

Introduction to In silico Methods

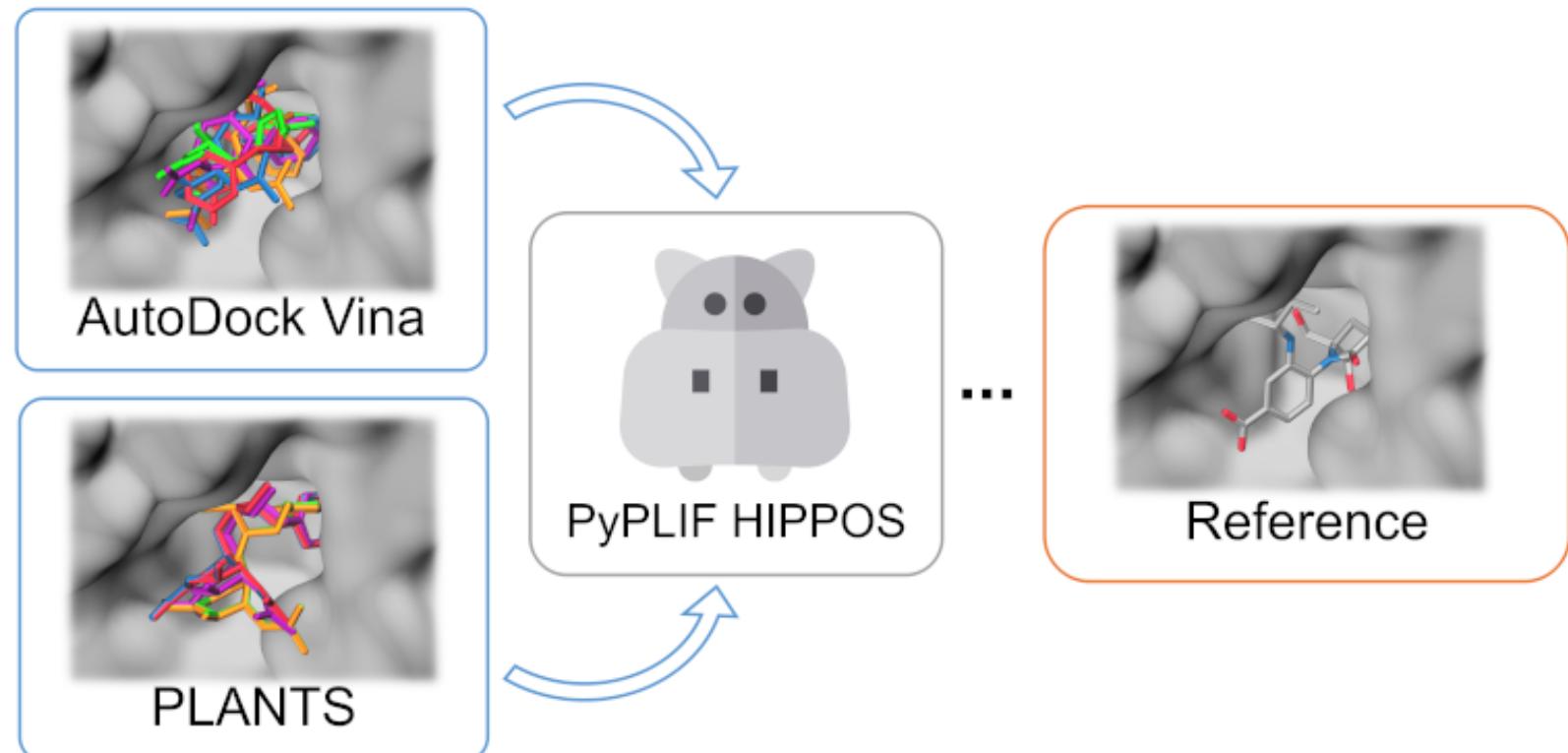
Muhammad Radifar

Guest Lecture @ Teknik Biomedis - Institut Teknologi Telkom Purwokerto

January 9, 2023



About Me



- 1 Pharmacy UGM 2005-2012
- 2 BME UGM 2012-2016
- 3 Lecturer @STIKES Gunabangsa Yogyakarta 2016-2019
- 4 Freelance Software Engineer
- 5 Co-author of PyPLIF & HIPPOS

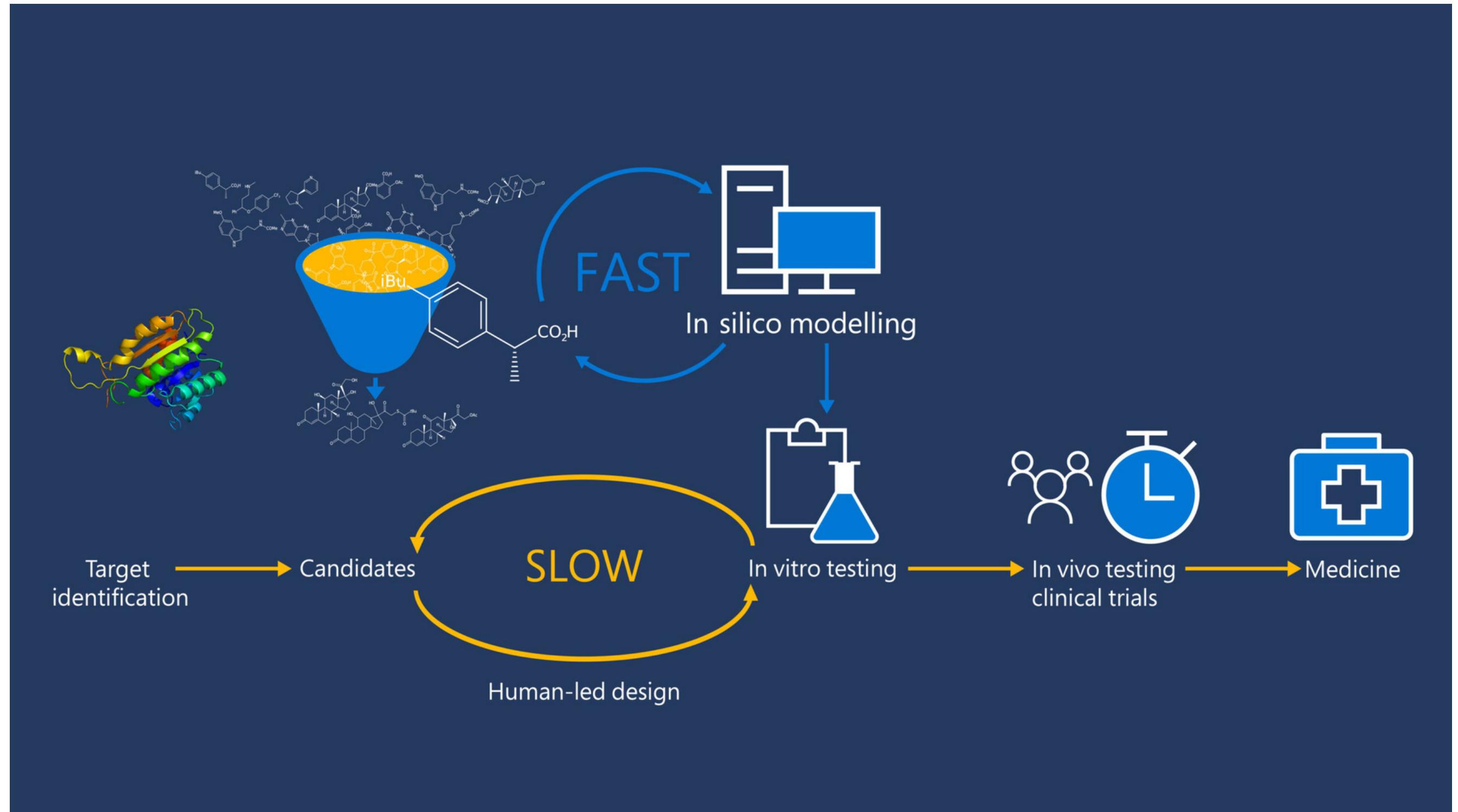
Pose #	1000101	0.500	0.365
1	0001010	0.050	-0.447
2	1000011	0.105	0.105
3	0001011	0.381	0.221
4	1110100	0.261	-0.084

Today you will learn

- 1 What is in silico in general
- 2 All you need to know about model
- 3 How drug is developed
- 4 How in silico method help drug development

github.com/radifar/open-lecture

In vivo - In vitro - In silico



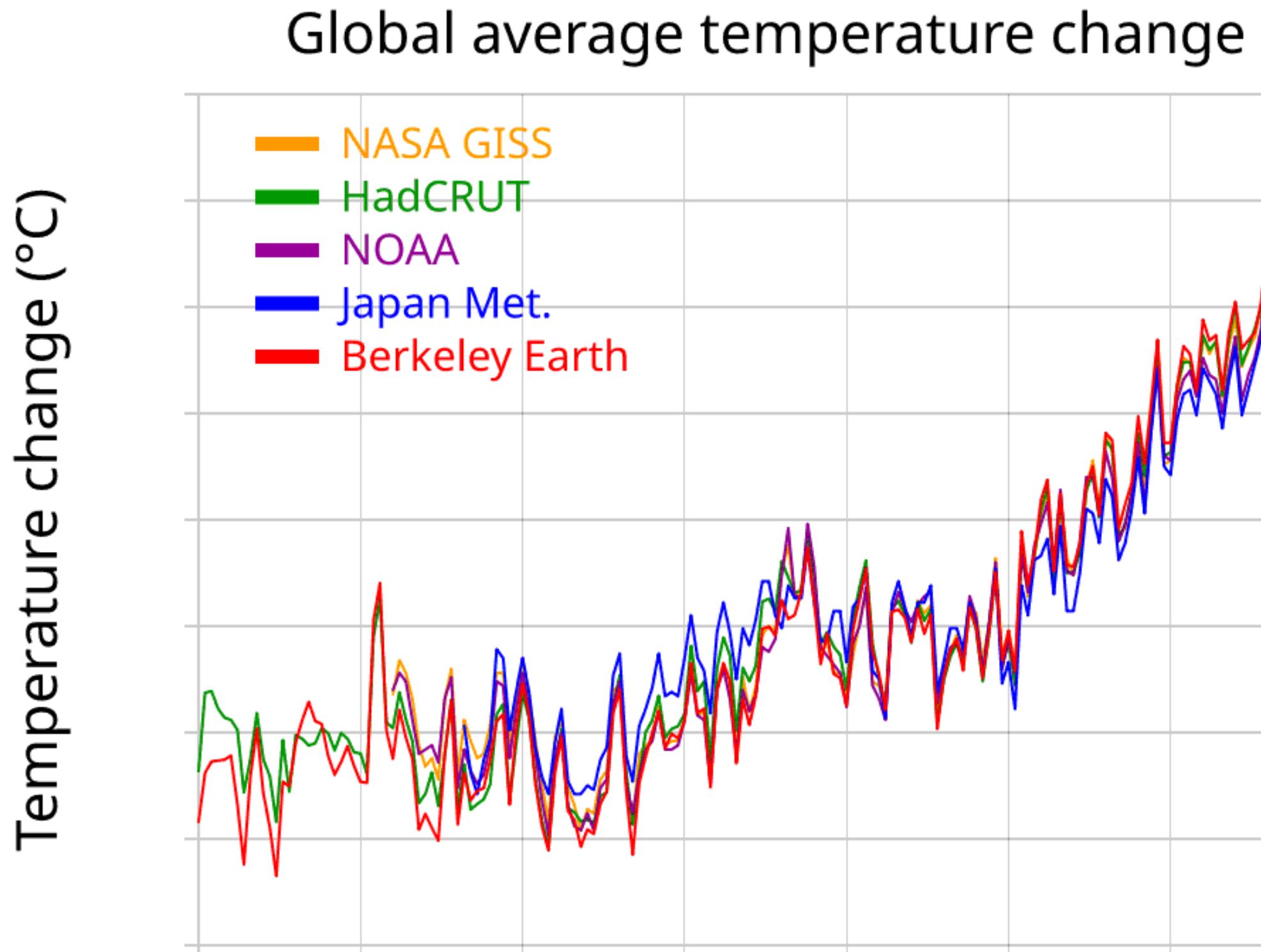
What is In silico?

In silico is a method that utilize **calculation** or/and **model** using **chips** to **simulate** or **predict** a phenomenon, or just simply to gain **better understanding** of a phenomenon.



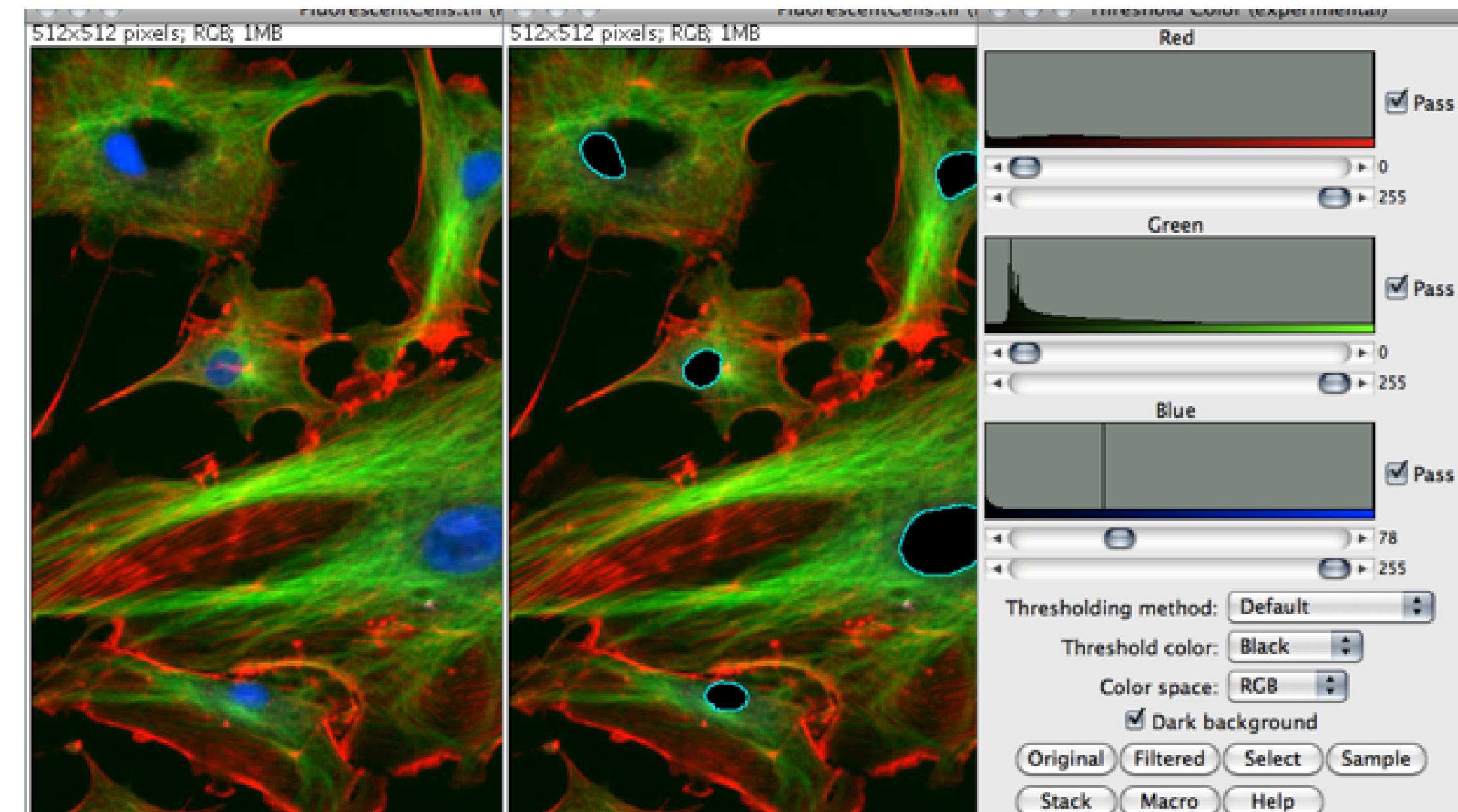
What is In silico?

Even something as simple as generating graph or chart to gain better understanding or insight towards your data



What is In silico?

Or... something as simple as
image editing and measurement



OK, I know that In silico uses calculation or/and model. I know what calculation is, but...

What is a Model?



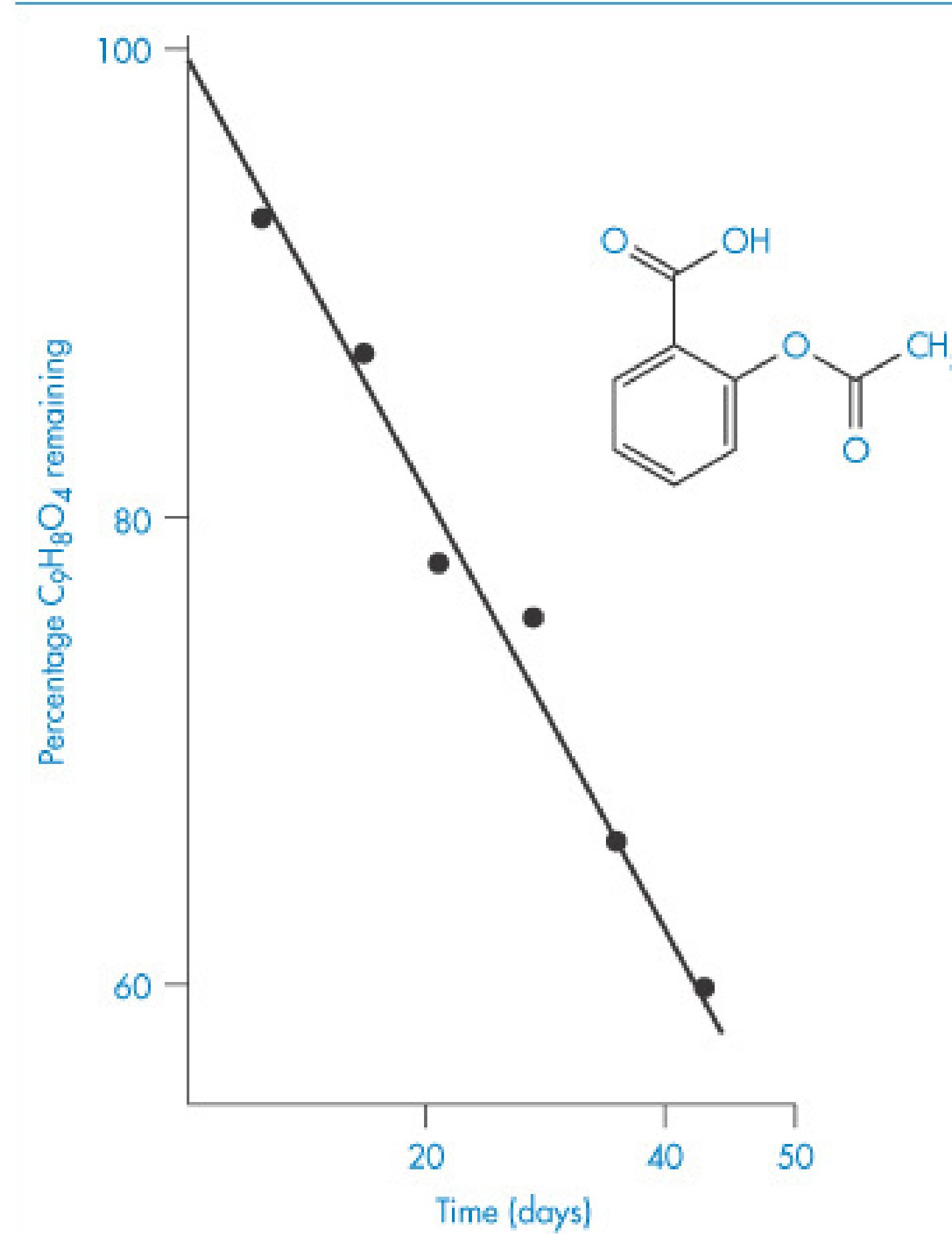


This Model?

