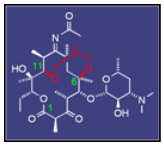
Chemistry 259

Medicinal Chemistry of Modern Antibiotics

Spring 2008



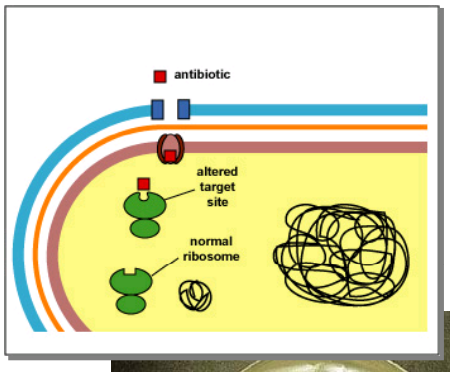
Lecture 3: *Drug Discovery, Development & Approval*
Part I

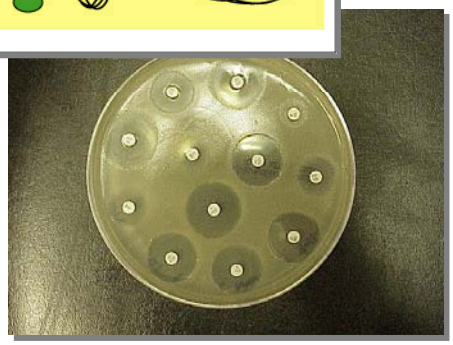
Thomas Hermann

Department of Chemistry & Biochemistry
University of California, San Diego

Drug Discovery & Development Process: What it Takes

Target based



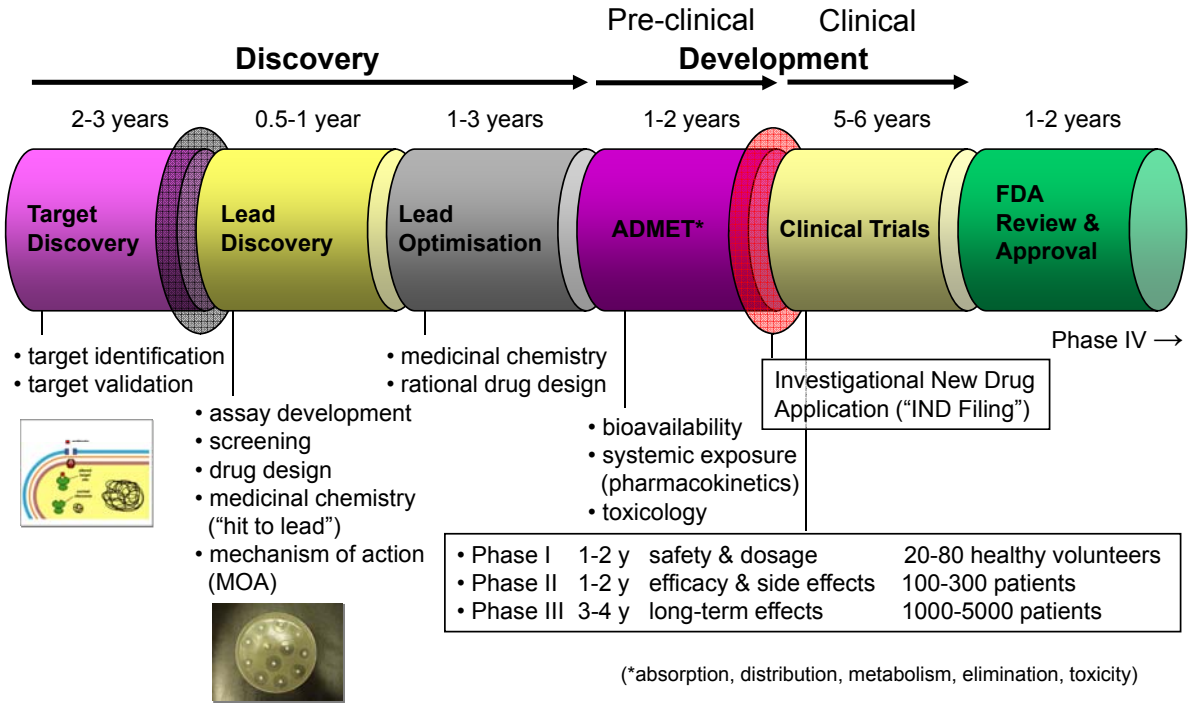


?

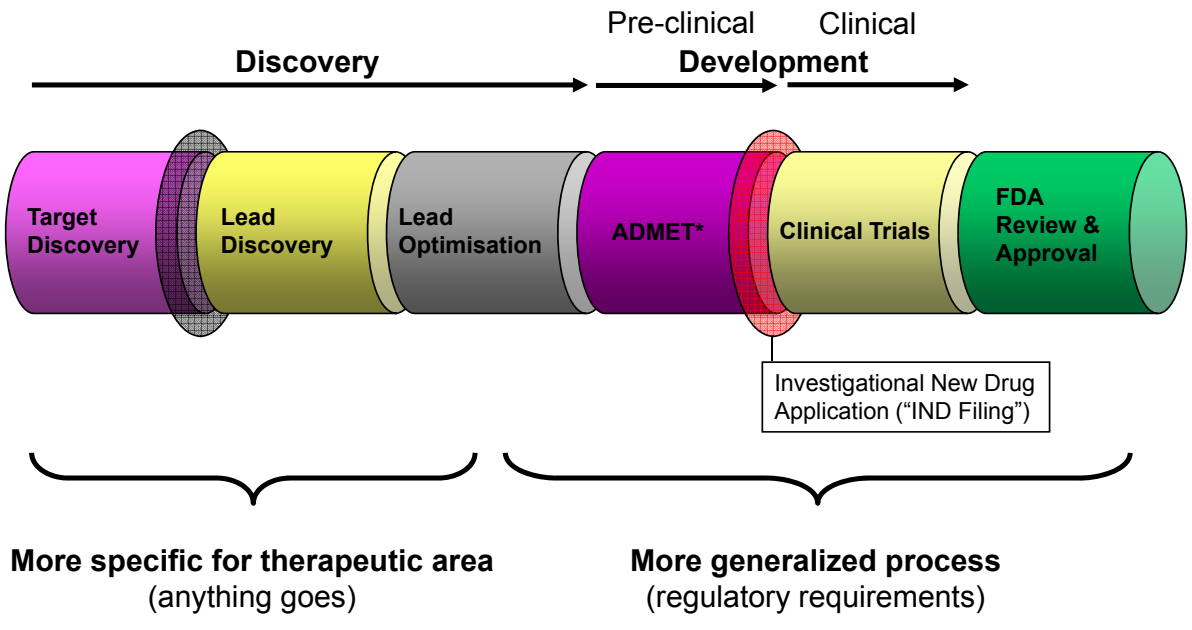
Screening based

Drug Discovery & Development Process: General Overview

- typically a 10-17 year process to get to market
- < 10% overall success rate



Drug Discovery & Development Process: General Overview



More specific for therapeutic area
(anything goes)

More generalized process
(regulatory requirements)

