

INTRODUCTION TO DRUG DISCOVERY

SEMESTER 2, MODULES 06527 and 06509

Dr A.N. Boa

Department of Chemistry, University of Hull

a.n.boa@hull.ac.uk, Room C301

Brief course outline

What is drug research?

Principles of drug action

Programme selection

Finding a lead compound

Structure-based and rational drug design

A very brief overview of lead optimisation, drug development and clinical trials

Bibliography

Web links

This document is available along from <http://www.hull.ac.uk/php/chsanb/teaching> along with selected links to relevant papers on-line.

Library resources - Section RS 4xx of the library

1. *'Medicinal Chemistry, Principles and Practice'*
Ed. F.D. King, Royal Society of Chemistry, 1994, **RS 403 M4**
2. *'Medicinal Chemistry: A Biochemical Approach'*
T. Nogrady, OUP, 1988, **RS 403 N7**
3. *'Introductory Medicinal Chemistry'*
J.B. Taylor and P.D. Kennewell, Ellis Horwood, 1981, **RS 403 T2**
4. *'Combinatorial Chemistry'*
N.K. Terrett, OUP Oxford Chemistry Masters Series, 1998, **RS 419 T3**
5. *'An Introduction to Medicinal Chemistry'*
G.L. Patrick, 2nd Ed., Oxford University Press, 2001, **RS 419 P3**
6. *'Burger's Medicinal Chemistry and Drug Discovery'* (comprehensive but advanced)
4th Ed., Ed. M.E. Wolff, Wiley, 1995, **RS 403 B9** (also 5th Ed. in Ferens' library)
7. *'Medicinal chemistry : an introduction'*
G. Thomas, Wiley, 2000, **RS 403 T4**
8. *'Drug discovery : a history'*
W. Sneider, Wiley, 2005, **RS 61 S6**

Glossary of Common Terms used in Drug Discovery

ADME(T):

Absorption, distribution, metabolism, excretion (and toxicology).

Array synthesis:

Form of parallel synthesis in which the reaction vessels are maintained in a specified spatial distribution, e.g. the “wells” of a 96-well plate

Bead:

A particle, normally spherical, of a microporous polymeric material used as a support to attach reagents or substrates. Used in solid-phase organic chemistry.

Bioisostere: (<http://en.wikipedia.org/wiki/Bioisosterism>)

Bioisosteres are substituents or groups with similar physical or chemical properties that impart similar biological properties to a chemical compound. In drug design, the purpose of exchanging one bioisostere for another is to enhance the desired biological or physical properties of a compound without making significant changes in chemical structure.

Chemical diversity libraries:

A library of chemical compounds may be small (e.g. a few compounds) or large (e.g. thousands or even millions of compounds), and it may focus on a narrow or wide range of “diversity space.” These libraries offer unprecedented opportunities for the rapid identification of small molecules with significant physiological effects.

Combinatorial Chemistry:

Using a combinatorial process to prepare sets of compounds from sets of building blocks. In the early 1990's it was believed that combinatorial chemistry would revolutionize the drug discovery industry. Ten years later the route from design and synthesis of compound libraries to identification of lead structures is still long and costly. Synthesis of an almost unlimited number of organic compounds covering as much of “chemistry space” as possible is no longer the most cost effective and time saving approach to hit identification. Creating libraries, using biological target structure to inform chemical design, facilitated by quantum advances in structural genomics and computational capabilities, is a smarter, more efficient way to produce good initial leads. Considering solubility, permeability and other drug-like properties early in library design and introducing both target and lead structural constraints in lead development are further ways to ensure more compounds make it to trial.

Drug:

A drug is a molecule that interacts with a biological molecule triggering a physiological effect. Drugs for the treatment of “illnesses” produce positive effects relating to the condition.

Enzymes:

Enzymes are catalytic proteins that increase the rate of chemical reactions in the body.

Hit:

A “hit” is colloquial term used for a compound whose biological activity exceeds a predefined, statistically relevant threshold or a molecule with robust dose response activity in

a primary screen and known, confirmed structure. The precise definition of the following terms varies widely between drug discovery companies.

Hormones:

Chemicals released onto the bloodstream, they produce their physiological effect in tissues possessing specific hormone receptors.

Lead compound:

A lead compound is a representative of a compound series with sufficient potential (as measured by potency, selectivity, pharmacokinetics, physicochemical properties, absence of toxicity and novelty) to progress to a full drug development programme. The precise definition of the following terms varies widely between drug discovery companies.

Lead discovery:

Lead discovery is the process of identifying active new chemical entities (NCEs), which by subsequent modification may be transformed into a clinically useful drug.

Lead optimization:

The synthetic modification of a biologically active compound, to fulfil all stereoelectronic, physicochemical, pharmacokinetic and toxicological properties required for clinical usefulness. The new lead optimization paradigm demands that companies move to parallel processes that evaluate binding affinity, ADME, drug properties, etc. earlier in the process in order to cut the time and costs lost in failed leads.

Pharmacophore: (<http://en.wikipedia.org/wiki/Pharmacophore>)

Ehrlich described a pharmacophore as "a molecular framework that carries (*phoros*) the essential features responsible for a drug's (= *pharmacon*'s) biological activity". This definition was updated by Gund (1977) to "a set of structural features in a molecule that is recognized at a receptor site and is responsible for that molecule's biological activity".

Pharmacodynamics: (<http://en.wikipedia.org/wiki/Pharmacodynamics>)

Pharmacodynamics is the study of the biochemical and physiological effects of drugs and the mechanisms of drug action and the relationship between drug concentration and effect. It is often summarily stated that pharmacodynamics is the study of what a drug does to the body, whereas pharmacokinetics is the study of what the body does to a drug.

Pharmacokinetics: (<http://en.wikipedia.org/wiki/Pharmacokinetics>)

Pharmacokinetics is a branch of pharmacology dedicated to the study of the time course of substances and their relationship with an organism or system.

Prodrug:

A prodrug is drug which is given (taken) in an inactive form. Once administered, the prodrug is metabolised by the body into a biologically active compound.

QSAR: (<http://en.wikipedia.org/wiki/QSAR>)

Structure-activity relationship (SAR) is a process by which chemical structure is correlated with a well defined process, such as biological activity or chemical reactivity. For example, biological activity can be expressed quantitatively as in the concentration of a substance required to give a certain biological response. Additionally, when physiochemical properties or structures are expressed by numbers, one can form a mathematical relationship, or

quantitative structure-activity relationship (QSAR), between the two. The mathematical expression can then be used to predict the biological response of other chemical structures.

Receptor

Most drugs produce their effects by acting on specific protein molecules, usually located in the cell membrane. These proteins are called receptors and are normally activated by endogenous chemicals in the body (transmitter substances or hormones). For example acetyl choline is a transmitter substance release from motor nerve endings. It activates receptors on the skeletal muscle initiating a sequence of events that result in the contraction of smooth muscle.

SAR

Structure-activity relationship. See QSAR

Structure-based screening:

Structure- based screening combines the power of NMR spectroscopy, automatic docking, and X-ray crystallography and provides the means to apply structural information (NMR, modelling, and X-ray) early in the projects to identify hits, select targets, and optimize the hits in terms of their affinities and specificities.

Target validation:

Target validation is the determination that a molecular target is critically involved in a disease process and a potentially valuable point of intervention for new drugs.

Transmitter substances:

Transmitter substances are chemicals released from nerve terminals which diffuse across the synaptic cleft and bind to receptors. This activates the receptors (presumably by changing their conformation (shape) and triggers a series of events such as muscle contraction or glandular secretion. After release the transmitter is inactivated by either enzymatic degradation or re-uptake.

What is drug research?

Historical Perspective

Throughout history people have found by trial and error which berries, roots, leaves and barks could be used for medicinal purposes to alleviate symptoms of illness:

Willow bark –	contains <i>salicin</i> –	fever reducing in general
Cinchona bark –	contains <i>quinine</i> –	fever associated with malaria
Chinese herbal remedies –	many examples	

The Doctrine of Signatures (after Paracelsus and Böhme - 16th -17th century) introduced the idea that God had “specially marked everything to reveal its purpose”. So, iris petals were used to treat bruises, goldenrod to treat jaundice etc. These ideas influenced the writings of herbalist Nicolas Culpepper.

With the idea of the Doctrine of Signatures mind The Rev Edward Stone searched along a riverbanks (i.e. a cold and wet place) for a plant-based cure for the fevers associated with influenza. As a result of this work he found, in 1763, that the bark of the willow was effective in reducing fever. Native American Cherokees had in fact used the willow bark for such medicinal purpose for centuries.

Progress in understanding diseases and ailments was however incredibly slow right up until the middle of the 20th century. Some relevant milestones could be considered to be:

17th century: Anton van Leeuwenhoek –Bacteria identified by microscopy

19th century: Louis Pasteur – made link between bacteria and disease

19th century: Robert Koch – identified micro-organisms for tuberculosis, cholera and typhoid

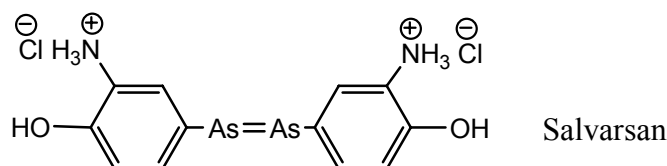
1867: Joseph Lister – a proponent of the “germ theory of disease.” This Edinburgh surgeon successfully used carbolic soap (containing phenol) to prevent infections during surgery.

1891: Paul Ehrlich – coined the term chemotherapy, used synthetic chemicals to try and cure disease.

At beginning of the 20th century anaesthetics (ether) and analgesics to relieve major pain (morphine / heroin) were really the only single chemicals in medicinal use, but there was still no effective treatment to cure infectious diseases and many major ailments.

Paul Ehrlich

- The so-called 'father of modern chemotherapy'
- Original proponent of the “Magic bullet” he aimed to use chemicals to treat disease
- 1910 first fully synthetic drug was made: ‘Salvarsan’ contained arsenic! It was not very good against bacteria but used for treating sleeping sickness (trypanosomiasis) and syphilis (a spirochaete disease caused by *Treponema pallidum*).



However Ehrlich was ignorant of the relevant biochemistry and even now sophisticated methods for predicting drug action do not foretell toxicity and side effects, nor drug transport or drug metabolism *in vivo*.

Modern Drug Research

This course is aimed to provide an **overview** of the key steps in modern drug discovery, starting from the initial decisions regarding project selection through to identification of a drug candidate and the production of an “investigational new drug” (IND) application. Some of these stages are described in detail in other lecture courses (e.g. pharmacology and toxicology), whilst this one concentrates more on the chemistry aspects. A cursory discussion of the clinical trials, the new drug application (NDA) and final release of a drug will be made.

