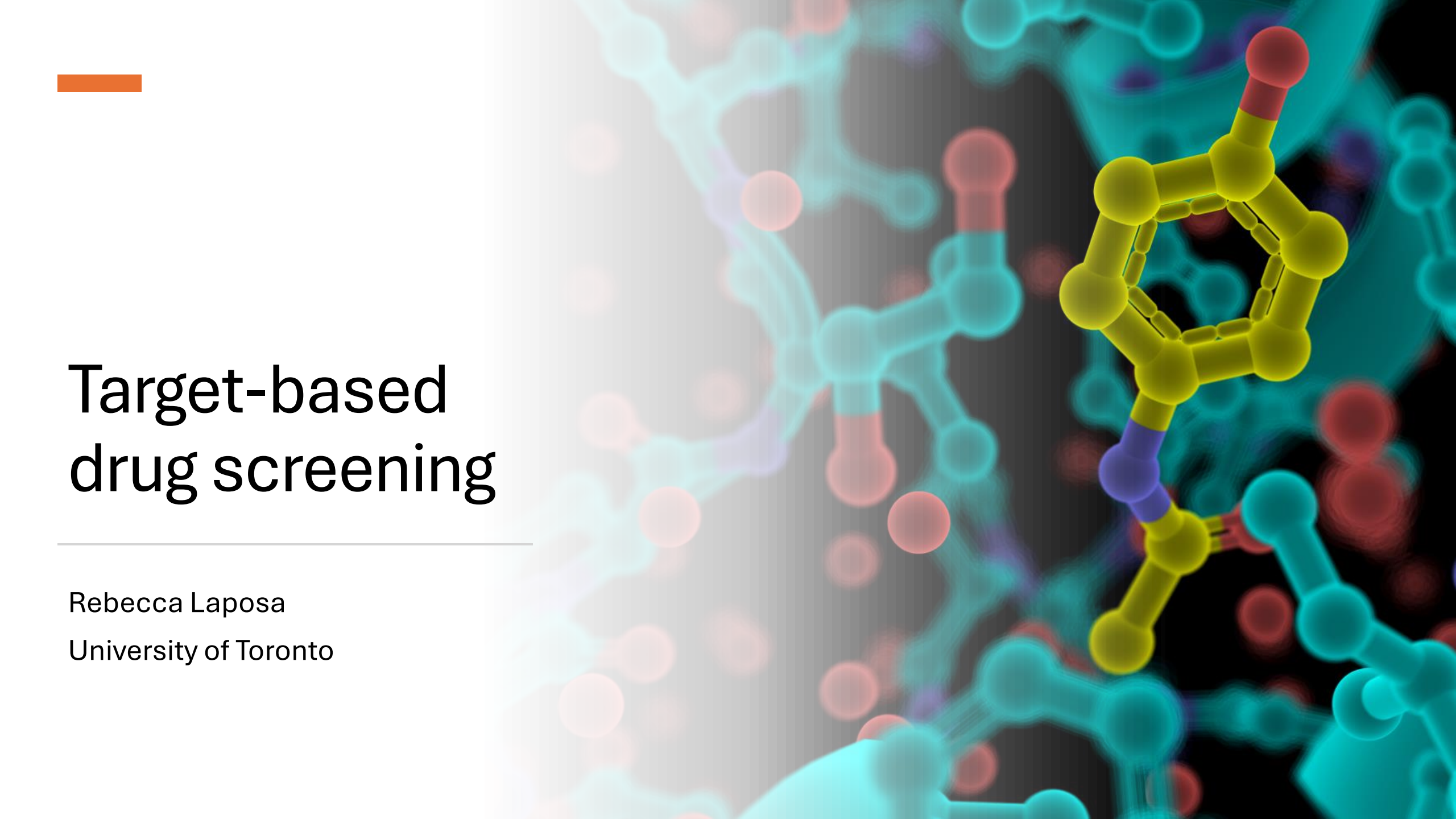
Target-based drug screening

Rebecca Laposa

University of Toronto



Learning objectives

In this lecture you will:

- Appreciate several major strategies for hit-finding in experimental drug screening
- Explore the contribution of chemical library choice to hit-finding
- Gain awareness of common artefacts of drug screens
- Be introduced to the complexity of drugs that intentionally target multiple proteins

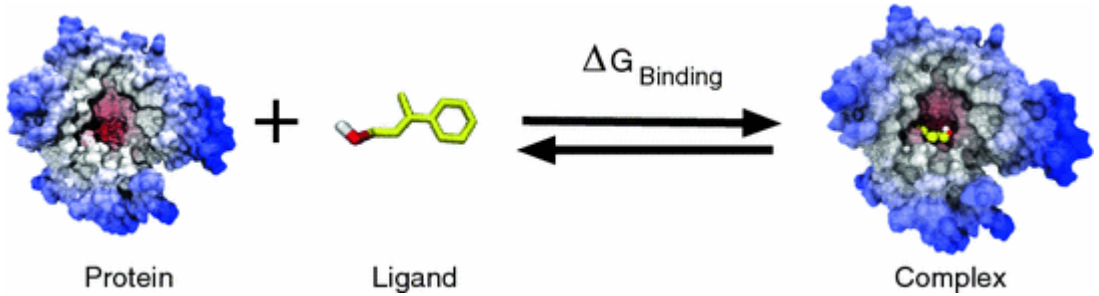
Early-stage drug discovery: **Hit-finding**

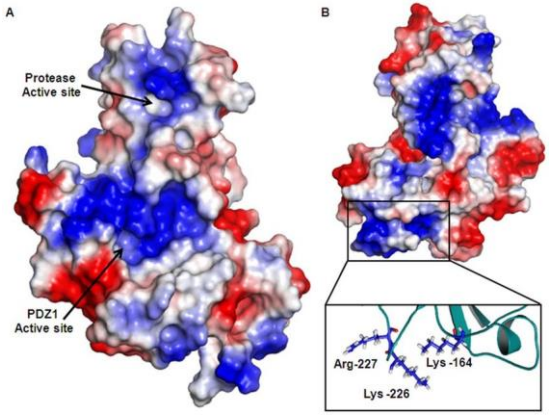
- In small-molecule (chemical) drug discovery, a **“hit” compound**:
- Binds to the protein target
- Modulates the function of the protein target
- Has relatively poor potency (eg 10 uM); the screen will often be conducted at a single concentration of 10 uM

- The “hits” will be improved by medicinal chemists to improve:
  - potency (10 uM hit to 20 nM clinical drug)
  - Pharmacokinetics (chemical properties of drug determines how drug moves in people, eg crosses membranes etc)

Molecular docking

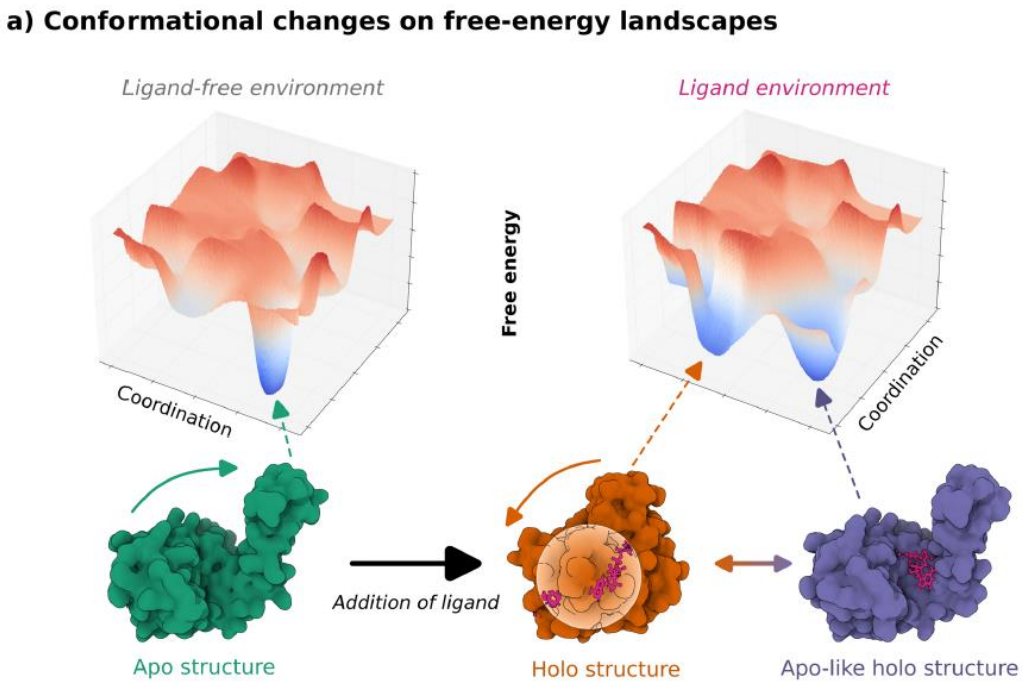
- The process of simulating the interaction between a small molecule (ligand or drug) and a macromolecule (eg a protein)
- Appropriate when a **protein's structure** is known (or well-predicted)
- Most intuitive:
  - Appropriate geometry (in 3D)
- Beyond just geometry,
- Multiple constraints:
  - Chemical interactions
    - Electrostatics and charge (see red/blue lower right)
    - Hydrogen bonding
    - Water – fat solubility (hydrophobicity, hydrophilicity)





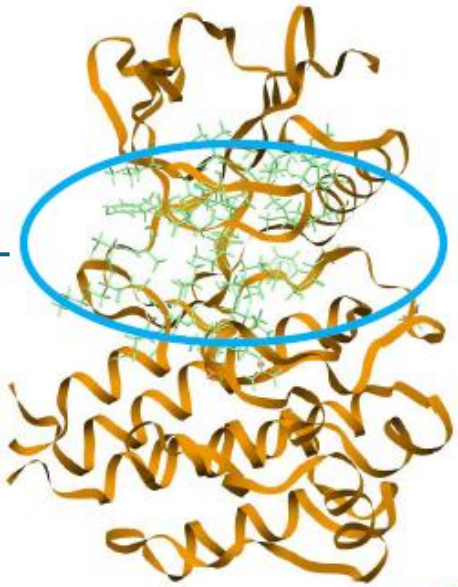
Protein-ligand complexes: distinct from the protein or ligand alone

- Last session: idea of induced fit. Chemicals change the protein structure (“induce” a fit)



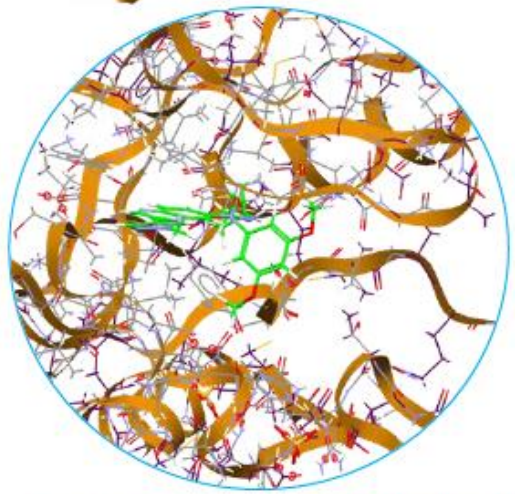
Considerations

- Binding site: a region of the protein to which a drug or chemical binds
- Pose: a specific position (often a modeled position) of the ligand-protein complex



