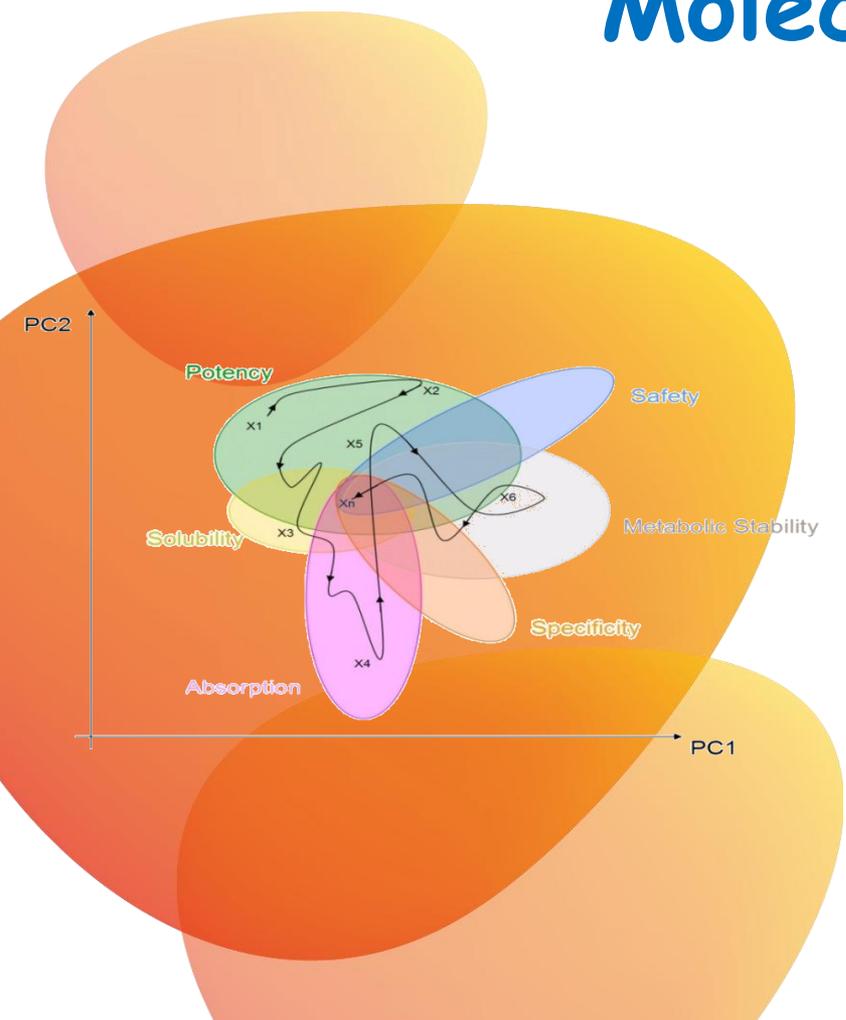


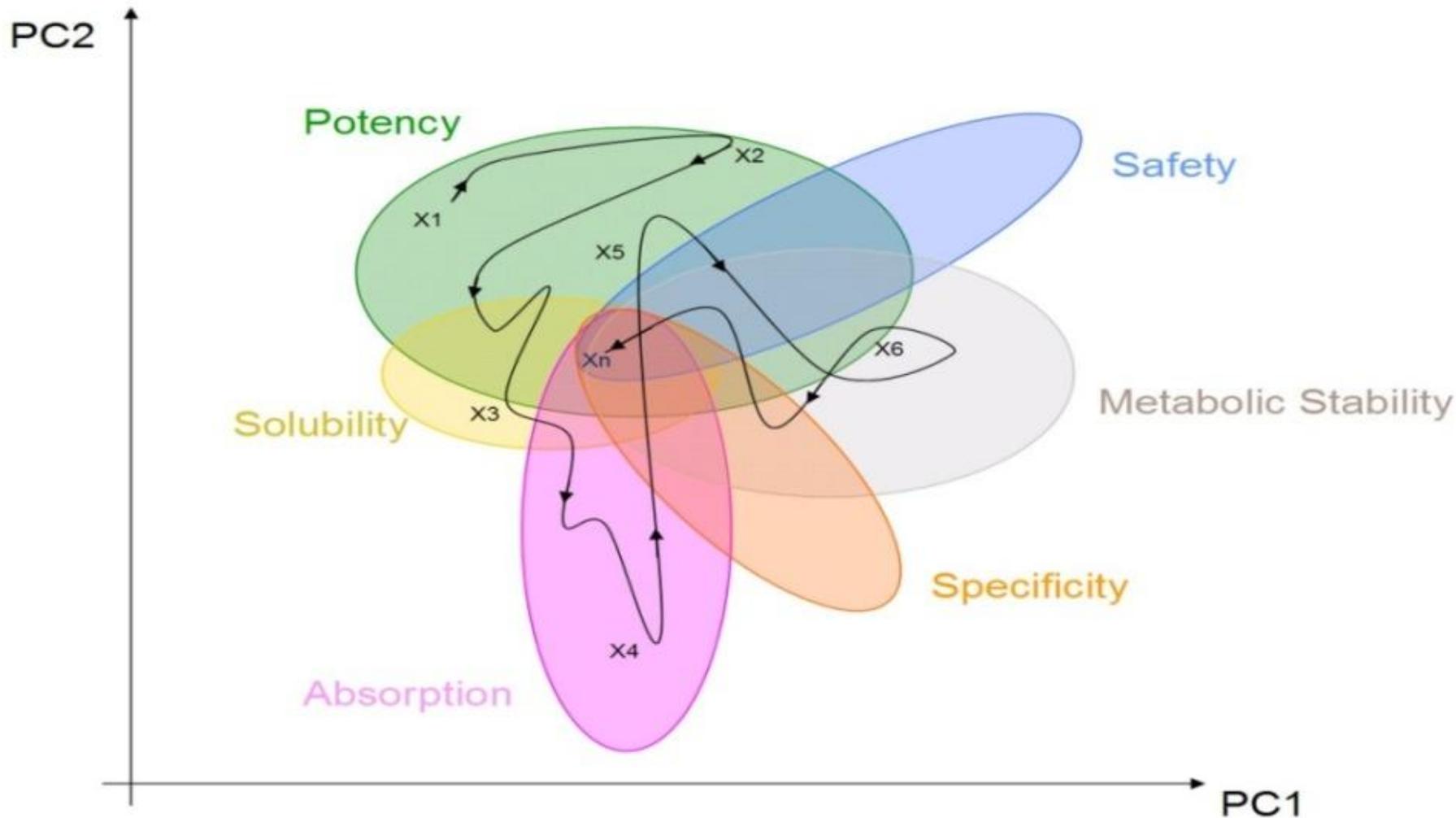
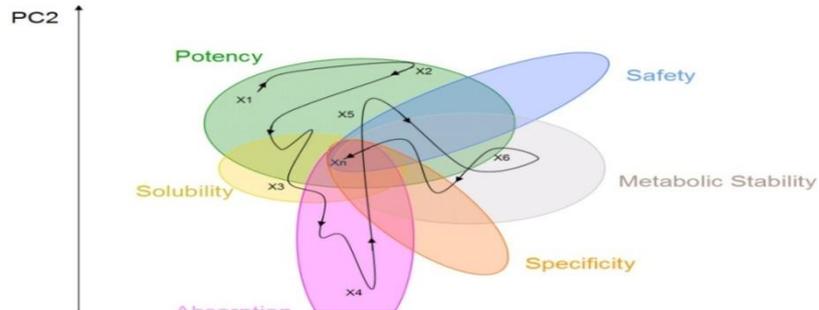
Molecular Obesity, Potency and other Addictions in Drug Discovery



Mike Hann
Bio-Molecular Structure
MDR Chemical Sciences
GSK Medicines Research Centre
Stevenage, UK

Mike.M.Hann@gsk.com

The challenge of drug discovery



Learning from our mistakes - the resurgence of reason based on metadata from big Pharma

Emergence of rules of thumb as guidance

– **Permeability/solubility:** Pfizer analysis of existing drugs and oral absorption profile

– (Lipinski, *Adv. Drug. Del. Revs.* 1997, 23, 3)

– Mol Wt <500, LogP <5, OH + NH count <5, O + N count <10: 90% of oral drugs do not fail more than one of these rules.

– Lipinski Rule of 5

– **Receptor Promiscuity:** AZ analysis of 2133 compounds in >200 Cerep Bioprint® assays

– (Leeson & Springthorpe, *Nat. Rev. Drug Disc.* 2007, 6, 881)

– cLogP < 3 decreases risk; > 4 increases risk; bases/quats >> neutrals > acids

– *Lipophilic Ligand Efficiency* LLE = pIC₅₀ – cLogP >5 for toxicity risk reduction

– AZ LLE >5 rule.

– **Receptor Promiscuity:** Roche analysis of 213 compounds profiled at Cerep

– (Peters et al, *ChemMedChem* 2009, 4, 680-686)

– Pronounced promiscuity not observed below a threshold cLogP of 2. Increased promiscuity with increased calculated basicity.

– **Toxicity:** Pfizer *in vivo* tolerability data on 245 compounds

– (Hughes et al, *Bioorg. Med. Chem. Letts.* 2008, 18, 4872)

– cLogP < 3 & TPSA > 75 give 6-fold reduced *in vivo* toxicity vs. >3 & <75; 24-fold for bases

– Pfizer 3/75 rule

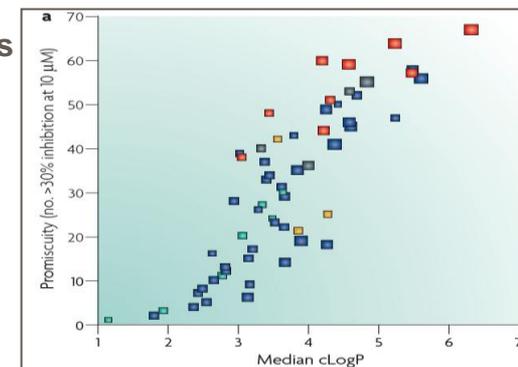
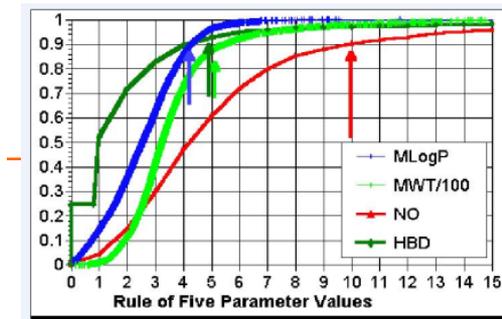
– **ADMET:** GSK analysis of ~30,000 GSK compounds yielded simple rules

of thumb for the effect of physchem parameters on solubility, permeability, bioavailability, volume of distribution, clearance, hERG inhibition, PGP efflux & P450 inhibition

– (Gleeson, *J. Med. Chem.* 2008, 51, 817.)

