Grand Challenges
African Drug Discovery Accelerator

Drug Discovery and Development Course

Principles of Compound Optimization

Greg Basarab

Supported by

Gates Foundation LifeArc H3D FOUNDATION

Learning Objectives

- Make a molecule great
- Consider key parameters associated with drug-likeness
- Employ computational tools to assess compound parameters
- Measure key compound parameters
- Validate compound parameters relative to project goals

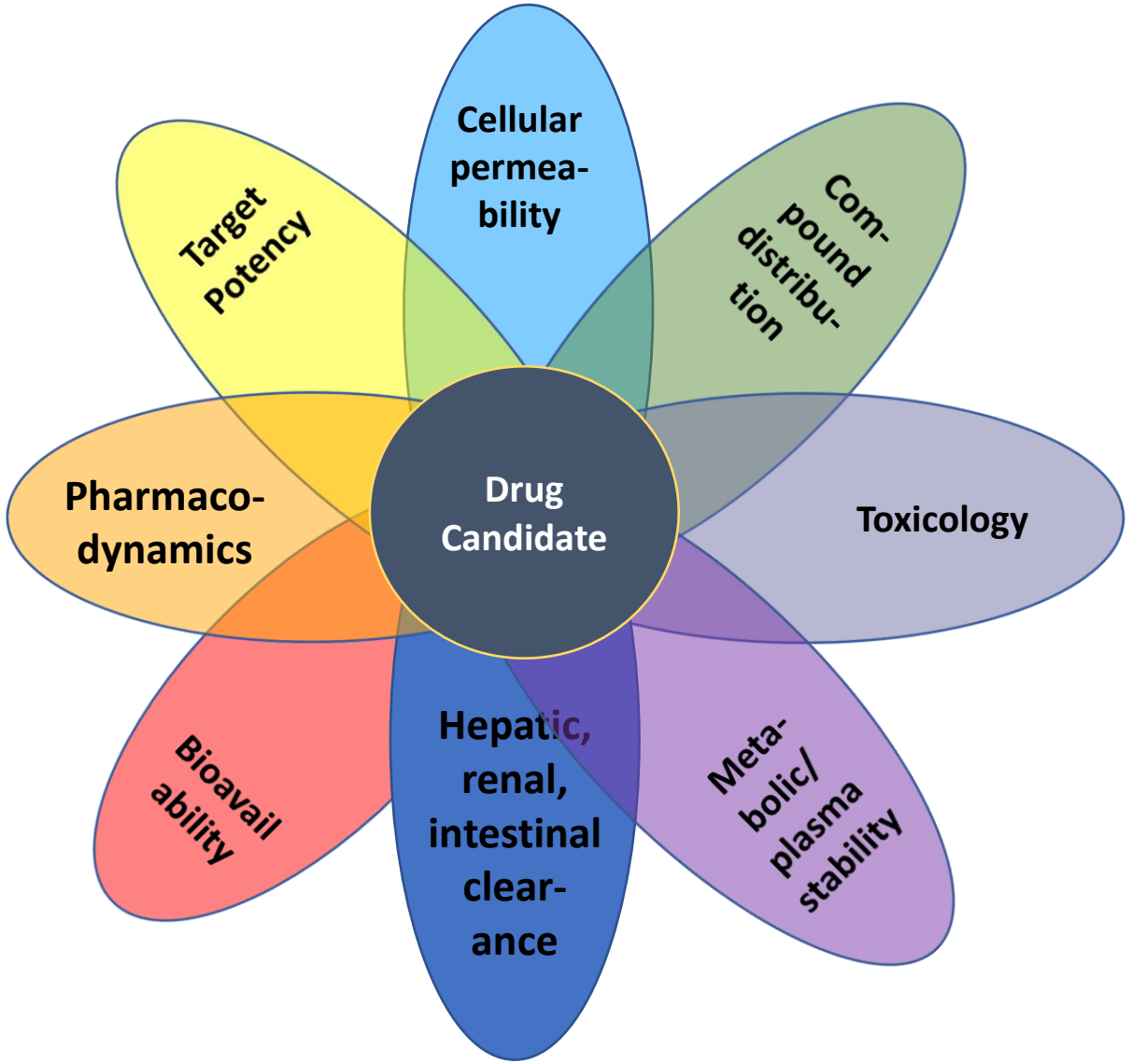
Druglikeness

Similar compound properties to existing drugs

or…

Think inside the box

Compound biological attributes

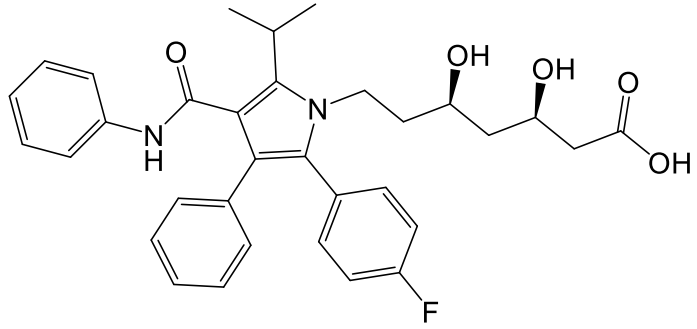


Lipinski ‘rule of five’

Oral drugs do not violate more than one of the following parameters:

- ≤ 5 H-bond donors
- ≤ 10 H-bond acceptors
- MW ≤ 500 Da
- LogP ≤ 5

Med Chemist @ Pfizer for 34 years



Atorvastatin (Lipitor)

3 H-bond donors

5 H-bond aceptors

MW = 558

ClogP = 4.5

Rot. bonds = 13

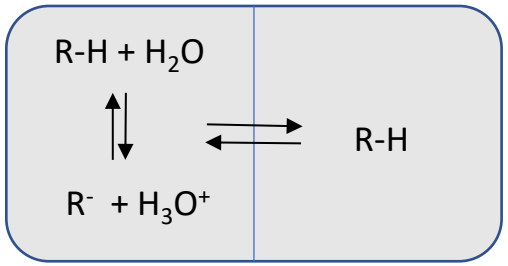
PSA = 110 Å^2

F = 14%

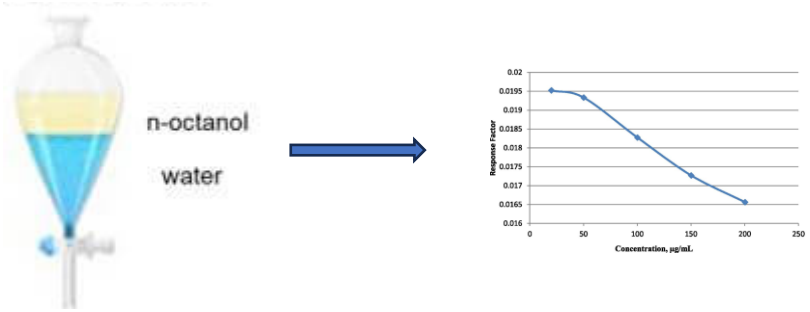
Best used in early hit identification stage to insure a viable scaffold is pursued

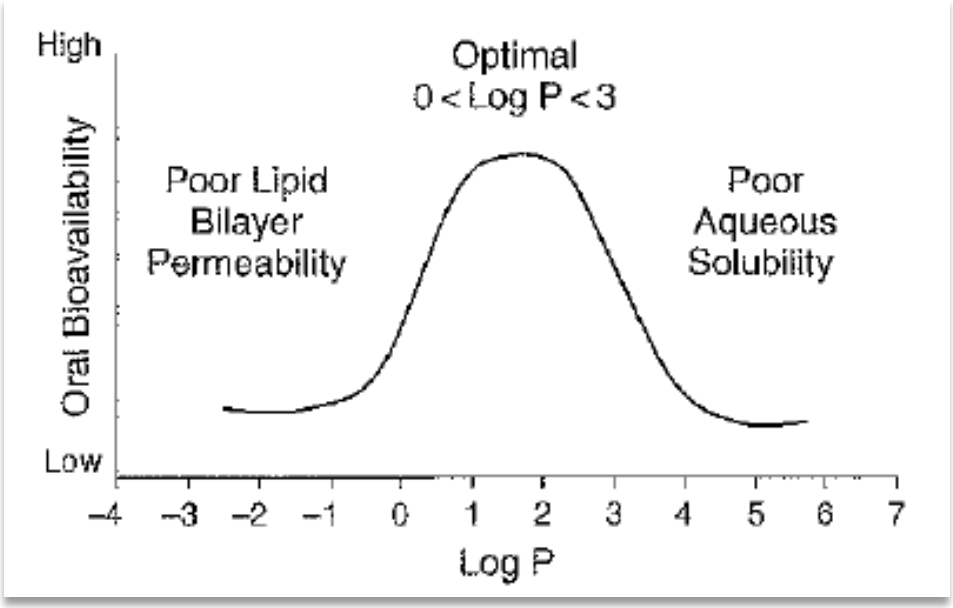
Lipinski, C. Drug Disc. Today, Tech. 1(4), 2004, 337.