Antibacterial Target Discovery: Genomics

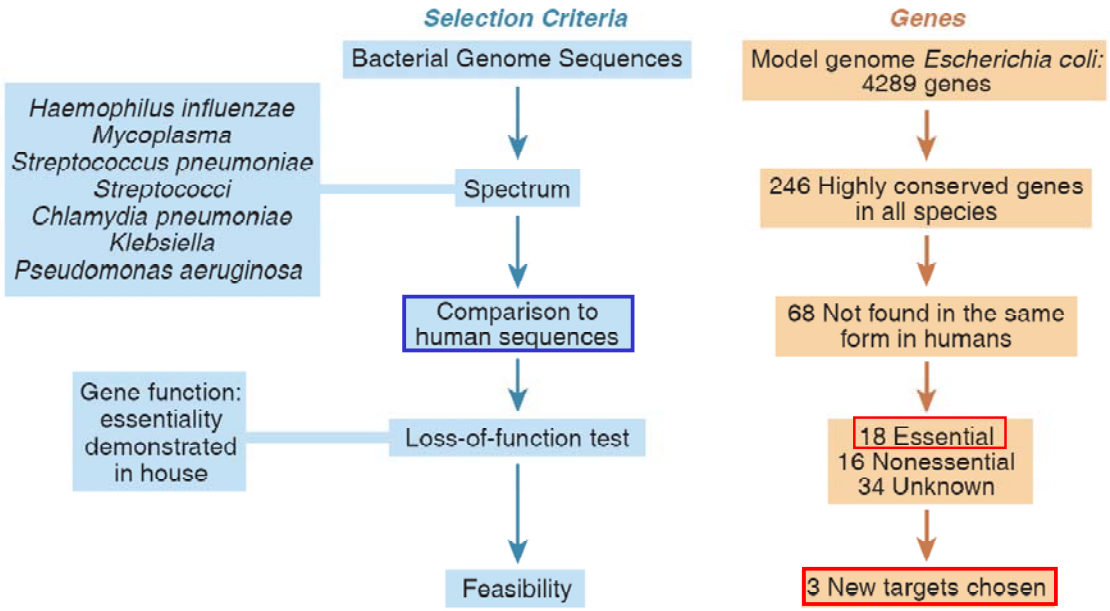
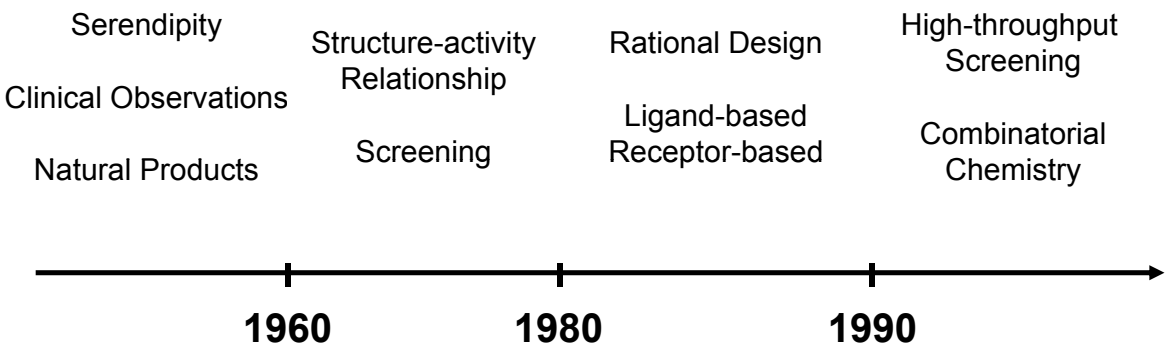


Fig. 3. An example of a target identification cascade for a respiratory tract antibacterial drug.

Antibacterial Lead Discovery: Historical



Antibacterial Lead Discovery: Screening

Table 1. Comparison of the screening strategies for novel antimicrobial compounds.

Whole-cell screening (looking directly for compounds which kill microorganisms)	Target-based screening (looking for biochemical inhibitors)
Advantages	
Selection for compounds which penetrate cells	More sensitive (can detect weak or poorly penetrating compounds suitable for chemical optimization)
Antimicrobial properties established	Easy screening
Highly reproducible	Different approach
Has been used successfully historically	Can target new areas of biology
	Facilitates rational drug design
Disadvantages	
Insensitive	Need to turn an in vitro inhibitor into an antibacterial drug (complicated by penetration issues)
Most active compounds are toxic	Genetic validation of targets (by gene knockout or reduced expression) can be misleading
No rational basis for compound optimization (target unknown)	
Mixed mechanisms of action	
In recent years has failed to deliver	

Lead Discovery: Screening

Natural sources (soil, plant extracts, etc.) and combinatorial chemistry provide a large number of molecules that can be tested by automated high throughput screening systems.

Sequencing of genomes may open new prospects to these techniques as new potential targets will be discovered.

Screening contributed to the discovery of many valuable leads; however, with automated high-throughput screening, the situation is more complex →

Lead Discovery: HTS

High Throughput Screening (HTS) established as a routine method around 1995.

Based on the use of robotics to screen large libraries of compounds onto an isolated target, a cell or a tissue so as to identify the molecules that are able to bind (affinity screen) or elicit a biological effect (functional screen).

The more advanced techniques enable to screen 100,000 compounds per day.

HTS depends on the development of quantitative tests which are pharmaceutically significant and adapted to the target and which can be reproduced on a large number of samples.

384 well plate

Lead Discovery: HTS

Drugs that evolved from structures discovered through HTS:

- nevirapine, delavirdine , efavirenz (HIV non-nucleoside RT inhibitors)
- bosentan (Tracleer, endothelin receptor antagonist; pulmonary arterial hypertension)
- gefitinib (Iressa, tyrosin kinase inhibitor; antineoplastic, lung cancer)

Companies are now aware that the original concept does not deliver to the expected extent.

Limited solubility, deposition after dilution, compound decomposition, as well as unknown concentrations, coloured impurities, fluorescence of some compounds, etc., produce false negatives and false positives.

In many cases, re-testing does not confirm any primary hits.

In other cases, re-testing of analogs uncovers their activity, although they were initially found to be inactive.

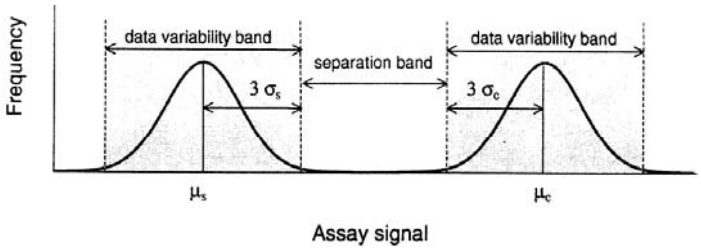
Re-testing is time-, labor- and cost-intensive.

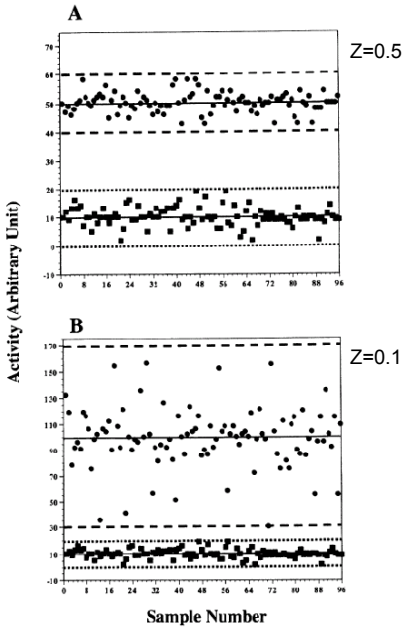
Lead Discovery: HTS

HTS Assay Validation: Z Score (Zhang et al. *J Biomol Screen.* 1999, 4, 67)

$$Zfactor = 1 - \frac{3 \times (\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$ (mean (μ) and standard deviation (σ) of the positive (p) and negative (n) controls)

Z-factor	Interpretation
1.0	Ideal. Large dynamic range with small standard deviations. Z-factors can never actually equal 1.0 and can never be greater than 1.0.
0.5 – 0.99	Excellent assay.
0 - 0.5	Marginal assay.
< 0	The signal from the positive and negative controls overlap, making the assay essentially useless for screening purposes.





Lead Discovery: Compound Sources

- Natural product libraries
- Existing compound libraries
- Combinatorial chemistry libraries
- Virtual libraries

Lead Discovery: Natural Product Libraries

- Natural product libraries continue to be an important source of lead compounds for drug discovery.
- Extracts of organisms from various sources are typically fractionated into samples containing just a few compounds per fraction.
  - Plant extracts
  - Marine organisms
  - Animal toxins
    - Cone snails
    - Snake and spider venoms
    - Frog and toad skin toxins and antimicrobials
    - …
- If a fraction has evidence of biological activity, it is characterized in more detail to identify the structure of the compound with biological activity.
  - Mass spectrometry, NMR, x-ray crystallography

Lead Discovery: Existing Compound Libraries

- Most pharmaceutical companies have large libraries of compounds (10^4-10^6) that have been generated by their medicinal chemists over the years (“legacy compounds”).

- Many smaller companies specialize in synthesis of custom libraries and distribution of legacy libraries of various origin (academic, …).