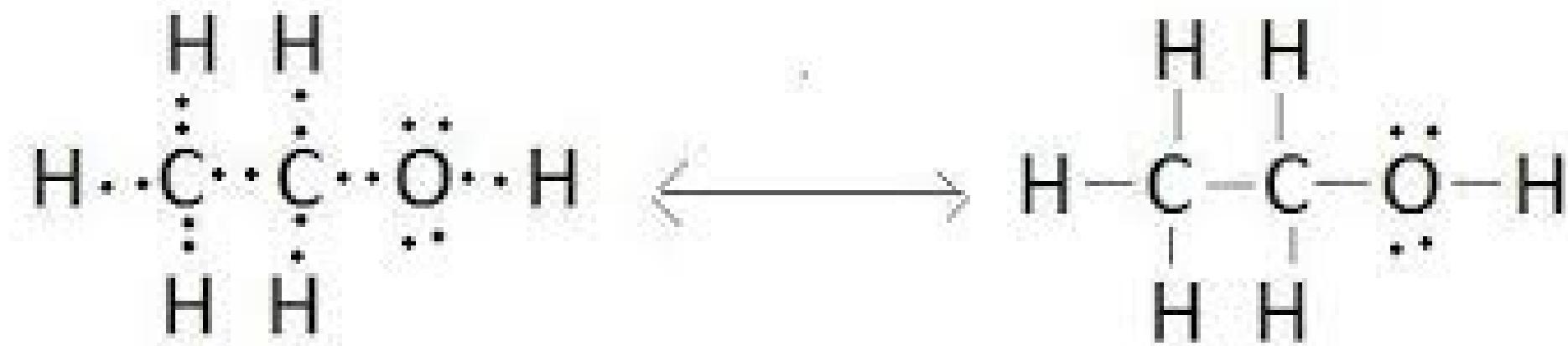
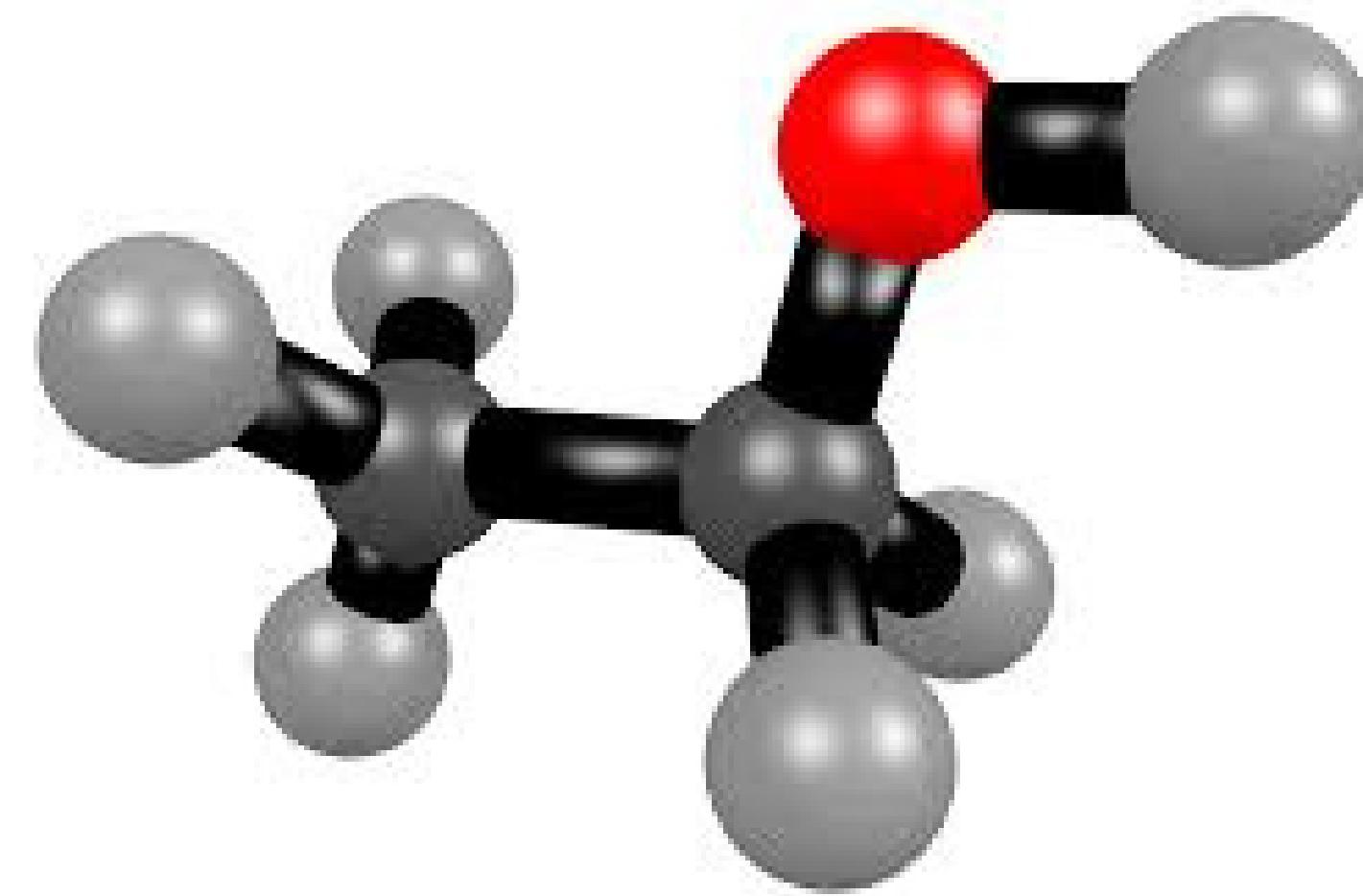
What is a Model?

Model is a simpler representation of reality.

It can even be as simple as simple linear or logarithmic equation.



Model
example

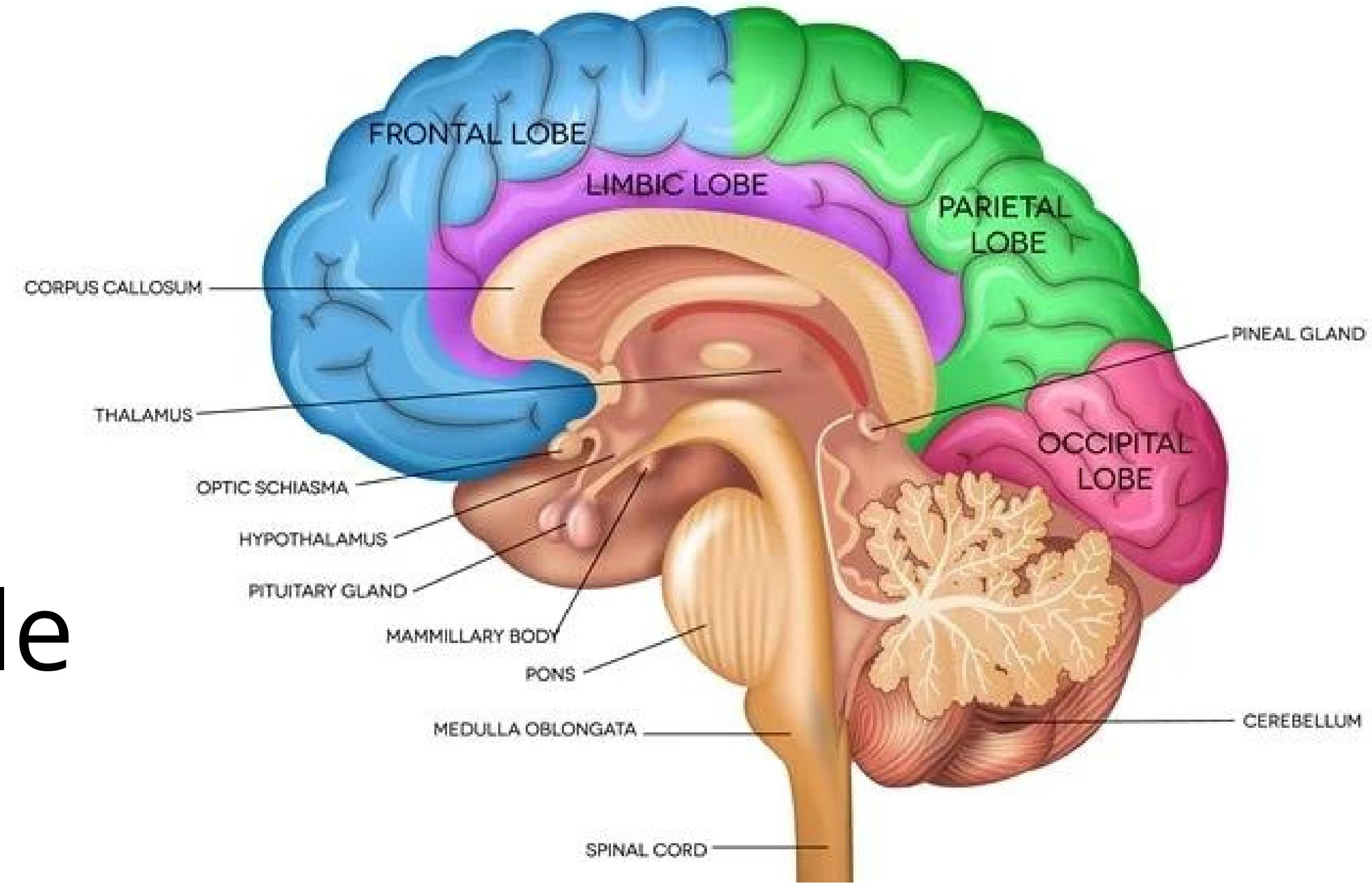


Model example



ANATOMY OF THE BRAIN

Model
example



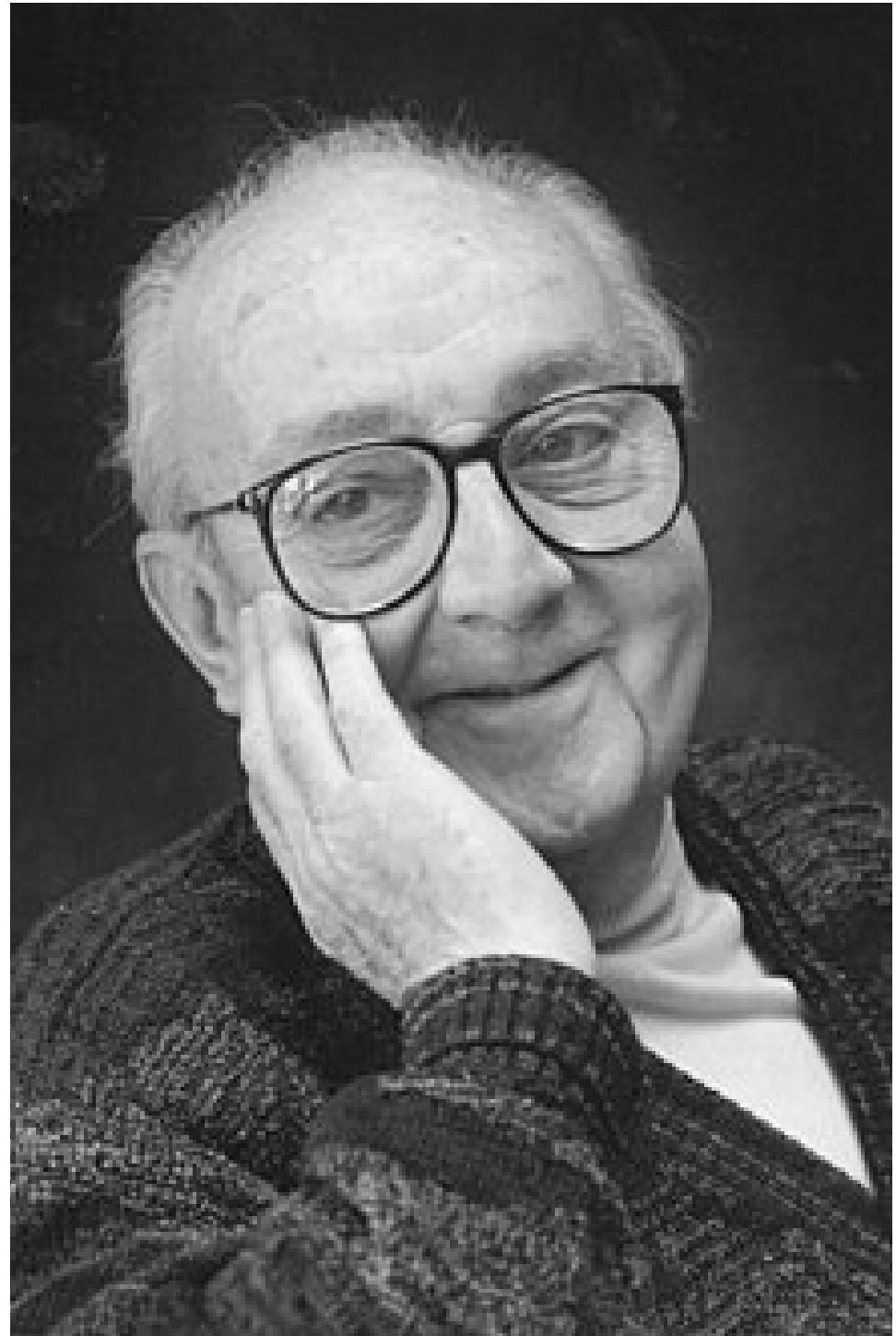
The World Map

Model
example

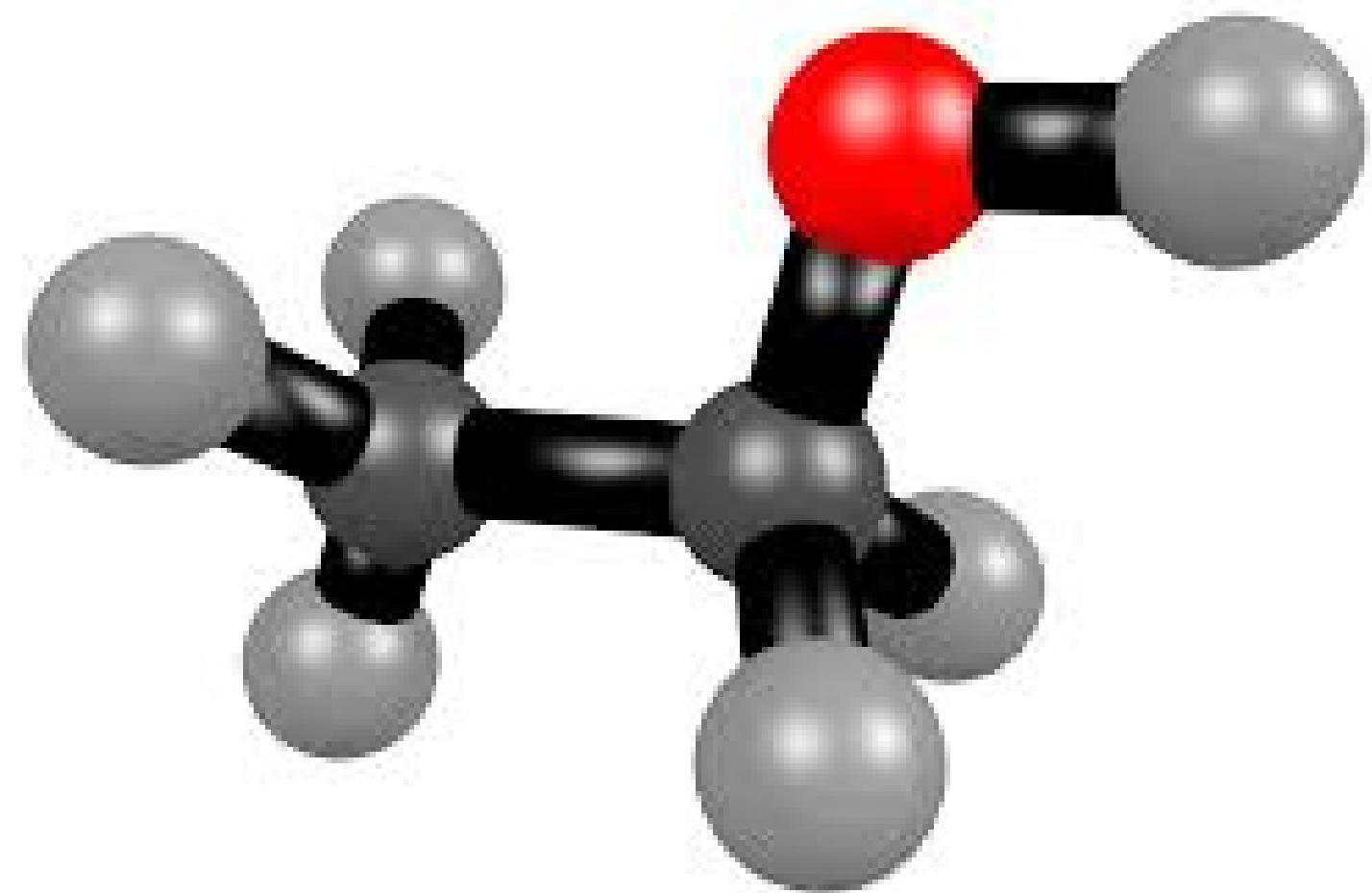


All models are wrong but some are useful

George E.P. Box



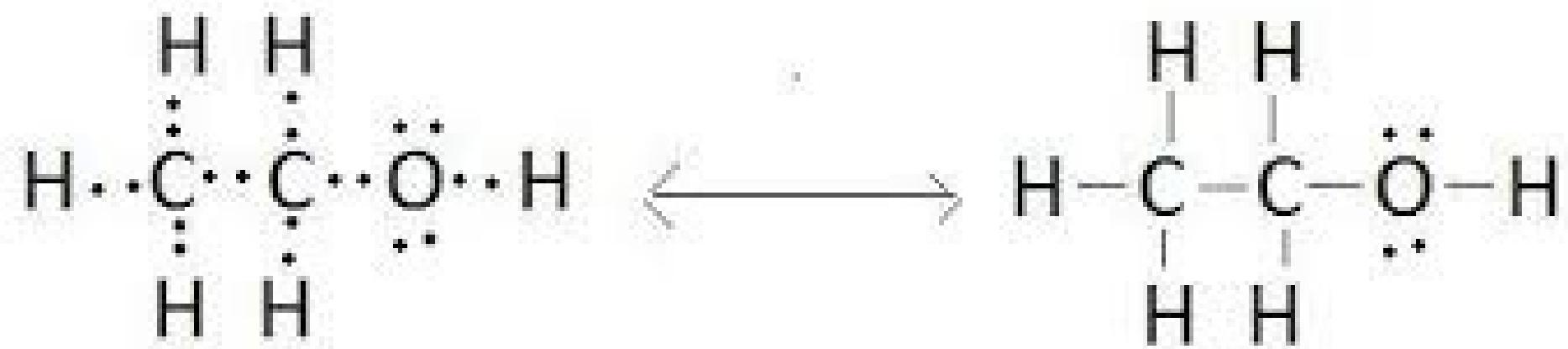
https://en.wikipedia.org/wiki/All_models_are_wrong



Which part is wrong?

Why is it useful?

Model example



Model example

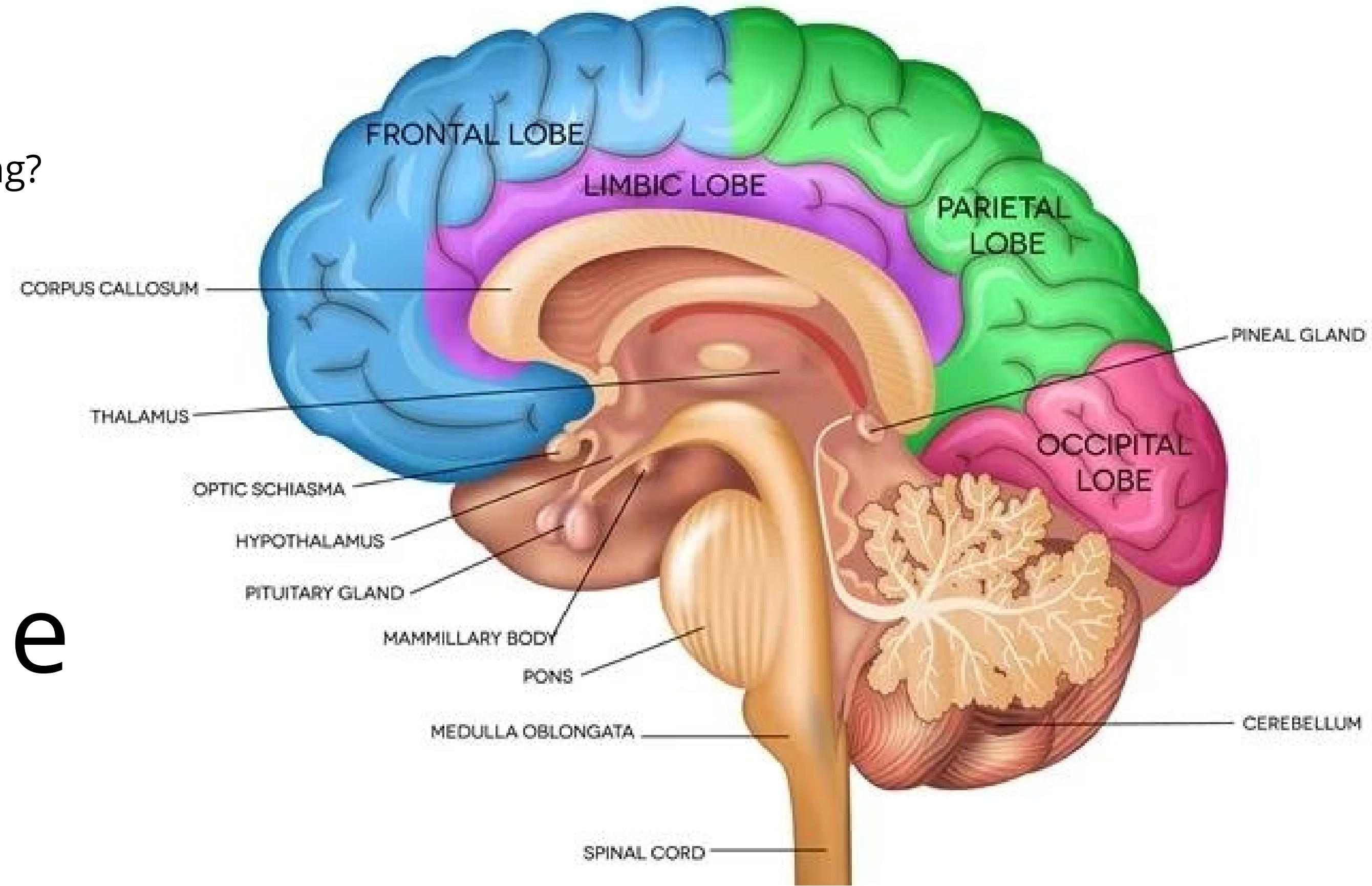
Which part is wrong?
Why is it useful?



