  - Bind to soil surface
- **High solubility** (>50 ppm):
  - Move with water.
  - Higher chance of leaching and movement with surface run off
  - More potential to be washed off crop

- Agrochemicals with high or low solubilities have less chance of reaching the target – a balance is required
- Active ingredients with a low melting point, that are charged or contain high levels of polar functional groups tend to be more water soluble

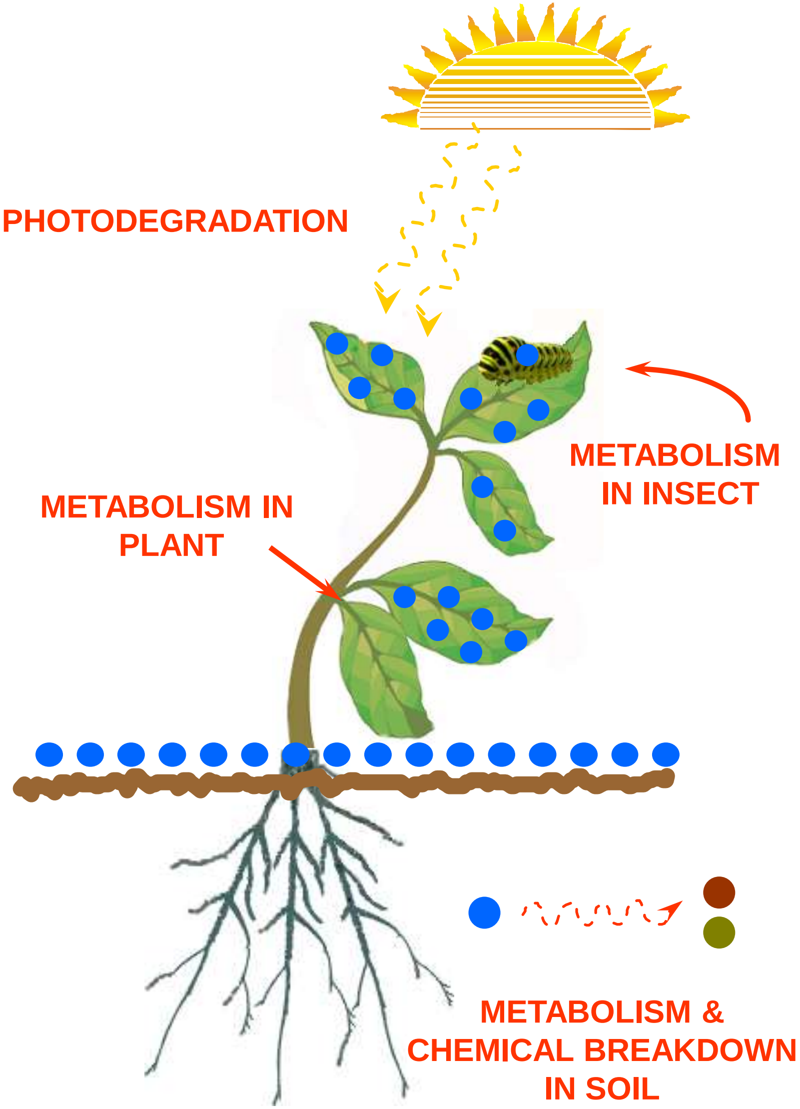
# Degradation or Breakdown Processes

- An agrochemical may break down through three primary modes of degradation:

- **Photochemical breakdown** is the breakdown of agrochemicals by sunlight
  - Extended conjugation can lead to photochemical instability

- **Chemical breakdown** is the breakdown of agrochemicals by chemical reactions
  - Sensitive functionality can be broken down by simple chemical processes

- **Biological metabolism** is the breakdown of agrochemicals within a living organism (e.g. soil microbes, plant, fungi or insect)



# Photostability

- Photodegradation occurs on foliar and soil surfaces and in water bodies
- One of the most destructive post-application pathways - but can be beneficial
- Can be measured by recording loss of parent compound in an artificial sunlight simulator (or outside if the sun ever shines!)

- NB – some very interesting photolytic degradation pathways can be observed

# Metabolism

- **Metabolism** is the molecular modification or degradation of an active ingredient resulting in a change in its structure
- The main purpose of metabolism by an organism is to make it more polar and hence more water soluble, to accelerate excretion or sequestration
- Metabolism normally results in reduced potency but in some cases agrochemicals may be activated by metabolism (procides)
- Even minor changes in structure can have a dramatic effect on activity

  e.g. Simazine is a herbicide which is hydrolysed to hydroxysimazine, a metabolite that is 1000x less biologically active than the active ingredient.

Simazine, Novartis

Hydroxysimazine

# Metabolism

- Some major metabolic pathways for agrochemicals are:
  - **Hydrolysis** of esters, amides, carbamates and nitriles
  - **Oxidation** of alcohols, aldehydes, sulfides and electron rich aromatic rings
  - **Reduction** of ketones, aldehydes and halogenated compounds
  - **Conjugation** of enones or halides to glutathione

Glutathione

Active ingredient

Inactive ingredient

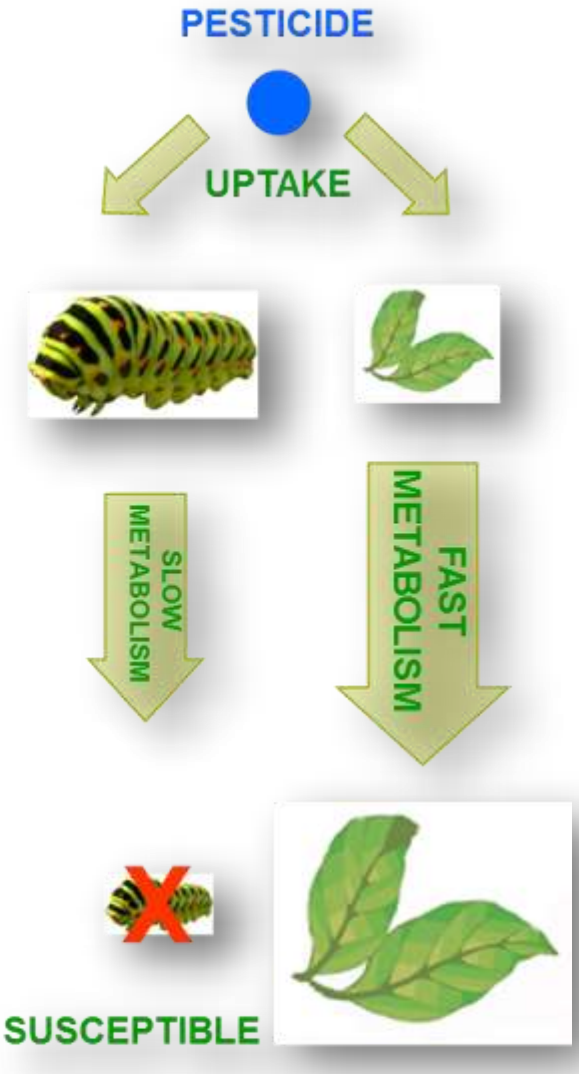
GST = glutathione S-transferase

- This can result in:

A **change in structure** so that it no longer interacts at the active site or a **change in physical properties** so that it can be excreted or sequestered more easily

# Metabolism – Selectivity

- **Metabolism** is one of the most important ways an organism can escape the toxic effects of an agrochemical
- Agrochemical-tolerant plants (crop) metabolise the chemical to inactive compounds before it has a chance to build up to toxic levels at the site of action
- Susceptible pests are unable to metabolise (detoxify) agrochemicals
- Crop **selectivity** may arise from:
    - Selective metabolism of agrochemical by crop
    - Differences at the active site
    - Reduced uptake of agrochemical by crop
    - Reduced distribution of agrochemical within crop



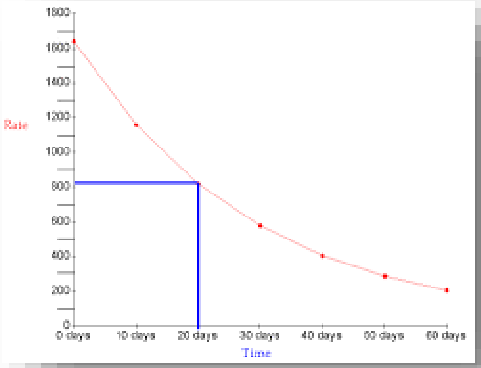
# Agrochemical Persistence

- Agrochemical persistence is measured in terms of **half-life** ($t_{1/2}$), or the time in days required for an agrochemical to degrade to one-half its original amount

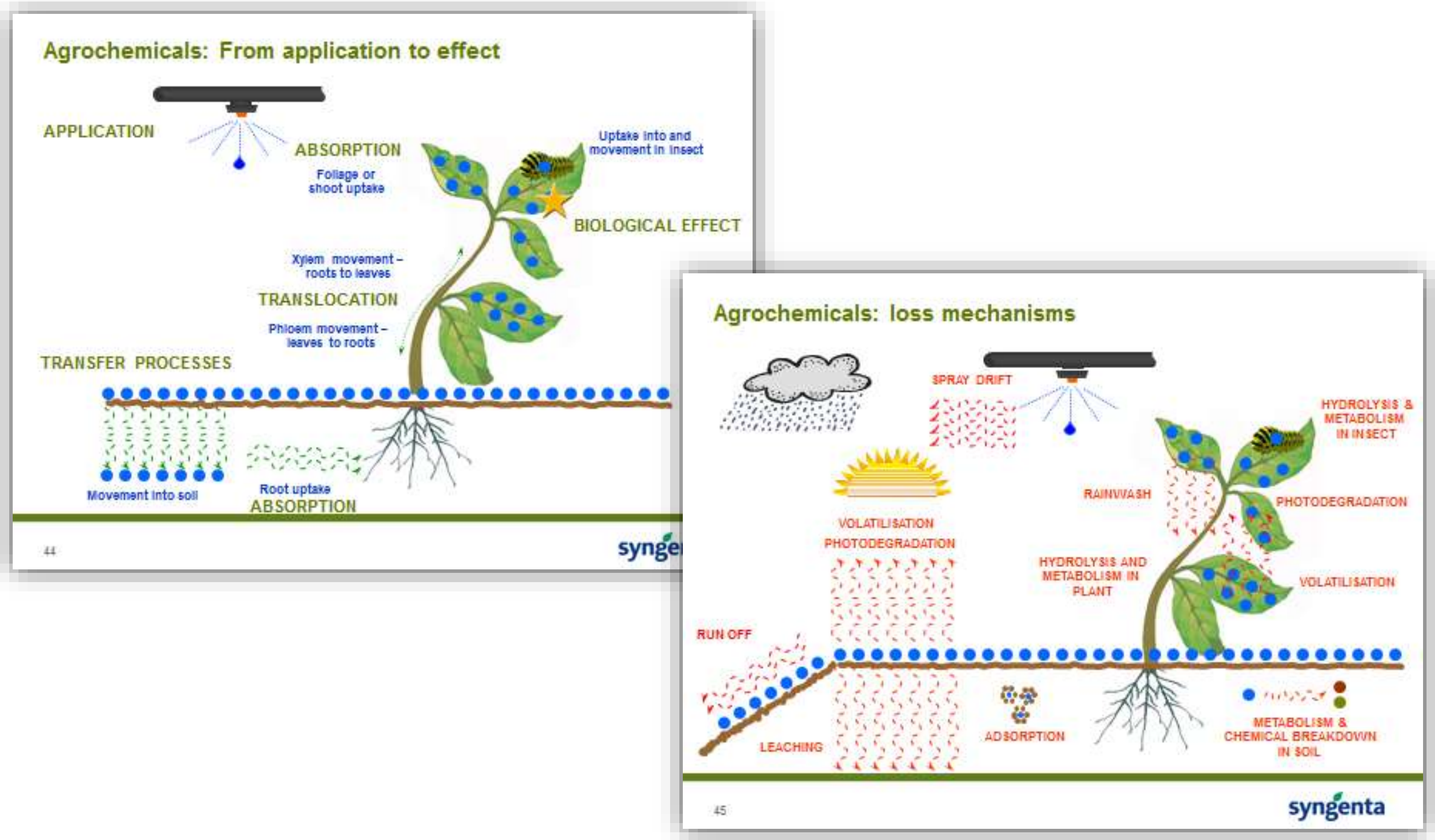
- If the agrochemical is **too persistent** ($t_{1/2}$ > 100 days) it may:
  - leach or move with surface run off or;
  - cause damage to next seasons crops

- If the agrochemical is **non-persistent** ($t_{1/2}$ < 30 days), degradation may occur before it has the desired effect

- The ideal half-life depends upon the application but in general a balance is required



# Physical Properties of Agrochemicals



## Agrochemicals: From application to effect

APPLICATION

ABSORPTION

Foliage or shoot uptake

Uptake into and movement in insect

BIOLOGICAL EFFECT

Xylem movement – roots to leaves

TRANSLOCATION

Phloem movement – leaves to roots

TRANSFER PROCESSES

Movement into soil

Root uptake

ABSORPTION

## Agrochemicals: loss mechanisms

SPRAY DRIFT

HYDROLYSIS & METABOLISM IN INSECT

RAINWASH

PHOTODEGRADATION

VOLATILISATION

PHOTODEGRADATION

HYDROLYSIS AND METABOLISM IN PLANT

VOLATILISATION

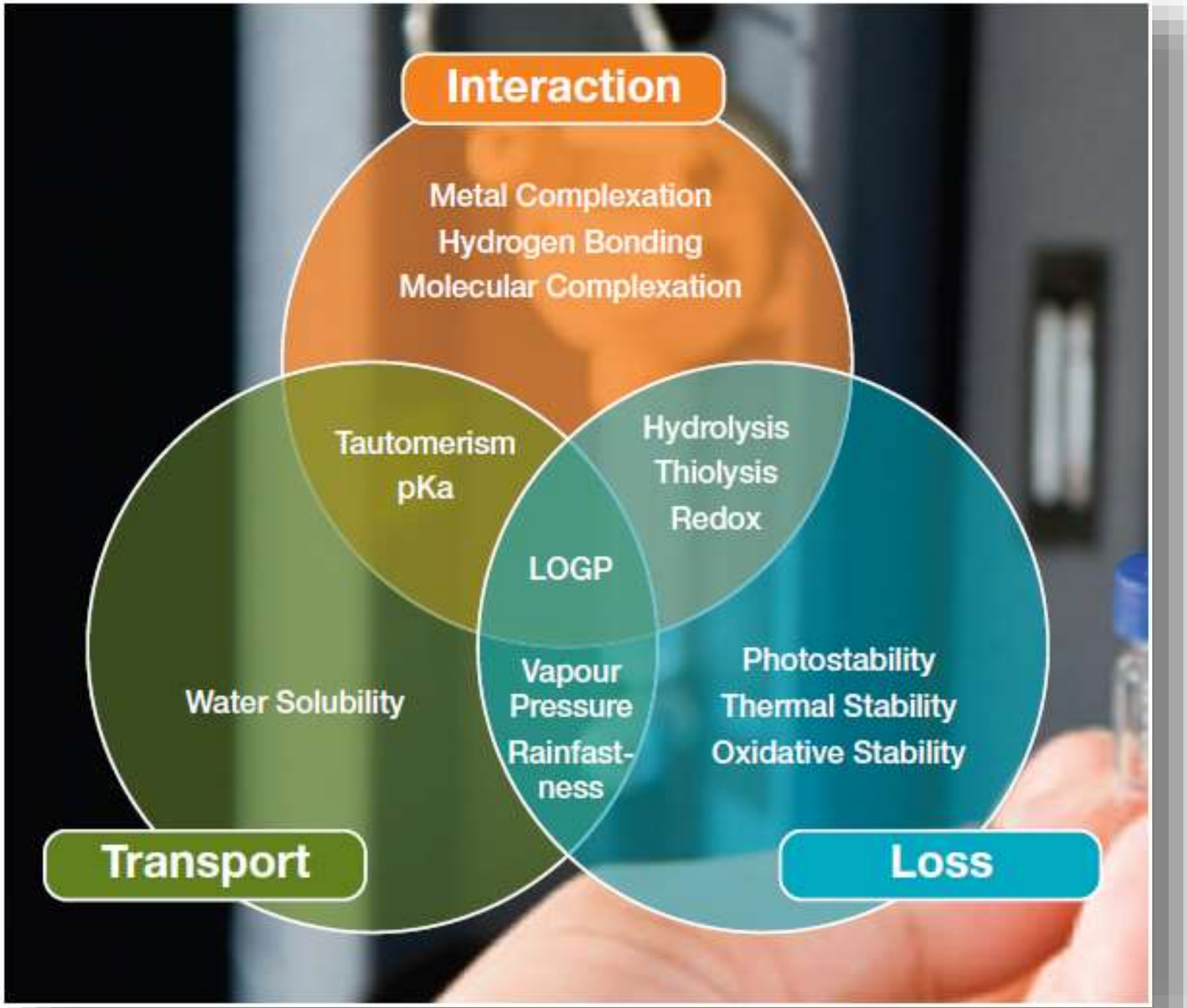
RUN OFF

LEACHING

ADSORPTION

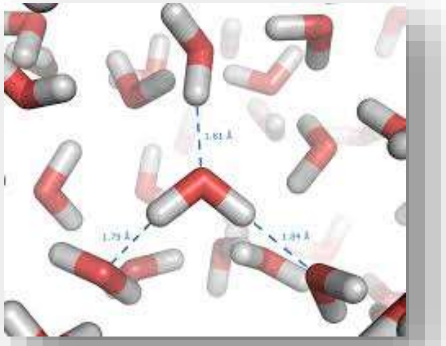
METABOLISM & CHEMICAL BREAKDOWN IN SOIL

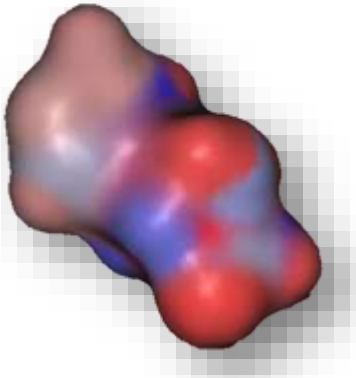
# Physical Properties of Agrochemicals



# Lipophilicity and Hydrophilicity (log P)

- **Lipophilicity** is a key physical property for predicting the bioavailability of agrochemicals and is related to **log P**.
- To get into target organisms, agrochemicals usually rely on passing through lipophilic membranes and then moving to their site of action in a more polar aqueous environment.
- It is essential therefore to have a balance between **lipophilicity** for penetration and **hydrophilicity** for effective mobility within the organism.

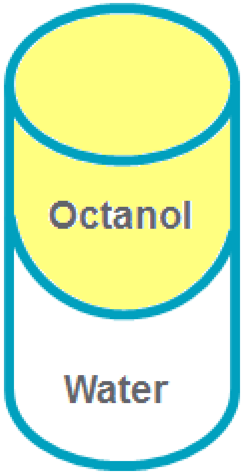


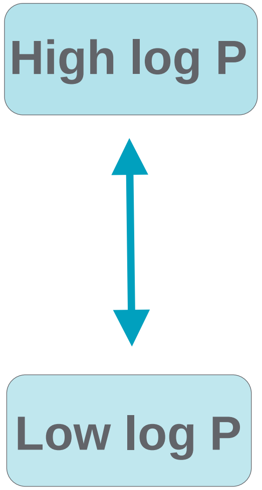


# Partition Coefficient: log P

- The lipophilicity of a molecule can be quantified experimentally by measuring the relative distribution of the **neutral** molecule in octanol/water.
- The relative distribution is known as the partition coefficient (P).