Lead Discovery: Combinatorial Chemistry

Synthesis of a large number of compounds, a library, combining in a systematic way, the representatives of two or more families of building blocks.

N aldehydes + M amines --> NxM products

First applied in 1963 when Merrifield performed sequential synthesis of a tetrapeptide.

Method was then extended to the organic, organo-metallic and inorganic chemistry with industrial applications in pharmacochemistry, catalysis, material sciences, dyes.

Feeds the HTS monster.

But …

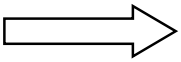
Lead Discovery: Combinatorial Chemistry

Even more disappointing than HTS results was the success rate of combinatorial libraries, especially in the early years.

Huge libraries of ill-defined mixtures of most often lipophilic and too large compounds were tested, without any positive result.

The hit rate of libraries generally decreases with an increase in the number of “over-decorated”, i.e. too large and too complex molecules.

Successful only after introduction of rules for drug-like properties. (⟶)

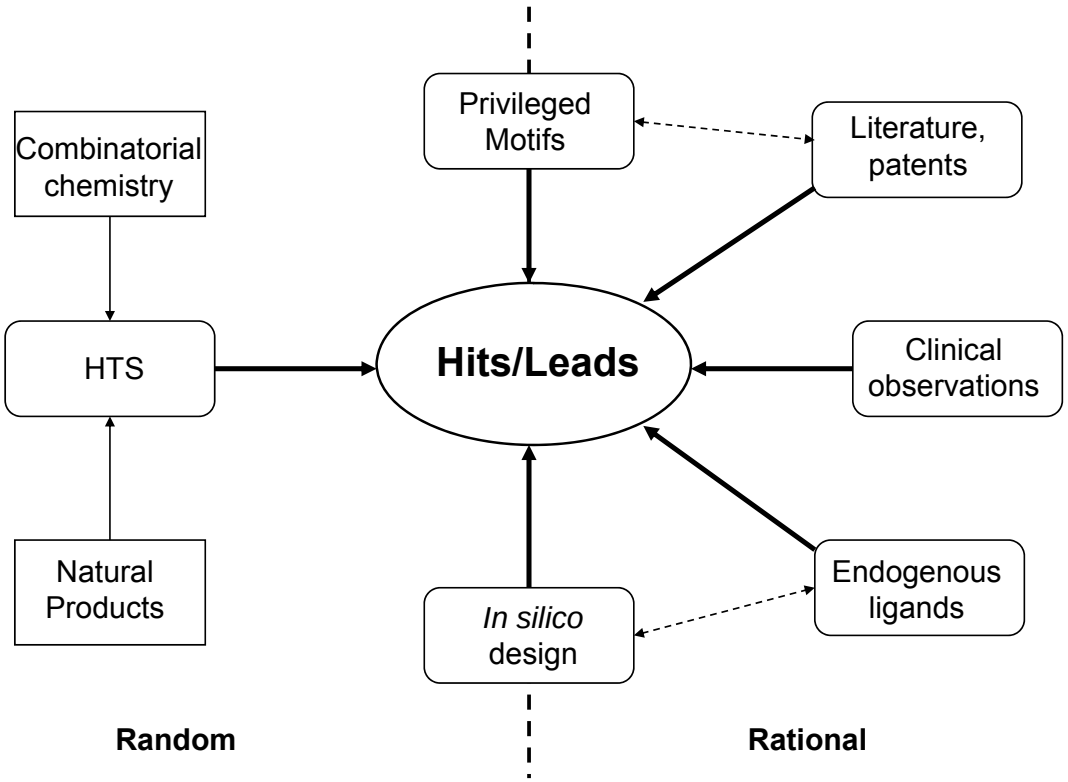


Change strategies in the synthesis of libraries.

Automated parallel synthesis of much smaller libraries of single and pure (or purified) compounds (often as “focused library”).

Today its main application is not so much in lead structure search but in lead validation and in the early phases of lead optimisation.

Lead Identification



Properties of a Lead

• Desirable biological activity, although it may be weak and even non-selective.

• Related analogs, indicating that structural modification will modulate biological activity as well as other properties ("SAR-ability").

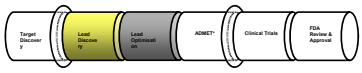
These properties might be improved during optimization:

• Absence of toxic groups or groups that will produce toxic metabolites ⟶

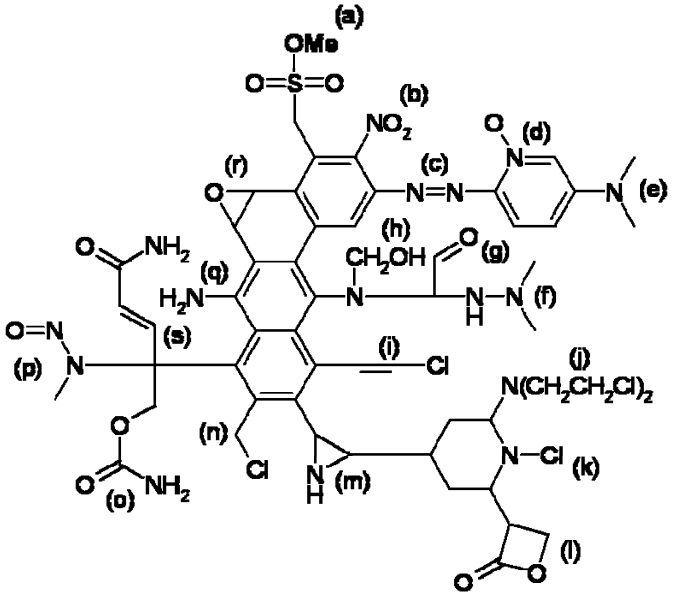
• Physicochemical Properties ⟶

• "Drug-likeness" ⟶

Toxic “Red Flags”

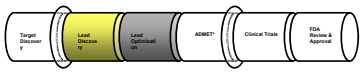


The groups illustrated on this theoretical molecule are associated with genotoxicity.



(t) - halogenated methanes $C(X)_4$, where X=H, F, Cl, Br, I

Physicochemical Properties



• Lipophilicity (hydrophobicity) – key parameter linking membrane permeability with the route of clearance. Hydrophilic compounds generally show poor permeability and hence low absorption, whereas highly lipophilic compounds exhibit poor aqueous solubility and slow dissolution rates, leading to poor oral absorption.

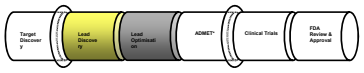
– logP – partition coefficient in octanol/water
– logD – distribution coefficient in body
(in aqueous buffer at pH 7.4 (blood pH) or 6.5 (intestinal pH))

• Solubility – low solubility is detrimental to good and complete oral absorption.

• H-bonding – important determinant of permeability. High H-bonding is related to low permeability and absorption due to the energy required to break H-bonds with solute molecules.

• Ionisation state – affects the solubility, lipophilicity, permeability and absorption of a compound. Charged compounds do not pass through membranes.

Drug-likeness: Lipinski's Rules (Rule of Five)



- Proposed by C. Lipinski to describe 'drug-like' molecules.
- Molecules displaying good oral absorption and /or distribution properties are likely to possess the following characteristics:
  - mass < 500Da
  - logP < 5.0
  - H-donors < 5
  - H-acceptors (number of N and O atoms) < 10

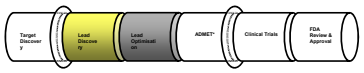
C.A. Lipinski (Pfizer)

- Rules used as a guide to inform drug design, but are not unequivocal.
- **Antibiotics (⟶)**, antifungals and vitamins (drugs that are injectable or substrates for membrane transporters) often do not adhere to these rules.
- Modified rules recommended to predict Blood Brain Barrier penetration:
  - mass < 450Da
  - PSA < 100Å^2 (polar surface area – related to logP)
  - H-donors ≤ 3
  - H-acceptors ≤ 6

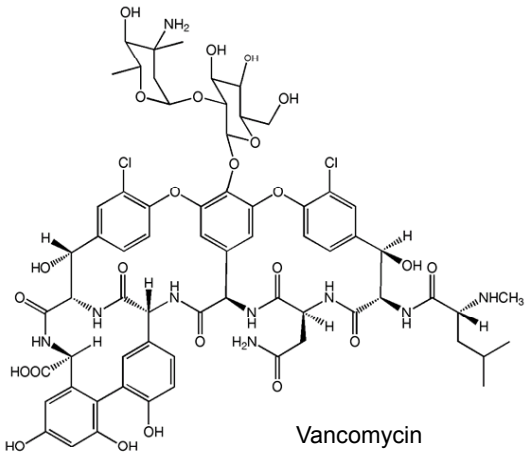
High risk of poor bioavailability if 2 or more of these conditions are violated.

