DRUG DESIGN

Optimizing Binding Interactions and Decision Making in Medicinal Chemistry

DRUG DESIGN: OPTIMIZING BINDING INTERACTIONS

Aim: To optimize binding interactions with target

Reasons

- To increase activity and reduce dose levels
- To increase selectivity and reduce side effects

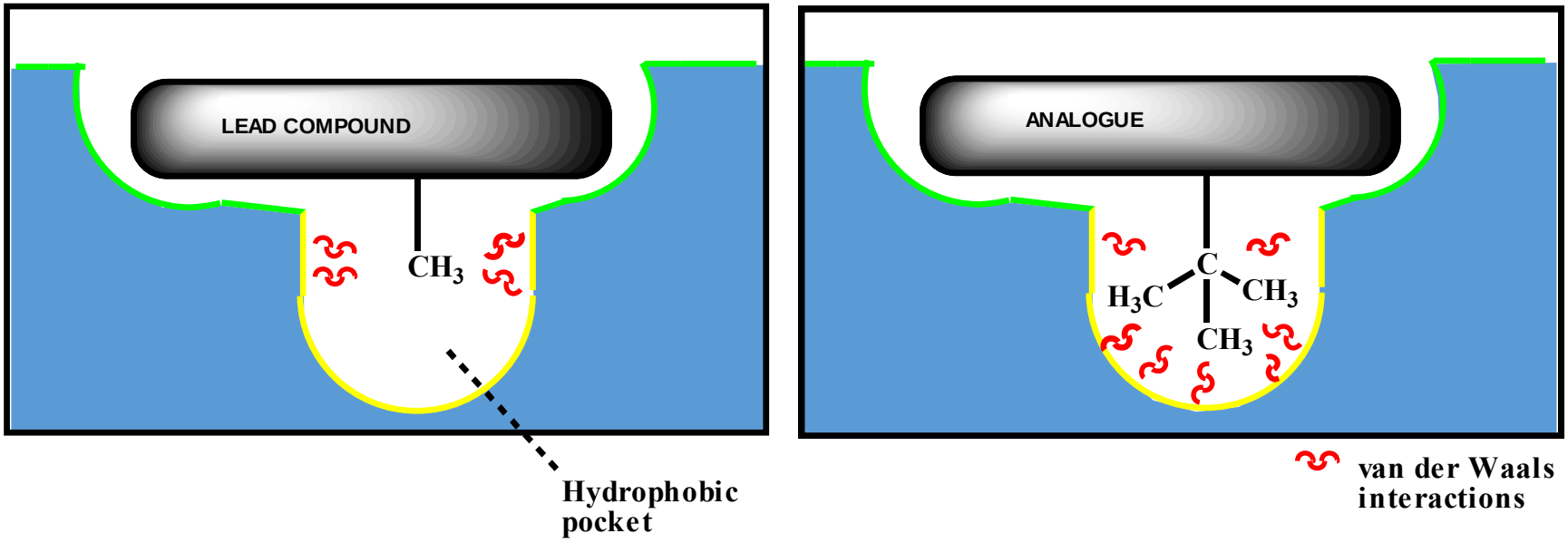
Strategies

- Vary alkyl substituents
- Vary aryl substituents
- Extension
- Chain extensions / contractions
- Ring expansions / contractions
- Ring variation
- Isosteres
- Simplification
- Rigidification

Vary Alkyl Substituents:

Rationale:

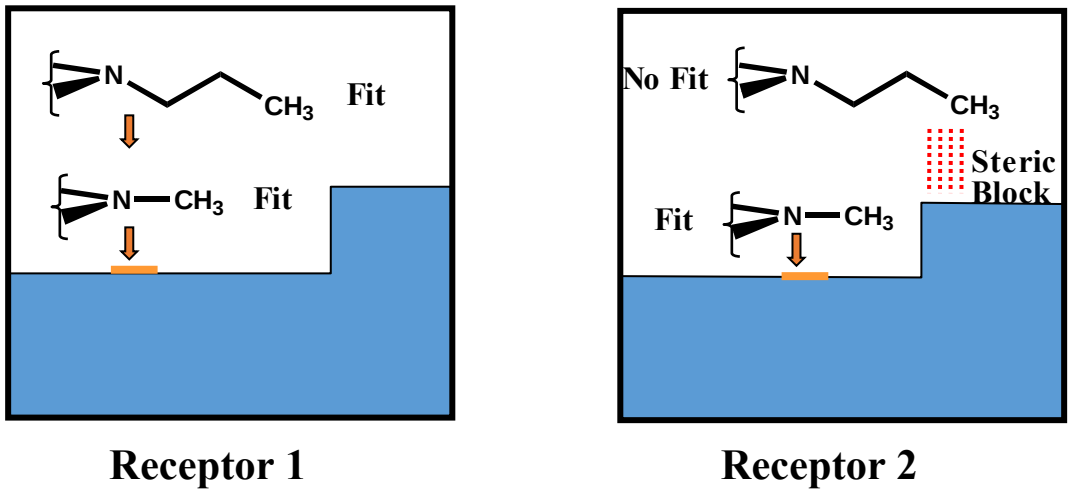
- Alkyl group in lead compound may interact with hydrophobic region in binding site
- Vary length and bulk of group to optimise interaction



Vary Alkyl Substituents:

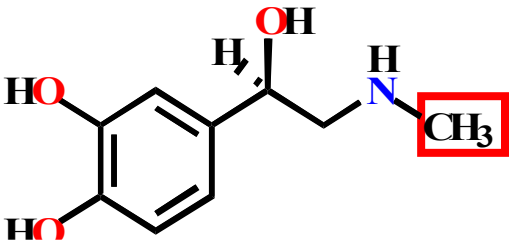
Rationale:

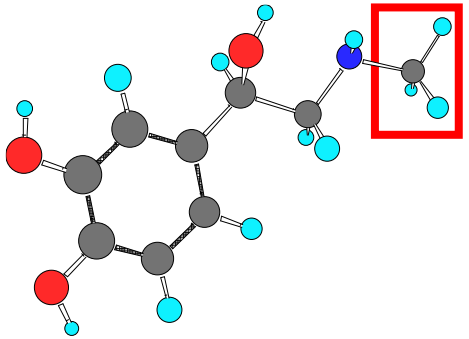
Vary length and bulk of alkyl group to introduce selectivity



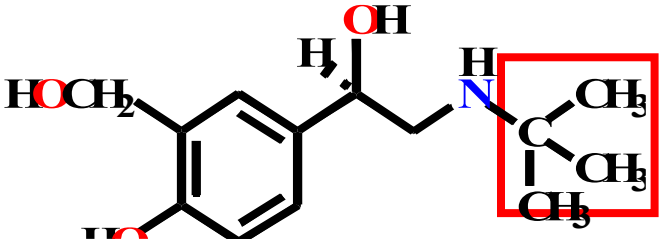
Example: Selectivity of adrenergic agents for β-adrenoceptors over α-adrenoceptors

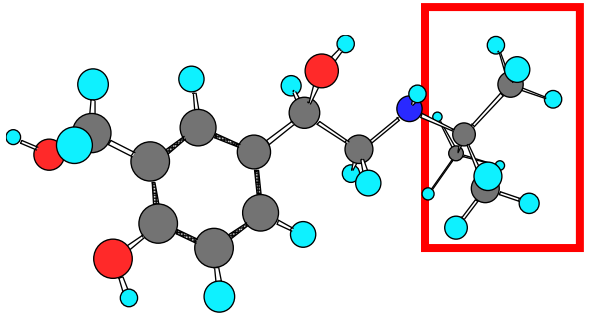
Adrenaline



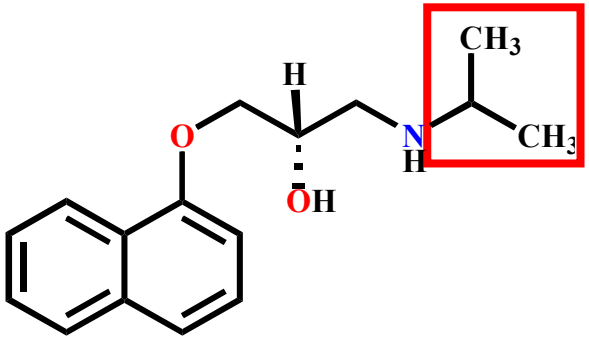


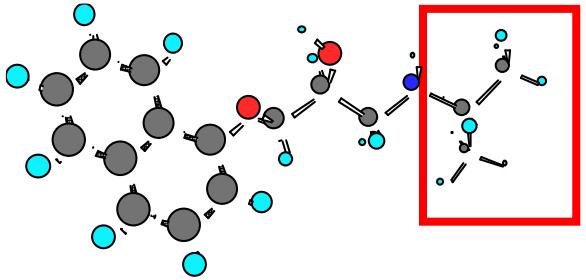
Salbutamol (Ventolin)
(Anti-asthmatic)





Propranolol
(β-Blocker)





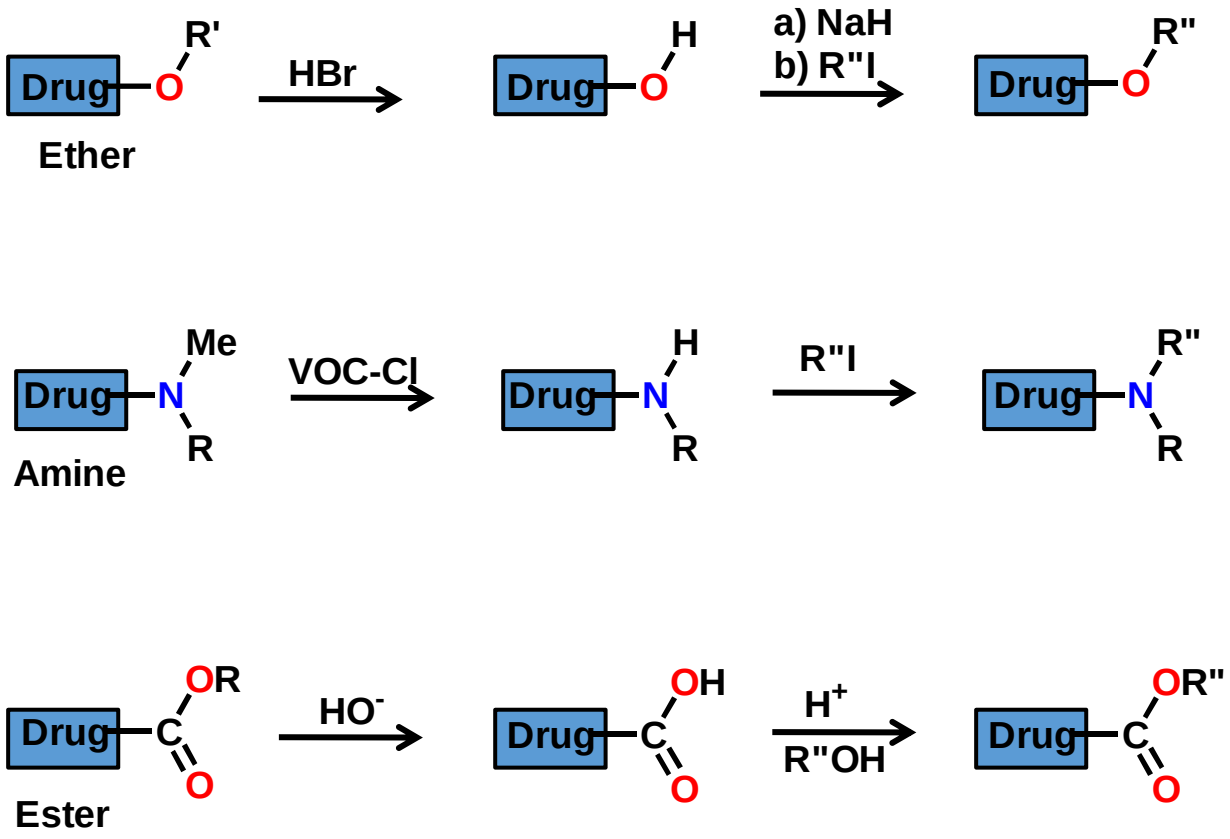
Vary Alkyl Substituents:

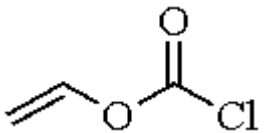
Synthetic feasibility of analogues:

- Feasible to replace alkyl substituents on heteroatoms with other alkyl substituents
- Difficult to modify alkyl substituents on the carbon skeleton of a lead compound.
- Total synthesis usually required to vary alkyl substituents that are on the carbon skeleton

Vary Alkyl Substituents:

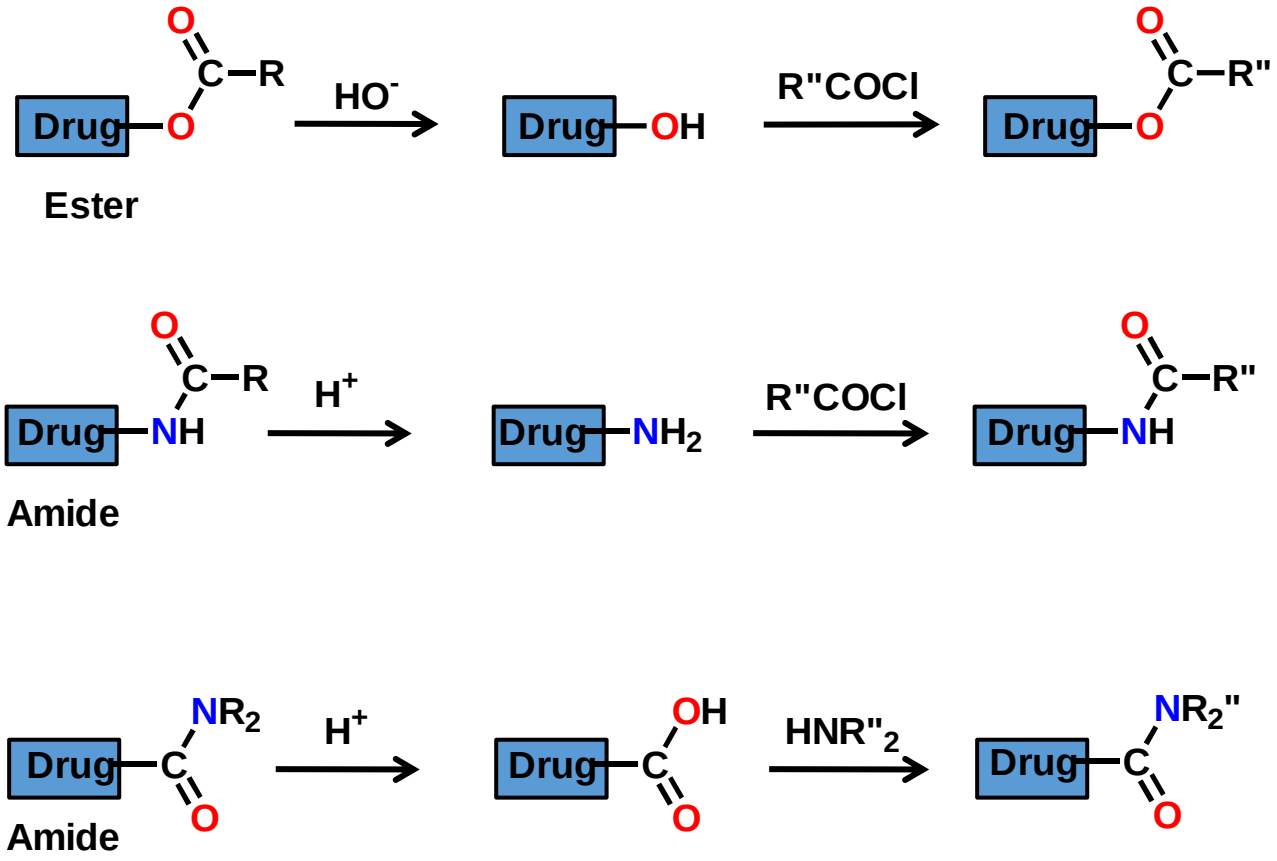
Methods



VOC-Cl =  Vinyl Chloroformate

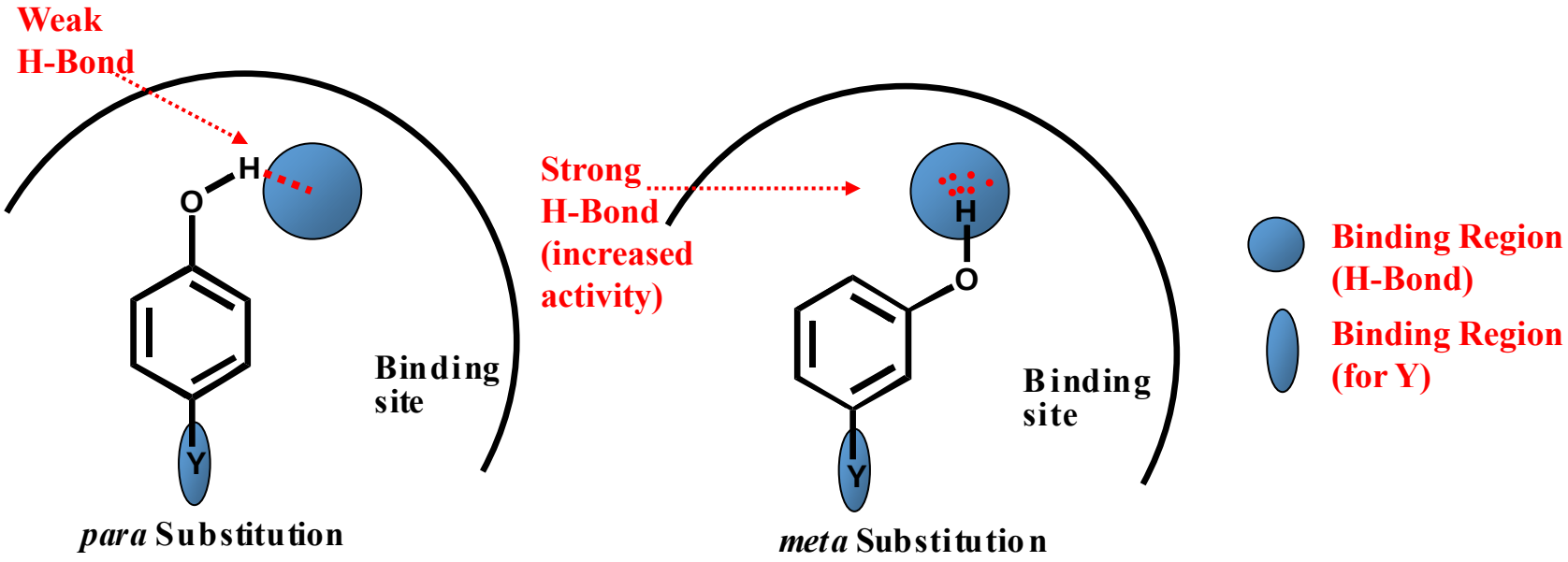
Vary Alkyl Substituents:

Methods



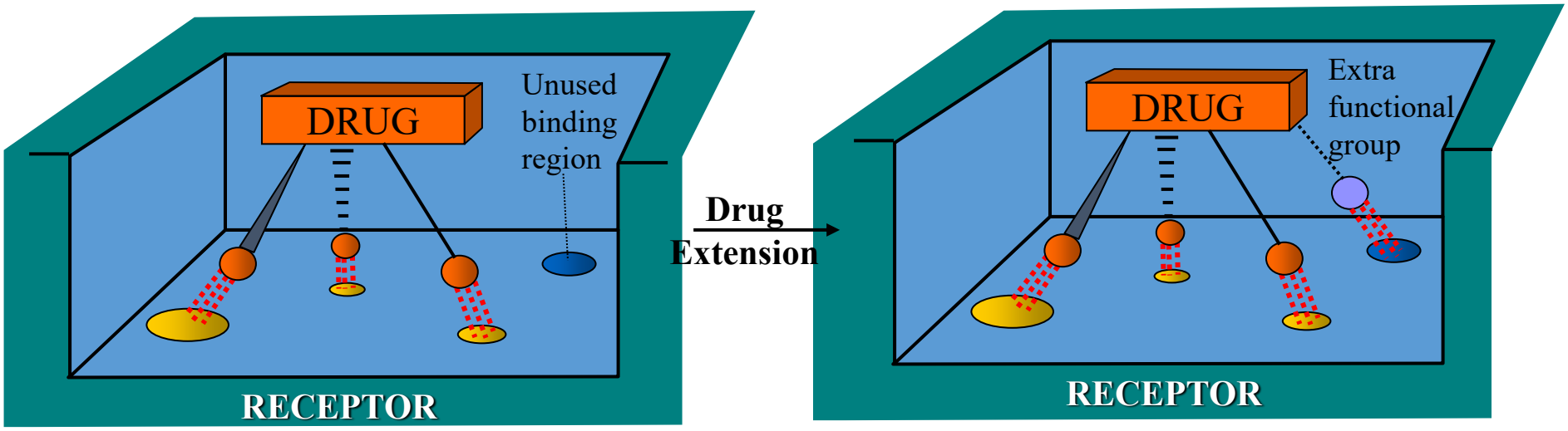
Vary Alkyl Substituents:

Vary substituents
Vary substitution pattern



Extension - Extra Functional Groups

Rationale : To explore target binding site for further binding regions to achieve additional binding interactions



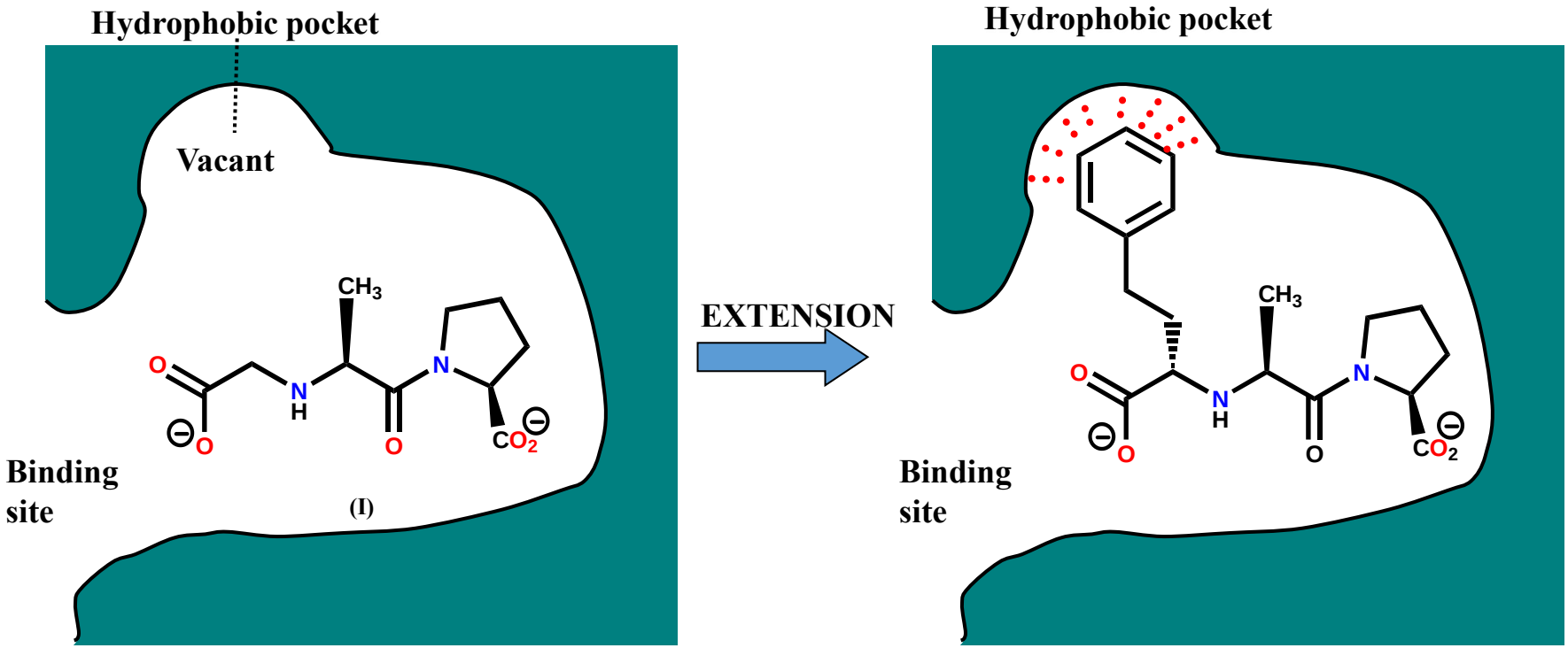
Binding regions

Binding groups

Extension - Extra Functional Groups

Example: Angiotensin-converting enzyme (ACE) Inhibitors

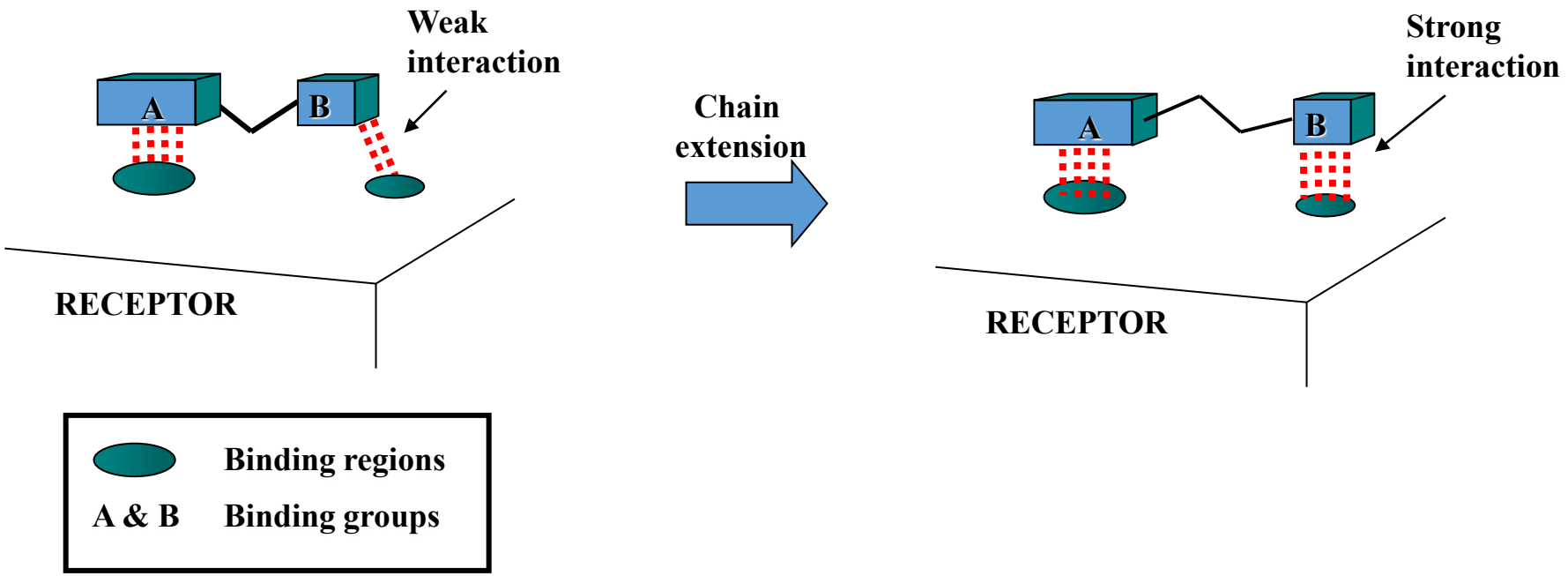
- ACE is a central component of the renin-angiotensin system (RAS), which controls blood pressure by regulating the volume of fluids in the body. ACE inhibitors help relax blood vessels.