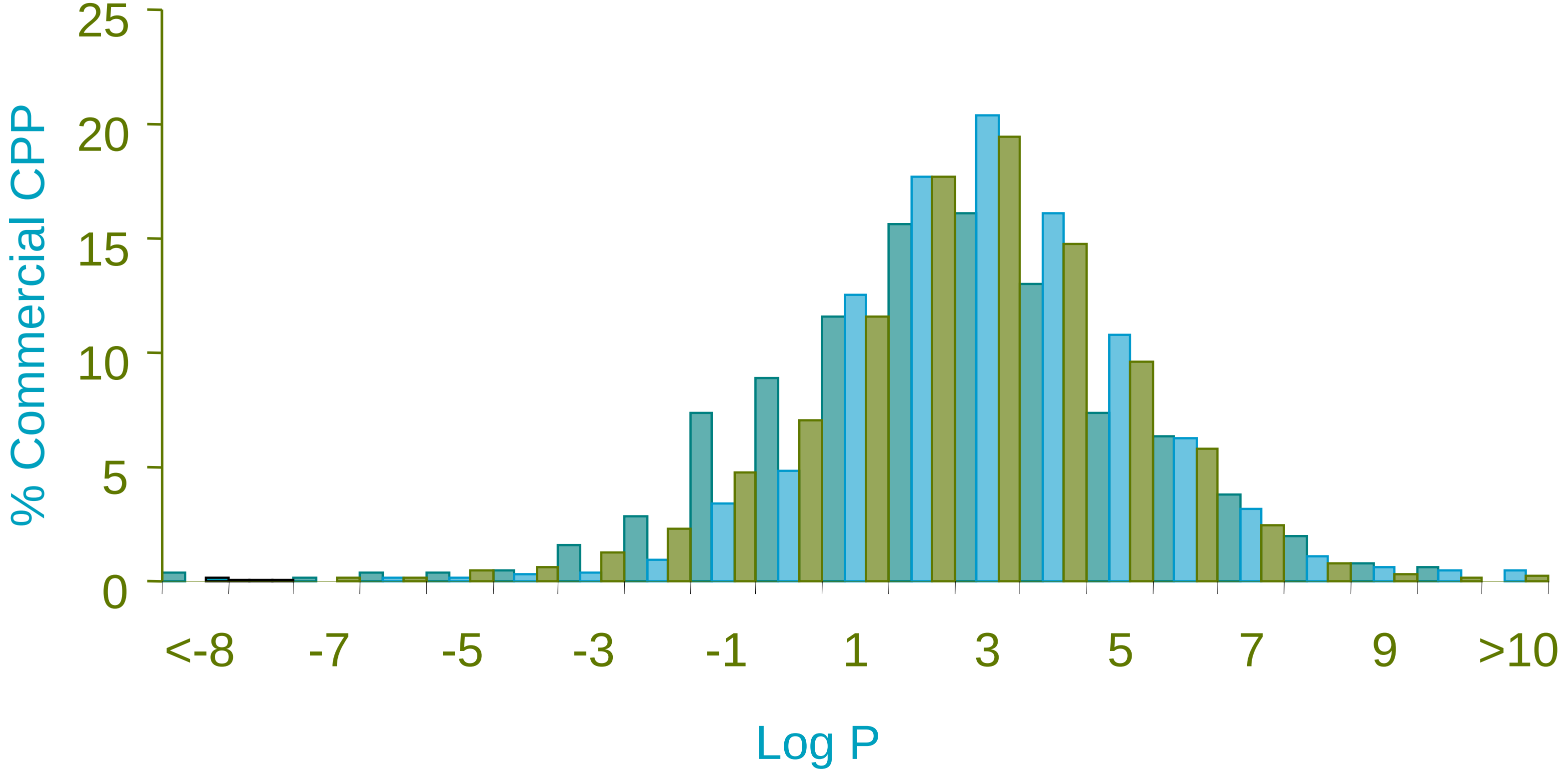
$$P = K_{ow} = \frac{[\text{octanol}]}{[\text{water}]}$$





- P can be very high or low so usually expressed as log P.
- Log P = $\log_{10}$ (P).
- High log P means [oct] > [water] i.e. Molecule is more lipophilic.

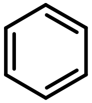
# ‘Effective’ Crop Protection Products: Log P Profile



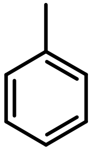
- The majority of agrochemicals have a log P of around 3, but the optimum depends on the target and clearly there is a wide range!

# Factors Affecting log P

- Addition of substituents/fragments to a molecule can dramatically change its overall log P
- Effects are *usually* additive

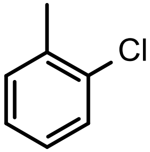


Log P = 2.2



Log P = 2.7

$\Delta = +0.5$



Log P = 3.4

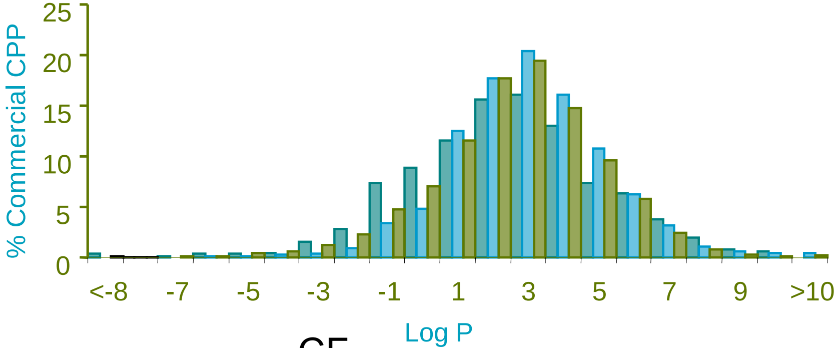
$\Delta = +0.7$

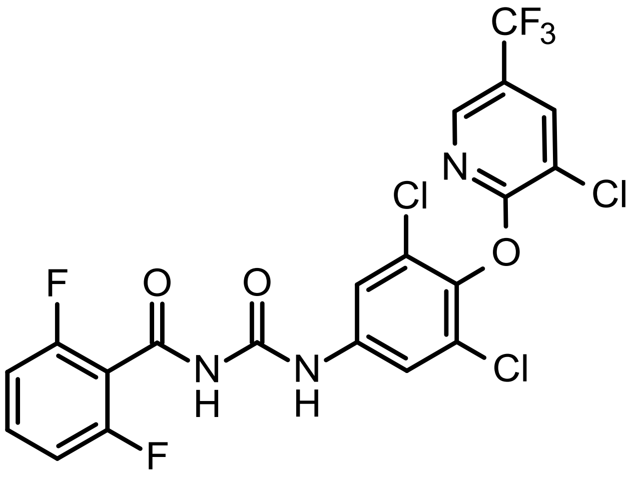
- Log P can be calculated by knowing the contribution that the various substituents make to lipophilicity. This contribution is known as the **lipophilic substituent constant** ($\pi$).

$$\boxed{\pi_X = \log P_X - \log P_H}$$

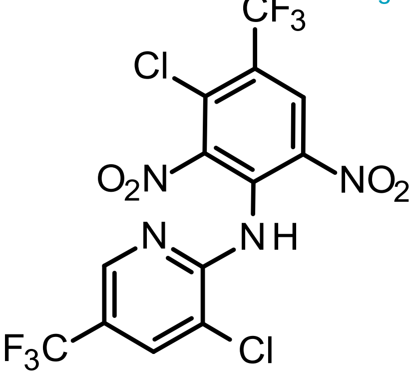
- Where X = analogue and H = parent compound.

# logP of some Herbicides

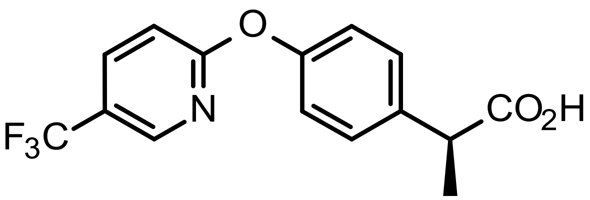




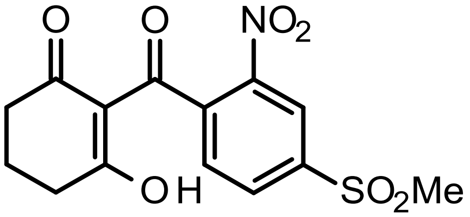
**Chlorfluazuron** **logP 5.8**



**Fluazinam** **logP 4.03**



**Fluazifop** **logP 3.18**



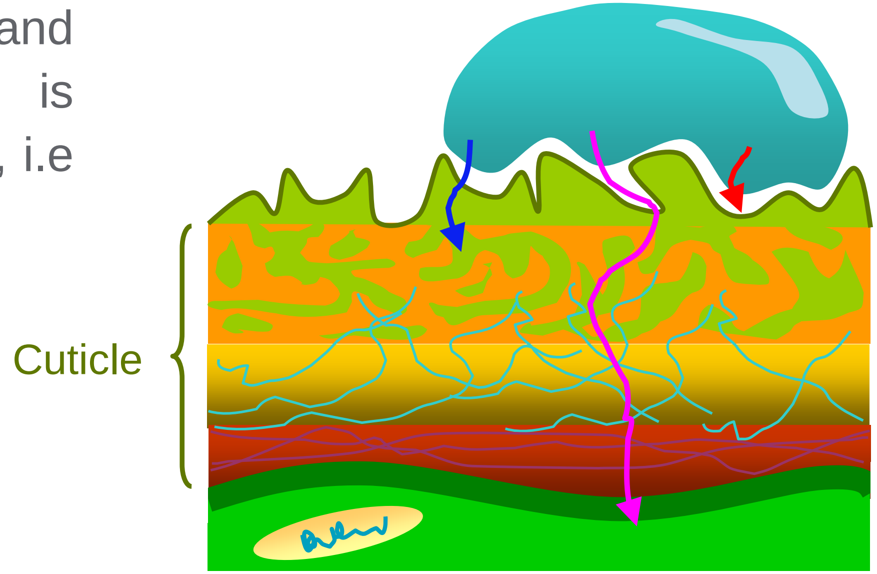
**Mesotrione** **logP 3.12**

# Log P Summary

- Compounds with low log P: absorption will be slow because the chemical can only cross the lipid bilayers slowly

- Compounds with high log P: the chemical will be held up in the lipid bilayers

- Transfer across internal tissues and movement through vascular tissue is maximal at an intermediate lipophilicity, i.e where log P ~3.

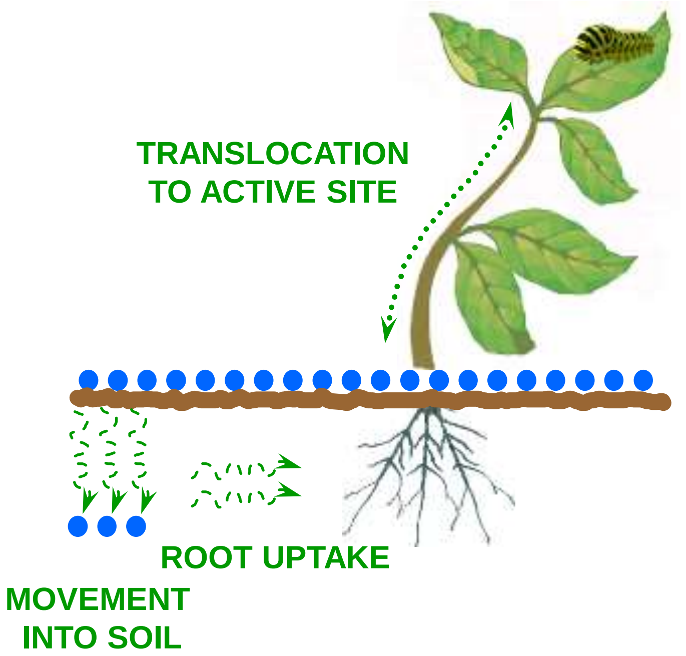


# Water Solubility

- **Water solubility** is a measure of how much an agrochemical will dissolve in water and is measured in ppm (parts per million)

- The higher the solubility the greater the availability for absorption by roots from the soil and translocation within the plant

- The lower the solubility, the lower the availability for absorption by roots from the soil and translocation within the plant
- Active ingredients with a low melting point, that are charged or contain high levels of polar functional groups tend to be more water soluble
- Environmental impact through leaching limits the use of highly soluble agrochemicals
- A balance is therefore required



# Dissociation Constant: p$K_a$

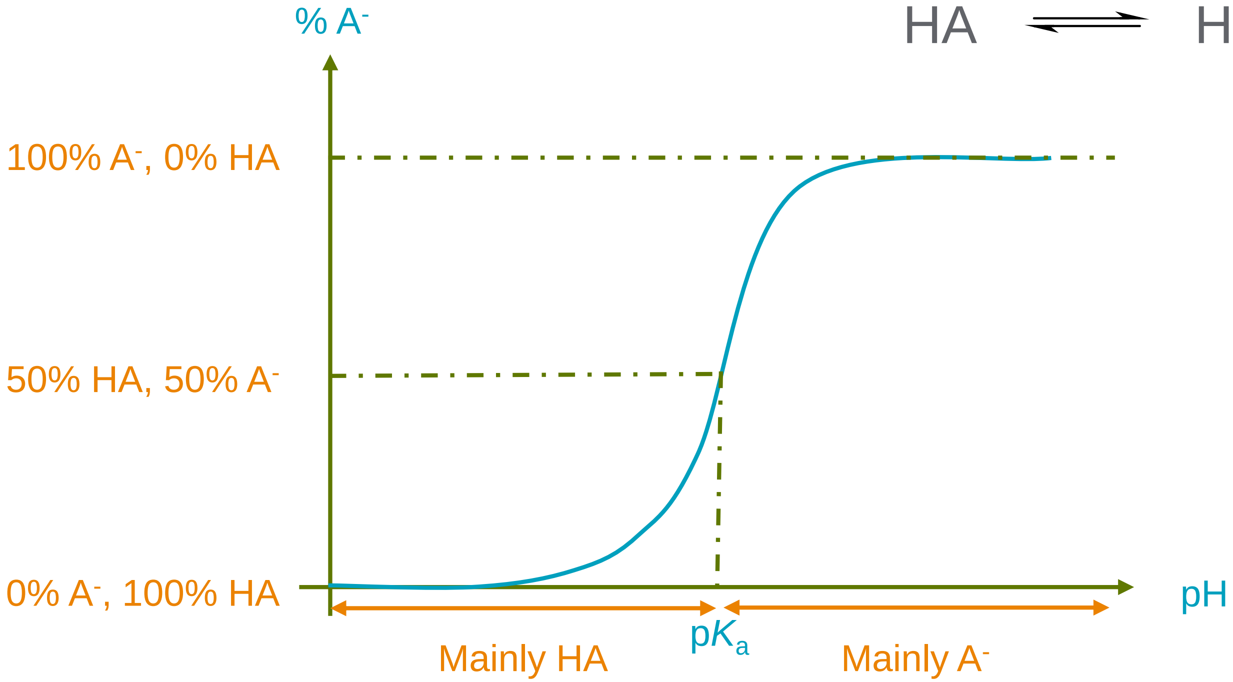
- $K_a$ is an equilibrium constant (**acid dissociation constant**) for the dissociation of an acid HA into its conjugate base $A^-$ and $H^+$ and is a quantitative measure of the strength of an acid in solution.

$$K_a = \frac{[A^-]\,[H^+]}{[HA]} \qquad\qquad HA \rightleftharpoons H^+ + A^-$$

- $K_a$ is often very high or low so is usually expressed as p$K_a$.
- p$K_a$ = -log$_{10}$ ($K_a$)
- It follows from the minus sign in the p$K_a$ definition that the lower the p$K_a$, the larger the equilibrium constant and therefore the stronger the acid in solution.

# p$K_a$ and pH

$$HA \rightleftharpoons H^+ + A^-$$



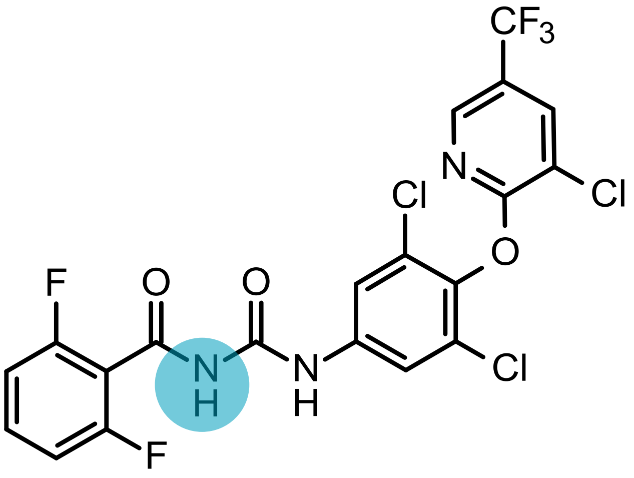
When [HA] = [A$^-$]:

$$K_a = \frac{[A^-][H^+]}{[HA]} = [H^+]$$

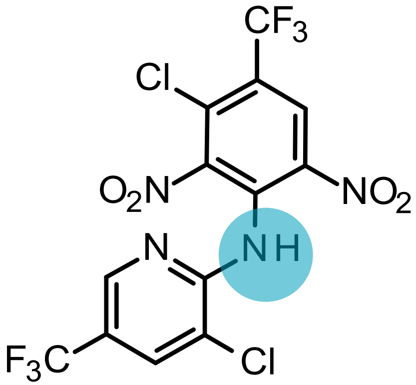
$$pK_a = -\log_{10}[H^+] = pH$$

- The p$K_a$ of the acid is the pH where it is exactly half dissociated.
- At pH above p$K_a$, the acid exists mainly as A$^-$.
- At pH below the p$K_a$, the acid exists mainly as HA.

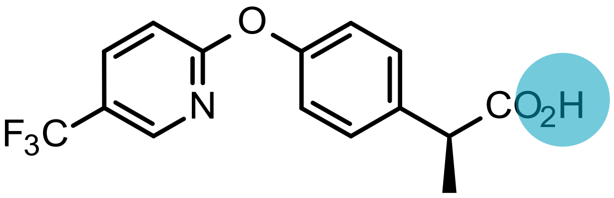
# p$K_a$ of some Herbicides



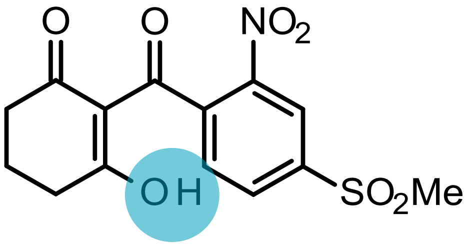
**Chlorfluazuron** p$K_a$ 8.10



**Fluazinam** p$K_a$ 6.83



**Fluazifop** p$K_a$ 3.20

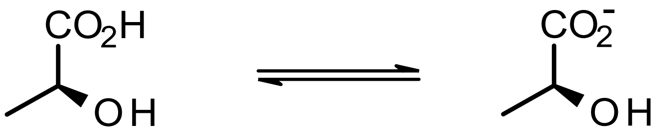


**Mesotrione** p$K_a$ 2.97

# Lipophilicity of Charged Species

- The p$K_a$ of a compound therefore determines which species (ionised or unionised) is likely to be present at a certain pH

- This has consequences on the lipophilicity of a species



More lipophilic

Less lipophilic

- Important point: **log P** refers to the partition coefficient of the **neutral** compound

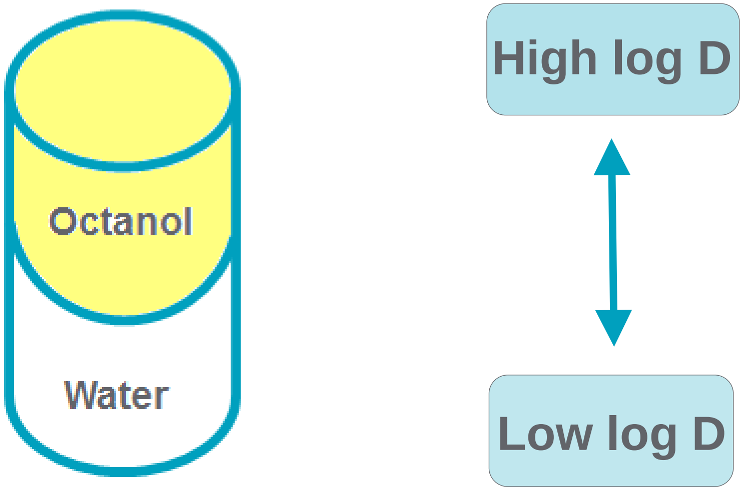
$$\log P = \log \left( \frac{[HA]_{oct}}{[HA]_{water}} \right)$$

- **Log D** must be used when referring to the partition coefficient of a charged species and must be quoted along with pH

# Distribution Coefficient: log D

- The **distribution coefficient** is the ratio of the sum of the concentration of all species of the compound (ionised plus unionised) in octanol to the sum of the concentration of all species of the compound (ionised plus unionised) in water.