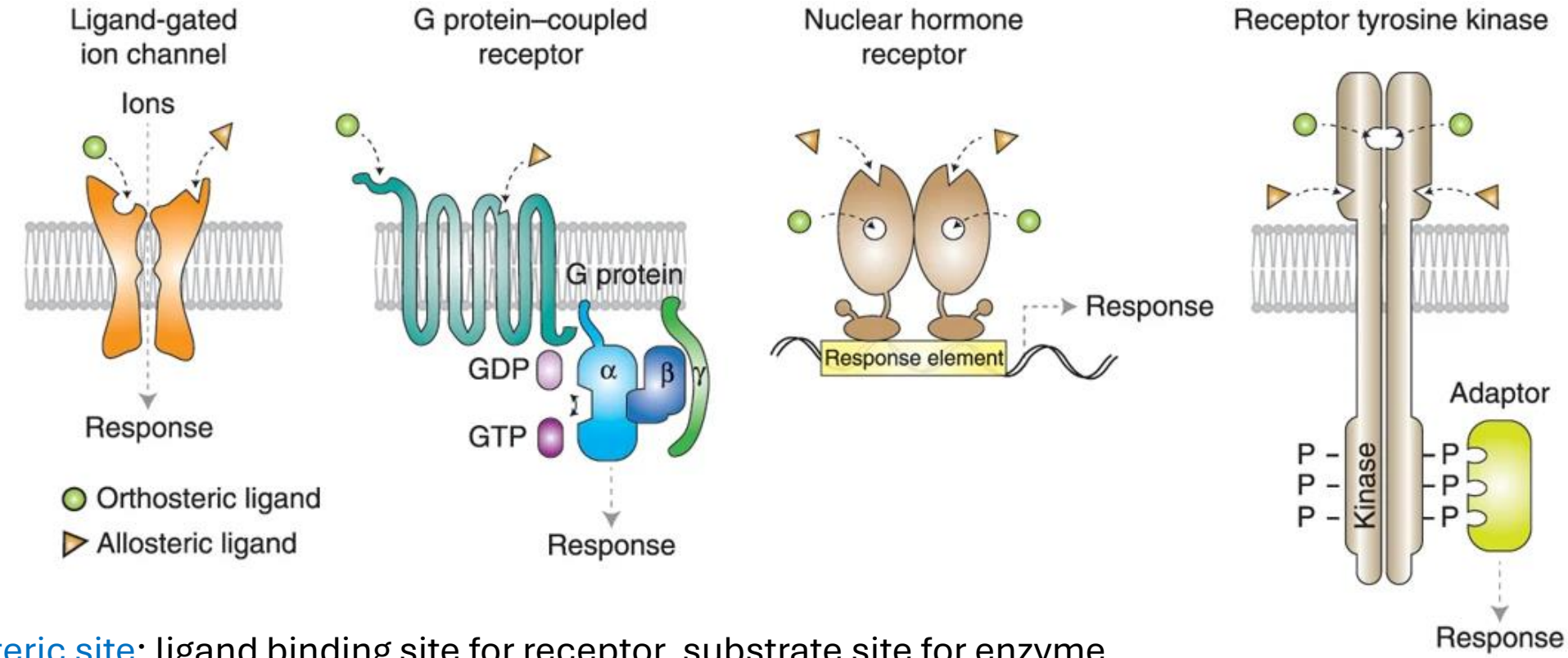
A protein displayed as a ribbon with its binding site represented in capped sticks style in green colour.

- Affinity: how tightly the chemical binds to the protein
- Affinity expressed as Kd (dissociation constant); low numbers = high affinity
- Screening assay:
  - Signal / no signal at single concentration of chemicals in a library
- Potency: drug's biological activity, expressed in units of concentration (IC50 = inhibitory concentration 50%)



Docking calculates a pose for the ligand and a score for the pose.

All major classes of receptors can be allosterically modulated



Orthosteric site: ligand binding site for receptor, substrate site for enzyme

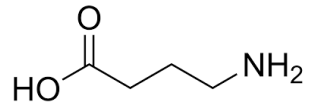
Allosteric site: site other than orthosteric site, causes shift in protein structure that indirectly influences protein function

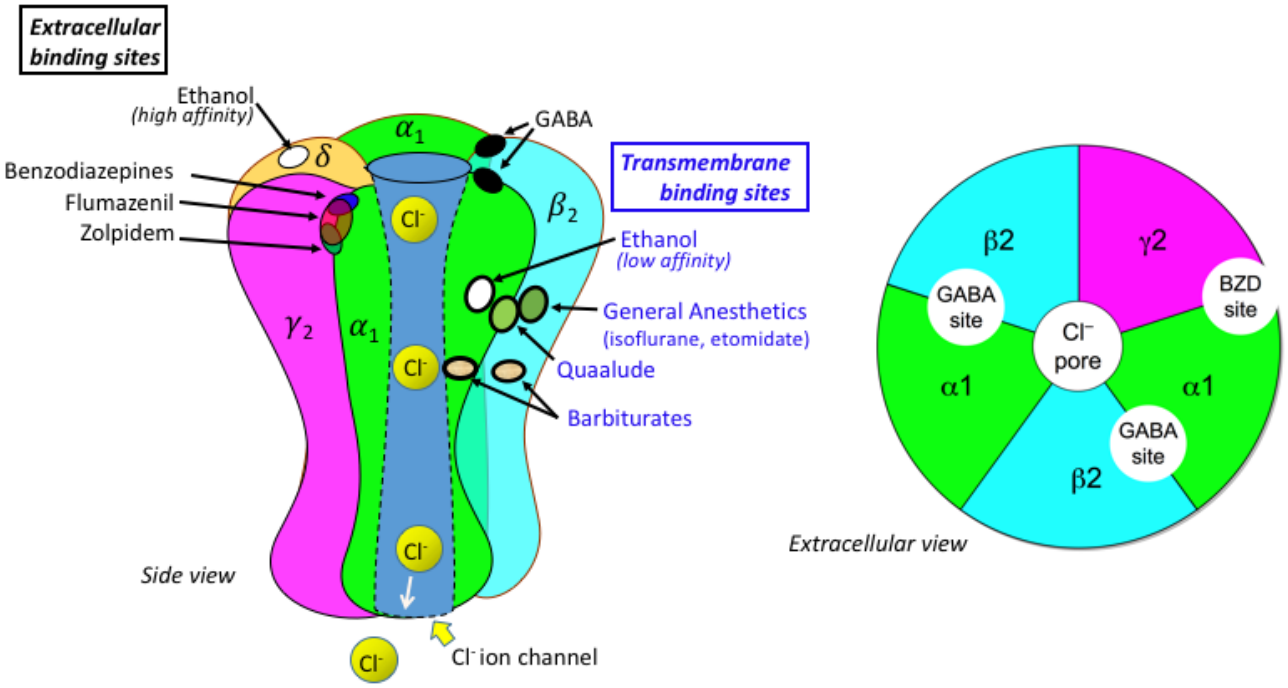
De Smet et al, Nat Biotechnol

Allosteric sites are important in pharmacology

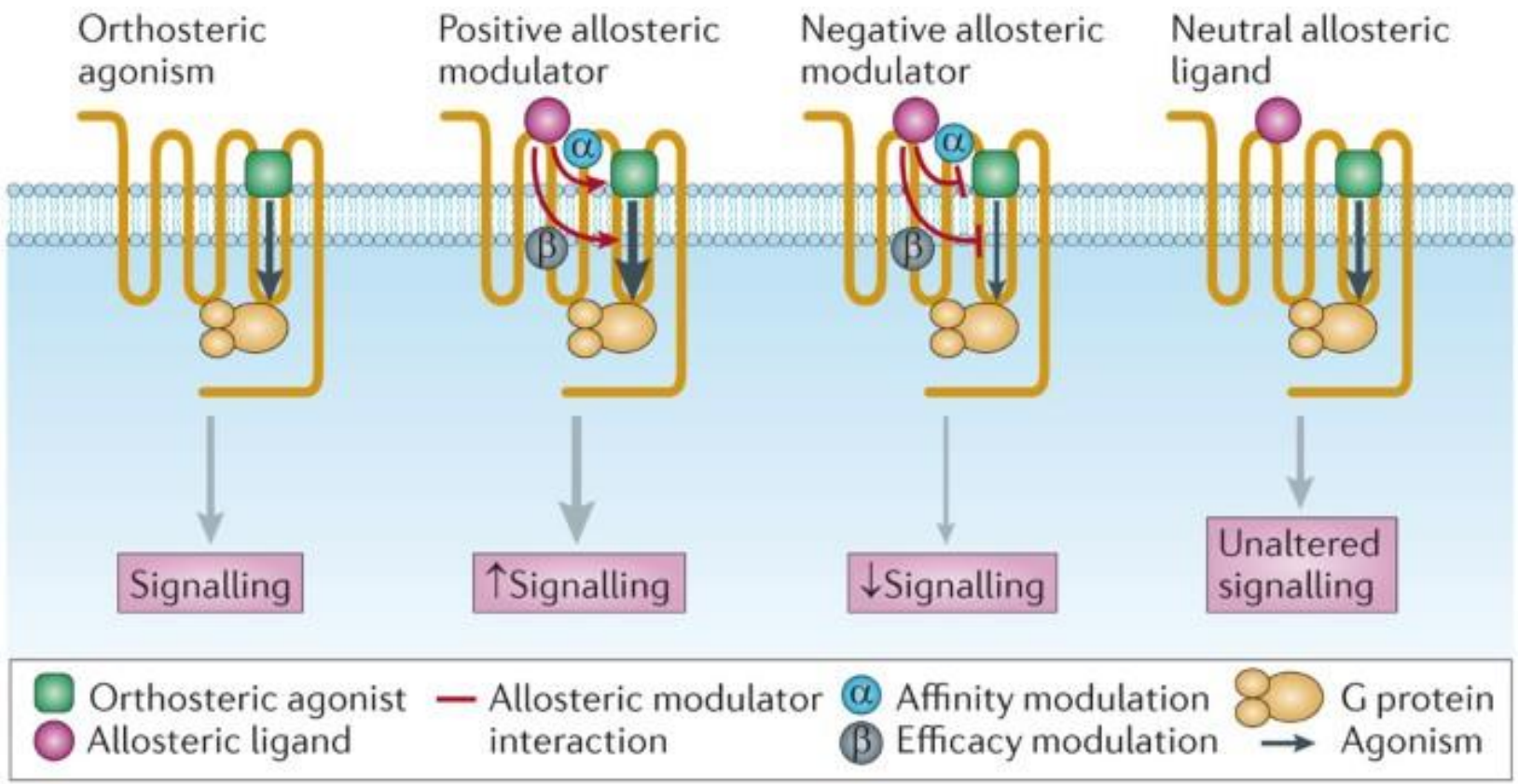
- Ligand-gated ion channel (GABA receptor)
- Endogenous ligand: GABA (neurotransmitter)
- Inhibitory neurotransmitter associated with mood
- Variety of allosteric sites with clinical impact
  - Benzodiazepines
  - Ethanol
  - Anesthetics