CA Lipinski, *Adv. Drug Del. Rev.,* 1997, 23, 3

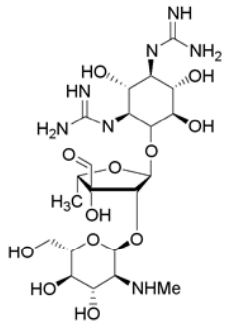
Drug-likeness: Antibiotics Are Often Misfits



Lipinski's rules describe <u>orally available</u> drugs!

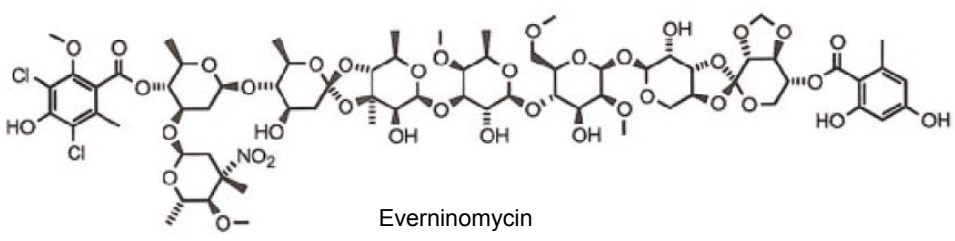


Vancomycin



Streptomycin

(i.v. antibiotics)



Everninomycin

Drug-likeness: Oral Availability – A Closer Look

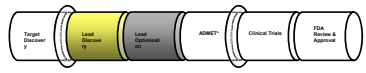


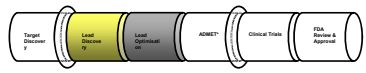
Table 5. Differences in Means for Selected Properties between Oral and Nonoral Drugs[a]

descriptor	oral mean (median) $n = 1202$	absorbent mean (median) $n = 118$	p-value	injectable mean (median) $n = 328$	p-value	topical mean (median) $n = 113$	p-value	SAR mean (median) $n = 113\,937$	clinical mean (median) $n = 1817$
			0.0016		**<0.0001**		**0.092**		
MW	343.7	392.3	0.49	558.2	**<0.0001**	368.5	**0.0094**	447.5	422.5
	(322.5)	(332.4)	0.43	(416.4)	**<0.0001**	(379.1)	**0.017**	(414.6)	(390.5)
			0.0059		**<0.0001**		**0.032**		
CLOGP	2.3	1.6	**0.02**	0.6	**<0.0001**	2.9	**0.001**	3.4	2.8
	(2.3)	(2.0)	**0.18**	(0.7)	**<0.0001**	(3.3)	**0.0002**	(3.5)	(3)
			0.073		**<0.0001**		0.06		
ONs	5.5	6.5	0.99	11.3	**<0.0001**	5	0.02	7.1	7
#O and N	(5)	(5)	0.27	(8)	**<0.0001**	(4)	0.12	(6)	(6)
			<0.0001		**<0.0001**		0.76		
OHsNHs	1.8	3	**0.007**	4.7	**<0.0001**	1.9	0.25	2.1	2.2
#OH and NH	(1)	(2)	**0.03**	(2)	**<0.0001**	(1)	0.38	(2)	(2)
			0.055		**0.0002**		**0.2**		
NRING	2.6	2.5	0.053	3.2	**0.0007**	2.9	**0.026**	3.5	3.3
	(3)	(2)	0.65	(3)	**<0.0001**	(3)	**<0.0001**	(3)	(3)
			<0.0001		**<0.0001**		0.57		
rotbond	5.4	7.9	0.15	12.7	**<0.0001**	5.3	0.36	8.4	8
	(5)	(4.5)	0.89	(7)	**<0.0001**	(5)	0.62	(7)	(6)
			0.21		**<0.0001**		0.71		
ACC	3.2	3.6	0.48	6.2	**<0.0001**	3.2	0.74	4	3.9
#H-accept.	(3)	(3)	0.63	(5)	**<0.0001**	(3)	0.16	(3)	(3)
			0.38		**0.087**		**<0.0001**		
HALOGEN	0.5	0.6	0.84	0.4	**0.0003**	0.9	**<0.0001**	0.6	0.5
	(0)	(0)	0.64	(0)	**<0.0001**	(0)	**0.0002**	(0)	(0)

[a] Within a p-value cell, the top p-value is from the two-sample t-test, the middle p-value is from the Wilcoxon test, and the bottom p-value is from the median test. For count-based descriptors, the t-test was performed on a (count + 0.5)$^{1/2}$ transformation. All p-values for the SAR group were <0.0001 and are not included in the table. All p-values, except for the halogen count, were <0.0001 for the clinical group and are not included in the table. Values in bold indicate that at least two of the three p-values are <0.05.

Vieth et al., *J. Med. Chem.*, 2004, 47, 224

Drug-likeness: Oral Availability – A Closer Look



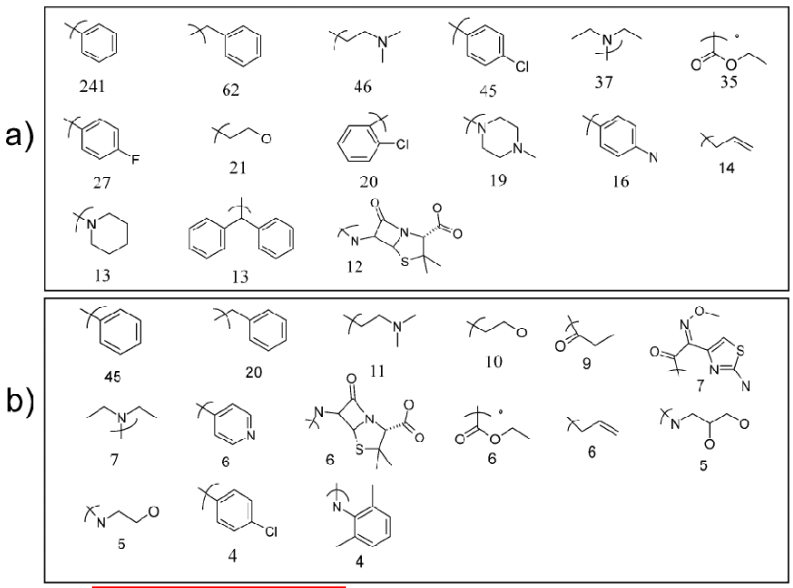
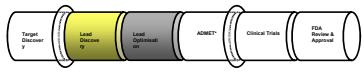


Figure 5. Comparison of the most frequent side chains: (a) oral; (b) injectable. The numbers indicate the count of the drugs containing that fragment. The means of properties are not significantly different for (a) and (b).