(c)LogD and (c)LogP



buffer 1-octanol





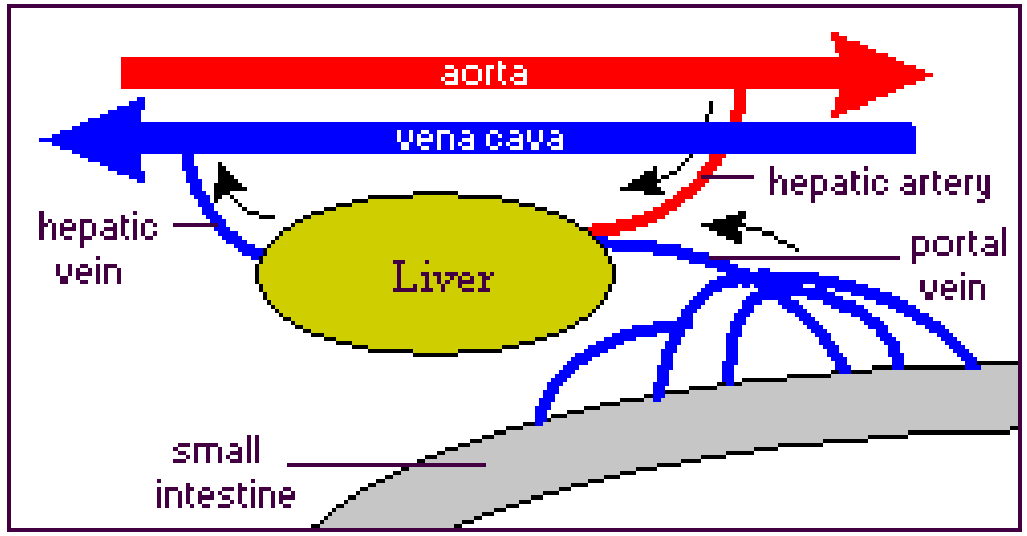
D (Distribution Coefficient) = $\frac{[drug]_{octanol}}{[drug]_{buffer\ (7.4\ usually)}}$

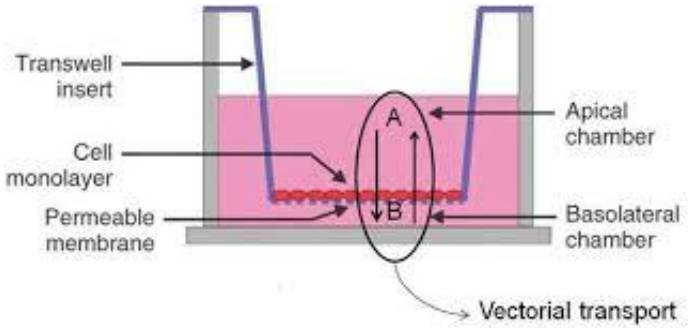
P (Partition Coefficient) = $\frac{[drug]_{octanol}}{[unionized\ drug]_{buffer\ (7.4?)}}$

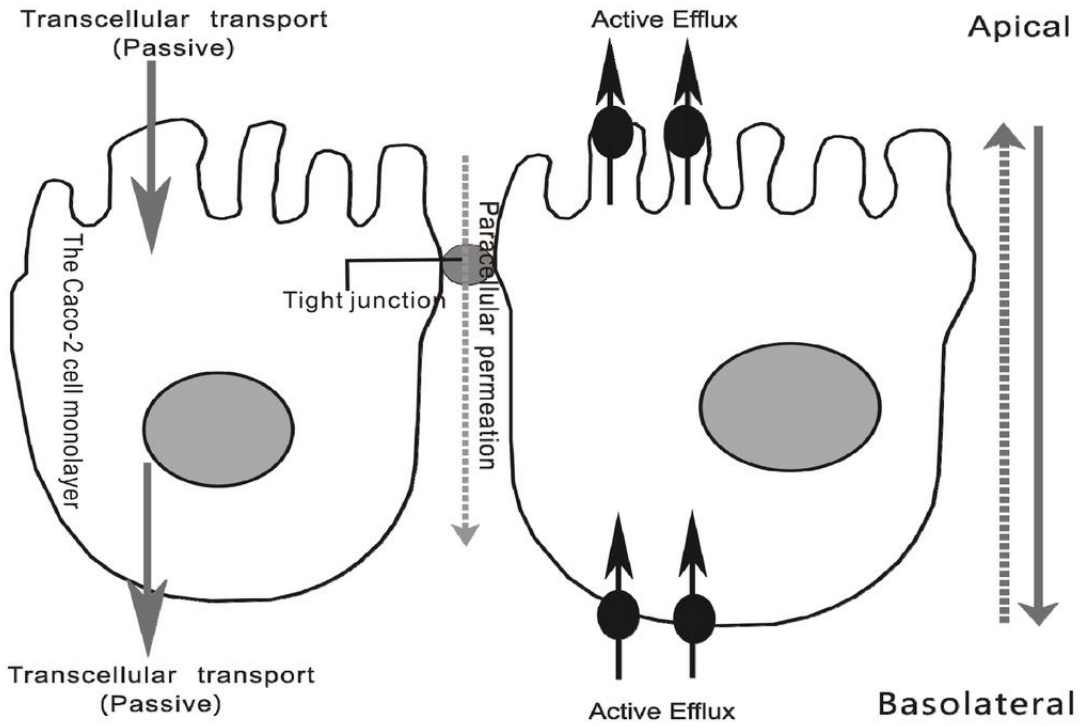
Neutral species optimally permeate cellular lipid bilayers (passive diffusion)
- Expression of activity against intracellular targets
- Intestinal drug absorption
- CNS penetration

Transcellular permeability

Oral drugs transit small intestine epithelium to bloodstream & portal vein







- Cancer coli (Caco-2) cells: human epithelial colorectal adenocarcinoma cells that aggregate directionally on filter surfaces
- Madin-Darby canine kidney (MDCK) cells: derived from canine kidney tissue

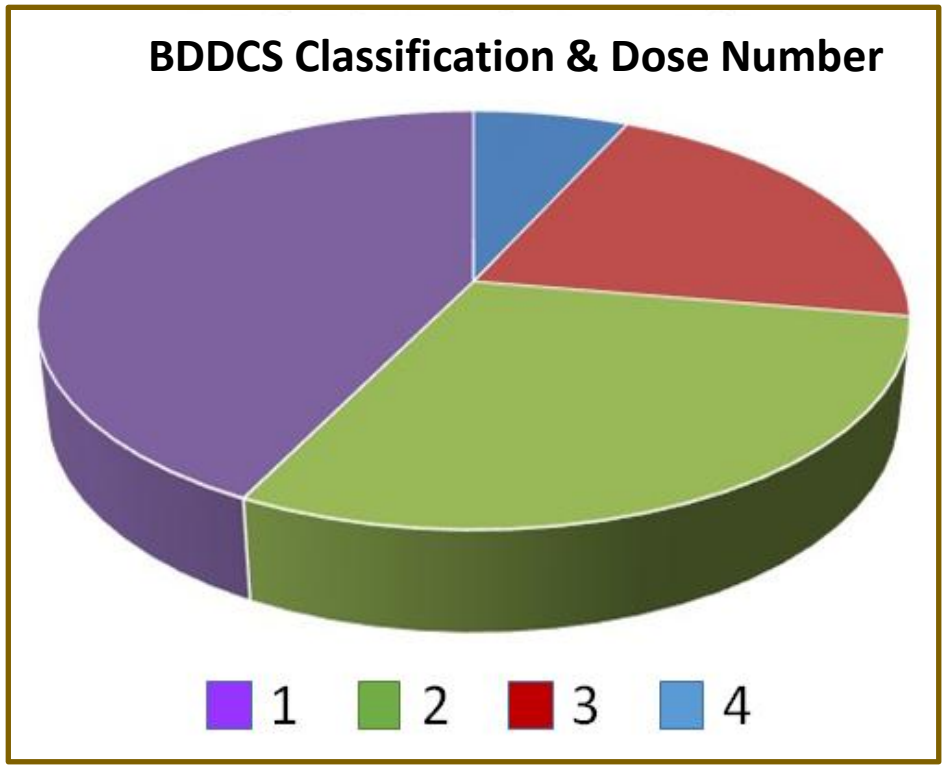
Solubility

- Solubility is perhaps the most important ADME property
- Valid *in vitro* data relies on compound being soluble under the conditions

Comment on the data for this compound	
Solubility	< 1 μM
IC_{50} (HepG2 cells)	> 200 μM