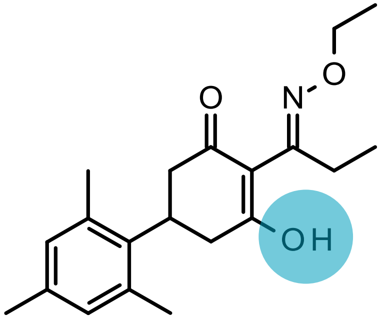
$$D = \frac{[AH]_{oct} + [A^-]_{oct}}{[AH]_{water} + [A^-]_{water}}$$



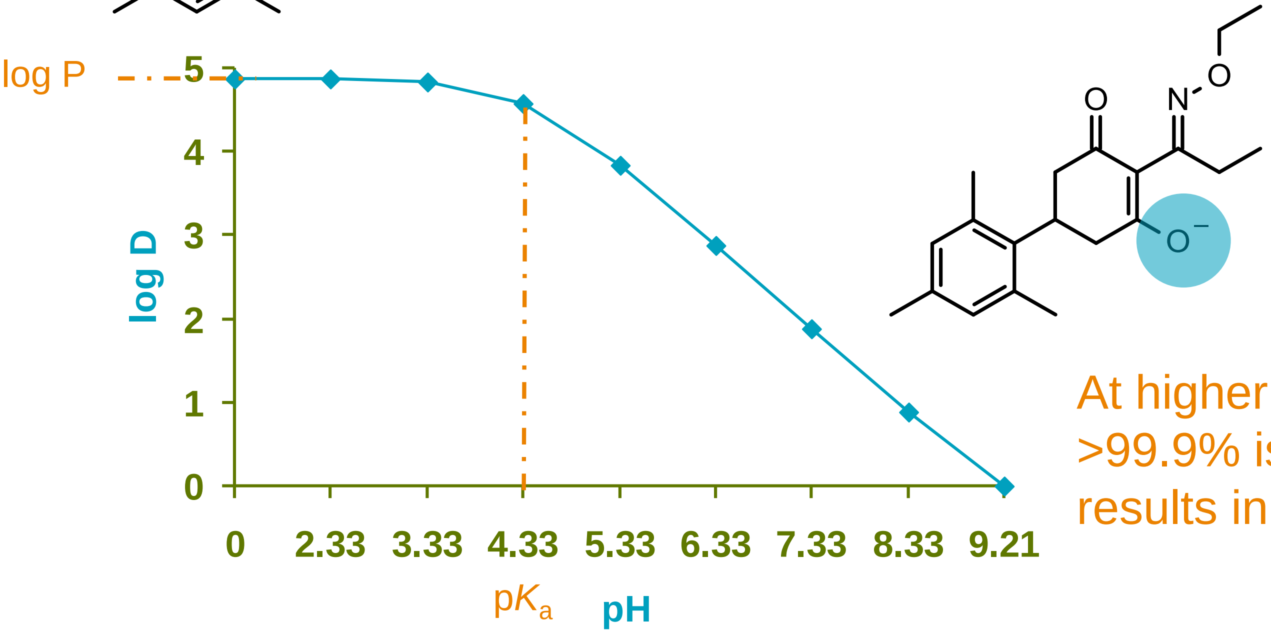
- Log D = $\log_{10}$ (D).
- High log D indicates high lipophilicity whereas low log D indicates low lipophilicity at a given pH.
- Log D is often quoted at pH 7 to give an indication of lipophilicity under physiological conditions (pH ~6-8).

# Tralkoxydim: p$K_a$ 4.3 & log P = 4.9

- Tralkoxydim is a member of the cyclohexanedione family of herbicides.

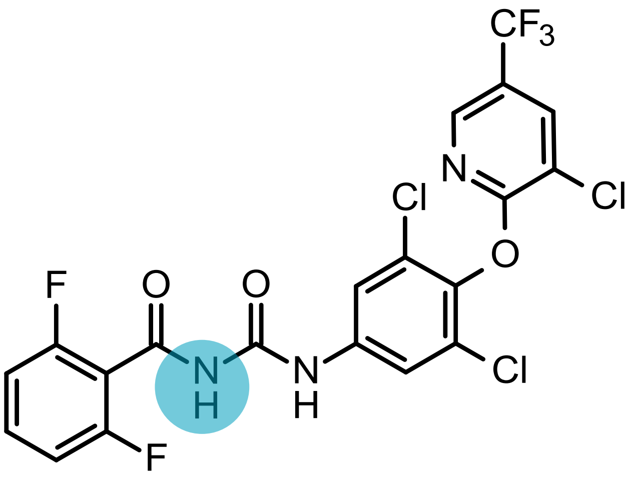


At low pH (acidic) <10% is deprotonated which results in a high log D value.



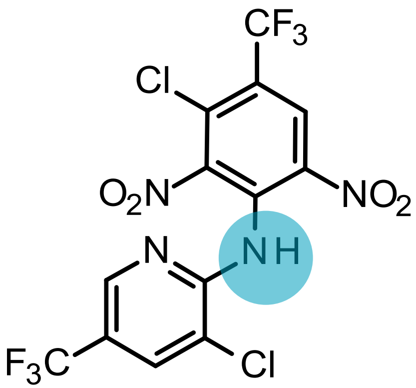
At higher pH (physiological pH ~6-8) >99.9% is deprotonated which results in a low log D value.

# logD$_{(pH7)}$ of some Herbicides



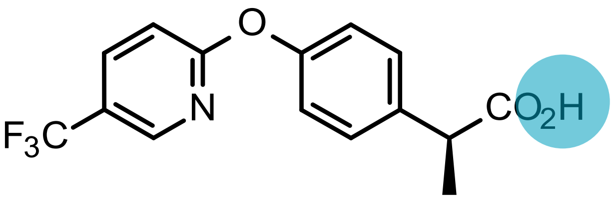
**Chlorfluazuron**

**logP 5.8, pK$_a$ 8.10; logD 5.77**



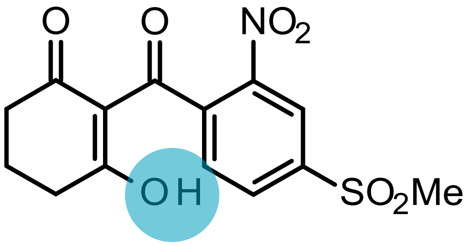
**Fluazinam**

**logP 4.03, pK$_a$ 6.83; logD 3.64**



**Fluazifop**

**logP 3.18, pK$_a$ 3.20; logD –0.62**

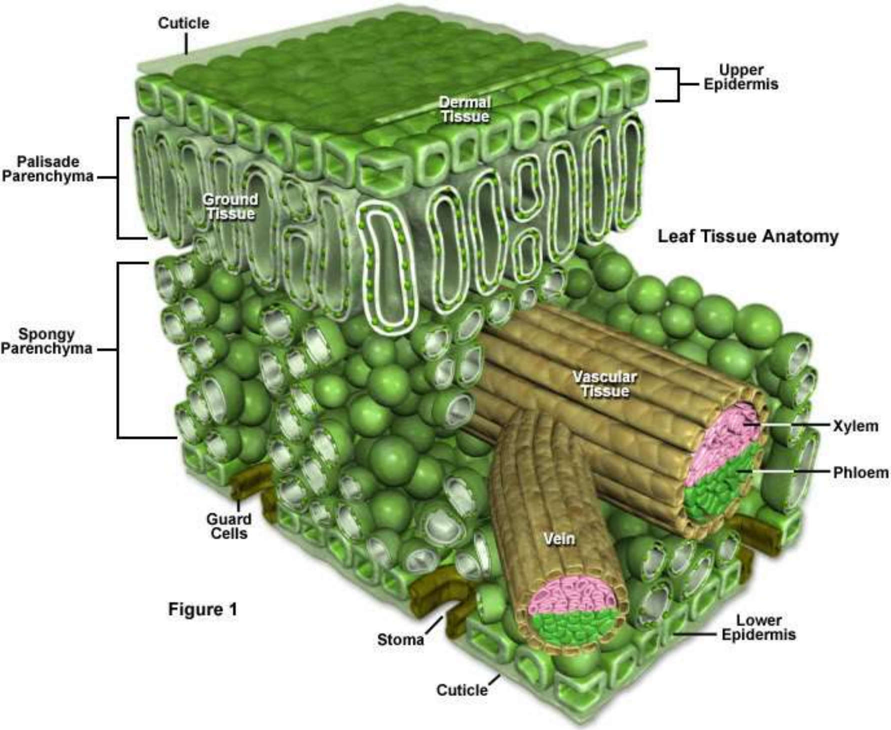


**Mesotrione**

**logP 3.12, pK$_a$ 2.97; logD – 0.91**

# AI localisation versus pest spectrum – Trapp model

- A simplified model represents the leaf as a series of compartments
- The compartments have differing properties e.g. size, pH, lipid content
- The localization of the active ingredient depends on it’s physical properties
- Anke Buchholz and Stefan Trapp; *Pest Man Sci, 2015*



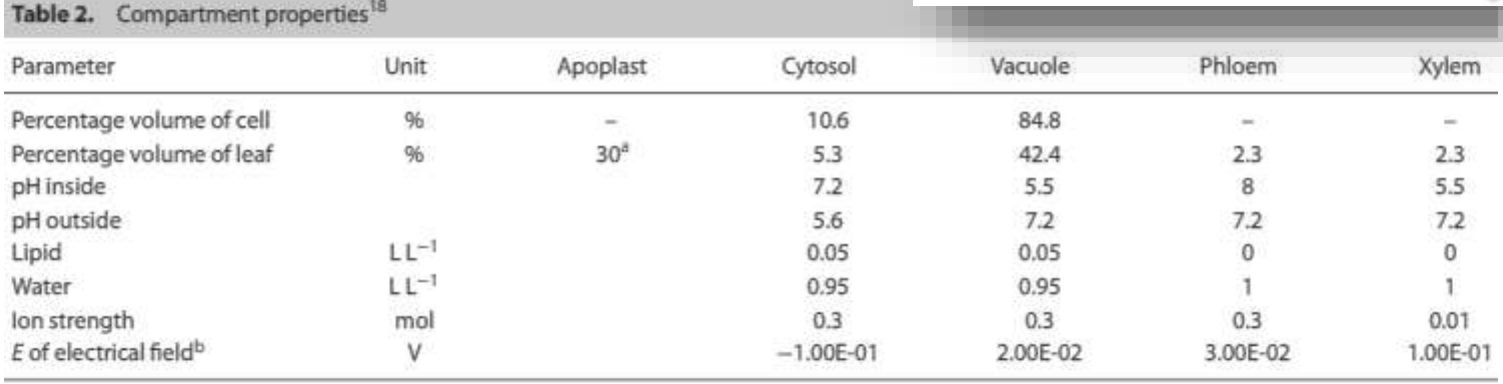
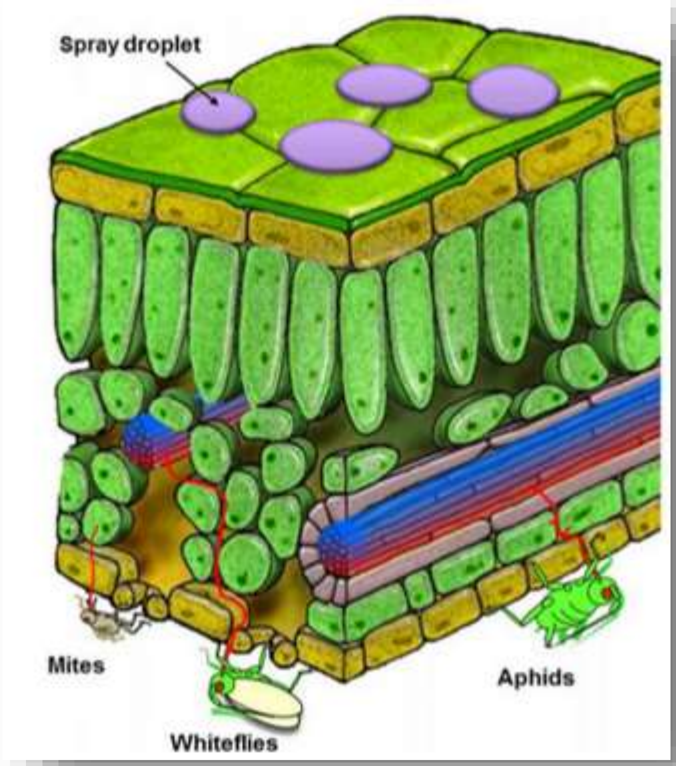


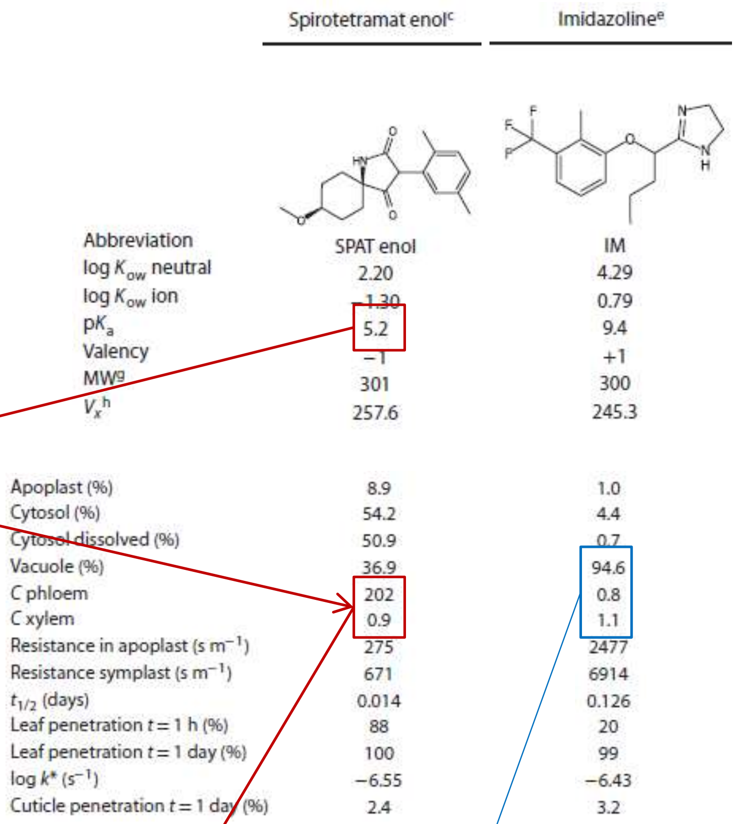
**Table 2.** Compartment properties[18]

| Parameter | Unit | Apoplast | Cytosol | Vacuole | Phloem | Xylem |
|---|---|---|---|---|---|---|
| Percentage volume of cell | % | – | 10.6 | 84.8 | – | – |
| Percentage volume of leaf | % | 30[a] | 5.3 | 42.4 | 2.3 | 2.3 |
| pH inside | | | 7.2 | 5.5 | 8 | 5.5 |
| pH outside | | | 5.6 | 7.2 | 7.2 | 7.2 |
| Lipid | $L\,L^{-1}$ | | 0.05 | 0.05 | 0 | 0 |
| Water | $L\,L^{-1}$ | | 0.95 | 0.95 | 1 | 1 |
| Ion strength | mol | | 0.3 | 0.3 | 0.3 | 0.01 |
| *E* of electrical field[b] | V | | −1.00E-01 | 2.00E-02 | 3.00E-02 | 1.00E-01 |

# AI localisation versus pest spectrum – Trapp model

- Different pests feed on different compartments of the plant
- Combined with knowledge of the physical properties of the AI, this allows prediction of pest spectrum





| | Spirotetramat enol$^c$ | Imidazoline$^e$ |
|---|---|---|
| Abbreviation | SPAT enol | IM |
| log $K_{ow}$ neutral | 2.20 | 4.29 |
| log $K_{ow}$ ion | −1.30 | 0.79 |
| p$K_a$ | 5.2 | 9.4 |
| Valency | −1 | +1 |
| MW$^g$ | 301 | 300 |
| $V_x$$^h$ | 257.6 | 245.3 |
| Apoplast (%) | 8.9 | 1.0 |
| Cytosol (%) | 54.2 | 4.4 |
| Cytosol dissolved (%) | 50.9 | 0.7 |
| Vacuole (%) | 36.9 | 94.6 |
| C phloem | 202 | 0.8 |
| C xylem | 0.9 | 1.1 |
| Resistance in apoplast (s m$^{-1}$) | 275 | 2477 |
| Resistance symplast (s m$^{-1}$) | 671 | 6914 |
| $t_{1/2}$ (days) | 0.014 | 0.126 |
| Leaf penetration $t = 1$ h (%) | 88 | 20 |
| Leaf penetration $t = 1$ day (%) | 100 | 99 |
| log $k^*$ (s$^{-1}$) | −6.55 | −6.43 |
| Cuticle penetration $t = 1$ day (%) | 2.4 | 3.2 |

**pH inside phloem = 8**

**Good activity on aphids and whiteflies**

**Good activity on mites**

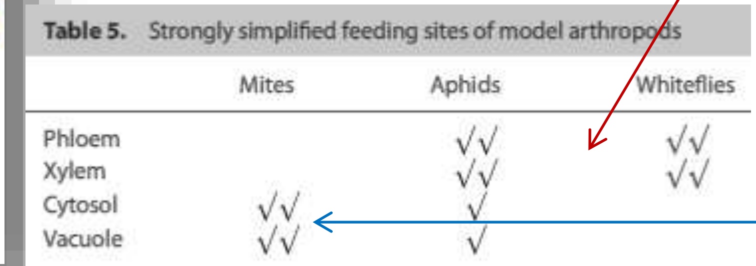
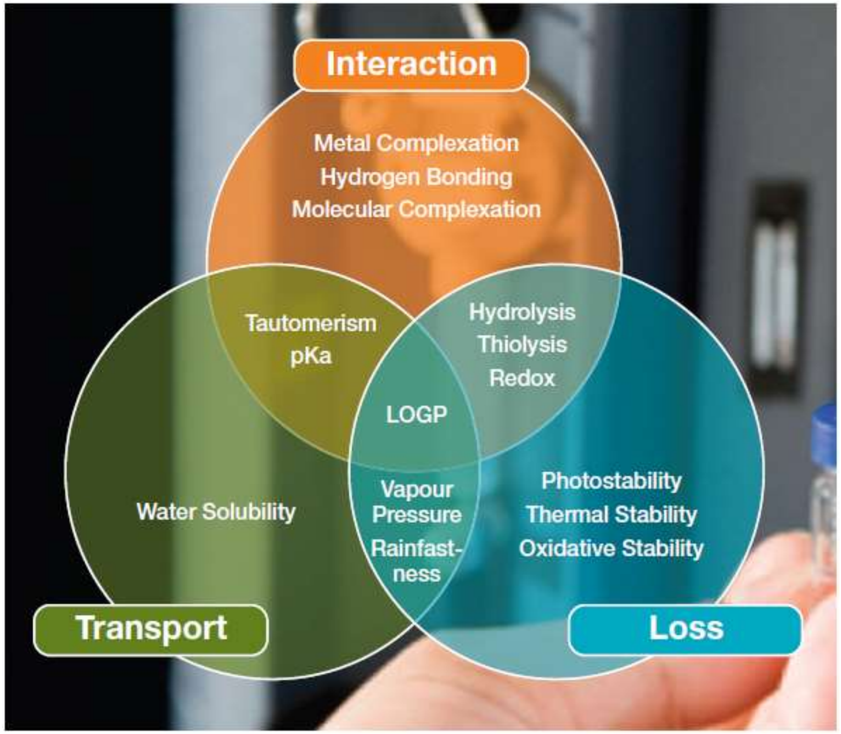


**Table 5.** Strongly simplified feeding sites of model arthropods

| | Mites | Aphids | Whiteflies |
|---|---|---|---|
| Phloem | | √√ | √√ |
| Xylem | | √√ | √√ |
| Cytosol | √√ | √ | |
| Vacuole | √√ | √ | |

# Physical Properties of Agrochemicals

- We have only scratched the surface of the processes that are involved in a molecule having the desired biological effect
- Having an understanding of the importance of physical properties to each of these processes is key to designing effective agrochemicals



# Agrochemical Lead Optimisation

**Taught Course**

Classification: PUBLIC

# The Route to Market

Lead Generation

Optimisation

Development

Registration

Commercial Pesticide

Number of compounds

140,000* → 5000 → 30 → 1-3 → 1 → 1 → 1

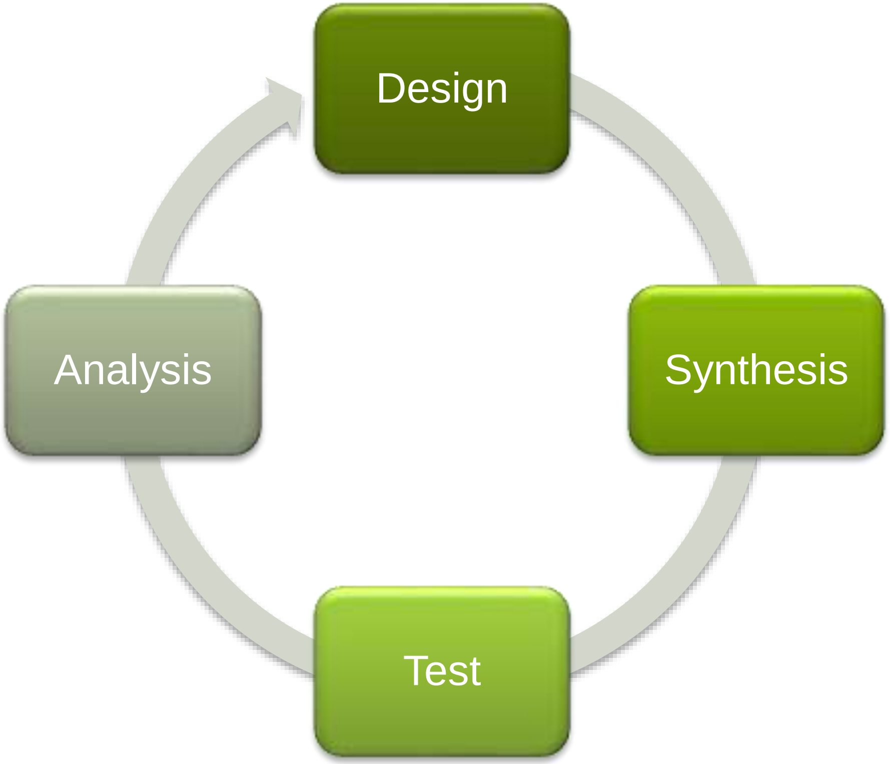
*average number of compounds synthesised to deliver one new market introduction in 2005

# Optimisation – The DSTA cycle

The new lead may not have:

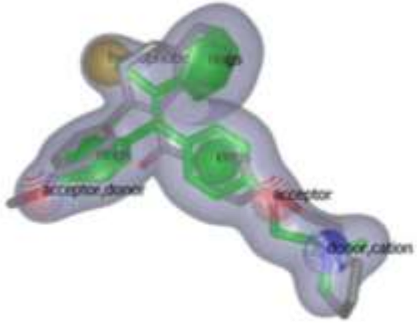
- The required level of **potency**
- Optimum **physical properties** for good bioavailability and environmental profile
- **Selectivity** for the desired target

- Optimisation is an iterative process
- Each DSTA cycle provides information about the **toxophore** and helps build the structure activity relationship (**SAR**)
- **Toxophore**: the minimum set of structural features that is recognised at a receptor site and is responsible for a molecule's intrinsic biological activity



# Optimisation – The DSTA cycle

**Design**



**Intuitive**

Relies on experience and prior knowledge

Fast – only need limited data

When does intuition become bias?