Vieth et al., *J. Med. Chem.*, 2004, 47, 224

Drug-likeness: Oral Availability – A Closer Look



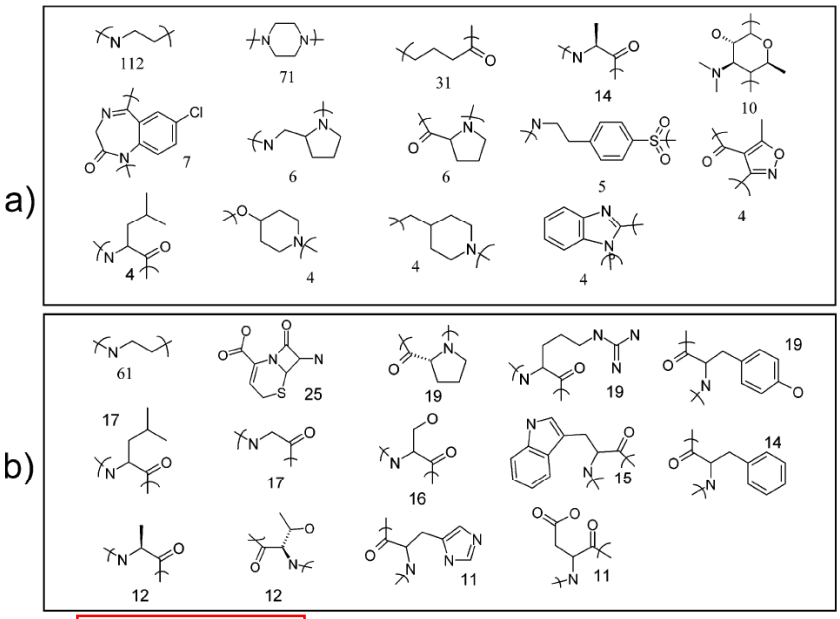
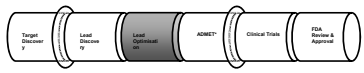


Figure 6. Comparison of most frequent scaffolds: (a) oral; (b) injectable. The numbers indicate the number of drugs containing the fragment. The means of physical properties (CLOGP, ON, rotbond) are significantly different for (a) and (b).

Vieth et al., *J. Med. Chem.,* 2004, 47, 224

Lead Optimisation



Lead structure optimisation is an evolutionary procedure, in which every minor or major improvement in certain properties leads to a new analog, which is further optimised until the final candidate has all desired properties to start its clinical investigation.

- Traditional Medicinal Chemistry
- Combinatorial/Parallel Synthesis (focused libraries)

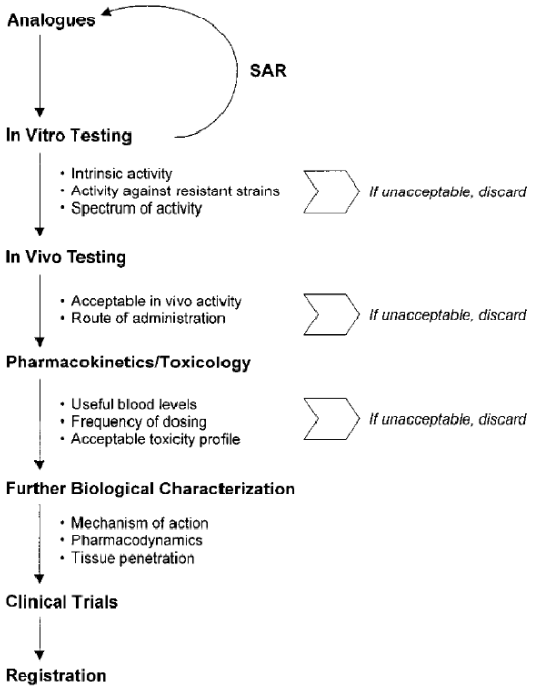
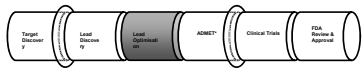


Figure 1. Generalized testing scheme for oxazolidinones showing incorporation of the structure–activity relationship (SAR) componen

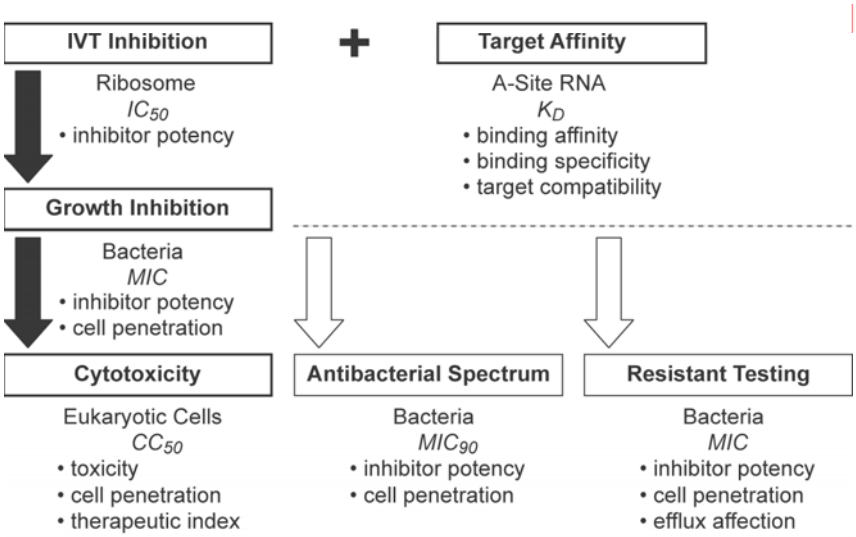
(Barbachyn & Ford, *Angew. Chem. Int. Ed.,* 2003, 42, 2010)

Lead Optimisation: Assays for Antibacterials

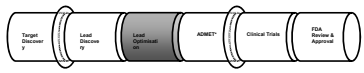


- Target-based affinity or functional assay (if target is known).
- MIC (minimum inhibitory concentration)
- Spectrum of activity (MIC90)

Example for Ribosome-Directed Antibacterial Discovery Program:

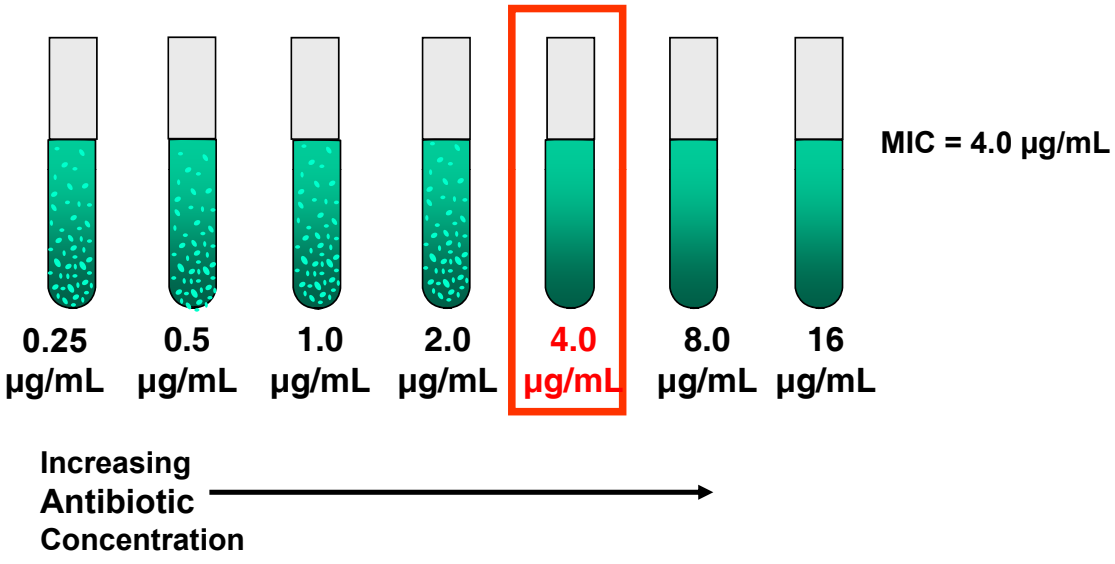


Lead Optimisation: MIC



MIC (minimum inhibitory concentration): lowest concentration of an antimicrobial that results in the inhibition of visible growth of a microorganism after overnight exposure

Known bacterial inoculum placed into each tube



Lead Optimisation: MIC

- **Breakpoint**: concentration above which the isolate is described at resistant and below which is susceptible
  e.g. S < 8mg/L R ≥ 8mg/L
  Breakpoint = 8mg/L

- **MIC_{50}** Median for series of MICs

- **MIC_{90}**
  - MICs of population ordered from lowest to highest
  - MIC value of the strains that appears 90% up the series.
  - Antibiotic considered to be successful if > 90% of population inhibited.
  - Also show if resistance is emerging i.e. 10% of population resistant.