- Poorly soluble compounds are difficult to formulate
  - Assessments of PK, PD, toxicology become problematic
- High solubility is a must for IV administration
- Thermodynamic solubility should be measured in:
  - pH 7.4 buffer (IV), 6.5 and/or 7.4 (PO)
  - Biorelevant media: FaSSIF – Fasted State Simulated Intestinal Fluid
    (PO cmpds) FeSSIF – Fed State Simulated Intestinal Fluid
    FaSSGF – Fasted State Simulated Gastric Fluid

Solubility and Permeability Together (PO drugs)



BSC Classification – solubility 250 mg & permeability (≥90% absorption drug + metabolites)

BDDCS = Biopharmaceutics Drug Disposition Classification System

$$\text{Dose Number} = \frac{\text{Max. dose}}{250_{*}\text{clogP}_{*}\text{Sol}}$$

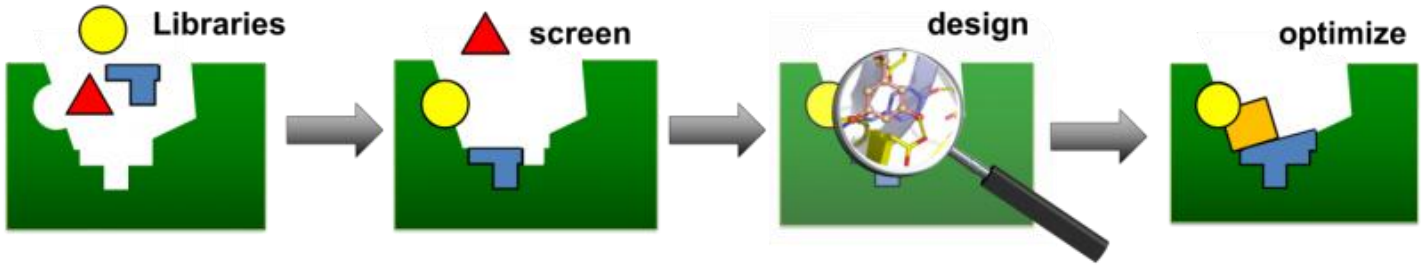
BDDCS 1 – high solubility & high permeability
Ave. Max. Dose = 113 mg
BDDCS 2 – low solubility & high permeability
Ave. Max. Dose = 204 mg
BDDCS 3 – high solubility & low permeability
Ave. Max. Dose = 276 mg
BDDCS 4 – low solubility & low permeability
Ave. Max Dose = 392 mg

Max Absorbable Dose =

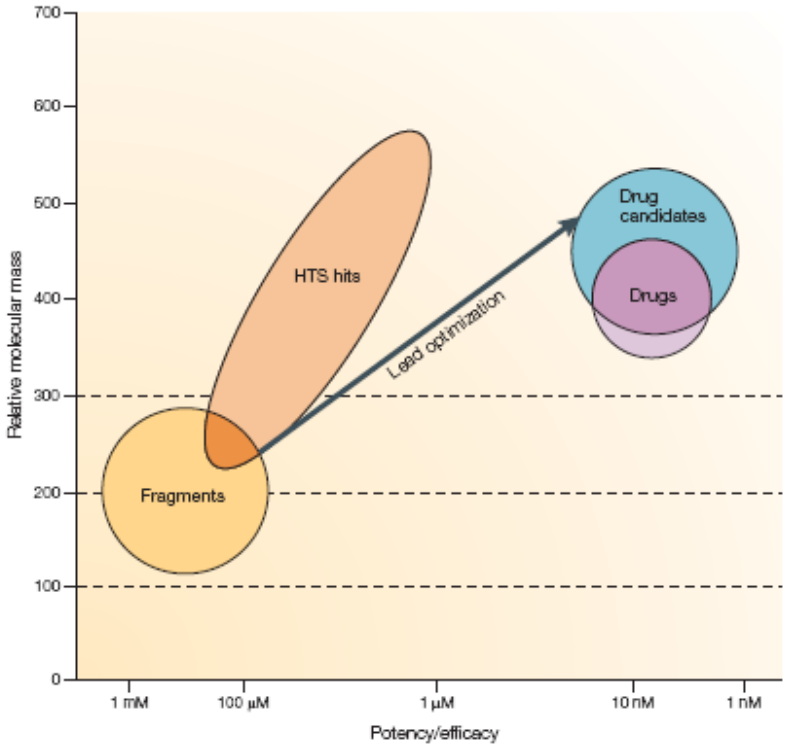
$$\textbf{Sol} \times \textbf{K}_{\textbf{a}} \times \textbf{SI}_{\textbf{WV}} \times \textbf{SI}_{\textbf{TT}}$$

Fragment based lead discovery

Rule of three: begin small because growth is (seemingly) inevitable



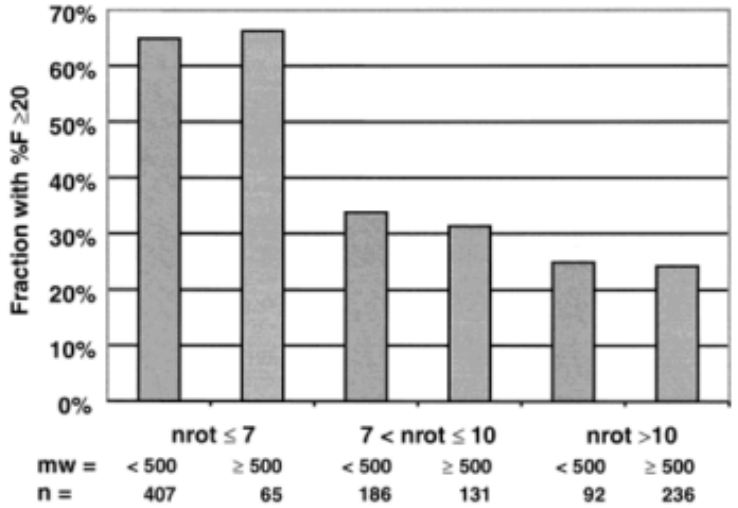
- **≤ 3 H-bond donors**
- **≤ 3 H-bond acceptors**
- **MW ≤ 300 Da**
- **LogP ≤ 3**
- # rot bonds ≤ 3
- PSA ≤ 60 Å^2

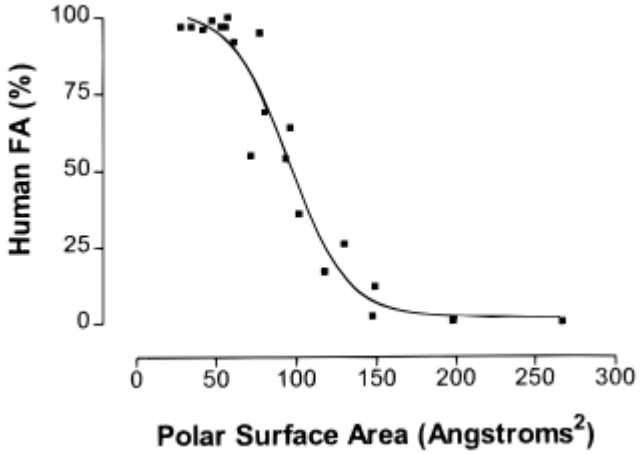


Congreve, M et al. H. *Drug Disc. Today*, 8(19), 2003, 876; Jhoti et al. *Nat. Rev. Drug Disc.* 12(8), 2013, 644; Mortenson et al. *J. Med. Chem.62(8)* , 3857, 2019.