**Systematic**

e.g. Topliss set or Design of Experiment approaches

Robust data, but compounds may not be easily accessible

**Data driven**

Data analysis e.g. QSAR (see later) may direct chemistry

Relies on robust data

***In silico* modelling**

Modern computational techniques allow for rapid 3D modelling e.g. *in silico* docking or scaffold-hopping

Need structural information

# Intuitive Design – Bioisosteres

“Bioisosteres are groups or molecules which have chemical and physical similarities producing broadly similar biological properties.” – Thornber

“Compounds or groups that possess near-equal molecular shapes and volumes, approximately the same distribution of electrons, and which exhibit similar physicochemical properties...” – Burger

8 considerations:

- Size: molecular weight
- Shape: bond angles and hybridization states
- Electronic distribution: polarizability, inductive effects, charge, and dipoles
- Lipid solubility
- Water solubility
- pKa
- Chemical reactivity, including likelihood of metabolism
- Hydrogen bonding capacity

# Intuitive Design – Bioisosteres

**Bioisosteric replacements** may be targeted at one or more of the following:

- **Structure**
  - may be involved in maintaining or 'locking' a preferred conformation
  - size and bond angles will play a key role
  - scaffold-hopping is a related concept (see later)

- **Receptor Interactions**
  - may be involved in mimicking or enhancing existing molecular interactions
  - size, shape, electronic properties and H-bond acceptor/donor properties may be important

- **Biokinetics**
  - may be involved in enhancing uptake, distribution or soil properties
  - lipophilicity, hydrophilicity, p$K_a$, hydrogen-bonding all important

- **Metabolism**
  - may be involved in blocking or enhancing metabolism
  - chemical reactivity may be key

# Bioisosteres – Classical examples

**Classical bioisosteres** have the same valency or are ring equivalents, for example

**1) Monovalent atoms or groups**

D and H
F and H
$NH_2$ and OH
RSH and ROH
F, OH, $NH_2$ and $CH_3$
Cl, Br, SH and OH

**2) Divalent atoms or groups**

C=C, C=N, C=O, C=S
$-CH_2-$, -NH-, -O- , -S-

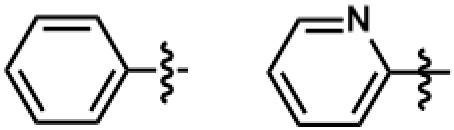
**3) Trivalent atoms or groups**

-CH=, -N=

**4) Tetrasubstituted atoms**

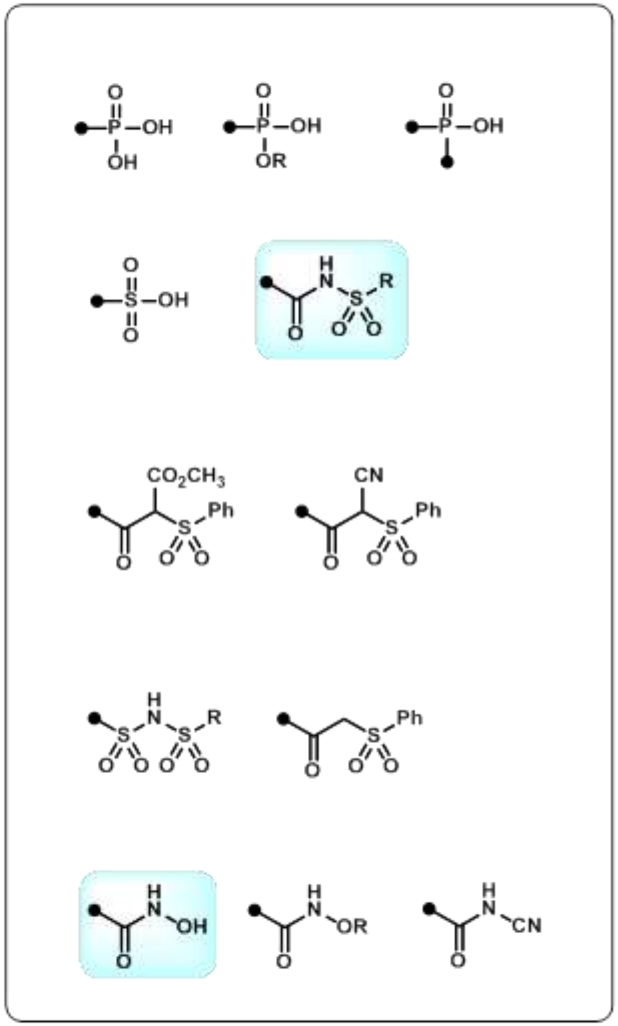
$R_4C$, $R_4Si$, $R_4N^+$

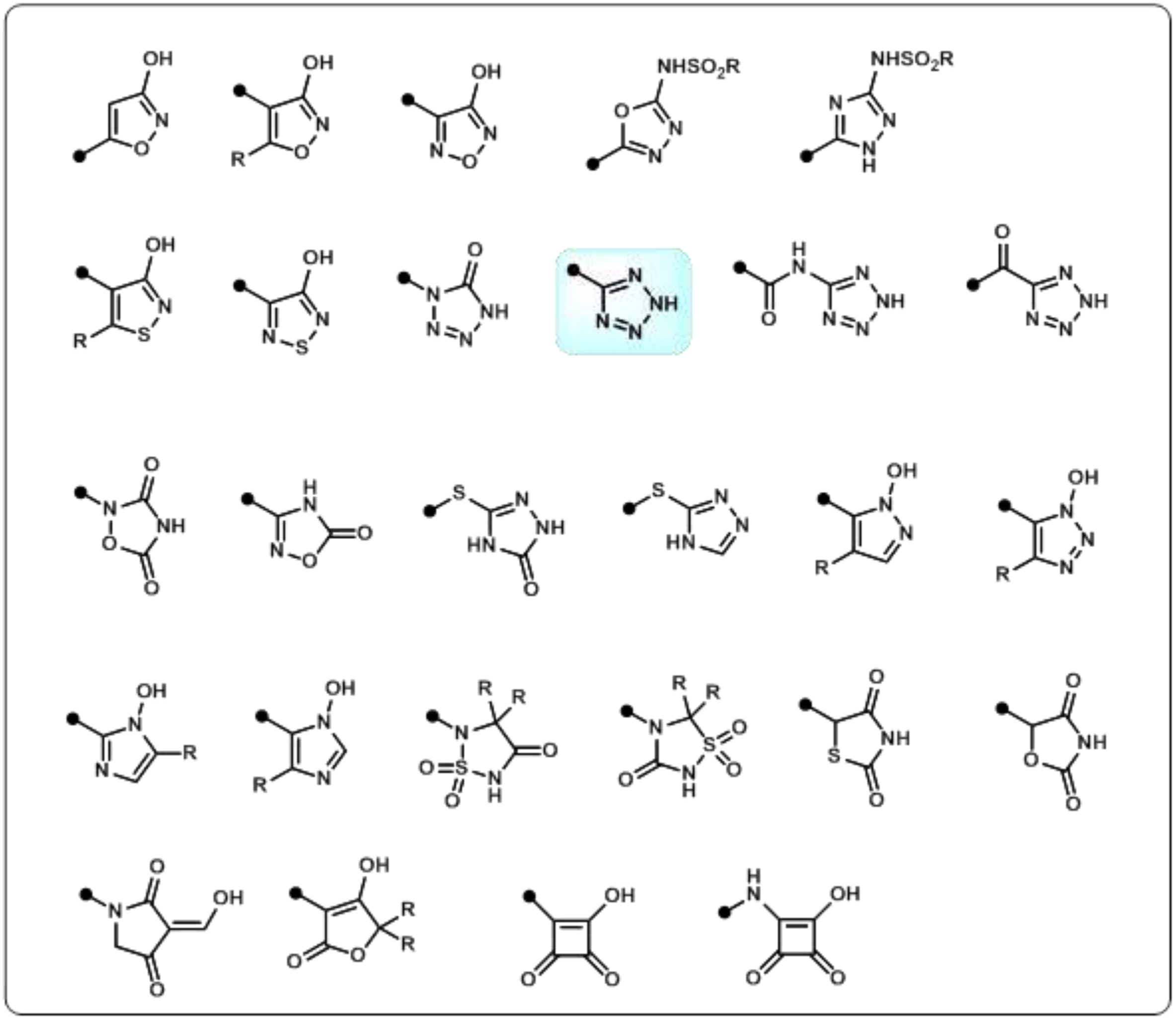
**5) Ring equivalent**



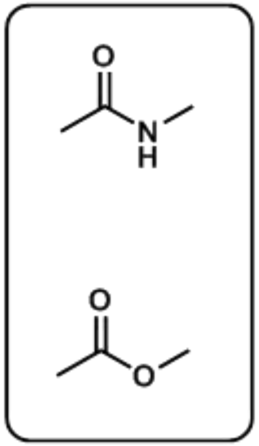
**Non-classical bioisosteres** are generally structurally distinct…

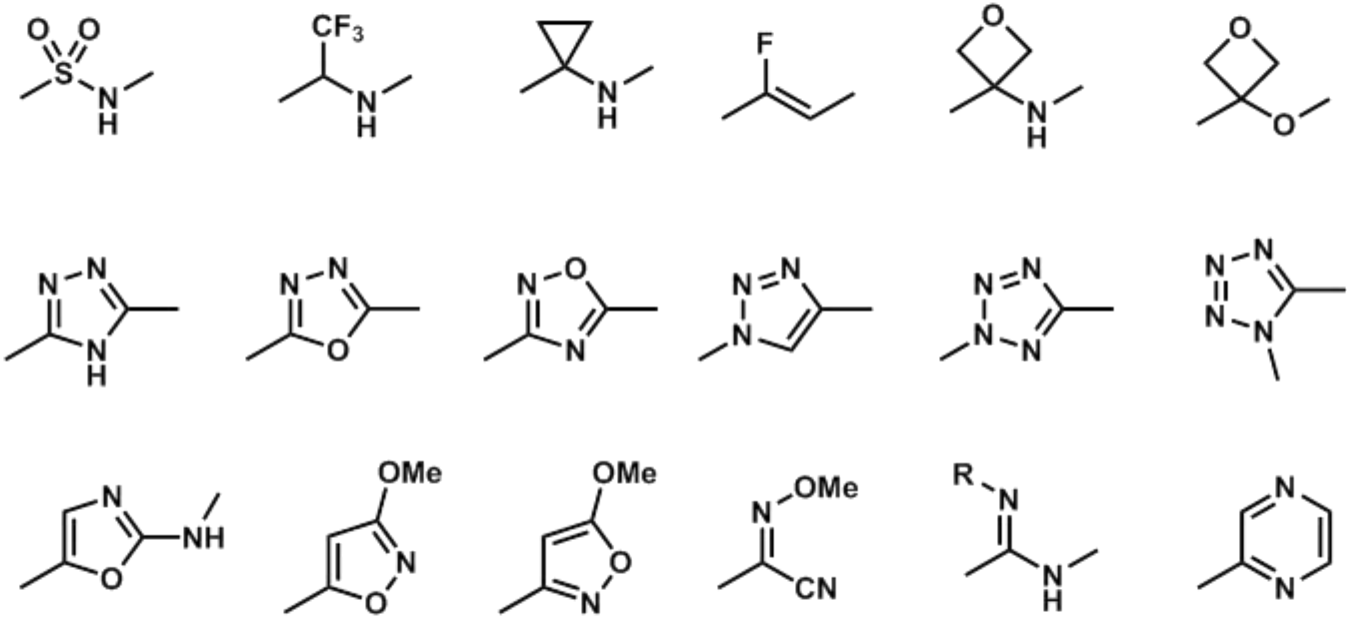
# Carboxylic acid bioisosteres





# Amide and ester bioisosteres

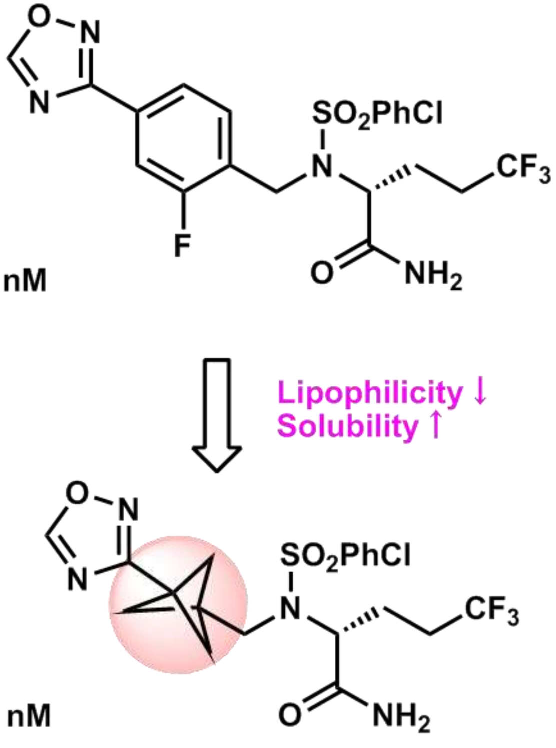




# Phenyl ring bioisosteres

Phenyl rings are often replaced by heterocycles to improve biokinetics (and often to reduce metabolism) however some more unusual replacements can also be made (*J Med Chem*, **2012**, *55*, *3414* – Pfizer)

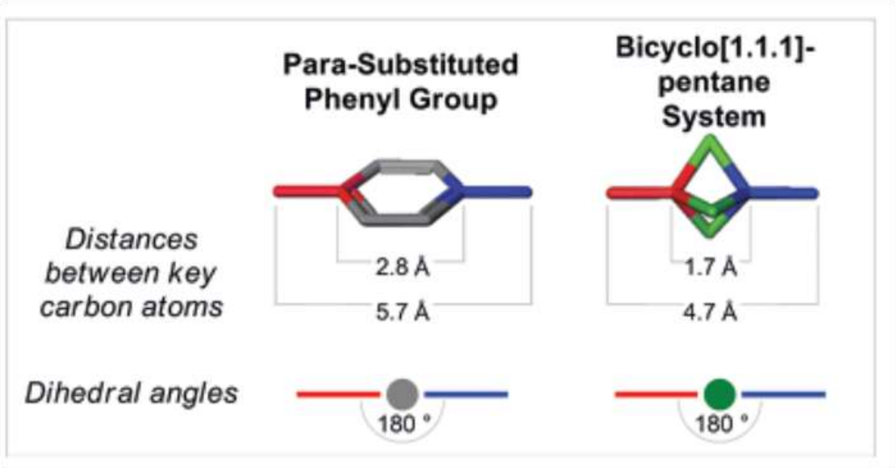
**γ-Secretase inhibitor** (Alzheimer's disease)



$IC_{50}$ = 0.225 nM
LogD = 4.7

Lipophilicity ↓
Solubility ↑

$IC_{50}$ = 0.178 nM
LogD = 3.8

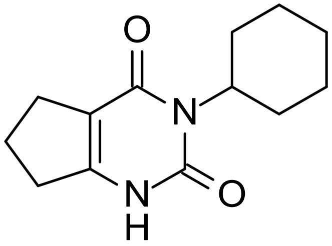


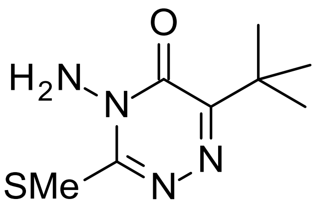
Bicyclo[1.1.1]pentane and Ph ring include comparable dihedral angles and similar distances between substituents

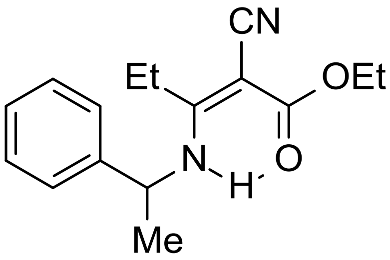
# *In Silico* Design - Scaffold-hopping

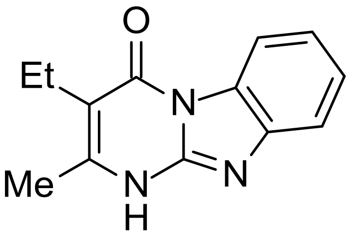
**Scaffold hopping** is the discovery of "structurally novel compounds starting from known active compounds by modifying the central core structure of the molecule" – Stahl *et al*, *Drug Discovery Today: Technologies*, **2004**, 217

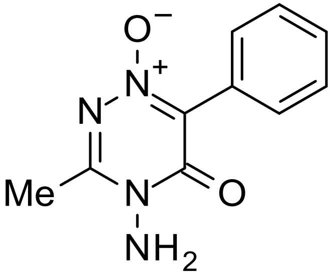
For example, the **azinone herbicides** all contain different heterocyclic cores however molecular modelling reveals that they have similar electrostatic profiles…