ANATOMY OF THE BRAIN

Model example

Which part is wrong?
Why is it useful?



The World Map

Which part is wrong?
Why is it useful?

Model example



Therefore ...

In silico (theoretical) result is no replacement for experimental result. And always need to be verified with experiment.



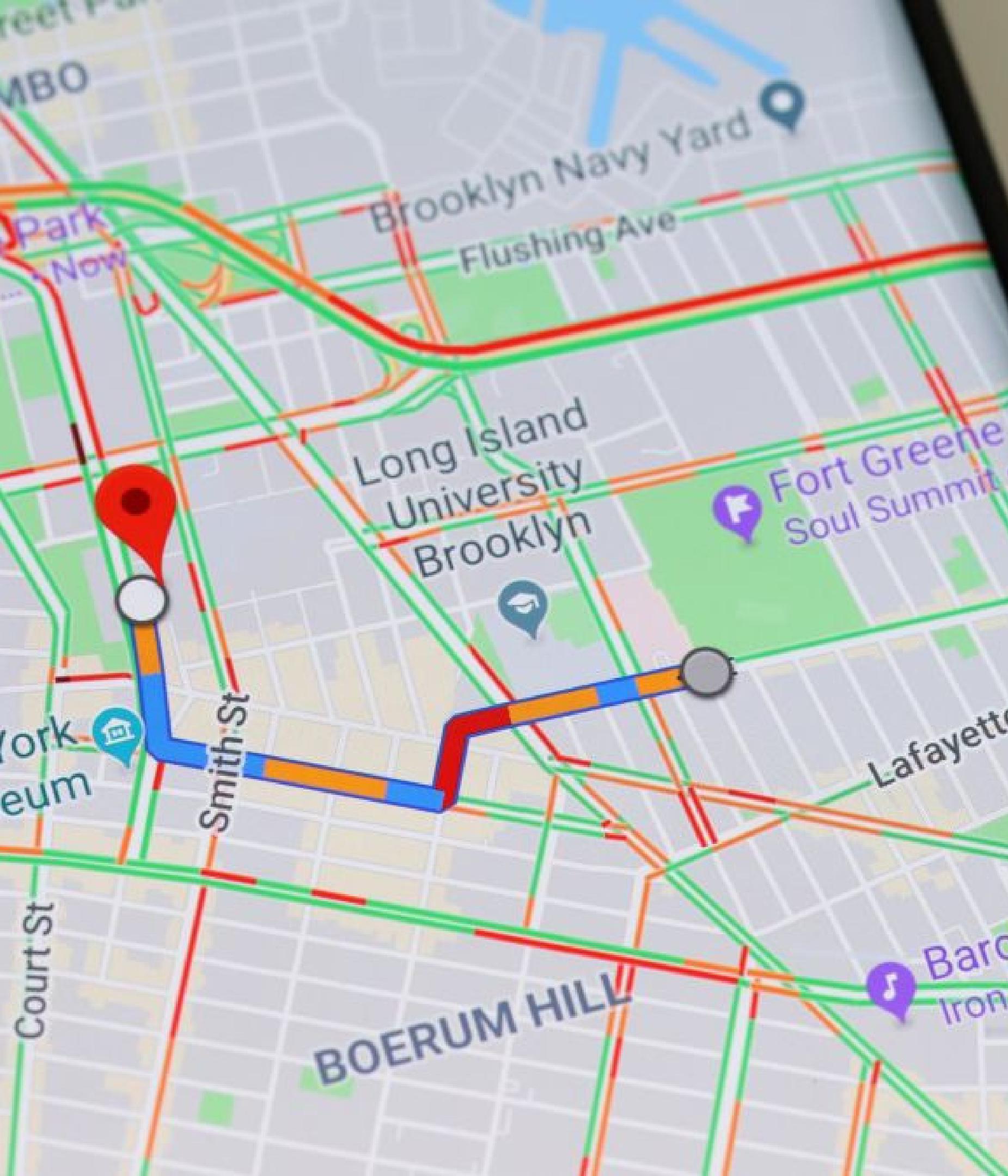
Model is everywhere

Starting from weather forecast



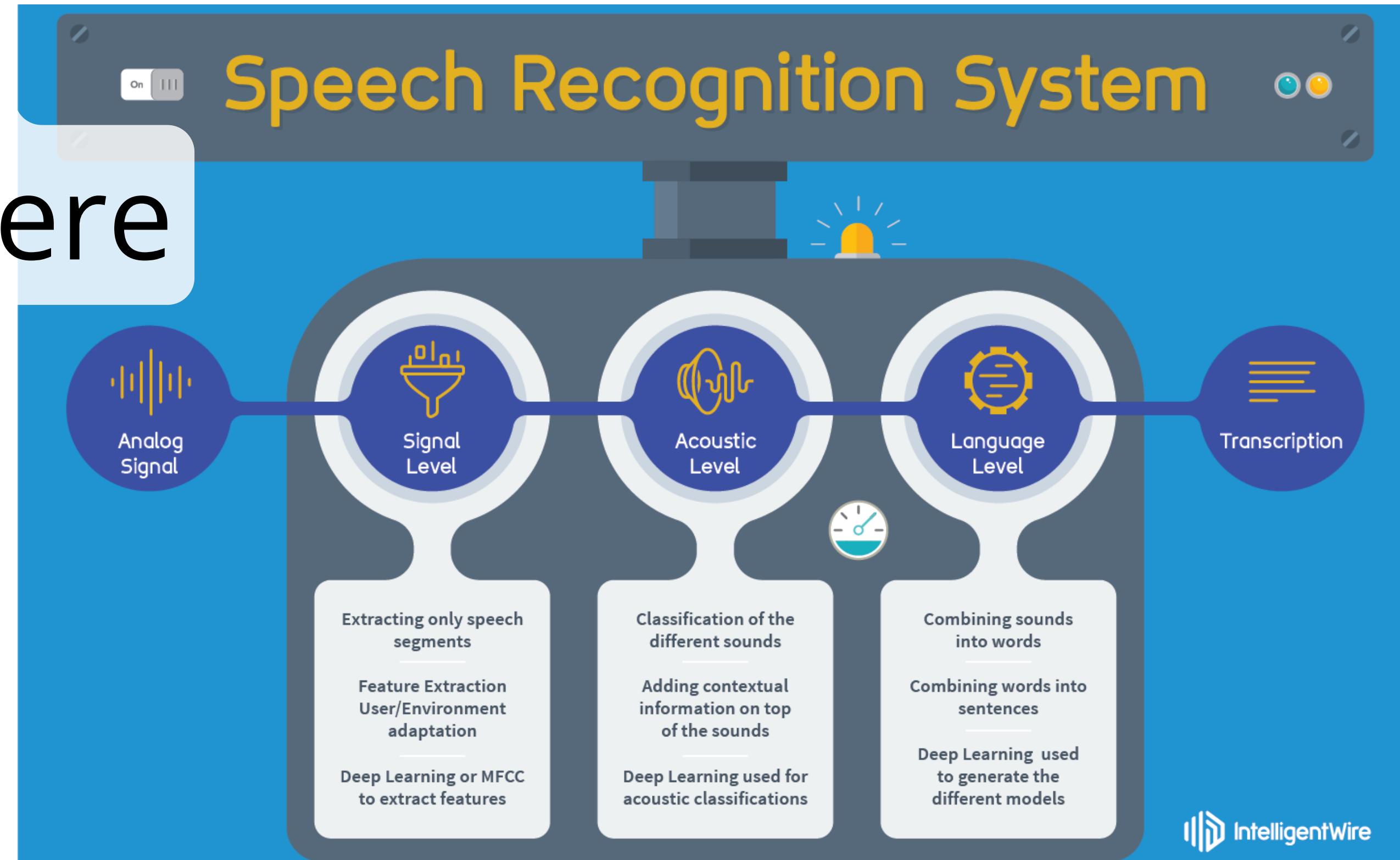
Model is everywhere

Google Maps...



Model is everywhere

Speech recognition...