Other correlated parameters

- ≤ 10 rotatable bonds





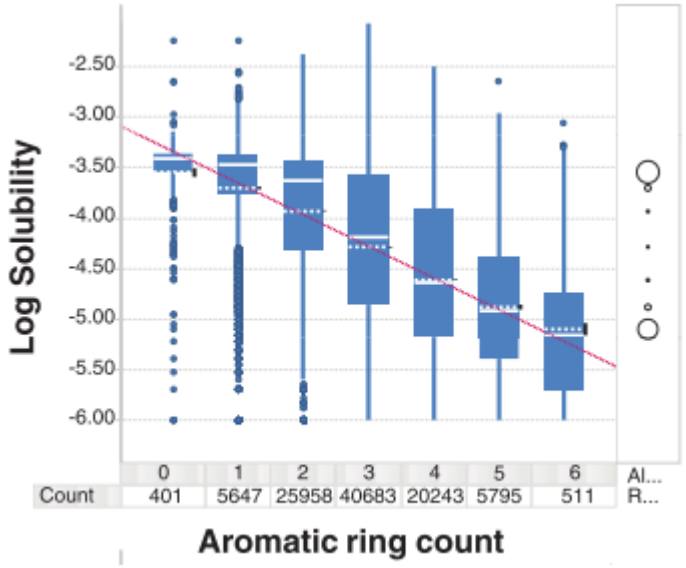
<u>Topological polar surface area</u>

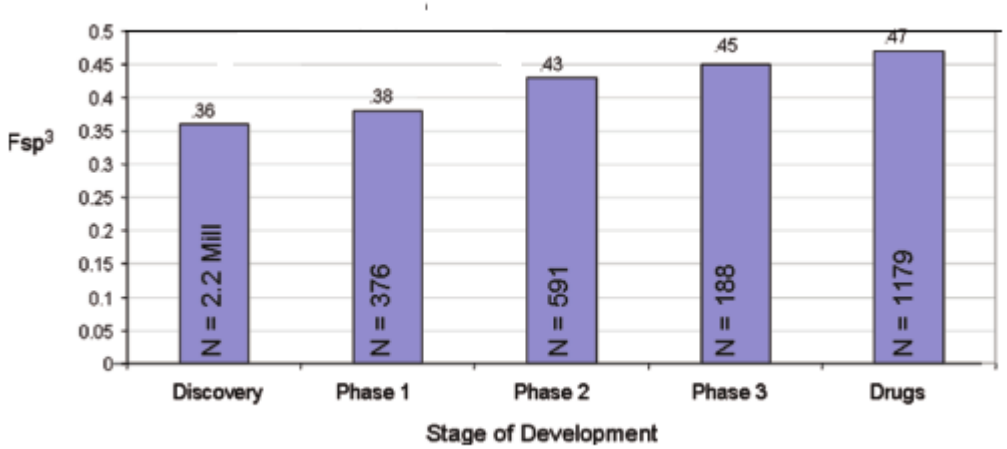
- TPSA ≤ 140 Å^2

Veber, D.F. et al *J. Med Chem.* 45(12), 2002, 2615; Clark, D. et al *J. Pharm. Sci.* 45(12), 2002, 2615.

sp^3 versus sp^2 (globular versus flat)

Number or aromatic rings - sp^3 count - Aromatic vs heavy atoms - Aromatic Proportion

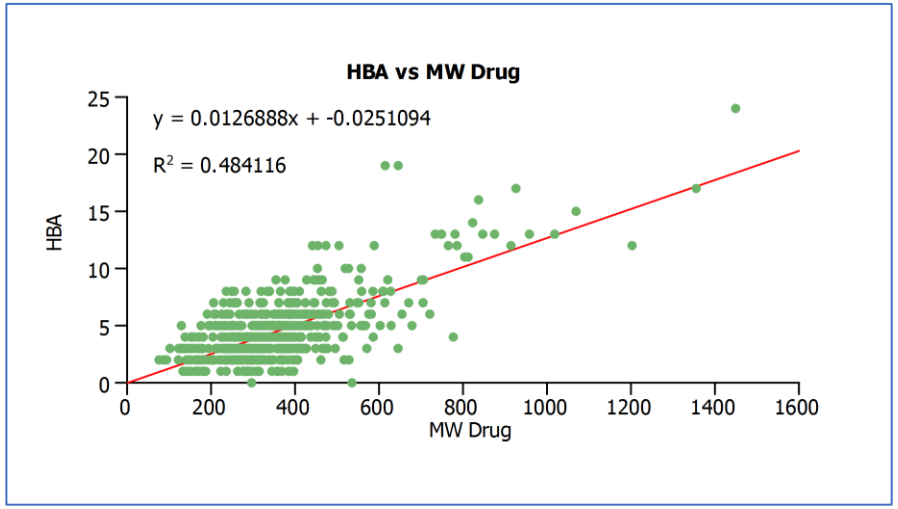


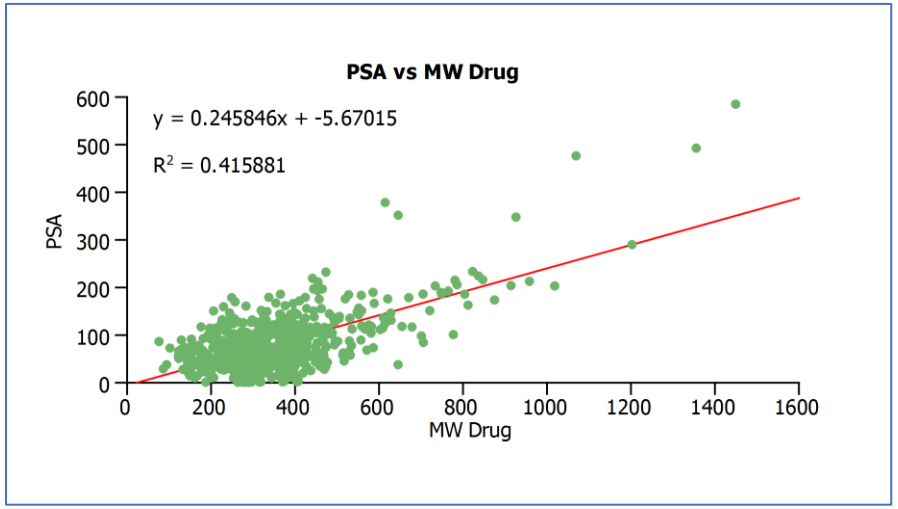


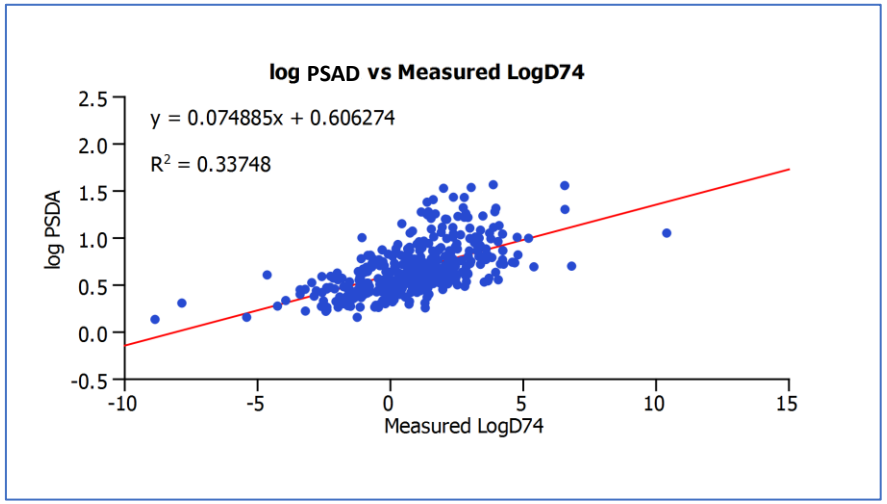
- Increased # aromatic (flat) rings ⇒ decreases solubility ⇒ associated w/ promiscuity: inc. ppb, cyp inhibition, hERG inhibition, *et al.*
- Heteroaromatics better than carboaromatics
- Increasing # of stereocenters ⇒ associated with positive clinical outcomes
- See 3D-Fragment Consortium (http://www.3dfrag.org/)

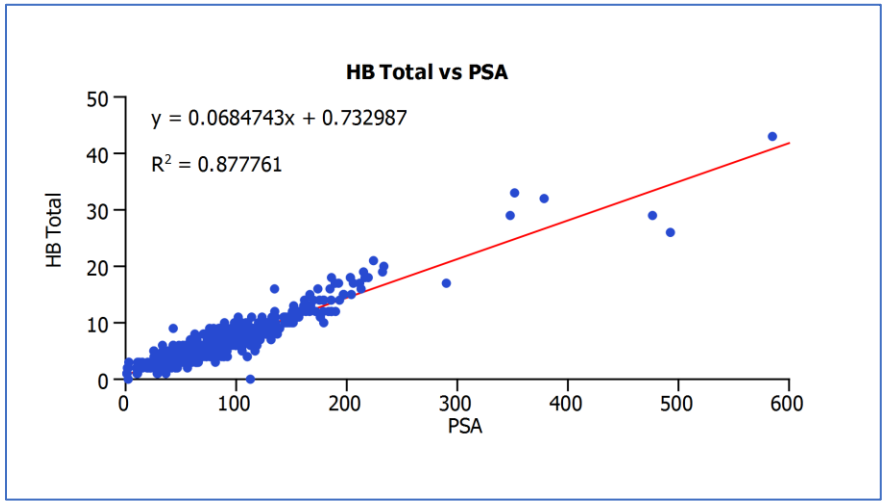
Lovering, F. et al *J. Med. Chem.*, 52(21), 2009, 6752; Ritchie, T. M. et al *Drug Disc. Today*, 16(3/4), 2011, 164; Maccari, L. et al *J. Med. Chem.*, 51(10), 2008, 2891.

