Lecture Sept. 6 and 11, 2018

Where Do Recent Small Molecule Clinical Development Candidates Come From? | Journal of Medicinal Chemistry

Dean G. Brown and Jonas Bostrom

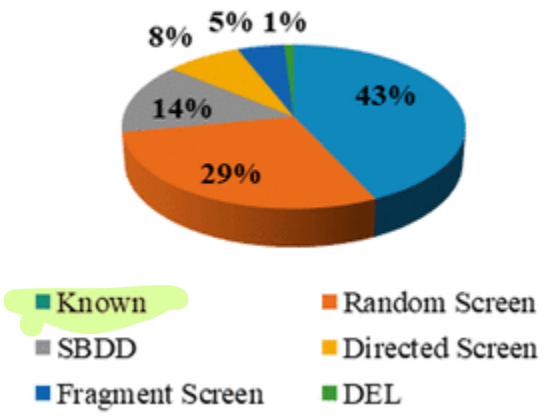
J. Med. Chem. **2018**, 61, 21, 9442–9468

Abstract

An analysis of 66 published clinical candidates from Journal of Medicinal Chemistry has been conducted to shed light on which lead generation strategies are most frequently employed in identifying drug candidates. The most frequent lead generation strategy (producing a drug candidate) was based on starting points derived from previously known compounds (43%) followed by random high throughput screening (29%). The remainder of approaches included focused screening, structure-based drug design (SBDD), fragment-based lead generation (FBLG), and DNA-encoded library screening (DEL). An analysis of physicochemical properties on the hit-to-clinical pairs shows an average increase in molecular weight (ΔMW = +85) but no change in lipophilicity (ΔclogP = −0.2), although exceptions are noted. The majority (>50%) of clinical candidates were found to be structurally very different from their starting point and were more complex. Finally, several reports of noncovalent scaffolds modified by a covalent warhead using SBDD approaches are discussed.

Key concepts

Lead generation strategies:



Target classes:

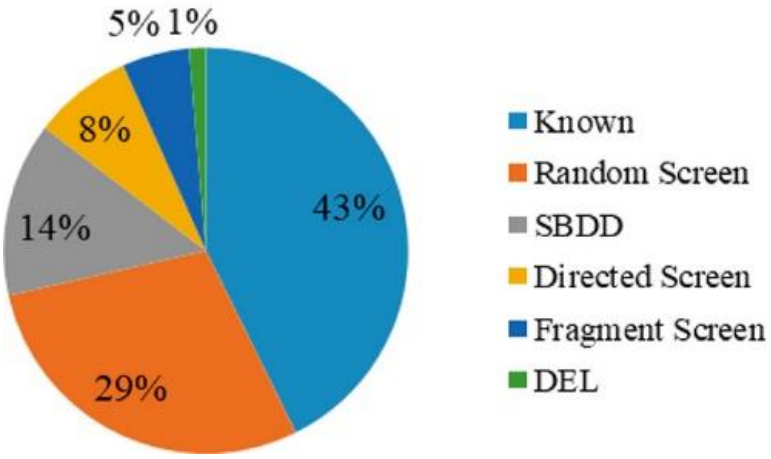
- Kinases (30%)
- Other enzymes (23%)
- GPCR's (17%)
- Epigenetic (9%)

Disease areas:

- Oncology (30%)
- CNS/Pain (18%)
- Infection (14%)
- Metabolic (12%)

[Epigenetics=the study of changes in organisms caused by modification of gene expression rather than alteration of the genetic code itself.]

Lead Generation Strategy (%)



Explain:

(1) Random screen or high throughput screen (HTS): screening of small molecules through collections of random compounds that can be large databases for large pharma companies (10^5-10^6)

(2) Directed screen: smaller set of compounds that are selected on prior knowledge of the target or chemical class. a.k.a. targeted, focused or biased screening.

(3) Structure-based drug design (SBDD): in silico screening of compound collections, including the use of protein 3-D structure.