Model is everywhere

Expiration date



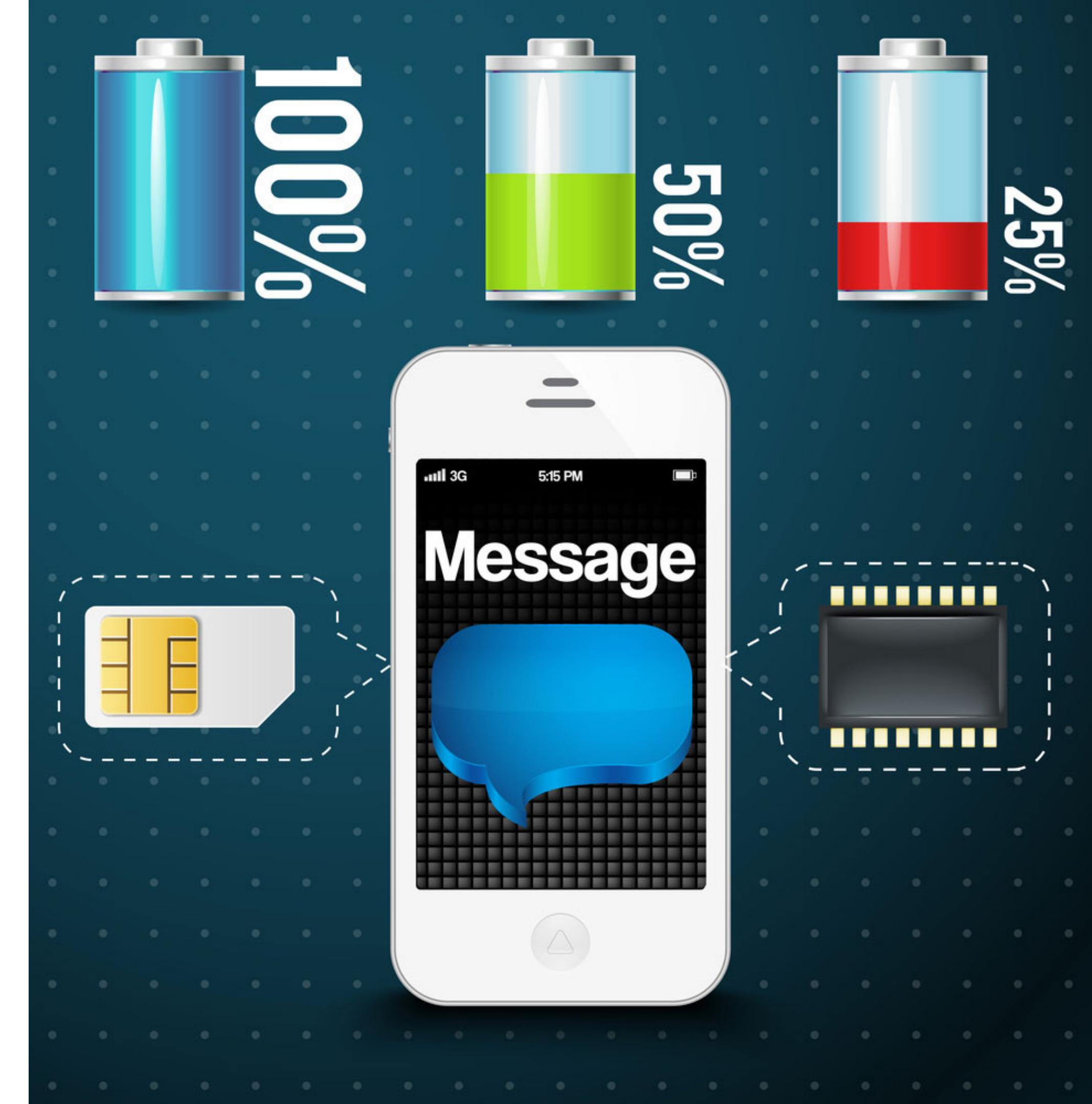
Model is everywhere

Racing game,
first-person shooter game



Model is everywhere

Battery indicator is a model too



Model is everywhere

Battery indicator is a model too

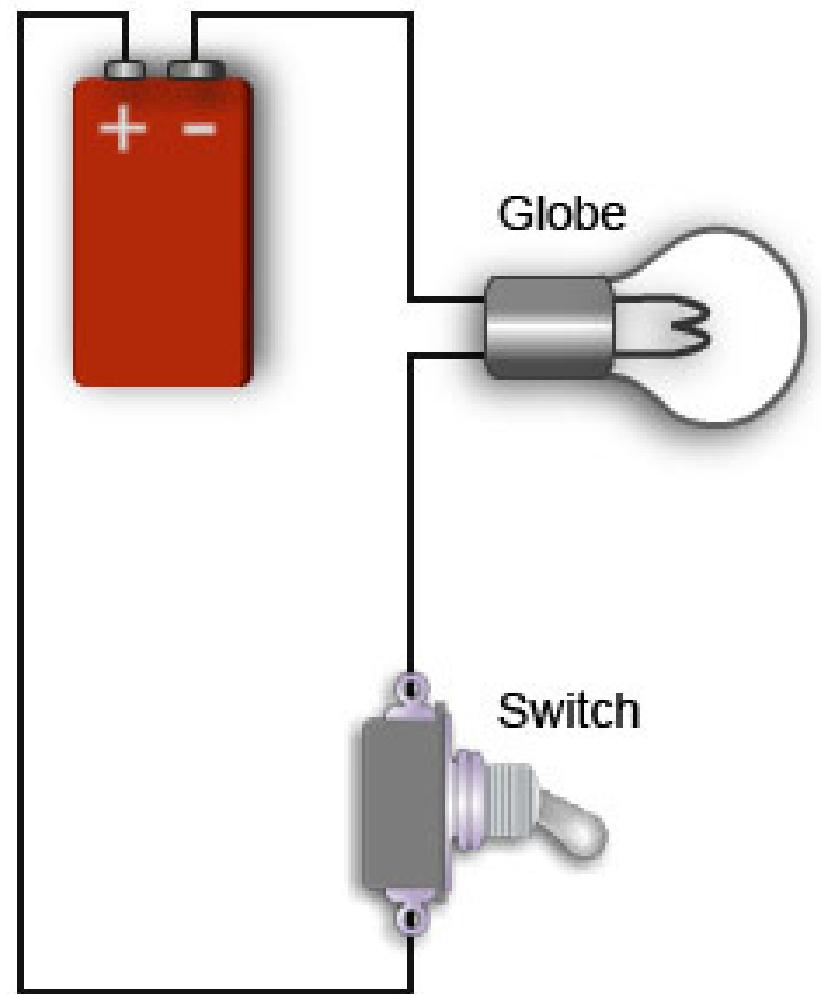


Figure 1: Electrical

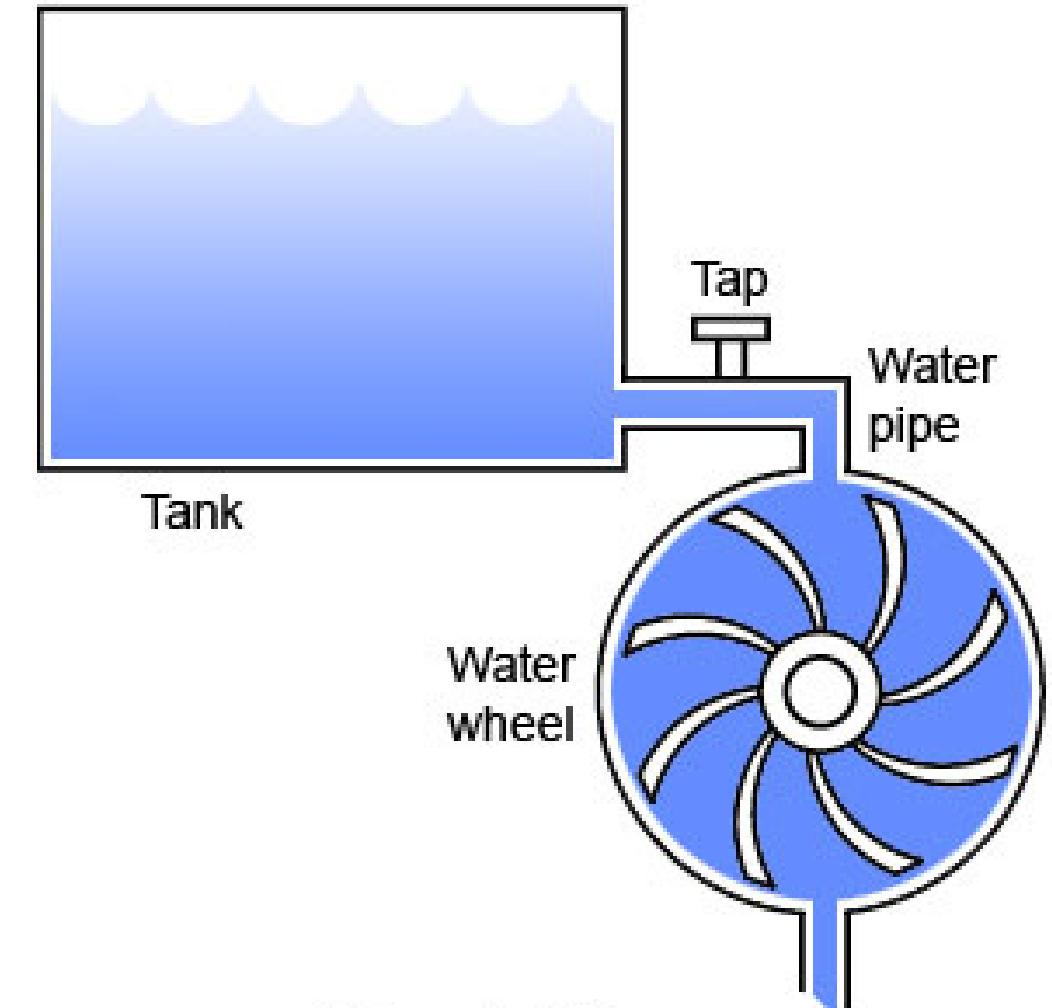


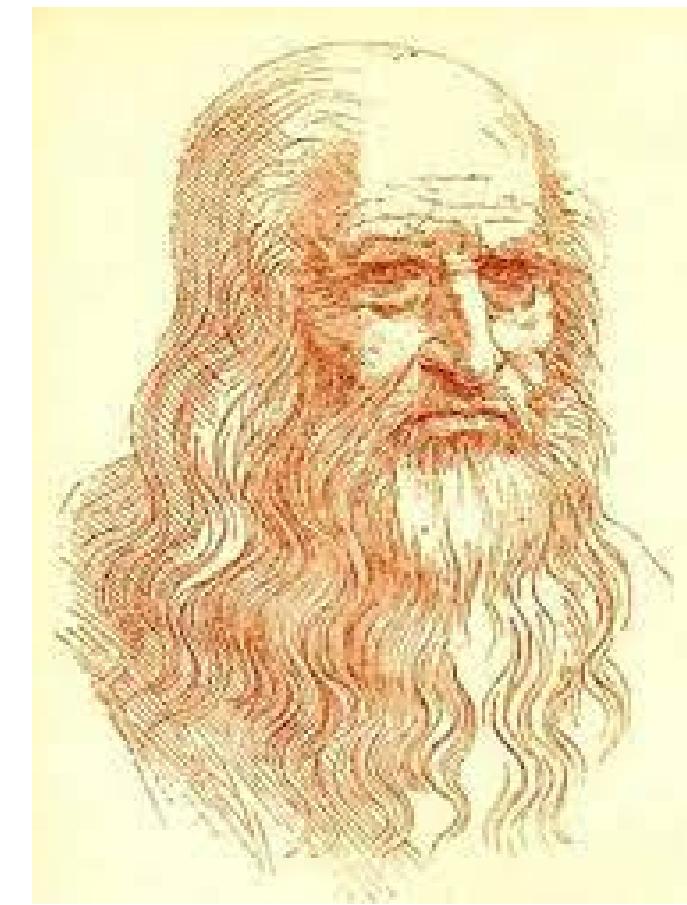
Figure 2: Water

Model is everywhere

Battery indicator is a model too

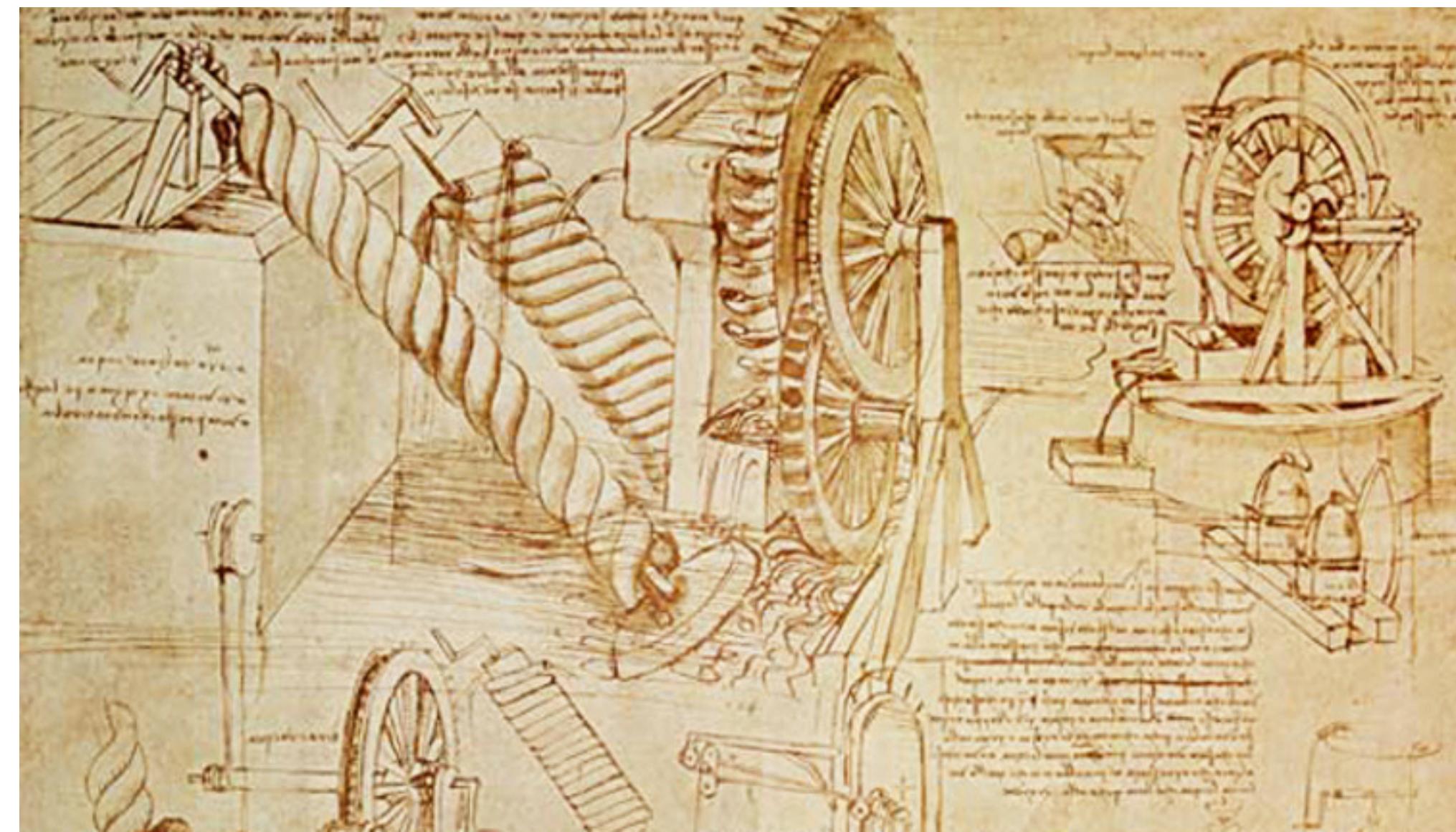
State of Charge 12V FLA & SLA	Flooded Lead Acid (FLA)	AGM (SLA)	Ranges & Warnings for all 12V Deep Cycle Batteries	State of Charge 12V LiFePo	LiFePo (LFP)
100%	12.7	13.0	Cycling in this zone will ensure a reasonable life expectancy. Depth of Discharge= 50% for FLA & SLA 80% for LiFePo	100%	14.4
90%	12.5	12.75		100%	13.6
80%	12.42	12.5		99%	13.4
70%	12.32	12.3		90%	13.3
60%	12.20	12.15		70%	13.2
50% (DoD)	12.06	12.05		40%	13.1
40%	11.9	11.95		30%	13.0
30%	11.75	11.81		20% (DofD)	12.9
20%	11.58	11.66		17%	12.8
10%	11.31	11.51		14%	12.5
0%	10.5	10.5	Dropping into this zone will permanently damage battery.	9%	12.0
				0%	10.0

Model can always be described using mathematical language



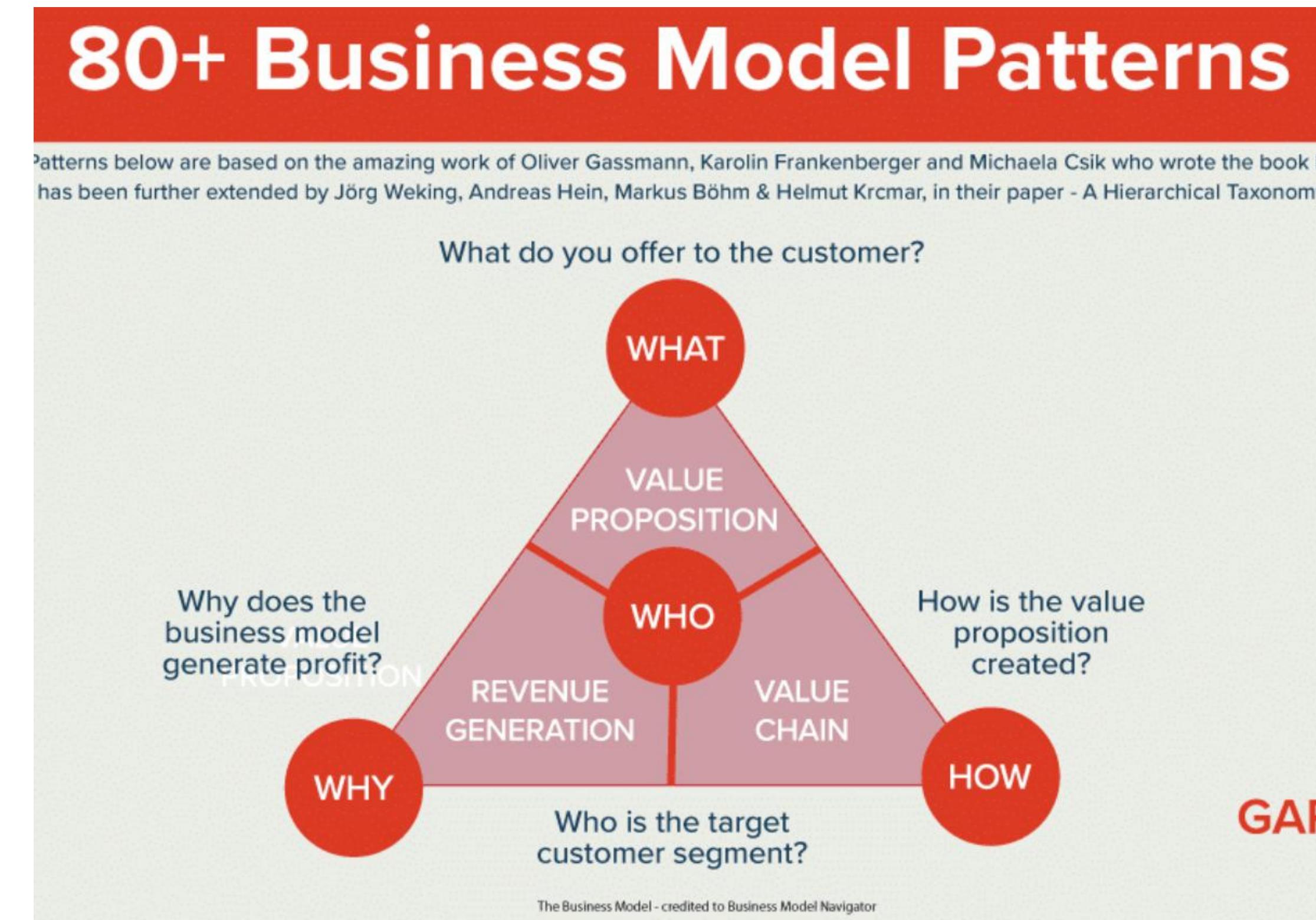
"... niuna umana investigazione si può dimandare vera scienza, d'essa non passa per le matematiche dimostrazioni..."

(no human research can be defined as true science if it cannot be mathematically demonstrated). Leonardo da Vinci (1651)

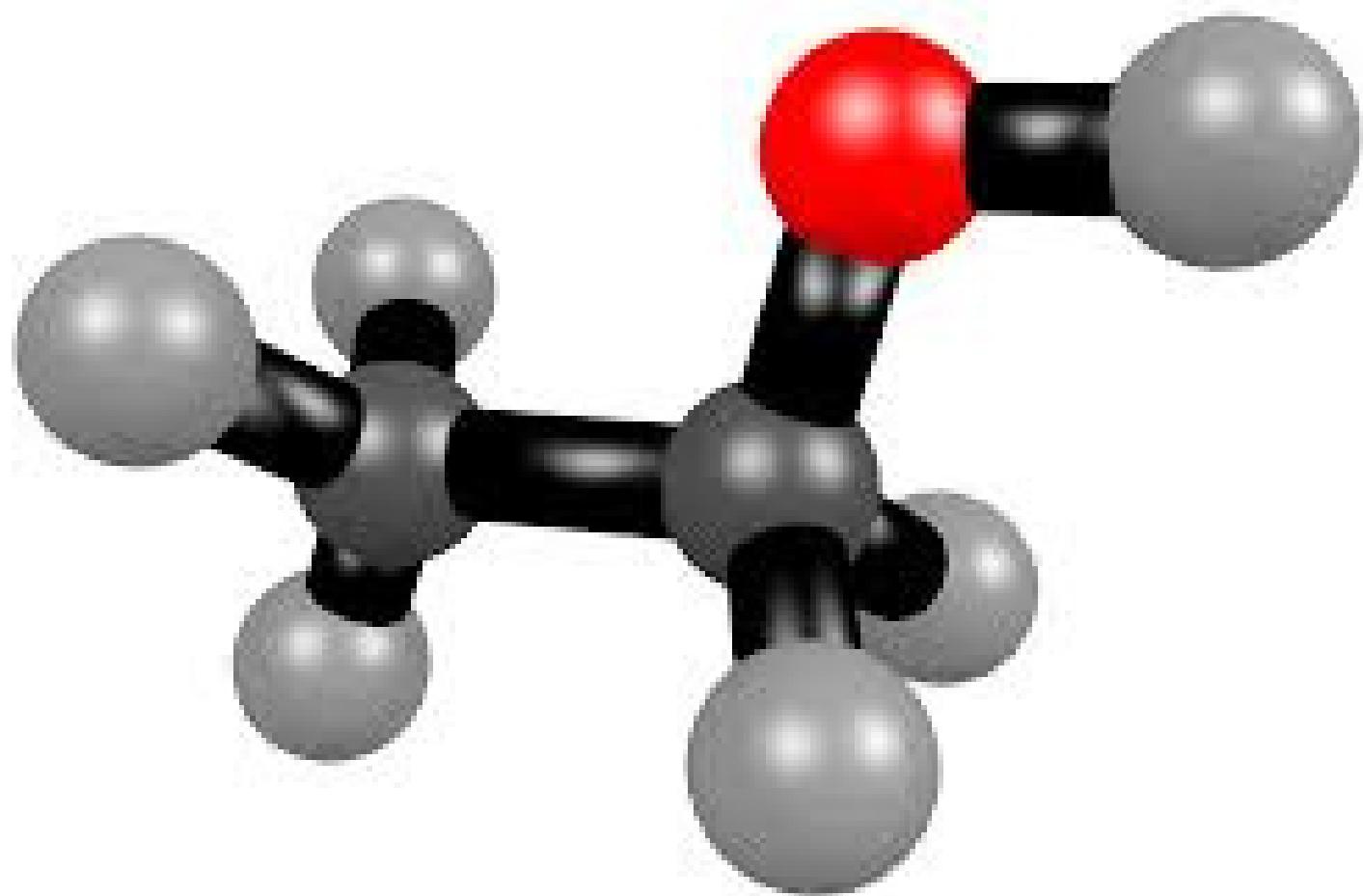


Model can always
be described using
mathematical
language

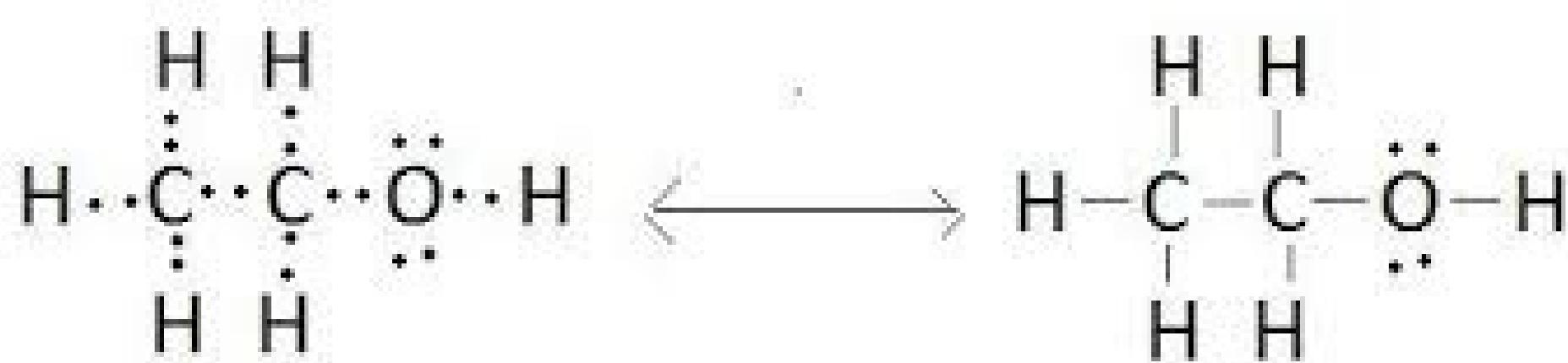
Even something as simple as 4
components connected with line



Model can always
be described using
mathematical
language



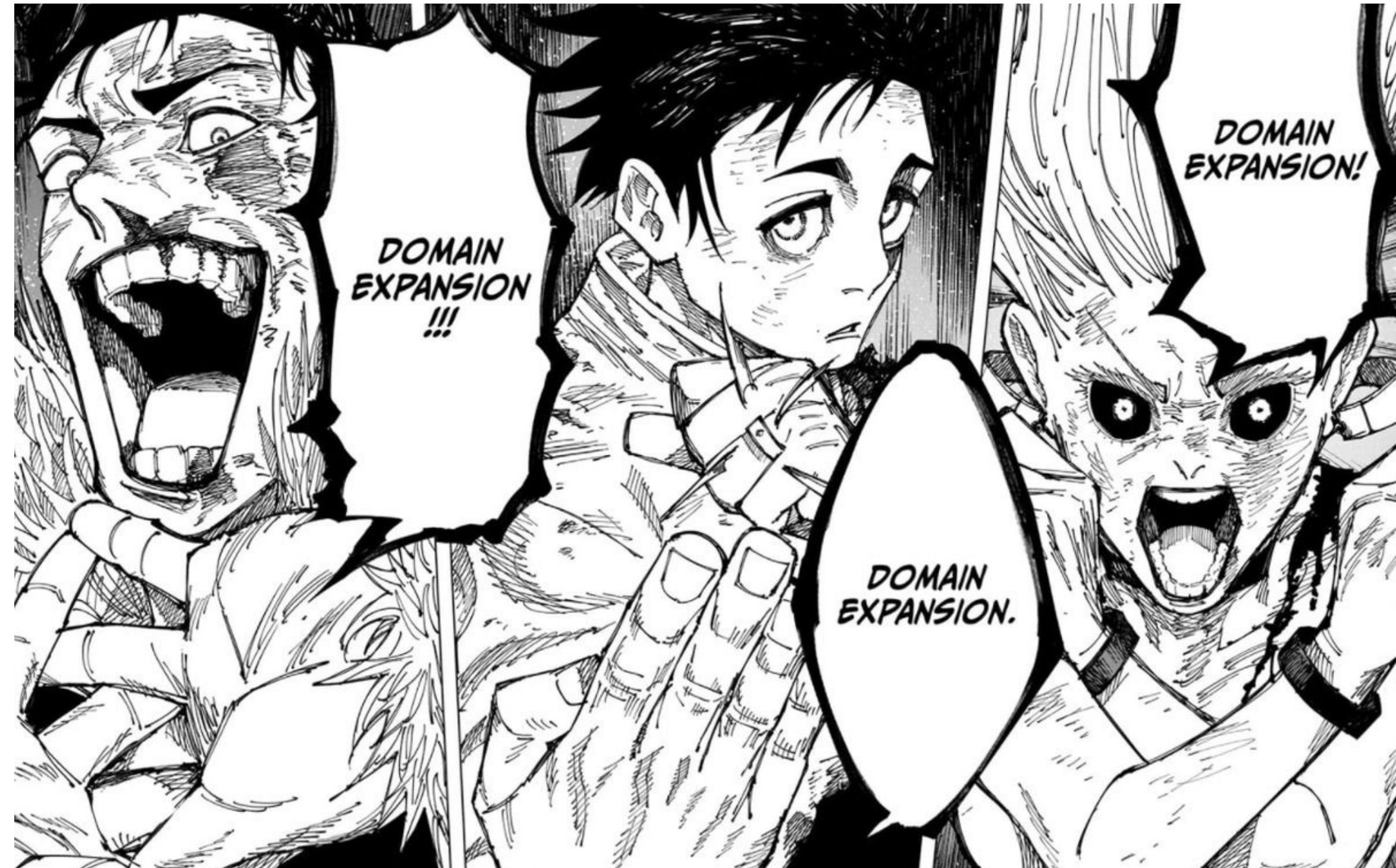
There are maths in the ball and
stick figure, and on Lewis
representation too!