– Mol Wt <400 & cLogP <4 reduces ADMET risks compared to >400 & >4

– GSK 4/400 rule



Observed odds for toxicity versus ClogP/TPSA

Toxicity	Total-drug		Free-drug	
	TPSA > 75	TPSA < 75	TPSA > 75	TPSA < 75
ClogP < 3	0.39 (57)	1.08 (27)	0.38 (44)	0.5 (27)
ClogP > 3	0.41 (38)	2.4 (85)	0.81 (29)	2.59 (61)

neutral molecules	MWT < 400 and clogP < 4	MWT > 400 and/or clogP > 4
solubility	average	lower
permeability*	higher	average/higher
bioavailability	average	lower
volume of Dist.**	average	average
plasma protein binding	average	higher
CNS penetration***	higher/average	average/lower
brain tissue binding	lower	higher
P-gp efflux	average	higher/average
in-vivo clearance	average	average
hERG Inhibition	lower	lower
P450 inhibition****	lower 2C9, 2C19, 2D6 & 3A4 inhibition	higher 2C9, 2C19 & 3A4 inhibition
P450 inhibition****	higher 1A2 inhibition	lower 1A2 inhibition
P450 inhibition****		average 2D6 inhibition

Some other things we have learnt



• Size & permeability:

The larger a “small” molecule is, the more lipophilicity it is likely to need to permeate membranes.

Permeability rules defining AZlogD limits required to achieve >50% chance of high permeability for a given molecular weight band

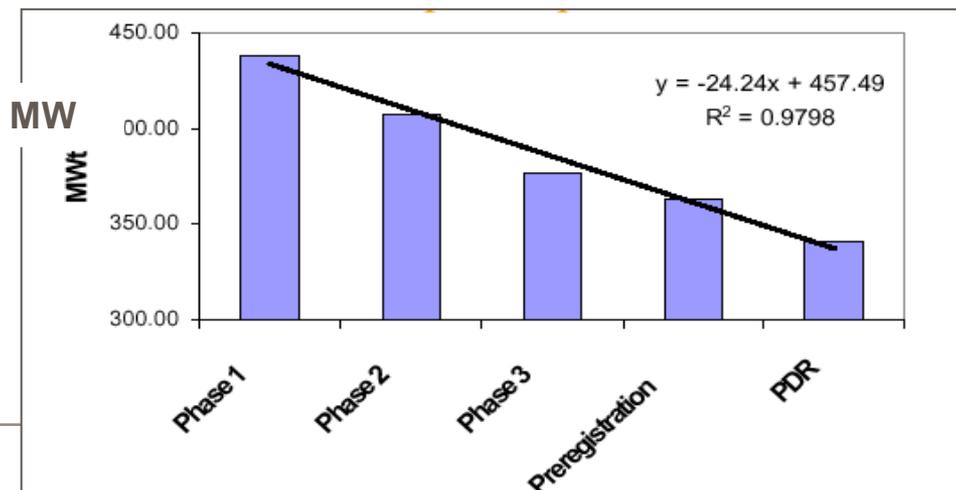
Molecular weight	AZlogD
<300	>0.5
300-350	>1.1
350-400	>1.7
400-450	>3.1
450-500	>3.4
>500	>4.5

- Defining optimum lipophilicity and MW ranges for drug candidates – MW dependent logD limits based on permeability. Waring, *Bioorg. Med. Chem. Lett.*, 2009, 19, 2844
- Lipophilicity in drug discovery. Waring. *Expert Opin Drug Discov.* (2010) 5(3) 235

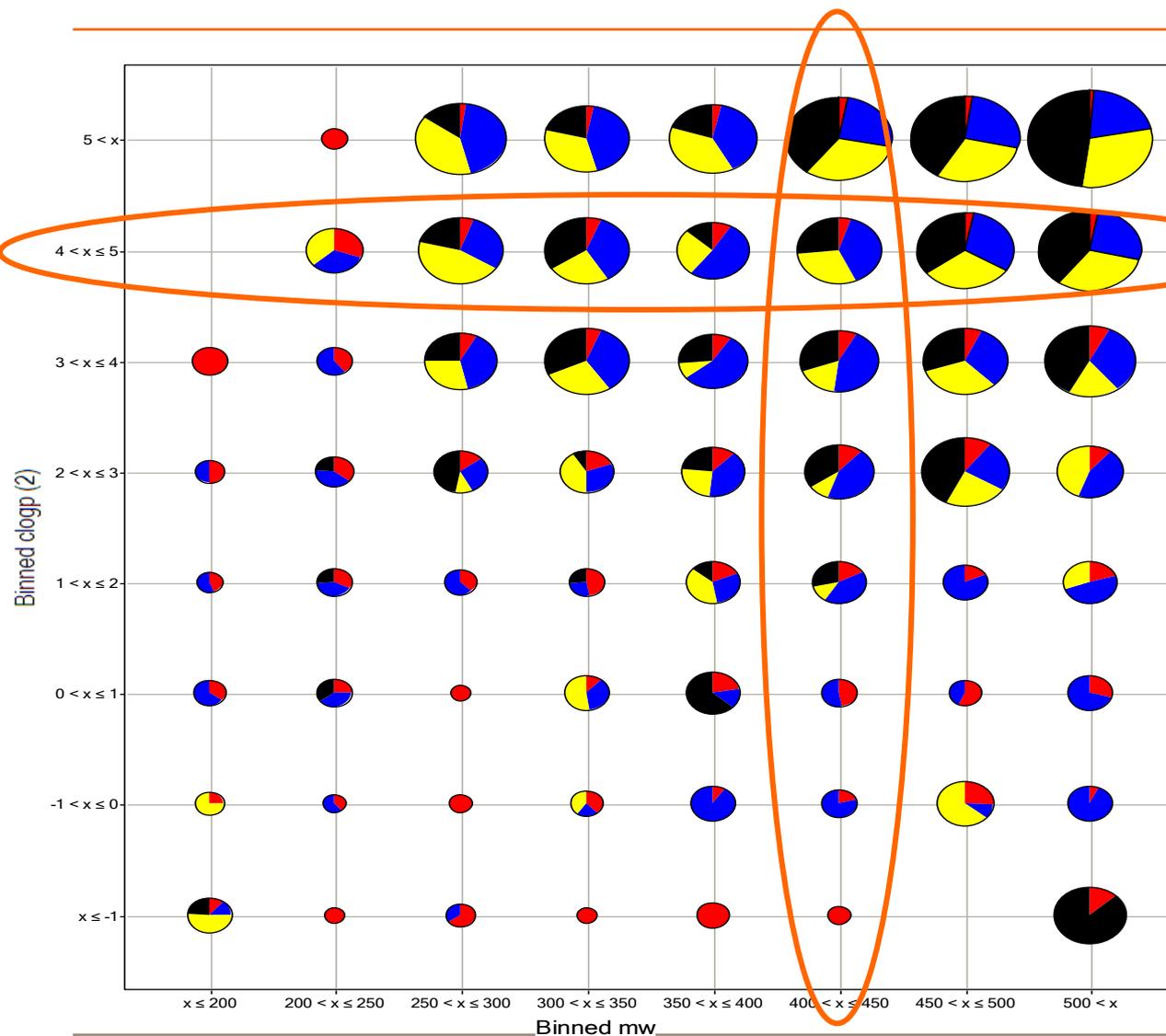
• Pre-clinical & clinical survival:

Larger & more lipophilic molecules have reduced chances of survival in pre-clinical & clinical phases

- Wenlock et al. *J. Med Chem.*, 2003, 46, 1250. A comparison of Physicochemical Property Profiles of Marketed Oral Drugs



Is MW or logP the source of promiscuity?



Graph showing series of pie charts in different cLogP and MW bins for a set of approximately 2500 compounds tested in more than 490 assays. The size of each pie chart represents the average number of hits for compounds in that pie, where a hit is defined as a pXC50 value of 5 or higher. The colours indicate the proportion of compounds within each pie having particular numbers of hits (red: <5; blue: 5-15; yellow: 15-25; black: >25), where a hit is defined as activity greater than 10mM in any of the ~490 assays

Property Forecast Index PFI - a useful overall guide to where to look for developable compounds



Drug Discovery Today • Volume 16, Numbers 17/18 • September 2011

REVIEWS

$$\text{PFI} = \text{mChromLogD}_{7.4} + \#\text{Aromatic rings}$$

TABLE 2

Percentages of compounds achieving defined target values in the various developability assays categorised by PFI or iPFI bins^a

Assay / target value	PFI = mChrom log D _{pH7.4} + #Ar								
	<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
2C9 pIC ₅₀ <5	97	90	83	68	48	32	23	22	38
2C19 pIC ₅₀ <5	97	95	91	82	67	52	42	42	56
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
Papp >200 nm/s	20	30	46	65	74	77	65	50	33
	iPFI = mChrom log P + #Ar								
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

Sweet spot for permeability is in conflict with other desirable properties!

^a Colouring refers to the % chance of achieving benchmark value in that PFI bin: green, ≥67%; yellow, 34-67%; and red, <33%.

POTENCY



What we have come to know (or rediscover!)



- Large and particularly lipophilic molecules are increasingly seen as bad - again!
- Cell penetration of larger molecules needs increasing lipophilicity
- Lipinski's 500/5 for oral bioavailability is increasingly seen as far too lenient when it comes to the wider ADMET issues.
 - We should be thinking 400/4 or PFI <6 as better indicators of the space with highest probability of successfully developing a drug
 - **and even smaller for leads as starting points!**

Average property values for the Sneader lead set, average change on going to Sneader drug set and percentage change

Av # arom	Δ arom	%	Av ClogP	Δ ClogP	%	Av CMR	Δ CMR	%
1.3	0.2**	15	1.9	0.5**	26	7.6	1.0**	14.5

Av # HBA	Δ HBA	%	Av # HBD	Δ HBD	%	Av # heavy	Δ heavy	%
2.2	.3**	14	0.85	-.05 ⁺	-4	19	3.0**	16

Av MW	Δ MW	%	Av MV	Δ MV	%	Av # Rot B	Δ Rot B	%
272	42.0**	15	289	38.0**	13	3.5	.9**	23