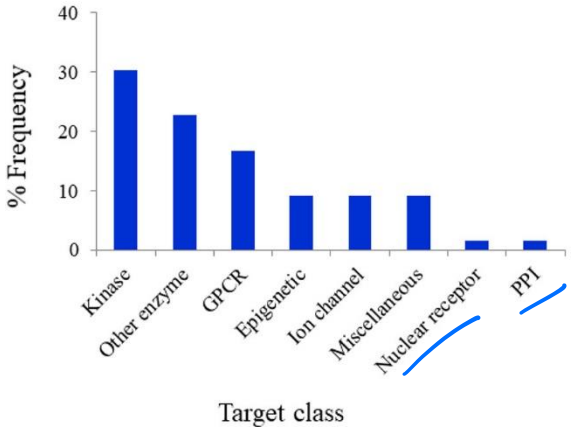
(4) Fragment screen or Fragment based lead generation (FBLG): typically libraries with a few thousand compounds of low molecular weight (mw<200 g/mol), screened at high concentrations.

(5) DNA-encoded library screening (DEL): Screening of very large collections (typically $>10^8$) of small molecule compounds using a technology that involves the conjugation of chemical compounds to DNA fragments.

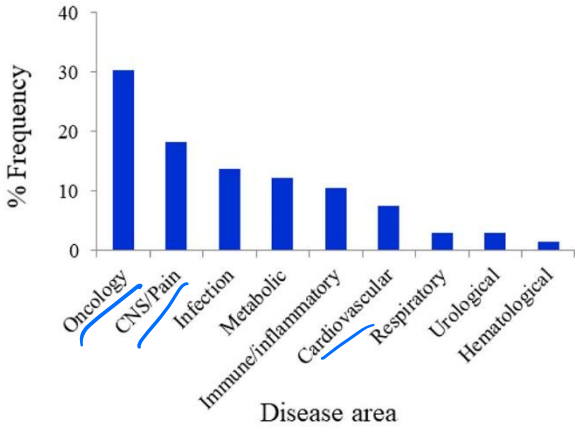
DNA → "barcode descriptor of chemical compound structure" → assay

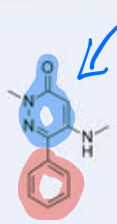
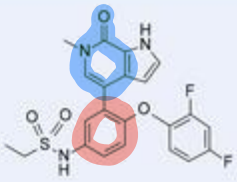
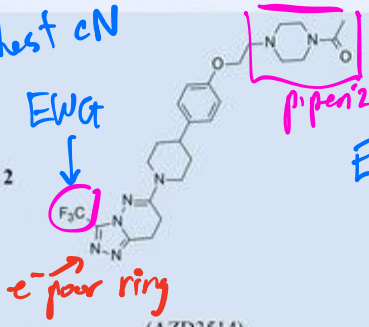
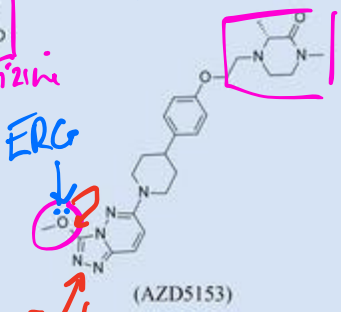
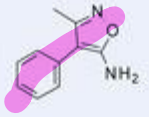
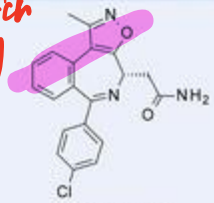
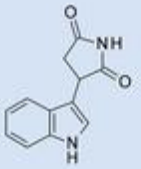
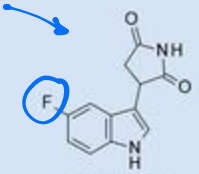
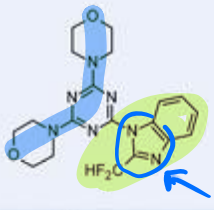
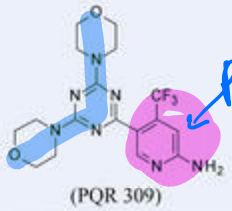
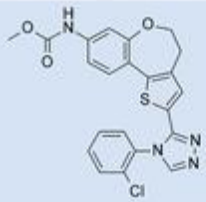
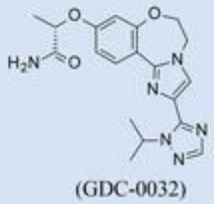
(6) Known compounds: project started based on a prior disclosure of an active compound, endogenous ligand or previous program