Caveat 1: Domain Applicability

Model can only work within its own domain (In Javanese: *menang kandang*)

Usually based on the data where the model come from



Caveat 1: Domain Applicability

Model can only work within its own domain (In Javanese: *menang kandang*)

Usually based on the data where the model come from



Caveat 1: Domain Applicability

Expiration date



Caveat 1: Domain Applicability

weather forecast



Caveat 2:
Model only reveal
correlation not
causation



Caveat 2: Model only reveal correlation not causation



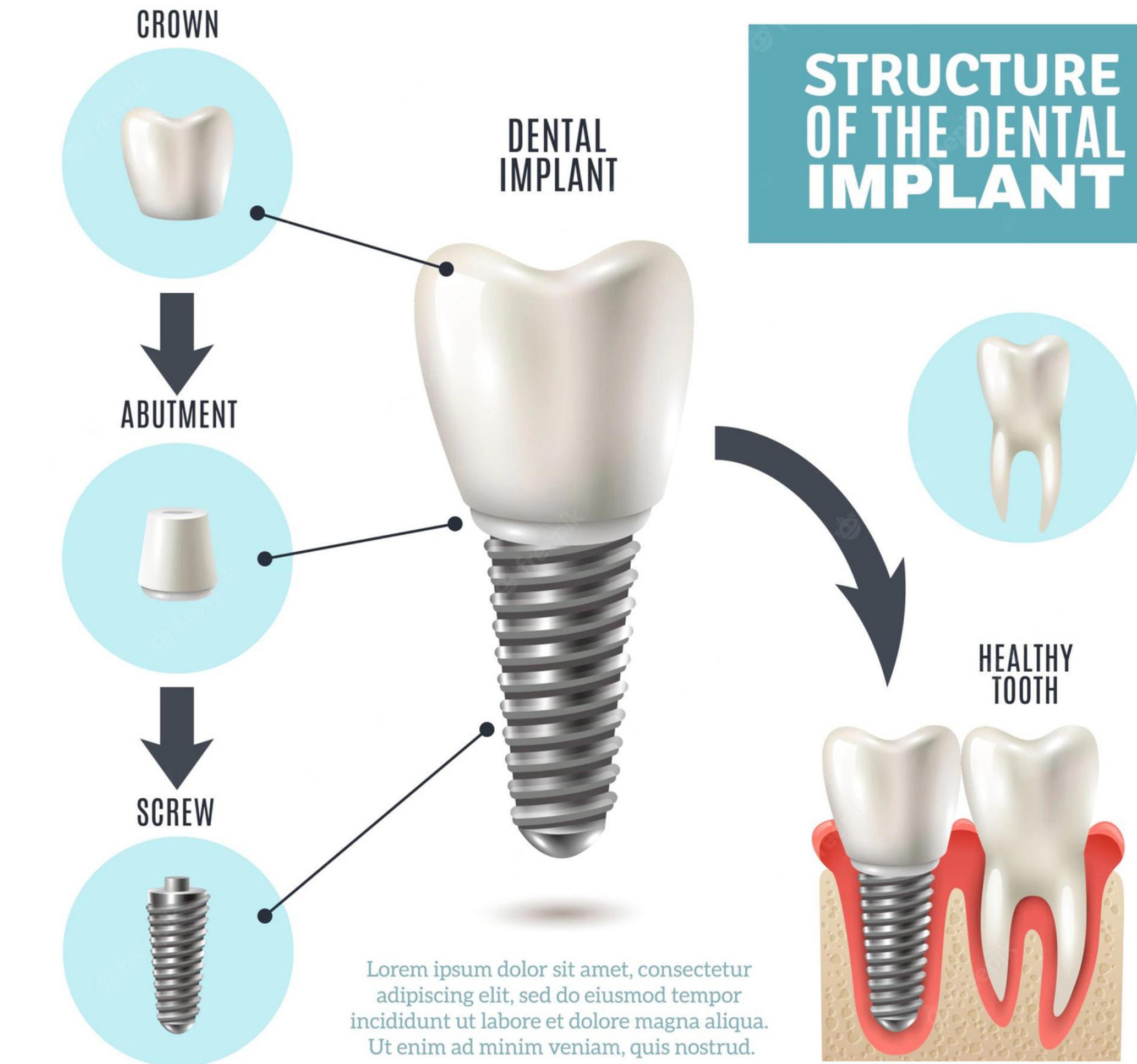
Consider this fun fact: Ice cream sales and violent crime are statistically related. As ice cream sales increase, so does violent crime. When ice cream sales decrease, so does violent crime. So it seems pretty clear that eating ice cream causes people to be violent, right? But that seems ridiculous.¹ That can't possibly be correct, can it? We will return to this example and answer that

How we use In silico?

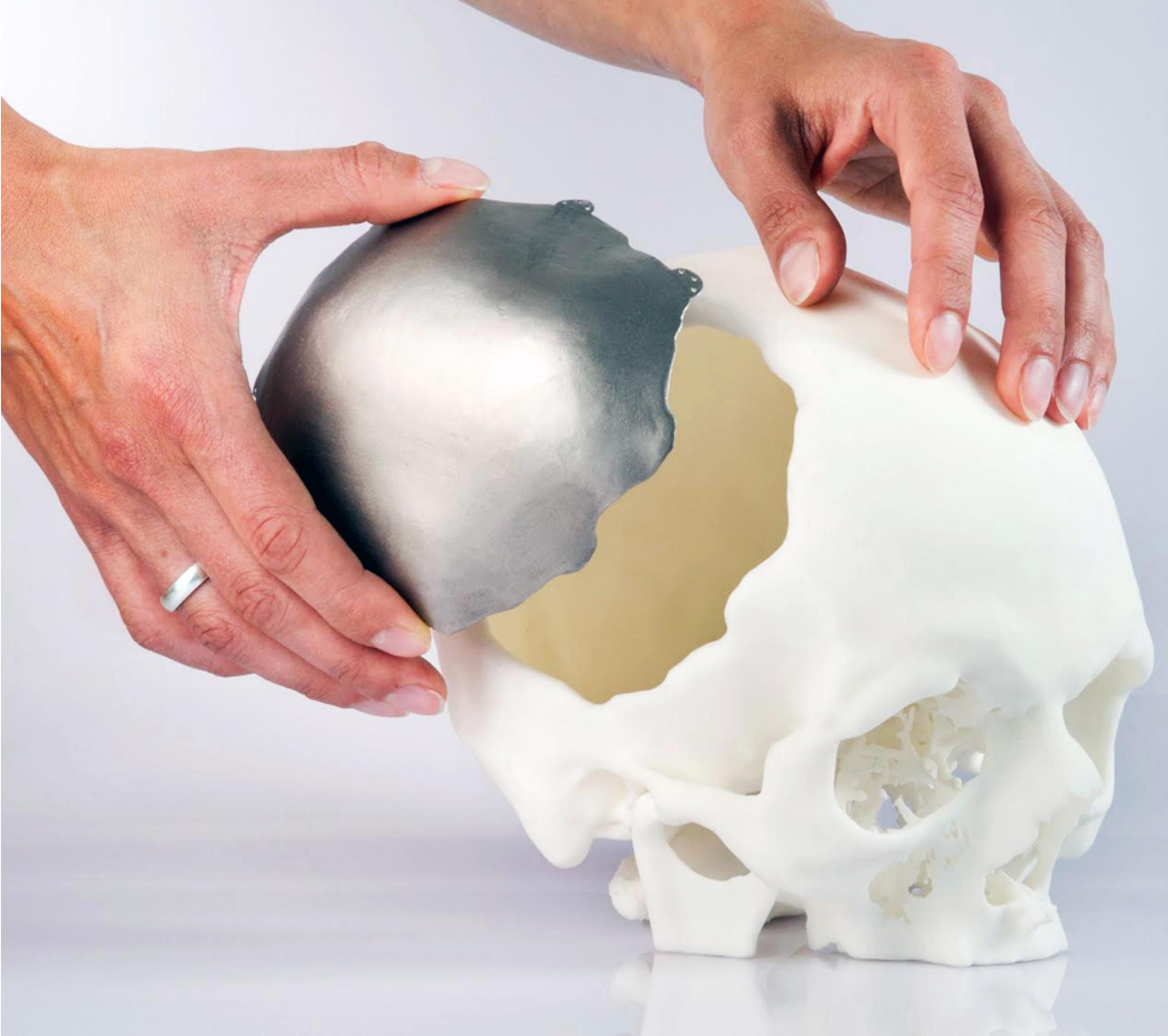
In silico has many usage in
Biomedical engineering research