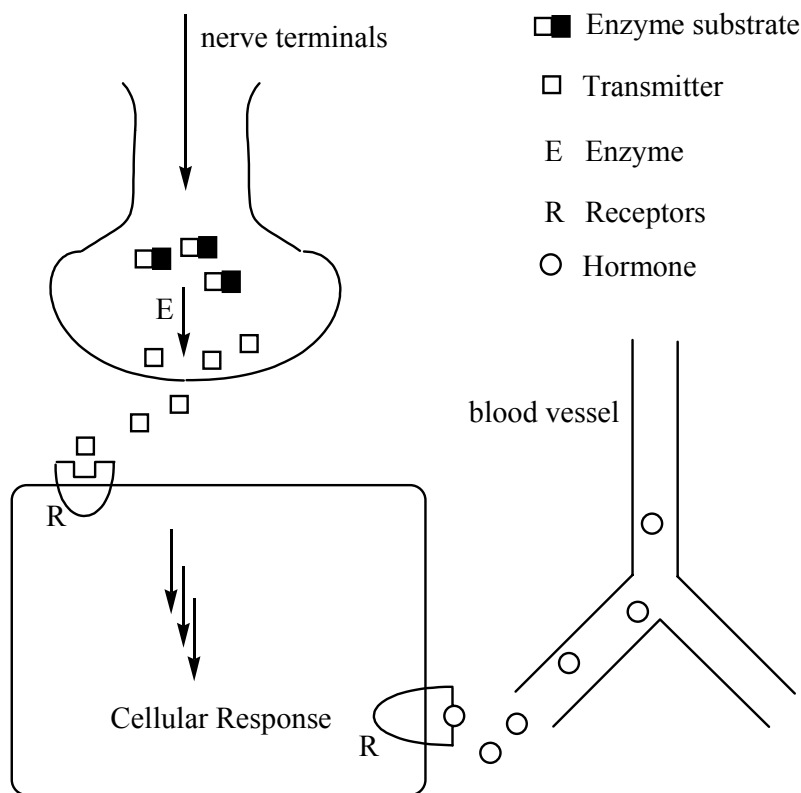
The key stages may be summarized as:

- 1 Programme selection (choosing a disease/condition to work on)
- 2 Identification and validation a drug target
- 3 Assay
- 4 Identification of a lead compound
- 5 Lead optimization
- 6 Identification of drug candidate
- 7 Clinical trials
- 8 Release of the drug
- 9 Follow-up monitoring

In years gone by the ADME and toxicology aspects were often left until a potential drug candidate had been identified. However many compounds failed at this stage so wasting years of effort and large sums of money. Nowadays initial ADMET studies are being conducted earlier and earlier in the process, in order to try and identify and discard problematic compounds as soon as possible.

Principles of drug action

Proteins, whether an enzyme, receptor or ion channel, and nucleic acids form critical links in the great chain that are “biochemical processes”. When one of these links is malfunctioning (not being activated properly, overworking or not working at all) then a disease state may arise. However the body often has ways around such problems. A bio(macro)molecule may be involved in a disease process, but to be a “drug target” it has to be “validated” - in other words shown to be critical in the disease process.



Such drug targets may be (adapted from www.kubinyi.de):

Target	Mechanism
Enzyme	Inhibitor - reversible or irreversible
Receptor	Agonist or antagonist
Nucleic acid	Intercalator (binder), modifier (alkylating agent) or substrate mimic.
Cell membrane ion channels	Blockers or openers
Transporters	Uptake inhibitors

How specific does the interaction between a drug and its target have to be?

A “back-of-the-envelope calculation”.....

Consider: 1 mole = 6×10^{23}
 an active compound MW 200 g mol^{-1}
 therefore 1 mg substance
 gives $(6 \times 10^{23}) \times (10^{-3} / 200) = \underline{3 \times 10^{18} \text{ molecules}}$

A human has approximately 3×10^{13} cells giving.....

$(3 \times 10^{18}) / (3 \times 10^{13}) \sim 10^5$ molecules of active substance per cell

An erythrocyte (a red blood cell) contains approx 10^{10} molecules.

\therefore 1 molecule of active substance per 10^5 cellular molecules.

Validation of a drug target

A bio(macro)molecule may be involved in a disease process, but to be a drug target it has to be validated. In other words shown to be critical in the disease process.

Useful techniques available to validate a target are:

Gene knockout:

does removal of the gene that encodes the target protein result in, for example, the death of a pathogen (disease causing microorganism)?

RNA interference (RNAi):

involves double-stranded ribonucleic acid (dsRNA) interfering with the expression of genes with sequences complementary to the dsRNA. Results in a reduction of the production of the protein (target) in question.

Programme Selection – What area?

The **development of a new compound is complex, time consuming and very expensive**. Success rate in getting an initial target compound to an approved and commercially available product is very low. Less than 2% of new compounds may show suitable biological activity in the laboratory, and structural modifications of an existing class can yield as little as 1% suitable compounds. Less than 10% of these compounds result in successful human clinical trials and reaches the market place.

This course focuses on the **scientific and operational aspects** of medicinal chemical research, but because of the above factors it is vitally important to decide on the right topic of research at the outset. **Before experimental work, a strategy is needed** to ensure that research is focussed on an area of significant medical need and commercial opportunity.

DECIDING ON AN AREA FOR RESEARCH

First of all the medicinal chemist needs to consider **selection of a disease target**.

In the **early days** of the pharmaceutical industry, there were **many therapeutic opportunities** which were successfully exploited and provided the medicines on which today's industry was founded. With antibiotics and psychotropics, for example, it was possible to produce breakthrough remedies for conditions which were not previously treatable. By contrast, in the 1990s, whilst there are still many serious disease challenges, we must also recognise that **acceptable therapies are available for many conditions**. Therefore, one task facing the pharmaceutical industry is to provide new agents which have **clear advantages over existing therapy**.

How to evaluate an opportunity before any new project is commenced? Those proposing a new research project should consider:

- A) the medical need**
- B) current therapy availability**
- C) competitor activity**
- D) commercial opportunity**

These factors can be thought of as a summary of the need for a new or improved therapy. Thus, evaluation of the opportunity should include a consideration of the following:

A) Medical Need

- Severity of the condition.
Is it self-limiting? Patients may not bother buy drugs for conditions that are mere minor irritations but if it's life threatening then the wallet comes out more quickly!

B) Existing Therapies

- Current therapies: is the level of satisfaction high or low?

It is necessary to show some advantage compared with existing therapy - difficult to assess in situations where there is already a high degree of satisfaction with existing therapy. *For example*, clinical efficacy would be straightforward to demonstrate for an oral antibiotic which is intended for the community market, but showing an advantage over available agents would be much more resource intensive, since current therapy is generally effective. I.e. the **feasibility of running a successful clinical development programme must also be considered.**

Other reasons why a new drug may have advantages could be:

- by virtue of increased selectivity for a particular biological mechanism.
- provide a new dosage form which results in a particular advantage to the patient. E.g.azole antifungal drugs for vaginal thrush: old azoles such as clotrimazole came in the form of pessaries, whereas fluconazole (Diflucan) is available in tablet formulation.

C) Competitor activity

What are your competitors doing?

- Will your proposed new drug permit a novel approach to the management of the disease? Optimum agents of a particular mechanistic class may have been identified; next therapeutic advance will require an alternative pharmacological approach.
For example, in the gastric ulcer market the H₂ antagonists (antihistamines) such as ranitidine (Glaxo) were challenged by the new "proton pump inhibitors" such as omeprazole from Astra.

Thus, an important issue is to decide whether to seek improvements within an existing drug class, or follow a novel mechanistic approach.

D) Commercial Opportunity

- The company must consider potential market (patient numbers).....
-and duration of the proposed therapy. Is the condition acute or chronic?

Also it is likely that, at the project selection stage, the medicinal chemist should be aware that a simple demonstration of efficacy and safety is unlikely to be sufficient for approval of most potential therapies.

The product must of course be commercially viable. Assuming the project is successful, about ten years are likely to elapse before an agent can be marketed.

EXAMPLE OF RESEARCH AND DEVELOPMENT TIMESCALE

SYNTHESIS	IND*	NDA ^{\$}	LAUNCH	PATENT EXPIRY
1982	1986	1992	1994	2002
DISCOVERY RESEARCH	CLINICAL DEVELOPMENT	REGULATORY REVIEW	POST MARKETING DEVELOPMENT	



*INVESTIGATIONAL NEW DRUG APPLICATION
^{\$}NEW DRUG APPLICATION

Finally, it is worthwhile to consider whether there are special needs such as **gaps in your corporate portfolio**, or special situations which argue for a therapeutic area. Thus, a particular case might be made to develop a follow-on product with improved properties for an agent nearing the end of its patent life. Also have the advantage here in that probably a lot of groundwork will have been undertaken already.

Starting in the Lab

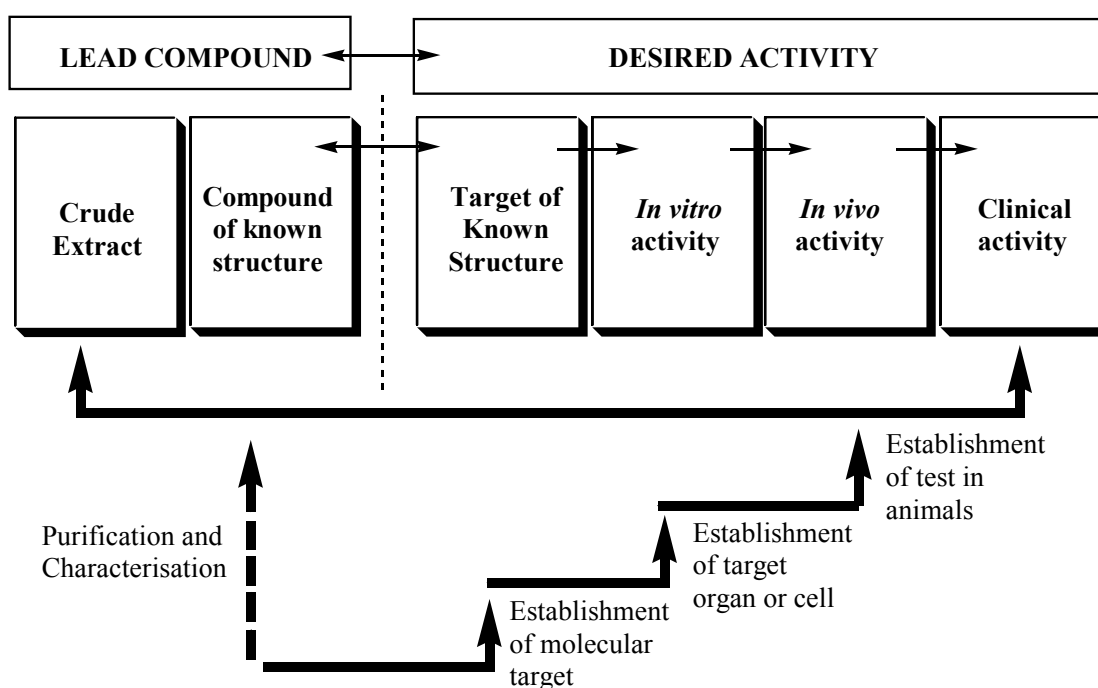
New projects can be divided into those:

- ...which already have a “lead compound” on which to base the design of novel analogues (e.g. existing drugs or those already in advanced clinical trials)
- ...those which do not. For these a lead compound needs to be identified.

DEFINITION: *Lead compound*

A lead compound is a compound from a series of related compounds that has some of a desired biological activity. This molecule can be characterised and modified to produce another molecule with a better profile of wanted properties to unwanted side effects.

A lead compound is a first foothold on the drug discovery ladder. It takes much more effort to make a lead compound into a drug candidate. Identification of a lead compound may arise from a variety of different routes and in modern drug discovery this is likely to be very different from in the past.