Cross-correlations – 700 oral drugs



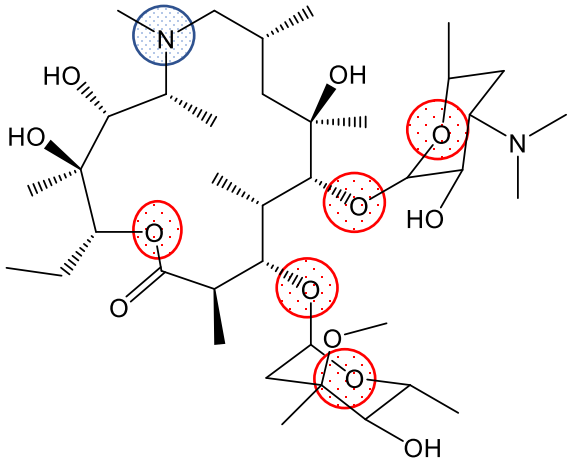


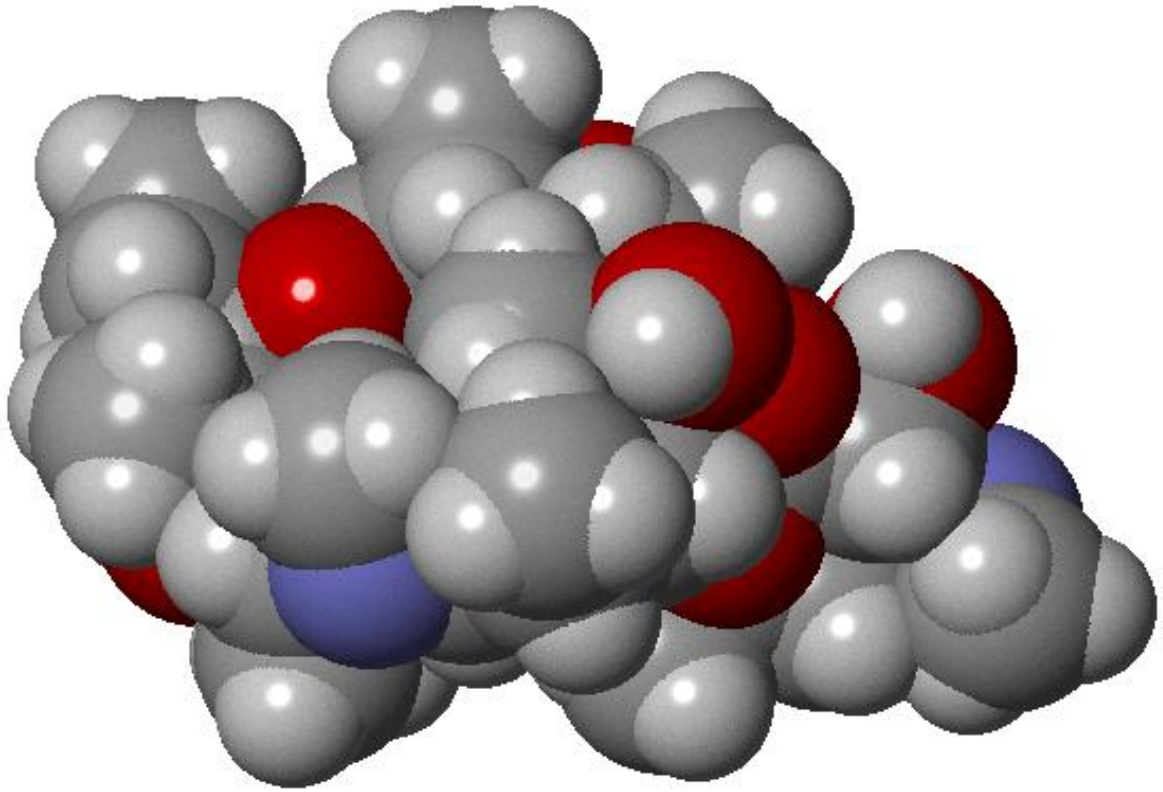




Data from Benet et al. *AAPS J.* 13*(4)*, 519, 2011.

Anti-Lipinski





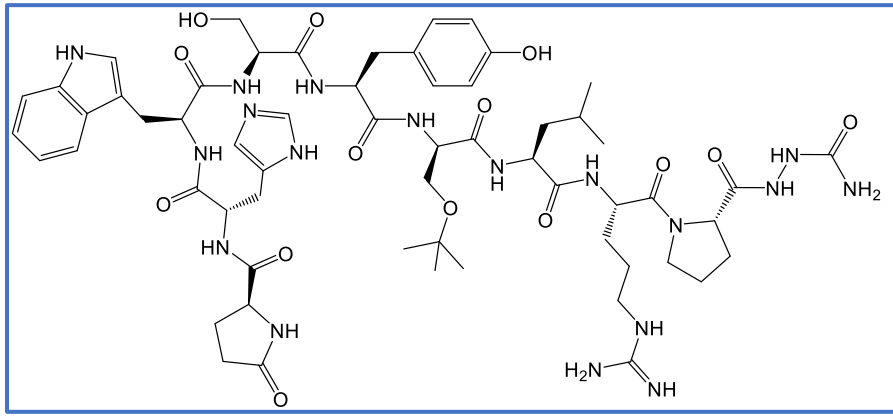
Azithromycin

F = 38%; 250 mg
MW = 749
PSA = 180 Å^2
cLogP = 2.9
cLogD = 0.47
pK_a = 8.74; 9.45
H-acceptors = 14
H-donors = 5-7
Rot. Bonds = 7

Exposed PSA = 131 Å^2
Solvent accessible H-bond acceptors = 8
Solvent accessible H-bond donors = 4
logP = 4.1
logD = 0.61

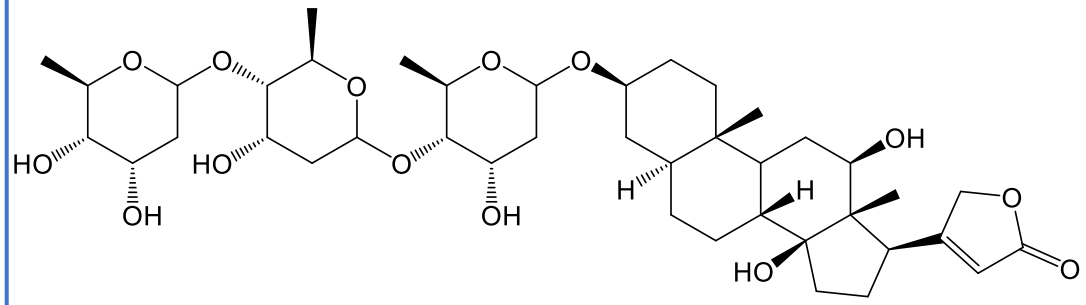
McFarland et al *J. Med. Chem.* 40(9),1997, 1340.

Anti-Lipinski



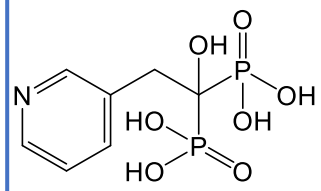
Goserelin (Zoladex) - GnRH agonist breast and prostrate cancers

18 H-bond donors
18 H-bond aceptors
MW = 1269
ClogP = -2.9
Rot. bonds = 42
PSA = 483 Å^2
injectible drug



Digoxin - atrial fibrillation

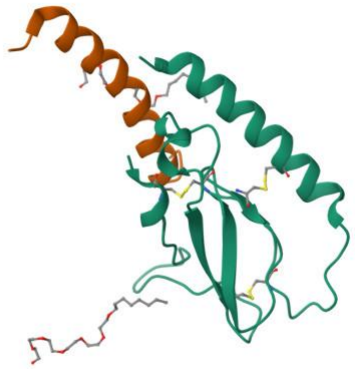
6 H-bond donors
13 H-bond aceptors
MW = 781
ClogP = 1.4
Rot. bonds = 7
PSA = 215 Å^2
F = 70%



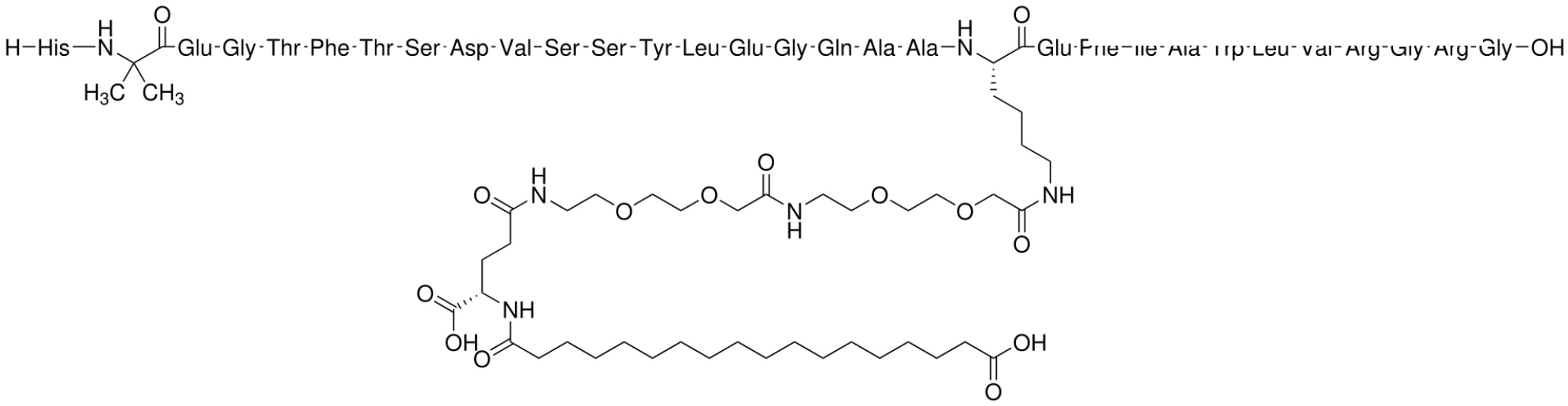
Risedronic acid - osteoporosis

5 H-bond donors
8 H-bond aceptors
MW = 283 PSA = 161 Å^2
ClogP = -2.62 F = 0.63%
Rot. bonds = 3 10 mg dose

Semaglutide (Ozempic®)