GABA



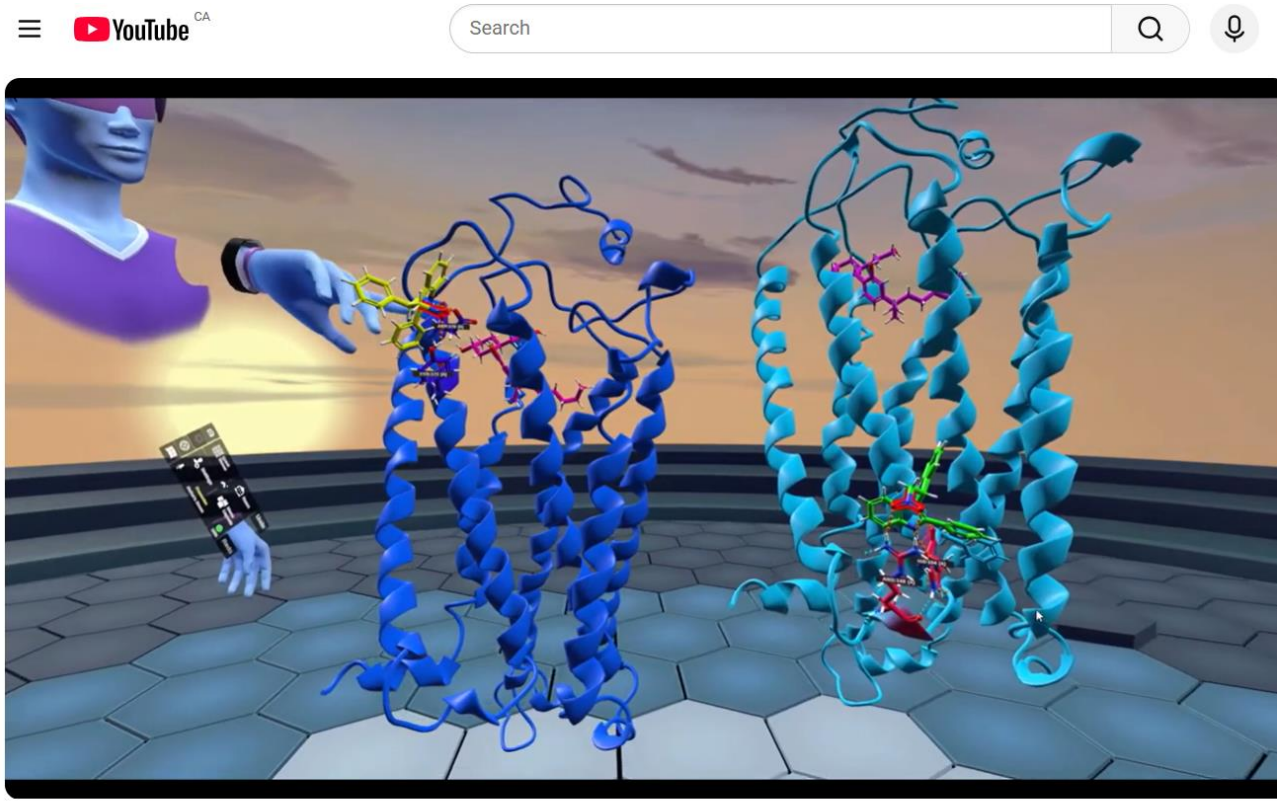


Orthosteric and allosteric drugs: G protein-coupled receptors



Nature Reviews | Drug Discovery

Exploring allosterics in Virtual Reality: Cannabinoid CB1 Receptor



Nanome Virtual Reality - Allosteric Modulation of CB1 Receptor

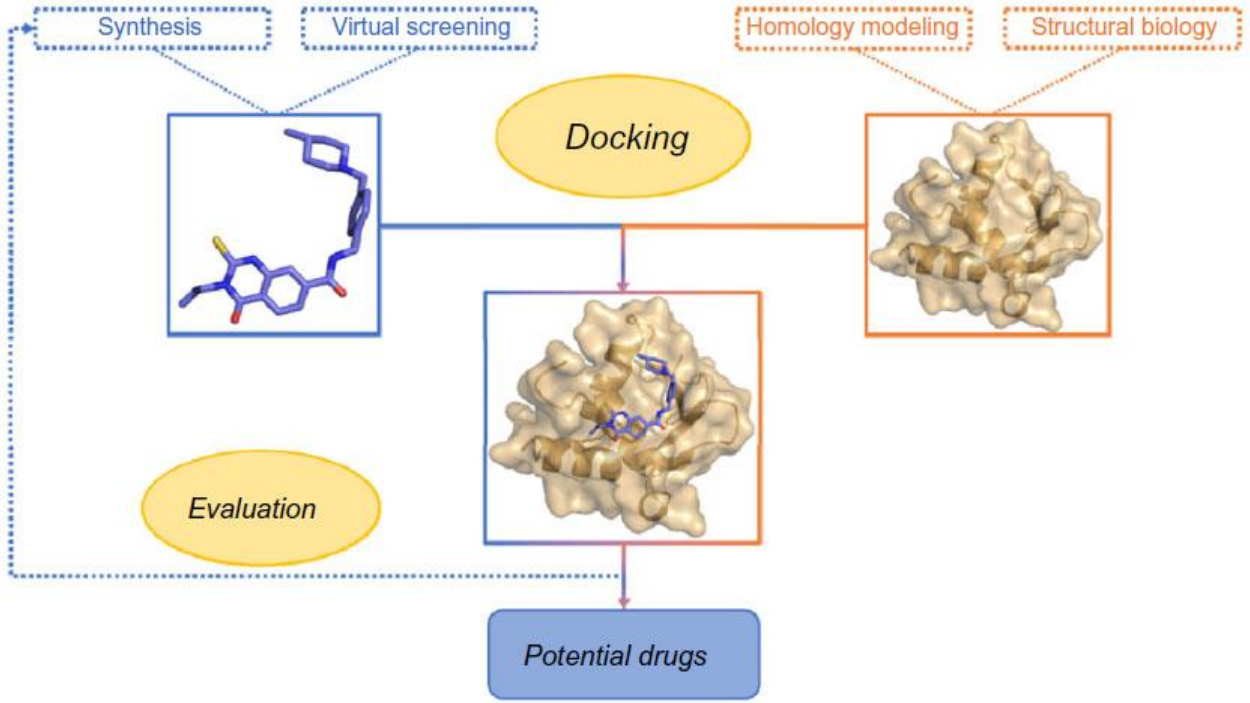
- https://www.youtube.com/watch?v=uVvDgSvrvYI

Modeling allosteric interactions

- Classical docking models assume stiff receptors

- Ensemble models: Implicit consideration of induced fit
  - Ensemble = set of all possible protein conformations
- Some models includeartial consideration of flexibility
  - May consider flexibility of amino acid side chains but not the protein backbone
- “Blind Docking” location of the chemical’s binding site is not determined ahead of time
  - E.g. DiffDock
- Models that simulate molecular dynamics: true induced fit
  - Motion is accounted for
- Emerging models: jointly model the protein-ligand complex, allowing each component to move
- Computational cost …

Returning to drug screening:
Iterative virtual – empirical cycles in drug discovery



deRuyjk et al., 2016

Traditional High-Throughput Screening

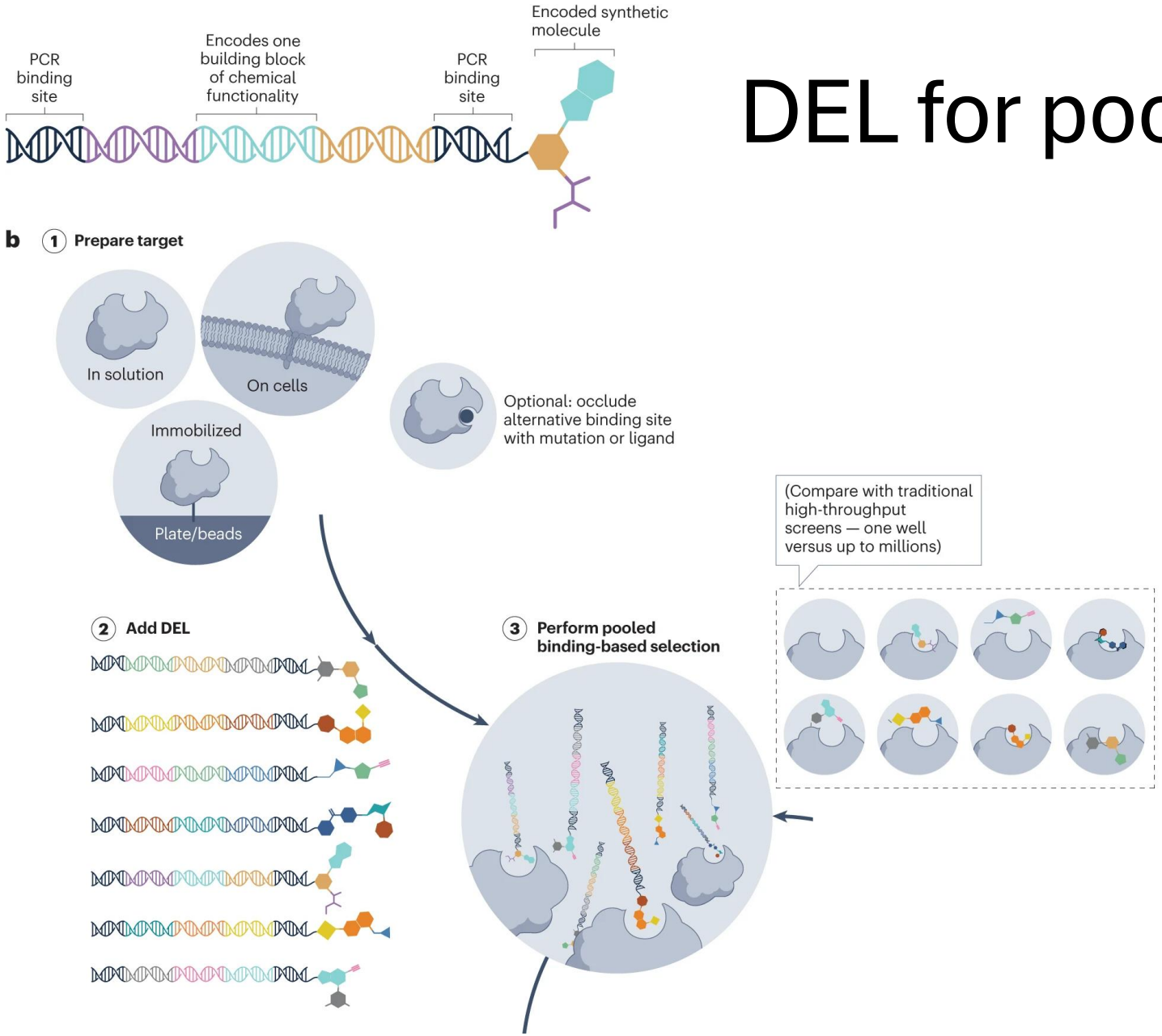
Each well: a single chemical from the chemical library
Test compounds at a single concentration, usually 10 uM

	96 wells	384 wells	1536 wells
Compounds/day:	10.000	40.000	200.000
Volume/well:	100-200 µl	30-100 µl	2.5-10 µl