<<<<<<< A lesson from history

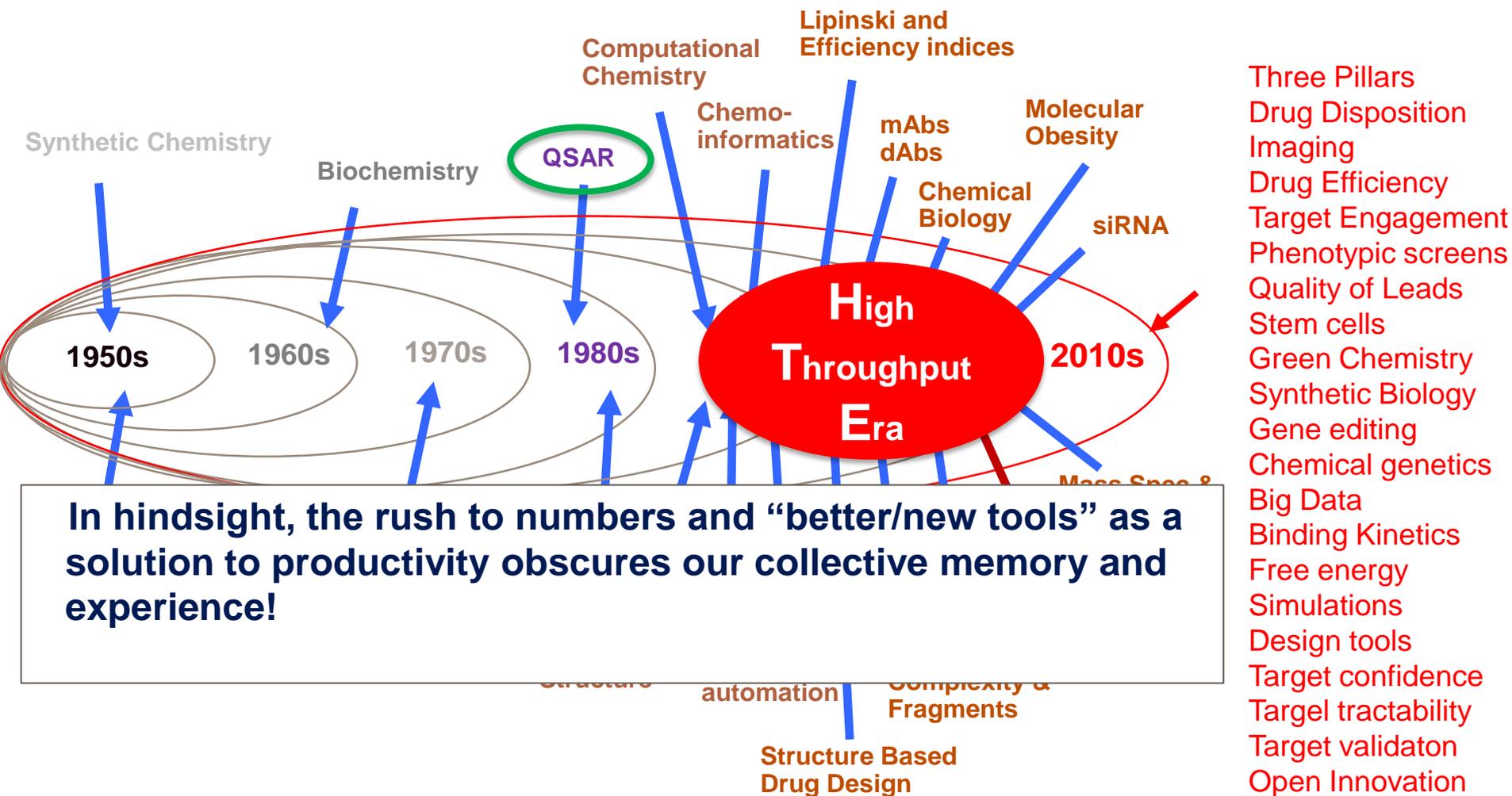
– **Where did it all go wrong?**

Data Sneader, W. Drug Prototypes and their Exploitation; John Wiley and Sons Ltd.: 1996.

Where have we come from and why has it got like this?



And what are the current attempts to try and improve?



The expanding “sciences” of Medicinal Chemistry and drug discovery

The curse of Molecular Obesity

- The tendency for drug discovery molecules to become too large and too lipophilic for their own good during lead optimisation through the quest for potency and specificity.
 - It presents a high risk to the future “health” of the compound as a drug candidate.
- As with medical obesity, which is measured by Body Mass Index BMI, we now make use of indices such as Ligand Efficiency Index LE and Lipophilic Ligand Efficiency Index LLE to help identify and control the problem.

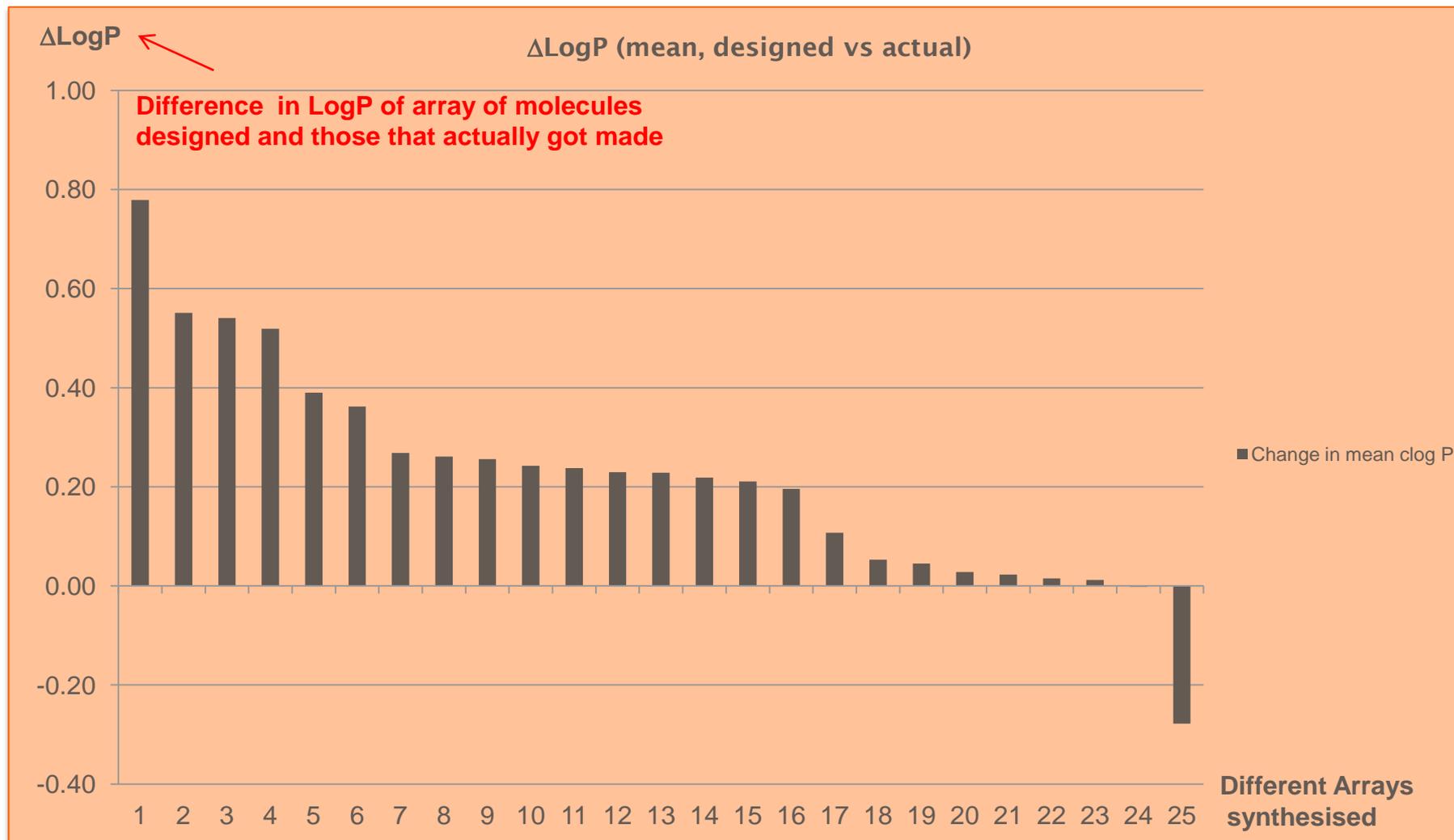
Indices as guideposts for life and drug discovery

- **Body Mass Index (BMI) = human weight / height²**
 - **Ligand Efficiency Index (LE)**
 - Potency in kcal/mol ($=-1.37\log K_d$) normalised by the number of heavy atoms
 - An *'idealised'* compound with 1nm pIC50 and 30 heavy atoms has LEI = 0.42
 - An *'okay'* compound with 10nm pIC50 and 38 heavy atoms (MW 500) has LEI = ca. .3
 - **Ligand Lipophilicity Efficiency Index (LLE)**
 - Potency normalised by lipophilicity
 - LLE = pIC50 – clogP (typical good value are 5-7 for nanomolar potency)
 - During optimisation potency should increase more than just that due to bulk logP effects. Particularly true with membrane bound targets.
 - **$LLE_{Astex} = 0.11 * \ln(10) * RT(\log P - \log(K_d \text{ or } pK_i \text{ or } IC_{50}) / HA$**
 - Lipophilic efficiency assessment for fragments
 - Scale fixed to be similar to LE so .3 is a base level number to aim for.
 - **Binding Efficiency Index (BEI)**
 - Potency (pIC50) normalised for MW
 - An *'idealised'* compound with 1nm pIC50 and MW of 0.333 kDA has BEI = 27
 - **Surface Binding Efficiency Index (SEI)**
 - Potency normalised for Polar Surface Area
 - An *'idealised'* compound with 1nm pIC50 and PSA of 50A² has SEI = 18
-

The link of potency and molecular obesity

- Potency can improve the **therapeutic index**, **specificity** and help **reduce dosage**
 - all good things but if we grow potency in the wrong way molecules can get very obese
 - We can easily measure and optimise against **potency**
 - Potency results come back quickly and we react to them with decisions as to what to make next
 - It satisfies the “we are making progress” paradigm!
 - Unreasonable time pressures can make this seem like an end its own right!
 - Potency tends to correlate with **increasing MW and logP** for most series because we make more interactions.
 - Size needs lipophilicity to pass through membranes
 - Adding MW is easier than subtracting in synthetic chemistry!!
 - Most medicinal chemists are **synthetic** organic chemists!
-