Anti-diabetic (type2)/Anti-obesity medication
Dose: 0.25 mg once a week (first 4 weeks) – IM
3.0 mg daily for 30 days (PO)
$T_{1/2}$ = 7 days

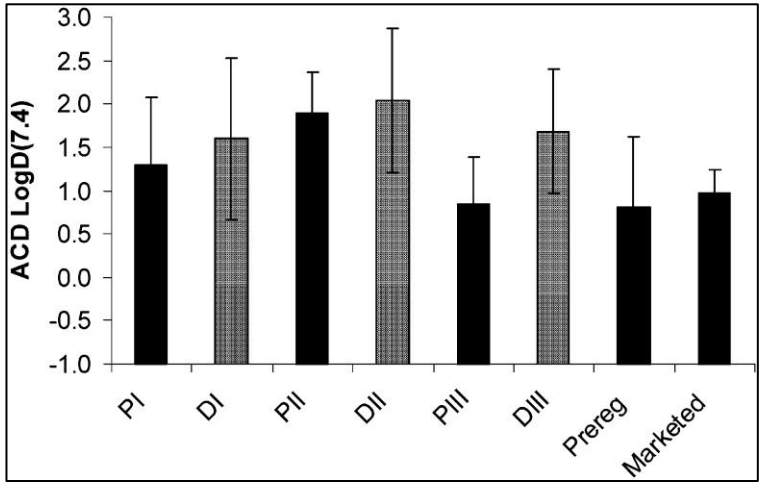


F = 0.8 – 1.4%
MW = 4114

Composite parameter – ligand efficiency

Ligand efficiency (LE):

pIC_{50} ÷ HAC (heavy atom count)

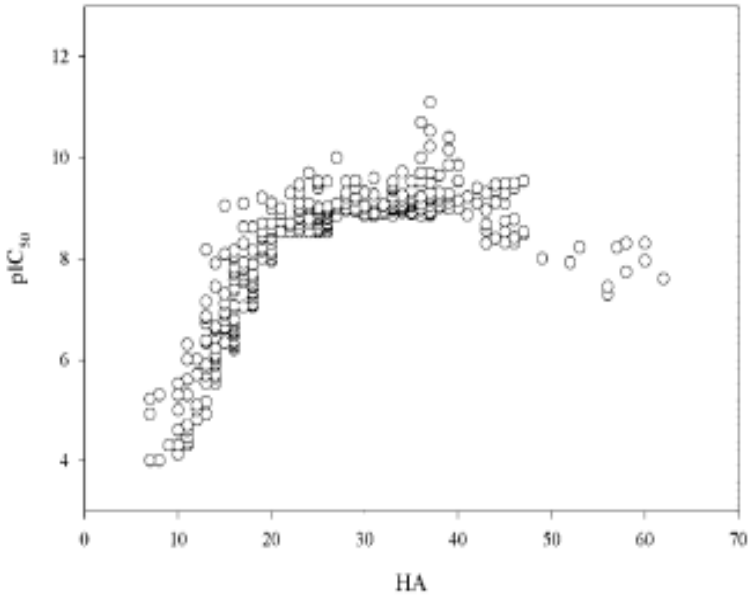


Others:

group efficiency (GE): **$\Delta pIC_{50} \div \Delta HAC$**

size independent ligand efficiency (SILE): **$pIC_{50} \div HAC^{0.3}$**

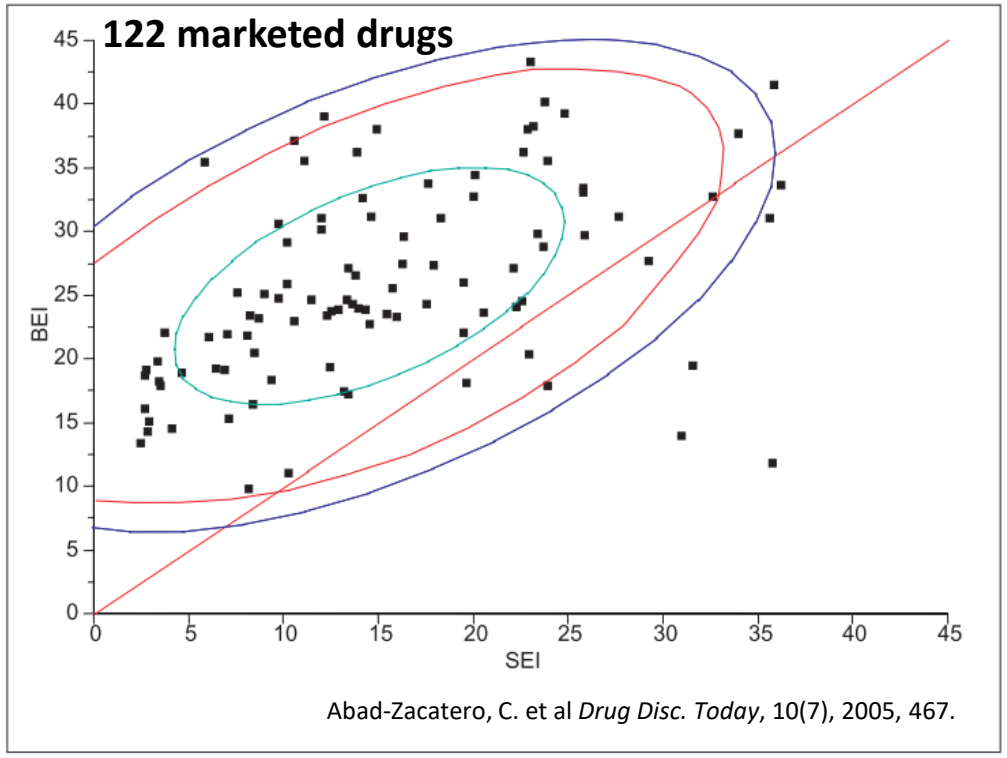
- Target potency increases with increasing size
- Tends to select for more lipophilic compound
- **Larger LE ⇒ better the hit or lead matter**



Wentlock et al *J. Med. Chem. 46(7)*6, 1250, 2003; Reynolds et al *J. Med. Chem.* 51(8), 2432, 2008.

De-convolute hits from a screening library

- percentage efficiency index (PEI): **[Cmpd@50% inh] ÷ MW**
  - alternative to LE (better accounting for atomic weight)
- binding efficiency index (BEI): **pIC_{50} ÷ MW**
  - similar to PEI w/ IC_{50} data
- surface binding efficiency index (SEI): **pIC_{50} ÷ TPSA**
  - normalize for polar atoms

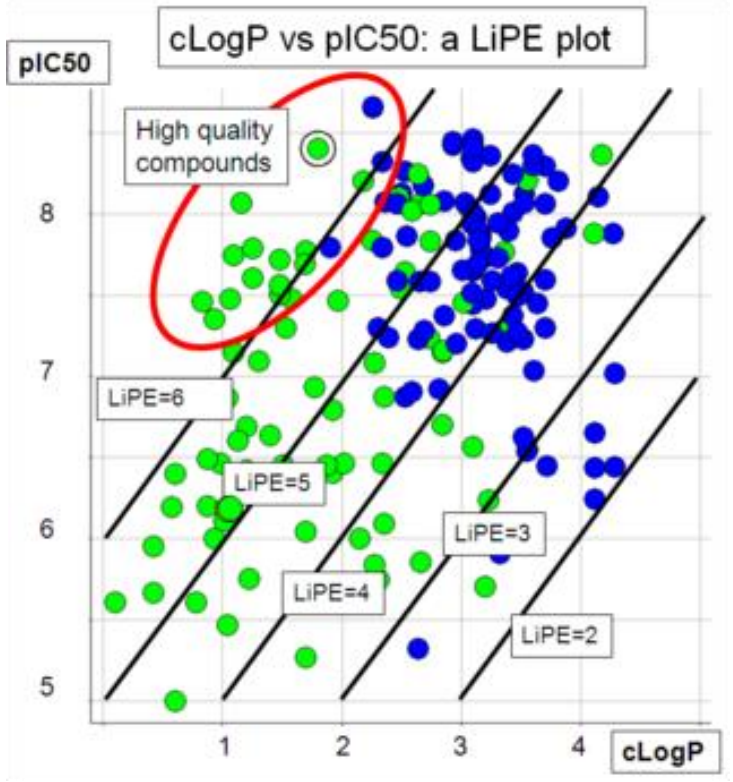


Abad-Zacatero, C. et al *Drug Disc. Today*, 10(7), 2005, 467.