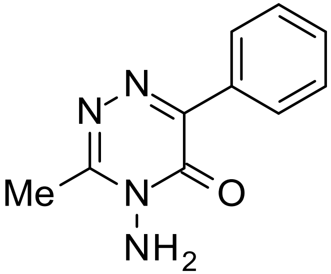

Lenacil


Metribuzin

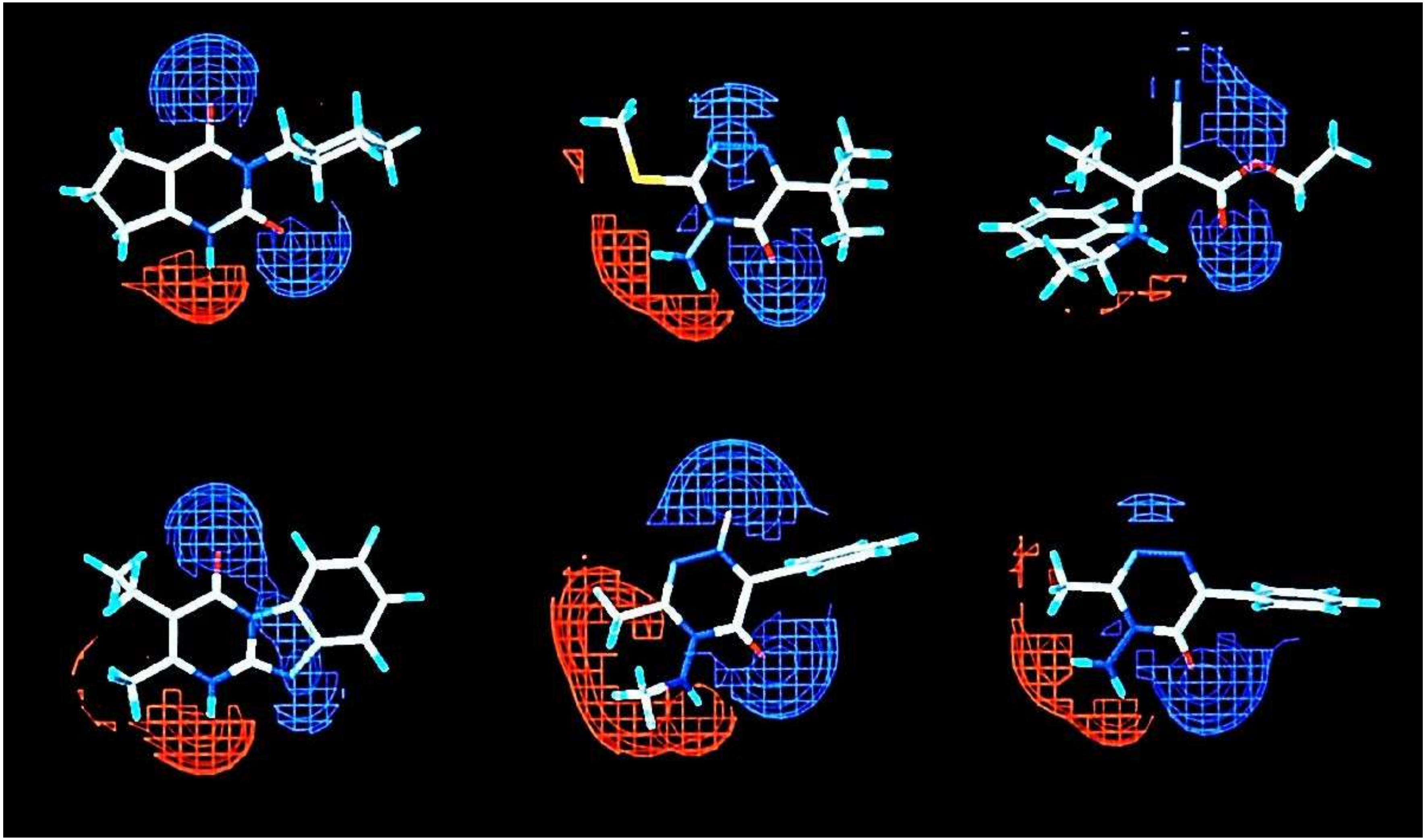

Cyanoacrylate


ALG6948


DRW3871


Metamitron

# Molecular Modelling – Electrostatics/Activity Correlation

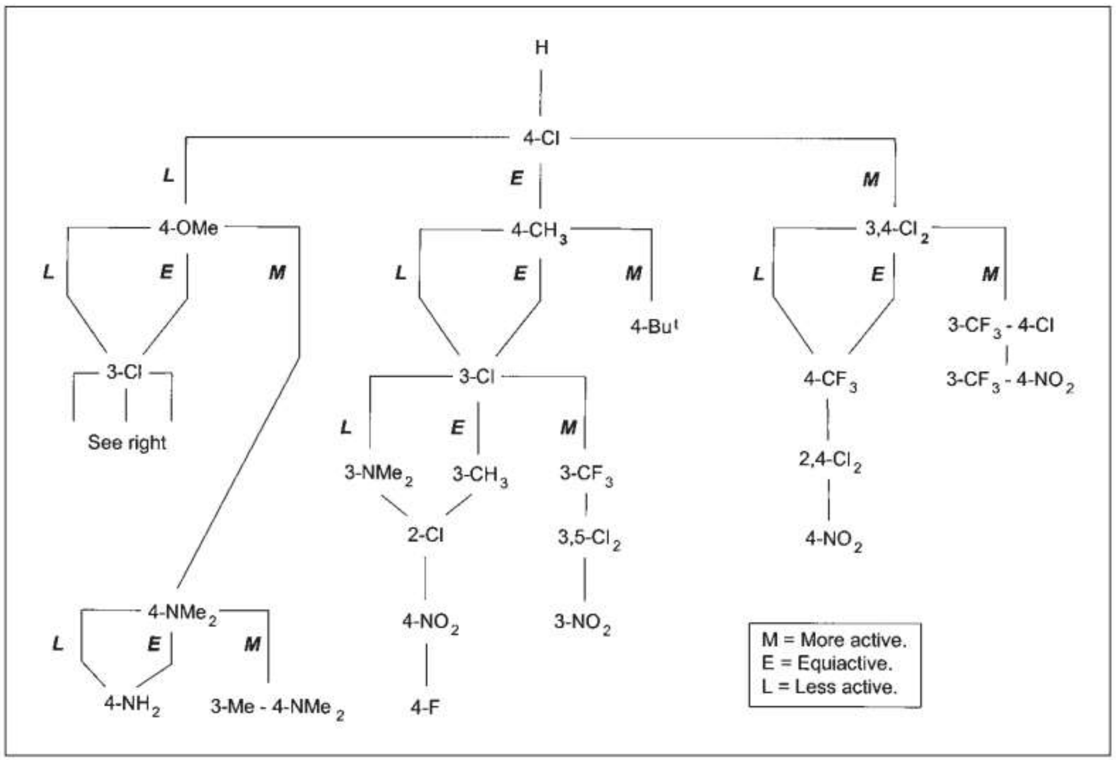


# Systematic Design – Topliss Trees

By analysis of structure-activity relationships (see later), some general rules can be derived for the optimisation of lipophilic, steric or electronic properties

In the 1970’s, Professor J G Topliss formalised some of these rules leading to the **Topliss Decision Trees** that can be used today (link)

Figure 1: Topliss scheme for aromatic substituents



By following the decision tree, more potent analogues should be achieved

In addition, information about the properties required for activity will be gained and can inform further systematic and intuitive design

An alternative approach is to use statistical methods such as Design of Experiments (link)

# Optimising for Selectivity

- **Crop selectivity** may arise from:
  1. Differences at the active site between species
  2. Selective metabolism of the agrochemical by the crop
  3. Reduced uptake or distribution of an agrochemical by the crop

1. Selectivity due to **differences at the active site** should be deduced from the structure activity relationship of the analogue series. Changes in the compounds shape and electronic distribution (e.g. by use of bioisosteres) can capitalise on this difference.

2. **Differing rates of metabolism** - Herbicide selectivity can arise when the crop metabolises the active compound at a faster rate than the weeds, resulting in a product that is safe to the crop

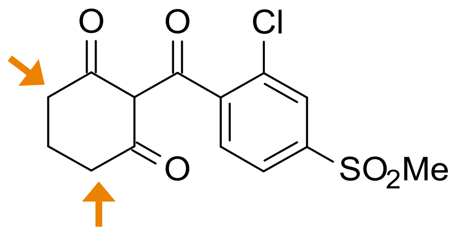
   - The introduction of a metabolically susceptible functional group can induce selectivity, e.g. sulfides, benzyl groups, olefins and electrophilic groups.

# Optimising for Selectivity

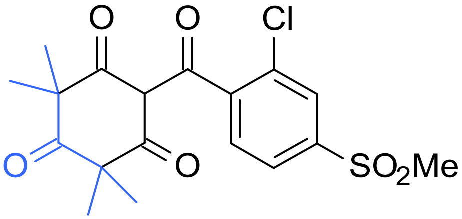
## 2. Differing rates of metabolism (cont.)

- A more selective compound may be less potent if it is de-activated more easily
- e.g. **Sulcotrione** is a herbicide product, but it was not the most active compound made during the optimisation process

*Metabolism Occurs*



*Metabolism Blocked*

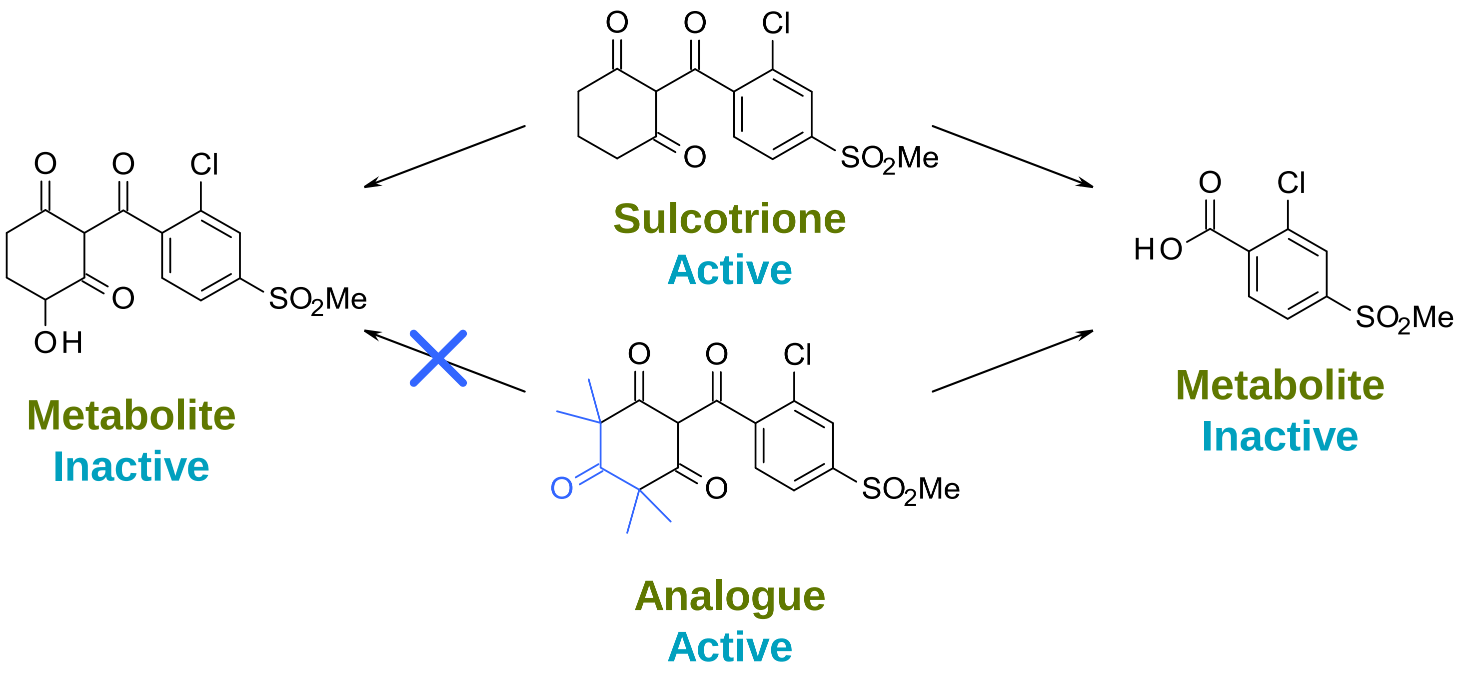


| Sulcotrione | Analogue |
|---|---|
| Active ($ED_{50}$ = 143 (g/ha))* | Highly Active ($ED_{50}$ = 21 (g/ha))* |
| Selective | Non-selective |

- Introducing functionality on the dione enhances grass activity, however this is accompanied by loss of selectivity and increased soil persistence

# Optimising for Selectivity

- The methyl substituents block metabolic de-activation in the plant:



**Sulcotrione**
**Active**

**Metabolite**
**Inactive**

**Metabolite**
**Inactive**

**Analogue**
**Active**

- Some selectivity for grass weeds over grass crops can be achieved by Sulcotrione, however it is more effectively used to control broad leaf weeds

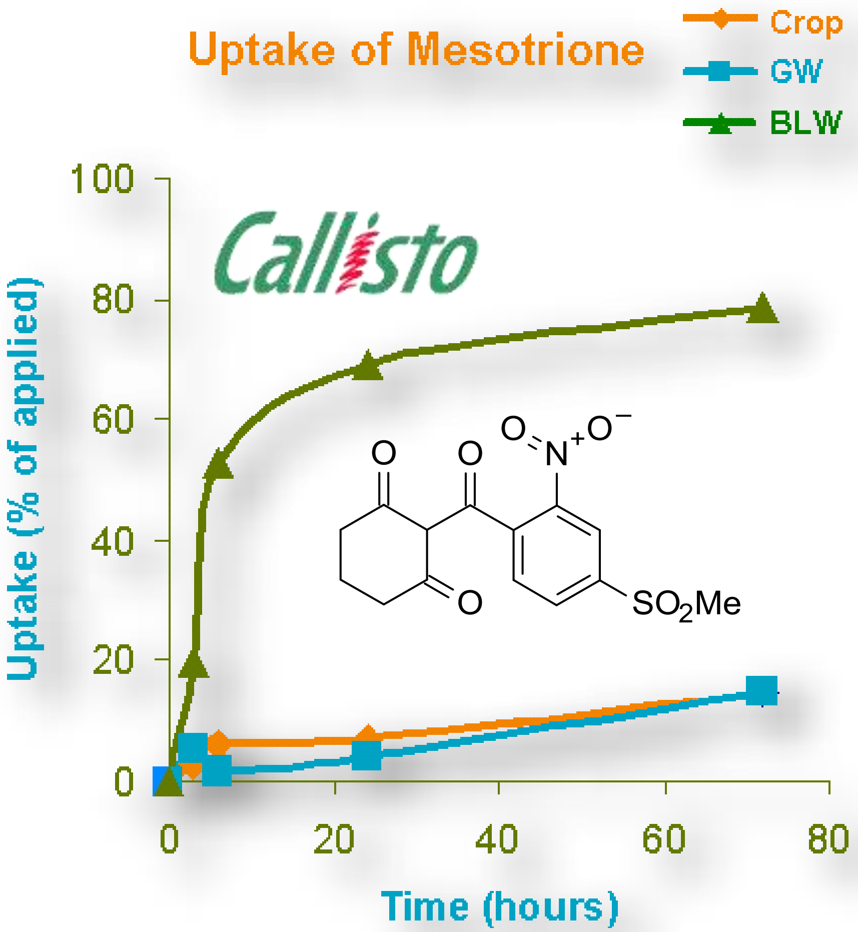
# Optimising for Selectivity

3. Selectivity may also arise from a **difference in uptake or distribution** of the compound, therefore introducing functionality that affects the log P (important for absorption) and water solubility (important for root uptake and systemicity) can induce selectivity

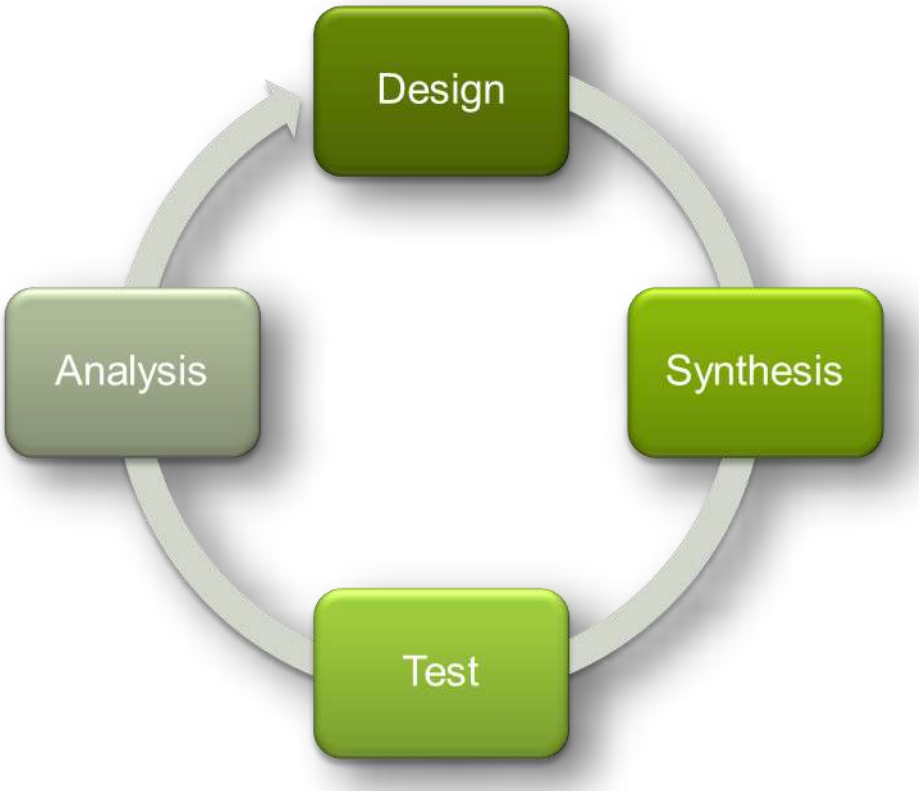
- Example: Mesotrione is used to control broad leaf weeds in grass crops.
- The difference in uptake of a broad leaf plant (in this case the weed) and a grass plant (crop) allows selectivity to be achieved.

**Mesotrione:**

**Maize (Crop) selective**
**Good Broad leaf weed (BLW) control**
**Poor grass weed (GW) control**



# The DSTA cycle



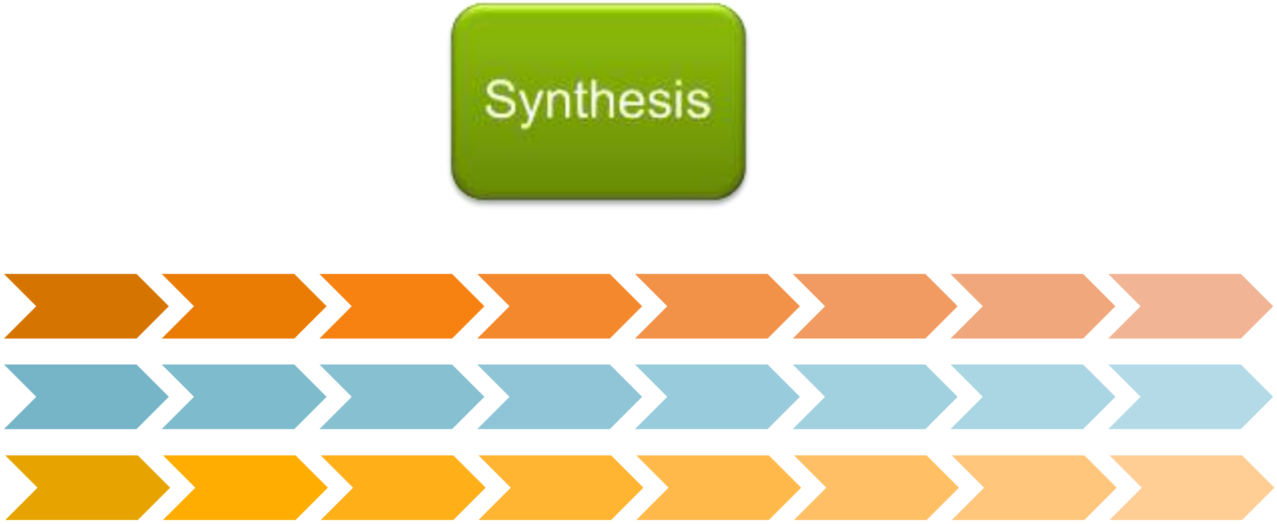
Design

Synthesis

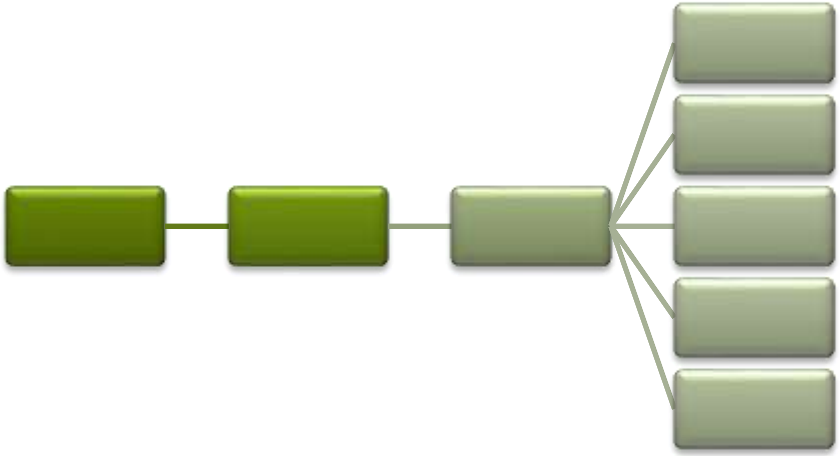
Test

Analysis

# Optimisation – The DSTA cycle



Linear ✗

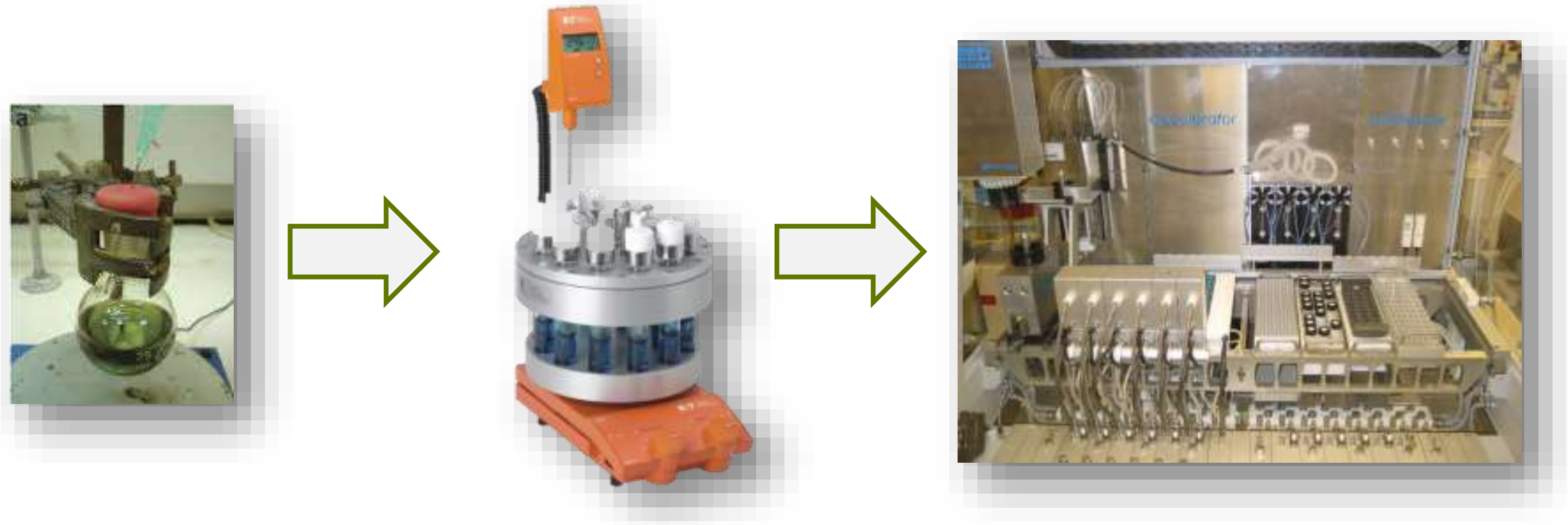


Branched ✓

Where possible, late-stage functionalization from a common intermediate is preferred