Lead Optimisation: Spectrum – MIC90

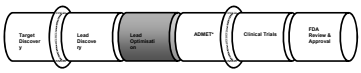
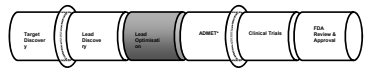


Table 1
Antimicrobial activity of dalbavancin compared with 15 other antimicrobials against 2644 Gram-positive cocci associated with SSTI and CR-BSI isolated in 2004 (United States)

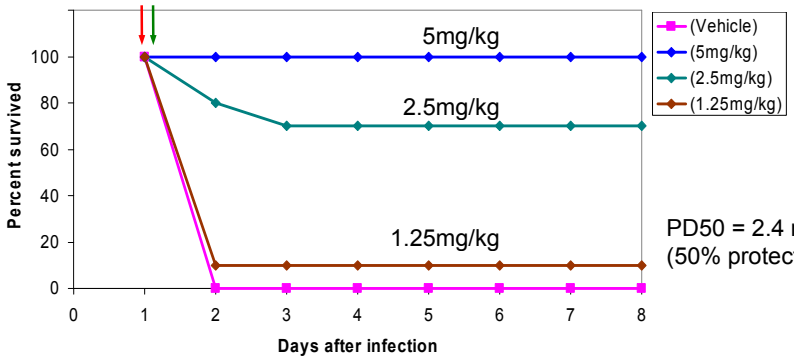
Organism (no. tested)/antimicrobial agent	Cumulative % of isolates inhibited at MIC (μg/mL)								MIC (μg/mL)	
	≤0.03	0.06	0.12	0.25	0.5	1	2	4	50%	90%
Staphylococcus aureus (2102)										
Dalbavancin	46	98	99	99	100	100	100	100	0.06	0.06
Oxacillin	–[a]	–	–	11	45	50	*51*[b]	–	1	>2
Ceftriaxone	–	–	–	0	<1	<1	7	49	8	>32
Clindamycin	–	20	70	71	*71*	71	71	71	0.12	>8
Daptomycin	0	0	1	47	99	*>99*	100	100	0.5	0.5
Erythromycin	–	<1	1	37	*38*	38	38	39	>8	>8
Gentamicin	–	–	–	–	–	–	95	*96*	≤2	≤2
Levofloxacin	<1	4	39	56	58	*58*	60	68	0.25	>4
Linezolid	–	0	0	0	<1	28	100	*100*	2	2
Rifampin	–	–	–	–	97	*98*	98	–	≤0.5	≤0.5
Synercid®	–	–	–	43	97	*>99*	>99	–	0.5	0.5
Teicoplanin	–	–	<1	6	64	96	>99	>99	0.5	1
Tetracycline	–	–	–	63	92	94	95	*95*	≤0.25	0.5
TMP/SMX[c]	–	–	–	–	96	98	*98*	–	≤0.5	≤0.5
Vancomycin	–	–	0	<1	14	98	>99	*>99*	1	1

Lead Discovery: Antibacterial *In vivo* Testing



Mouse Protection Model (systemic infection):

IV route (single dose)

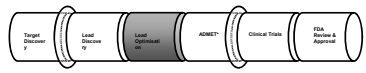


PD50 = 2.4 mg/kg
(50% protective dose)

- 10 mice per group
- Infection with *E. coli* ATCC-25922 by IP route

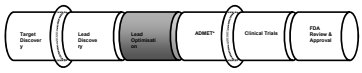
(Zhou et al., *Antimicrob. Agents Chemother.*, 2005, 49, 4942)

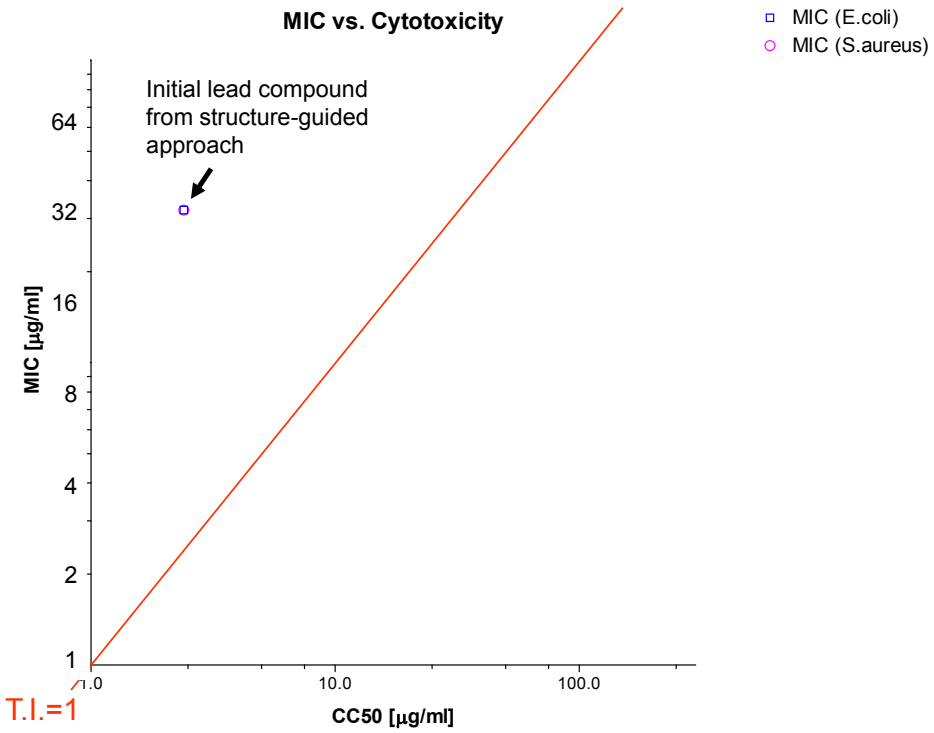
Lead Optimisation: Focused Libraries



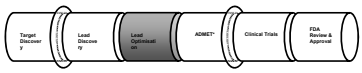
- After identification or generation of a lead compound for a specific target, it is often desirable to use **focused libraries** of **high complexity**, but relatively **low diversity** to optimize the lead.
- Structural analysis of the lead compound(s) will suggest key pharmacophores that are critical for target binding
  - The focused library should consist of molecules that contain various combinations and/or forms of these key pharmacophores
- Focused combinatorial libraries can be synthesized by:
  - Combining the pharmacophores in different ways
  - Adding various substituents to the pharmacophores
  - Chemically modifying the pharmacophores in defined ways

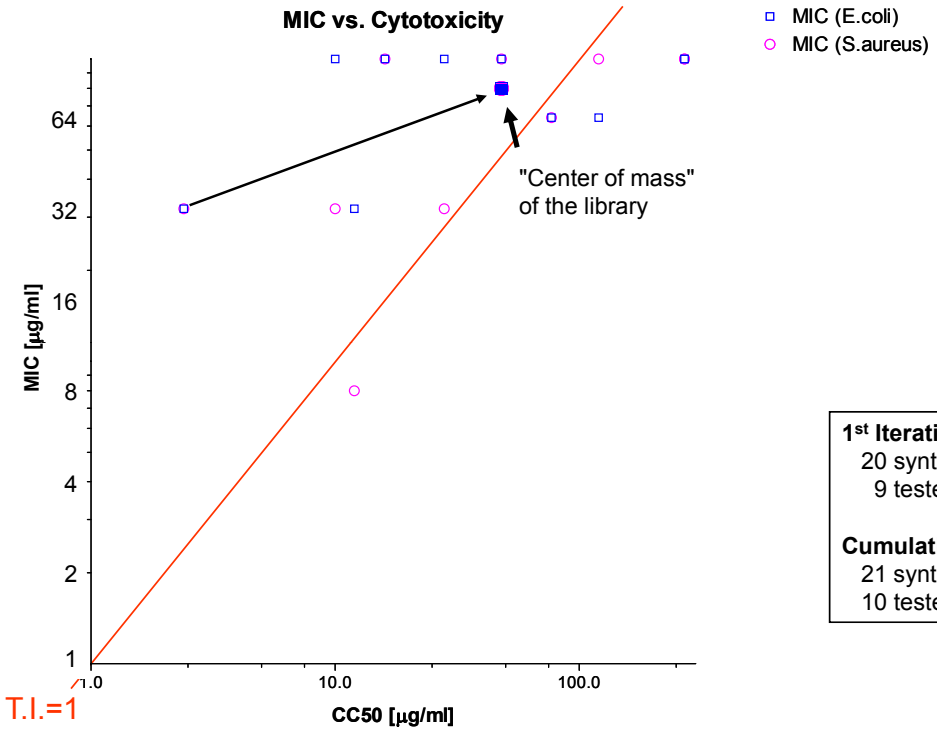
Lead Optimisation: Example of Antibacterial Lead