Organic synthesis and purification favours lipophilic molecules



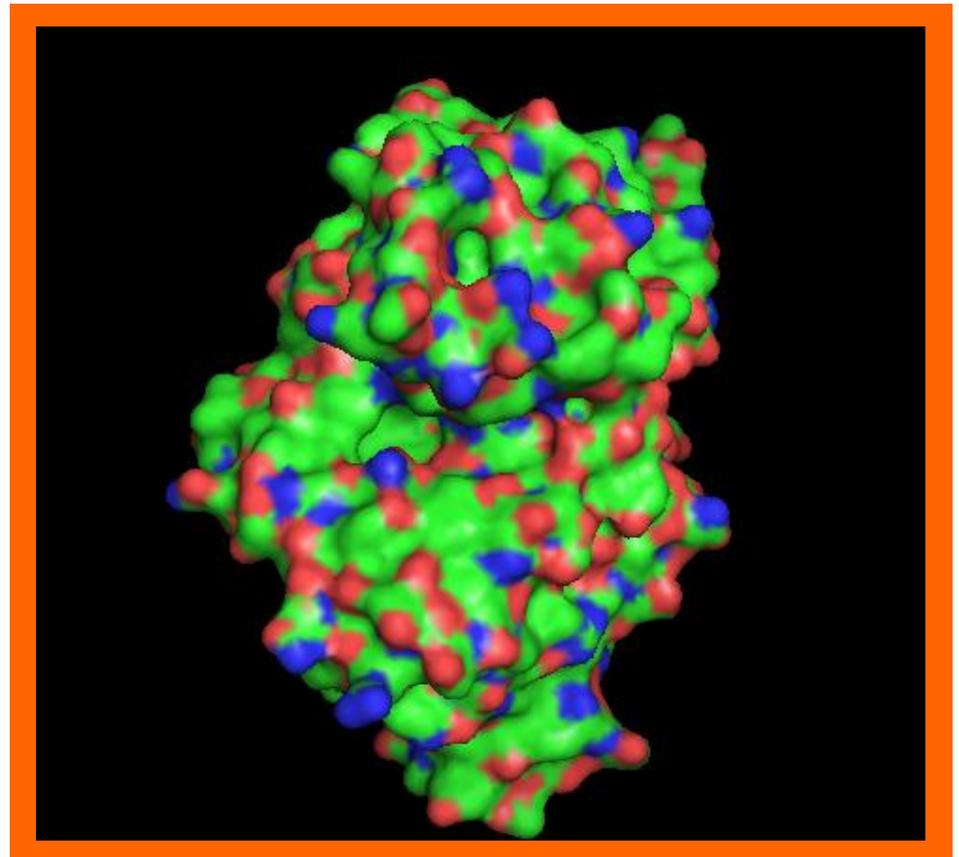
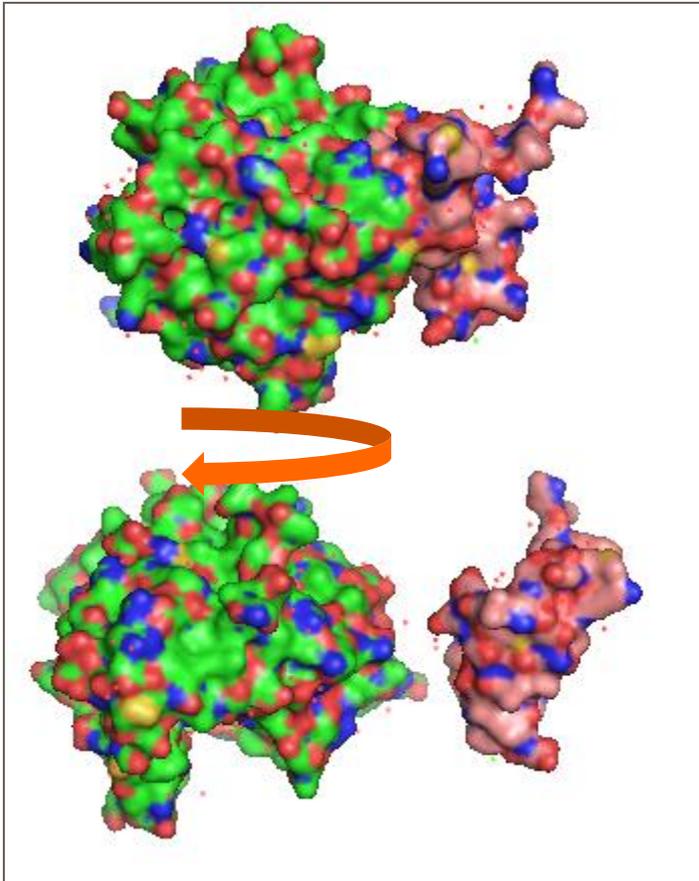
The link of potency and molecular obesity - cont'd?



- We often start with **isolated protein** in a biochemical assay with none of the environment of more phenotypic assays to help balance the physicochemical properties.
- We look for early signs of **cellular potency** in our screening cascades - this needs both some intrinsic potency and **cellular penetration**
 - Both of these are very easily driven by increasing logP.
 - Once we get cellular activity the damage may already be done if we do not revisit to look at how we got there.

The link of potency and molecular obesity - cont'd?

- **Structure based design** using crystal structures is a fantastic tool but it can easily draw you into the specifics of building potency rather than looking at the wider challenges at the same time.



The origins of potency – enthalpy and entropy

$$\Delta G = \Delta H - T \Delta S = -RT \ln K_d$$

- **Measurements of Free Energy show that for synthetic ligands, potency correlates with buried apolar surface area (ie size of interface and it's lipophilicity)**
- Buried **apolar** surface area (lipophilicity) is an easier way to get potency than through buried **polar** surface area
- We need to be very careful that we are not drawn down the path of using too much lipophilicity as a quick fix for potency!

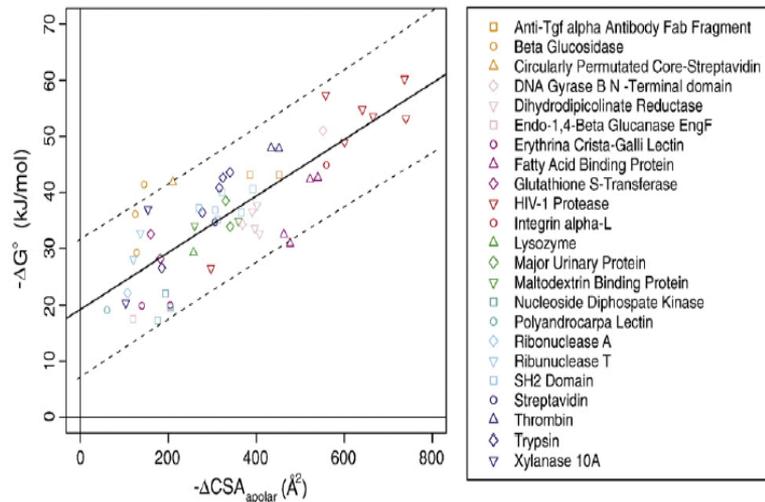


Fig. 2. The Gibbs free energy of binding for protein–ligand interactions correlates well with reduction in hydrated apolar surface area upon complex formation ($R^2=0.65$). The key indicates the proteins involved in each interaction. A linear least-squares fit to the data gives an intercept of $19.4 \pm 1.8 \text{ kJ mol}^{-1}$ and a slope of $0.049 \pm 0.005 \text{ kJ mol}^{-1} \text{\AA}^{-2}$. The thin dotted lines represent the 95% confidence intervals of the fit.

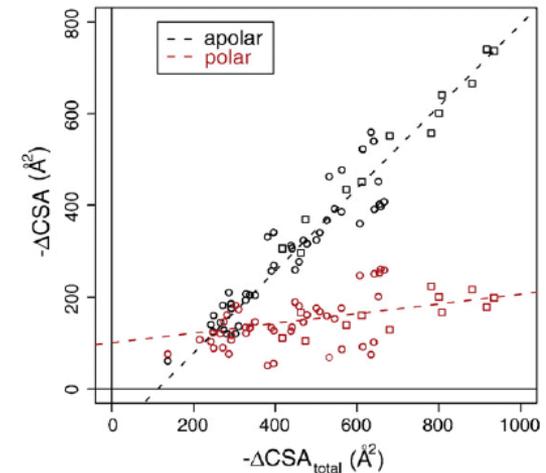


Fig. 3. (a) The apolar ($\Delta \text{CSA}_{\text{apolar}}$) and polar ($\Delta \text{CSA}_{\text{polar}}$) contributions to the total reduction in solvent ASA ($\Delta \text{CSA}_{\text{total}}$) upon complex formation diverge substantially with increasing extent of the binding interface. The interactions between proteins and synthetic ligands are represented by squares, whereas biological and other interactions are represented by circles. The intercept, slope and R^2 of $-\Delta \text{CSA}_{\text{apolar}}$ are -78.0 ± 15.3 , 0.86 ± 0.03 and 0.93, whereas those of $-\Delta \text{CSA}_{\text{polar}}$ are 78.0 ± 15.3 , 0.14 ± 0.03 and 0.29, respectively. (b) Ligands themselves show similar but less marked trends in $\text{CSA}_{\text{apolar}}$ and $\text{CSA}_{\text{polar}}$.

The Thermodynamics of Protein-Ligand Interaction and Solvation: Insights for Ligand Design, Olsson, Williams, Pitt & Ladbury, *J Mol Biol* (2008) 384, 1002-1017

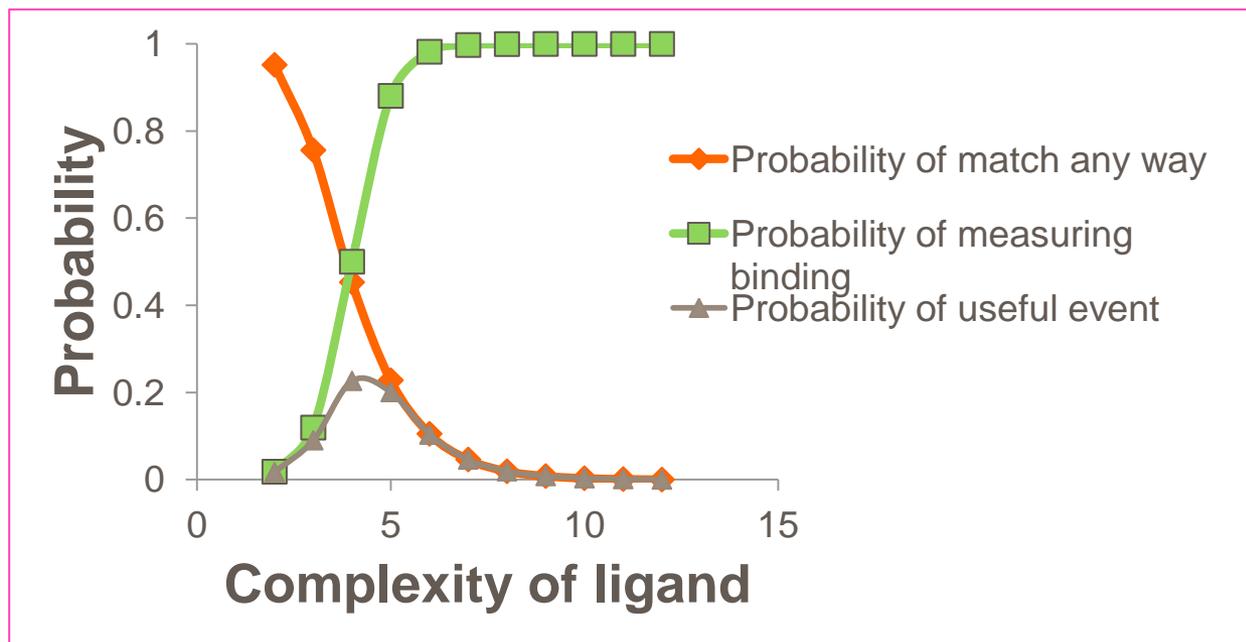
Why is it so difficult?