Newer experimental libraries for hit-finding

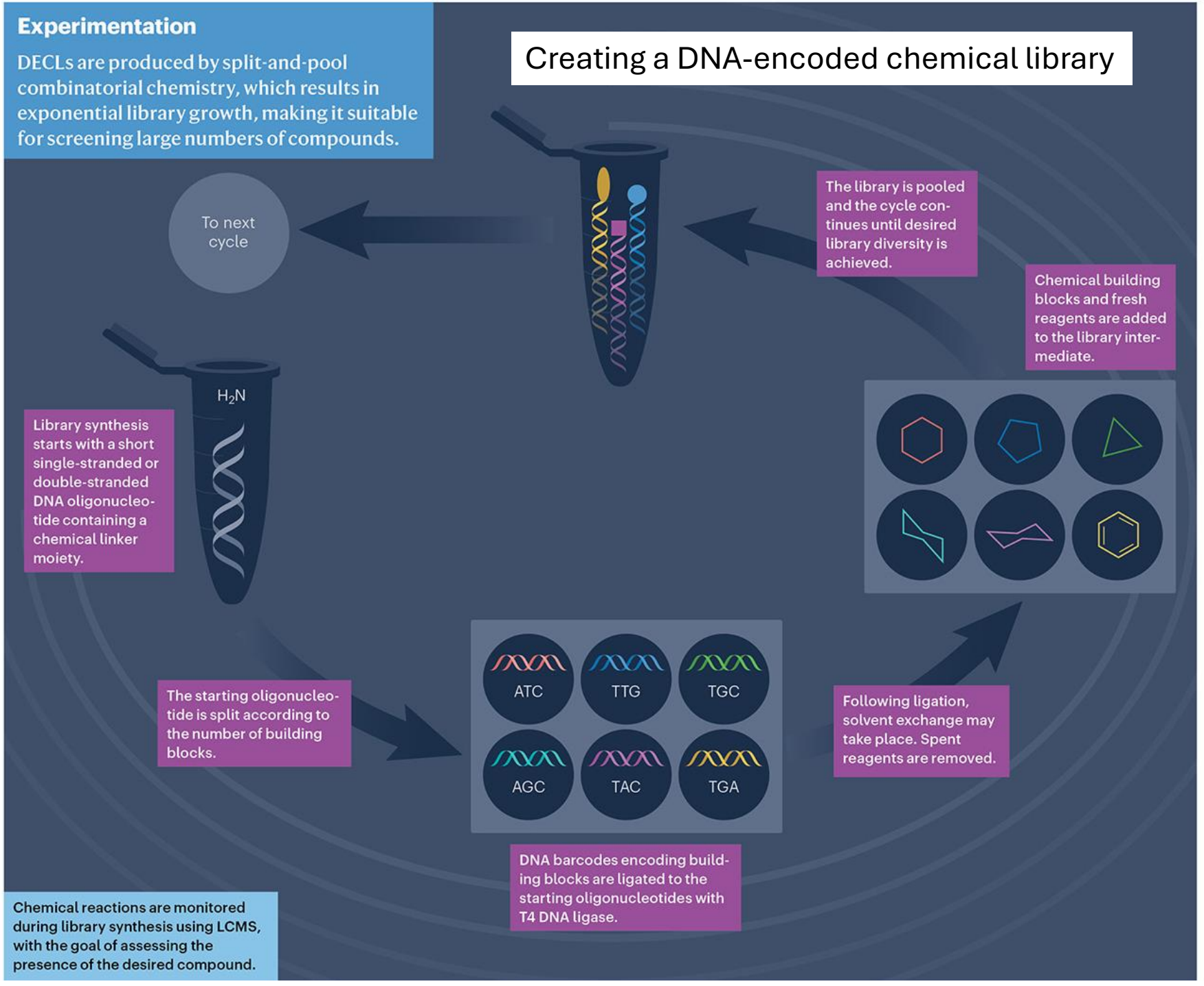
- DNA-encoded libraries (DELs) have a DNA tag attached to each chemical
- Why DNA as a tag?
  - DNA is highly **specific** (a sequence of ~20 bases maps uniquely to a position in the 3-billion-base human genome
  - DNA is highly **sensitive** – we can detect very tiny amounts of it.
    - Forensics, polymerase chain reaction (PCR) to amplify
  - Consider DNA tag as the “Read” operation.
- Advantage of DNA-tagged chemicals: pool the chemicals together
  - Miniaturization, much higher throughput

DEL for pooled screening



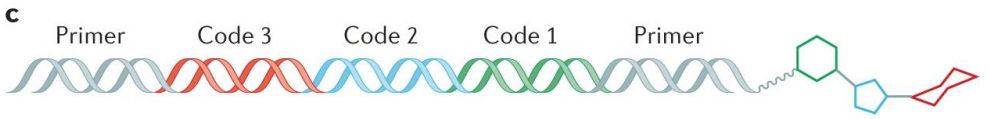
Peterson and Liu, NRDD 2023

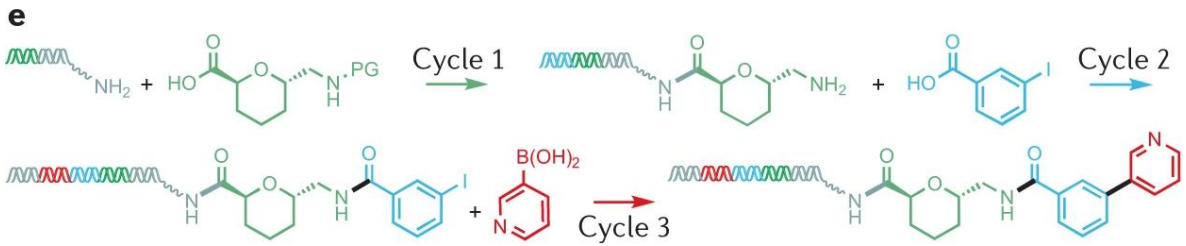
Creating a DNA-encoded chemical library

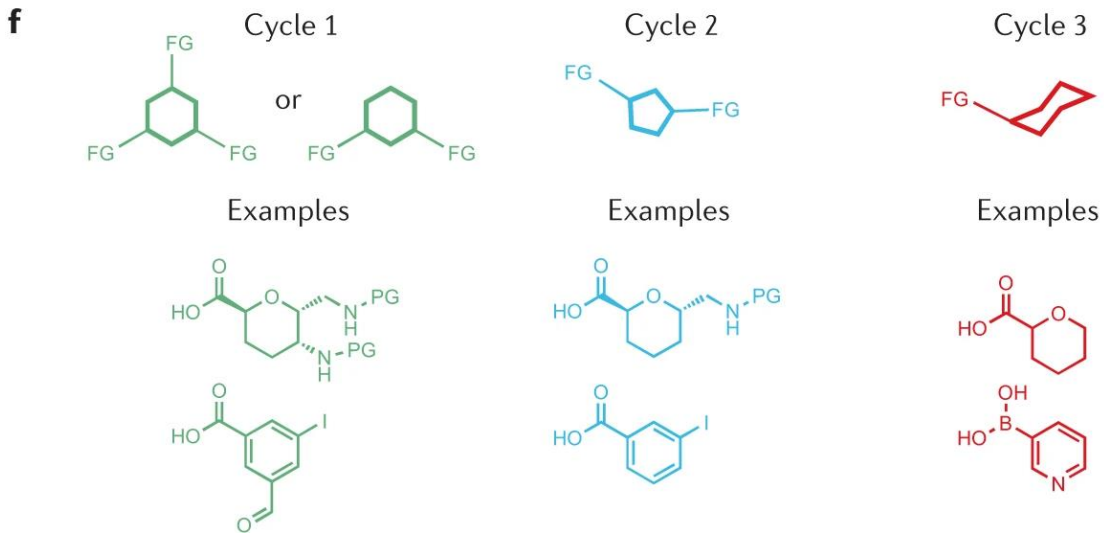


Nat Rev Meth Primers 2022

Combinatorial chemistry to create DNA-encoded chemical library

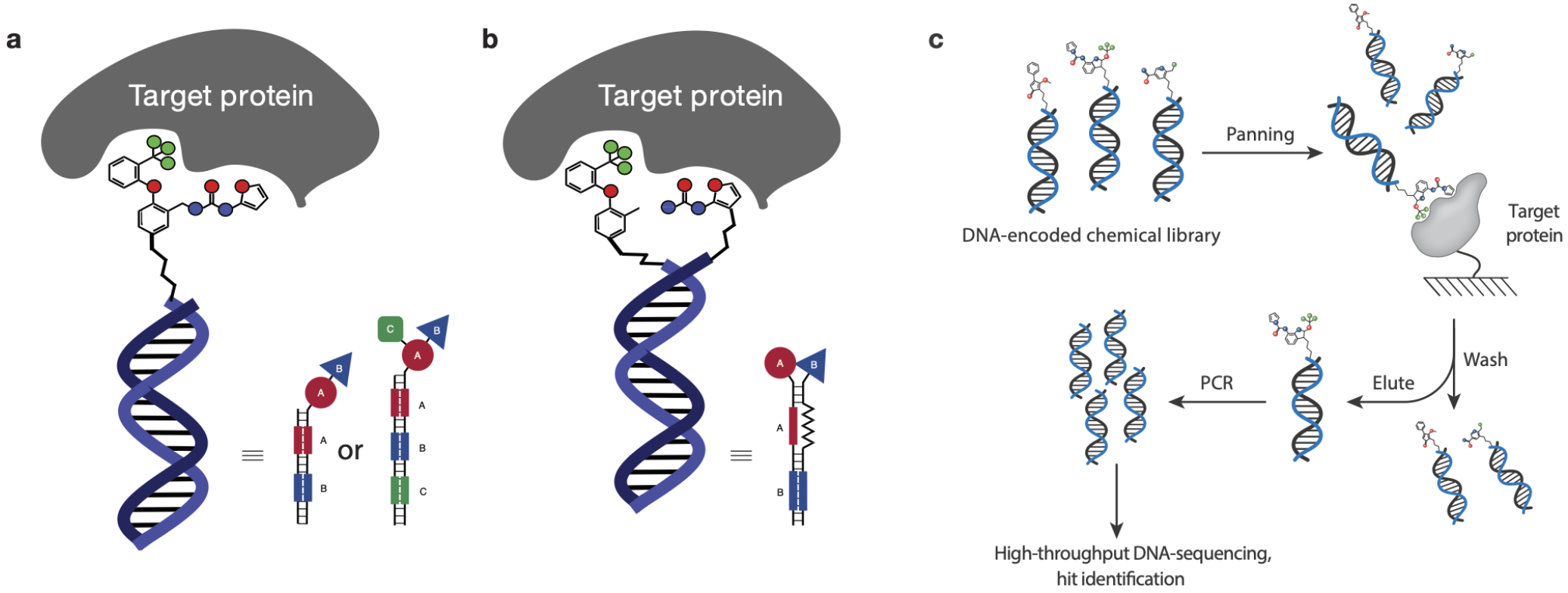






Satz et al., Nat Rev Meth Primers 2022

Screening with DNA-encoded libraries (DELs)



Source: ETH Zurich

Screening with DNA-encoded libraries

- Tagging of small molecules (chemicals) with DNA “bar tags” allows for miniaturization of high-throughput screens
- Millions of DNA-tagged chemicals in a single 1 mL tube
- Sequence the DNA to identify the “hit” chemicals
- Requires computational approach based on DNA sequences, not small molecule chemistry

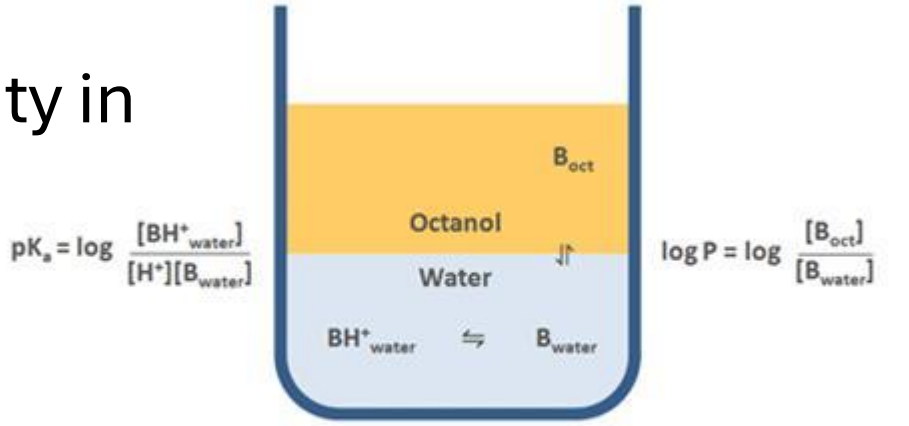
Advantages of DNA-encoded libraries for drug screening