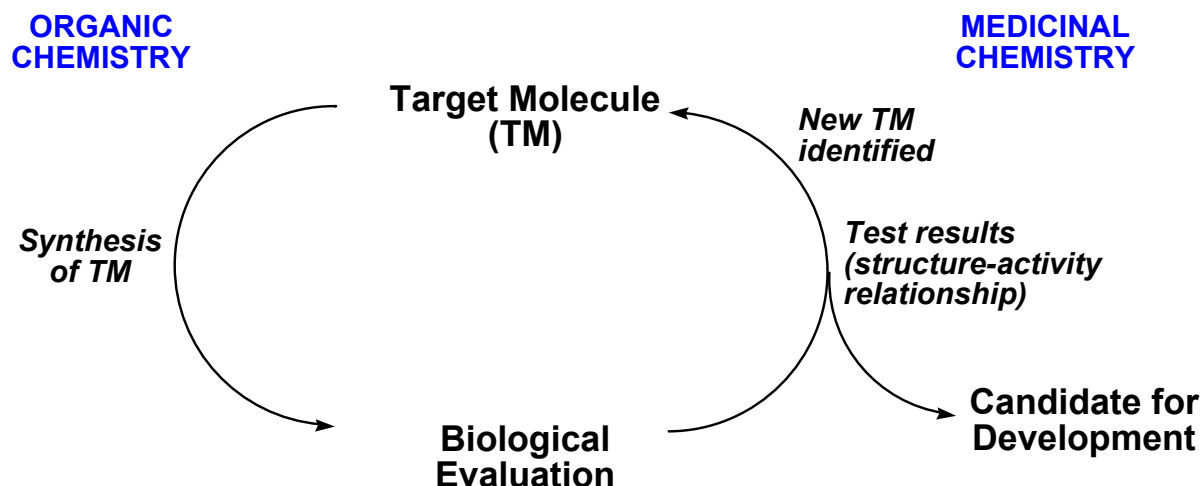


It is obvious, but worth stating for emphasis, that the screening / assay of your compounds must reflect the properties you are seeking.

DEFINITION: *Screening (assaying)*

The testing of a (series of) molecule(s) against a known biological **target** that correlates with a cellular, pharmacological activity is known as screening, e.g. enzyme inhibition or receptor binding.

This is feasible in modern research because: a) targets can be identified, b) targets are available in through isolation of the protein in large quantities (molecular biology techniques), c) robotised high-throughput screening technologies have become available.



Assuming synthesis is ongoing there **must be primary screens** available which have **sufficient capacity to enable an acceptable rate of progress**, and which are capable of discriminating between compounds which are worthy of further investigation and those which are not. Thus potency and / or selectivity at a given receptor or enzyme may be primary selection criteria.

Biological Evaluation

A) Choice of an *in vitro* screen is helpful in most cases in ensuring that new compounds act by the chosen mechanism and meet minimum levels of potency and selectivity.

B) In many areas there will be more detailed *in vitro* assays, but it is important that these primary selection screens are followed up with meaningful *in vivo* animal models. These should be designed to confirm the relevance of compounds found active with *in vitro* assays. It is again helpful to have the capacity to evaluate a significant number of analogues so that rapid turn-round is achieved to guide chemistry. Typically these screens might provide data on *in vivo* potency, dose response, duration of action, and pharmacokinetics.

C) At the third level, there is a place for a more complex efficacy model, though this might be reserved for compounds of substantial interest. In some areas it is impractical or impossible to employ an efficacy model and compounds may be advanced to the clinic on the basis of animal pharmacology and pharmacokinetics. However, although advances in the study of transgenic animals may enable the development of animal models of diseases which exclusively affect humans, their use must be approached cautiously and with thoughtful choice of controls to validate their use.

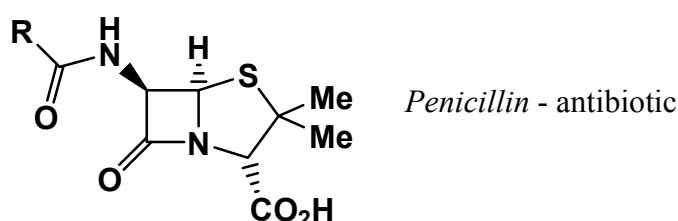
Sources of Lead Compounds

How do we get started? How do we find a new lead compound? The identification of a new lead compound can fall into two categories, depending on whether a biological target relevant to the disease/condition in question is known or unknown. Below are some examples of how lead compounds have been discovered in the past.

Chance Observation

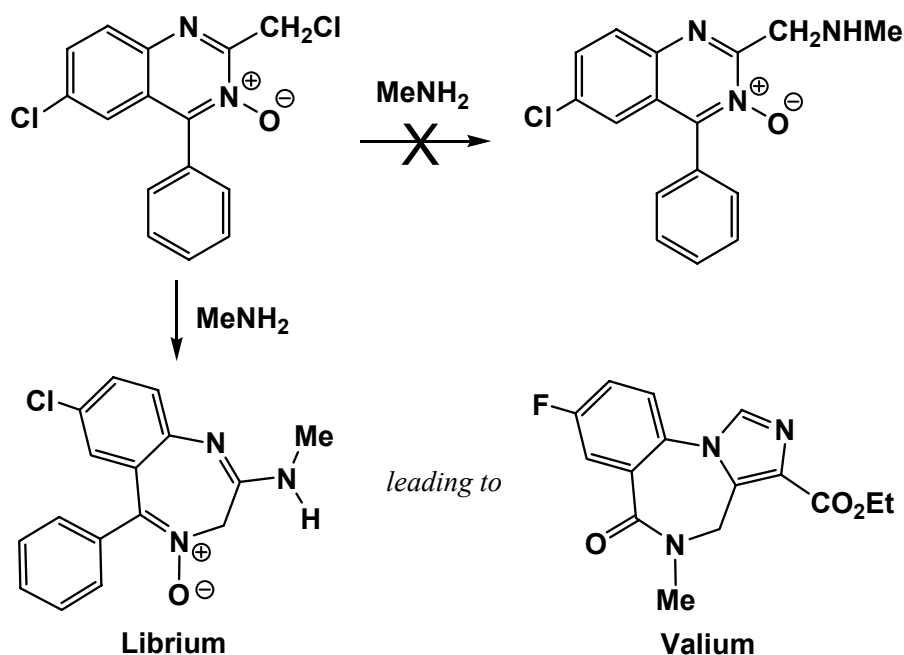
Penicillin

Fleming discovered that mould from one culture caused bacteria in its vicinity to undergo lysis. Penicillin G and V isolated; antibiotic activity discovered; new β -lactams discovered and made through modification to give antibiotics with improved activity.



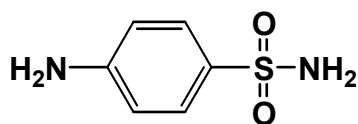
Chlordiazepoxide (Librium)

Sternbach, searching for new drugs to treat anxiety, revisited some quinazoline oxides which were first made 20 years ago and thought to have the benzheptoxadiazinone structure. Forty or so products were made and found to be inactive. One last reaction with a primary amine, methylamine, produced a white solid that was shelved as the project was winding up. The product was subsequently rediscovered in a lab tidy up, shown to have the 6,7 ring benzo-1,4-diazepine structure and biological evaluation revealed its potent activity. The product was subsequently modified to give benzodiazepines such as temazepam, lorazepam, dornonoct etc. used for the short term treatment of anxiety.

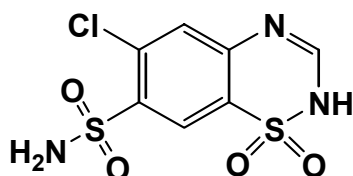


Clinical Observation of Side effects

Some drugs have been found to have side effects. Thus we can aim modify structures to reduce the primary indication and optimise side effects. E.g. sulphanilamide is the active form of the sulphonamide class of antibacterial agents and led to the identification and development of diuretics, such as chlorothiazide.

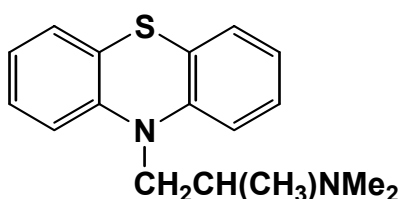


sulphanilamide

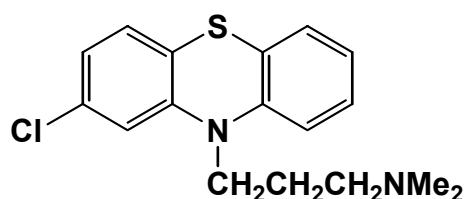


chlorothiazide diuretic

Another case is the development of phenothiazine antipsychotic-tranquillizers such as chlorpromazine which was the first drug used for the treatment of schizophrenia. Initially phenothiazines were being investigated as antihistamines by Laborit, a French Navy surgeon, for blocking the action of the vasodilator histamine (for preventing surgical shock), but the relaxed unconcerned nature of the patients undergoing surgery was noticed by the surgeon.



Promethazine



Chlorpromazine

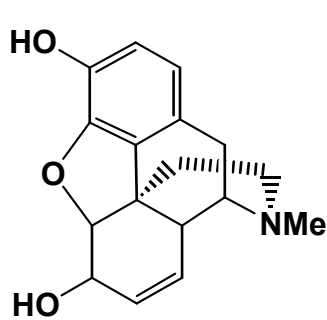
These “successes” have led to the SOSA (selective optimisation of side activities) approach to lead discovery. This approach takes existing drugs, which of course have already been approved for use in humans and therefore extensively tested, and screens them for activity against new pharmacological drug targets. The Prestwick Chemical Library is a collection of 1120 off-patent compounds available for this purpose. (<http://www.prestwickchemical.com/index.htm>).

3.1.3 Natural Product Activity in Herbal Remedies / Folklore

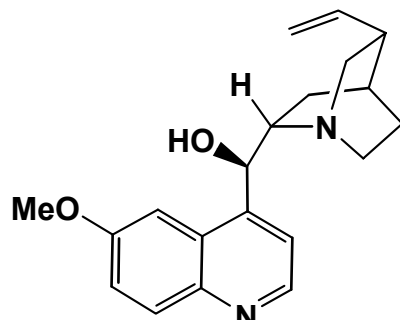
Natural products are the basis of many folk remedies:

<http://www.life.umd.edu/classroom/bsci124/lec29.html> and also: <http://www.life.umd.edu/classroom/bsci124/lec30.html> have a useful summary of plants and their biological effects.

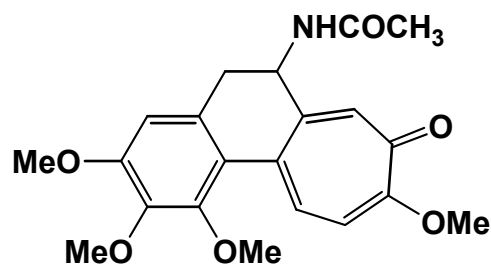
The actual remedy is invariably a complex mixture of compounds from which the active product can be isolated and characterized. This then acts as a lead compound.