a) Target class distribution of hit-to-clinical pairs

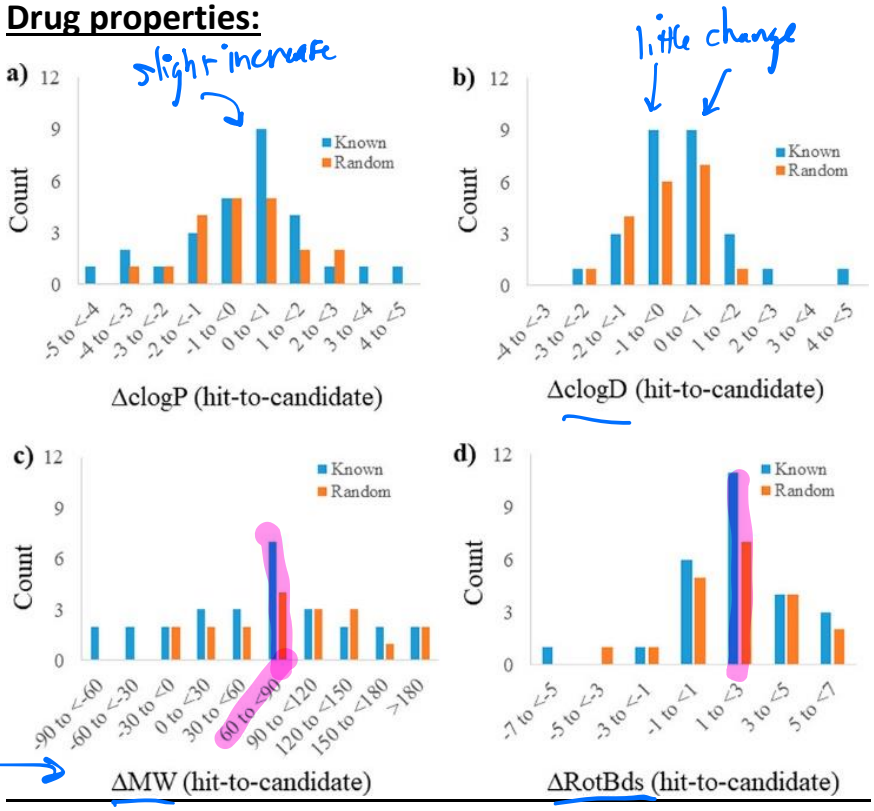


b) Disease area distribution of hit-to-clinical pairs



Entry	Hit	Clinical candidate	Hit finding method	Target	Disease area	Ref
1		 (ABBV-075, mivebresib)	Fragment Screen	BRD4	Oncology	13,14
2	 (AZD3514)	 (AZD5153)	Known	BRD4	Oncology	15
3		 (CPI-0610)	Fragment Screen	BRD4	Oncology	16
4		 (EOS200271, PF-06840003)	Random Screen	IDO-1 inhibitor	Oncology	17
5		 (PQR 309)	Known	PI3K/ mTOR	Oncology	18
6		 (GDC-0032)	Known; SBDD	PI3kα	Oncology	19

Drug properties:



slight increase

little change

adding weight enhances potency

1. c log P = calculated log of the partition coefficient (P) = $\frac{[Drug]_{octanol}}{[Drug]_{water}}$

mimics lipids or cell membranes

any bio aqueous solution

(P) =

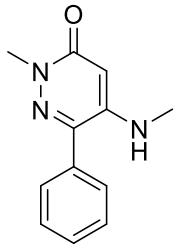
(note c = calculated versus experimentally determined)

-measure of solubility of drug (solute) in two layers, one organic and one aqueous → mimicking experience of drug passing through….