1. enormous library size (trillions of compounds) → much better coverage of chemical space
2. small physical space for compound storage
3. miniaturization of assays
4. low-cost tools for academic institutions and small pharmaceutical companies
5. track record of success in eventually producing drug-like compounds

Pursuing challenging drug targets: antibiotic drug discovery

Chemical characteristics of drugs and compounds in chemical libraries

- Since human cells have membranes and hydrophilic regions, drugs need chemical characteristics of both fat solubility (lipophilicity=hydrophobicity) and water solubility (hydrophilicity)
- Partition coefficient: ratio of solubility in solvent vs water

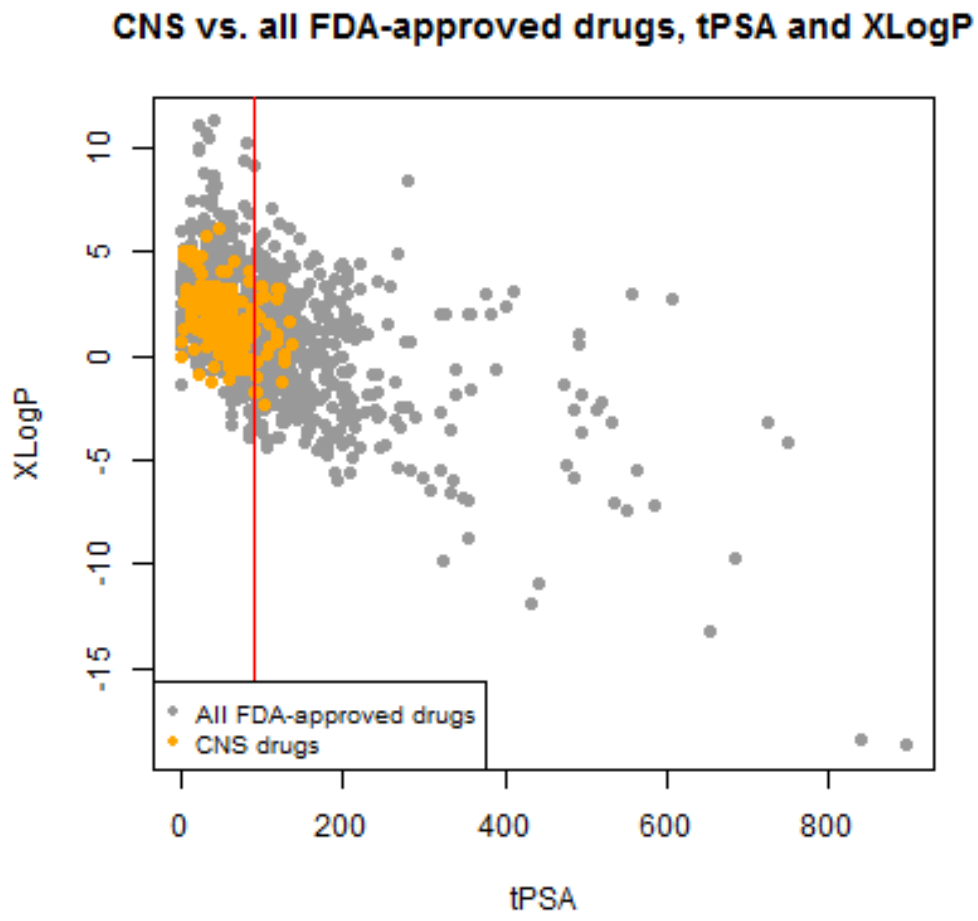


$$P = \text{Partition Coefficient} = \frac{\text{Concentration of neutral species dissolved in partition solvent}}{\text{Concentration of neutral species dissolved in water}}$$

Image: Enamine

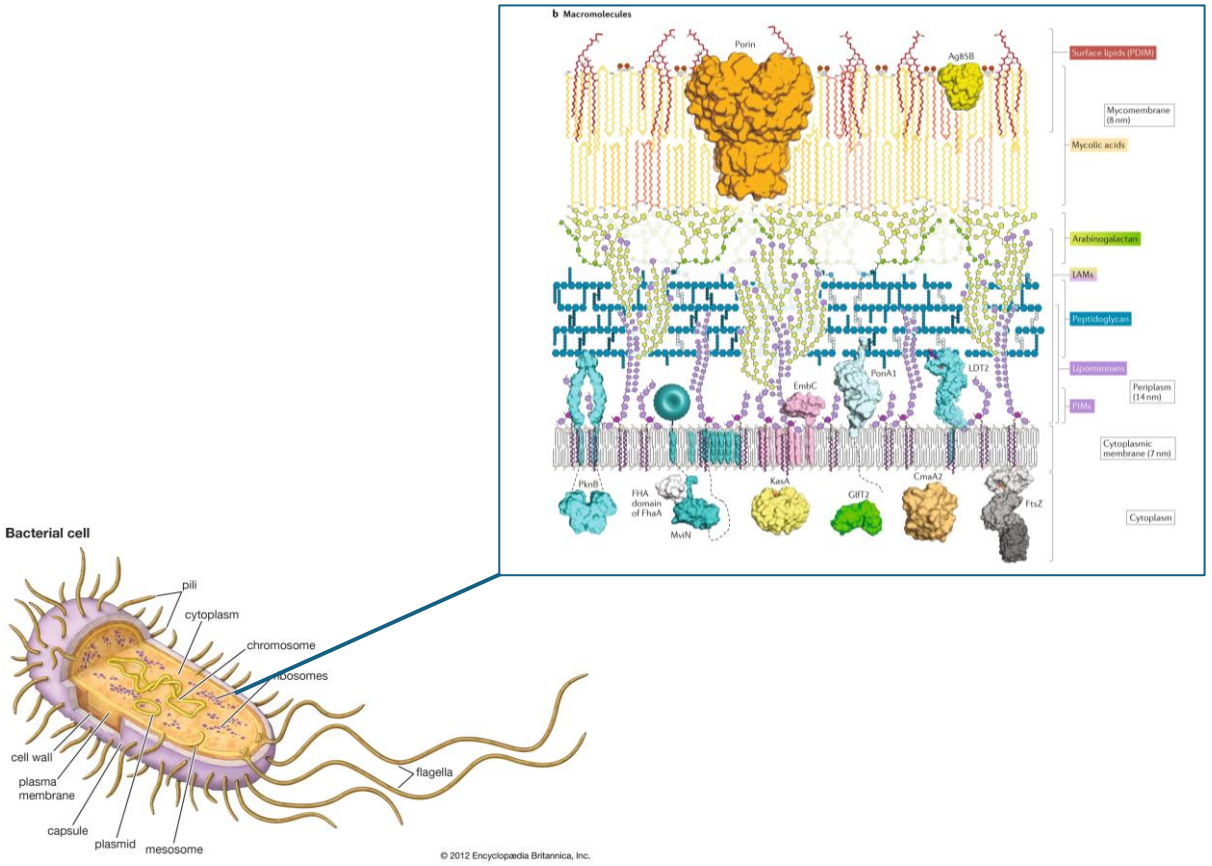
Chemical libraries – choosing an appropriate subsection of chemical space

- Are the chemical characteristics of your library relevant to your clinical goal?
- CNS drugs occupy a subset of all FDA-approved drugs
- CNS drugs tend to be:
  - Fat-soluble (higher logP)
  - Smaller (lower polar surface area)



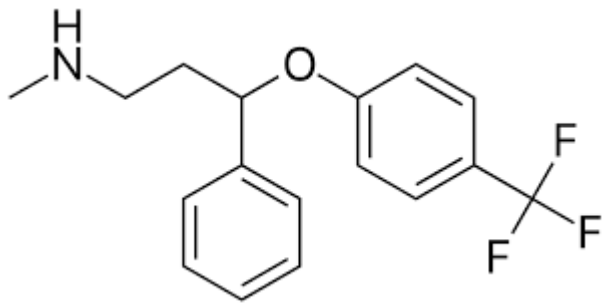
Chemical libraries for Antibiotic drug discovery: different region of chemical space

- Bacterial cell wall:
  - highly impermeable
- Relevant drug target may be inside the bacteria, so drug needs to cross the cell wall
- Tuberculosis (MTb)
  - Bacterial cell wall has the permeability of candle wax
  - Challenging for protein targets inside the bacteria