Chain Extension / Contraction

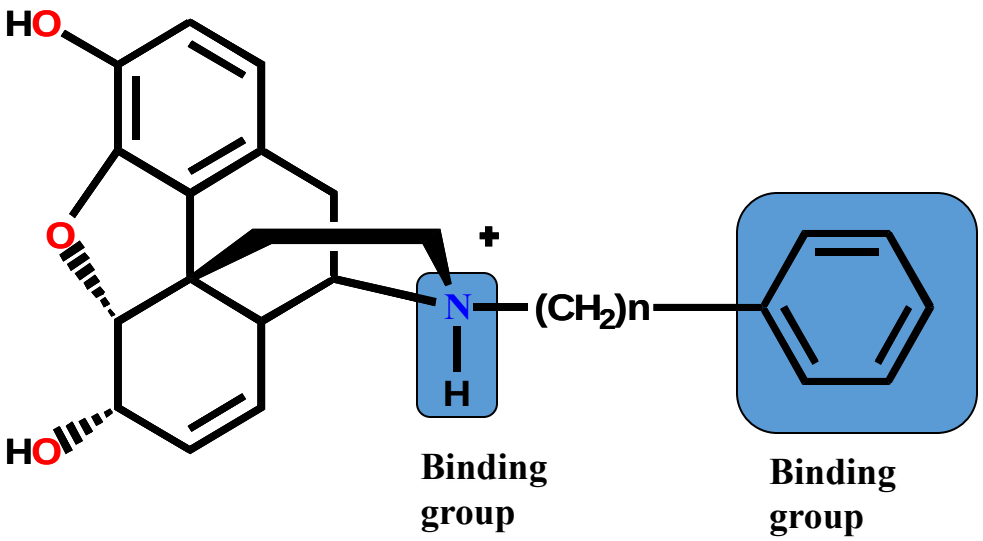
Rationale:

- Useful if a chain is present connecting two binding groups
- Vary length of chain to optimize interactions



Chain Extension / Contraction

Example: *N*-Phenethylmorphine, μ-opioid receptor antagonist.



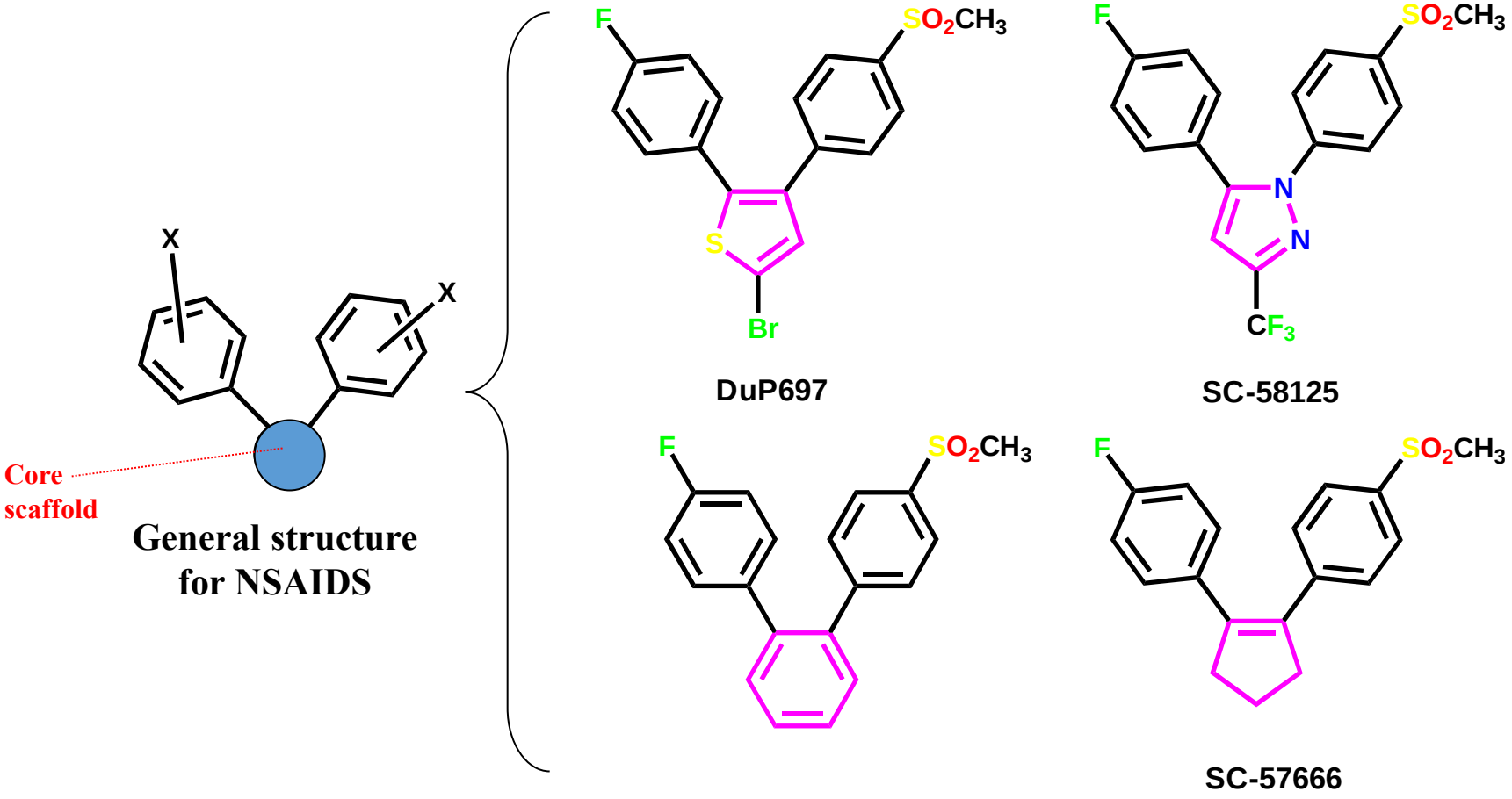
Optimum chain length = 2

Phenethyl group extends out to reach an additional binding point deeper inside the μ-opioid receptor pocket.

Ring Variations:

Rationale:

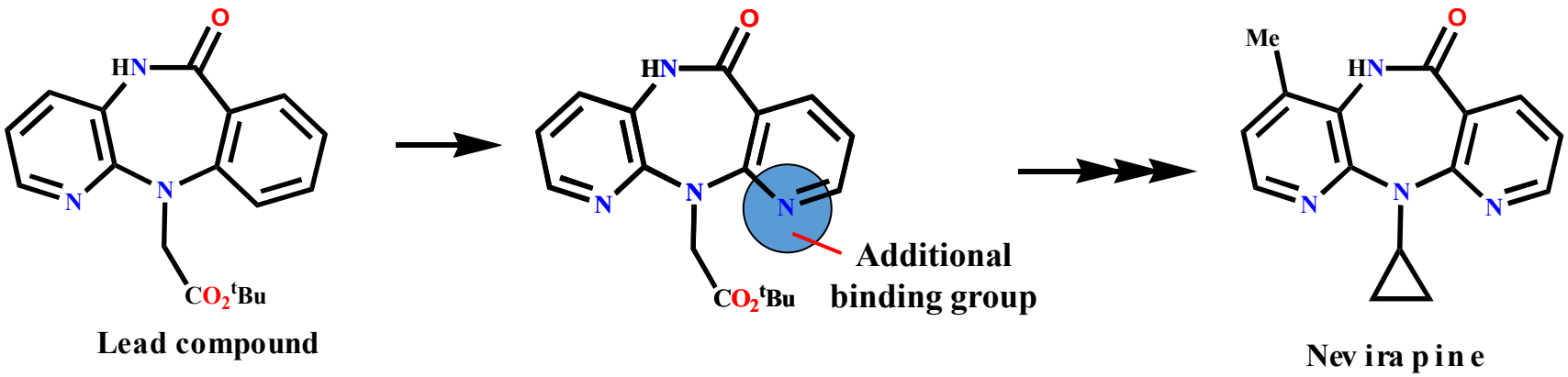
- Replace aromatic/heterocyclic rings with other ring systems
- Often done for patent reasons



Nonsteroidal anti-inflammatory drugs (NSAIDs) block the COX enzymes and reduce prostaglandins throughout the body.

Ring Variations:

Example: Nevirapine (antiviral agent)



Isosteres and bio-isosteres

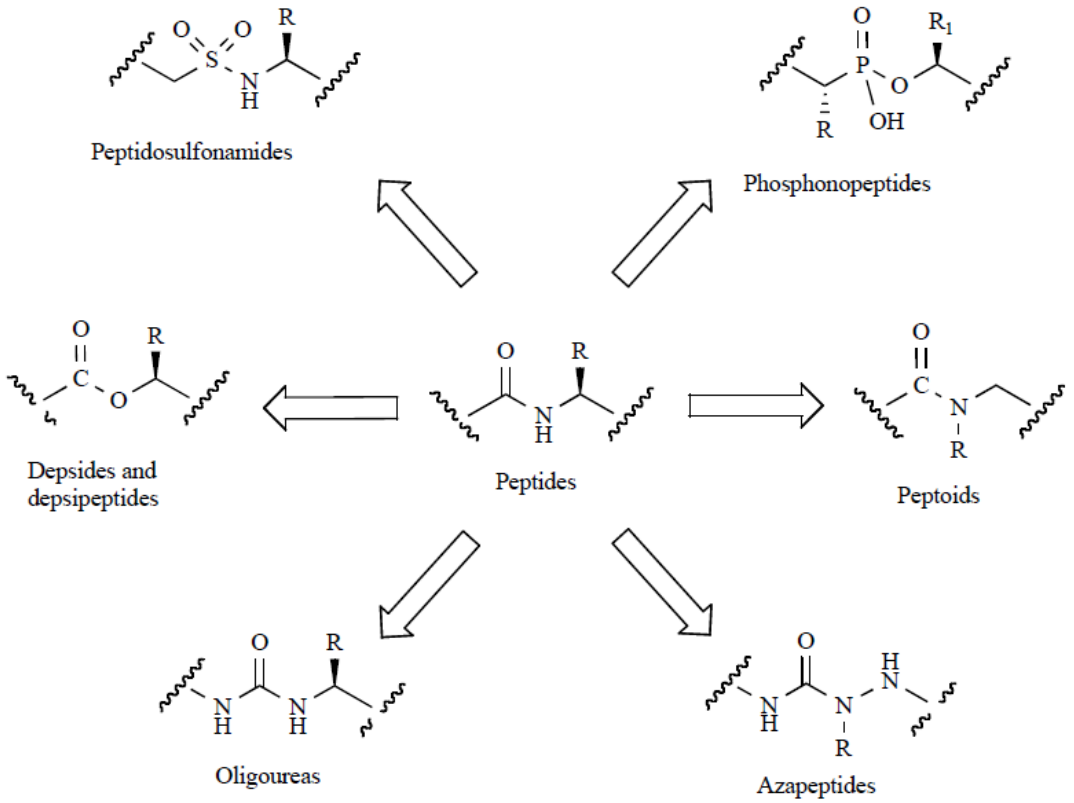
Rationale for isosteres:

- Replace a functional group with a group of same valency (isostere)
- e.g. OH replaced by SH, NH_2, CH_3
- e.g. O replaced by S, NH, CH_2
- e.g. H replaced by F

- Leads to more controlled changes in steric/electronic properties

- May affect binding and/or stability

Isosteres and bio-isosteres



Schematic presentation of peptidomimetic containing amide surrogates that are isosteric with the natural peptidic amide bonds.

Mini Rev. Org Chem. **2007**, 4:268-280

Simplification

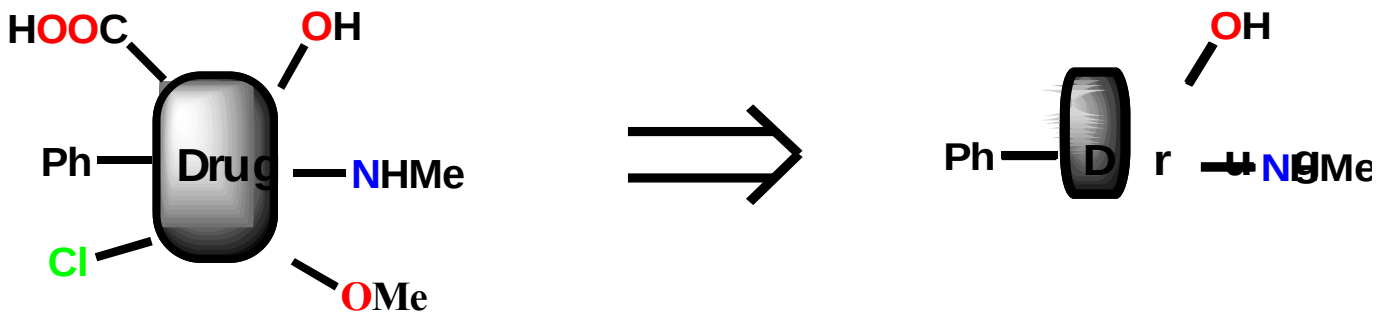
Rationale:

- Lead compounds from natural sources are often complex and difficult to synthesise
- Simplifying the molecule makes the synthesis of analogues easier, quicker and cheaper
- Simpler structures may fit the binding site easier and increase activity
- Simpler structures may be more selective and less toxic if excess functional groups are removed.