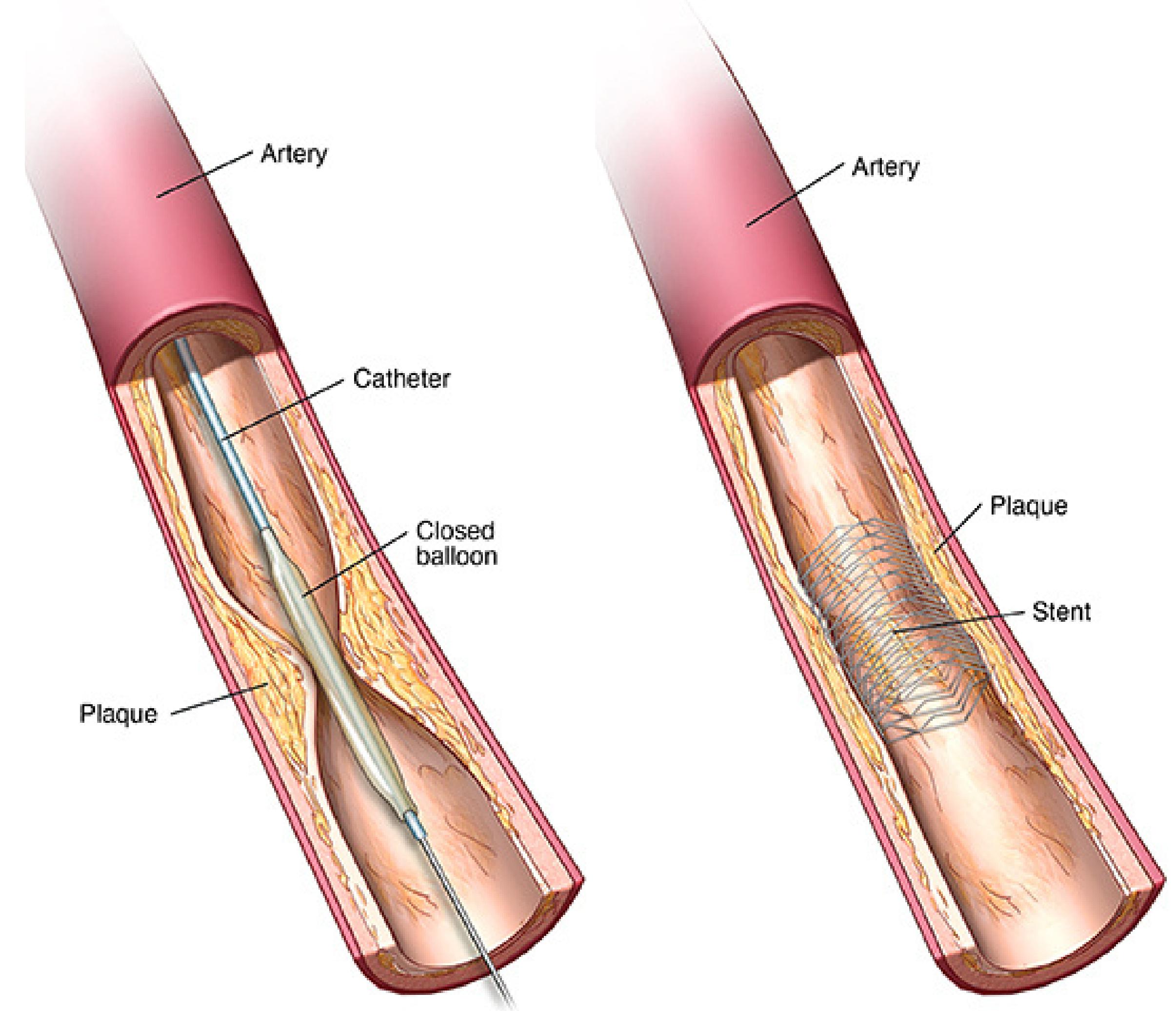
Dental implant design



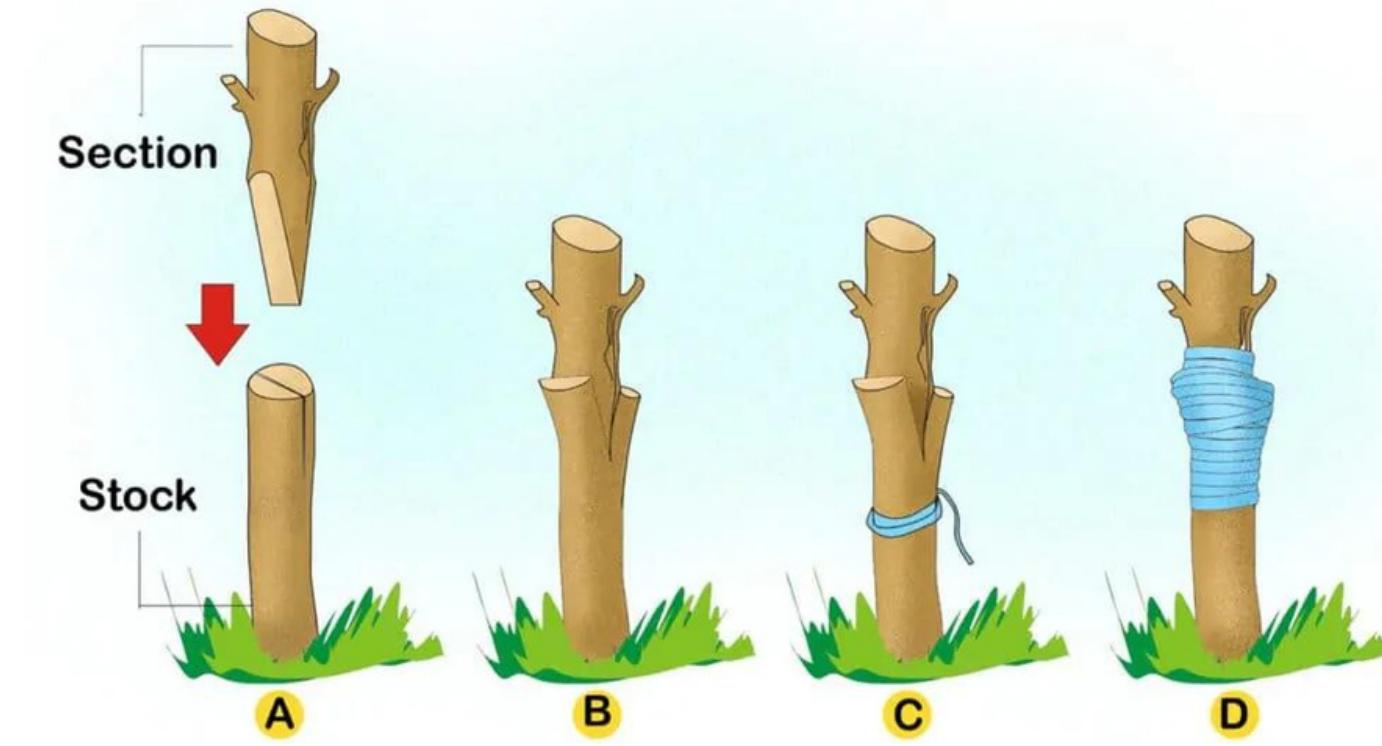
Cranial implant design



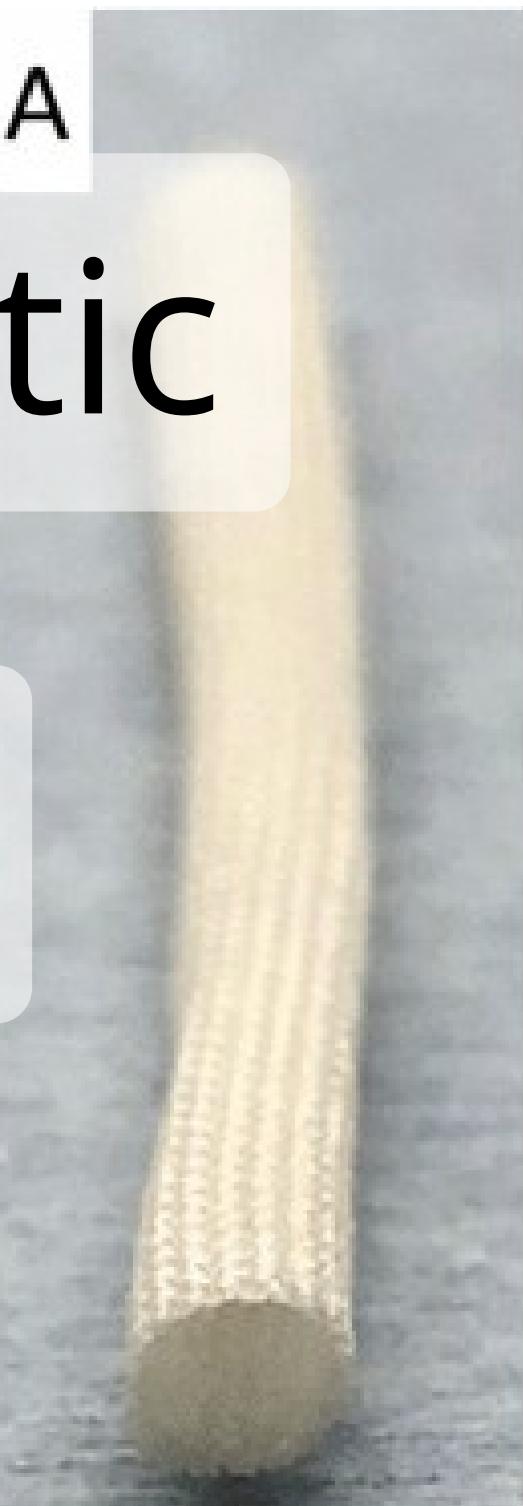
Stent design



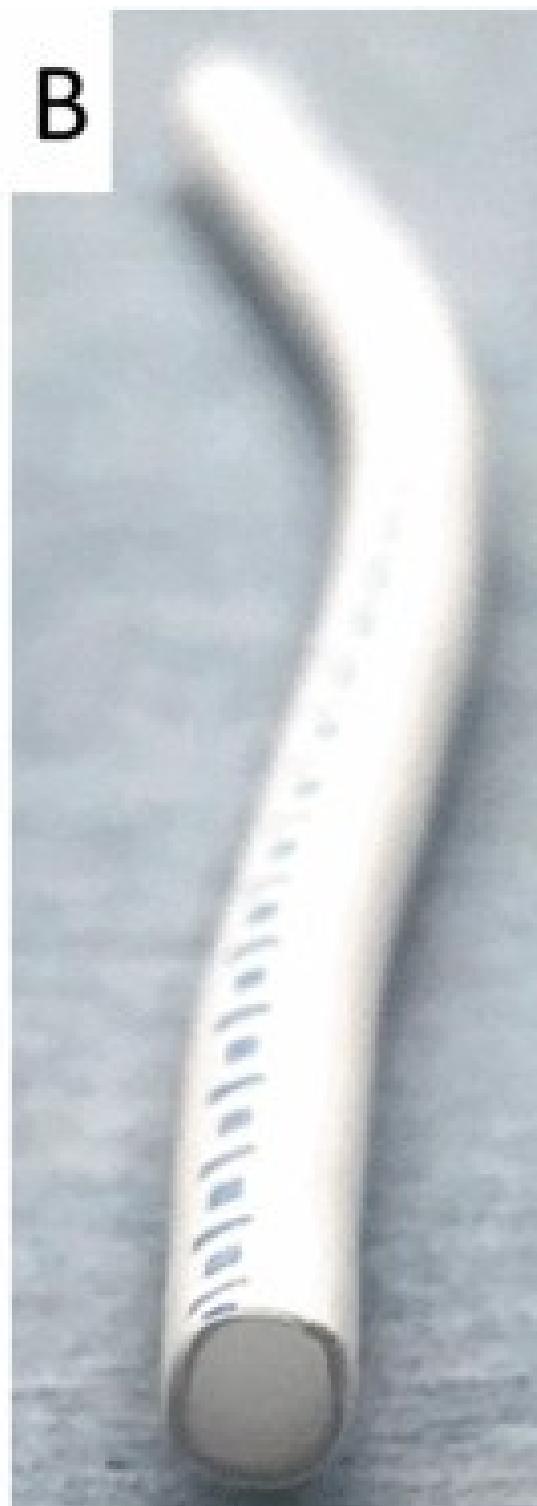
Synthetic graft design



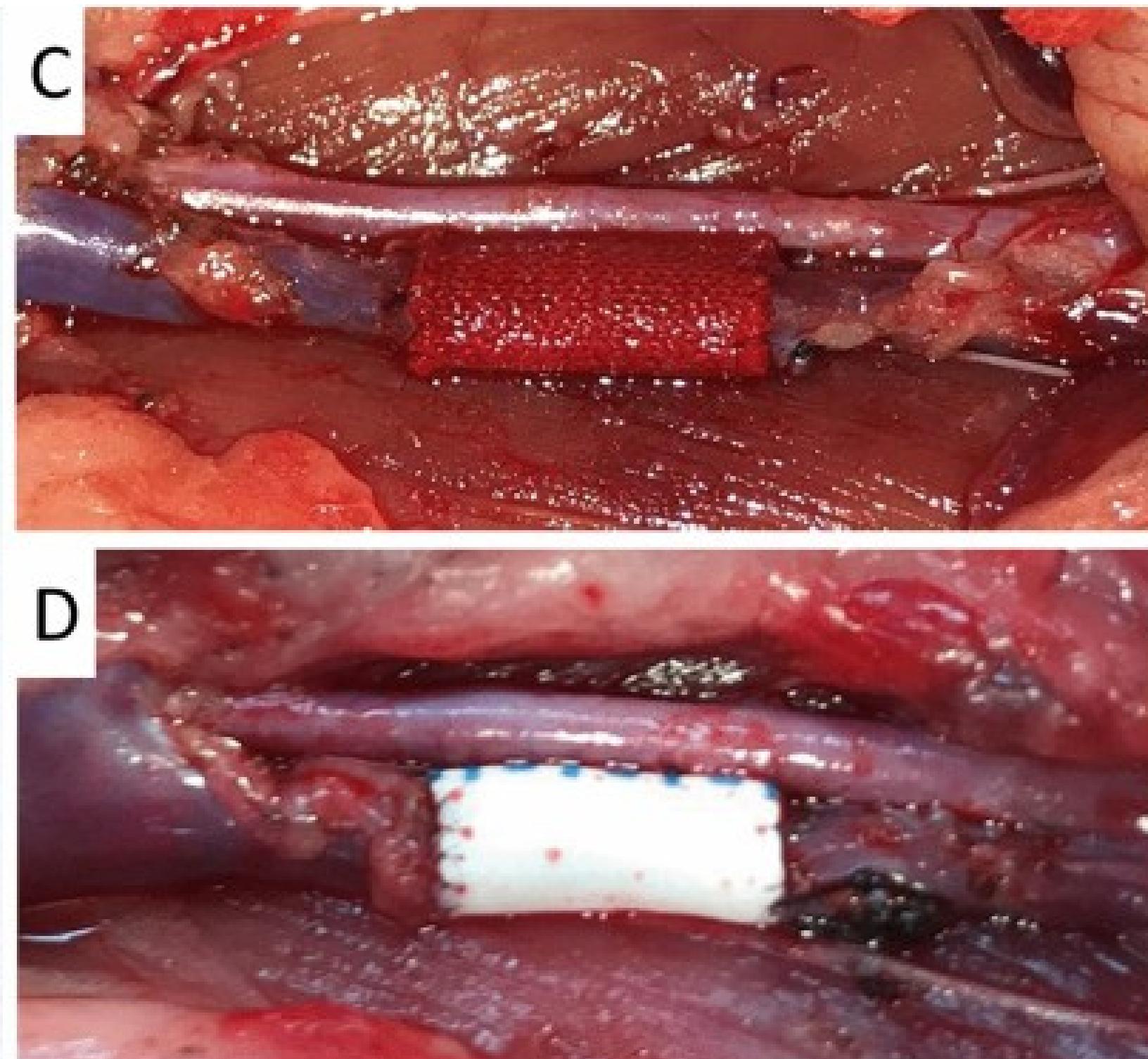
Synthetic graft design



3mm

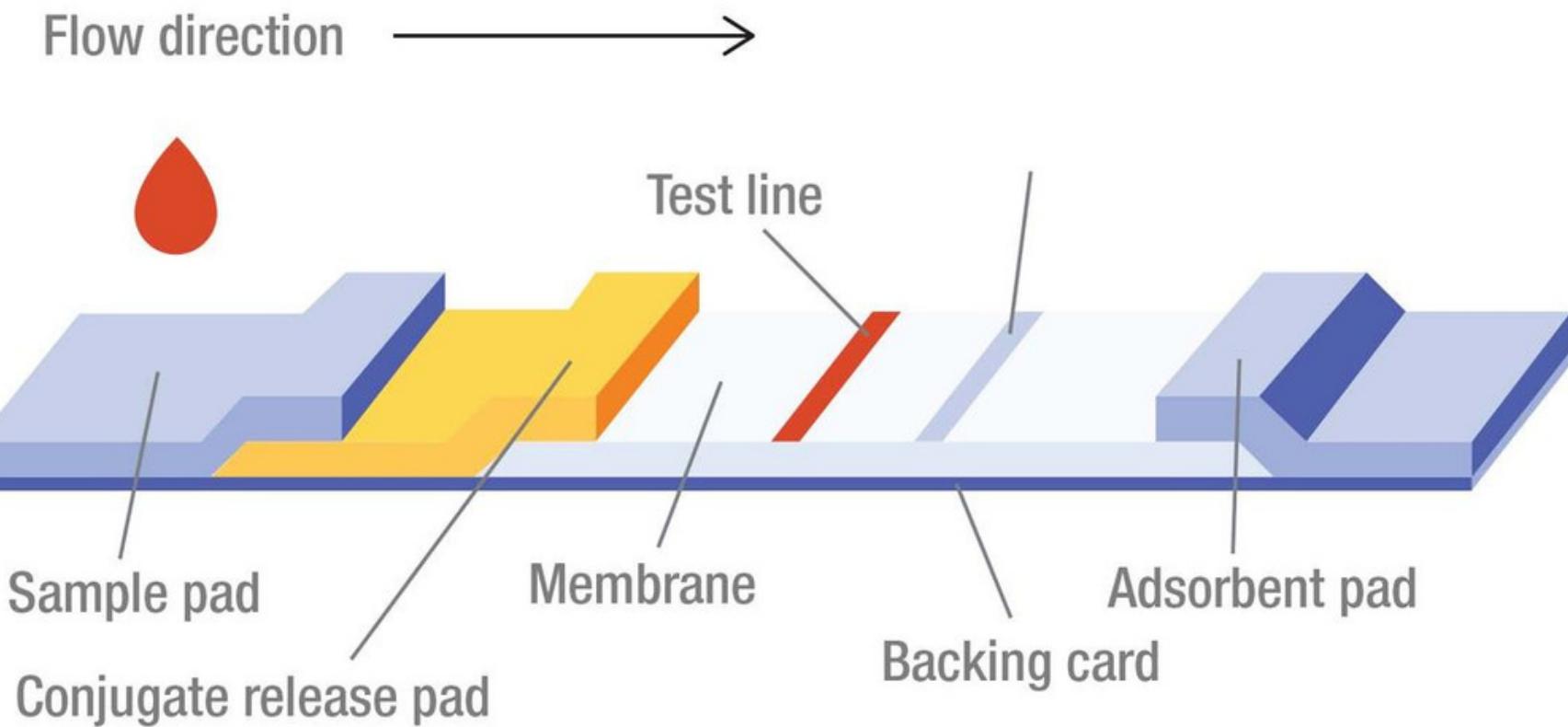


3mm



10mm

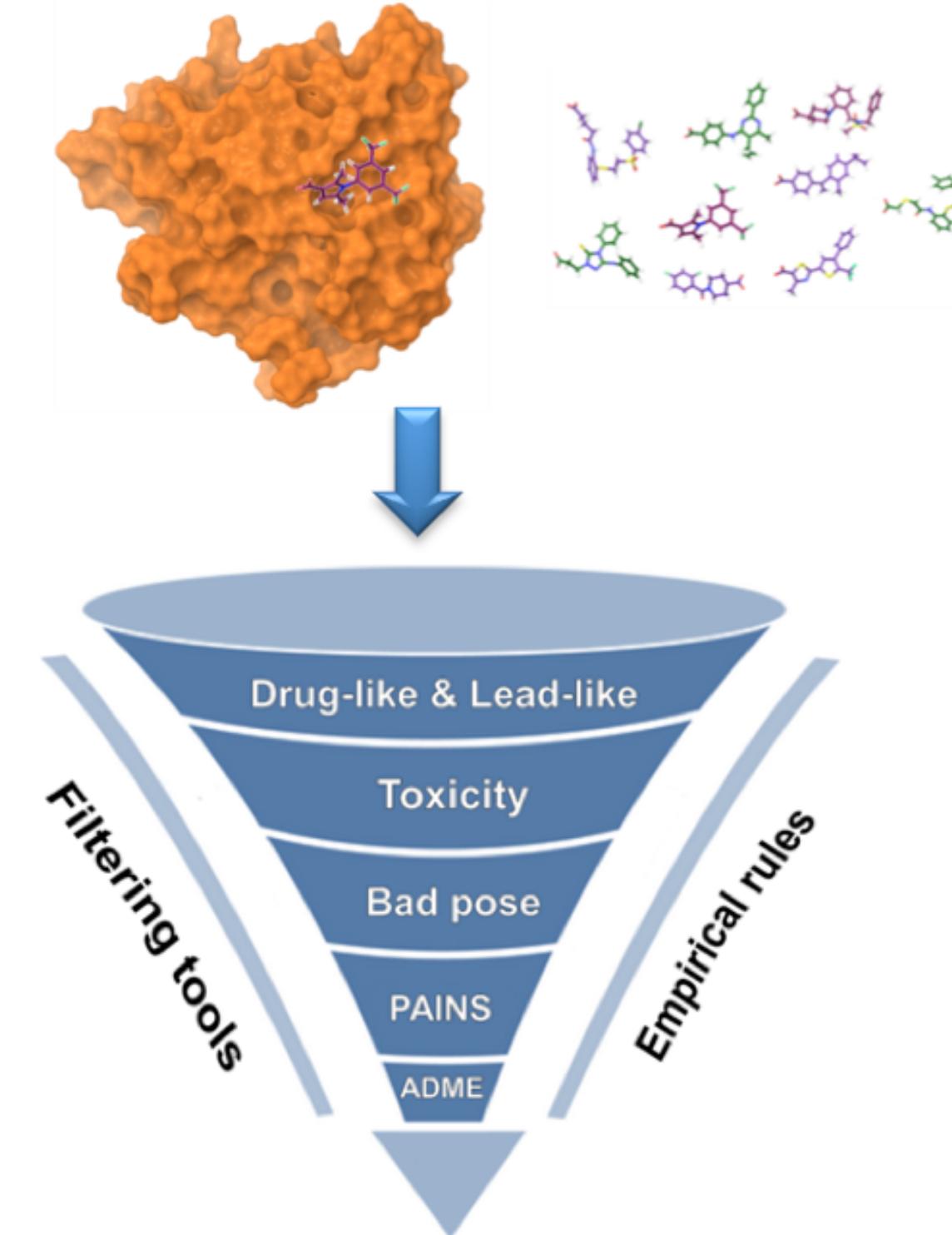
Rapid test (or any other diagnostics)



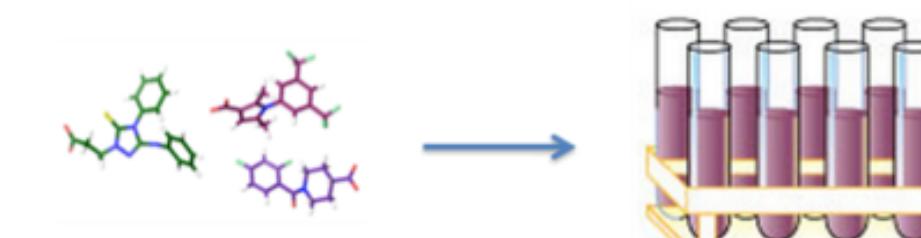
Last but not least

Computer-Aided Drug Discovery
(CADD)

Virtual Screening & Scoring

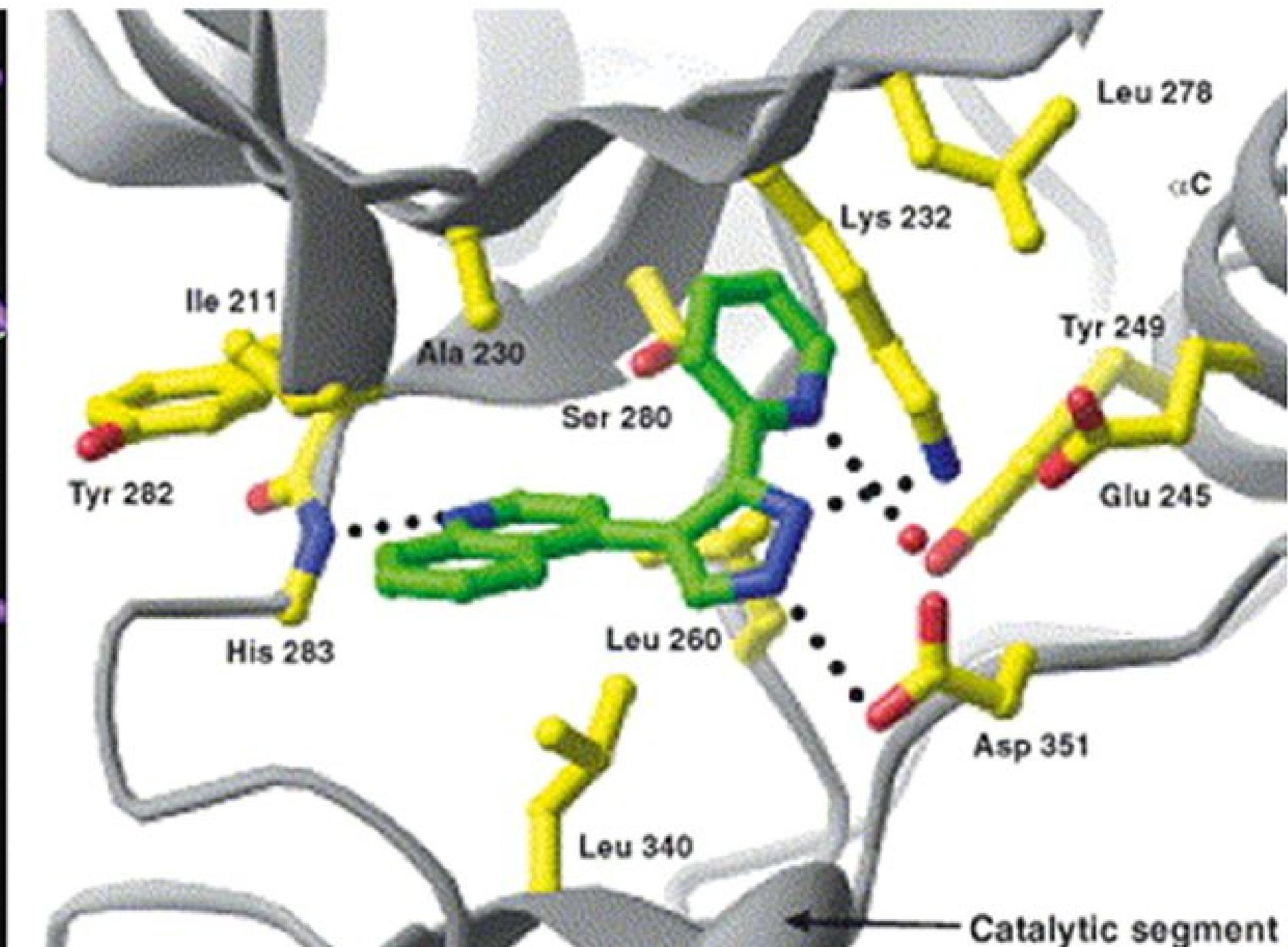
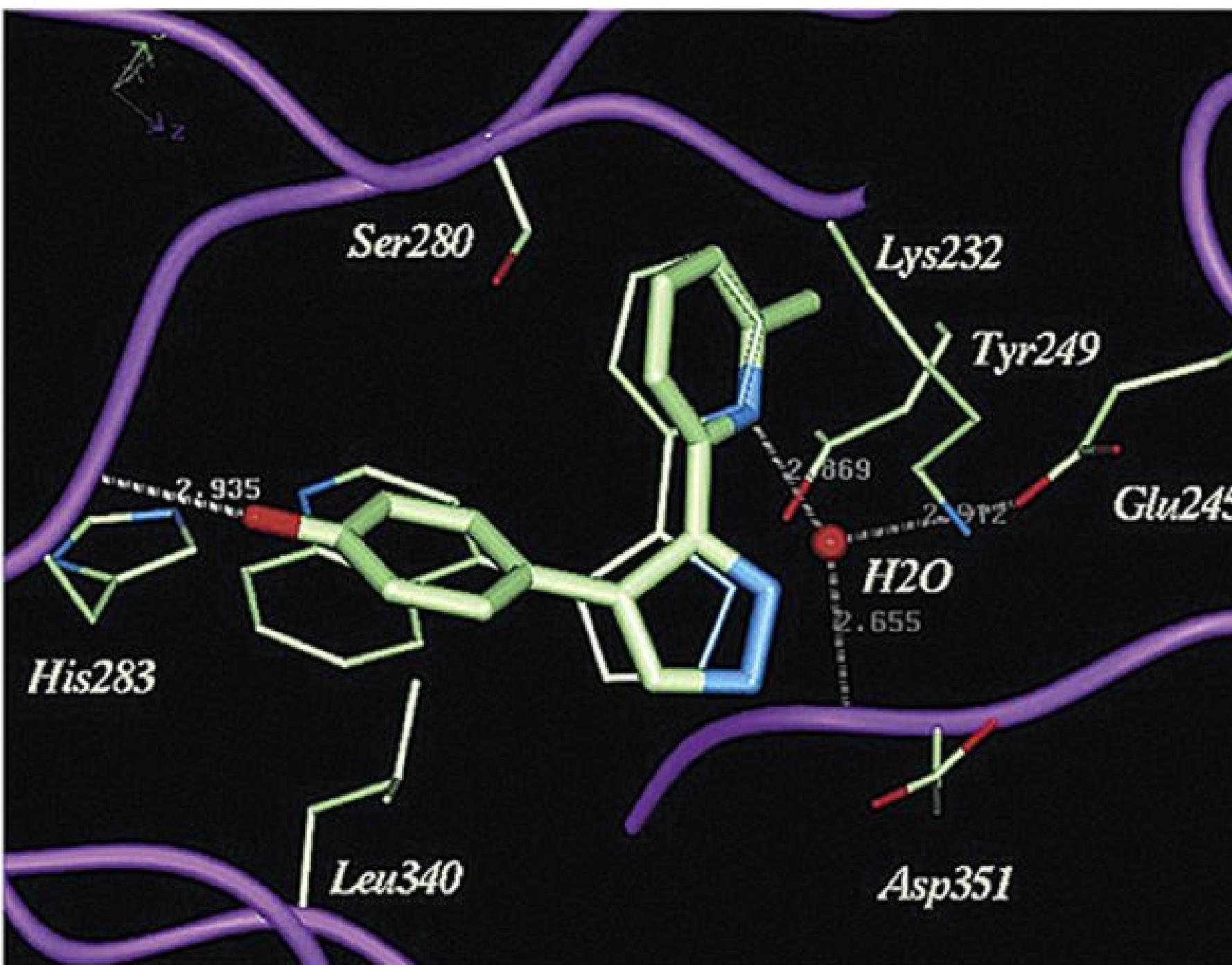


Re-ranking Virtual Screening results
Compound selection & *in vitro* assays



Enough with the theory

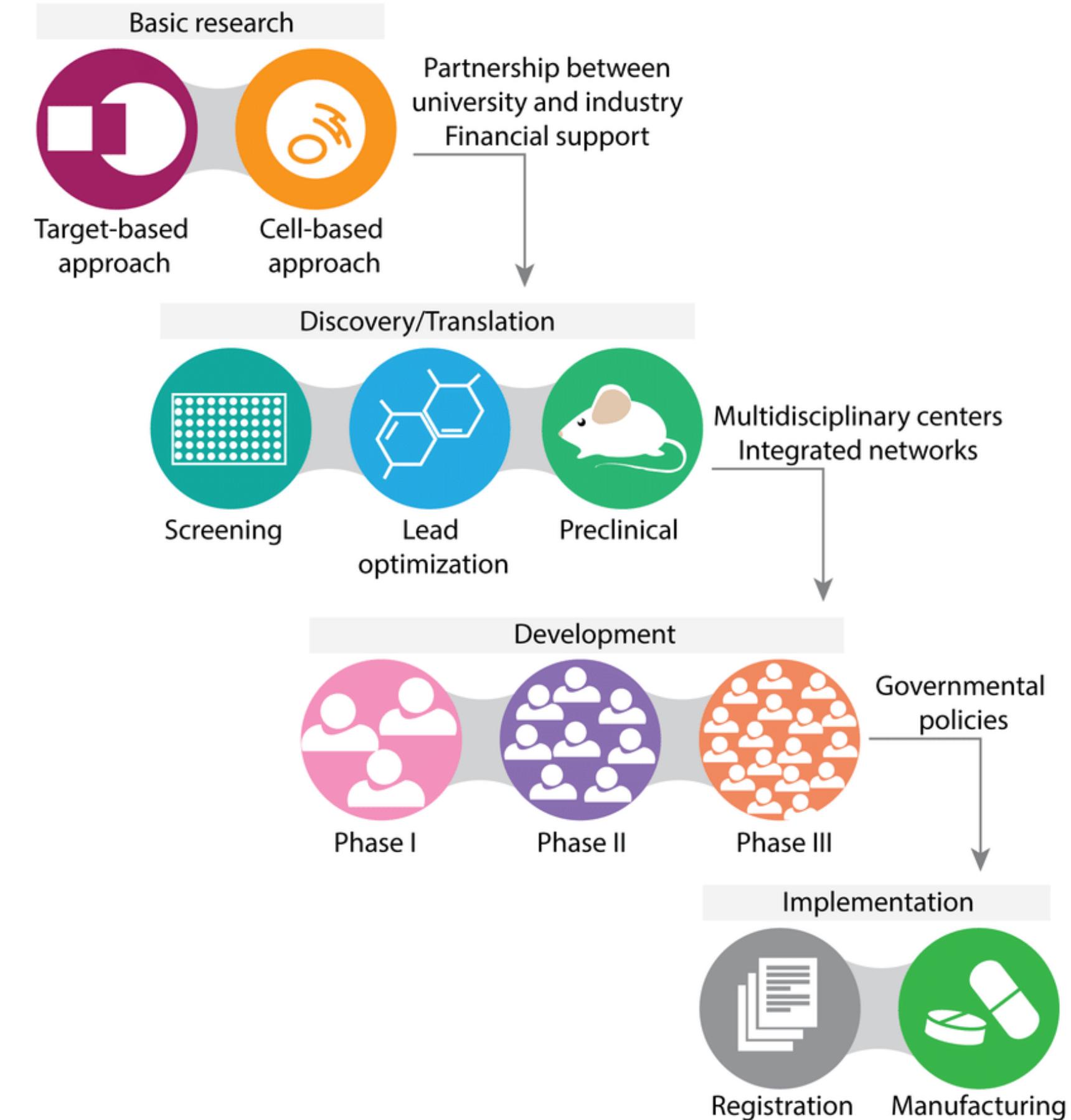
Lets move on to the real world application



Where does our drug comes from?