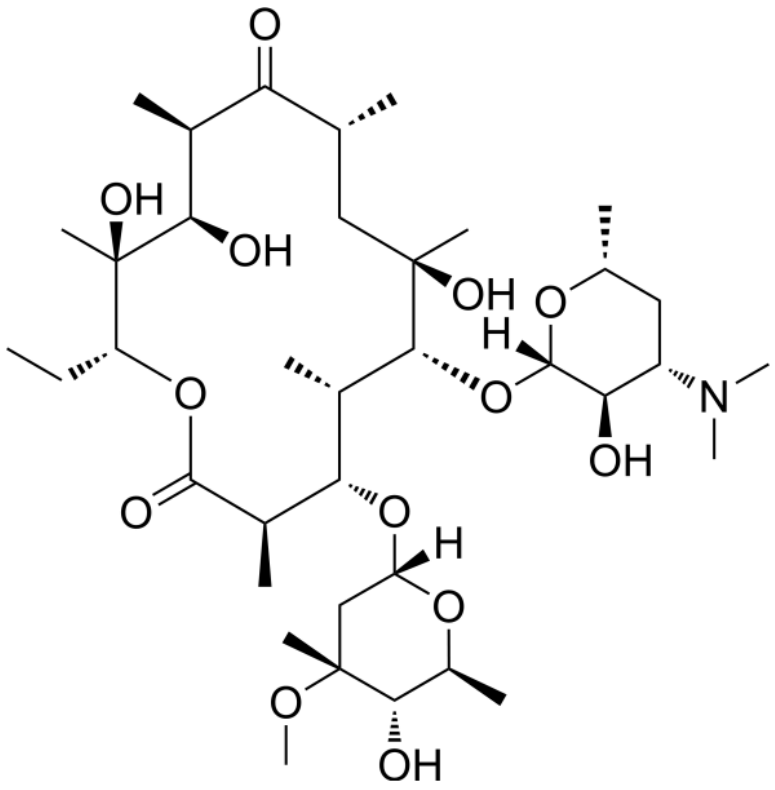


Dulberger et al., 2019 Nat Rev Microbiol

Representative chemical structures

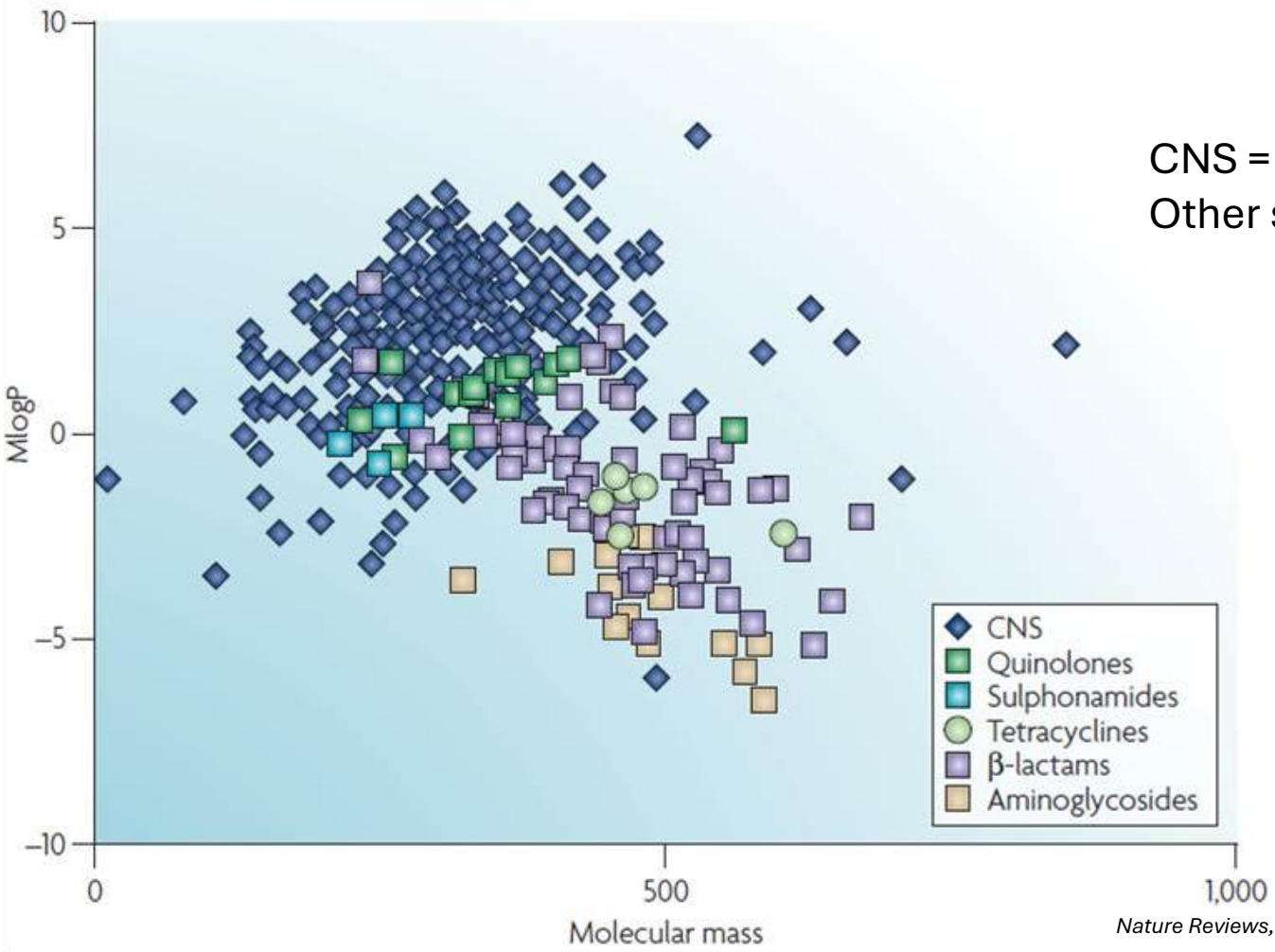


Antidepressant: fluoxetine



Antibiotic: erythromycin

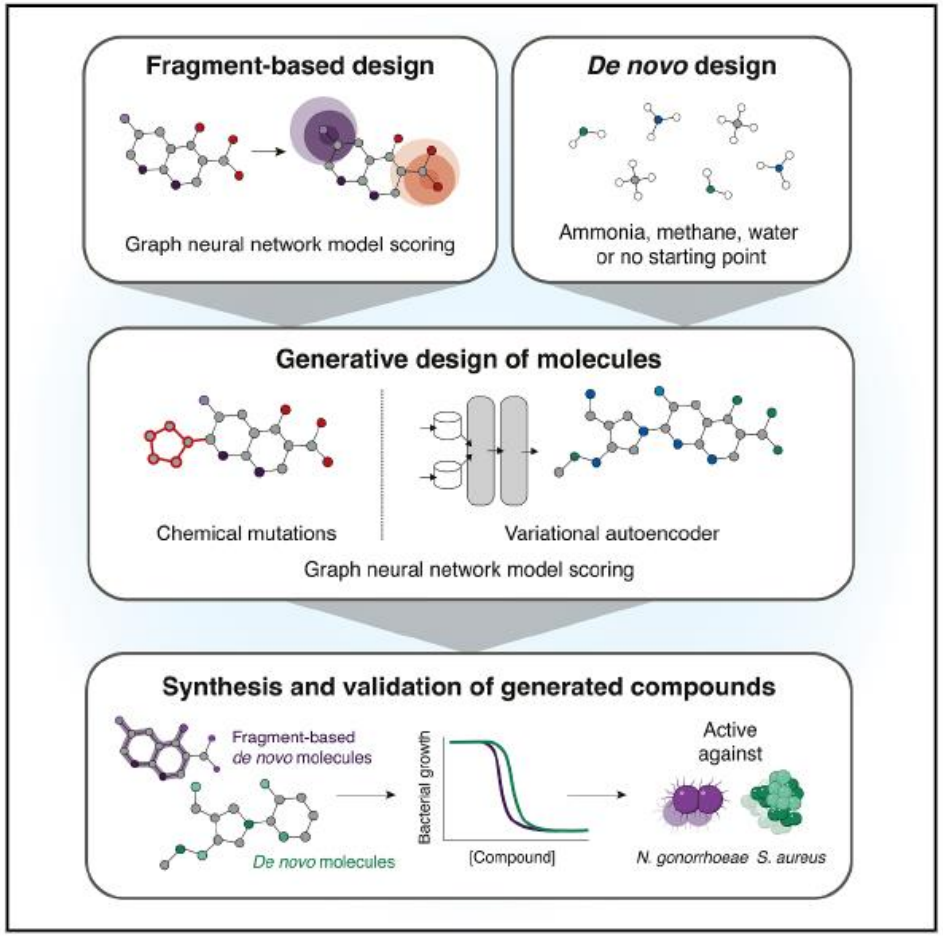
Depending on your goal, your compound likely inhabits a particular region of chemical space

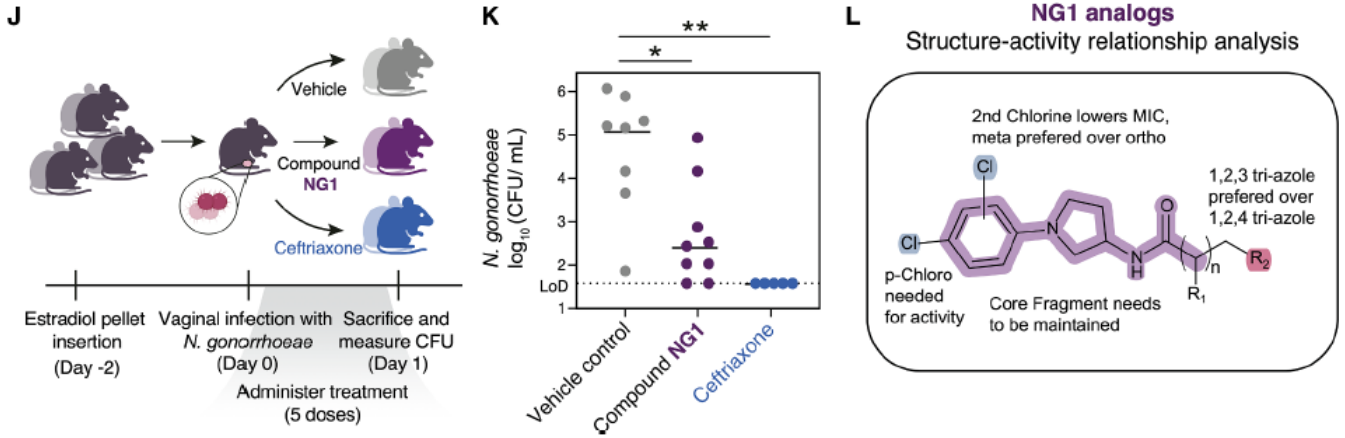


CNS = central nervous system
Other squares, circles = antibiotic drugs

*Nature Reviews, Drug Discovery (2013), **12**, 371-387. Platforms for antibiotic discovery*

Antibiotic drug design: generative deep learning

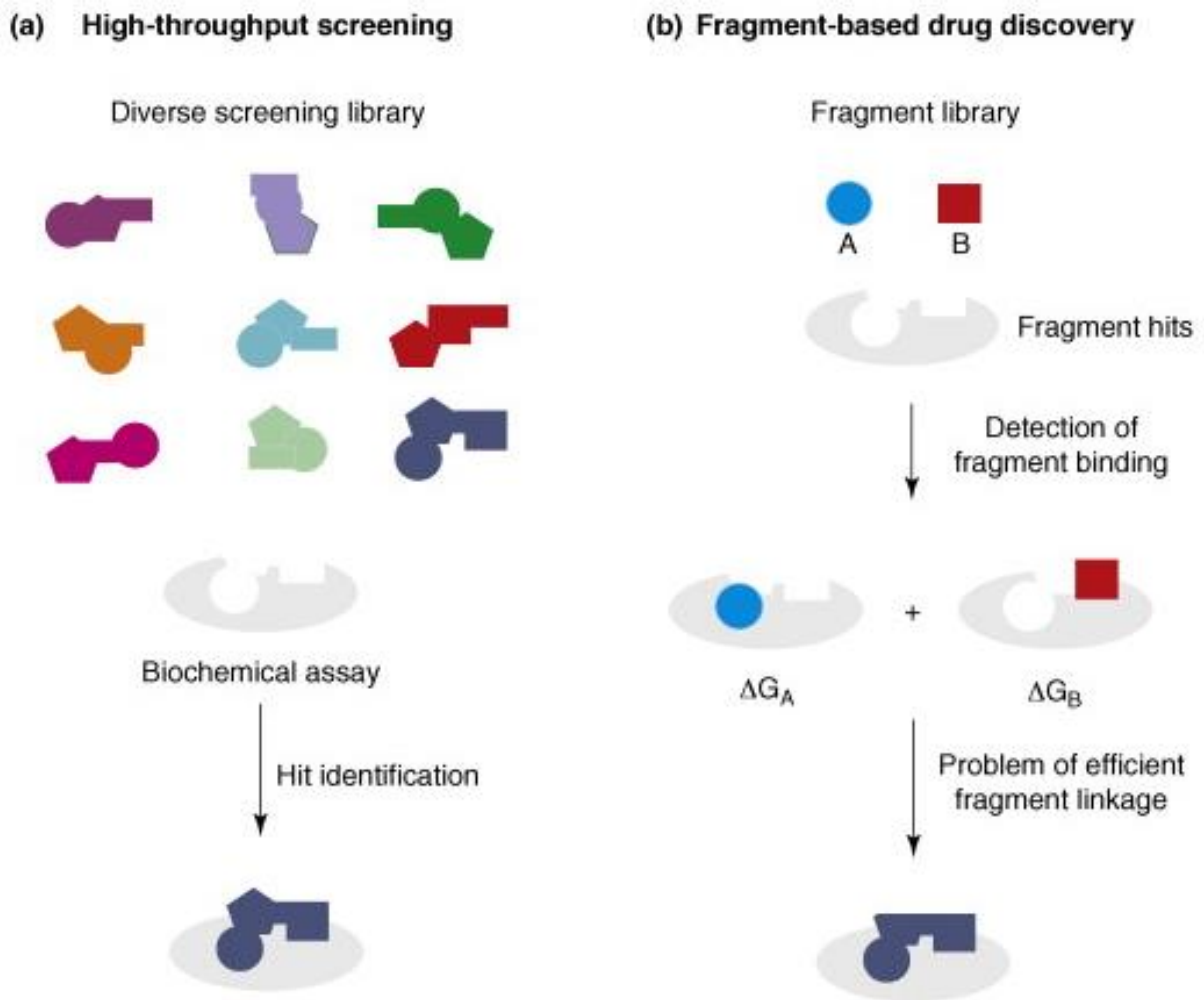




- Numerous research groups working on AI for antibiotics
- This particular paper leveraged a **fragment-based** approach
- Instead of “large” chemicals, this project worked with much smaller chemical fragments