Drugs dominated by non-polar atoms (MW ➢ PSA)

Lipophilicity Efficiency

Lipophilicity Efficiency (LiPE)
or Ligand Lipophilicity Efficiency (LLE)
(LiPE): **pIC_{50}– cLogP**



- Battle the proclivity to increase target potency by increasing lipophilicity
- Useful for lead optimization

Leeson et al Nature Rev. 2007, 6(11), 881-90; Abad-Zapatero Drug Disc. Today 2010, 15, 804-11

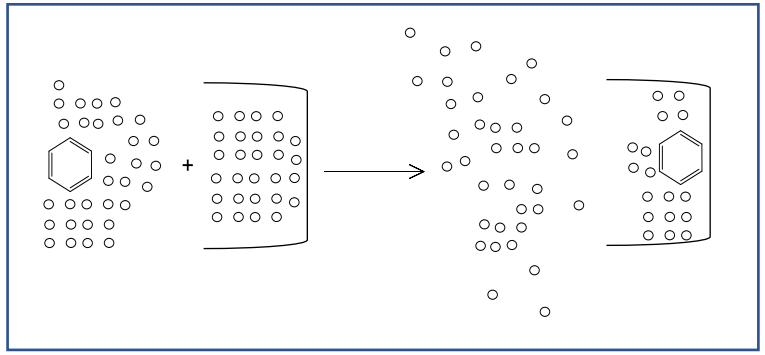
Enthalpy versus Entropy

H-bond interactions, electrostatic interactions, van der Waals interactions, etc.

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

Measurable via Isothermal Titration Calorimetry (ITC)

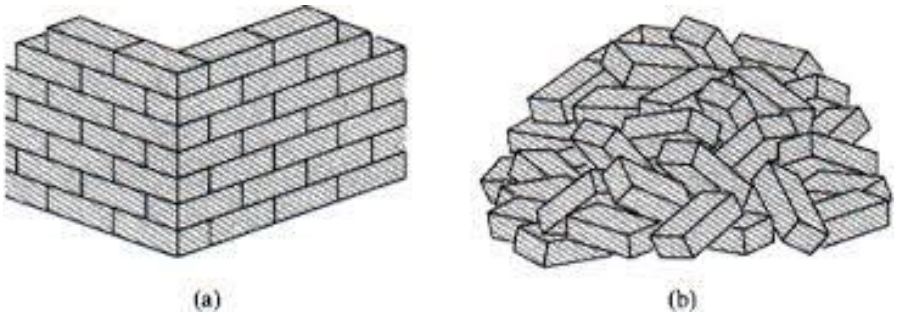
Primarily hydrophobic interactions – due to desolvation

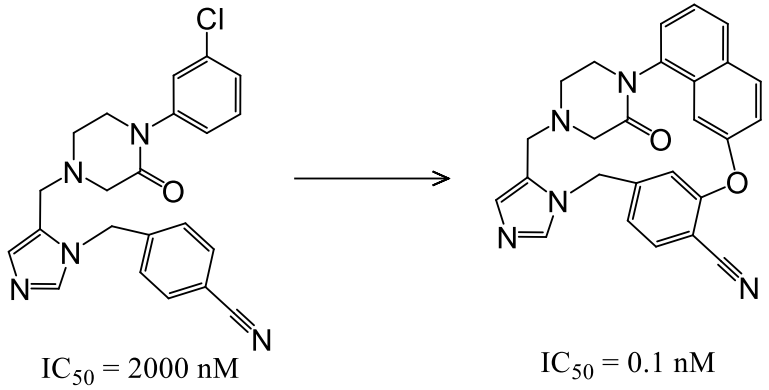


Entropy

- Tied-ups, tied-backs, tie-dyed, cyclized, constrained,

Which took more energy to create???





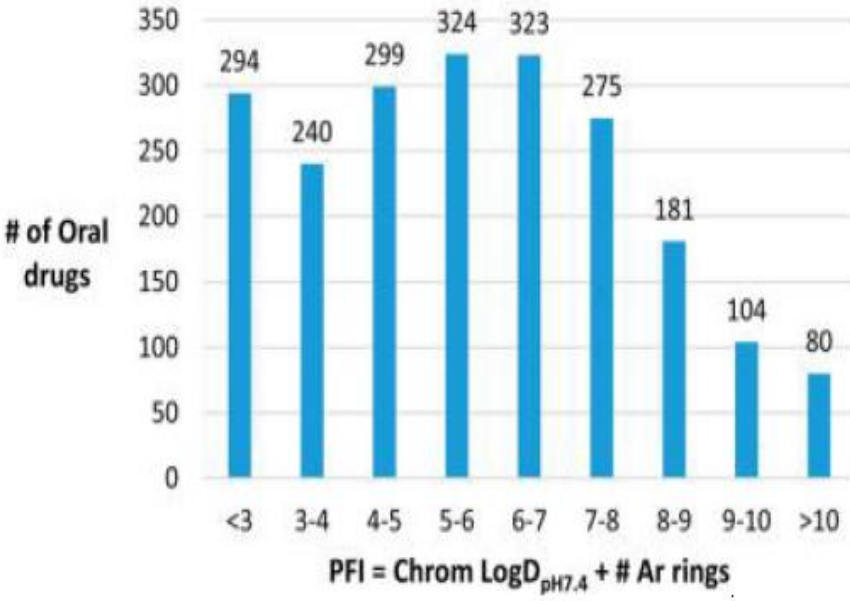
Conformationally locked farnesyltransferase inhibitors

- **Both ligands and targets sites become more ordered on association (binding) – costs associated**
- **Pre-organization into binding conformations can greatly improve potency – money in the bank**

Propert Forecast Index (PFI)

PFI = logD + #Aromatics

Assay / target value	PFI = Chrom $LogD_{pH7.4}$ + #Ar rings								
	<3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	>10
Solubility >200 µM	89	83	72	58	33	13	5	3	2
%HSA <95%	88	80	74	64	50	30	17	8	4
Cyp 2C9 pIC_{50} <5	97	90	83	68	48	32	23	22	38
Cyp 2C19 pIC_{50} <5	97	95	91	82	67	52	42	42	56
Cyp 3A4 pIC_{50} <5	92	83	80	75	67	60	58	61	66
Cl_{int} <3 ml/min/kg	79	76	68	61	54	42	41	39	52
P_{app} >200 nm/s	20	30	46	65	74	77	65	50	33
	iPFI = Chrom LogP + #Ar rings								
hERG pIC_{50} <5 (+1 charge)	86	93	88	70	54	36	29	21	11
Promiscuity <5 hits with pIC_{50} >5	85	78	74	65	49	30	20	13	7



Young, R. J. *Nat. Rev. Drug Disc.* 16(17-18), 822, 2011

- Optimal for permeability ⇒ intermediate value
- Optimal for potency ⇒ intermediate value
- Optimal for all else ⇒ go small and go polar

Leeson, P.; Young, R. J. *ACS Med. Chem. Lett.* 6, 722, 2015

Other composite parameters

- ligand efficiency dependent lipophilicity (LELP): c**logP ÷ LE**
  - accounts for price of LE paid in clogP
- polar surface area density (PSAD): **MW ÷ PSA**
  - normalize for PSA going up as molecule size increases
- dose number (Do): **Max. Dose ÷ (250 x cLogP x Solubility)**
  - Lower solubility can follow low dose drug
- ligand efficiency scale (LE_Scale):