Simplification

Methods:

- Retain pharmacophore
- Remove unnecessary functional groups



Simplification

Methods

Remove excess rings

Example

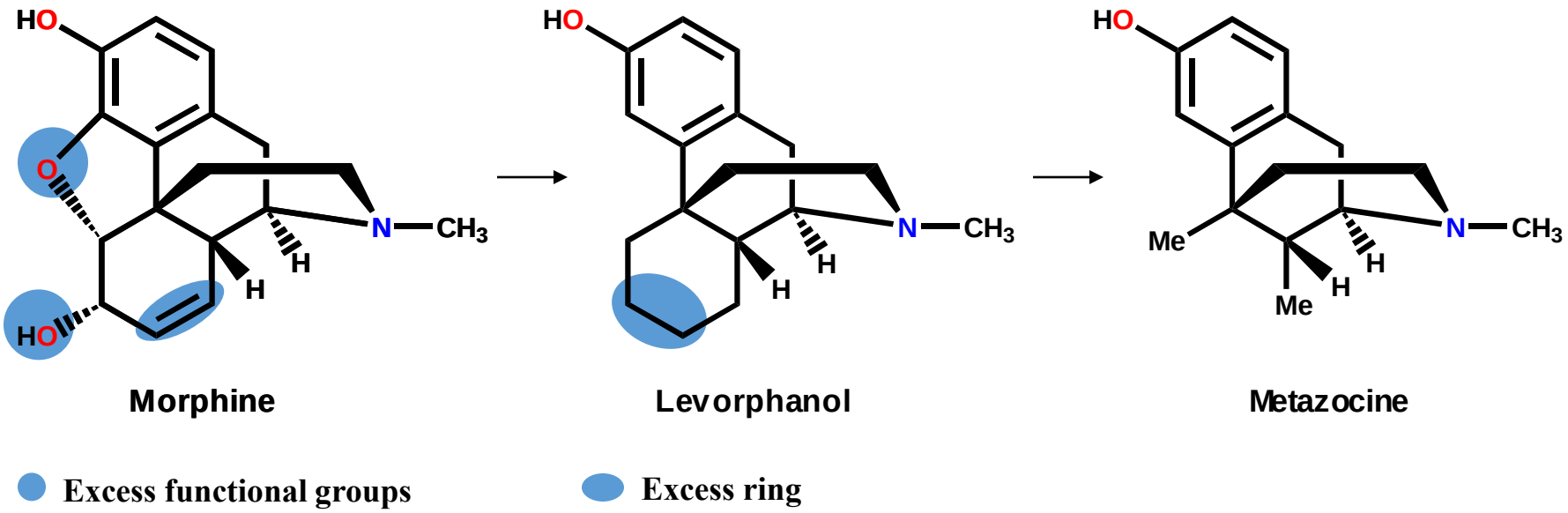
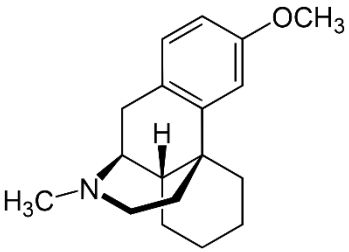


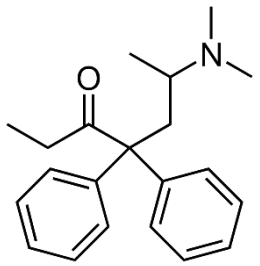
Table 1

Receptor binding affinity or inhibition of neuronal reuptake (K_i, nM, except nAChR, IC_{50}) and antinociceptive potency (ED_{50}, s.c., mg/kg).

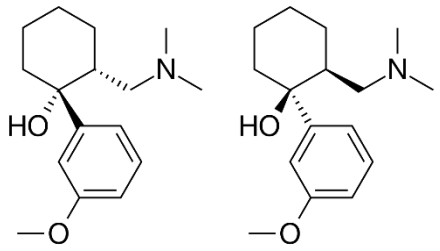
Compound	MOR	DOR	KOR	NRI	SRI	NMDA	nACh	ED_{50}
Morphine	1	145	23	IA	IA	IA	—	2.4
Methadone (±)	2	435	405	—	—	≥850	—	0.9
L isomer	1	371	1,860	702	14	—	—	—
D isomer	20	960	1,370	12,700	992	—	2,500	—
Levorphanol	0.1–0.4	4–5	2–4	1,210	86	630	—	0.4
Dextromethorphan	1,280	11,500	7,000	240	23	1,720	—	—
Tramadol (±)	2,120	57,700	42,700	785	992	—	—	—
(+) enantiomer	1,330	62,400	54,000	2,510	528	—	—	—
(–) enantiomer	24,800	IA	53,500	432	2,350	—	—	—



dextromethorphan



methadone



(1*R*,2*R*)-tramadol (1*S*,2*S*)-tramadol

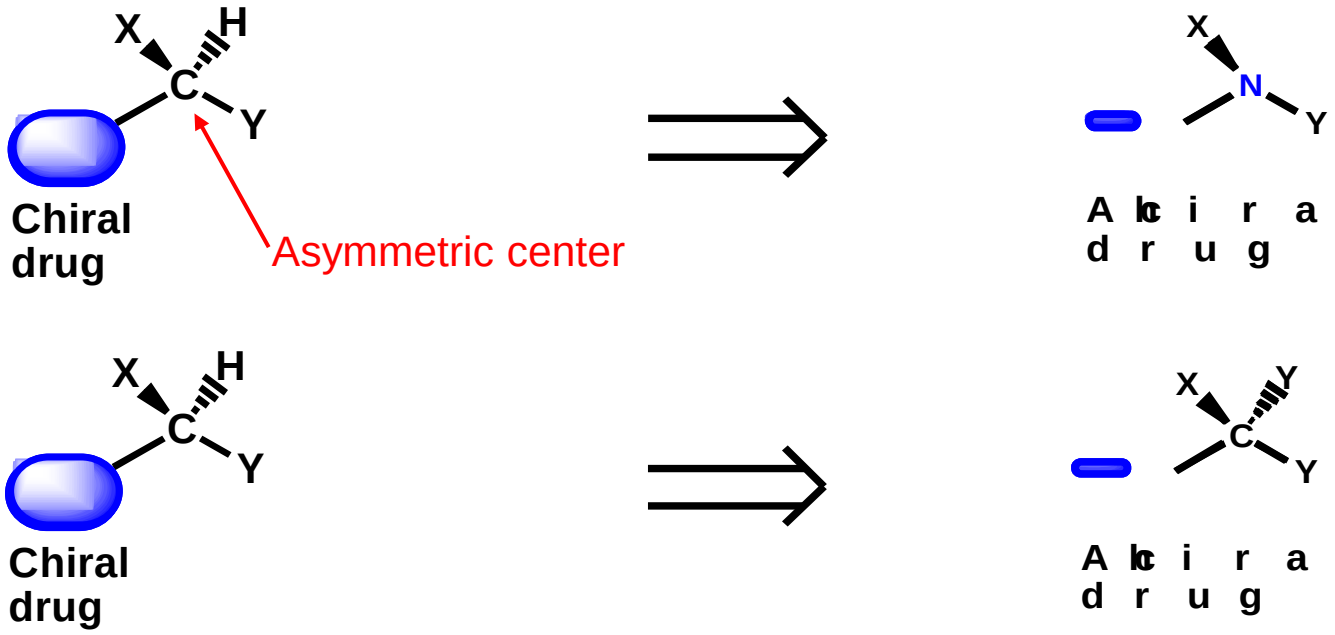
MOR, DOR, KOR = μ, δ, κ opioid receptor type, respectively; NRI = neuronal norepinephrine reuptake inhibition; SRI = neuronal serotonin (5-HT) reuptake inhibition; NMDA= N-methyl-D-aspartate receptor (glutamate receptor and ion channel protein found in nerve cells), nACHR = α3β4 nicotinic acetylcholine receptor; ED_{50} = rat tail-flick test; IA = inactive (>100,000 nM); NT = not tested.

Pain Medicine, Volume 16, Issue 9, 1 September 2015, Pages 1673–1679,

Simplification

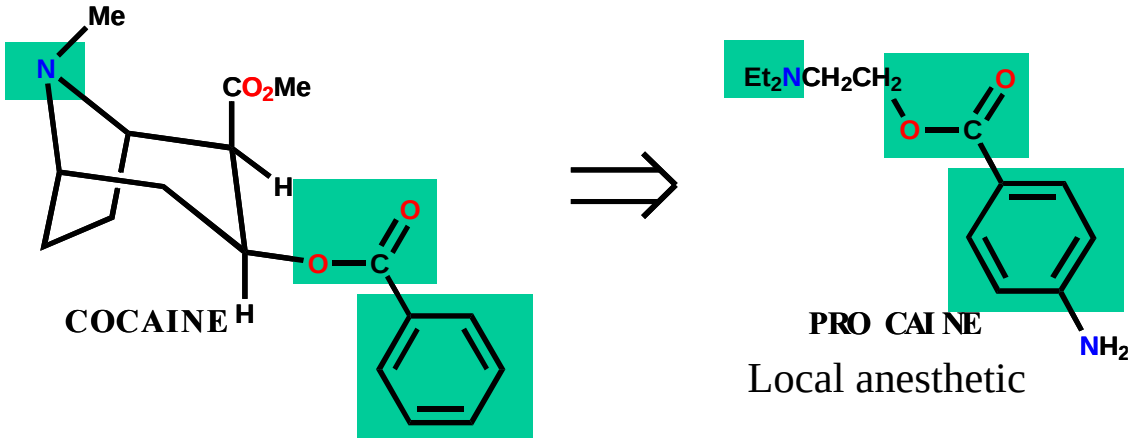
Methods:

Remove asymmetric centres



Simplification

Example:



Pharmacophore

- Important binding groups retained
- Unnecessary ester removed
- Complex ring system removed

Simplification

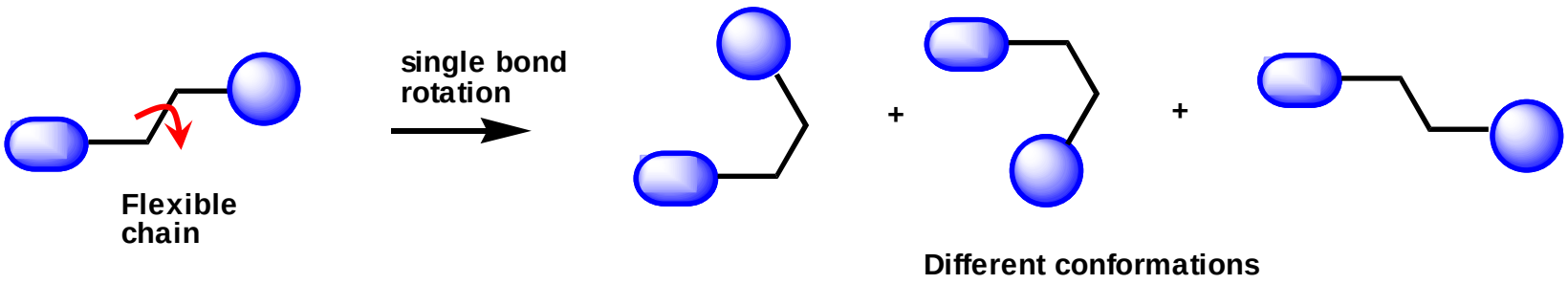
Disadvantages:

- Oversimplification may result in decreased activity and selectivity.
- Simpler molecules have more conformations.
- More likely to interact with more than one target binding site.
- May result in increased side effects.

Rigidification

Note

- Endogenous lead compounds are often simple and flexible
- Fit several targets due to different active conformations
- Results in side effects



Strategy

- Rigidify molecule to limit conformations - conformational restraint
- Increases activity - more chance of desired active conformation being present
- Increases selectivity - less chance of undesired active conformations