Lead Optimisation: Example - 1st Iteration

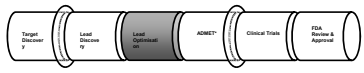


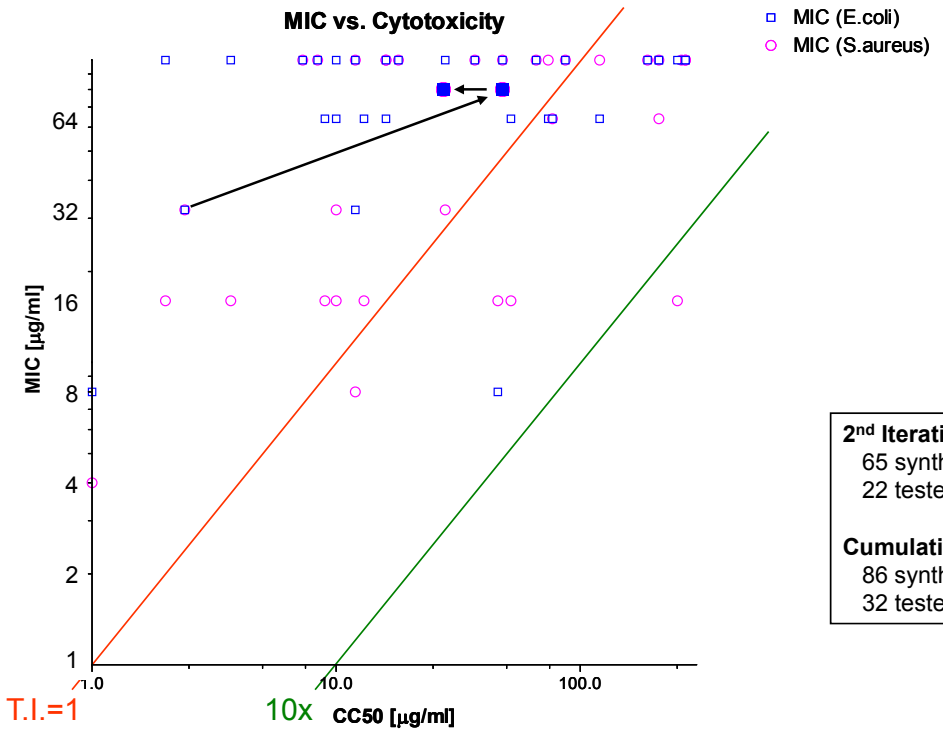


1st Iteration:
20 synthesized
9 tested

Cumulative:
21 synthesized
10 tested

Lead Optimisation: Example – 2nd Iteration

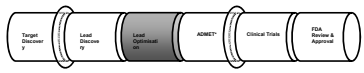


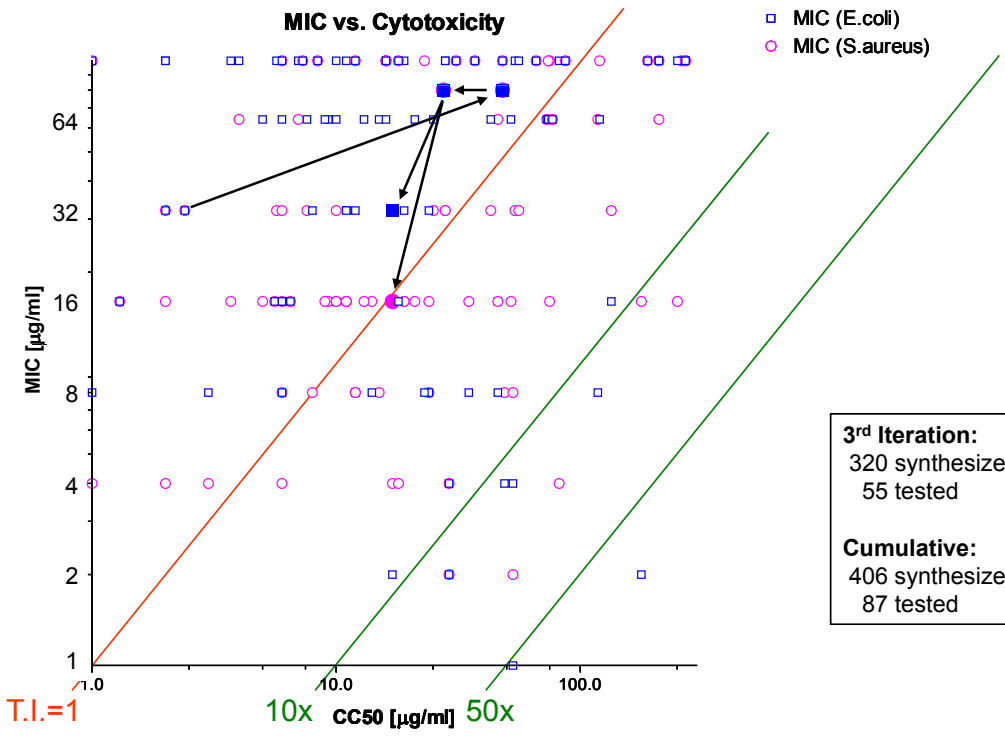


2nd Iteration:
65 synthesized
22 tested

Cumulative:
86 synthesized
32 tested

Lead Optimisation: Example – 3rd Iteration

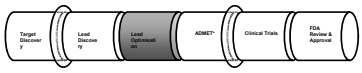


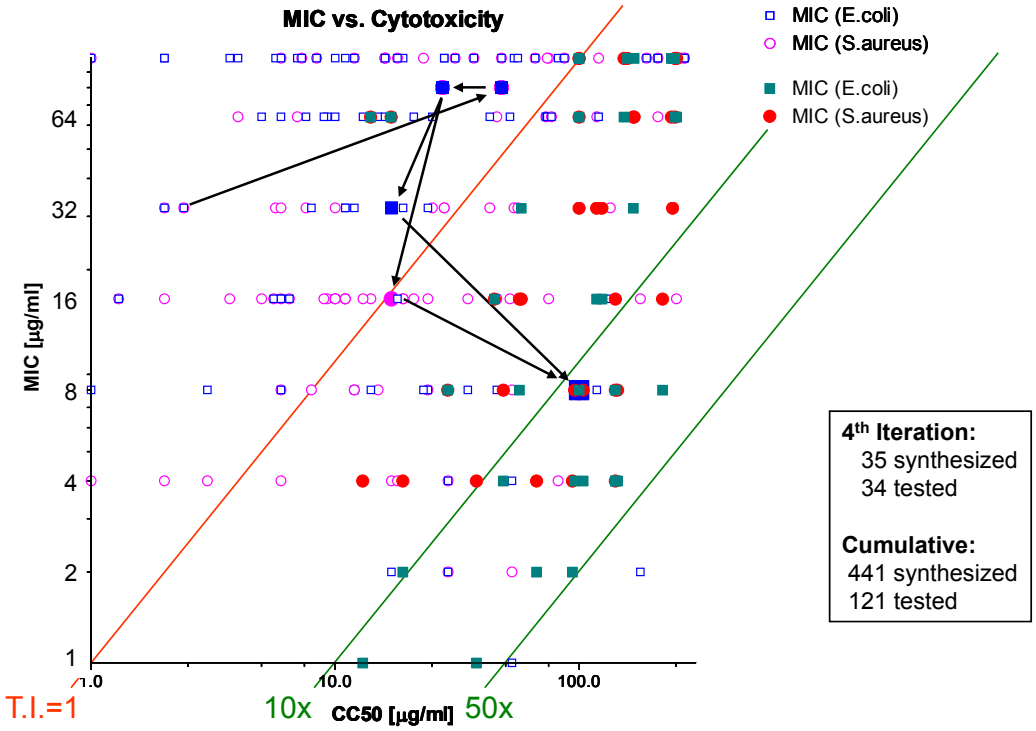


3rd Iteration:
320 synthesized
55 tested

Cumulative:
406 synthesized
87 tested

Lead Optimisation: Example – 4th Iteration

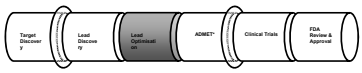


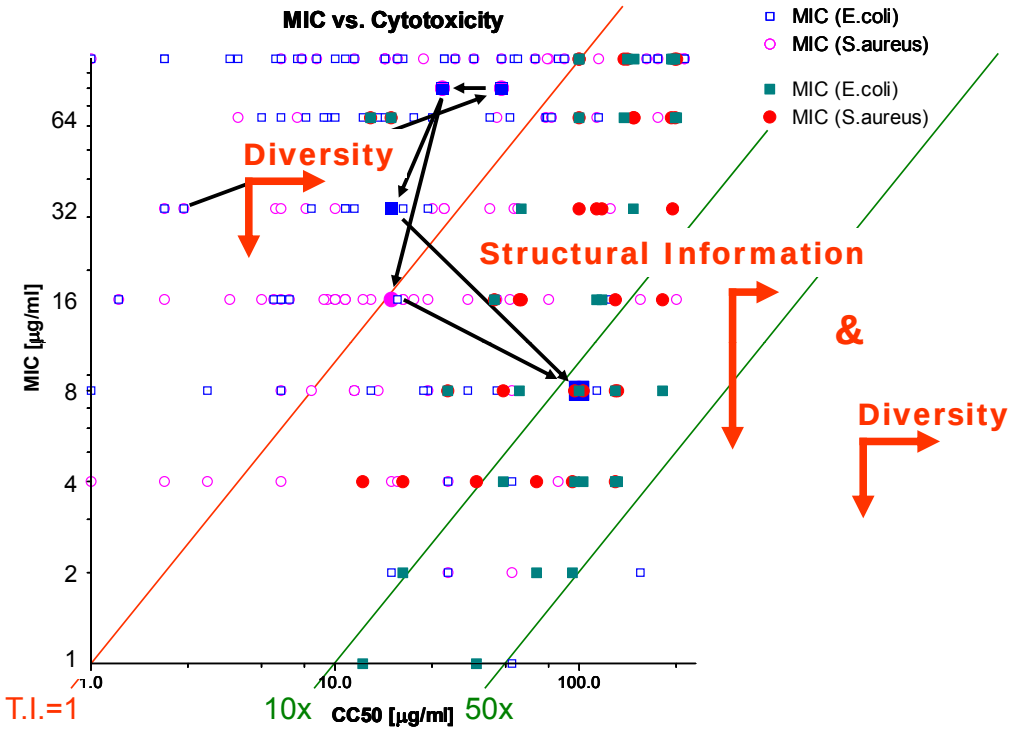


4th Iteration:
35 synthesized
34 tested

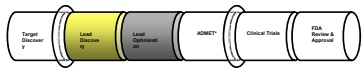
Cumulative:
441 synthesized
121 tested

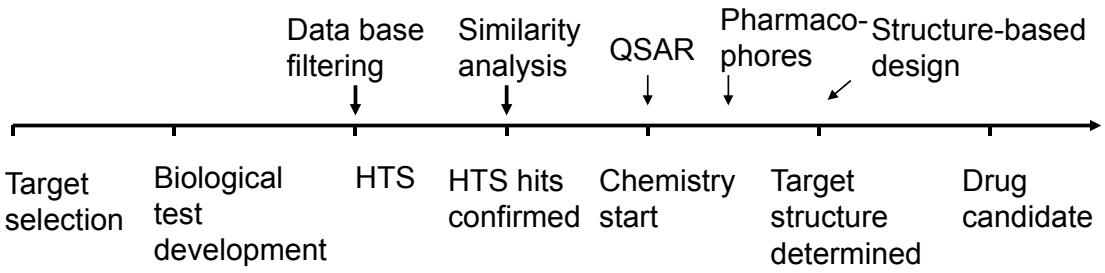
Lead Optimisation: Example – Driving Forces





Lead Discovery/Optimisation: Comput. Design





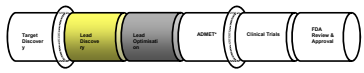
• After target identification and the setup of an activity test: filtering databases of existing molecules to narrow down molecules for screening.

• After HTS: similarity analysis of available ligands comparable to the positive molecules obtained by screening.

• After the phase of synthesis chemistry has started: QSAR (qunatitative structure activity relationship) as well as pharmacophore screening.