Krishnan et al., Cell 2025

HTS vs Fragment-Based Drug Discovery



Schmit and Rademan Trends in Biotech 2009

Fragment-Based Drug Discovery

Disadvantages

- Fragment-sized chemicals are weaker binders due to their small size
  - Fewer possible interactions possible with protein target
- Fragment binding is often non-specific
  - Bind to off-targets as well as desired targets
  - Fragments have the potential for being useless if they are too non-specific (too small)
- Need a biophysical assay for weak binding (instead of traditional HTS assays)
- Need to “build out” or join the fragments
  - After identifying fragment, there are more steps before you arrive at a testable compound

Advantages of Fragment-Based Drug Discovery

- An option for challenging drug targets, especially when you don't have an appropriate chemical library for meaningful high-throughput screening (eg antibiotic drug discovery)
- FBDD libraries are smaller (fewer possibilities)
  - Estimated 166 billion possibilities for compounds with 17 heavy atoms
  - Compare to 10^{63} possibilities for "drug-like small molecules"
- Screening a fragment library covers a much greater fraction of the relevant chemical space

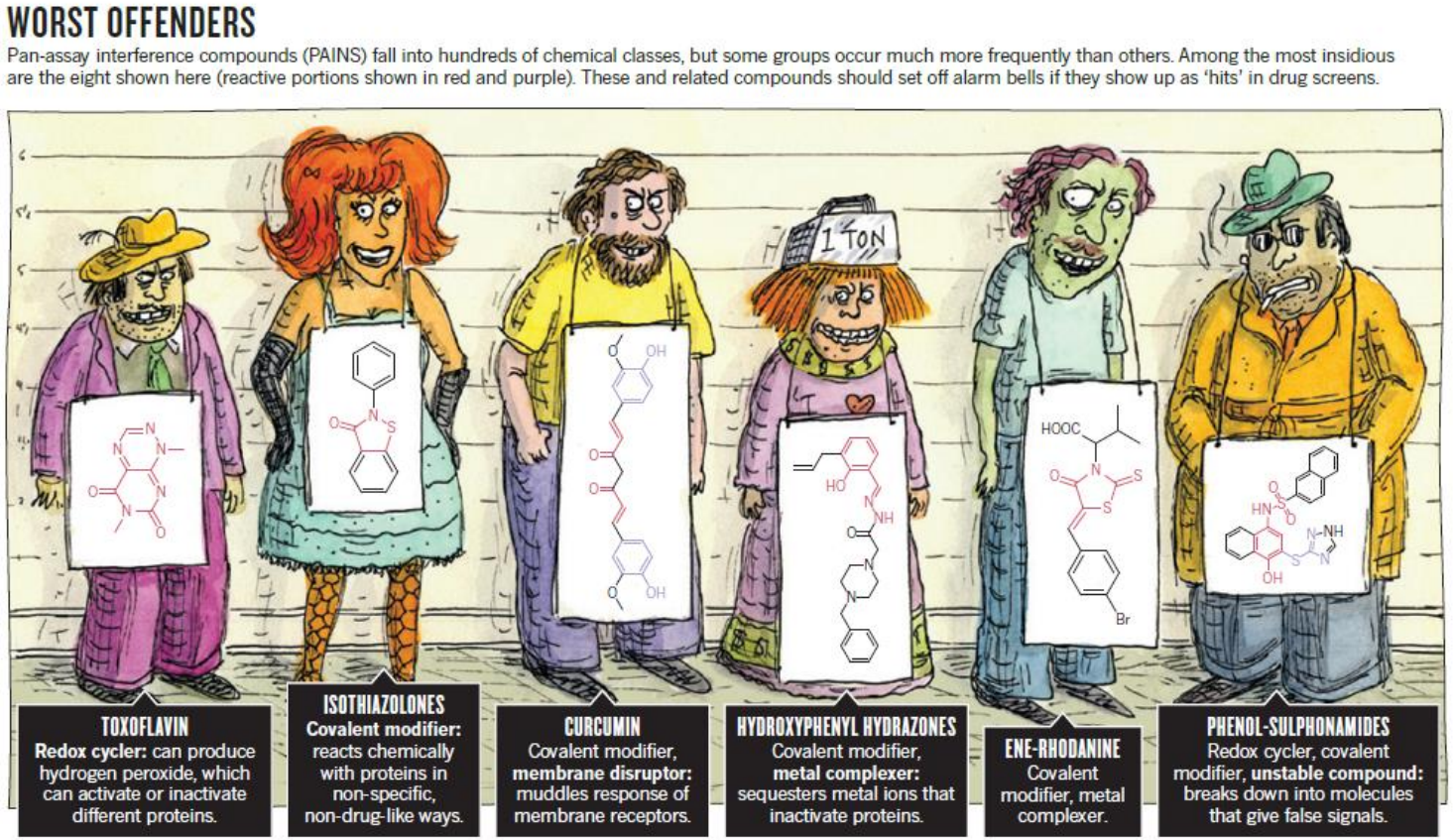
Advantages of Fragment-Based Drug Discovery

- Chances of **novelty** are much greater
  - Even from the same hit against the same target, no two companies are ever likely to end up with the same compound
  - Patents require intellectual property – require novelty
- Optimal properties can be built in at each stage of expansion
  - Reduced lipophilicity giving fewer PK/tox issues
- Selectivity can be built in from an early stage
- Potentially allows far greater chance of finding the rare compound which will hit difficult targets
- Avoidance of complexity reduces chances of unfavourable interactions
  - As complexity increases probability of matching a binding site actually decreases
- Commercially-available building blocks and availability of analogues can allow for rapid elucidation of SAR and hit expansion

Drug screening: Beware of artefacts

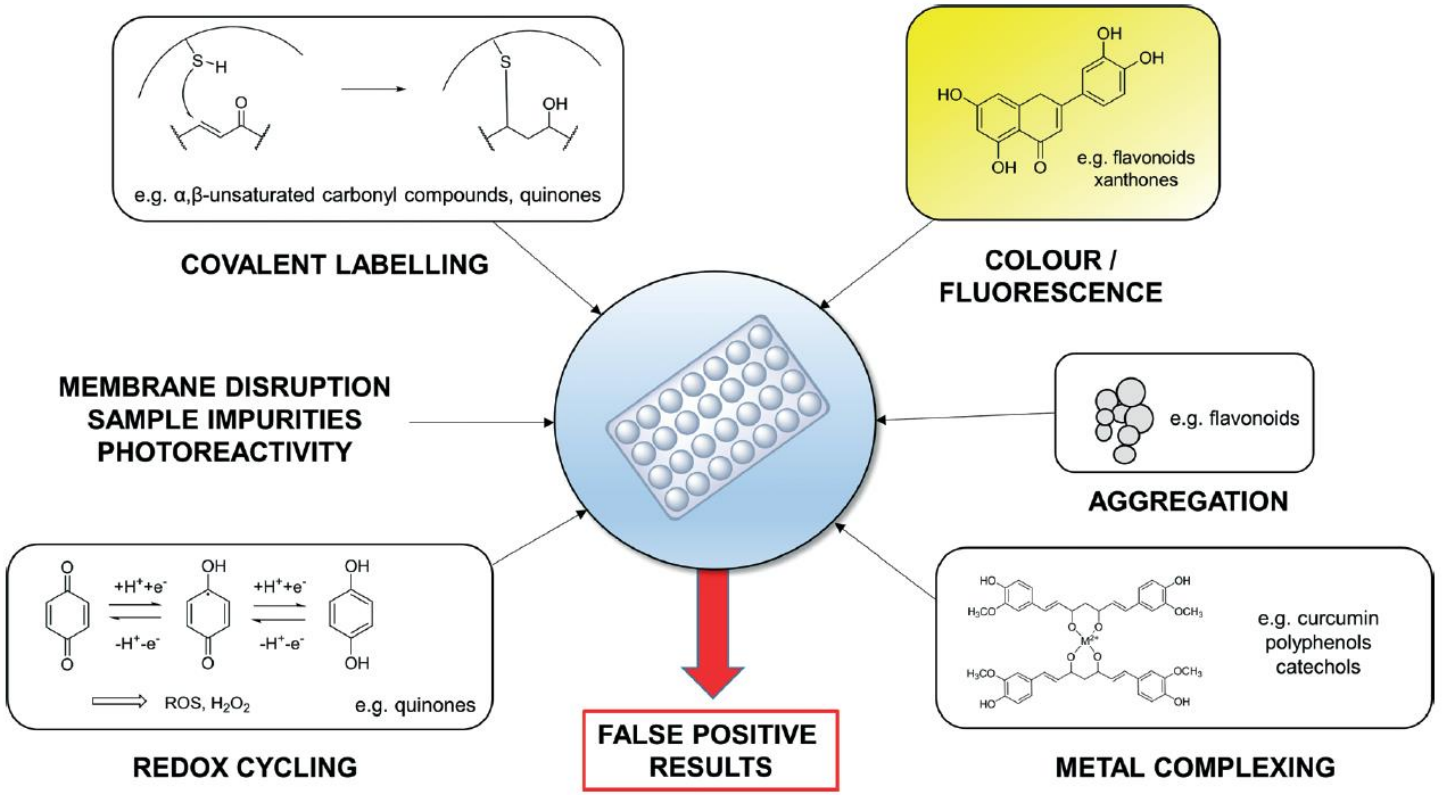
PAINS compounds:
false positive “hits” in numerous drug screens

- If you identify these chemical classes as “hits” in your screen and announce them as true hits → naive



Pan-Assay interference compounds – variety of mechanisms of non-specificity

- Covalent modifiers – a frequent mode of artefacts, others possible



Glazer and Holzgrabe Med Chem Commun 2016

How to deal with PAINS compounds?

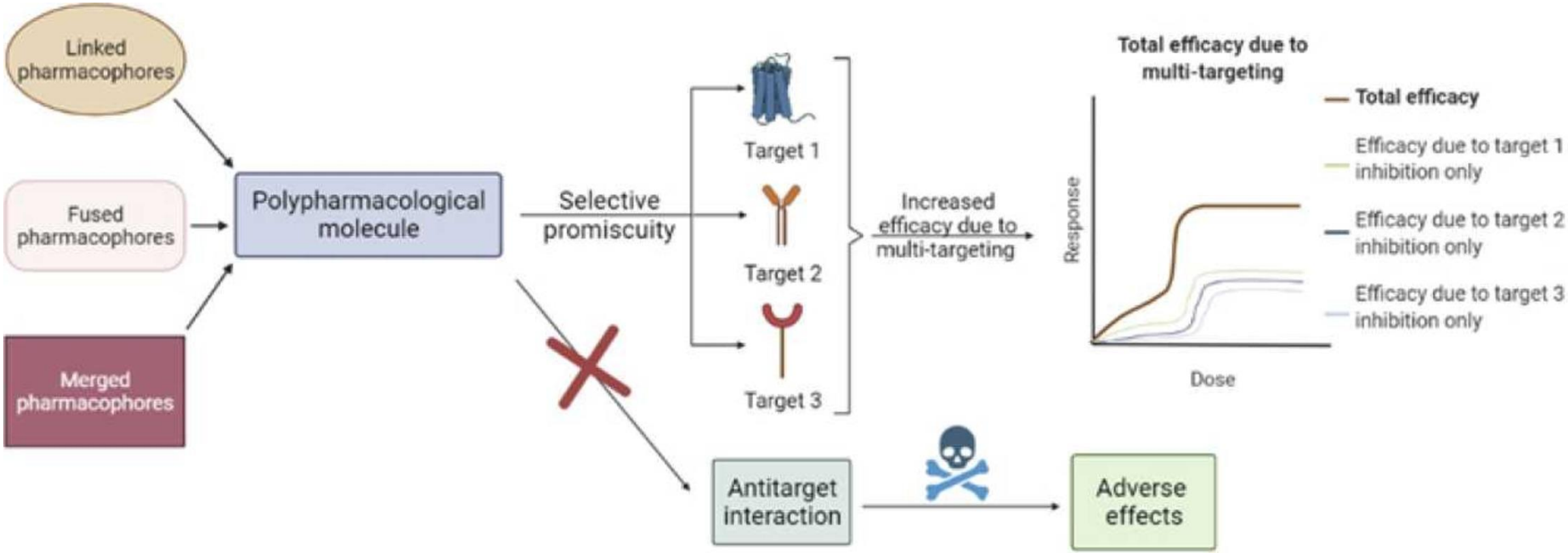
- Remove the PAINS “hits” from a screen data prior to moving forward.
  - May require cross-comparison from other screens since by definition these compounds show up in many different assays
- Don’t train your model on PAINS…
- Model may take problematic shortcuts when identifying “binders”
- Analogy: neural networks trained to find tanks – or clouds?