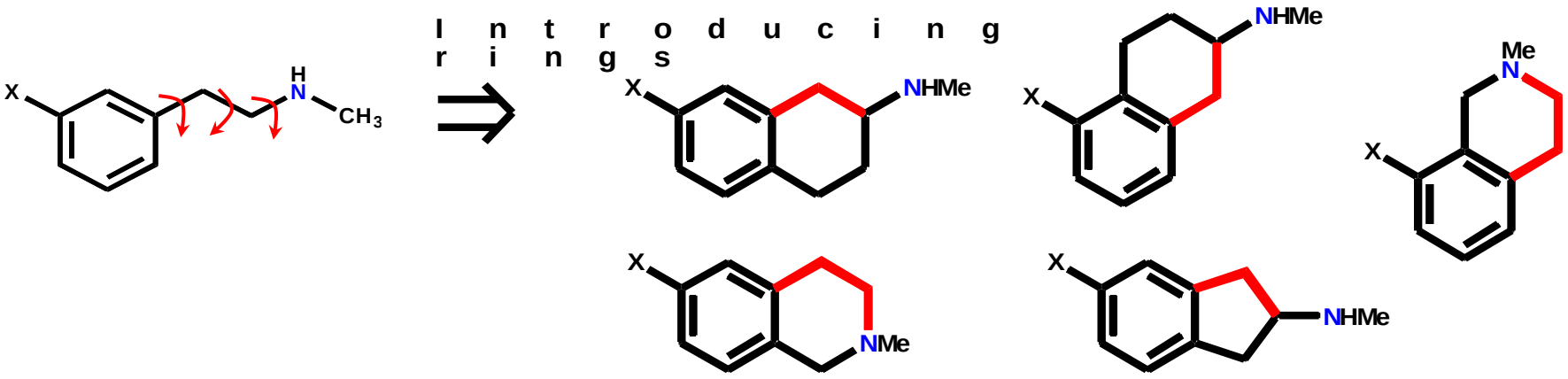
Disadvantage

- Molecule is more complex and may be more difficult to synthesise

Rigidification

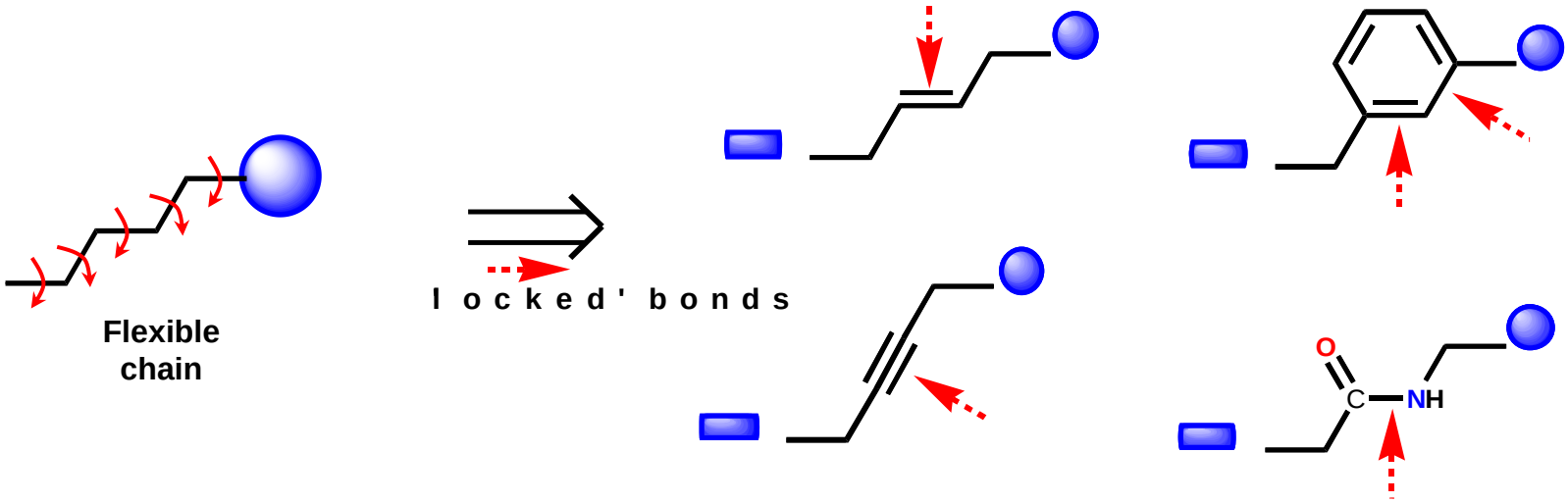
Methods - Introduce rings

- Bonds within ring systems are locked and cannot rotate freely
- Test rigid structures to see which ones have retained active conformation



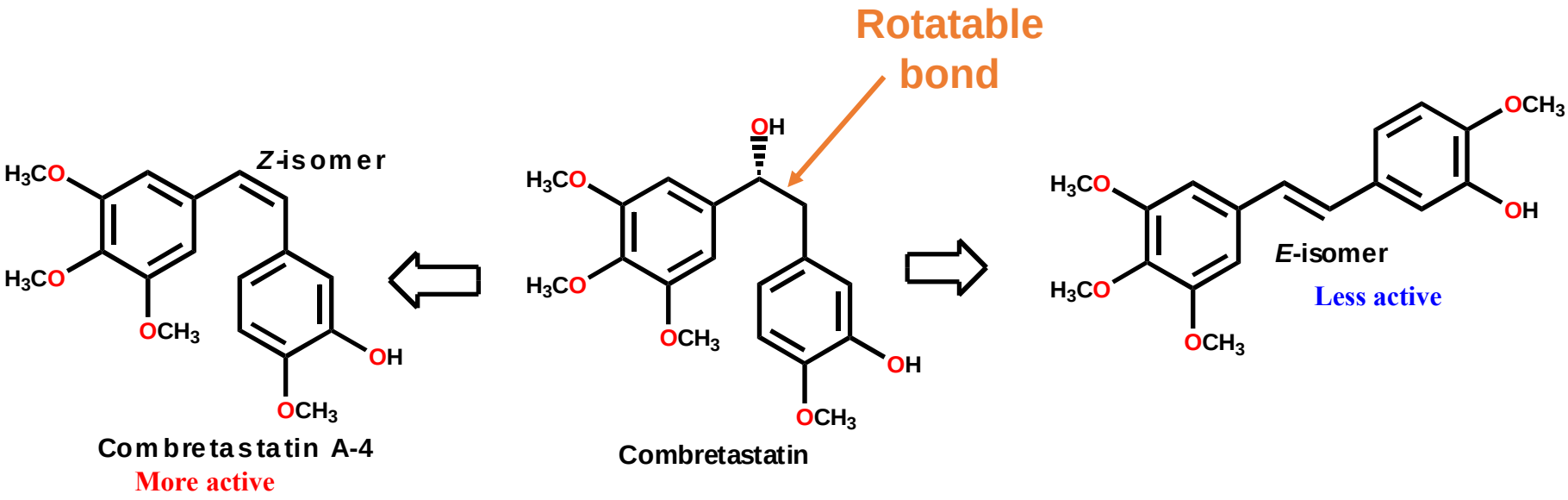
Rigidification

Methods: Introduce rigid functional groups



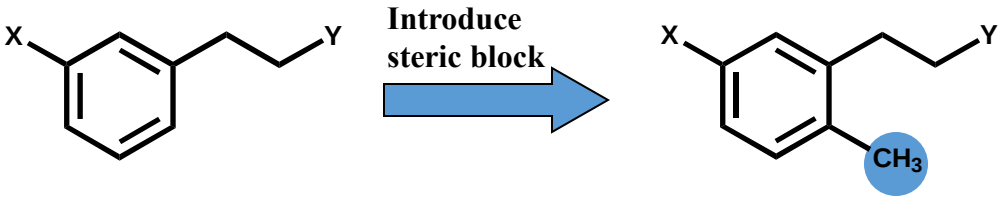
Rigidification

Examples: Combretastatin (anticancer agent)



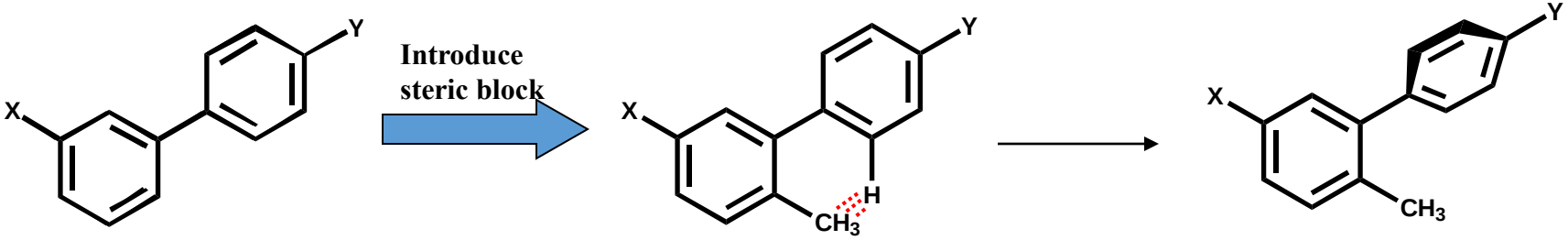
Rigidification

Methods: Steric blockers



Flexible side chain

Steric block



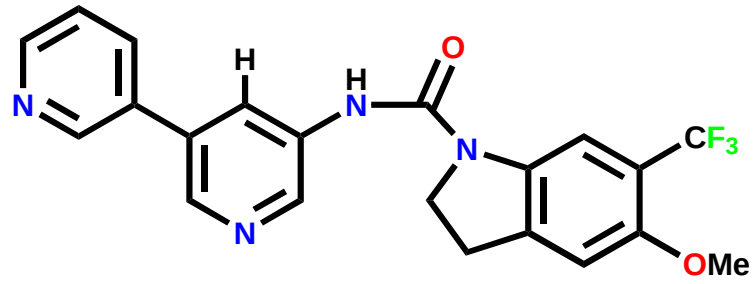
Coplanarity allowed

Steric clash

Orthogonal rings preferred

Rigidification

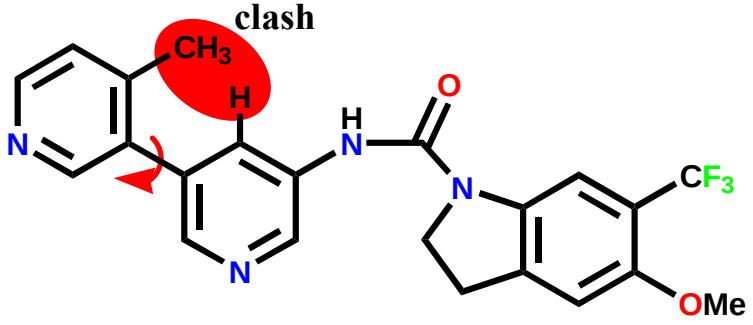
Methods: Steric blockers

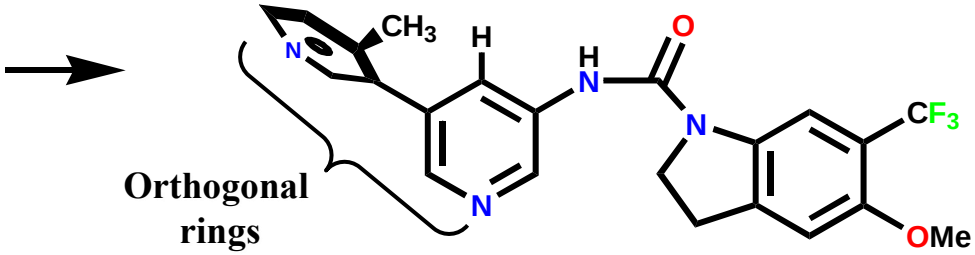


Serotonin antagonist

Introduce methyl group

Steric clash





Orthogonal rings

Structure-based drug design: *De Novo* drug design

Strategy

Carry out drug design based on the interactions between the lead compound and the target binding site.

Procedure

- Crystallise the target protein with a bound ligand
- Acquire the structure by X-ray crystallography
- Download to a computer for molecular modelling studies
- Identify the binding site
- Identify the binding interactions between ligand and target
- Identify vacant regions for extra binding interactions
- Remove the ligand from the binding site *in silico*
- 'Fit' analogues into the binding site *in silico* to test binding capability
- Identify the most promising analogues
- Synthesise and test for activity
- Crystallise a promising analogue with the target protein and repeat the process

Decision making in medicinal chemistry

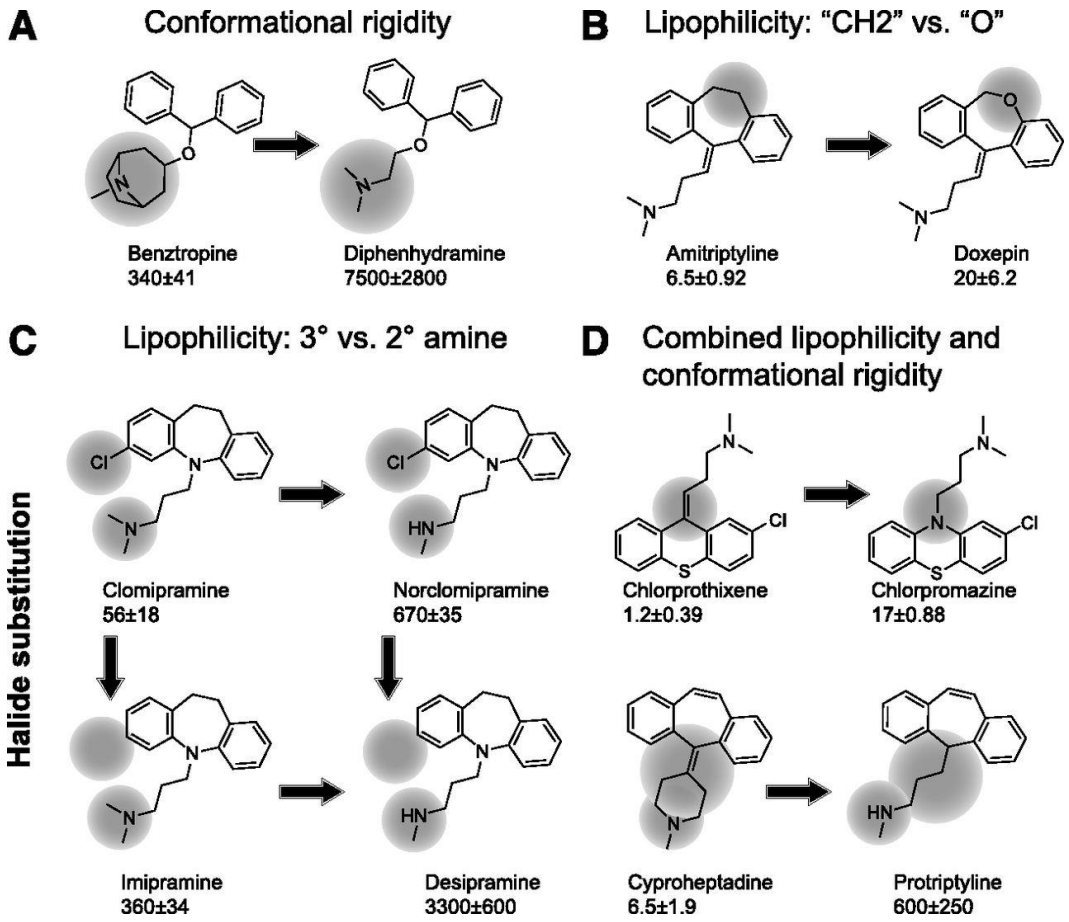
Drug design: optimizing target interactions

Lead compound is a compound that has a desirable biological activity with therapeutic relevance, but typically has some shortcoming that is likely to be overcome through the development of analogs.