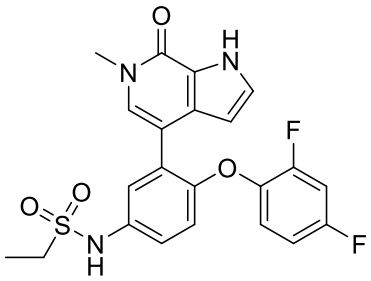
lipophilic proteins and fats, and aqueous solutions in body.

-Drug form is non-ionized

e.g. ChemDraw example.... from entry 1



initial hit c log P = 1.85



mivebresib (ABBV-075) c log P = 3.08 more lipophilic

$\frac{[Drug]_{octanol}}{[Drug]_{H_2O}} = \frac{100}{1} =$ log 10^2 $\log_{10} 10^2 = 2$

(references for compounds: J. Med. Chem. 2017, 60, 3828 and 2017, 60, 8369)

Note: final drug candidate increased in lipophilicity by more than one order of magnitude!

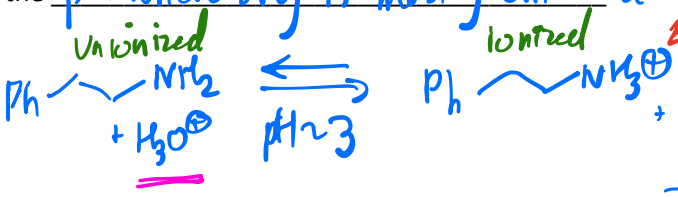
*most drugs have a log P = 0-4 blood-brain barrier

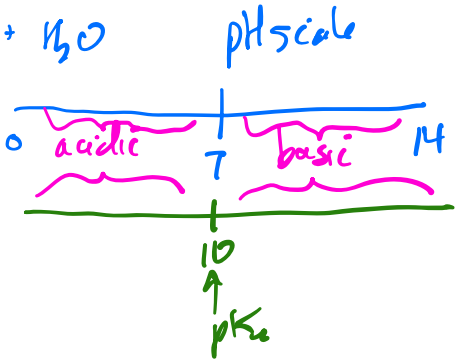
-barbituates (sedatives) all have a log P ~ 2 (need BBB penetration)

2. log D = Distribution coefficient, which is also the ratio of [Drug]octanol/[Drug]water, but it includes ionized forms of the drug (solute) so the pH of the aqueous solution matters. Note: most ionized form of drug will be in the aqueous layer

-log P = log D at the pH where drug is mostly unionized +/-

-e.g. ....





3. molecular weight... ✓

4. rotatable bonds= like it says

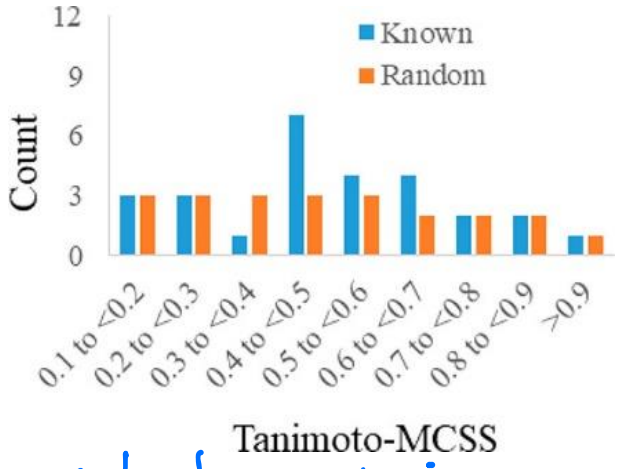
-measure of drug flexibility, but also entropy (freedom or disorder)

To understand this, consider...