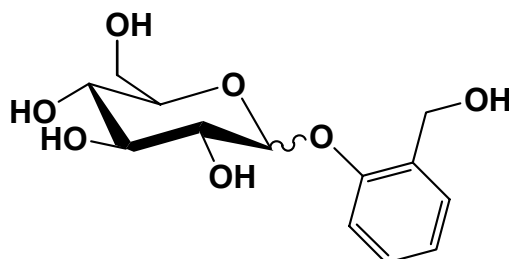
Morphine
poppy - pain relief



Quinine
cinchona bark - antimalarial



Colchicine
crocus - gout



Salicin
willow bark - fever and pain reducing

Lead compounds have been found in almost every category life form, for example:

Life form	Species	Lead Compound/Drug
Moulds	<i>Cephalosporin acremonium</i>	cephalosporins (β -lactam antibiotic)
Plants	yew tree	taxol (antitumour agent)
Marine Organism	deep water sponge	discodermolide (antitumour agent)
Reptile	snake venoms	teproptide (ACE inhibitor)

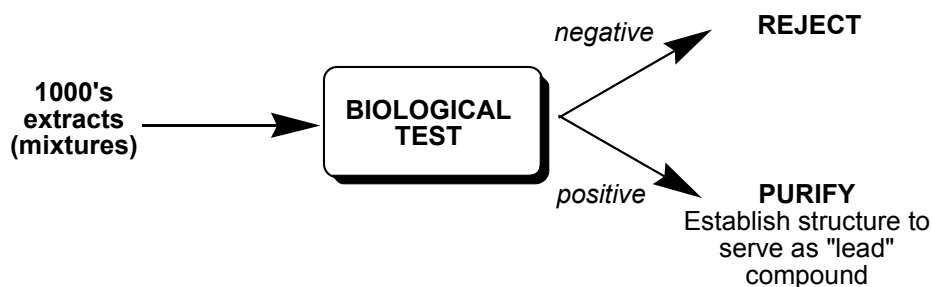
Natural Product Screening

We have seen that chance observation, side effects of current drugs and folk / herbal remedies serve as useful starting points for the identification of *lead compounds*.

It is worth noting:

- 80% of the world's population uses drugs exclusively from natural sources.
- 35% of drugs contain 'principles' (key structure elements) of natural origin
- Less than 5% of the 500,000 higher plant species have undergone biological pharmacological screening.
- Each plant has potentially 10,000 different constituents.

Based on the fact that many existing drugs are based upon natural products and that natural products have served as excellent lead compounds, general natural product screening is widely used as a method of finding lead compounds.



Thus extracts from plants, marine organisms, animal toxins, microbial broths are all used in biological activity screening tests (assays) in the search for biological activity. For this a relevant target needs to have been identified and a screen developed. If an extract gives a positive 'hit' then the active constituent is isolated to hopefully serve as a lead compound.

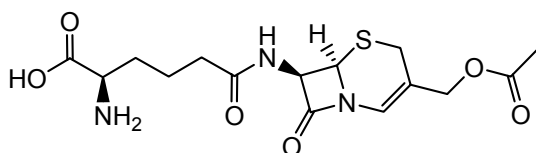
Advantages of natural product screening

- Molecules are structurally diverse
- Much precedence as a source of lead compounds.

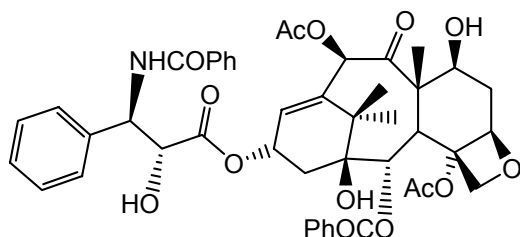
Problems with natural product screening

The mixtures are often very complex and contain many large macromolecules (e.g. carbohydrates, lipids, proteins etc). This range of products can often hide biological activity.

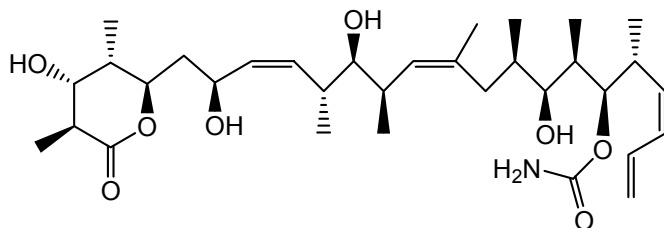
- Isolation of an active component present in a very small amount can be problematic given a large amount of background rubbish.
- Compound isolation and structure determination difficult
- Structures often complex, therefore difficult to synthesise and identify pharmacophore (the key structural element needed for a product to have activity).



Cephalosporin C
mould - antibiotic



Taxol
yew tree - antitumour



Discodermolide
marine sponge - antitumour
/ immunosuppressant

(pyro)Glu-Trp-Pro-Arg-Pro-Glu-Ile-Pro-Pro-OH

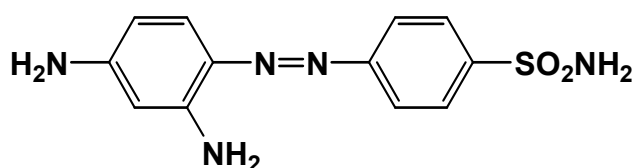
Teprotide
Venom from Brazilian viper - ACE
inhibitor (lowers blood pressure)

Some drugs which have been developed from natural compounds

Compound	Origin	Uses
Artemisinin	Sweet wormwood	Antimalarial derived from a traditional chinese medicine.
Acyclovir	Synthetic analogue of cytarabine from a marine source	Used to treat herpes infections.
Cyclosporin	Fungus	Used to prevent tissue graft rejection.
Digoxin	Foxglove	Digitalis has been used since 1775: digoxin remains an effective drug for heart failure.
Diosgenin	Mexican wild yam	Used in manufacture of steroidal contraceptives and in hormone replacement.
Etoposide	May apple	Synthetic analogue of podophyllotoxin. Used in chemotherapy to treat testicular and some lung cancers.
Galanthamine	Snowdrop	In trials for Alzheimer's disease.
Mevastatin	<i>Penicillium</i>	Used to reduce blood cholesterol levels.
Podophyllotoxin	May apple	Used against warts and skin cancers.
Pethidine	Synthetic analogue of atropine	Morphine-like analgesic.
Tirofiban	Synthetic analogue of snake venom peptide	A blocker of platelet aggregation used in angina.
Vinblastine and Vincristine	Periwinkle	Used to treat leukemias and lymphomas.

Dedicated Random Screening

Dedicated random screening is like natural product screening except that known, single compounds are tested at random in your biological assay / screen. It was in this fashion that Gerhard Domagk (IG Farbenindustrie, Germany, 1931) screened azo dyes (from the paint factory in which he worked) in the search for compounds with biological activity. This led to the discovery of the first truly effective sulphonamide antibacterial Prontosil (a red dye) in 1935.



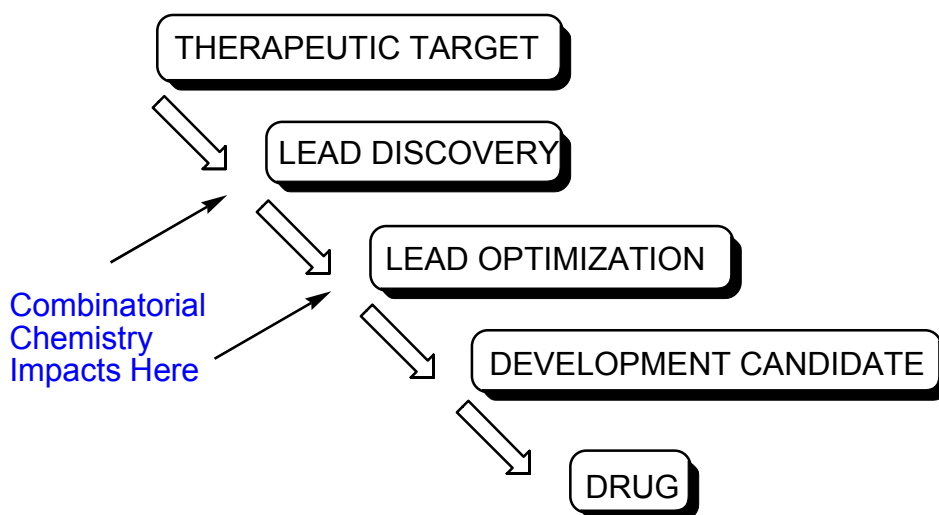
Compounds from past projects, University research collections can all be tested in a screen for the biological activity you are looking for in your particular project. Organisations such as Specs-BioSpecs (Holland, www.specs.net) and MDPI (Switzerland, www.mdpi.org) buy up compounds from Universities and other sources and make them available for company testing.

A '**compound library**' is a collection of compounds, just as we use 'library' for a collection of books. Like a book library the variety of compounds (diversity) in the collection can be very narrow (e.g. specialist Ferens library), very big but relatively narrow (Brynmor Jones, academic), big and diverse (city library). Compound libraries can therefore be the screened of lead compounds as often they have been made for totally different projects (e.g. antifungal agent naftifine was a by-product from chemical research into new agents for treating disorders of the central nervous systems).

A disadvantage of such libraries is that they are often quite small and therefore of **limited or restricted structural diversity**.

Combinatorial Chemistry

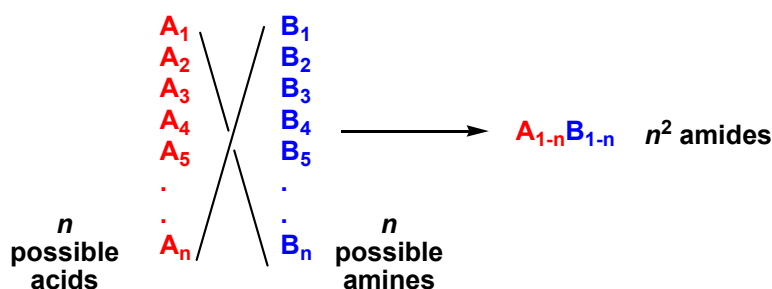
Obviously in searching for a compound with the desired biological activity, the more compounds we can screen the more likely we are going to find a useful lead compound. The technique of **combinatorial chemistry** ('CombiChem') has revolutionised the way chemists go about searching for lead compounds and also at the stage of lead optimization.



The essence of combinatorial chemistry (CombiChem) is the generation of large numbers of compounds very quickly.

The traditional method of medicinal chemistry concerned the preparation, purification and characterization of individual compounds. Referring back to the synthesis-test cycle we can see that if this process is speeded up then faster progress can be made. Combinatorial chemistry has revolutionized the way in which medicinal chemists in industry and academia go about making their compounds. After all only a few milligrammes are needed for a first run biological *in vitro* assay.

Consider the synthesis of a dipeptide by coupling of two amino acids (formation of an amide bond). There are twenty 'DNA encoded' amino acids and so there is the possibility of forming 20^2 dipeptides using the twenty different 'monomer' units. A hexapeptide could give rise to 64 million different variations using the 20 building blocks! To synthesize these "long-hand" would take a very long time.