Paracetamol, Ibuprofen,
Penicilline, CTM, Captopril,
Metformin, Salbutamol,
Atorvastatin, Hydrocortisone,
Betamethasone, etc.

Drug Discovery Pipeline



How hard is it to deliver one drug into the market?

The recent estimates of the mean cost of developing a new drug ranging from \$314 million to \$2.8 billion

Between 2009-2018:
FDA approved 355 new drugs and biologics
R&D expenditures were available for 63 (18%) products, developed by 47 different companies.

JAMA | Original Investigation

Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018

Olivier J. Wouters, PhD; Martin McKee, MD, DSc; Jeroen Luyten, PhD

How hard is it to deliver one drug into the market?

Accounting for the costs of failed trials:

- the **median** capitalized R&D investments to bring new drug to market was estimated at **\$1141.7 million**
- the **mean** investments was estimated at **\$1559.1 million**

(All costs are adjusted to 2018 US\$)

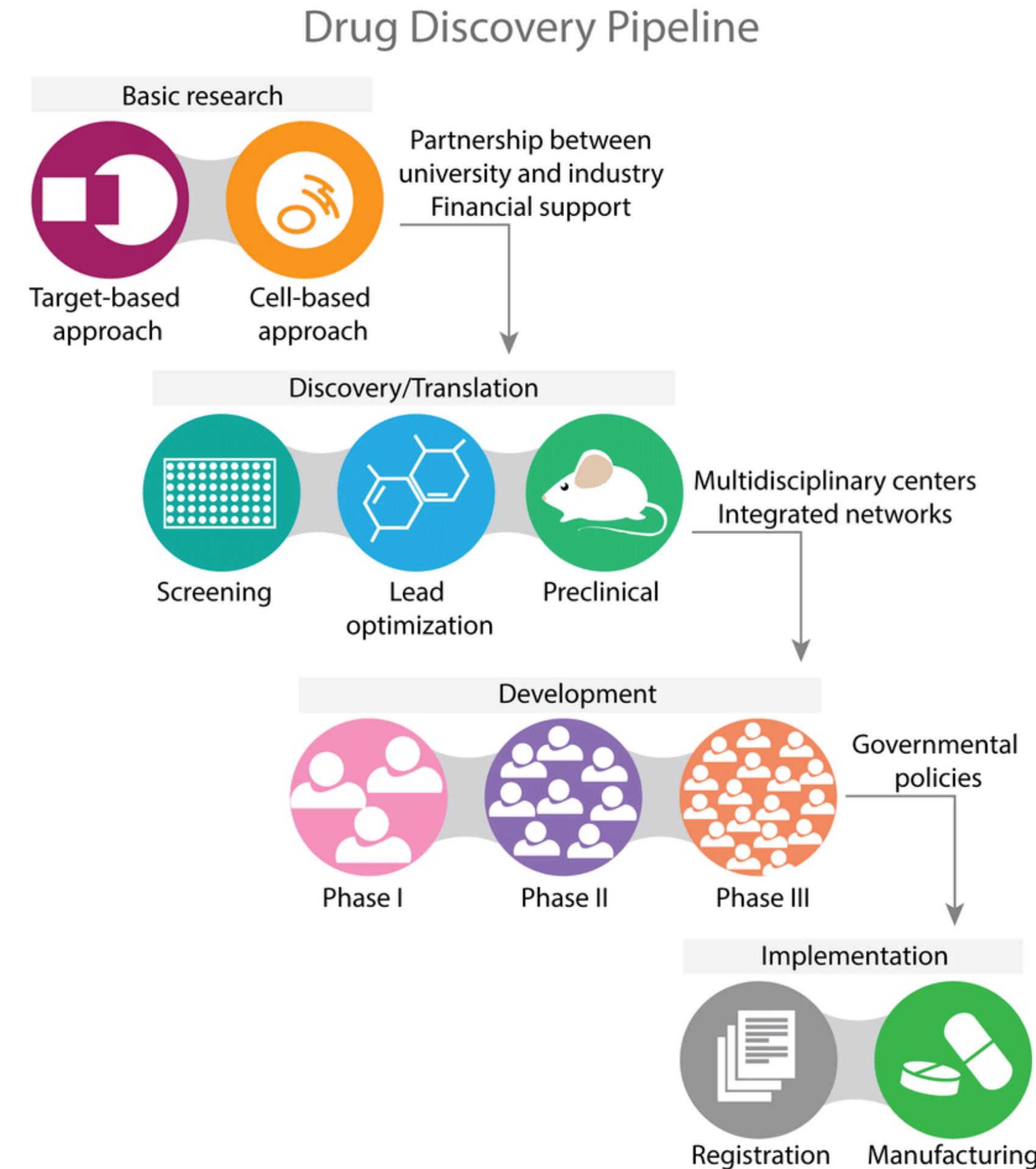
JAMA | Original Investigation

Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018

Olivier J. Wouters, PhD; Martin McKee, MD, DSc; Jeroen Luyten, PhD

How hard is it to deliver one drug into the market?

Not to mention that it takes **10-15 years** to complete all **3 phases of clinical trials** before the licensing stage.



How hard is it to deliver one drug into the market?

Clinical Trial Phases

Trials happen in four phases. Each phase serves a different purpose, and scientists collect important data at every step. Data must be reviewed and approved by the U.S. Food and Drug Administration (FDA) at each stage before a treatment can proceed to the next phase. Treatments must be approved through this process before they can be sold in the commercial market.

Phase I

In this phase researchers test a new product in a small group of people for the first time to evaluate **safety**, determine a safe dosage range, and identify side effects. This usually takes less than a year.



20 to 80 patients

Phase II

The product is given to a larger group of people to study **effectiveness** and further assess safety. This stage can last for months or years.



Hundreds of patients

Phase III

The product is given to **large groups** of people to confirm safety, monitor side effects, **compare it to other treatments**, and collect information that could make it safer to use. This phase usually takes years to complete.



Thousands of patients

Phase IV

This phase happens after a product is **approved for consumer sale**. The product is studied to find additional information about its risks, benefits, and best use.



Consumer market

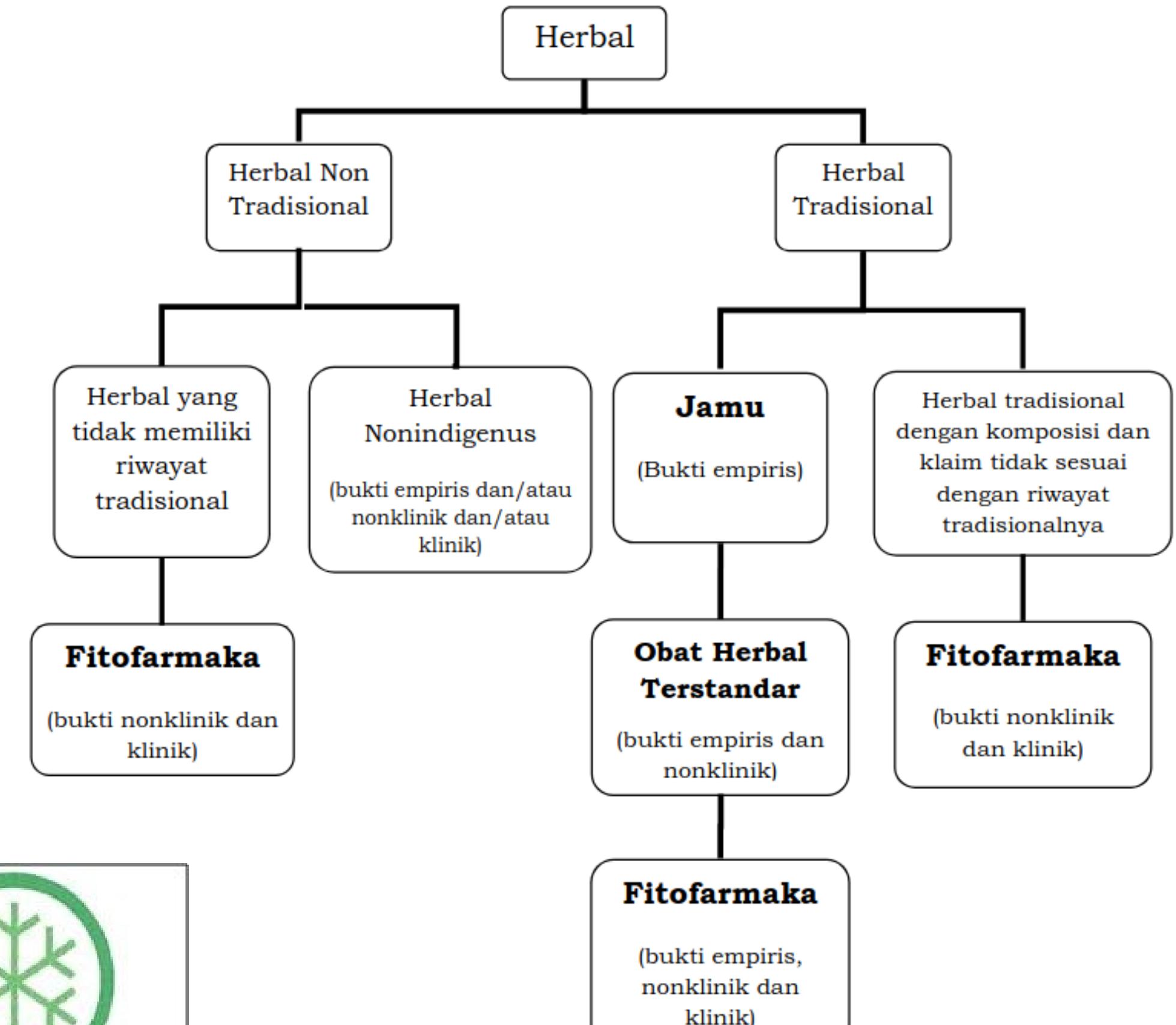
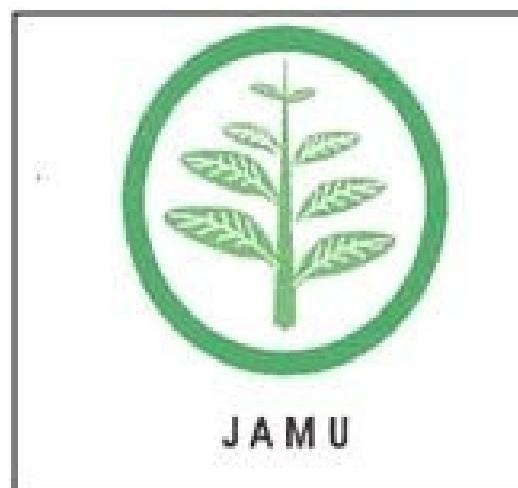
Things to consider:

- efficacy
- safety
- bioavailability (dosing)



How hard is it to deliver one drug into the market?

But what about traditional medicine in Indonesia?

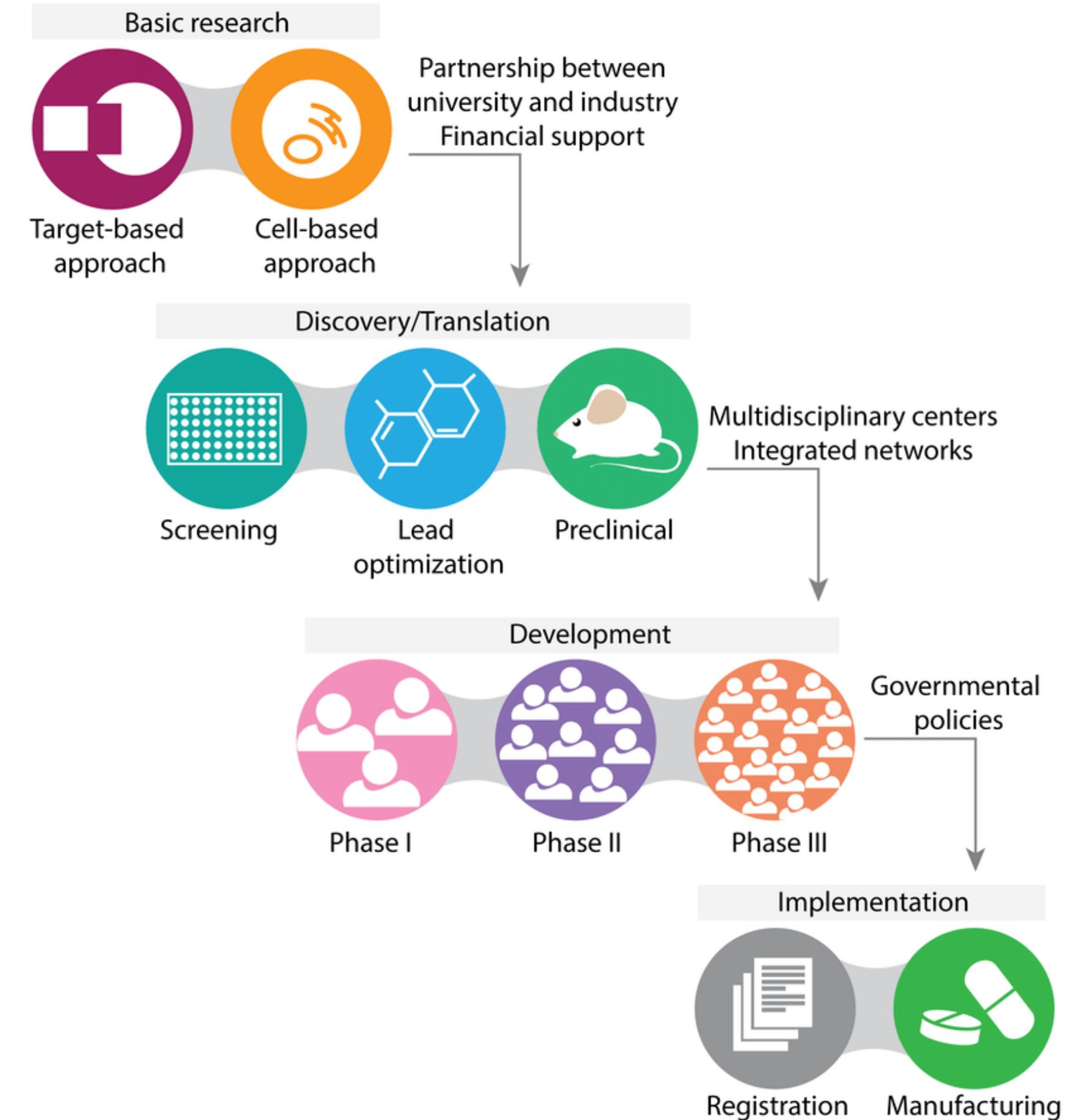


Peraturan Kepala BPOM RI No 13 Tahun 2014
Tentang
Pedoman Uji Klinik Obat Herbal

How hard is it to deliver one drug into the market?

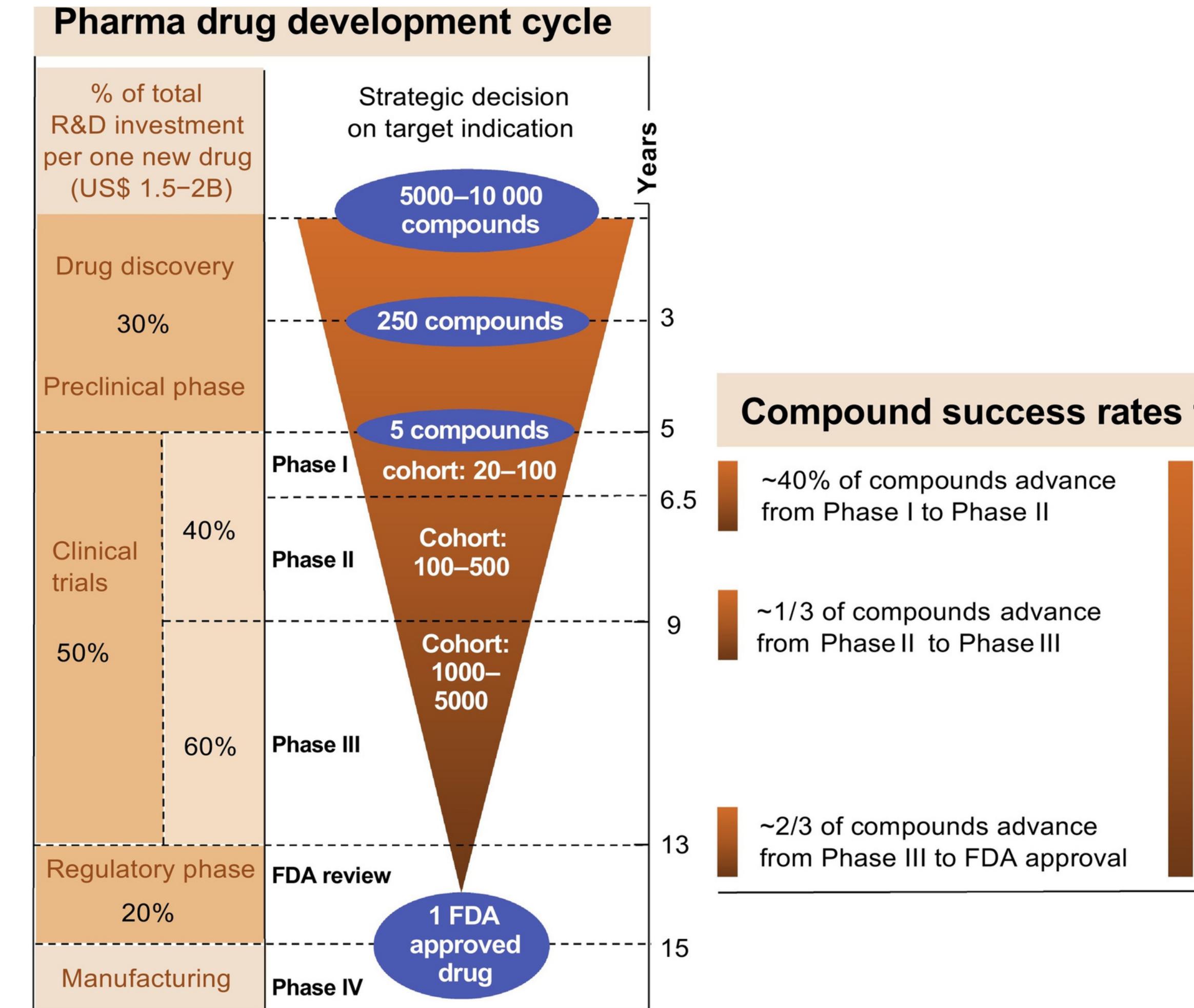
But why COVID-19 drug available so quickly?

Drug Discovery Pipeline



So what can we do to make it faster and cheaper?