1. Solvation accountancy is challenging
2. Enthalpic interactions are directional, have more information content, and are harder to get right due to their complexity

The molecular complexity approach to thinking about interactions

Receptor + - - + + + - - + -
Ligand ++ - -



•Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. Hann MM.; Leach AR.; Harper G. *JCICS* (2001), 41(3), 856

•Molecular complexity and fragment-based drug discovery: ten years on.. Leach AR1, Hann MM. *Curr Opin Chem Biol.* 2011 Aug;15(4):489-96

•Coping with complexity in molecular design, A.R.Leach and M.M.Hann chapter in "de novo Molecular Design" ed – G. Schneider, 2014 (Wiley-VCH)

Complexity and high information content

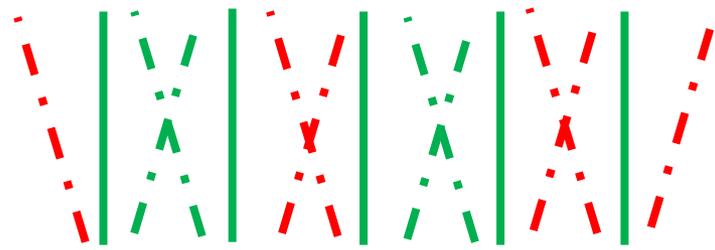


Receptor

+ - - + + - + -

Ligand

+ + - - +



= attractive primary interaction



or

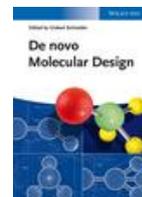
= attractive secondary interaction



or

= repulsive secondary interaction

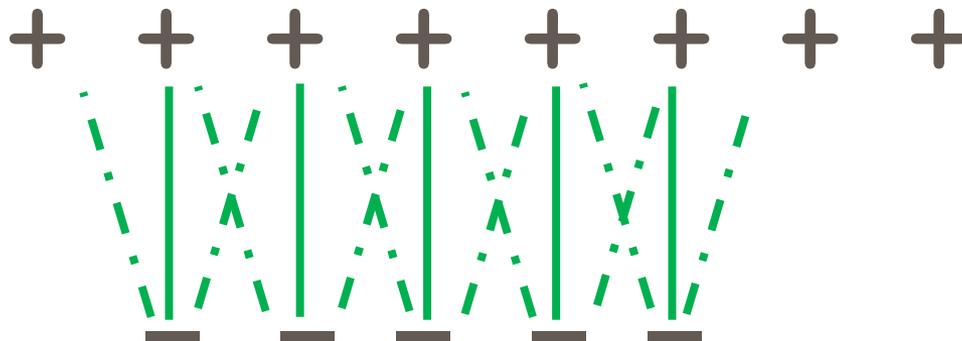
- High information content
- High Shannon entropy
- Difficult to shift < >
- Hard to get correct



Complexity and low information content



Receptor



Ligand

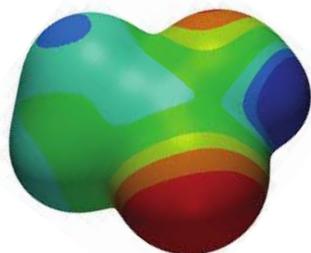
| = attractive primary interaction

/: or = attractive secondary interaction

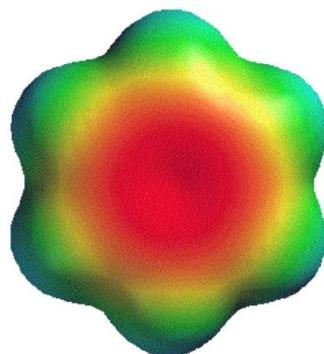
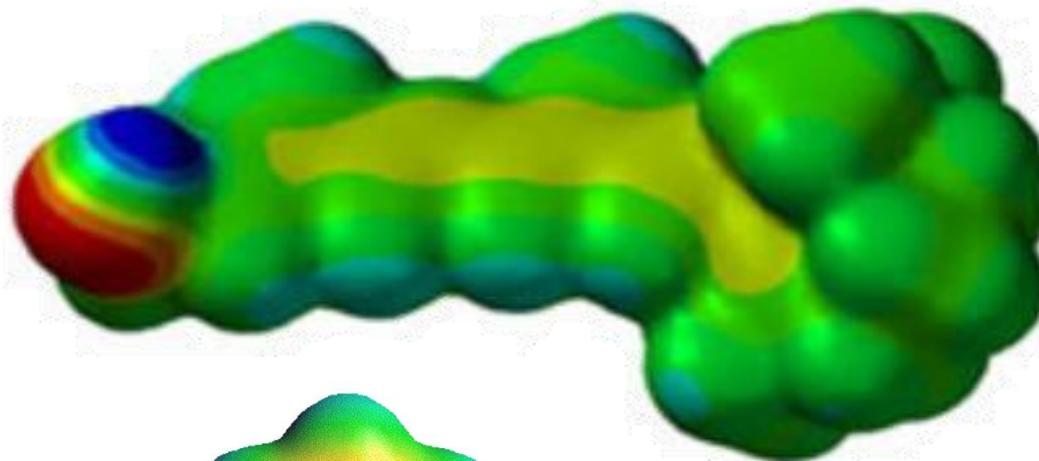
- Low information content
- Low Shannon entropy
- Easy to shift < >
- Easy to get correct



Information content per unit surface area



e.g amide
High information content
Directional interactions
Difficult



e.g. aromatic
Adaptable information content
Polarisable
Intermediate



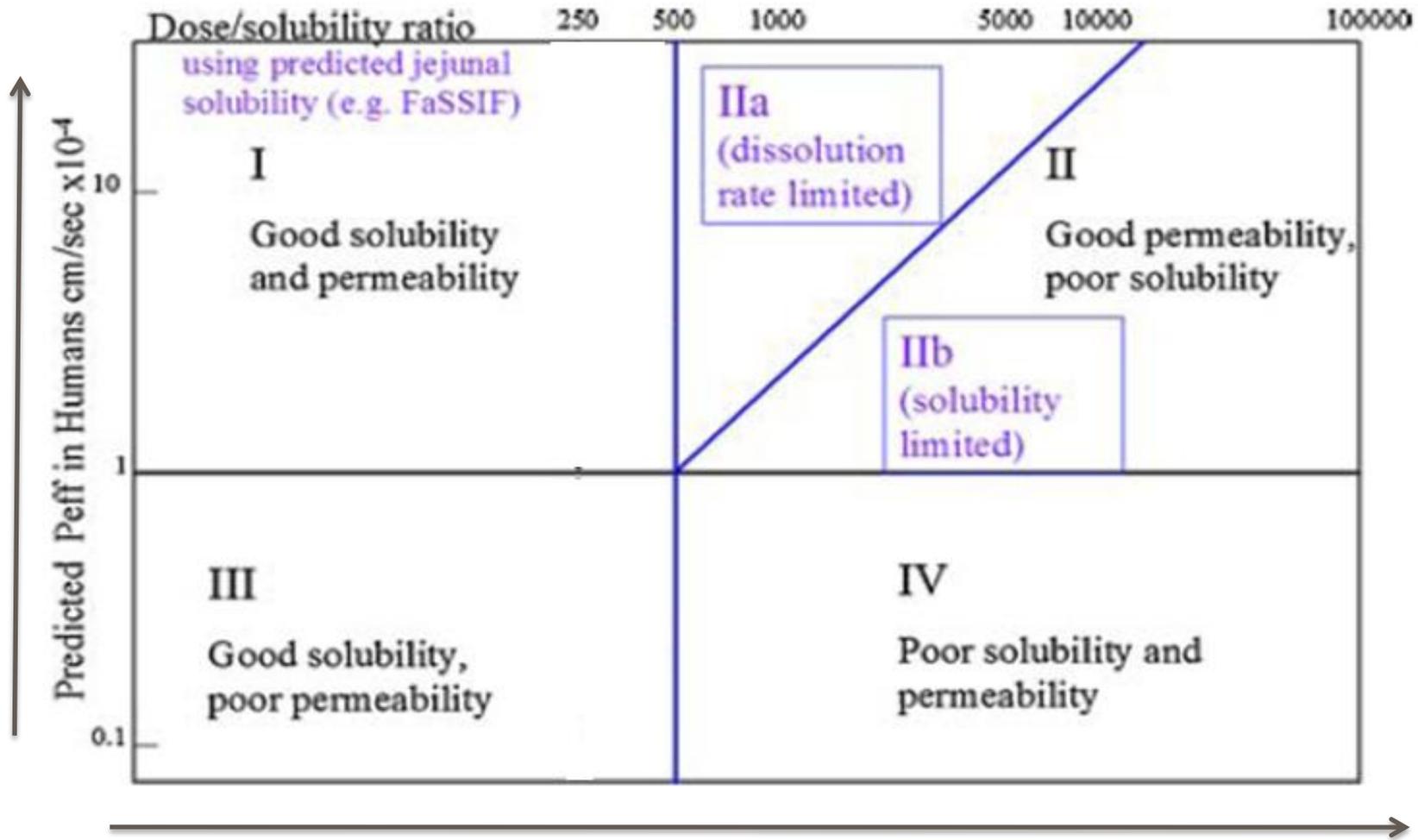
e.g. Aliphatic
Low information content
 δ +ve over entire surface
Lipophilic interactions
Easy!