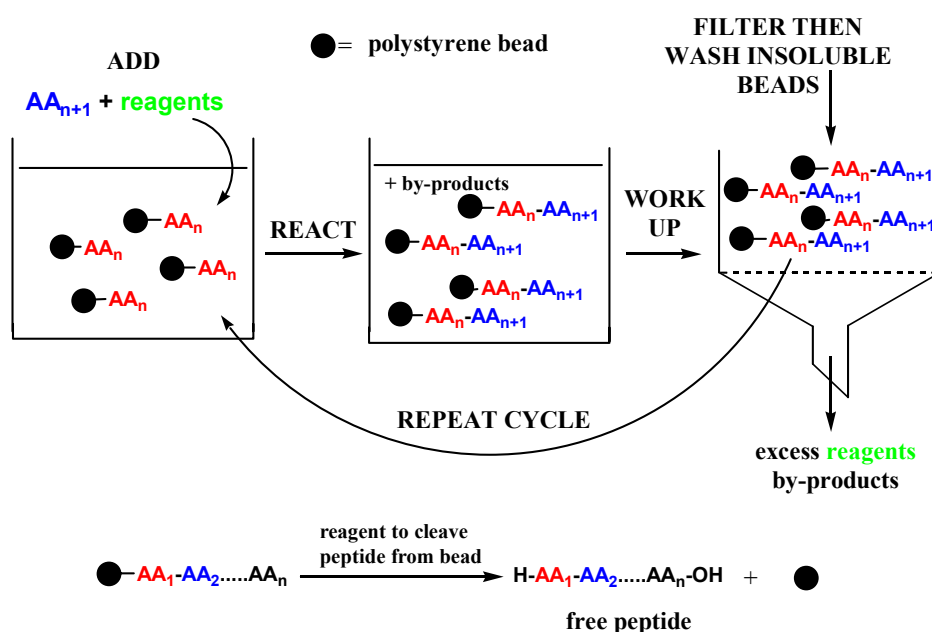
If we wanted to synthesize and test all these peptides, thus making a “library of amides”, how could we get around the laborious step-wise preparation of days gone by? How about making them all at once using CombiChem!?

Combinatorial Chemistry – consists of two fundamental approaches:

- A:** solid phase synthesis
- B:** solution phase parallel synthesis.

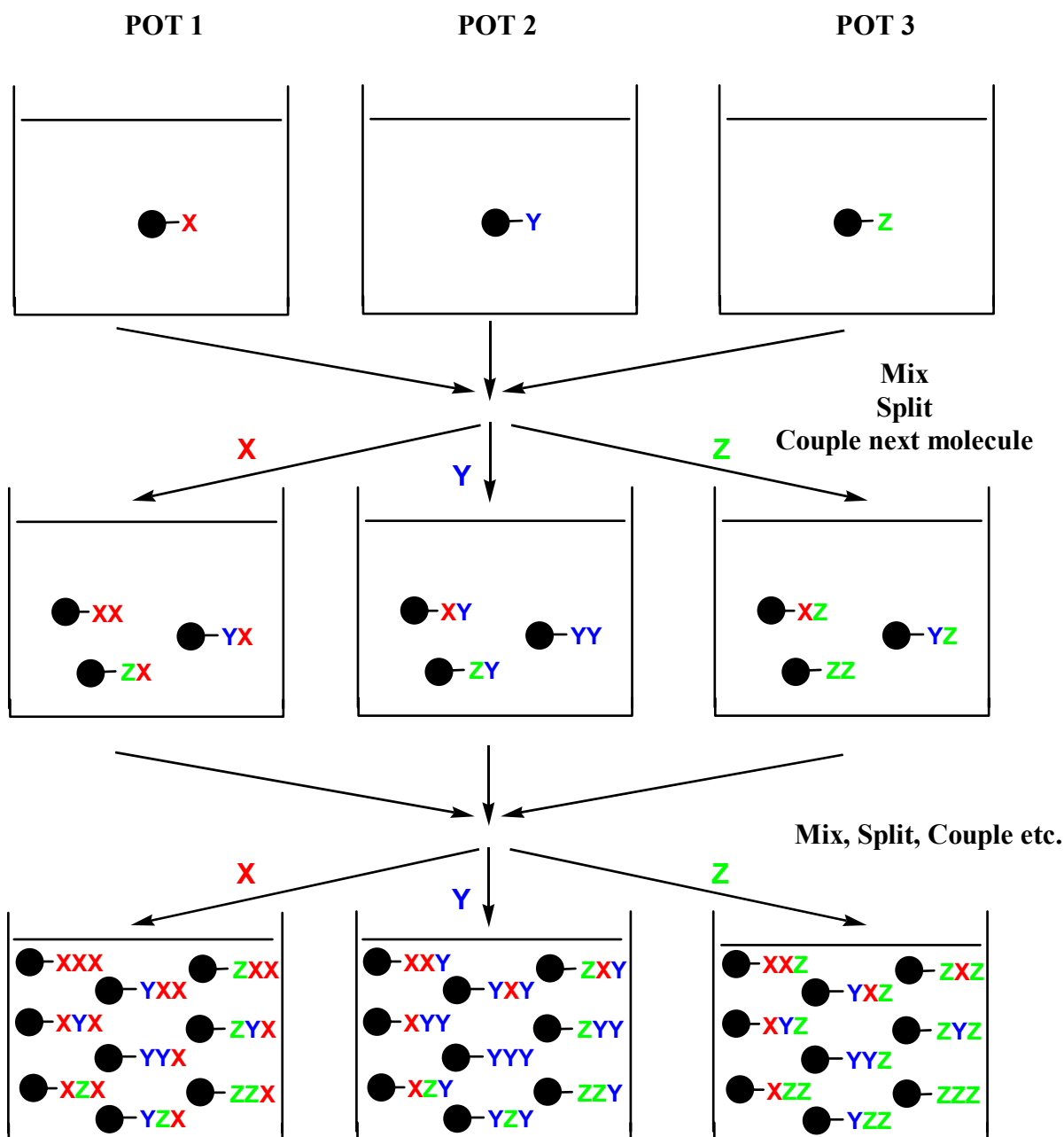
A: Solid Phase Organic Synthesis (SPOS)

SPOS has roots in R.B. Merrifield’s synthesis of peptides attached to microporous polystyrene beads (1963). The essence of SPOS is to attach your substrate to a solid "support" (known as beads or resin) rendering the reactions heterogeneous. This aids the work up and purification of your reaction products:



Merrifield Solid-Phase Peptide Synthesis (1963)

Attaching the substrate to a solid phase allows simple purification by simple filtration of the solution containing by-products and the solid bead. This process for making peptides works very well and is now an automated process.

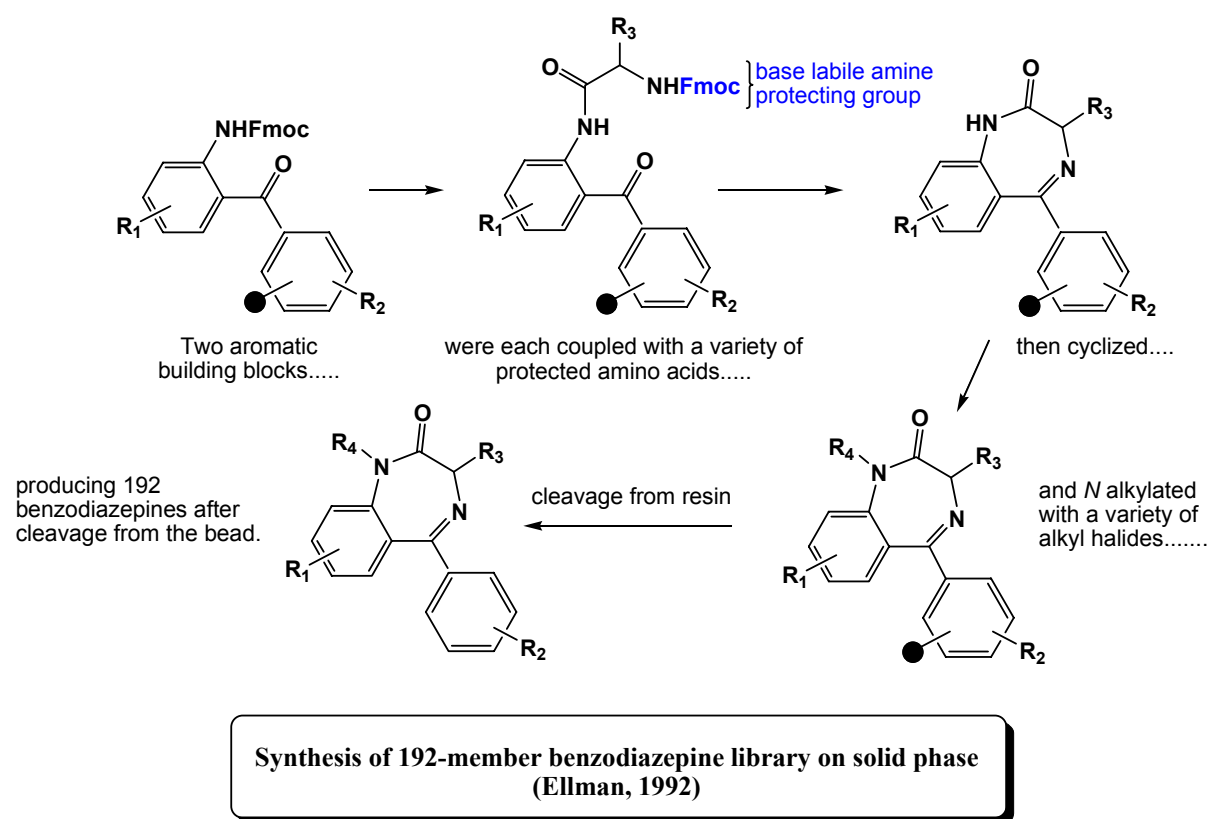


Split-Mix Synthesis (Furka, 1988)

The split-and-mix method (also known as the divide-couple-recombine (DCR) method) allows for the sequential reaction of compounds attached to beads, in which each bead contains only one type of compound.

This approach can be used to make compounds very rapidly, but there is the major problem of identifying which bead has which compound attached. Some ingenious methods have been developed to work out this problem (“deconvolution”) but these are outside the scope of this course.

Using this split-mix approach Jonathan Ellman (University of California, Berkeley) produced a 192- member library of benzodiazepines illustrating the wider application of SPOS to any organic compounds rather than just peptides.



Problems with this method–

- Quantities produced can be very low for large libraries (10s of nanomoles).
 - poses problems with full representation of molecules in very large libraries – not enough beads unless very large quantities (kgs) are used!
- Synthesis development
 - solution phase methods don't always smoothly apply to the solid phase
- Residual functional groups
 - diversity is limited in that we must always have a functional group (acid, alcohol, amine) with which we can attach or molecule to the bead.
- Characterisation of intermediates difficult
 - How can we tell if our reaction has worked?
- Deconvolution – how do you know *what* compound is attached to any one bead?

Advantages –

- handling of material is easy
- can be automated
- purification is easy - simple washing and filtration is usually all that is needed.

A more recent approach is to use supported *reagents* rather than *substrates*, with many of the advantages listed above, but with very few of the disadvantages.