# Optimisation – The DSTA cycle

Synthesis



- Method of synthesis depends on the quantity and complexity of the targets, and should take into account efficient design and availability of reagents and building blocks
- In the discovery phase, speed is normally prioritised over yield

# Optimisation – The DSTA cycle

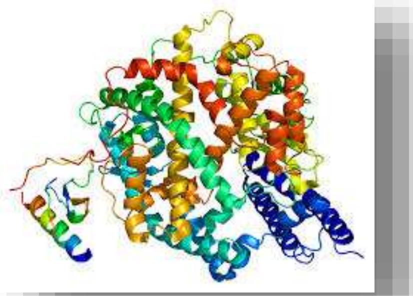
Test

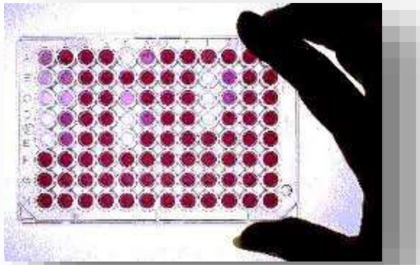
***In vitro* testing**:

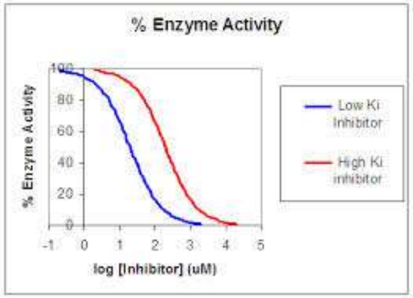
- Gives a measure of intrinsic activity against the enzyme

but…

- You don’t always know the enzyme involved and it may not always be amenable to a functional *in vitro* assay
- Translating into *in vivo* activity can be difficult







# Optimisation – The DSTA cycle

Test

***In vivo* testing**:

- Gives a realistic measure of activity on the whole organism

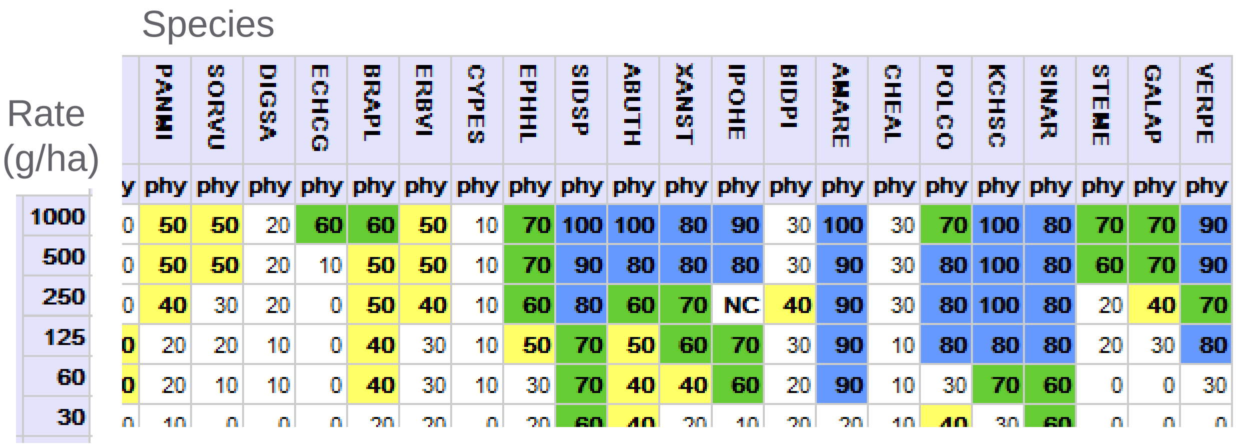
but…

- Much more complicated system – many factors may lead to a compound being inactive

# Optimisation – The DSTA cycle

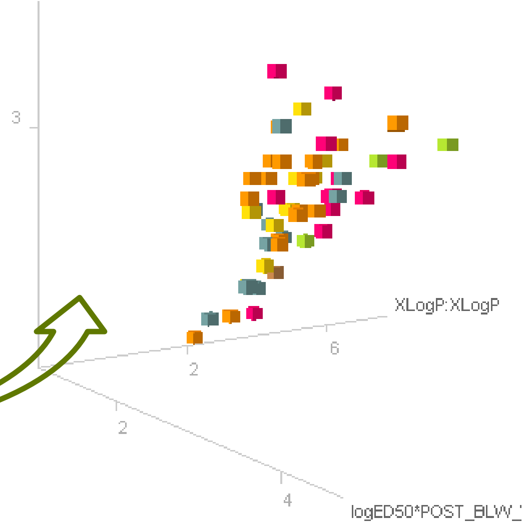
Analysis

## Biology results:



Species

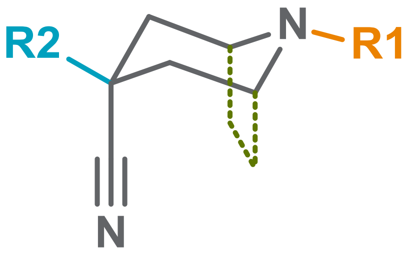
| Rate (g/ha) | | PANMI | SORVU | DIGSA | ECHCG | BRAPL | ERBVI | CYPES | EPHHL | SIDSP | ABUTH | XANST | IPOHE | BIDPI | AMARE | CHEAL | POLCO | KCHSC | SINAR | STEME | GALAP | VERPE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | y | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy | phy |
| 1000 | 0 | 50 | 50 | 20 | 60 | 60 | 50 | 10 | 70 | 100 | 100 | 80 | 90 | 30 | 100 | 30 | 70 | 100 | 80 | 70 | 70 | 90 |
| 500 | 0 | 50 | 50 | 20 | 10 | 50 | 50 | 10 | 70 | 90 | 80 | 80 | 80 | 30 | 90 | 30 | 80 | 100 | 80 | 60 | 70 | 90 |
| 250 | 0 | 40 | 30 | 20 | 0 | 50 | 40 | 10 | 60 | 80 | 60 | 70 | NC | 40 | 90 | 30 | 80 | 100 | 80 | 20 | 40 | 70 |
| 125 | 0 | 20 | 20 | 10 | 0 | 40 | 30 | 10 | 50 | 70 | 50 | 60 | 70 | 30 | 90 | 10 | 80 | 80 | 80 | 20 | 30 | 80 |
| 60 | 0 | 20 | 10 | 10 | 0 | 40 | 30 | 10 | 30 | 70 | 40 | 40 | 60 | 20 | 90 | 10 | 30 | 70 | 60 | 0 | 0 | 30 |
| 30 | 0 | 10 | 0 | 0 | 0 | 20 | 20 | 0 | 20 | 60 | 40 | 20 | 10 | 20 | 20 | 10 | 40 | 30 | 60 | 0 | 0 | 0 |



## Data analysis

# QSAR (Quantitative Structure Activity Relationship)

- **Free Wilson QSAR Analysis** can be used to look for trends in data to identify important factors, it relates structural features to biological properties
- In Free Wilson analysis a descriptor matrix is assembled where 1 indicates the presence and 0 the absence of a particular substituent. The molecule and substituent constants are then fitted by linear regression
- For example, the activity of **cyanopiperidines** (nicotinic acetylcholine receptor agonists) as insecticides was correlated to the 4 component parts of the molecule.



*Cyanopiperidine*

| Compound | Activity | R1=H | R1=Me | R1=Bn | R2=Ph | R2=3Py | Bridge |
|---|---|---|---|---|---|---|---|
| **1** | 0.4 | 1 | 0 | 0 | 1 | 0 | 1 |
| **2** | 2.4 | 0 | 0 | 1 | 1 | 0 | 0 |
| - - - | - | - | - | - | - | - | - |

***Linear Regression***

**Activity = Molecule Constant + Substituent Constant for R1 + Substituent Constant for R2 + Substituent Constant for Bridge**

# QSAR

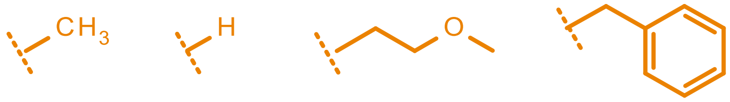
- Results for cyanopiperidines on aphids:

**Activity*** = **Molecule Constant** + **Substituent Constant for R1** + **Substituent Constant for R2** + **Substituent Constant for Bridge**

Molecule Constant = 1.55

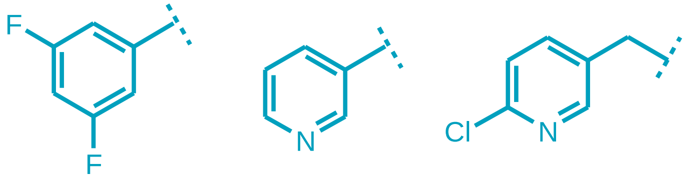
Constant for Bridge = -1.55

Substituent Constants for R1 =



**+0.0** **-0.18** **-0.26** **-0.04**

Substituent Constants for R2 =

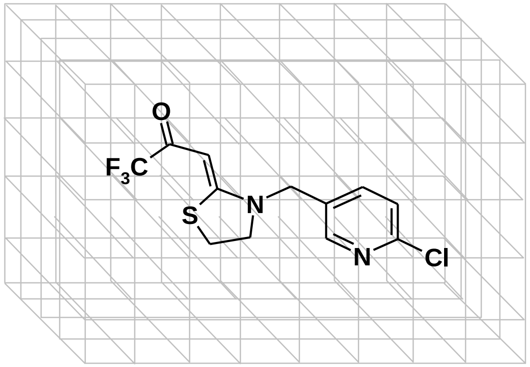


**+0.0** **-0.85** **+1.44**

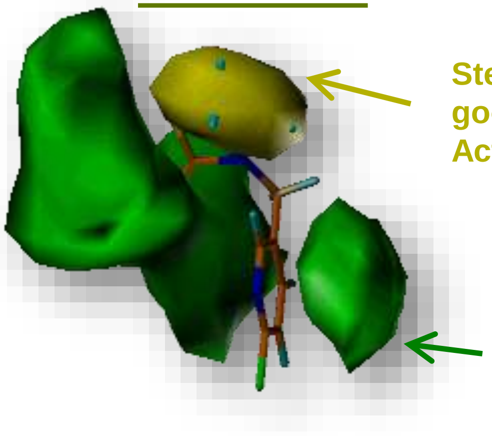
- Free-Wilson assumes *additive* QSAR model for activity - this is true if each substituent has an independent effect
- Easily understood and used by chemists, but predictions are limited to substituents in the training set.
- The model is validated by plotting predicted activity *vs* actual activity.
- This technique gives an overview of the SAR in a series.

# QSAR

- **CoMFA** – Comparative Molecular Field Analysis is a modern QSAR technique which relates steric or electrostatic fields to biological activity.
- All molecules are minimised, overlaid and embedded in a 3D grid of points:



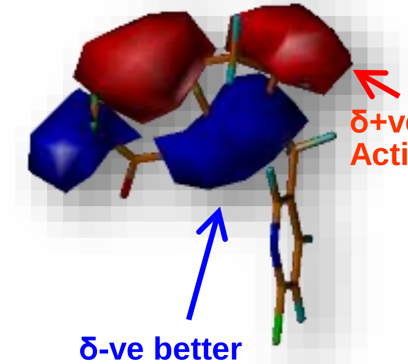
**Steric Field**



Steric Bulk good for Activity

Steric Bulk detrimental to Activity

**Electrostatic Field**



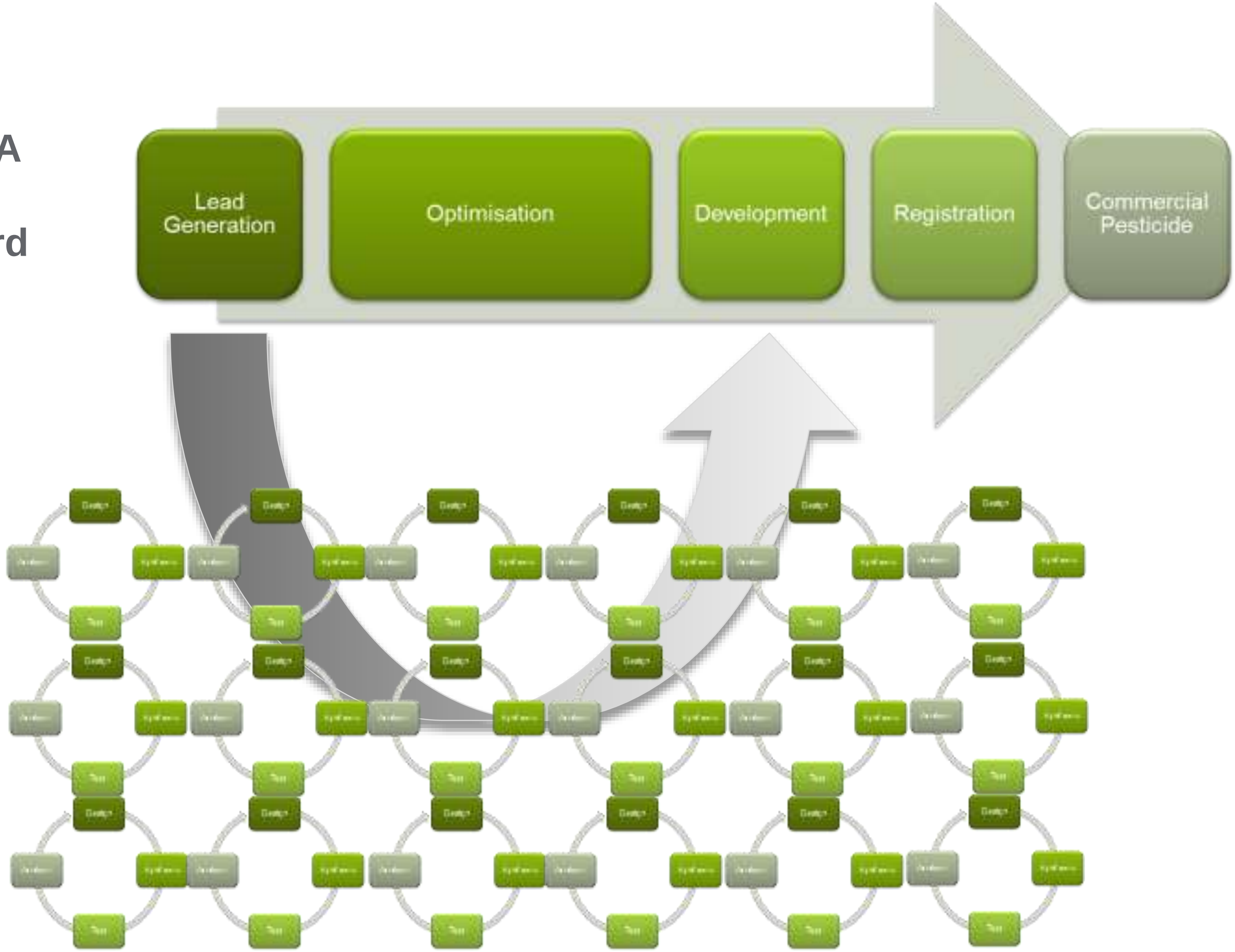
δ+ve better for Activity

δ-ve better for Activity

- For each molecule the electrostatic and steric fields are correlated at each grid point with biological activity. It indicates which regions allow bulky substituents and retain activity and which prefer electronegative or positive groups.
- This is used to qualitatively guide new analogue design and quantitatively predict activity of new analogues.

# Hit to Lead

**It takes many many DSTA cycles to find the ideal compound to take forward into development**

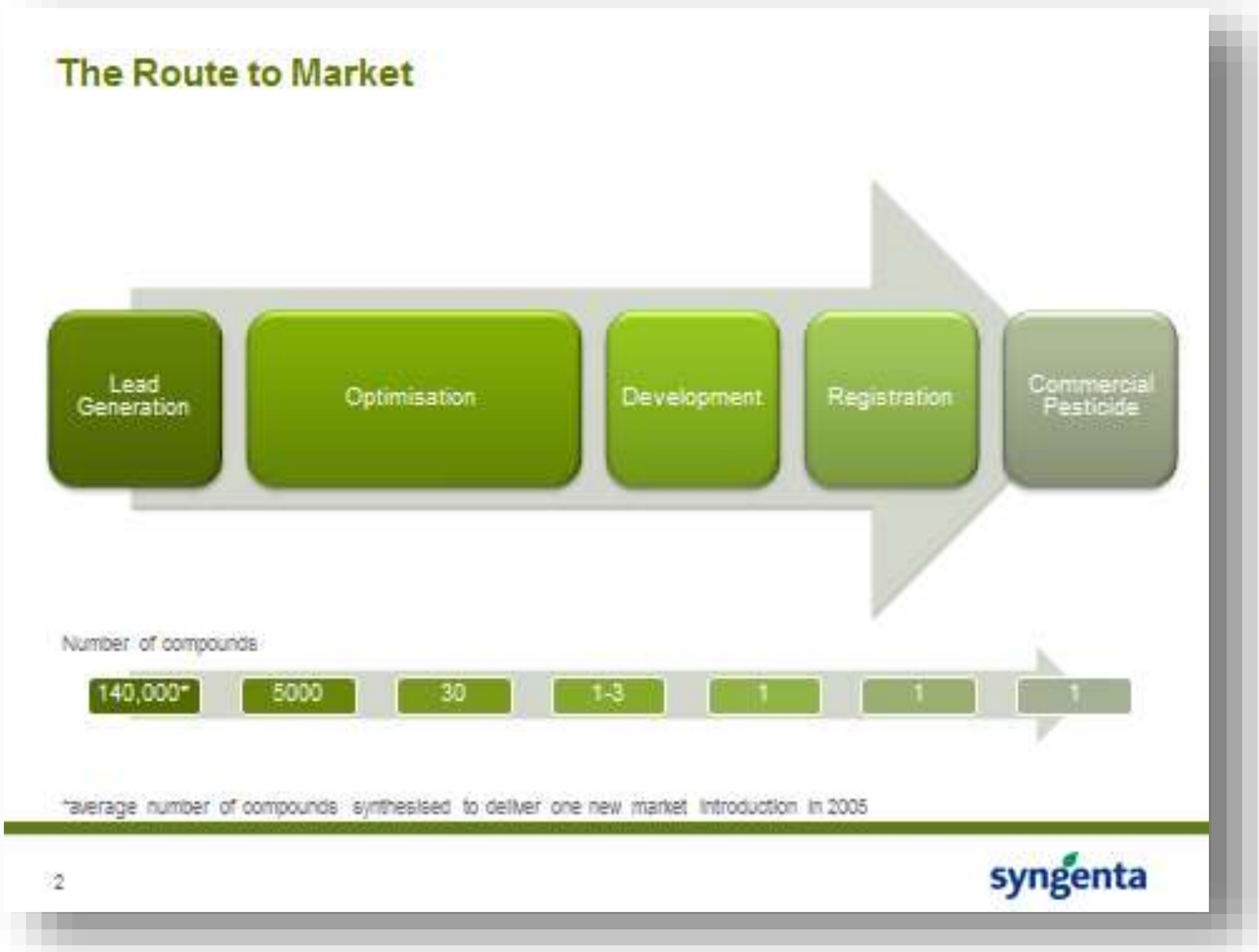


# Compound evaluation and development

**Taught Course**

Classification: PUBLIC

# Development candidates



The Route to Market

Lead Generation

Optimisation

Development

Registration

Commercial Pesticide

Number of compounds

140,000*

5000

30

1-3

1

1

1

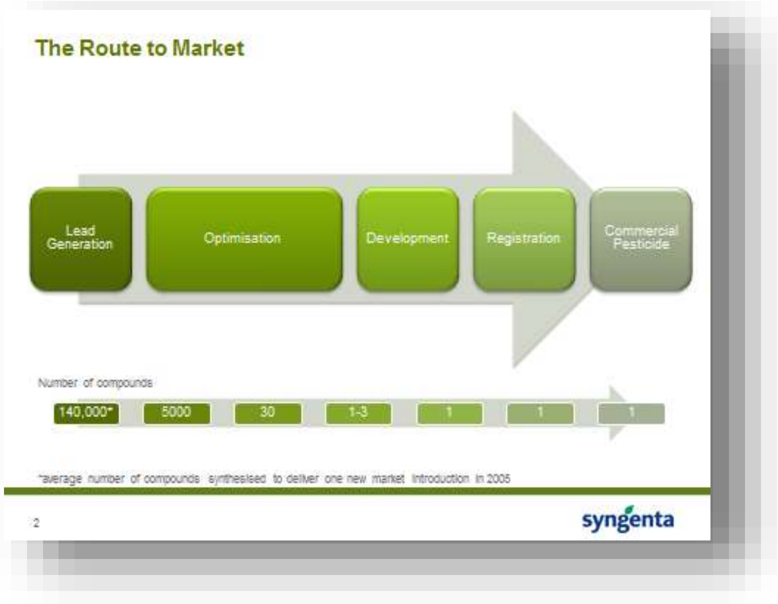
*average number of compounds synthesised to deliver one new market introduction in 2005