More Molecular Obesity related issues



- **Every increase of logP by one unit increases by one order of magnitude the amount of compound present in membranes or bound to lipophilic proteins, etc..**
 - Home to key signalling proteins (GPCRs, ion channels, transporters etc) reside. Likely local high concentrations play havoc with them. => promiscuity
- **Lipophilicity is the antithesis of solubility – relying on formulation to get insoluble compounds on board is only going to aggravate the body!**
 - Your body can't easily eliminate lipophilic compounds (they are too insoluble!) so it has to work harder to make them more polar with higher energy species => toxicity

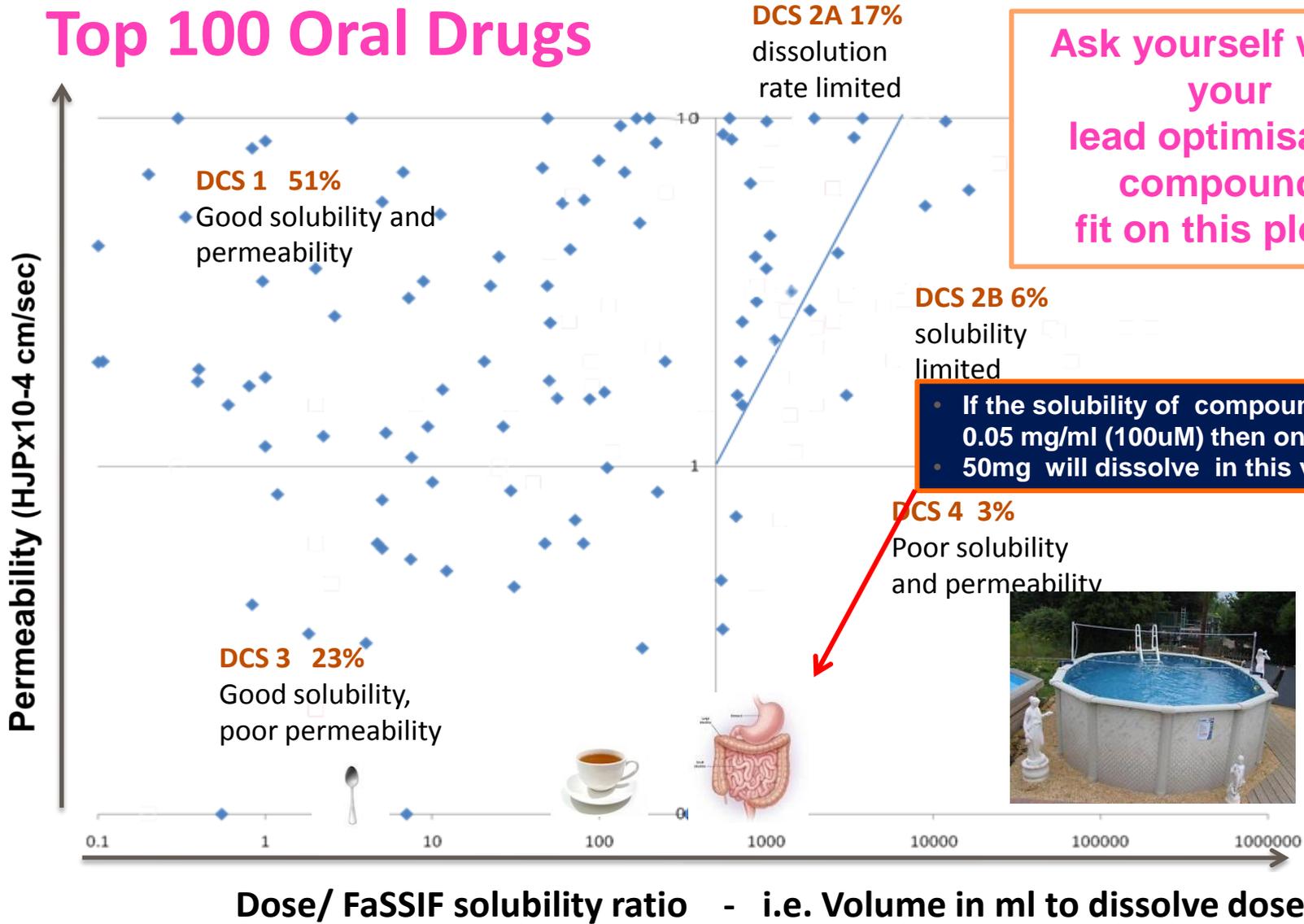
Developability Classification System DCS



Increasing dose/decreasing solubility = increasing Volume to dissolve the dose

Using knowledge of physchem properties & dose - the Developability Classification System DCS

Top 100 Oral Drugs



Historically potency is not everything either!

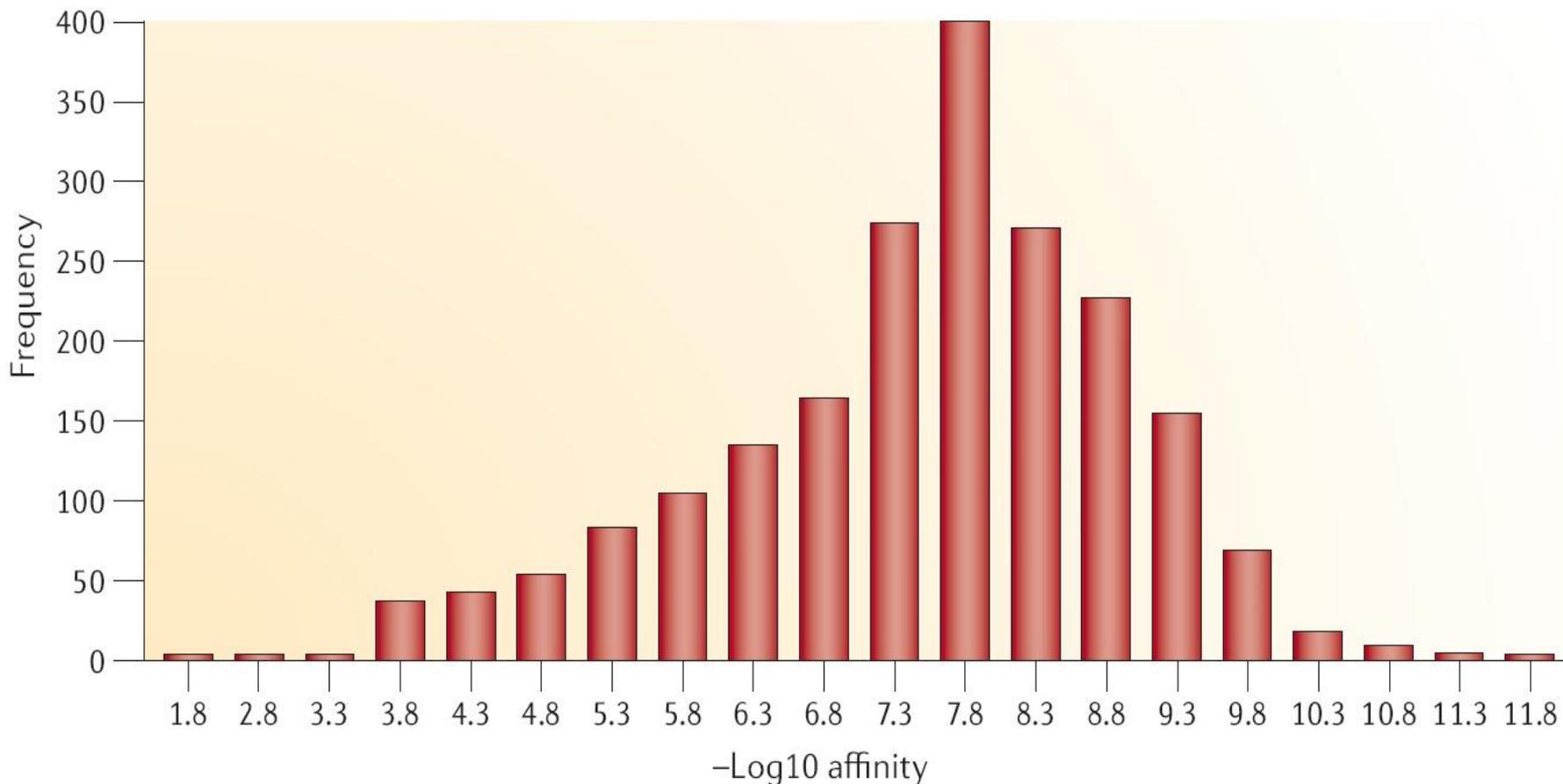
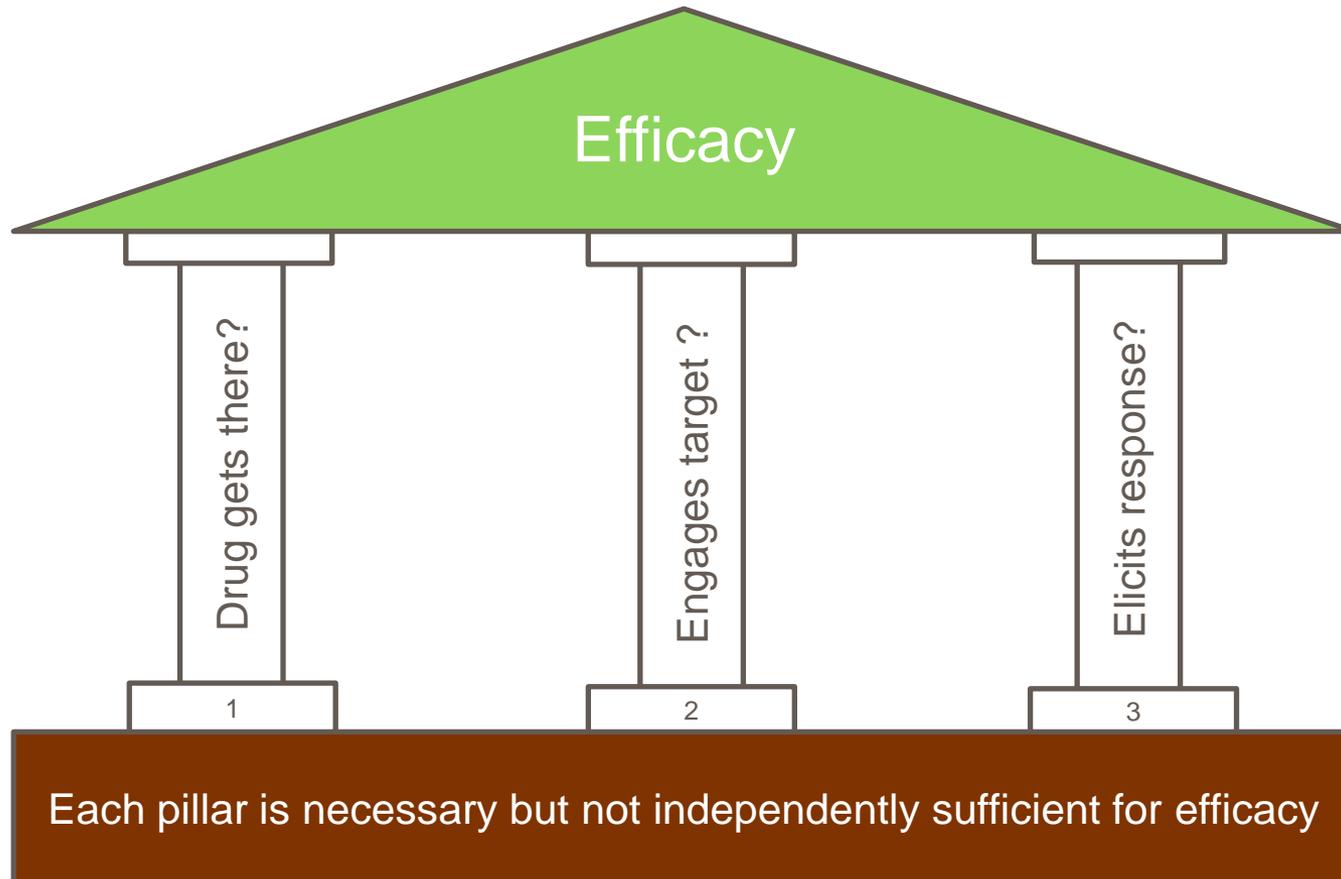