Structure-activity-relationship (SAR): the aim is to discover which parts of the molecule are important to biological activity and which not.

(A–D) Structure-activity relationship trends for AaDOP2 receptor antagonists. Compound names and in vitro IC_{50} values (nM) for AaDOP2 antagonism were included.

AaDOP2: *A. aegypti* D1-like dopamine receptor



Conley *et al.* J Pharmacol Exp Ther (2015) 352:53–60

Drug-like properties

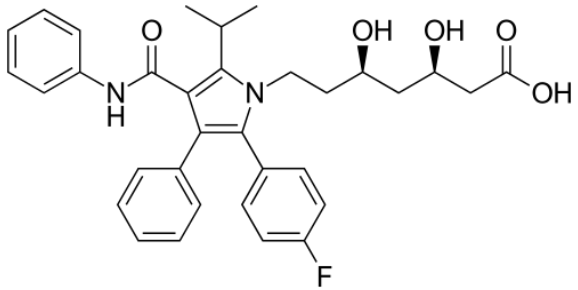
- "Drug-Like"-orally administered small molecules that have appropriate properties in terms of absorption-distribution-metabolism-excretion (ADME) and acceptable toxicity properties.

- Drug like properties (base on analysis of physicochemical properties of marketed drugs):
  - Lipophylicity
  - H-bond donor count

  } did not change significantly in drugs over the last decade

  - number of O and N atoms
  - H-bond acceptor count
  - rotatable bonds
  - molecular weight (Mw)
  - number of rings

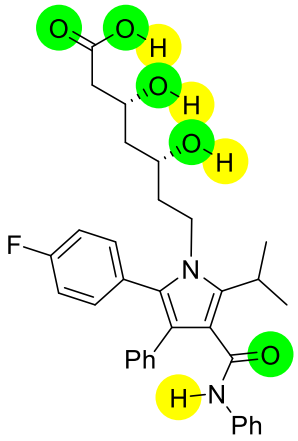
  } Increased in drugs between 1983 and 2002

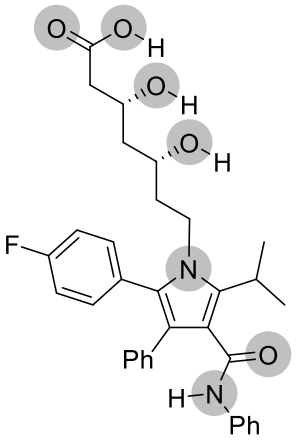


Lipitor (Atorvastatin): lipid lowering drug

Drug-like properties

- Most small molecule drugs <u>are administered orally</u>
- Orally administered drug needs to pass through intestinal lining after digestion, be carried in blood and penetrate lipid-based cell membrane to reach cellular target.
- Physicochemical descriptors:
  a) Lipophilicity (assessed by the logP and logD values).
  b) Molecular weight (Mw)
  c) Number of H-bond donors and acceptors
  d) Topological polar surface area (TPSA)





	Lipitor
Mw	559
logP	4.46
$logD_{7.4}$	1.7
HBD	4
HBA	5
TPSA	112 Å^2

Physicochemical properties for Lipitor (Atorvastatin). HBD HBA TPSA

Drug-like properties

- Lipophilicity (assessed by the logP and $logD_{7.4}$ values).
- 1-octanol: model of cellular membrane

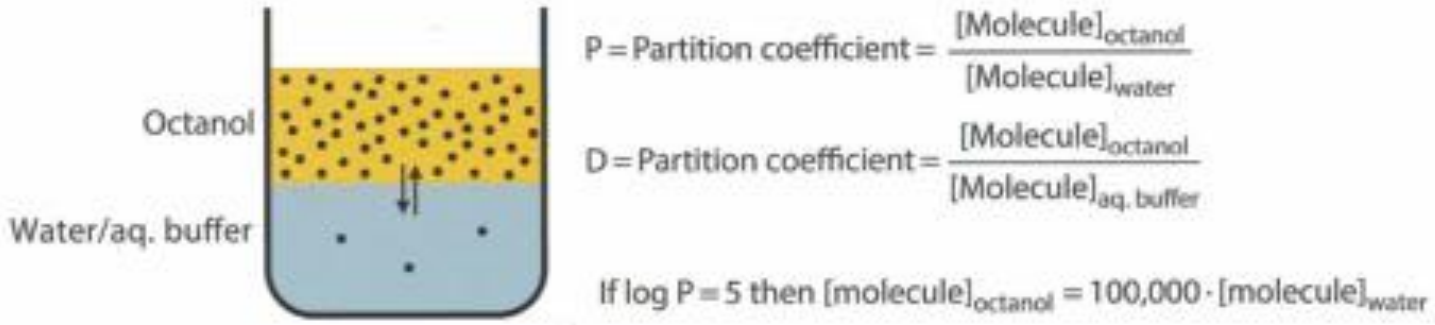


FIGURE 5.1 Definition of the partition coefficient P or D for the equilibrium distribution of a molecule between octanol/water and octanol/aqueous buffer, respectively.

For small molecule drugs log P is typically 3-5, meaning that the molecule has a preference for octanol over water by factor of 1,000-100,000.

The topological polar surface area (TPSA) of a molecule is defined as the surface sum of all polar atoms, primarily O and N, and is therefore roughly count of the number of polar atoms, each of which contributes with cca. 14 $Å^2$. Molecules with a TPSA of greater than140 $Å^2$ poorly permeate cell membranes. For molecules to penetrate BBB a TPSA less than 90 $Å^2$ is required.

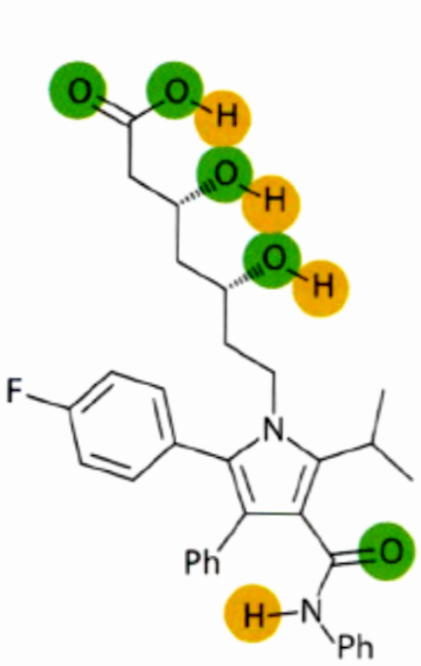
The Lipinski rule of five (Ro5)

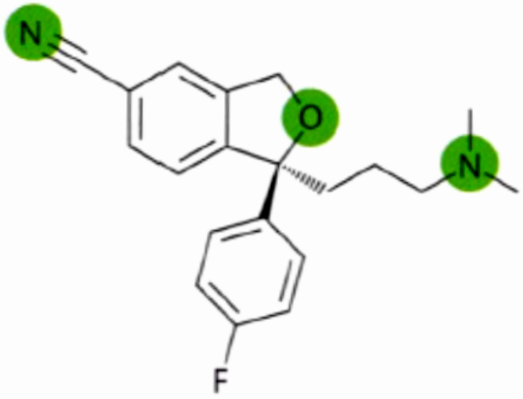
- Based on the analysis of the physicochemical properties of drugs in a database of about 2500 orally available small molecules that had entered at lease phase II clinical trials (Lipinski, C. A. *et al. Adv. Drug Deliv. Rev.* (1997) 23:3-25).
- Based on four essential physical properties :
  1. Mw ≤ 500
  2. log P ≤ 5
  3. HBD ≤ 5
  4. HBA ≤ 10

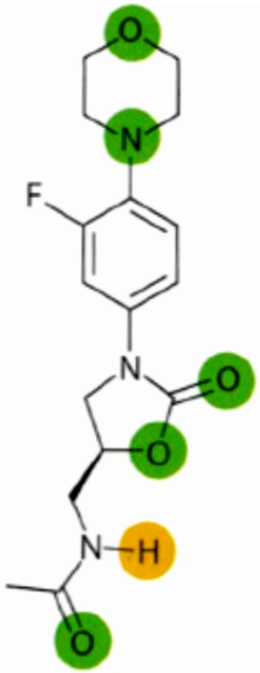
- Nine of ten orally active small-molecule drug candidate that achieve phase II clinical status are found within these boundaries.
- It is unlikely that compounds with two or more “violations” can be orally absorbed.
- Examples:
  - Lexapro (escitalopram)-antidepressant
  - Zycox (linezolid)-antibiotic

- Exemptions (subject to active up-take across the gut by transport protein):
  - Lipitor (atorvastatin)-lipid lowering drug

Examples:



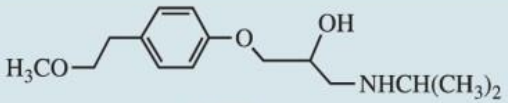




Lipinski Criteria	Lipitor (Atorvastatin)	Lexapro (Escitalopram)	Zyvox (Linezolid)
MW < 500	559	324	337
Log P < 5	4.46	3.13	0.17
HBD count < 5	4	0	1
HBA count > 10	5	3	5
Number of Ro5 violations	1	0	0

Note: The H-bond donors and acceptors are indicated by the orange and green colors, respectively.

TABLE 2.13 Change in log *D* as a Function of pH for Metoprolol (**2.107**)[1]



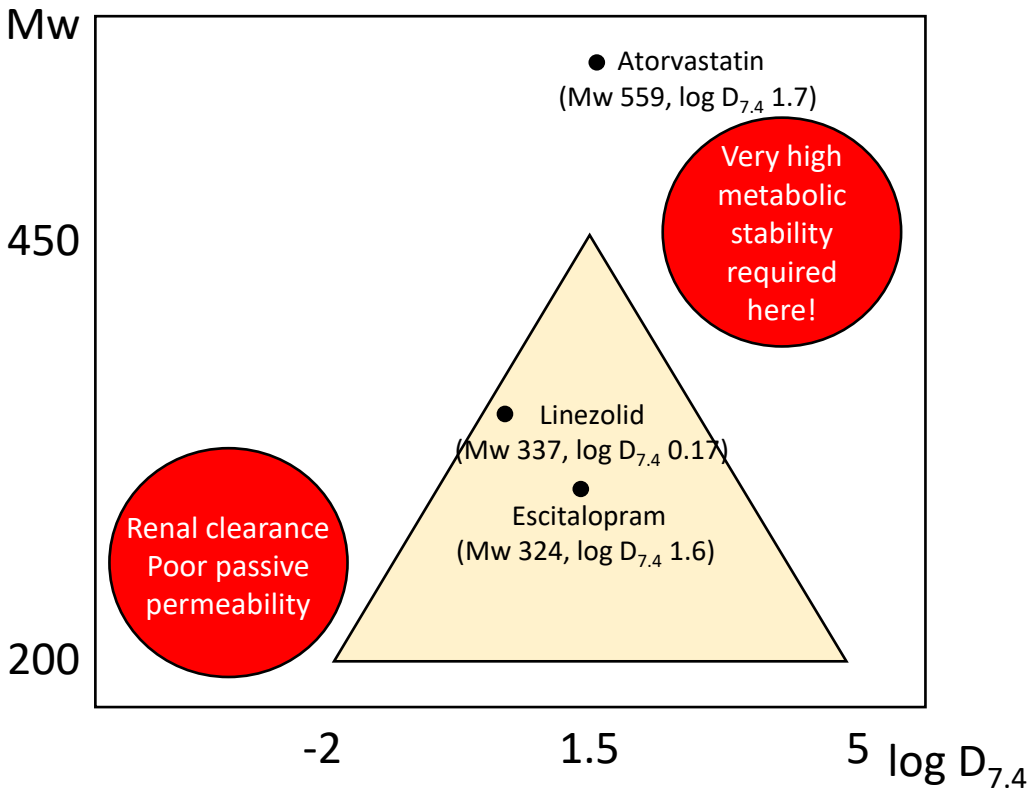
2.107

pH	log *D*
2.0	−1.31
3.0	−1.31
4.0	−1.31
5.0	−1.28
5.5	−1.21
6.0	−1.05
6.5	−0.75
7.0	−0.34
7.5	0.12
8.0	0.59
8.5	1.03
9.0	1.39
10.0	1.73

[1]*The authors are grateful to Karolina Nilsson and Ola Fjellström (AstraZeneca) for providing the log D values as a function of pH using ACD software.*

The Golden Triangle Model

- The guidelines developed from *in vitro* ADME data and computational data with the goal to aiding medicinal chemists to identify permeable and metabolically stable compounds.