In thermodynamic terms, what happens to a drug when it binds in a protein or enzyme? In particular, what effect does the binding have on entropy?

drug binding requires organization of drug's conformation to optimize binding in protein/target → loss in entropy

& more rotatable bonds allows more conformations but requires more organization or ordering on binding target.



-Tanimoto MCS = measure of molecular complexity (see J. Med. Chem. 2014, 57, 3186)

(MCS = maximum common substructure)

*Similar molecules have a Tanimoto number close to 1. Dissimilar molecules have a Tanimoto number close to 0.

LE = ligand efficiency = affinity of molecule for target →

(ΔG)/ number of heavy atoms (non-H's)

[great reference: "The role of ligand efficiency metrics in drug discovery" Andrew L. Hopkins, György M. Keserü, Paul D. Leeson, David C. Rees & Charles H. Reynolds, *Nature Reviews Drug Discovery* **2014**, *13*, pages 105–121]

-measure of how much binding each atom provides to overall molecule's affinity for target

-affinity measured as ΔG in kcal from K_i or IC_{50} → inhibitory constant at 50%

-Note: ΔG = $-RT\ln K_d$ or $-RT\ln K_i$

-the inhibition constant is a dissociation constant (K_D) for E-S or R-L. enzyme-substrate receptor-ligand

$ES \rightarrow E + S$ $\quad K_i = \frac{[E][S]}{[ES]}$ ← high number indicates strong affinity = very low value

$R\text{-}L \rightarrow R + L$

10^{-9} or 1 nM ← good for drugs

10^{-6} or 1 μM ← OK, weaker

-also sometimes IC_{50} is used in place of Ki

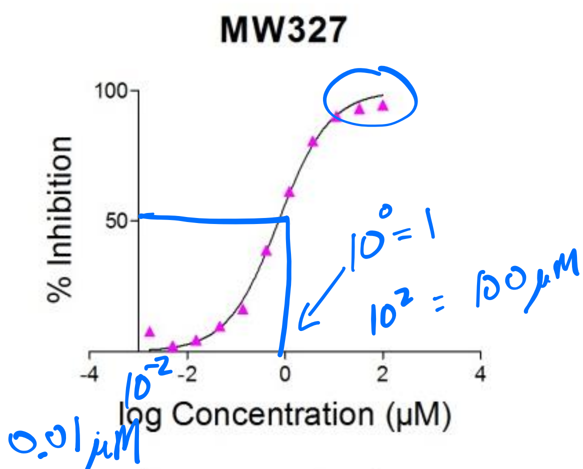
>>>SIDE BAR: IC_{50} = inhibitory concentration 50 %. >>

IC_{50}= inhibitory concentration 50%

- The concentration of drug which results in 50% inhibition of biological activity

-most common way to report drug inhibition of enzymes

-quick and dirty analysis: [S] held constant and [I] varied:



	327
HillSlope	0.8308
EC50	0.7322

IC_{50} ← ~1

-note: log Concentration is log[I], so 2 means 10^2

Potential flaws? ____

-calculation is done with energy units of ______________

Therefore, wit R=0.00198 kcal*K^{-1}*mol^{-1} and assuming T= 300 K, then…

ΔG = __________________________________ = -0.59 (2.303) log K_{eq} . = -_____________

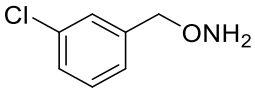
-number of heavy atoms (HA)= number of non-H atoms

Therefore equation can be written as

LE = -_____ or _____ or

_____

e.g.



K_i = 154 nM has _____ heavy atoms, therefore

LE = _____ → LE = _____

Optimal values?

"LE > _____ kcal per mole per heavy atom

From *Nature Reviews Drug Discovery* **2014**, *13*, pages 105–121

LLE = _____

=

-links _____ and _____in an attempt to estimate drug "likeness" (Wikipedia)

-c log P often substituted for log P; pIC50 is the –log (IC50)

e.g. IC50 ~ 10^{-9} M (nM level) and log P around 2, then…

LLE = _____

Optimal values?