2

# Development candidates

## 5 Key Questions:

- *Does it work?*
- *Is it safe?*
- *Is it ours?*
- *Can we make it?*
- *Can we sell it?*



### The Route to Market

Lead Generation → Optimisation → Development → Registration → Commercial Pesticide

Number of compounds

140,000* → 5000 → 30 → 1-3 → 1 → 1 → 1

*average number of compounds synthesised to deliver one new market introduction in 2005

2

# Candidate Evaluation – *Does it work?*

- Glasshouse tests are an imperfect model for screening compounds because the conditions are controlled:

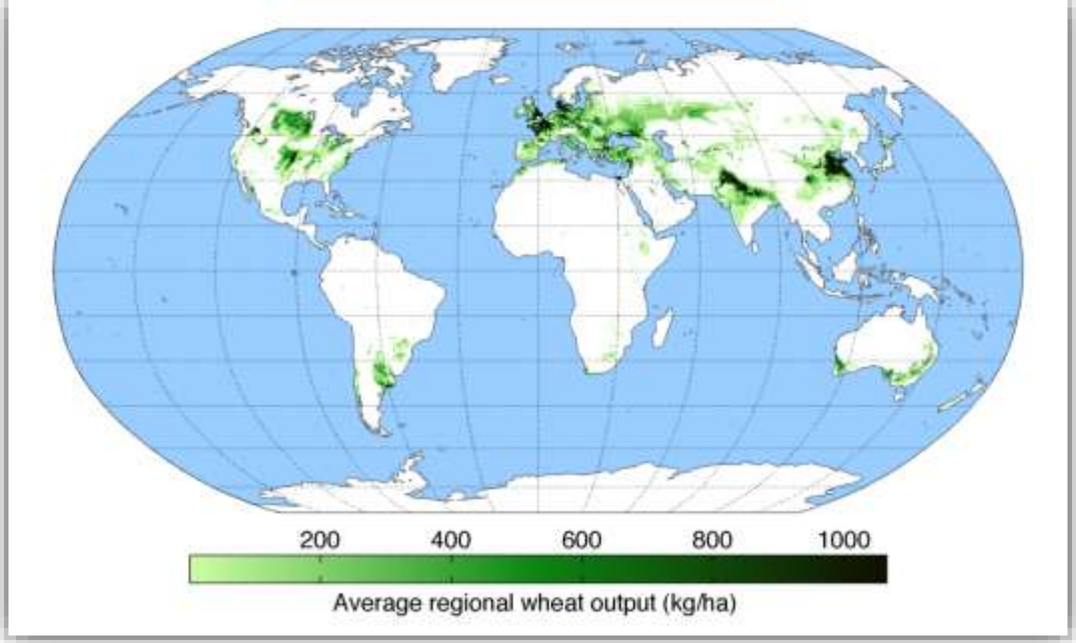
  - There are no other pests
  - Plants are well watered and cared for
  - It is hard to mimic real weather conditions

- **Field trials** give a more realistic picture of what would happen if a farmer used the product on his field

  - Methods of application are more realistic
  - Weather conditions and pests are typical of a given area

# Field Trials

- **Field trials** are carried out in key geographical areas depending on the crop e.g. corn in various US states, rice in China, Northern Italy, Indonesia and Japan, cereals in Northern Europe and Canada



- Field trials will be carried over **several growing seasons** around the world to fully understand the potential (and weaknesses) of a product

- In addition, the performance of the product is further optimised by consideration of the application rate and timing, potential mixture partners and the **formulation**

# Formulation

- Active ingredients (AI) are not applied directly to a field but are formulated or processed into a product that can be used easily, effectively and safely

- It is important that formulations:
  - Are stable in a bottle (in concentrated form) for a given shelf life
  - Are stable when diluted in water for up to 24 hours
  - Allow the AI to be transported to where it is needed within the plant
  - Maximise the biological activity and crop safety
  - Have good spray characteristics

# Formulation

- When selecting a formulation it is important to consider the following:
  - How the compound will be applied and route of uptake into the plant (i.e. roots/soil, into/onto leaf surface or into paddy water for rice)
  - Physical properties of the compound
    - Melting point
    - Water solubility and log P
  - Any chemical stability issues
    (i.e. photostability/hydrolytic instability)

# Formulation: Additives

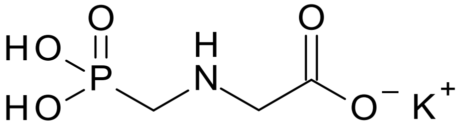
- Formulation additives can be used to enhance the bioavailability of the AI, make it safer to the crop, or improve the physical stability of the product
  - **Surfactants** (adjuvants) give improved distribution, retention and uptake into plant tissues. They also help to physically stabilise certain formulations, thus preventing crystal growth and sedimentation
  - **Stabilisers** improve chemical stability on storage or prolong the biological effect in the field and may be antioxidants, UV screeners or buffers
  - **Safeners** improve the selectivity of the product with respect to the crop
- e.g.

Pinoxaden (AI) Adigor Cloquintocet-mexyl

# Formulation Examples: Soluble Concentrates (SL)

- Homogeneous solution of the AI in water

- AI requirements:
  - Water soluble
  - Hydrolytically stable



orange squash glyphosate dilution

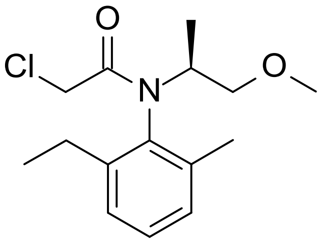
- Example products:
- orange squash, glyphosate (highly water soluble, log P = -3.2)

| Advantages | Disadvantages |
|---|---|
| Cheap (mostly water) | Low rain-fastness |
| Safer for users due to reduced dermal penetration wrt oil based formulations | Low temperature instability of formulation (i.e. freezes at 0 °C!) |
| Good dilution properties | Possible strength limitations |

# Formulation Examples: Emulsifiable Concentrates (EC)

- Spontaneous emulsification of AI and surfactants in water to give fine dispersion of droplets (1-2 microns)

- AI requirements:
    - High oil solubility
    - Liquid AI

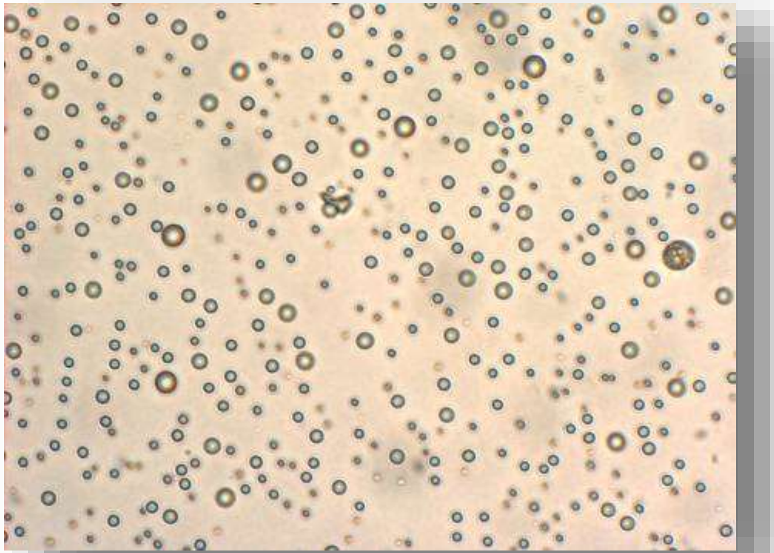


dettol S-metolachlor dispersion

- Example products:
    - dettol, *S*-Metolachlor (log P = 2.9, colourless oil)

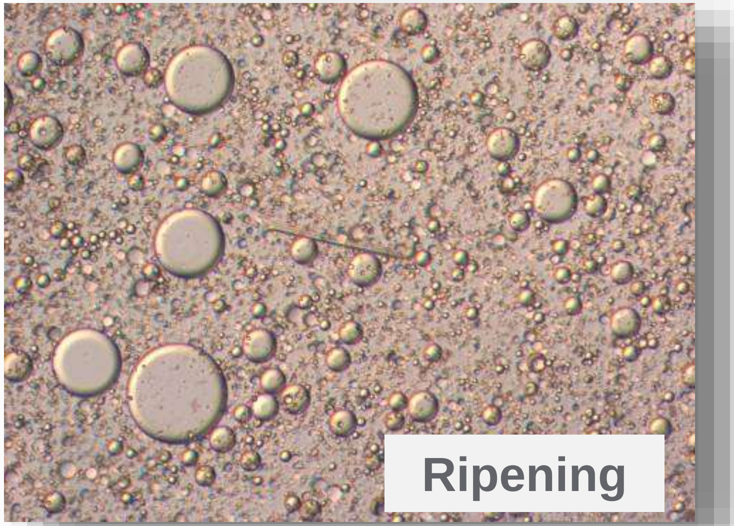
| Advantages | Disadvantages |
|---|---|
| Good bioefficacy | May get crystallisation at low temperature |
| Simple processing | Potentially less safe for operators |

# Formulation: Microscopy of EC Formulation Dilutions

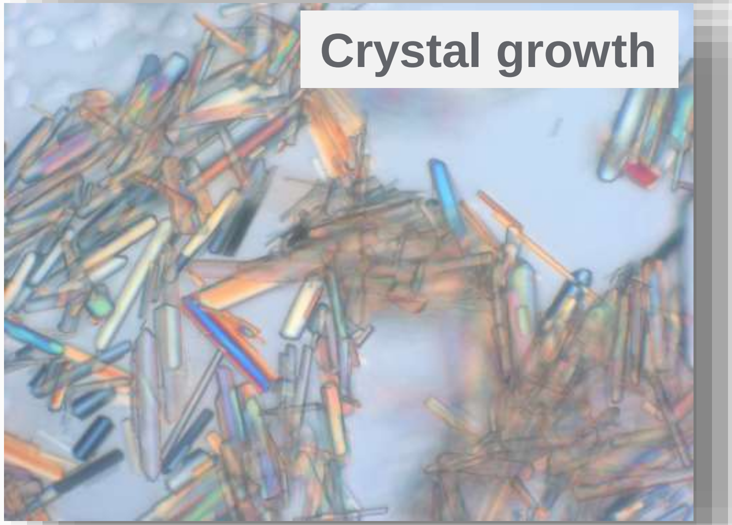


**Well dispersed droplets**

**~2 microns**



**Ripening**

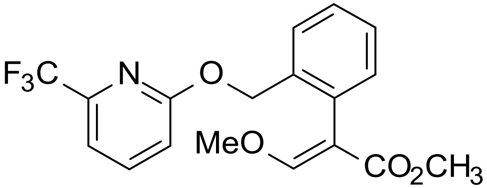


**Crystal growth**

# Formulation Examples: Suspension Concentrates (SC)

- Finely milled particles of AI (0.5-10 microns) suspended in water

- AI requirements:
  - Hydrolytic stability.
  - Melting point > 70 °C.
  - Low water solubility.
  - Stable crystalline form.



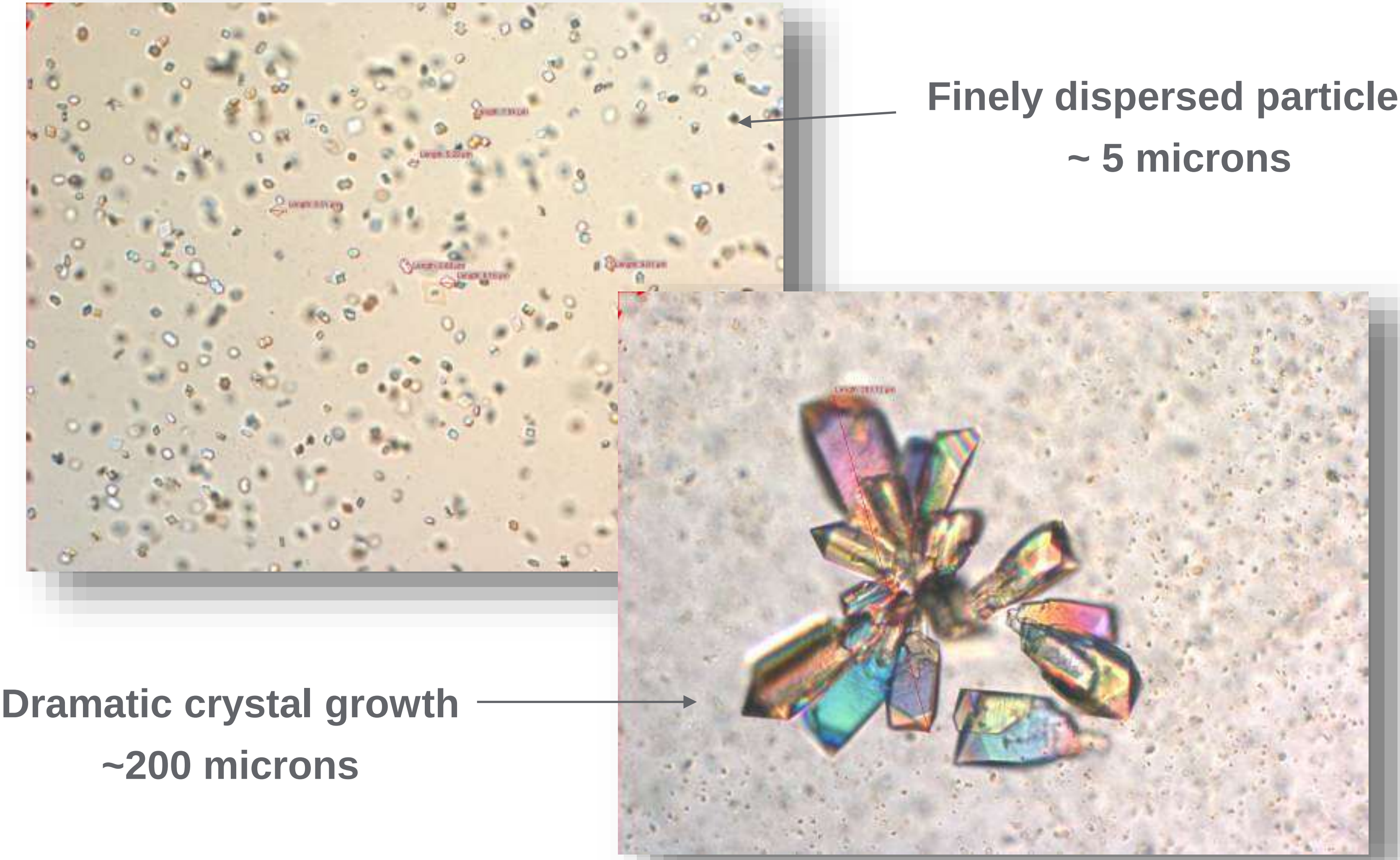
milk of magnesia

picoxystrobin suspension

- Example products:
  - milk of magnesia, picoxystrobin (colourless powder, mp = 75 °C, log P = 3.6)

| Advantages | Disadvantages |
|---|---|
| Operator safety | Complex to optimise |
| High strength possible | Physical stability issues due to crystal growth |
| Good biological activity | Sensitive to changes in AI e.g. polymorphs |

# Formulation: Microscopy of Milled Particles for SC



**Finely dispersed particles**

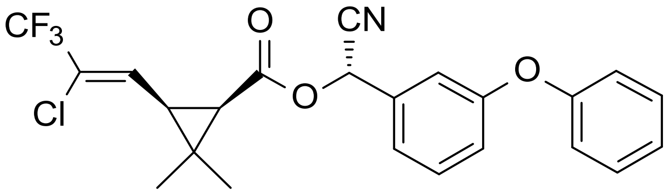
**~ 5 microns**

**Dramatic crystal growth**

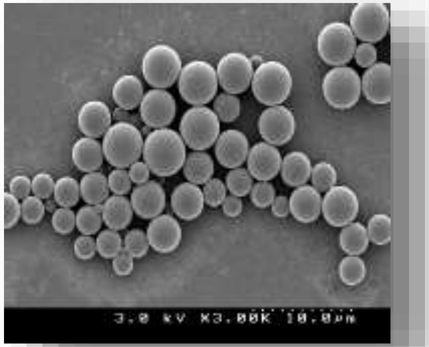
**~200 microns**

# Formulation Examples: Microcapsules

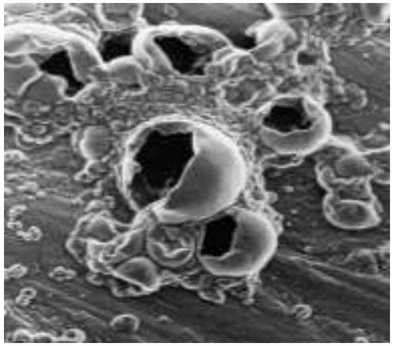
- Microcapsules are hollow spheres (made of polyurea) that have been filled with AI and are typically 1-20 micron
- Upon drying, the AI is rapidly released to give quick action and good residual pest control
- Example product: lambda-cyhalothrin (i.e. Karate with Zeon Technology)



Karate®
with Zeon technology



**Wet microcapsules**



**Karate Zeon capsules**
**5 mins after drying**

## Advantages

Improved safety for workers due to limited contact with active ingredient

High load capacities are possible making capsules an economically viable technology for controlled delivery of the active ingredient

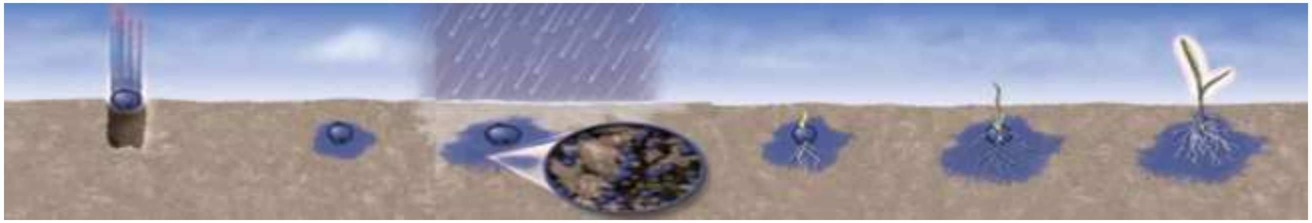
# Formulation Examples: Seed Treatment

- Involves the chemical treatment of the seeds/seedlings (usually with fungicide or insecticide, rarely used as a herbicide treatment) to protect it from a range of pathogenic organisms in the environment such as:
    - Soil-borne diseases
    - Early foliar diseases
    - Insect pests
    - Nematodes