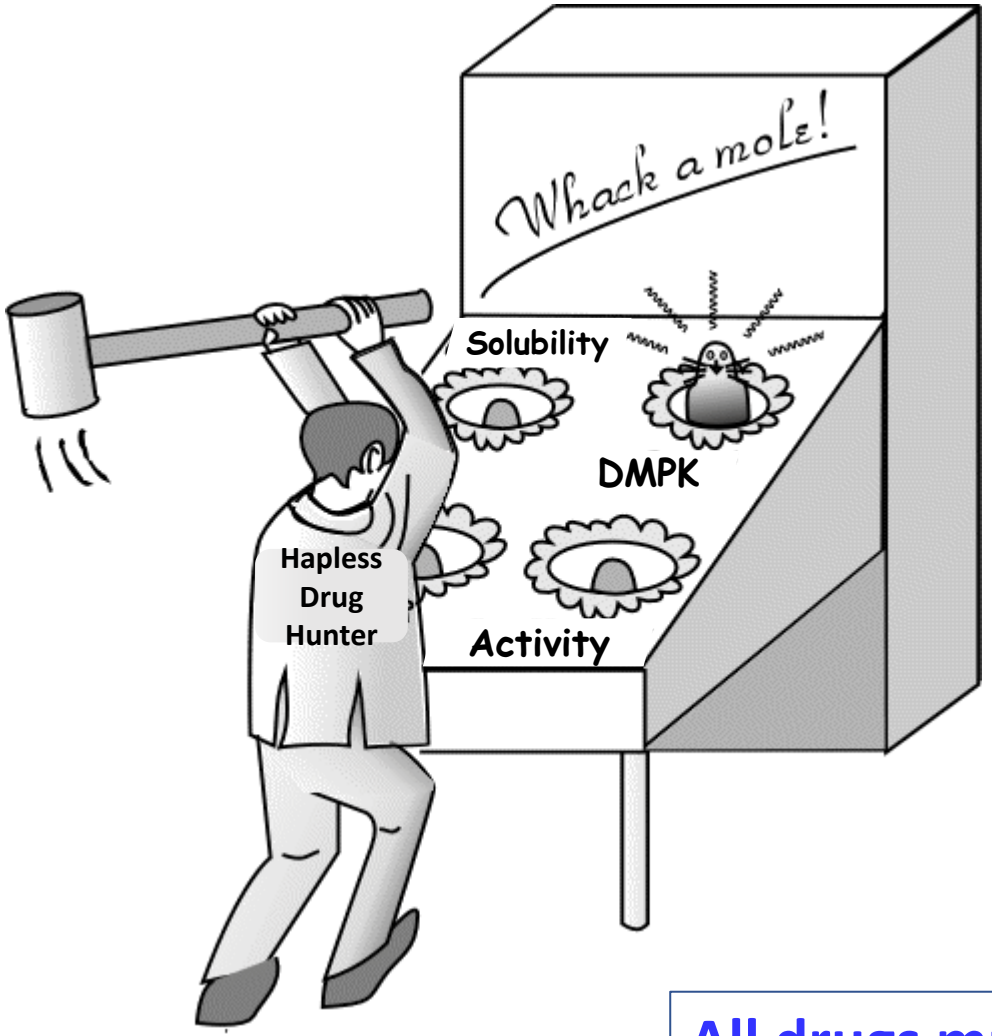
  $0.072+7.5/(HA)+25.7/(HA^2)-361.5/(HA^3)$

  - normalize for small molecules versus larger molecules
- ligand lipophilicity index (LLE_{AT}):

  $LLE_{AT} = 0.11 – \ln(10)\cdot RT(\log P - pIC_{50}) \div HAC$

  - subtract out lipophilicity component for ΔG of binding

Too many (often conflicting) optimizations



Simultaneous optimization (or mitigation):

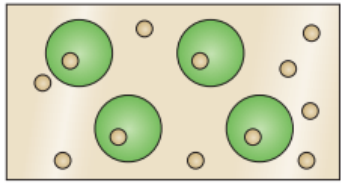
- Target activity
- Cell Permeability
- Solubility
- Clearance (biliary, metabolic, renal, etc.)
- Reactive metabolites
- Distribution
- Plasma protein binding
- Plasma stability
- Absorption/bioavailability (oral)
- Off-target activity
- Ion channel binding
- Genotoxicity
- Hepatotoxicity
- Mitochondrial toxicity
- Drug-drug interaction (Cyp inhibition, transporter inhibition)
- Cost-of-goods/synthetic feasibility
- etc.

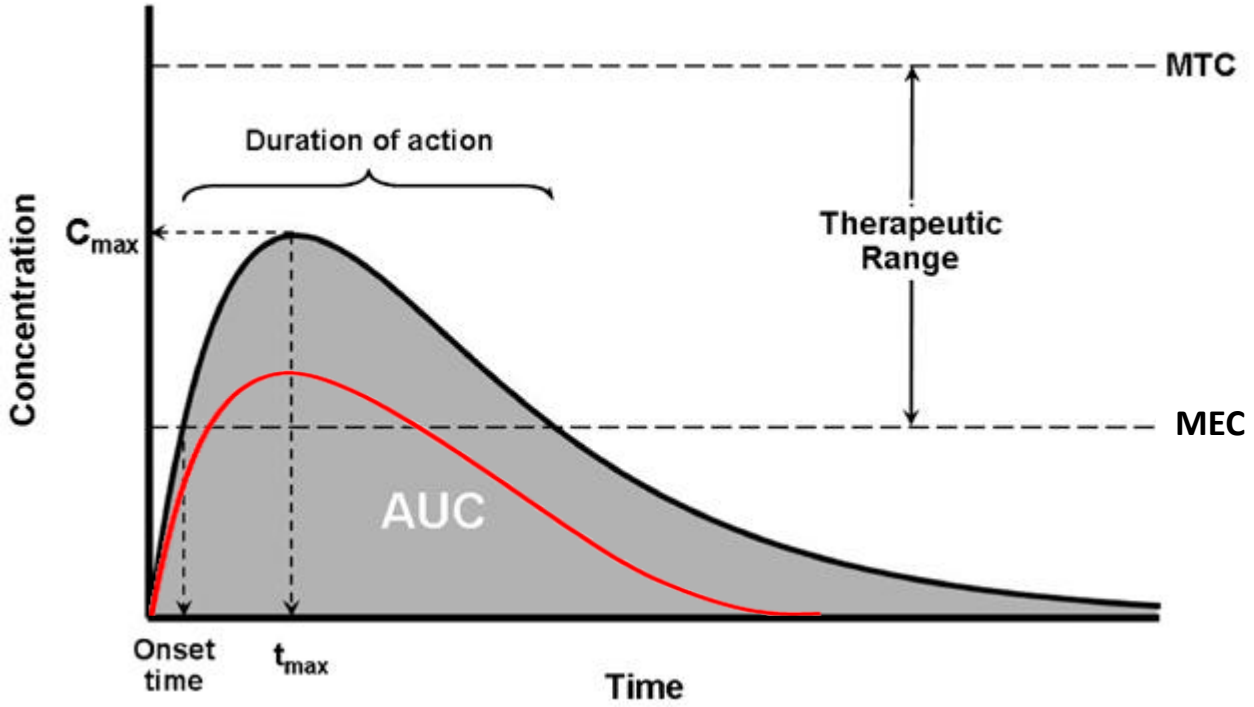
All drugs must make some compromises

Multi-Parameter Optimization (MPO)

$$AUC = \frac{Dose*F}{Cl}$$

$$Cl = \frac{0.693*V_d}{t_{1/2}}$$





$$\text{Ideal Cmpd} \propto \frac{fAUC}{MEC}$$

MPO

	Target pot. (IC_{50})	*Pf*NF54 IC_{50}	*PfK1* IC_{50}	Solu-bility	µ-some Cl_{int} (Mo)	µ-some Cl_{int} (Hu)	hERG	HepG2	Caco-2 (A⇒B)	Caco-2 (efflux)	TPSA	LogD
Hi-Low	Low	Low	Low	High	Low	Low	High	High	High	Low	Low	~2
Weight	2	10	8	10	7	4	3	6	5	5	3	3
A												
B												
C												
D												
E												
F												
G												
H												
I												
J												

MPO

	Target pot. (IC_{50})	*Pf*NF54 IC_{50}	*PfK1* IC_{50}	Solu-bility	µ-some Cl_{int} (Mo)	µ-some Cl_{int} (Hu)	hERG	HepG2	Caco-2 (A⇒B)	Caco-2 (efflux)	TPSA	LogD
Hi-Low	Low	Low	Low	High	Low	Low	High	High	High	Low	Low	~2
Weight	2	10	8	10	7	4	3	6	5	5	3	3
L												
W												
M												
Y												
A												
B												
F												
X												
N												
D												

MPO Excercise

Rank order these 4 compounds from best to worse. The aim is to put into an in vivo mouse efficacy model towards selecting a development candidate

	Target (IC_{50}, nM)	*Pf*NF54 (IC_{50}, nM)	FaSSIF Solu-bility (μM)	Plasma Cl_{int} (Mo, ml/min/kg)	Bioavail ability (Mo, %)	Plasma Protein Binding (Mo, %)	hERG (IC_{50}, μM)	HepG2 (IC_{50}, nM)
Hi-Low	Low	Low	High	Low	High	Low	High	High
Weight	???	???	???	???	???	???	???	???
A	24	32	36	2.6	76	93	100	270
B	8	25	36	3.2	62	86	10	120
C	23	5	3750	13	66	90	10	13
D	1	5	255	19	66	82	30	100

Grand Challenges
African Drug Discovery Accelerator

Parting thought: Rules only make sense if they are broken. Breaking the rule is one way of observing it.

Sir Thomas More

English lawyer, judge, social philosopher, author, statesman, amateur theologian, and noted Renaissance humanist

Supported by

Gates Foundation

LifeArc

H3D FOUNDATION