LLE > _____

based on a Kd <10 nM molecule having a heavy atom count of 38 (~500 Da) and a cLogP < 3.

Key point: You can drive your potency simply by making a drug _____, but that won't help in the long run with _____. Often greasy molecules have bad _____.

Important points highlighted in Results and Discussion section

Overall: MW increased and structural complexity increased (Tanimoto change significant)

1. Random HTS versus Known compound leads→ _____

Message: Whether you start with a _____or with a _____ _____ compound, you have similar _____ happening.

2. More _____ atoms added on average: _____

3. 2/3rds of clinical candidates added _____

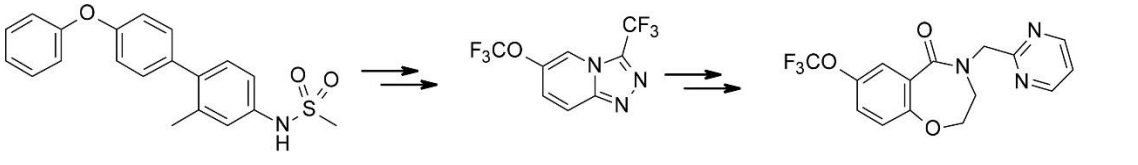
-properties of heterocyclic N? _____ _____

4. 1/3rd of clinical candidates _____

-properties of _____? _____

5. _____ common, but _____ also worked:

Fig. 5. Example of hit-clinical success from small, directed screening set



1

Hit (~800 cpds screened)
late I_{Na} block = 86% @ 10 μM

2

Lead GS-967
late I_{Na} block IC_{50} = 0.33 μM

3

Clinical candidate, GS-6615
late I_{Na} block IC_{50} = 0.88 μM

-not sure how the "hit" helped lead to the "lead" and then on to "clinical candidate"

6. Most screens used ____________________________________ biochemical screens

-isolated enzyme or receptor ____

-less common, but interesting: phenotypic screens→ _____________________________________

____

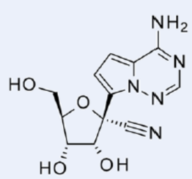
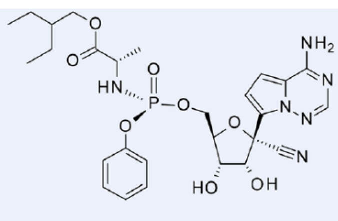
7. "Regulatory authorities ____"

**"Not all diseases can be sorted under the '________________________________' theory" !!

-prior to 2010, ___________ AstraZeneca internal drug projects lacked a clear link between main target and disease

8. Table 1, entry 55→ _____________________________ concept with an Ebola drug

-prodrug= ____

55		 (GS-5734)	Directed Screen	Antiviral	94

RANDOM INTERESTING FACT:

-2/3rds of all drugs have been evaluated for at least one other disease area (p. O, bottom right)

>>>

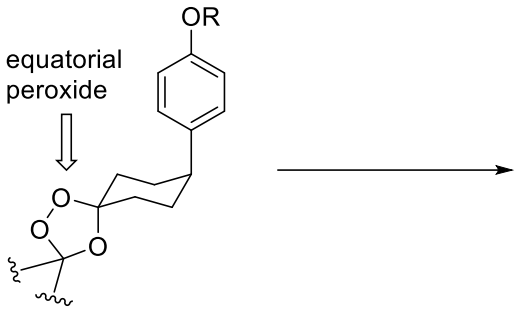
9. Concept: "____" drugs

-imitate another "_____________________________" drug

-1/3rd get ________________________________ by FDA, so there must be clear benefits

-an example of " _______________________________ " → _____________________________________ (IP) is important