B: Parallel Solution-phase (Array) Synthesis

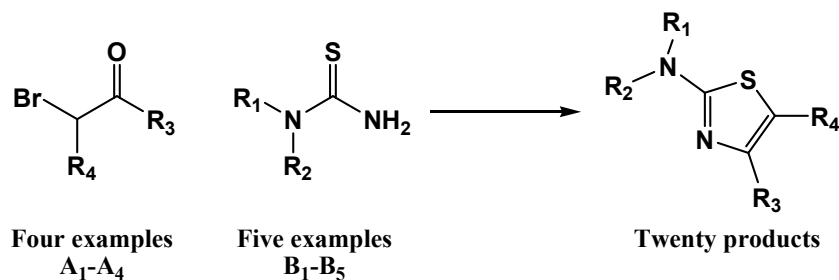
Obviously there are disadvantages to solid-phase library generation. A newer 'fashion' is now to use solution-phase combinatorial synthesis. New equipment, such as 'personal synthesizers' (e.g. Radley's reaction stations) and 'multi vial apparatus,' allows for parallel (at the same time) synthesis of many compounds simply and quickly by one chemist. These techniques lend themselves to robotized technology (e.g. Argonaut synthesizers) such as in the peptide synthesis seen above.

Solution-phase synthesis has advantages in that:

- No development work needed (compared to SPOS).
- Easy to identify active hit as its position (X,Y coordinate) in the array encodes the reagents and thus structure of the product.

Disadvantages include

- A build up of impurities can occur unless the by-products are volatile and the reactions very clean and high yielding.
- Most useful for one or two steps reactions.



Synthesis of Aminothiazoles via the Hantzsch Method (Glaxo-Wellcome, 1996)

	B_1	B_2	B_3	B_4	B_5
A_1	A_1B_1	A_1B_2	A_1B_3	A_1B_4	A_1B_5
A_2	A_2B_1	A_2B_2	A_2B_3	A_2B_4	A_2B_5
A_3	A_3B_1	A_3B_2	A_3B_3	A_3B_4	A_3B_5
A_4	A_4B_1	A_4B_2	A_4B_3	A_4B_4	A_4B_5

Analysis by automated
HPLC-MS

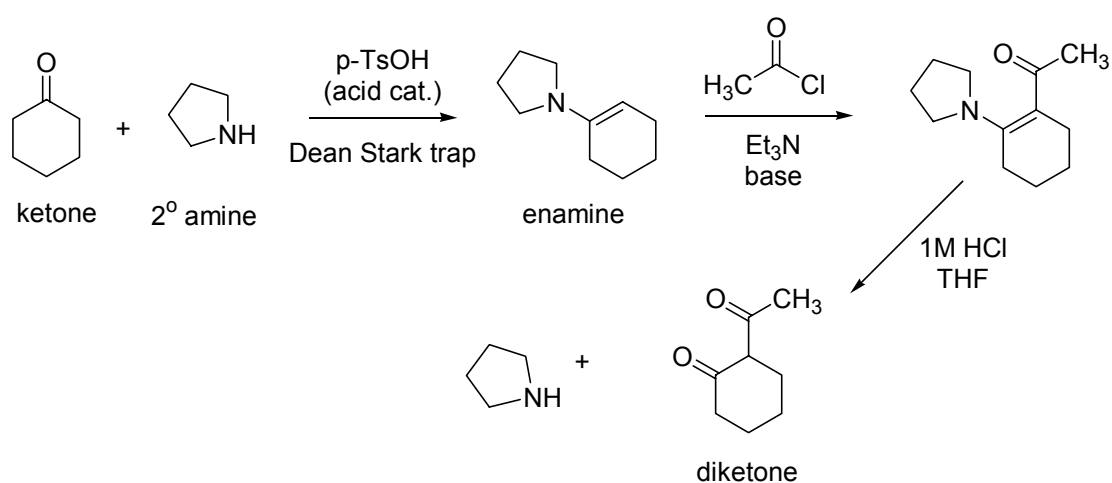
In a 'proof of concept' study the reliable Hantzsch synthesis was used by workers at GlaxoWellcome to prepare twenty thiazoles including fanetizole [$R_1 = \text{Ph}(\text{CH}_2)_3$; $R_2 = \text{H}$; $R_3 = \text{Ph}$; $R_4 = \text{H}$] a known anti-inflammatory agent. The compounds were prepared using robotic system. (70 °C, 5h in DMF; add diethylamine, solvent removed by nitrogen stream over 24 h! Structure analysis was performed by HPLC-mass spec.)

The impact of CombiChem on drug discovery is obvious. These techniques accelerate the process of new lead identification (through larger, more diverse libraries), but has found to be more of use in lead optimization using smaller or ‘directed’ (focussed) libraries.

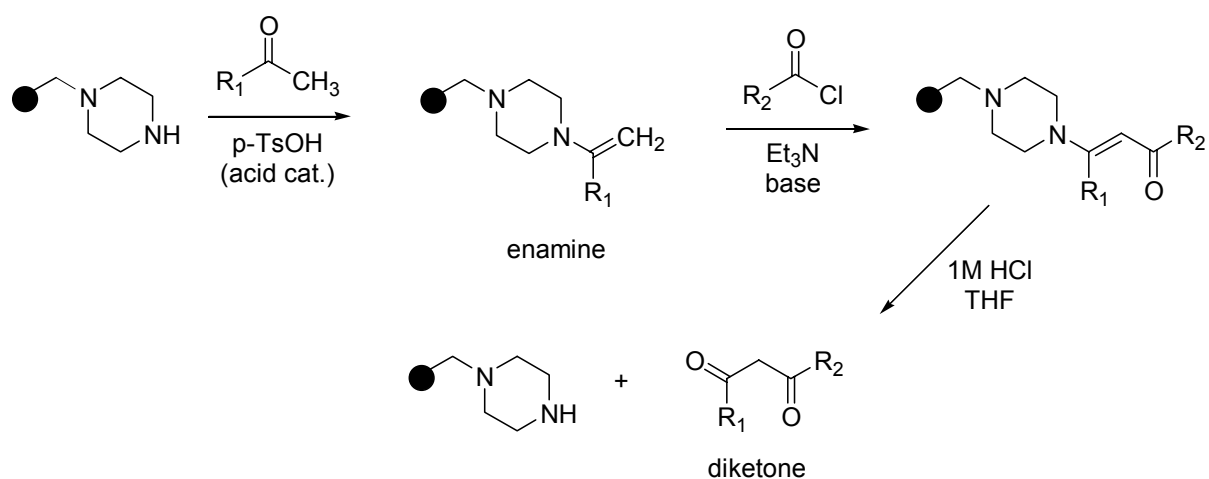
Designing a simple compound library: Pt 1

E.g. synthesis of a library of diketones (*Tetrahedron Lett.*, 2003, **44**, 1067-1069). This shows how a simple reaction may be “modified” to take advantage of supported reagents and intermediates allowing the synthesis of a range of diketones.

The solution phase reaction (c.f. Yr 2 Sem 2 Org Lab experiment).



The solid-phase reaction using a supported secondary amine

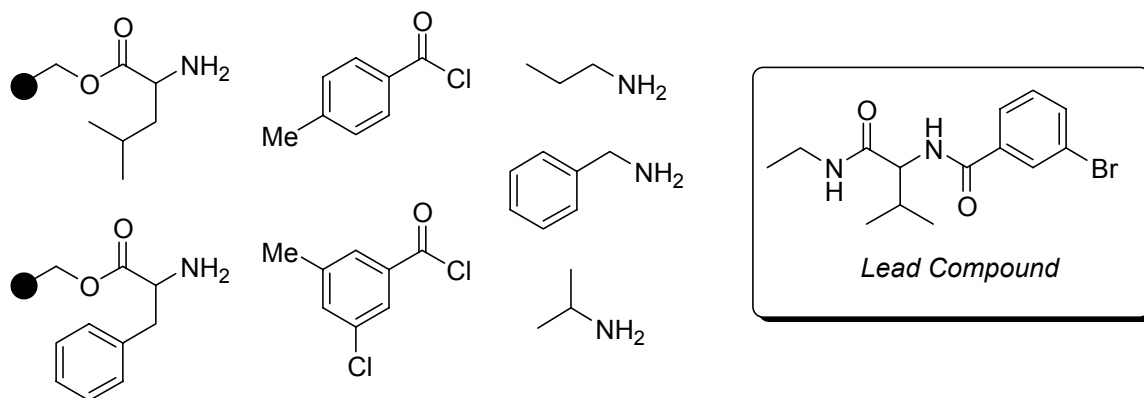


Using a variety of combinations of methyl ketones (step 1) and acid chlorides (step 2) a library of unsymmetrical 1,3-diketones can be made.

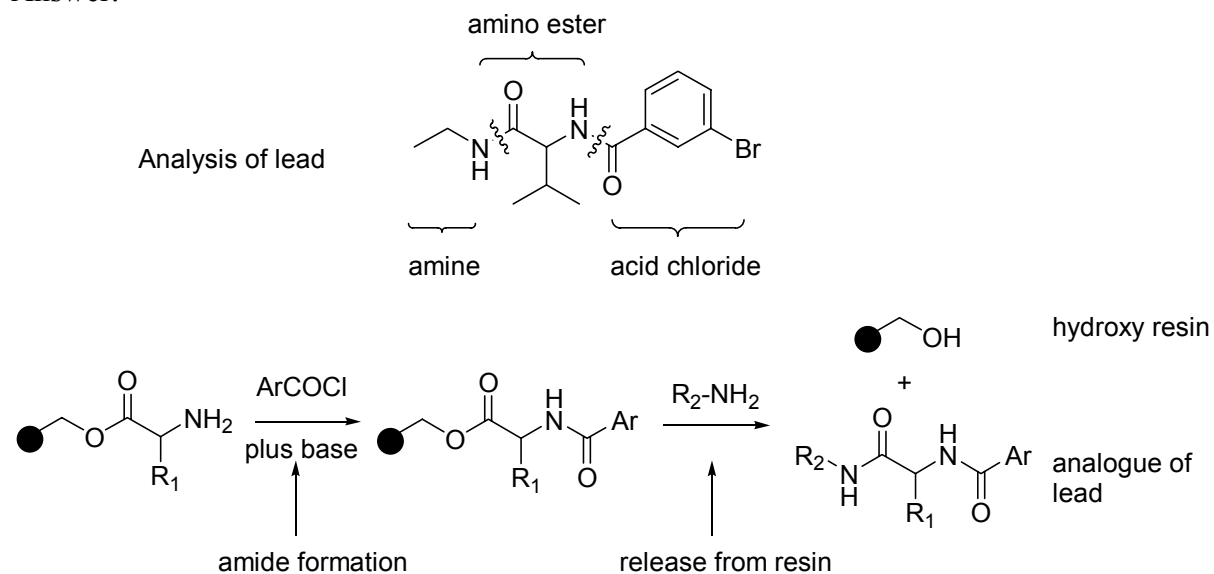
Designing a compound library: Pt 2

Synthesis of a library of lead compound analogues.