Beyond single drug targets

One drug → one target → clinical efficacy is an over-simplification

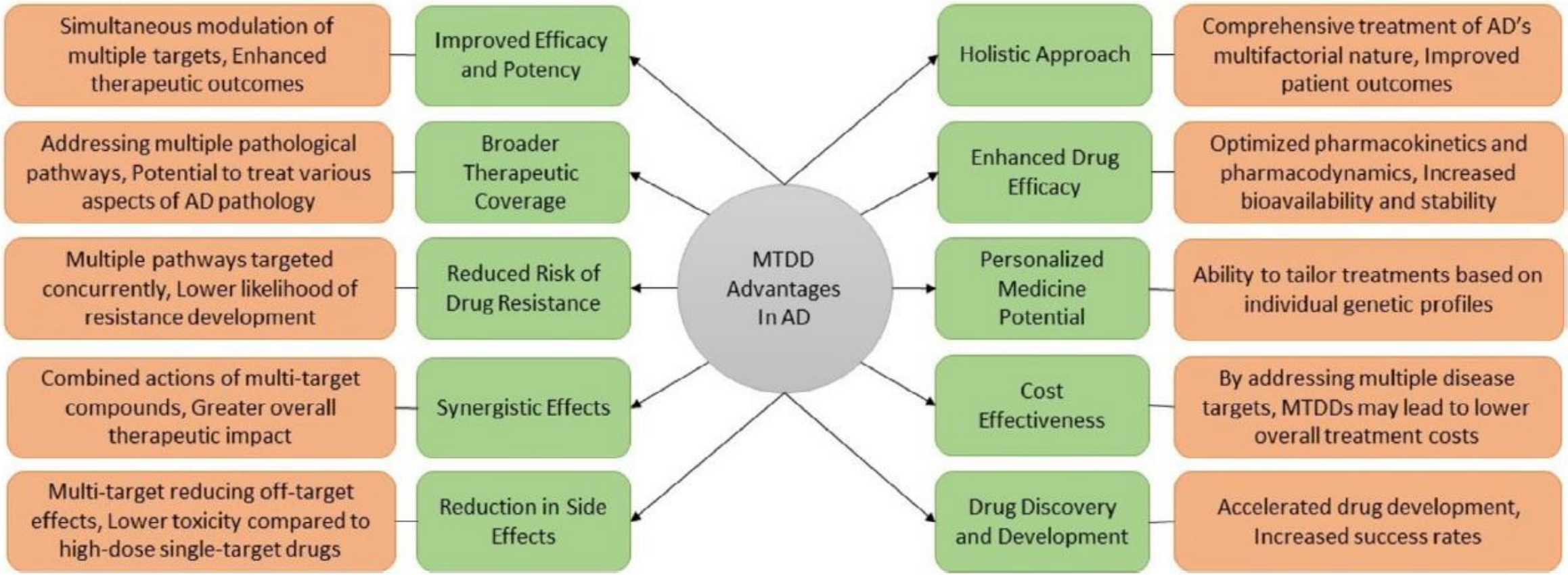
Drugs with multiple drug targets

- Clinical utility and efficacy



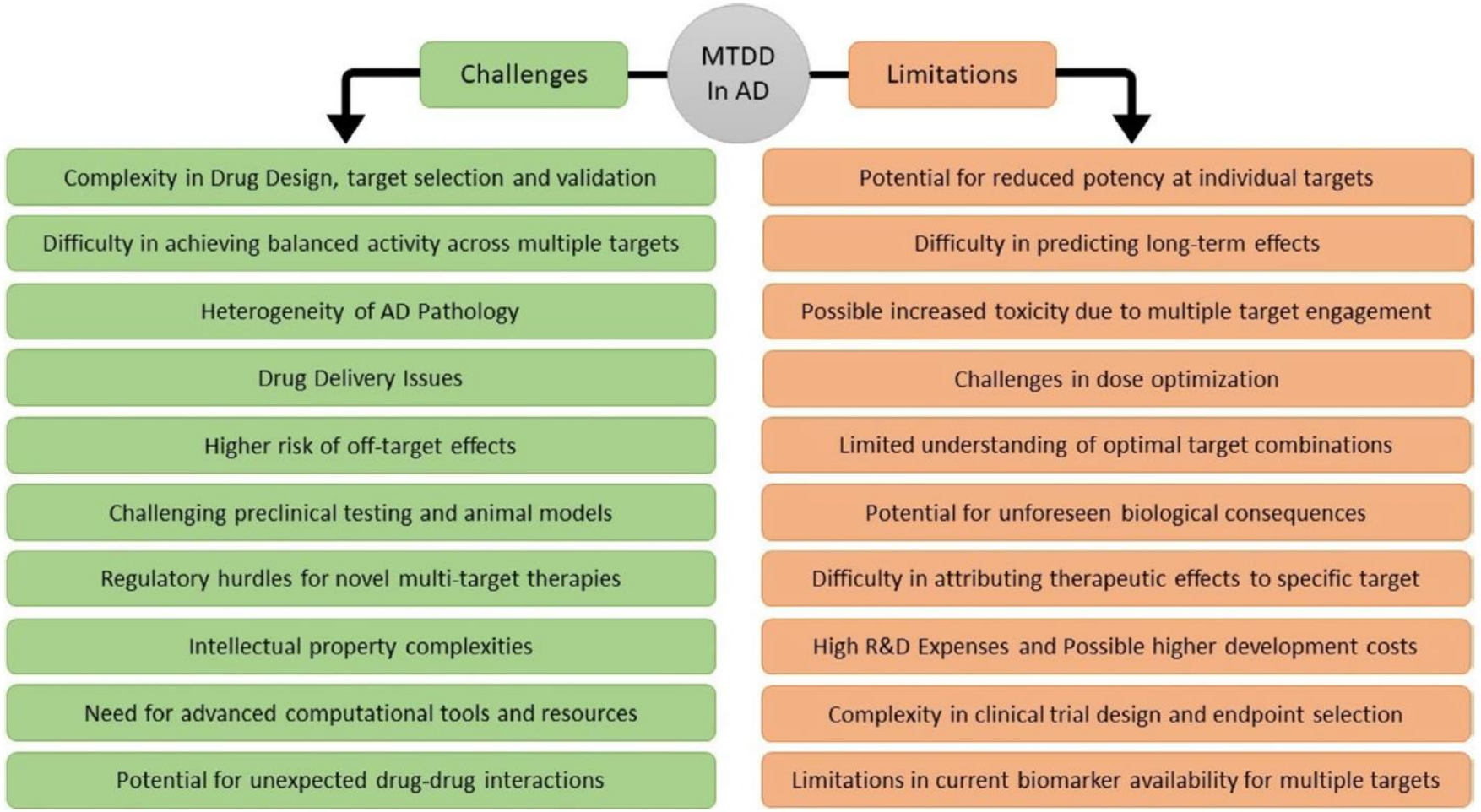
Kabir and Muth, Pharmacol Res 2022

Advantages of Multi-target drug development in Alzheimer's Disease



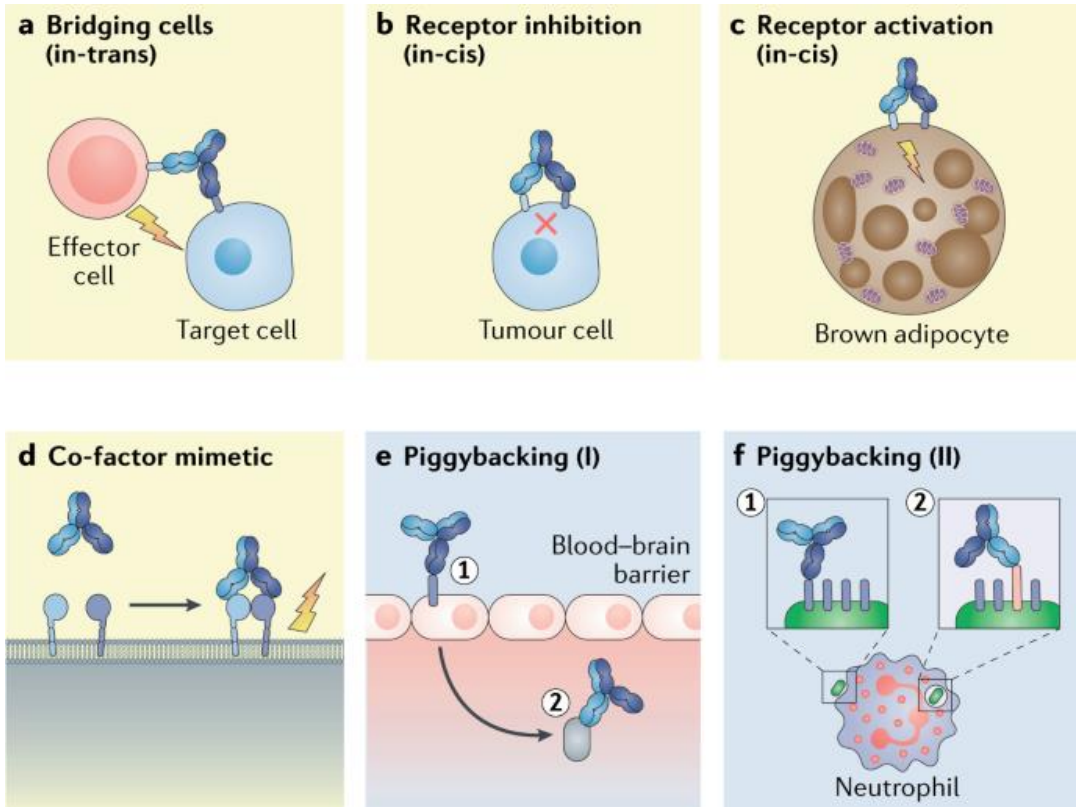
Hossain and Hussain Pharmacol Res Perspect 2025

Challenges and Limitations of Multi-target drug development in Alzheimer's Disease



Hossain and Hussain Pharmacol Res Perspect 2025

Bispecific antibody drugs: two intended targets



Labrjin 2019 Nat Rev Drug Discovery

Takeaways

- High-performance AI/ML models should address both orthosteric and allosteric interactions
- There is an opportunity for subject matter-informed AI/ML in the antibiotic space. Interrogate appropriate regions of chemical space
- Strategies to improve coverage of chemical space:
  - DNA encoded libraries
  - fragment-based drug discovery or DNA-encoded libraries
- Beware of Pan-assay interference compounds. Don’t be naïve and call these “hits”; remove them from your models
- Beyond single-target drugs. Clinical efficacy of multi-target drug development (intentionally targeting > 1 protein)