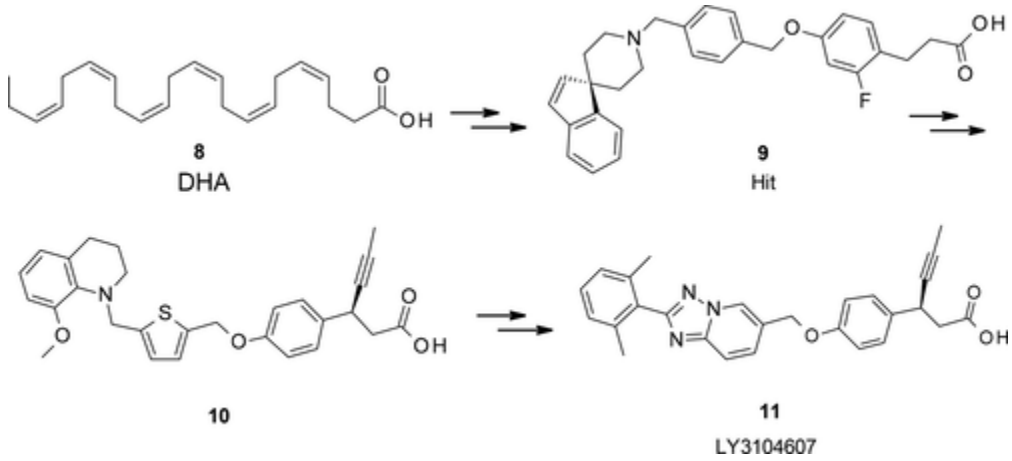
10. Table 1, entry 58: Chair _____ on peroxide structure



-better _____ for axial peroxide and improved _____

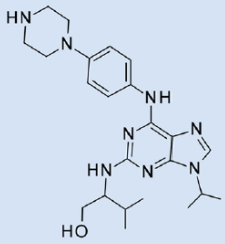
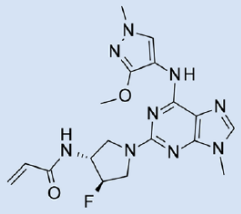
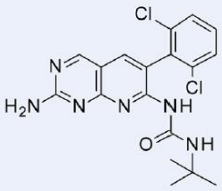
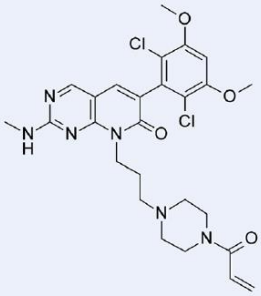
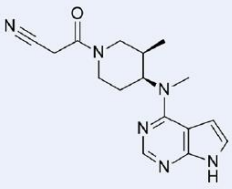
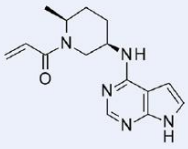
11. Table 1, entry 25 and Figure 8: natural substrate, DHA (docosahexanoic acid, an omega-3 fatty acid; a primary structural component of the human brain, cerebral cortex, skin and retina) is the natural ligand of GPR40 and was the _____.

-see similarity in _____ chain and benzene structures → both _____



12. IP issues _____! (p. Q, bottom right)

13. Table 1, entry 8, 11 or 61: _____ inhibitors

8		 (PF-06747775)	Directed Screen; SBDD	EGFR (mutant)	Oncology	23
11		 (PRN1371)	Known; SBDD	FGFR	Oncology	26, 27
61		 (PF-06651600)	Known; SBDD	JAK3	Inflamm. diseases	101

Source of covalent inhibition? _____

Entry 61 lead is _____ (trade name: Xeljanz, a JAK1 and JAK3 inhibitor) → clinical candidate (PF-06651600) is a JAK3 selective inhibitor. Development based on X-ray structure.

JAK= _____

>>

-From Young and Leeson [Mapping the Efficiency and Physicochemical Trajectories of Successful Optimizations | Journal of Medicinal Chemistry]

<u>4 major strategies that enable successful trajectories:</u>

1. seek necessary _____
2. _____
3. _____
4. _____