**Dressing**

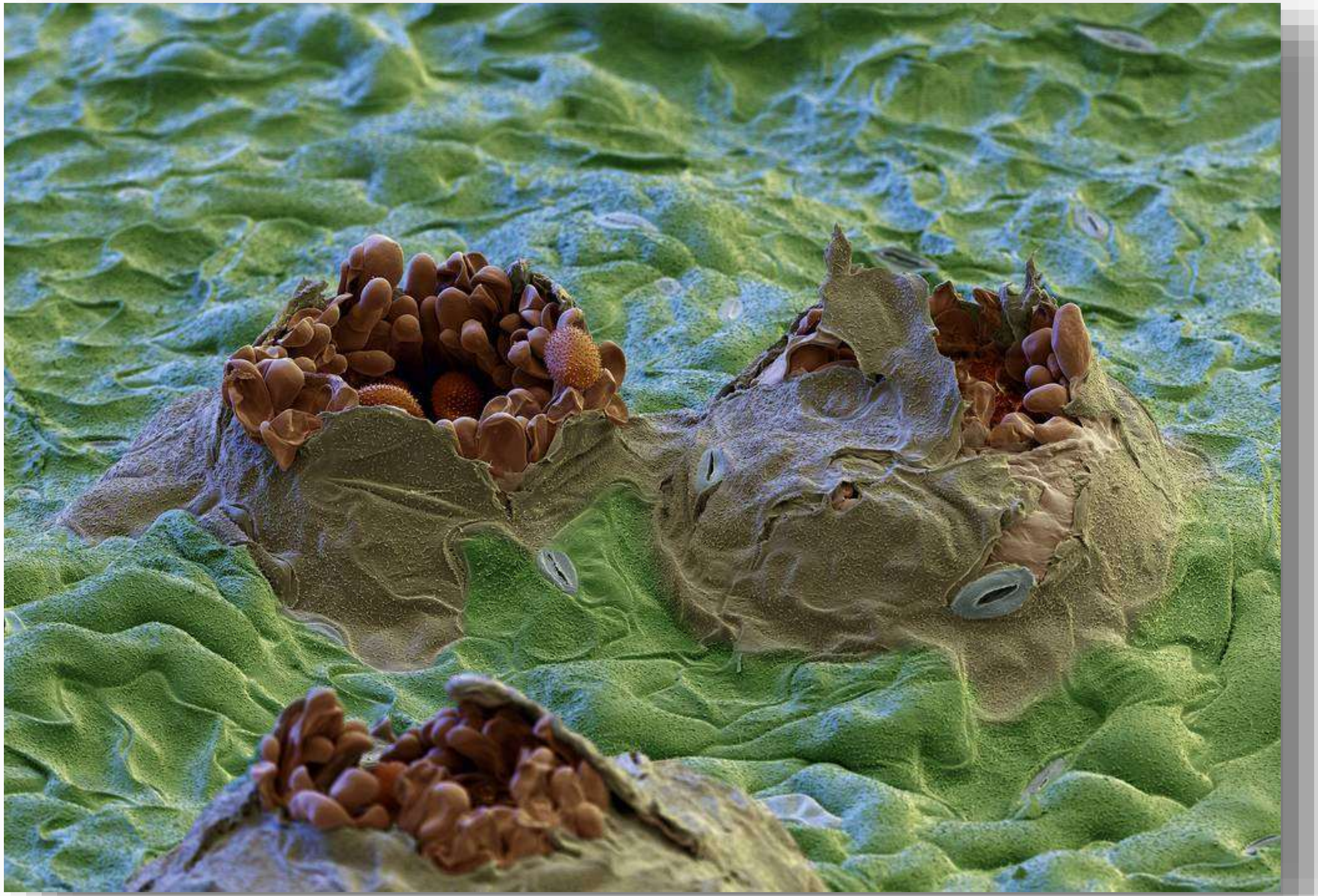

**Coating**


**Pelleting**



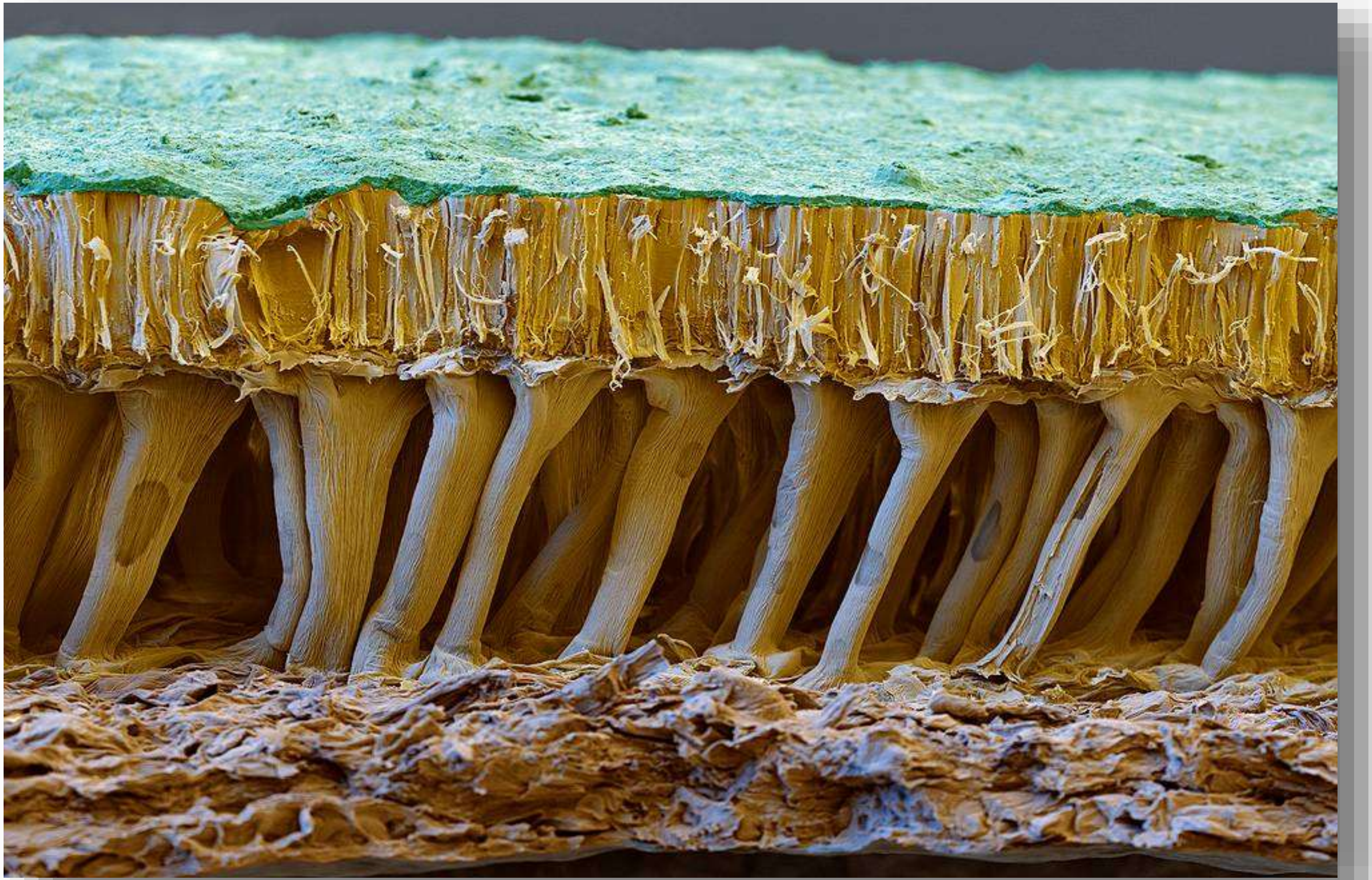
| Advantages (w.r.t. spraying) | Disadvantages (w.r.t. spraying) |
|---|---|
| Reduced exposure for the farmer | |
| Less AI per hectare required | Post emergence spray may also be required |
| Safer to non-target insects | |

# Seed Treatment: Soybean Rust



500:1 magnification of soybean rust fungus, which can decimate a soybean crop (Image: BASF)

# Seed Treatment: Soybean Rust



Soybean seed which has been coated with a fungicide to protect against soybean rust. The seed coating is roughly 1/10th the thickness of a human hair (Image: BASF)

# Formulation Additives: Safeners

- Herbicide **safeners** selectively protect crop plants from herbicide damage without reducing activity in the target weed species

- Used commercially to improve herbicide selectivity between crop and weed species

- Usually applied as part of the formulation with the herbicide but can also be applied as a seed treatment.

**Herbicide safener + herbicide on wheat**

A,B,C = 3 different safeners

D = unsafened wheat

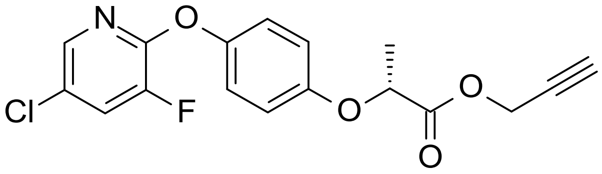
E/F = grass weeds

# Formulation Additives: Safeners

- **Safeners** work by increasing the enzyme activity in the crop with respect to the weed which allows faster **metabolism** and therefore deactivation of the herbicide

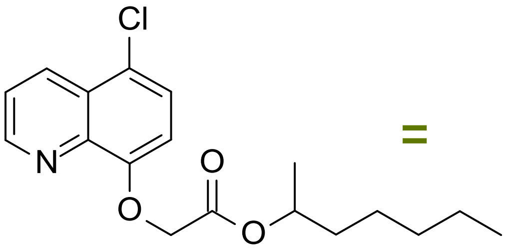
- For example: **Topik** is a 4:1 mixture of Clodinafop (ACCase herbicide) and Cloquintocet-Mexyl (safener). The combination controls post emergence annual grass weeds in rice, soybeans, wheat and rye

- The addition of the safener ensures good crop tolerance



**Clodinafop**

+



**Cloquintocet-Mexyl**

=

Topik®

# Candidate Evaluation – *Is it Safe?*

Strict regulations must be adhered to regarding the safety of a particular product before it can be registered for sale

It is important to consider whether the product will be safe to:

- **The environment**

  - Where the does the product go?
  - What does it break down to form in the environment?
  - What effect does it have on non-target organisms?

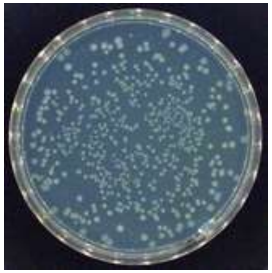
- **Human health**

  - How hazardous is it?
  - What metabolites does it form in plants and how much will reach the food chain?
  - What is the exposure during application?

Very sensitive analytical methods are developed, and used to measure levels of the parent compound and it's metabolites in the soil and in the food chain

In Europe, registration is controlled by the EU Commission and the European Food Safety Authority (EFSA). In the US, it is controlled by the US Environmental Protection Agency (EPA)

# Hazard + Exposure



**Hazards:**

- Extensive cell based and *in vitro* assays are carried out early on in a project to asses whether an active ingredient has any toxicity issues
- Assuming severe toxicity is not observed, further laboratory studies combined with a selection of animal studies can tell us exactly what effects the compounds might have and the dose rates at which they might occur
- Short term studies look for signs of acute toxicity, irritation and sensitisation while longer term studies focus on the potential to cause cancer, reproductive effects or other chronic illnesses

**Exposure issues that are considered include:**

- Operators may be exposed to the product as it is handled and applied
- Workers, bystanders and those living close to the area of application may come into contact with the product as it is applied or through entering treated areas
- Consumers may be exposed to residues in food or water

# Safety and Environmental Studies – Risk

- Acceptable levels of exposure are based on **additional safety margins** several orders of magnitude higher than the already safe levels measured in a range of animal studies

- **Acceptable Operator Exposure Levels** (AOELs) are defined for operators applying the products or for people working in treated areas. The use of suitable protective clothing will usually be recommended as an additional safety feature

- **Acceptable Daily Intake** (ADI) levels and the **Acute Reference Doses** (ARfDs) for consumers are also determined, for both the product and it's metabolites.

# Candidate Evaluation – *Is it ours?*

## Intellectual Property

- Once the AI, formulation etc has been established it must be protected by a **patent**
- A patent is a temporary exclusive right *to prevent others from commercially making, using, selling, importing, or distributing a patented invention without permission*
- Having a patent does not automatically grant the patentee **freedom to operate**
- A patent usually lasts for 20 years from the date of application and are normally published after 18 months
- The invention is defined, and limited, by the scope of the claims
- The basic requirements for a patent are that the claim:
    - Is **novel**: not already known in the literature
    - Is **adequately described**: so someone else could repeat it
    - Involves an **inventive step**
    - Is **useful**: capable of industrial application
- Patents may also cover processes for making a product, not just the use of the product itself

# Candidate Evaluation – *Can we make it?*

**Two key questions:**

- Can the compound be made at the required scale?
- Can the compound be made **economically** at the required scale?

**Discovery Chemistry** | **Process Research** | **Process Development** | **Manufacturing**

**10 - 100 mg** → **10 g - 1 kg** → **1 - 100 kg** → **1000+ tons**

**Greenhouse Screening** | **Field Trials** | **Toxicology and Environmental** | **Sales**

# Candidate Evaluation – *Can we sell it?*

- The average cost of developing a new agrochemical is **~$260 million** (not including the cost of failures)
- This cost (and more) must be recouped in future sales
- We need to understand what the market is for the product…but it takes ~10 years to get to market so assumptions must be made about what the market will be like in the future
- Are we trying to solve a problem that doesn't exist?
- Can we compete with existing products in the market?
- The price that you can sell the product at is determined by how much the farmer is willing to pay
- For a £1 loaf of bread, the farmer receives ~£0.07 (gross) *(Source: National Farmers Union)*

# Candidate Evaluation – *Can we sell it?*

- The bearable **Cost of Goods** (COGs) is related to the expected price and the application rate
- The greater the benefit to the farmer, the more he will be willing to pay
- The more potent the compound, the lower the use rate per hectare
- For example, if you have a target of $50/hectare (1 hectare = 10,000 m$^2$) the cost of goods depends upon the application rate as shown:

| Application rate | Bearable cost of goods |
|---|---|
| 1000 g/ha | 50 $/kg |
| 500 g/ha | 100 $/kg |
| 250 g/ha | 200 $/kg |
| 100 g/ha | 500 $/kg |

- NB – Developing a compound which is 2x as active may not be justified if it costs 3x as much to produce, however there may be other benefits to lower application rates such as environmental safety

# Process Chemistry

- The development of efficient, cheap and scalable routes to new agrochemicals is key to their development
- Process chemistry begins with **route development** followed by **optimisation**, **scale-up** and **troubleshooting**
- An in-depth understanding of the processes is crucial to anticipate and avoid potential problems

Discovery Chemistry

Process Research

Process Development

Manufacturing

10-100 mg → 10 g-1 kg → 1-100 kg → 1000+ tons

Greenhouse Screening

Field Trials

Toxicology and Environmental

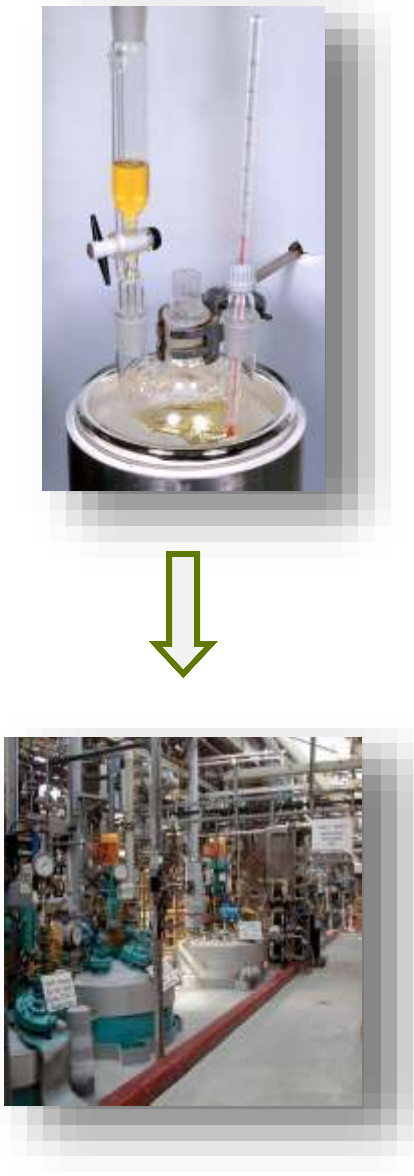
Sales

26

syngenta

# Route Selection: Process Research

- Once a compound has been selected for development, a suitable manufacturing route must be found
- There will always be too many routes to test them all
- A systematic approach can be used:
  - **Brainstorm** all possible route options
  - Decide on route **selection criteria**
  - Search all relevant **literature**
  - **Prioritise** the paper routes based on selection criteria
  - **Evaluate** best routes in the laboratory
  - Use lab **data** to rank the potential routes
  - **Select** the most promising route(s) for further development



# Possible Route Selection Criteria

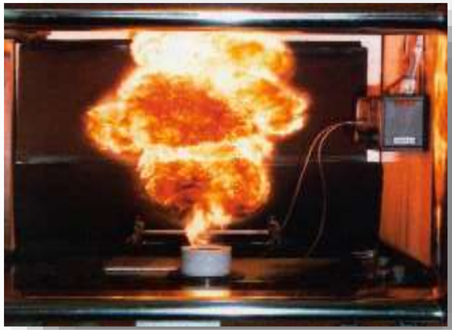
- A number of criteria may be used to assist in route selection:

| Criteria | Considerations |
|---|---|
| **Final product cost** | What are expected yields and raw material costs? How many stages in the process? Is any capital expenditure required? |
| **Ease of scale up** | Is there any chemical precedence? Are raw materials available in bulk? What purification is required? |
| **Hazards** | Which reagents and solvents will be used? What reaction conditions are required? |
| **Environmental impact** | Which reagents and solvents will be used? What yields are expected? What stoichiometry is required? Is there any scope for recycling material/waste solvent? |
| **Freedom to operate** | Is the route free from other companies intellectual property? |
| **Product quality** | How will the product be purified? How reliable/robust in the chemistry? Is there a consistent impurity profile? |

# Potential Scale Up Hazards

- When choosing a suitable route for scale up it is important to consider the following potential hazards
  - Use of toxic or highly reactive reagents
  - Use of flammable solvents
  - Use of dusty reagents or intermediates
  - Generation of reactive intermediates (e.g. diazonium compounds)
  - Generation of gas during the reaction
  - Generation of excess heat during the reaction.

**Flammable solvent**

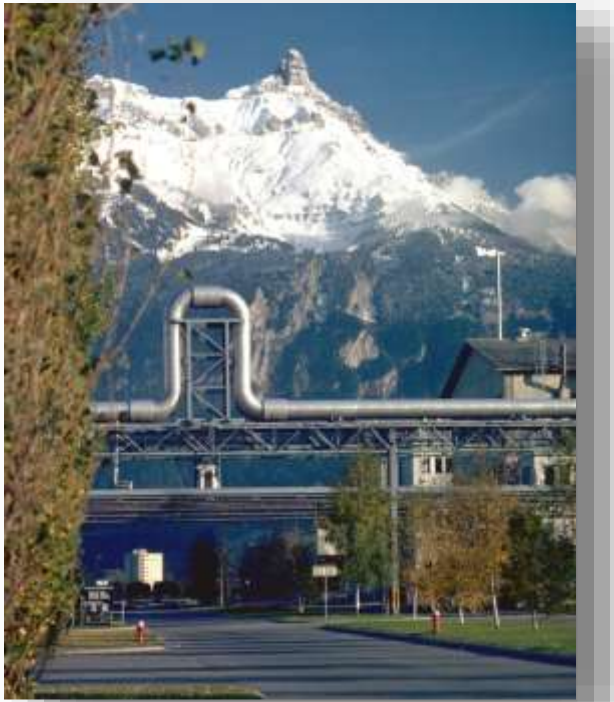


**Dust – ignition of 5g of anthraquinone**

**Explosion due to build up of reactive intermediates**

# Process Development

- Once the route has been selected, a manufacturing process must be established that:
  - Minimises the product **cost**
  - Meets the required **specification** (purity, impurity profile, colour, physical form, etc.)
  - Has the lowest possible capital expenditure
  - Is **safe** to operate
  - Is compliant with environmental legislation



# Process Development

- Process optimisation involves careful optimisation of each step within a chosen route including:
  - Optimising temperatures and pressures
  - Combining or telescoping successive reactions
  - Reducing excess stoichiometries to minimise waste
  - Maximising reaction concentrations
  - Determining rate and order of addition
  - Varying reagents and catalysts

# Process Research and Development: Summary

- Requires expertise in synthetic and physical chemistry combined with chemical engineering skills
- Aims to achieve **lowest possible cost** (product + capital)
- Processes must be **reliable and reproducible**

- Development of a product is ongoing through its lifetime (process optimisation):
  - Responding to changing environmental issues
  - Trouble shooting
  - Raw material supplies
  - Cost reduction