- Apex: Mw = 450 and log $D_{7.4}$ = 1.5
- Baseline: log $D_{7.4}$ = -2 to log $D_{7.4}$ = 5 at Mw = 200

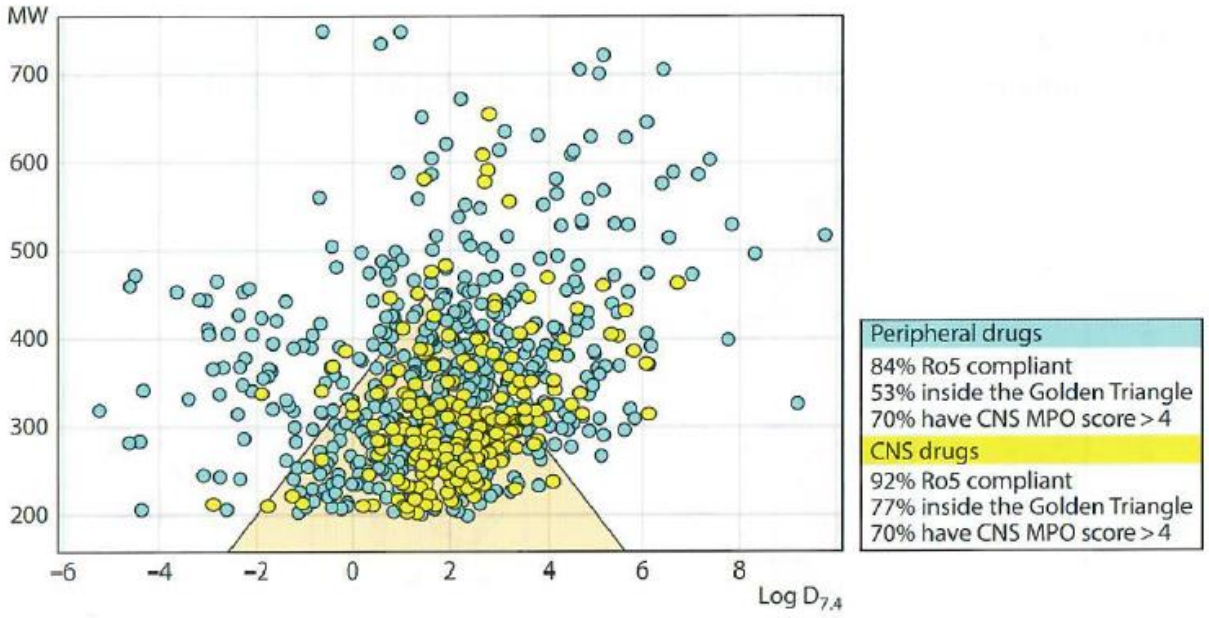
- Compounds with combined good liver microsomal stability and good permeability are typically found within the Golden Triangle.

Decision making in medicinal chemistry

- Ligand efficiency (LE)
- Ligand lipophilic efficiency (LLE)

Allow optimization of the potency averaged for molecular size and normalized for lipophilicity.

- The ADME properties of compounds deteriorate with increasing Mw and/or log P.



Plot of Mw vs. log $D_{7.4}$ for 591 peripheral drugs (cyan) and 273 drugs (yellow). The Golden Triangle is indicated.

Decision making in medicinal chemistry

Ligand efficiency (LE):

- LE is measure of ΔG of binding in relation to the molecular size (represented by the number of non H-atoms, called the heavy-atom-count, HAC).

$$[\text{Ligand}]_{\text{free}} + [\text{Target}]_{\text{free}} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} [\text{Ligand:Target}]_{\text{complex}}$$

$$\frac{1}{K_d} = K_a = \frac{k_a}{k_d} \qquad \Delta G = -RT \ln K_a$$

- LE is defined as:

$$LE = \frac{\Delta G}{HAC} = \frac{-RT \ln K_a}{HAC}$$

The majority of oral drugs have HAC between 10 and 30.

Decision making in medicinal chemistry

Ligand efficiency (LE):

$$\Delta G = -RT \ln K_a = 300K \times 1.98 \times 1^{-3}\ \text{kcalmol}^{-1}K^{-1} \times \ln K_a$$

$$\Delta G = -0.5961\ \text{kcalmol}^{-1} \times 2.303 \times \log K_a = -1.4 \log K_a$$

$$LE = -1.4\ \frac{\log K_a}{HAC}$$

- If K_a is not available then:

$$LE = -1.4\ \frac{\log IC_{50}}{HAC}$$

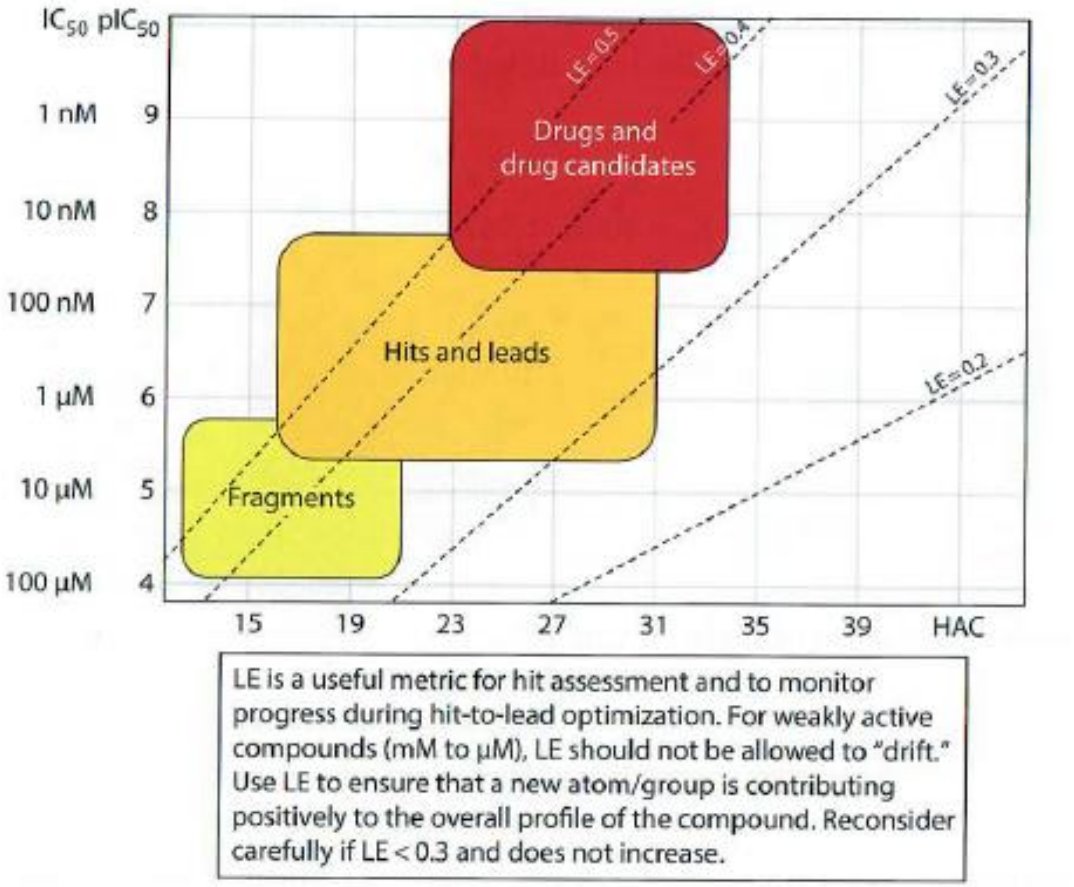
Potency

Size

A good starting point for a lead would be a LE above 0.3.

Decision making in medicinal chemistry

Ligand efficiency (LE):



Plot of pIC_{50} vs HAC. Lines of compounds with equal LE values are indicated. The typical location ranges from fragments, hits and leads, and drug candidates are illustrated.

Decision making in medicinal chemistry

Ligand efficiency (LE): weaknesses

- HAC treats all heavy atoms equally, e.g. introduction of CH_3 is identical to introduction of NH_2, OH, F, Cl or Br. It has been shown that C, N, O, S or halogen do not contribute to potency identically.

- LE is not independent of HAC. A 10 fold change in potency per heavy atom does not result in constant LE, e.g. 15 HAC compound with a pIC_{50} of 3 does not have the same LE as a 16 HAC compound with pIC_{50} of 4. $\Delta pIC_{50} = 1$, $\Delta HAC = 1$, $\Delta LE = 0.07$

- Does not take log P into account.

- Advisable not to focus solely on the metric LE but keep other factors in mind such as log P, solubility, and metabolic stability.

Decision making in medicinal chemistry

Lipophilic ligand efficiency (LLE): balancing potency with respect of lipophilicity

- High lipophilicity increases the likelihood that a compound will bind to multiple targets and result in off-target pharmacology that may limit the therapeutic window due to adverse events or dose-limiting toxicity. Increase in log P can lead to decreased selectivity and suboptimal ADME profile.
- Lipophilicity is related to low solubility and increased susceptibility to oxidative metabolism, typically in liver.
- LLE is a way of balancing potency normalized with respect to lipophilicity.

$$LLE = -\log K_a - \log P = pK_a - \log P$$

LLE=0 not selective
LLE=5 100,00 time more selective for target

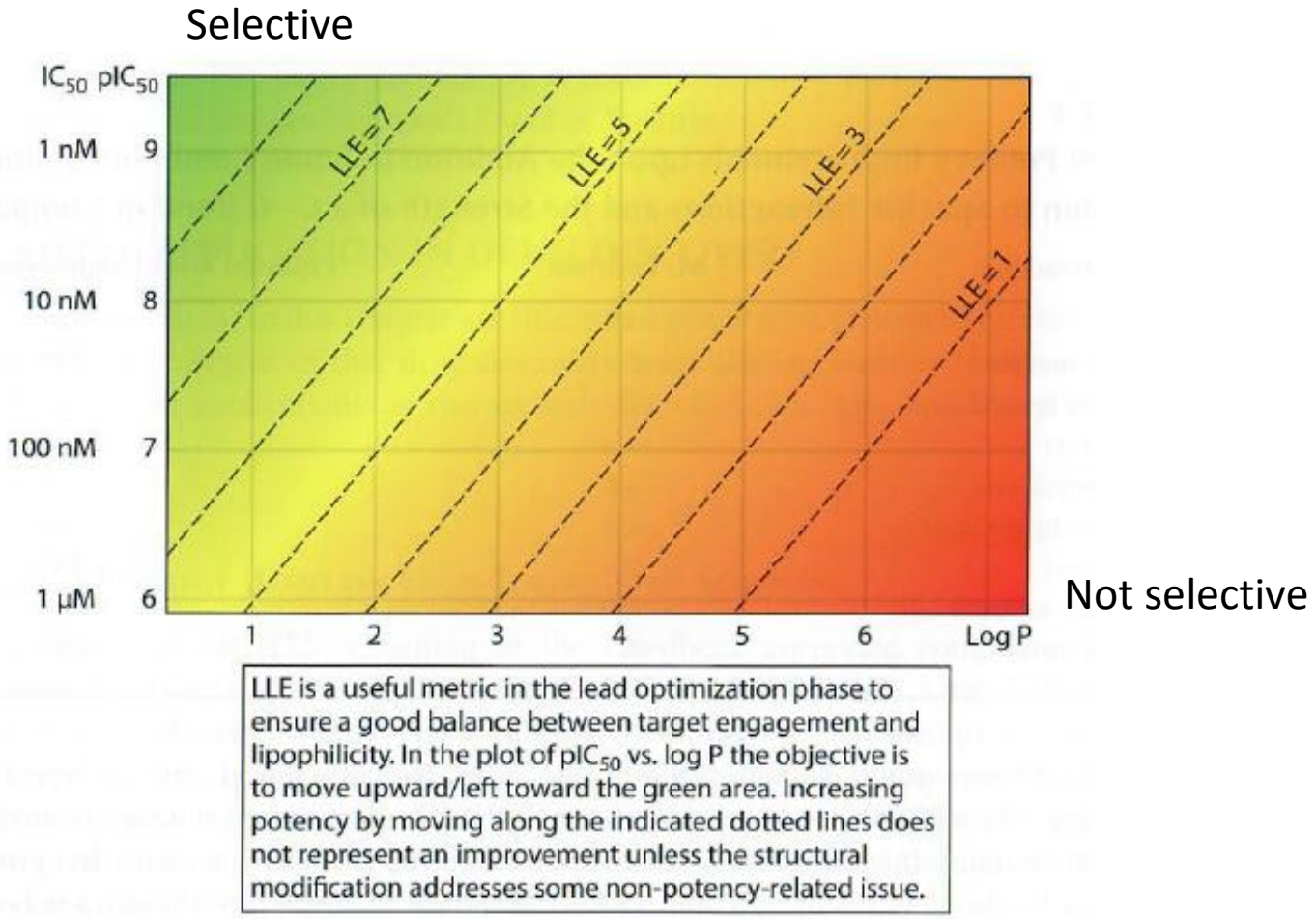
- If K_a is not available then:

$$LLE = -\log IC_{50} - \log P = pIC_{50} - \log P$$

- Compounds that have the potential to become development candidates should have LLE in the range of 5-7 or greater or compounds attractive as leads in the range of 3-5 (or greater).

Decision making in medicinal chemistry

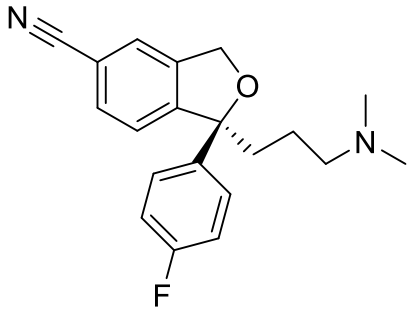
Lipophilic ligand efficiency (LLE):



LLE plot of pIC_{50} vs log P. Lines of compounds with equal LLE values are indicated.

Decision making in medicinal chemistry

Calculation of LE and LLE:



Lexapro
(escitalopram)

Parameter	Value	Calculation of LE and LLE
Target affinity	2.1 nM	$LE = \frac{-1.4 \times \log(2.1 \times 10^{-9})}{36} = 0.36$
HAC	36	$LLE = -\log(2.1 \times 10^{-9}) - 3.13 = 5.55$
log P	3.13	

- antidepressant (selective serotonin reuptake inhibitor)

Expected potency improvements upon the addition of some common functional group in relation to specific interactions and the strength of a C-C bond in comparison.