Question: Using the building blocks below, illustrate how you can construct a small library of analogues of the *lead compound* given using the Split-Mix Synthesis technique (see p. 20)

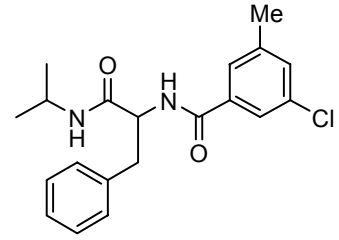
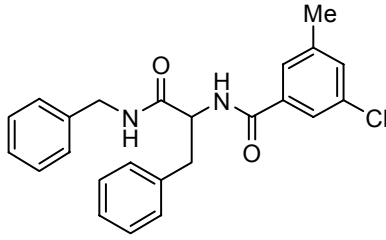
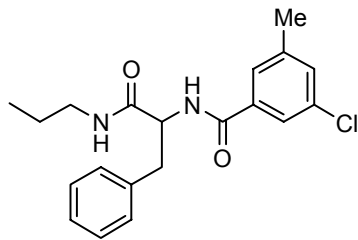
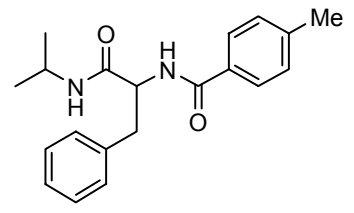
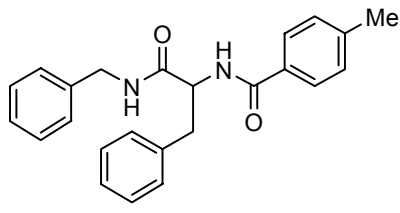
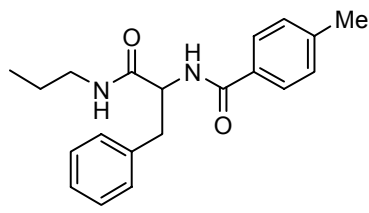
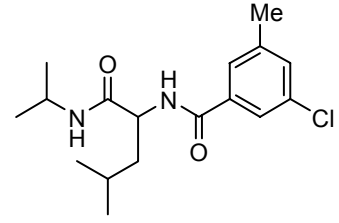
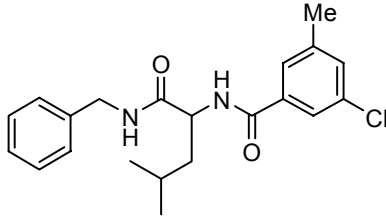
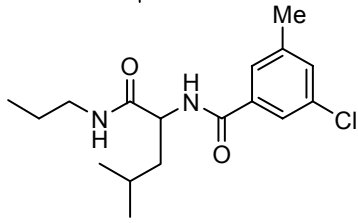
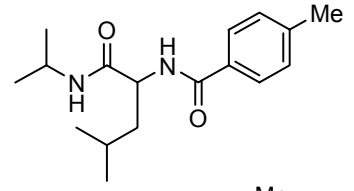
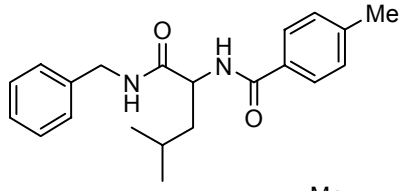
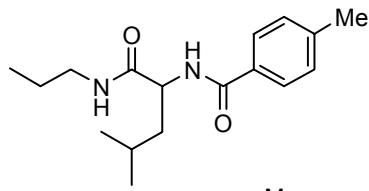


Answer:



With two amino esters and two acid chlorides a total of four intermediates could be made. Release from resin with one of three amines gives a total of 12 analogues.

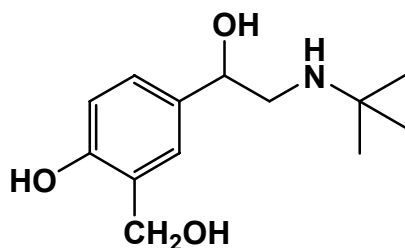
Library members



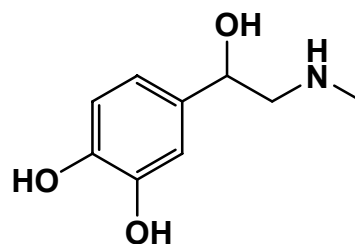
Substrate-Based Drug Design

Substrate analogues

As the name implies, the natural molecule (substrate) for a receptor or enzyme can serve as a lead. An example is salbutamol, an analogue of adrenaline, which stimulates adrenergic β_2 -receptors and is used to treat asthma. Salbutamol is said to act as an **agonist** of adrenergic β_2 -receptors in that it elicits the same response as the natural compound (an **antagonist** however blocks the receptor/enzyme without eliciting the physiological response).



Salbutamol



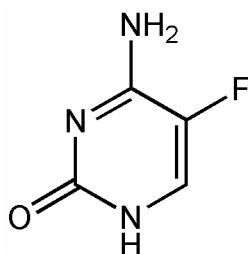
Adrenaline

With the recent advances in the biological sciences more and more natural transmitters and hormones have been identified which in due course may serve as useful lead compounds. Other examples can be found in '*Medicinal Chemistry, Principles and Practice*', Ed FD King, Royal Society of Chemistry, 1994, p 194, **RS 403 M4**)

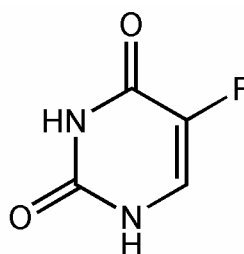
Nucleoside analogues

Suicide (irreversible) inhibitors are inhibitors that often, but not always, resemble the natural substrate but which contain functional groups that irreversibly bind to the enzyme or receptor active site [containing "reactive" functional groups, these molecules can however pose a question of possible toxic side effects unless they are very specific].

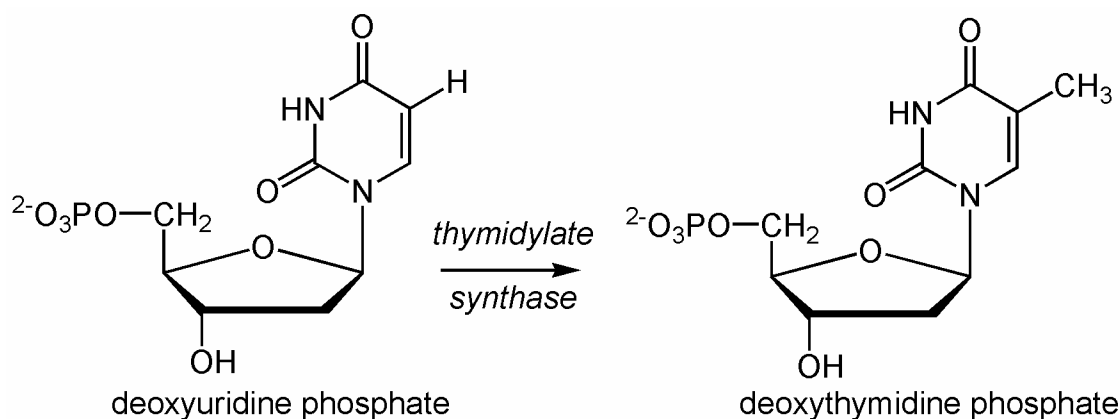
An example of such a drug is 5-fluorouracil (5-FU) which blocks the action of thymidylate synthase. This enzyme is essential for making the DNA building block thymidine monophosphate. 5-Fluorocytosine (5-FC) is an antifungal agent which is converted to 5-FU in *vivo*. The mode of action is then the same as 5-FU.



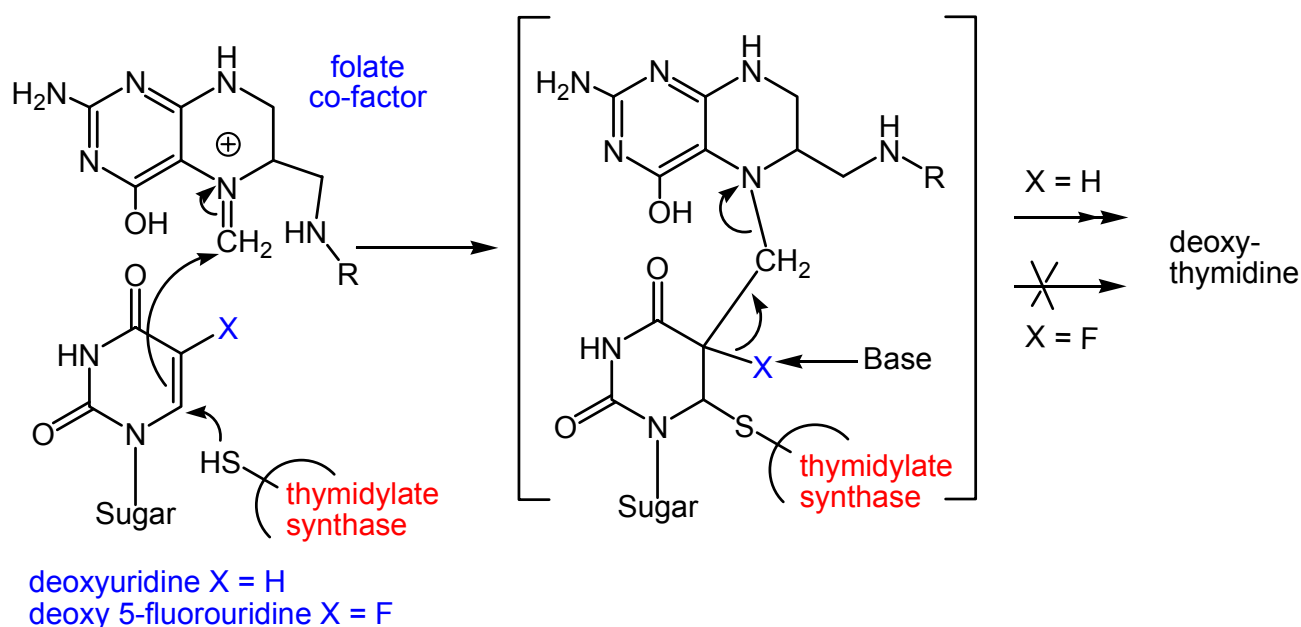
5-FC



5-FU



The presence of a fluorine atom on the heterocycle means that the enzyme, once covalently bound to the drug, cannot be recycled. So DNA production is halted due to a lack of thymidine.



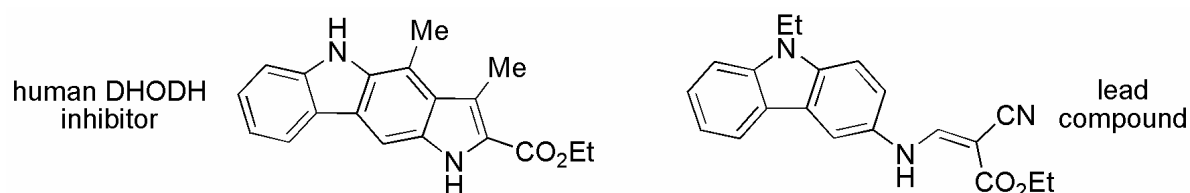
Rational Drug Design

With the advances in molecular biology techniques (for making and isolating “large” amounts of proteins) and X-ray crystal structure solving techniques, obtaining X-ray crystal structures of proteins and receptors is becoming common place. The Protein Data Bank (see <http://www.rcsb.org/pdb/home/home.do>) has data for hundreds of published structures which are all freely available on the WWW. Coupled with advances in computing power and molecular modelling (e.g., see the Tripos company web site: <http://www.tripos.com/> for information on some state of the art software), so called *rational drug design* has been advanced as “the way forward” in the search for new drugs.

Example: Inhibitors of dihydroorotate dehydrogenase (DHODH).

DHODH is an enzyme involved in making the heterocyclic portion of uridine monophosphate (see before). DHODH is a new antimalarial drug target. A screen of inhibitors of human

DHODH against the parasite DHODH came up with a hit which was developed into a lead compound (in this department).



The X-ray crystal structure of the parasite enzyme is known so a computational study has allowed the “docking” (superposition) of the lead structure into the active site of the enzyme. This study is directing optimization of the inhibitor structure through determination of the intermolecular forces between enzyme and inhibitor.