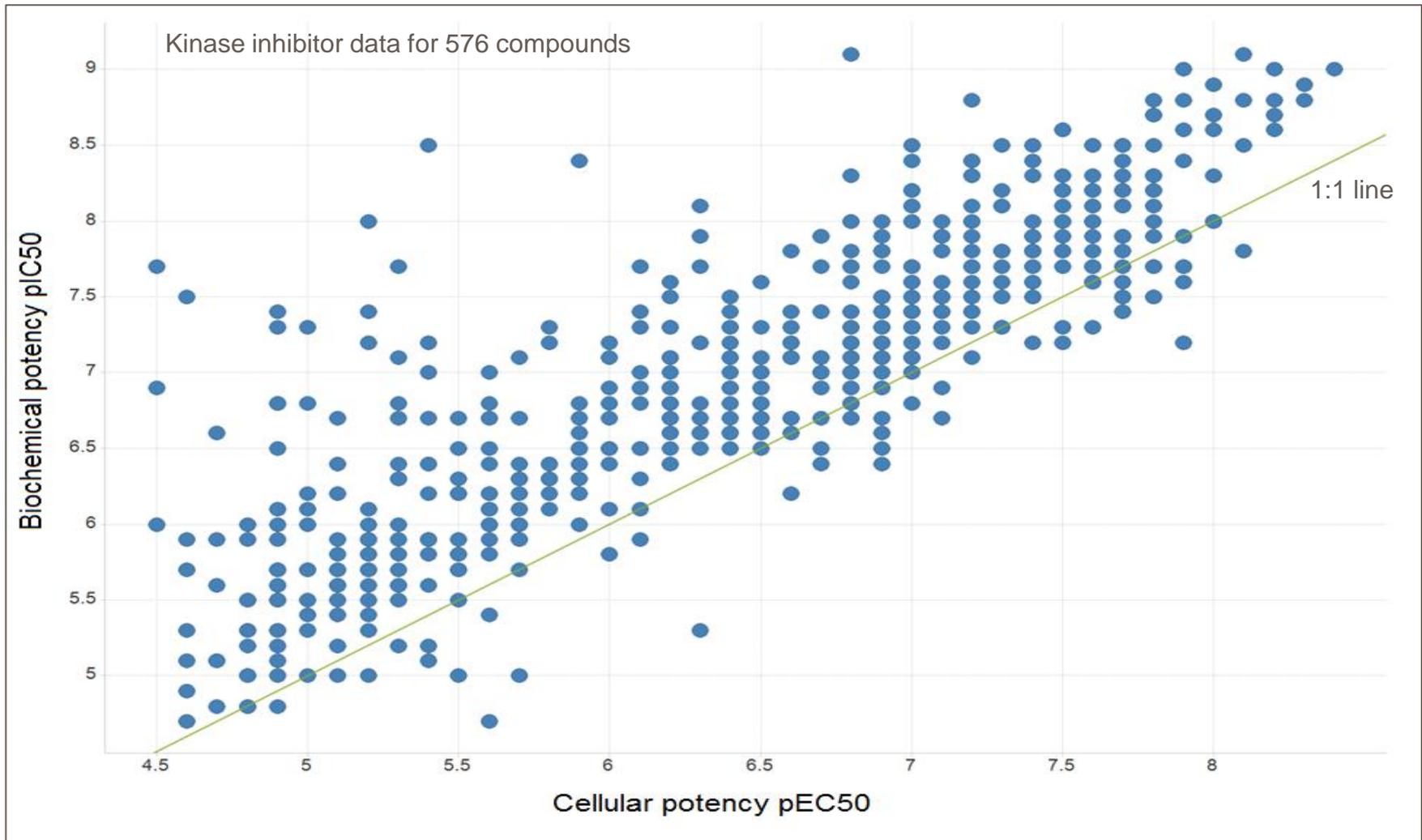
Figure 2 | **Frequency distribution for small-molecule drug potencies.**

How many drug targets are there? Overington, John P.; Al-Lazikani, Bissan; Hopkins, Andrew L.
Nature Reviews Drug Discovery (2006), 5(12), 993-996

Pfizer three pillars analysis



Typical “cell drop-off” effect compared to biochemical enzyme data - is your compound getting to the site of action



What really matters at the end of the day is dosage!

- Hence the interest in Drug Efficiency which tells you how much of your dose actually is available in the biophase of interest.

$$\text{DRUGeff} = \text{Biophase Concentration} * 100/\text{Dose}$$

Drug efficiency: a new concept to guide lead optimization programs towards the selection of better clinical candidates. Expert Opinion on Drug Discovery 2010, 5(7), 609-618; S Braggio, D Montanari, T Rossi & E. Rattl

- And more recently the use of Drug Efficiency Index as a strategy towards low therapeutic dose

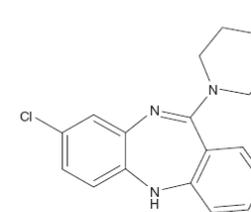
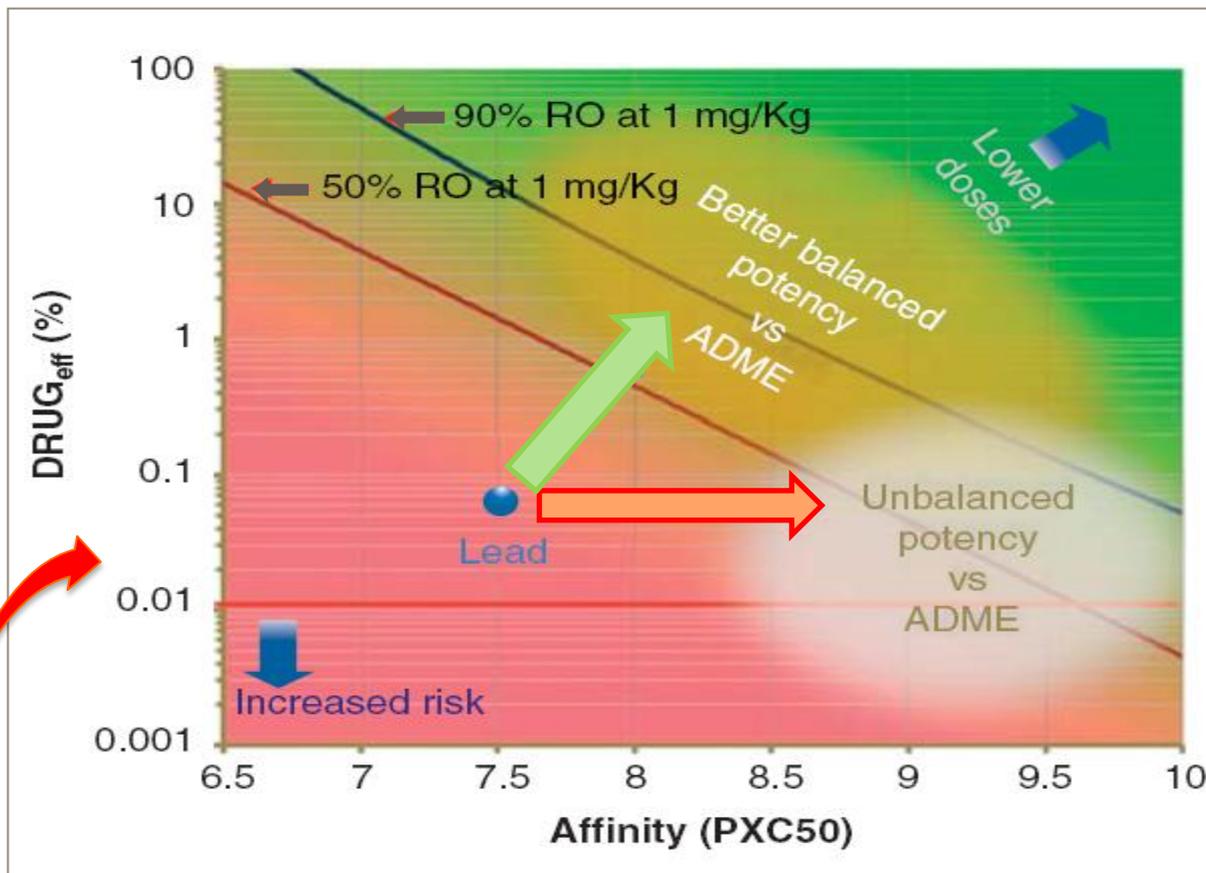
$$\text{DEI} = \text{Log}[\text{DRUGeff}(\%)] + \text{pKd}$$

DEI is a correction of the *in vitro* affinity by the *in vivo* pharmacokinetic potential.

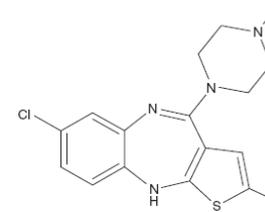
It is a simple descriptor directly connected to efficacy and therapeutic dose with the potential to probe the balance between *in vitro* affinity and ADME properties.

Application of drug efficiency index in drug discovery: a strategy towards low therapeutic dose. Montanari, Dino; Chiarparin, Elisabetta; Gleeson, Matthew Paul; Braggio, Simone; Longhi, Raffaele; Valko, Klara; Rossi, Tino. Expert Opinion on Drug Discovery, Volume 6, Number 9, September 2011, pp. 913-920(8)

Let's think more about what we are not using!



Clozapine



Olanzapine

Parameter	Clozapine	Olanzapine
Target affinity D2 pK _i	7.14	6.97
MW	327	312
clogP	3.71	3.01
Lipinski rule	Pass	Pass
LE	0.43	0.43
LLE	3.4	4.0
LELP	8.6	7.0
4/400 Rule	Pass	Pass
3/75 Rule	Fail	Fail
DRUG _{eff} (%)	0.03	0.85

- Clozapine = 99.97 % of drug is not being used for target engagement!
- What else is it getting up to???
- **High affinity tempts low Drug_{eff}**

And why do we waste compound?