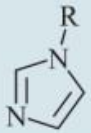
Group/interaction	ΔG kcal/mol	Expected X-Fold Improvement in Potency
H-bond	-1.4	10
Salt bridge (ion-pair)	-3.4	300
Methyl from lipophilicity	-0.7	3
Buried methyl	-1.4	10
Cl from lipophilicity	-1.0	5
Phenyl from lipophilicity	-2.5	60
Buried phenyl	-3.4	300
Ethane rotating barrier	3.0	-140
C-C bond energy	80	$1x10^{57}$

TABLE 2.16 Median Values of Calculated Experimental Properties for a Set of Marketed CNS Drugs

Parameter	How Determined	Median Value of Parameter
CLog *P*	Calculated	2.8
CLog *D* (pH 7.4)	Calculated	1.7
MW	Calculated	305.3
TPSA	Calculated	44.8
Hydrogen bond donors (HBD)	Calculated	1
pK_a	Calculated or by titration experiment	8.4
Passive permeability (Papp)	MDCK (canine kidney) cells	>10E-5 cm/s
Efflux to Influx Ratio	MDCK cells expressing MDR1	≤2.5
Liver microsome stability (CLint,u)	In vitro microsome preparations	≤100 mL/min/kg
Ligand efficiency (LE)	Potency assay, MW	0.46
LLE	Potency assay, CLog *P*	6.4
LELP	Calculated from LE and CLog *P*	5.9

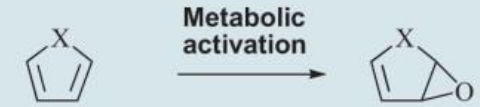
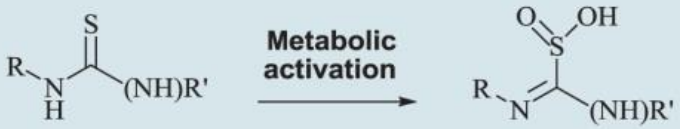
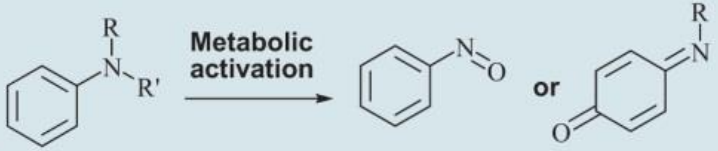
Elimination of functional groups viewed as undesirable in drugs

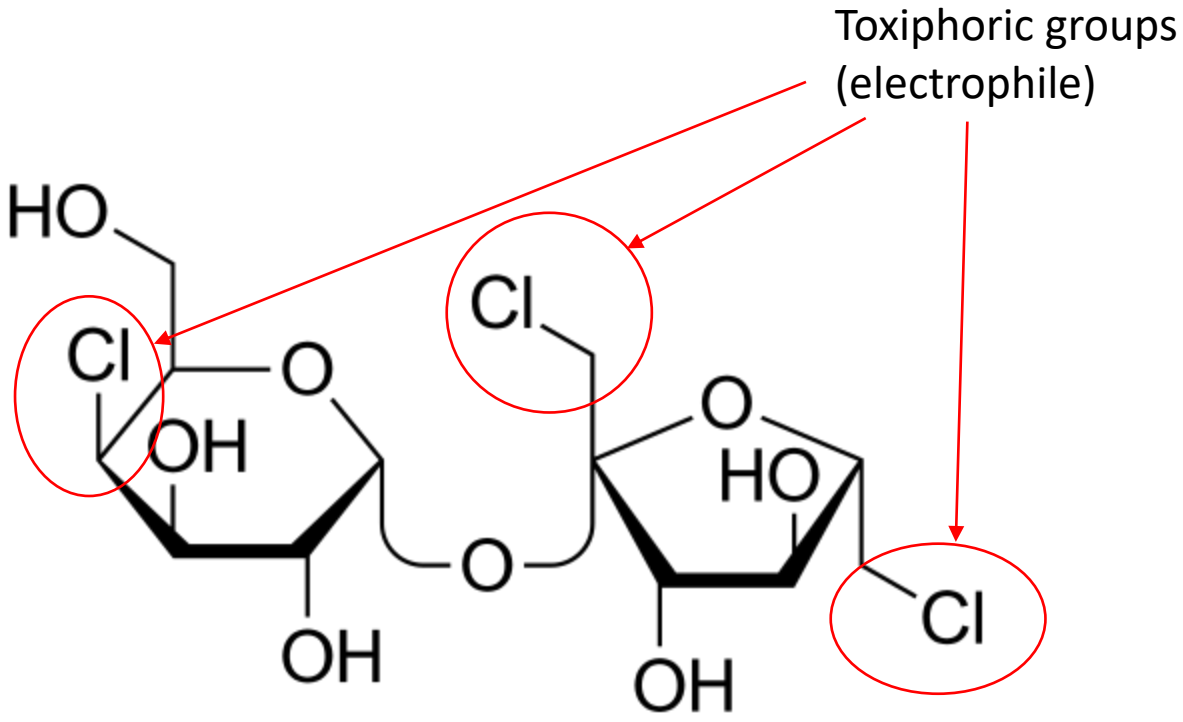
TABLE 2.4 Representative Groups Viewed as Toxicophoric Because of the Reactivity

Toxicophoric Group	Rationale
 **EWG = electron withdrawing group, e.g., carbonyl, cyano, etc.**	Michael acceptor; electrophilic group that can alkylate biological nucleophiles, for example, cysteine -SH
	Epoxide; electrophilic group that can alkylate biological nucleophiles
	Imidazole; can chelate metals, for example, iron in heme proteins such as cytochrome P450 enzymes

Elimination of functional groups viewed as undesirable in drugs

TABLE 2.5 Representative Groups Viewed as Toxicophoric Because They May be Metabolized to Undesirable Moieties

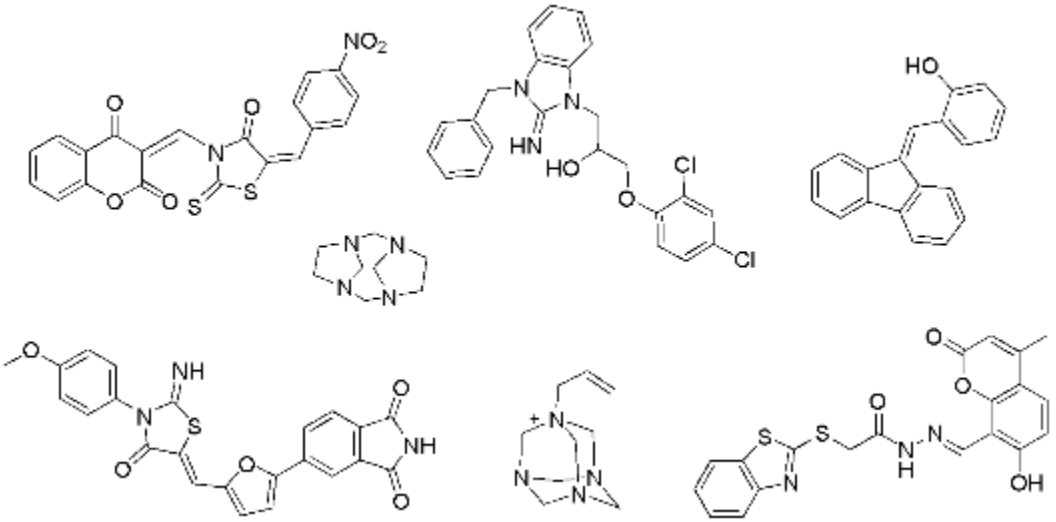
Toxicophoric Group	Rationale
 X = O or S	Furans and thiophenes; tend to be metabolized to electrophilic epoxides
	Thioamides and thioureas; tend to be metabolized to electrophilic imines
	Anilines; tend to be metabolized to electrophilic nitroso or quinone derivatives



Sucralose (Splenda) is an artificial sweetener and sugar substitute.

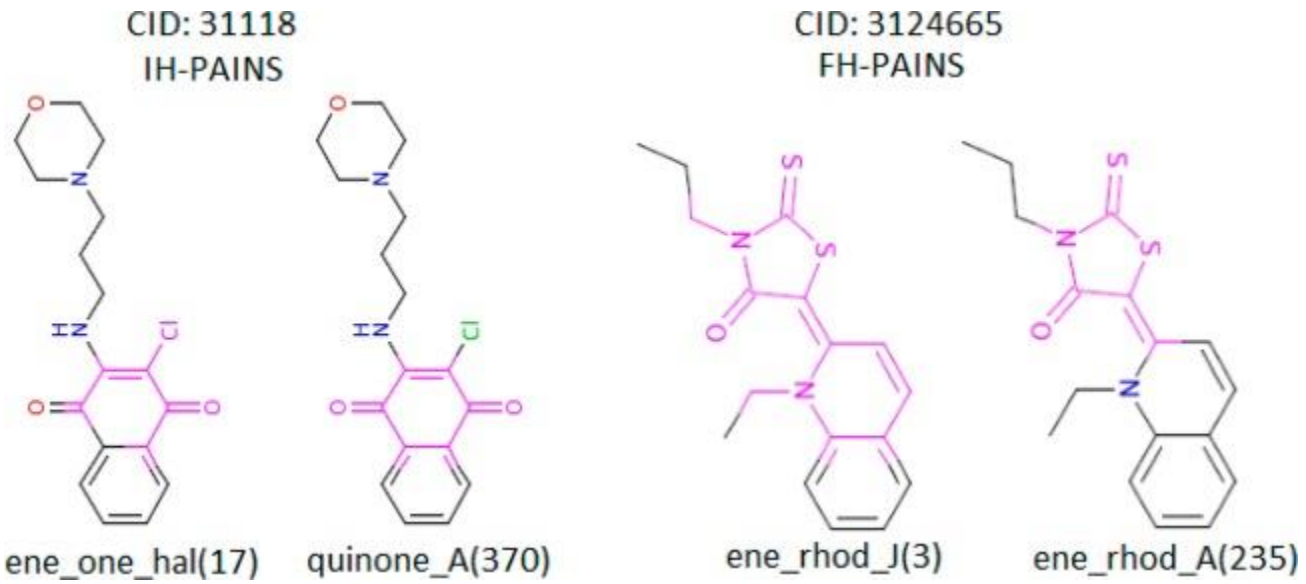
Pan-Assay Interference Compounds (PAINS)

Pan-Assay Interference Compounds (PAINS) are defined by their ability to show activity across a range of assay platforms and against a range of proteins. The most common causes of PAINS activity are metal chelation, chemical aggregation, redox activity, compound fluorescence, cysteine oxidation or promiscuous binding. Many PAINS have multiple functionalities, causing different types of interference and resulting in *in vitro* and *in vivo* activity.



PAINS encompass some 400 structural classes, but more than half of PAINS in a typical library fall into just 16 easily recognizable categories.

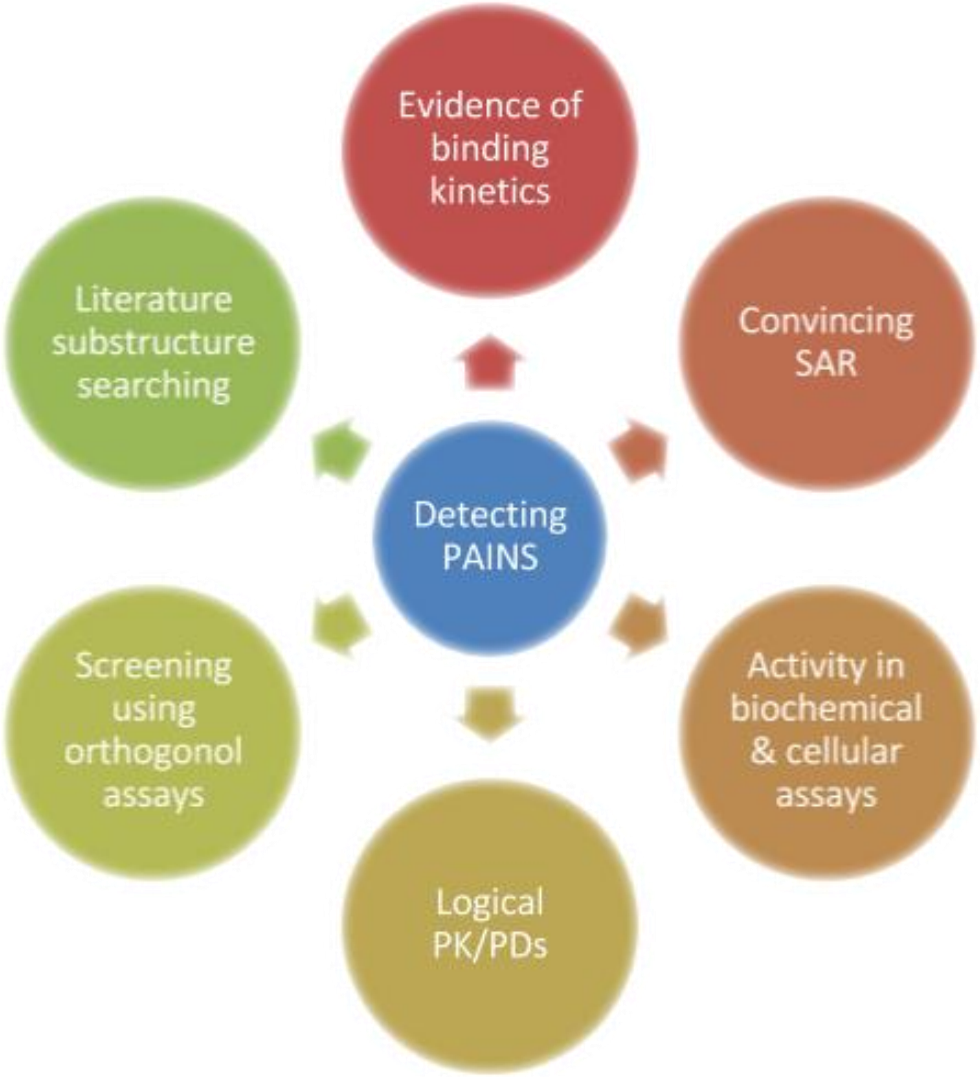
Software tools can filter PAINS from screening libraries, but they are no match for sharp-eyed scientists.



Compounds with multiple PAINS alerts

J Chem Inf Model. 2017 Mar 27; 57(3): 417–427.

How a PAINS compound can be identified.



https://chembiohub.ox.ac.uk/blog/2015/03/10/pains.html