The crystal data is freely available six months after publication. If you browse to the PDB web site above and type in “dihydroorotate dehydrogenase” into the search box (NB select the ‘structure title’ option from the drop down menu first), you are directed to a page giving the details of the different X-ray crystal structures of the enzyme, dihydroorotate dehydrogenase (the target enzyme). There are examples of structures of the enzyme from humans, bacteria, parasites and some even complexed with compounds (drug or drug-like molecules). Not only can pictures of these proteins be viewed, but atomic coordinates for the protein be downloaded and used in sophisticated molecular modelling packages. X-ray structural data like this can be used to develop “*in silico*” new and more selective drugs based on the lead.

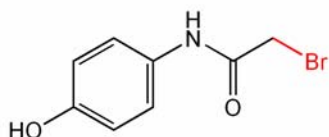
Having said that, no examples exist (yet) of drugs having been designed using these techniques alone. Chemo- and bioinformatics techniques have become very useful tools in *accelerating* the drug design and development process.

Assessing the Quality of the Lead

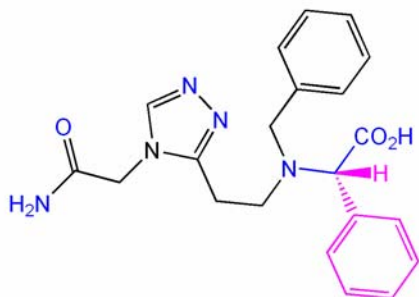
Assuming your compound has the desired potency and selectivity from your assay and it is effective *in vivo*, you should consider the following:

- Is the molecule easy to synthesise and is it amenable to chemical modification? Does it have awkward chiral centres? Is it difficult or lengthy to synthesise?
- How soluble is it? Could formulation (i.e. converting the drug into the suitable pill, gel, cream suitable for treating the condition) be a problem?
- Is it free from structural elements that are likely to engender toxicity? E.g. alkylating agents, Michael acceptors
- Does it resemble competitors’ compounds? Will there be patent problems?

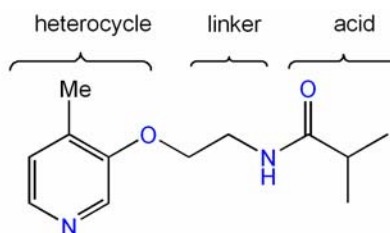
Some examples of possible ‘lead molecules’ and an assessment, considering the factors above, of their properties are given below.



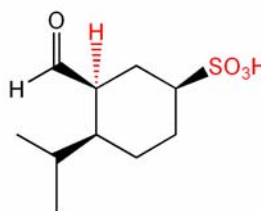
alkylating agent > toxicity problems!
(also Michael acceptors, C=C-C=O unit)



Chirality easily available from amino acid



simple structure
obvious synthesis:> easy to make analogues
H-bonding possibilities > water solubility



no obvious synthesis / source of chiral starting materials
may readily epimerise α - to aldehyde
sulfonic acid bad for oral absorption

(taken from 'Medicinal Chemistry, Principles and Practice', Ed FD King, Royal Society of Chemistry, 1994, p 186/7, **RS 403 M4**.)

Other well known groups which engender toxicity are planar three-ring aromatic systems. These intercalate (bind) well with DNA and so can cause interference with its function. Aromatic nitro or aromatic amines are also known to cause problems. These can be converted *in vivo* into aromatic hydroxylamine and nitroxide groups that can self-condense to form toxic nitrogen-containing heterocycles. These types of well-known "bad" groups are sometimes referred to as "toxicophores".

It is highly desirable that the difference between the effective dose (ED_{50}) and the toxic dose (LD_{50}) of a drug *in vivo* is generally a minimum of a 1000-fold. This ratio, ED_{50}/LD_{50} , is known as the "therapeutic index".

Lipinski's Rule of Five

Reference: C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings", *Adv. Drug Delivery Rev.*, 1997, 23(1-3), 3-25 (available from <http://www.sciencedirect.com> if on campus). Lipinski proposed four parameters that define the "drug-likeness" of potential drug candidates based on analysis of existing drug molecules. "The Rule of Five" got its name from the cut-off values for each of these parameters of which all have values of five or a multiple of five.

".... We found that the sum of Ns and Os in the molecular formula was greater than 10 in 12% of the compounds. Eleven percent of compounds had a MWT of over 500. Ten percent of compounds had a $ClogP$ larger than 5 (or an $MlogP$ larger than 4.15) and in 8% of compounds the sum of OHs and NHs in the chemical structure was larger than 5.

The “rule of 5” states that: poor absorption or permeation is more likely when:

- A. There are more than 5 H-bond donors (expressed as the sum of Ohs and NHs);
- B. The MW is over 500;
- C. The Log P is over 5 (or Mlog P is over 4.15);
- D. There are more than 10 H-bond acceptors (expressed as the sum of Ns and Os).”

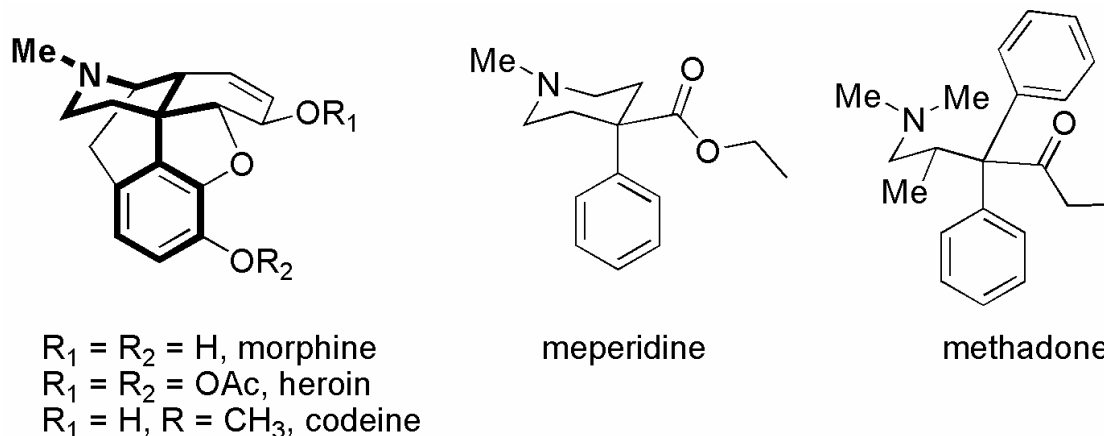
In rule C the parameter P is the water - octanol partition coefficient. A partition coefficient is the equilibrium concentration of solute in a non-polar solvent divided by the concentration of the same species in a polar solvent; i.e. a measure of hydrophilicity/phobicity of a molecule. This is an important factor in determining the bioavailability of a drug molecule.

(Mlog P is a molecule-based descriptor which describes an estimation of the log of the octanol-water partition ratio developed by Moriguchi (Moriguchi, I. *et al.* Chem. Pharm. Bull. 1992, 40, 127-130).

A continual process of assessing compound properties takes place in a drug discovery programme. A lead compound may not necessarily become a drug candidate. Lead optimization can be as key a process as lead identification.

Identification of a pharmacophore

We have defined a lead compound as “a compound from a series of related compounds.....”. The question is therefore posed what are the essential structural elements for biological activity? A pharmacophore is “a set of structural features in a molecule that is recognized at a receptor site and is responsible for that molecule's biological activity”.



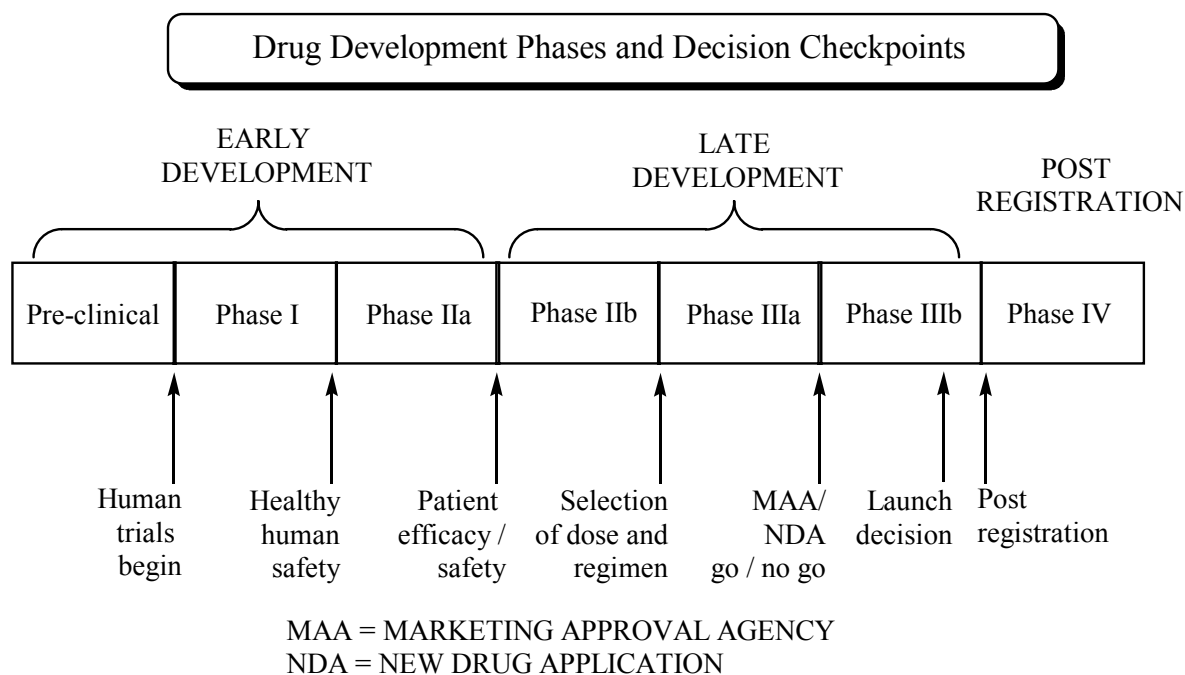
In this example the pharmacophore from the opioid drugs is highlighted in bold. Morphine, heroin and codeine have analgesic but also addictive properties, whereas meperidine has the analgesic properties (but only about 10% that of morphine) but very little addictive side effects. Methadone is as potent as morphine, and is used as a treatment for patients trying to give up heroin addiction.

Once a pharmacophore has been identified as series of related compounds must be made to improve potency and reduce toxicity. Determining a structure-activity relationship (SAR) is

the process by which chemical structure is quantitatively correlated with biological activity. When physicochemical properties or structures are expressed by numbers, one can form a mathematical relationship, or quantitative structure-activity relationship (QSAR), between the two.

These topics and more are covered in the continuation module 06508 “Drugs: from design to delivery”.

Clinical Trials



Clinical trials are designed to: determine safety and tolerance in man; pharmacokinetics (what the body does with the drug); bioavailability for a range of doses; determine the pharmacological profile.

- Pre-clinical animal studies. Submission of “Investigational New Drug” application to government bodies such as US FDA.
- Phase I normal, healthy human volunteers.
- Phase II to evaluate safety in efficacy of drug in patients.
- Phase III large patient number study to establish efficacy versus a placebo or comparator compound.
- Phase IV long-term surveillance / monitoring of adverse reactions.