• After structure of the target/target complexes have been determined: structure-guided optimisation and design of de novo ligands.

Lead Discovery/Optimisation: QSAR



Quantitative Structure-Activity Relationship (QSAR) is a mathematical relationship between a biological activity of a molecular system and its geometric and chemical characteristics.

QSAR attempts to find consistent relationship between biological activity and molecular properties, so that these "rules" can be used to evaluate the activity of new compounds.

Input: *n* descriptors $P_1,..P_n$ and the value of biological activity (*EC50* for example) for *m* compounds.

	Biol.Act.	P_1	P_2	...				P_n
Cpd1	0.7	3.7						
Cpd2	3.2	0.4						
…								
Cpdm								

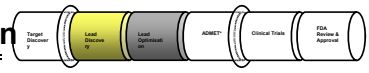
The problem of QSAR is to find coefficients *C0,C1,...Cn* such that:

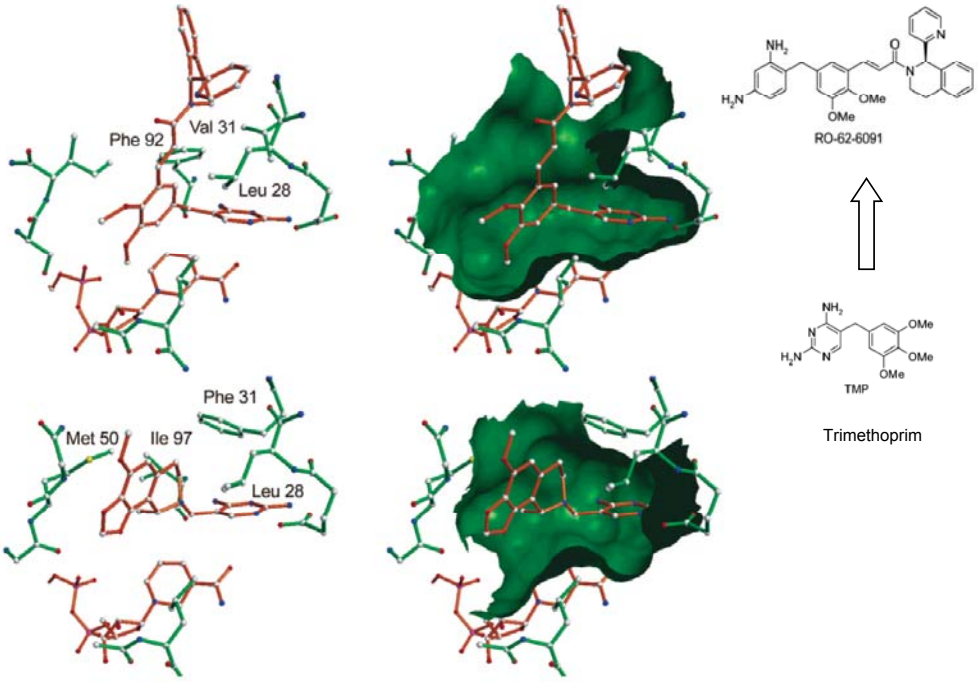
Biological activity = C0+(C1*P1)+...+(Cn*Pn)

and the prediction error is minimized for a list of given m compounds.

Partial least squares (PLS) is a technique used for computation of the coefficients of structural descriptors.

Lead Discovery/Optimisation: Struct.-Based Design





Trimethoprim

Figure 8. Binding mode of RO-62-6091 as determined by crystal structure analysis (top structures) complexed with *S. aureus* DHFR and the modeled binding mode of (*R*)-**9** (bottom structures) complexed with *S. pneumoniae* DHFR. The cofactor NADPH is partially visible at the bottom of the binding sites.

(Wyss et al., *J. Med. Chem.*, 2003, 46, 2304)