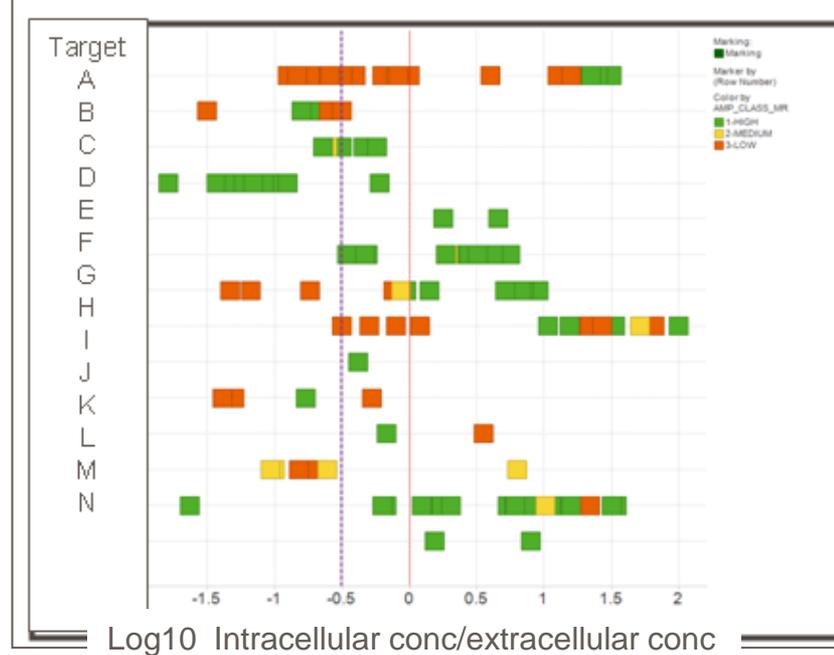
- We make very potent and lipophilic compounds which probably have very low free concentration at site of action (ie low $K_{p_{uu}}$)
- We assume the “free drug hypothesis” will allow compound to get to the site of action

- We measure blood concentration and then use AMPA/CACO2 measurements or logD models to guide our medicinal chemistry

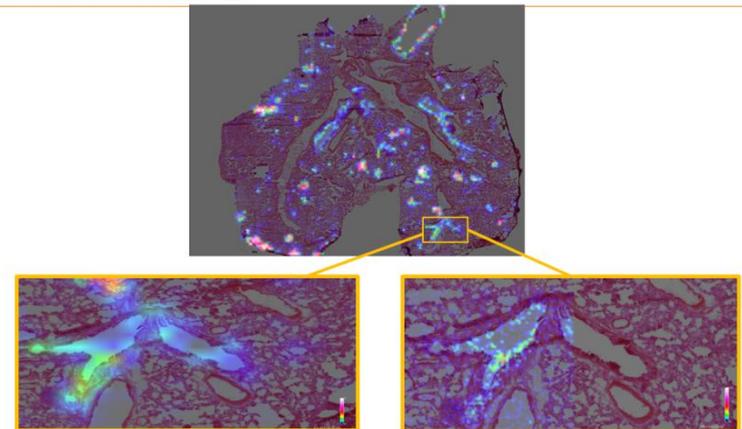
- But technology now exists to measure actual cellular concentration and disposition in early discovery

1. Incubate cells with compound, wash, rupture, extract, quantify by MS
2. MALDI/SIMS imaging

1. Methods to measure the intracellular concentration of unlabeled compounds within cultured cells using liquid chromatography/tandem mass spectrometry. L.M. Colletti et al. Analytical Biochemistry 383 (2008) 186–191
Rapid Measurement of Intracellular Unbound Drug Concentrations. A. Mateus et al. Mol. Pharmaceutics 2013, 10, 2467–2478
2. MALDI imaging in rodent lung slices showing compound distribution A. West and P. Marshall. GSK



Drug distribution in Rat lung MALDI MS imaging



Low resolution 200µm

High resolution 30µm

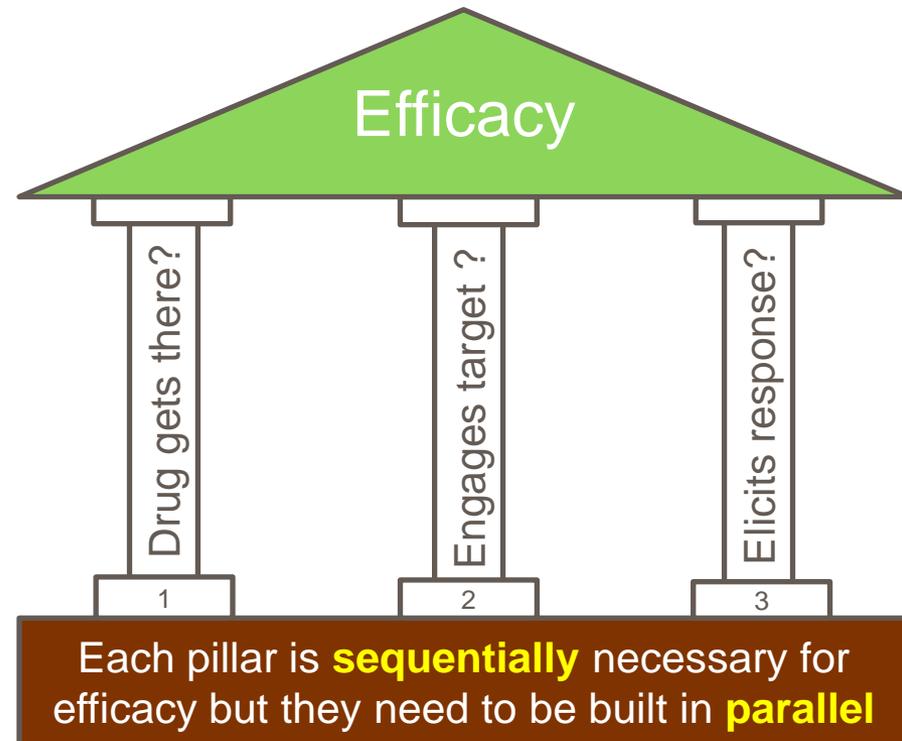
Moving from:

Typical order of events in a drug discovery screening cascade

1. Biochemical assay
2. Cellular assay
3. *in vivo* assay



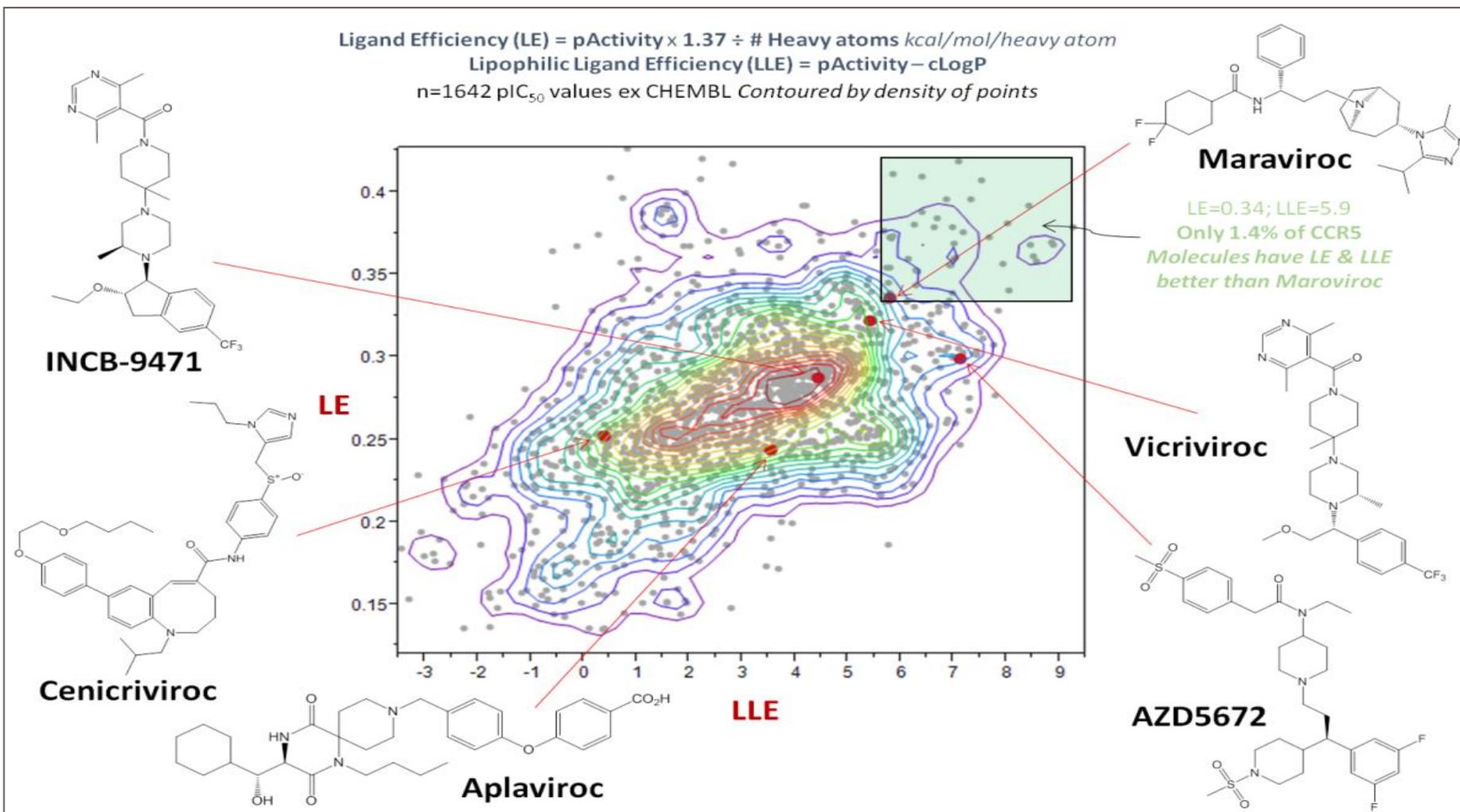
To:



Nobody said this was easy!



Distribution in LE/LLE space of a range of CCR5 antagonists



The role of ligand efficiency metrics in drug discovery . Andrew L. Hopkins, György M. Keserü, Paul D. Leeson, David C. Rees and Charles H. Reynolds. NRDD 13, 2014, 105-121

Summary



– **Medchem is a discipline and we should be Rigorous and Disciplined in making sure we make the very best molecules we can.**

- Not necessarily the most potent
- Not necessarily the easiest to make
- Not necessarily the quickest to find



– **Molecular Obesity has been killing us**

- We are addicted to quick wins – e.g. potency and its consequences

– **We have an increasing understanding of why and how to separate out the drivers to let molecules survive.**

- Start slim and stay fit! Control the risks!
- Know where your compound is going in lead optimisation when you can still do something about it!

– **Known knowns, known unknowns, unknown unknowns and**

The part that Donald Rumsfeld forgot

Unknown knowns

- Those things that are known but we don't know ourselves
 - Those things that are known but we have forgotten
 - Those things that are known but we choose to ignore

 - Lets try not to ignore the medchem knowledge that has been gained at very considerable expense over many years!
-

Acknowledgements

— Many at GSK present and past:
Darren Green, Klara Valko,
Andy Brewster, Ian Churcher,
James Butler, Rob Young,
Alan Hill, Stephen Pickett,
Anthony Taylor, Chris Luscombe,
Andy West, Peter Marshall,
Gareth Wayne, Bill Leavens,
Simon Readshaw, Laurie Gordon
Shenaz Bunally, Ian Reid
Martin Bayliss, et al ..

— Andrew Leach, Anne Hersey,
John Bradshaw, Gavin Harper,
Paul Gleeson, Dino Montanari,
Paul Leeson, Andy Hopkins et al

— Mike Waring, Chris Lipinski,
George Keseru
Per Artursson and Andre Mateus et al..