This is where in silico method useful



In silico method may elevate success rate by predicting many properties of drug candidates

Note:

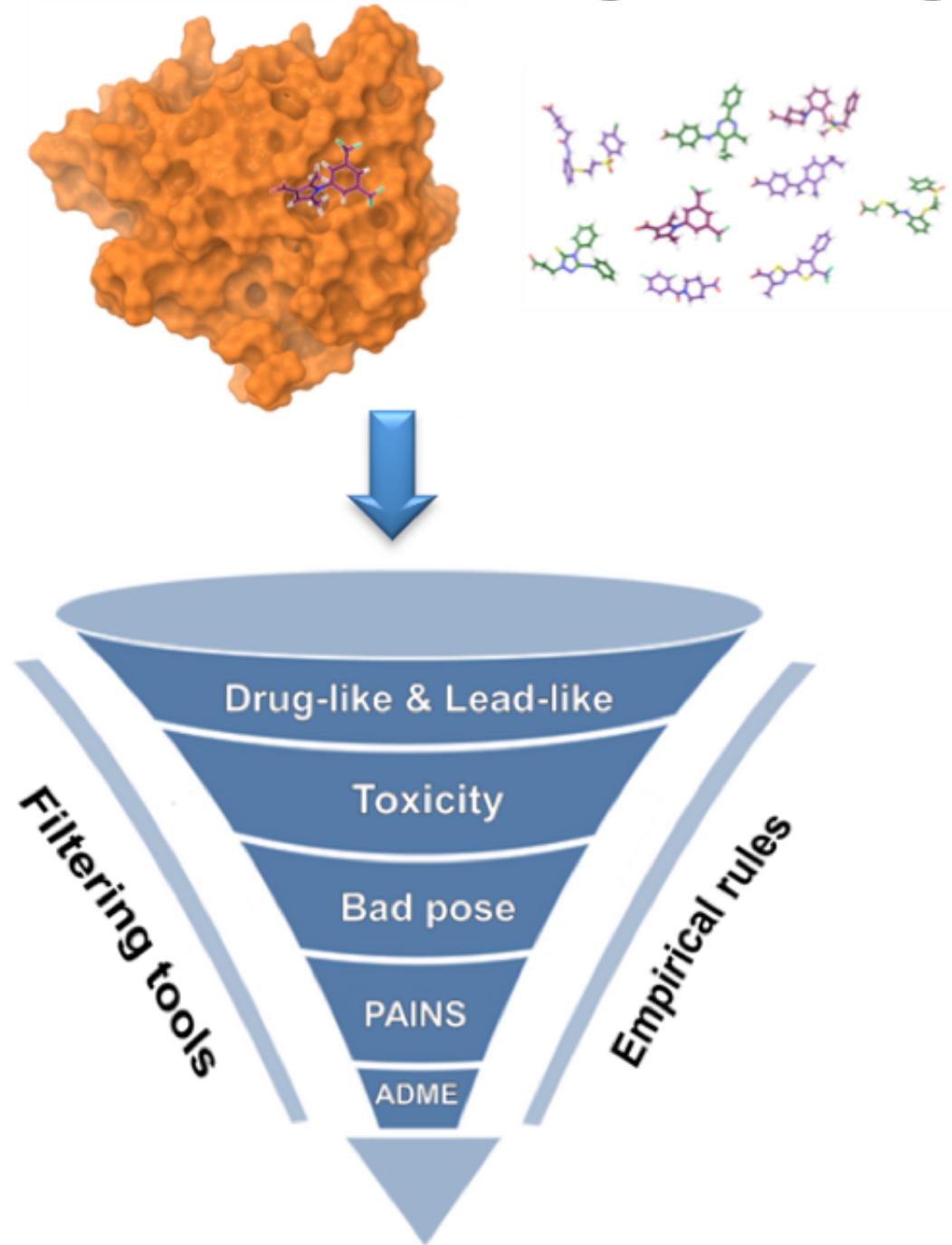
Drug-like = Lipinski rule (rule of five)

Lead = Has potential as drug candidate but needs to be improved (efficacy & safety)

PAINS = Pans-assay interference compounds

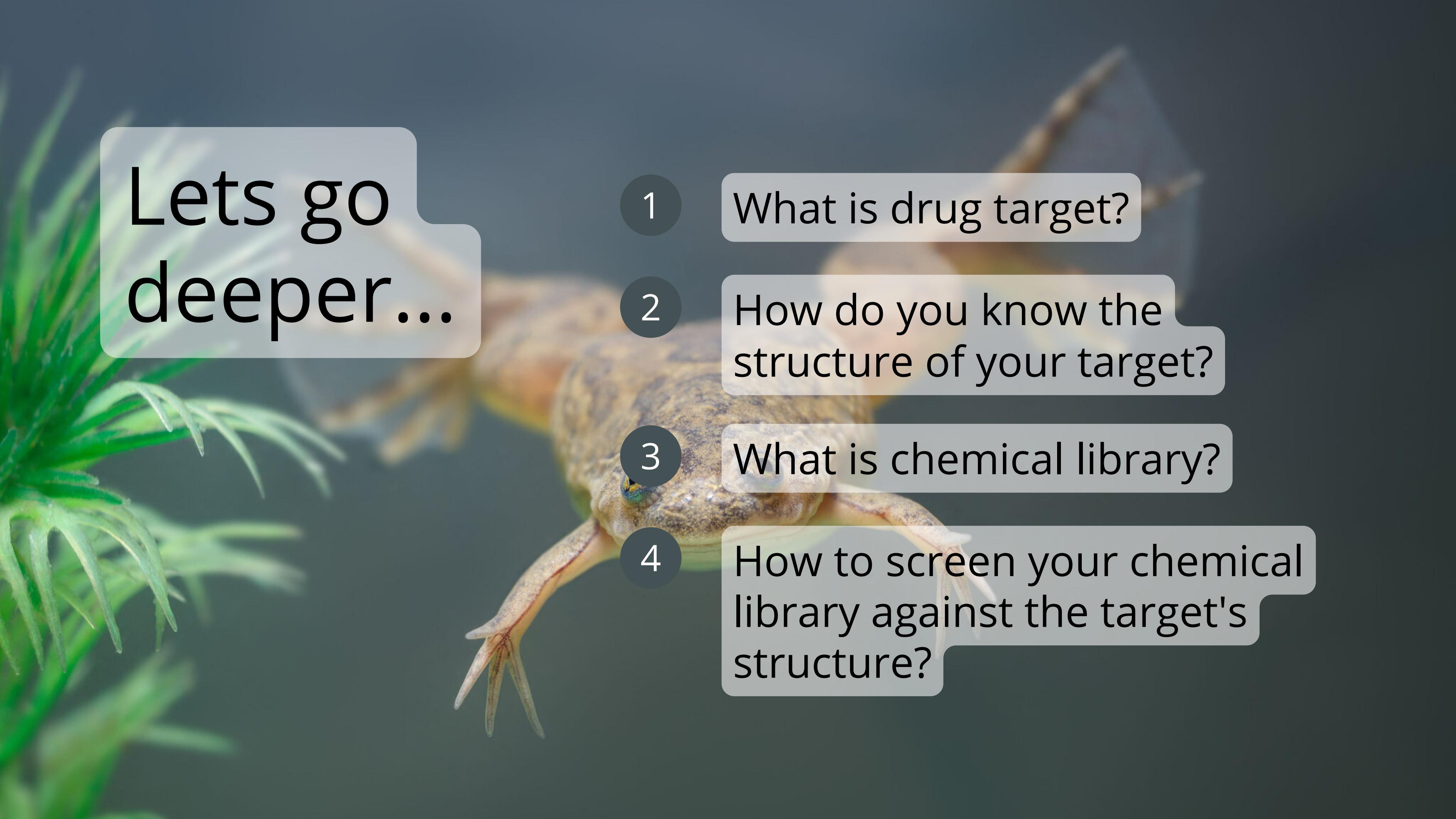
ADME = Absorption Distribution Metabolism Excretion
(In short: bioavailability profile)

Virtual Screening & Scoring



Re-ranking Virtual Screening results
Compound selection & *in vitro* assays





Lets go
deeper...

1

What is drug target?

2

How do you know the
structure of your target?

3

What is chemical library?

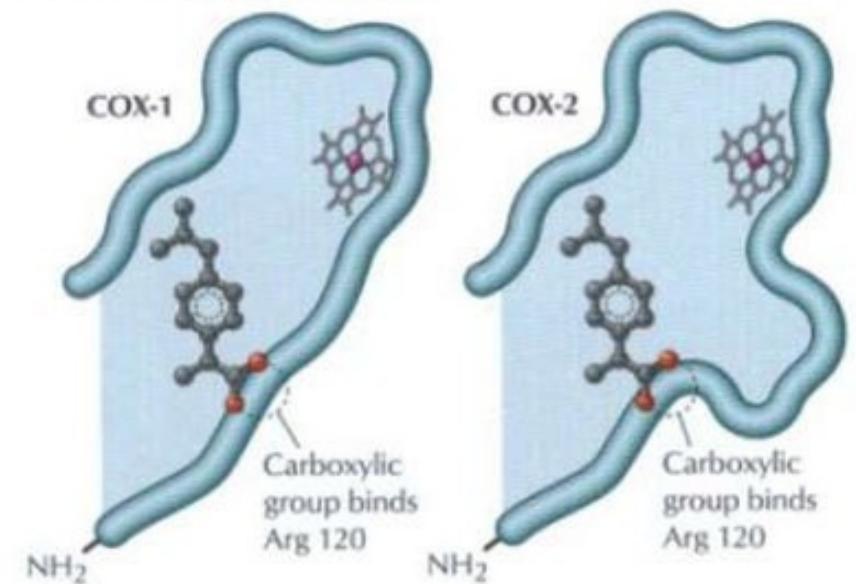
4

How to screen your chemical
library against the target's
structure?

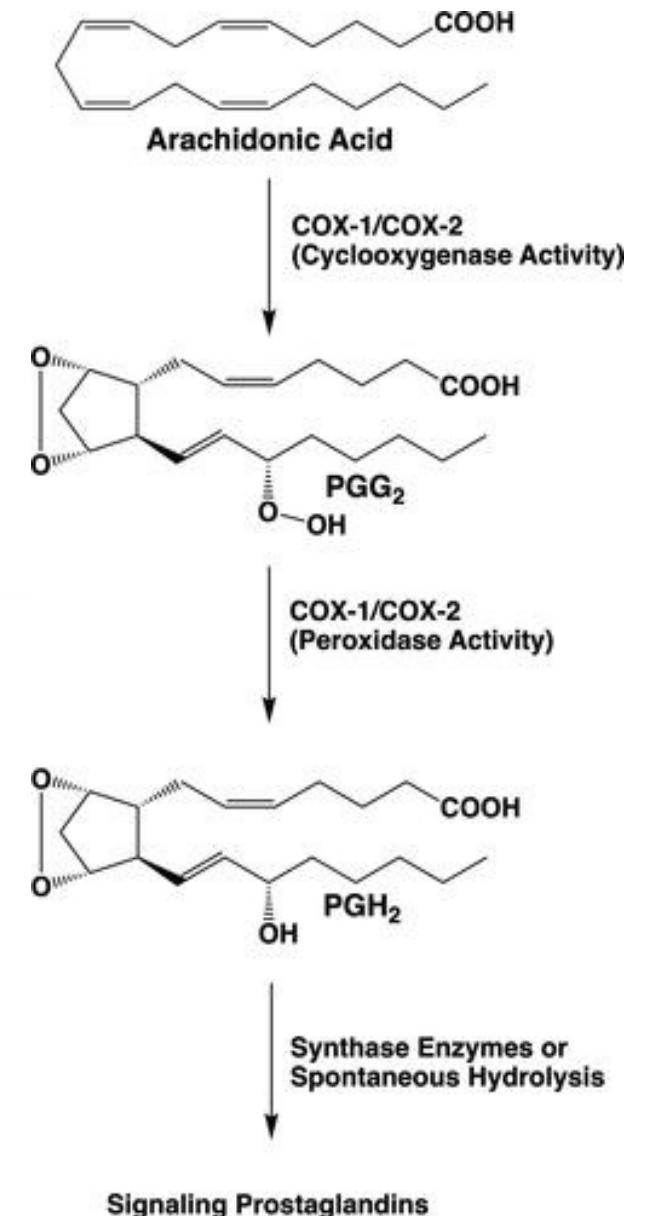
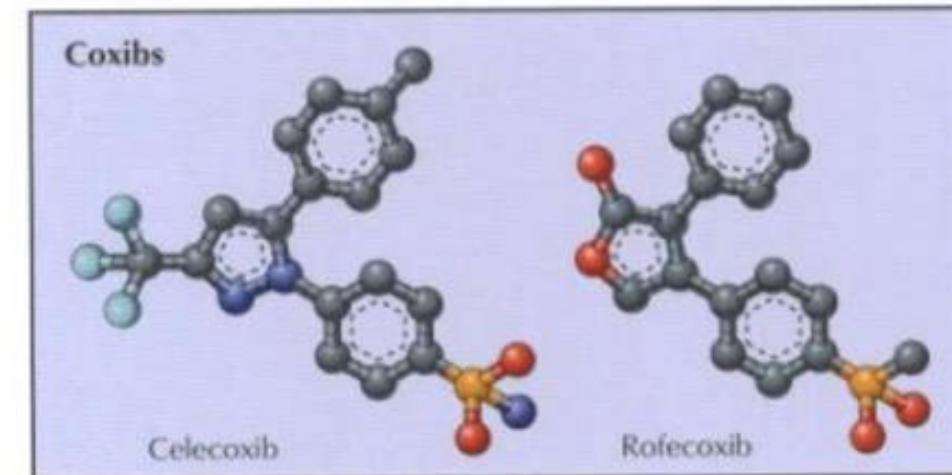
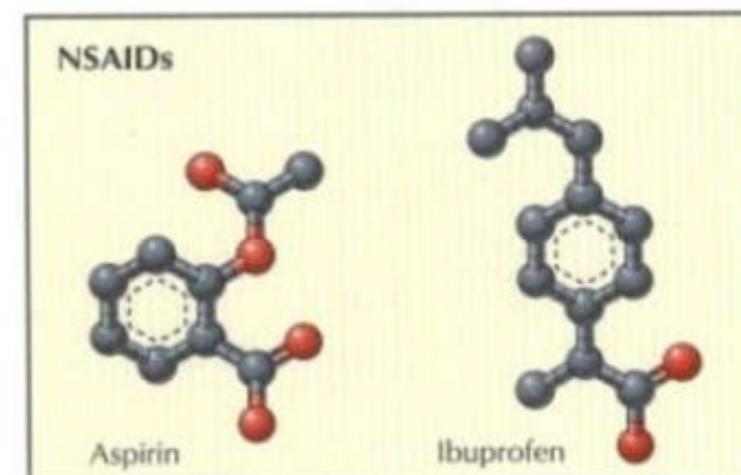
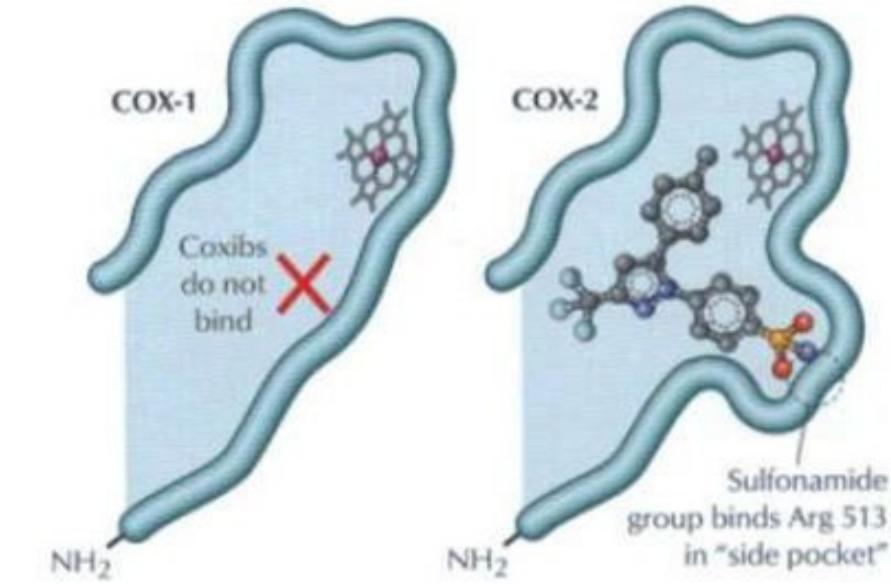
Drug target I

Mostly protein, one
of them is **enzyme**.

NSAIDs: Mechanism of Action



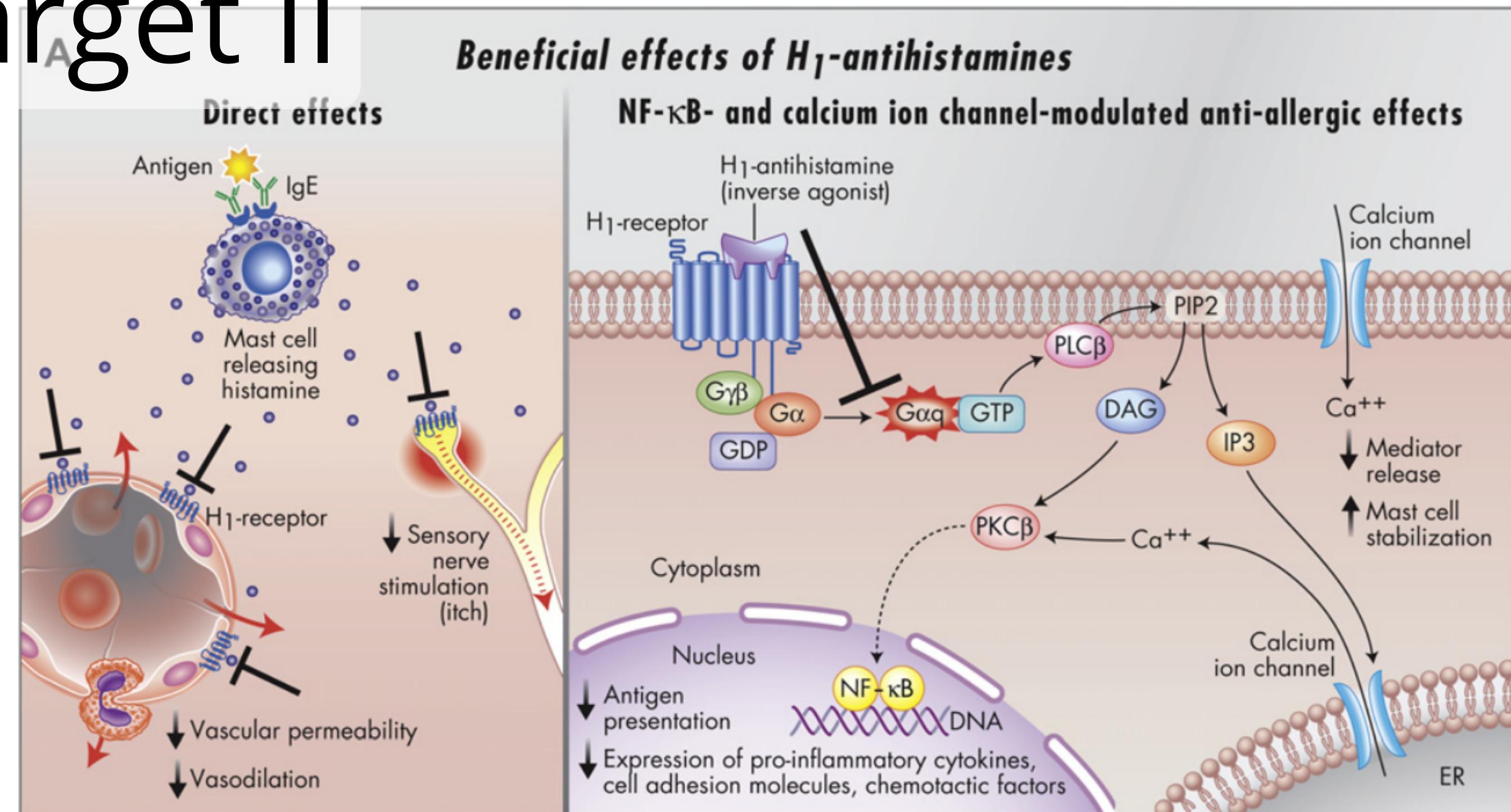
Coxibs: Mechanism of Action



Drug target II

Receptor

When activated by certain ligand will activate chain of reaction inside cell



Drug target III

Transporter

When activated by certain ligand will increase the I/O flux of certain molecule.

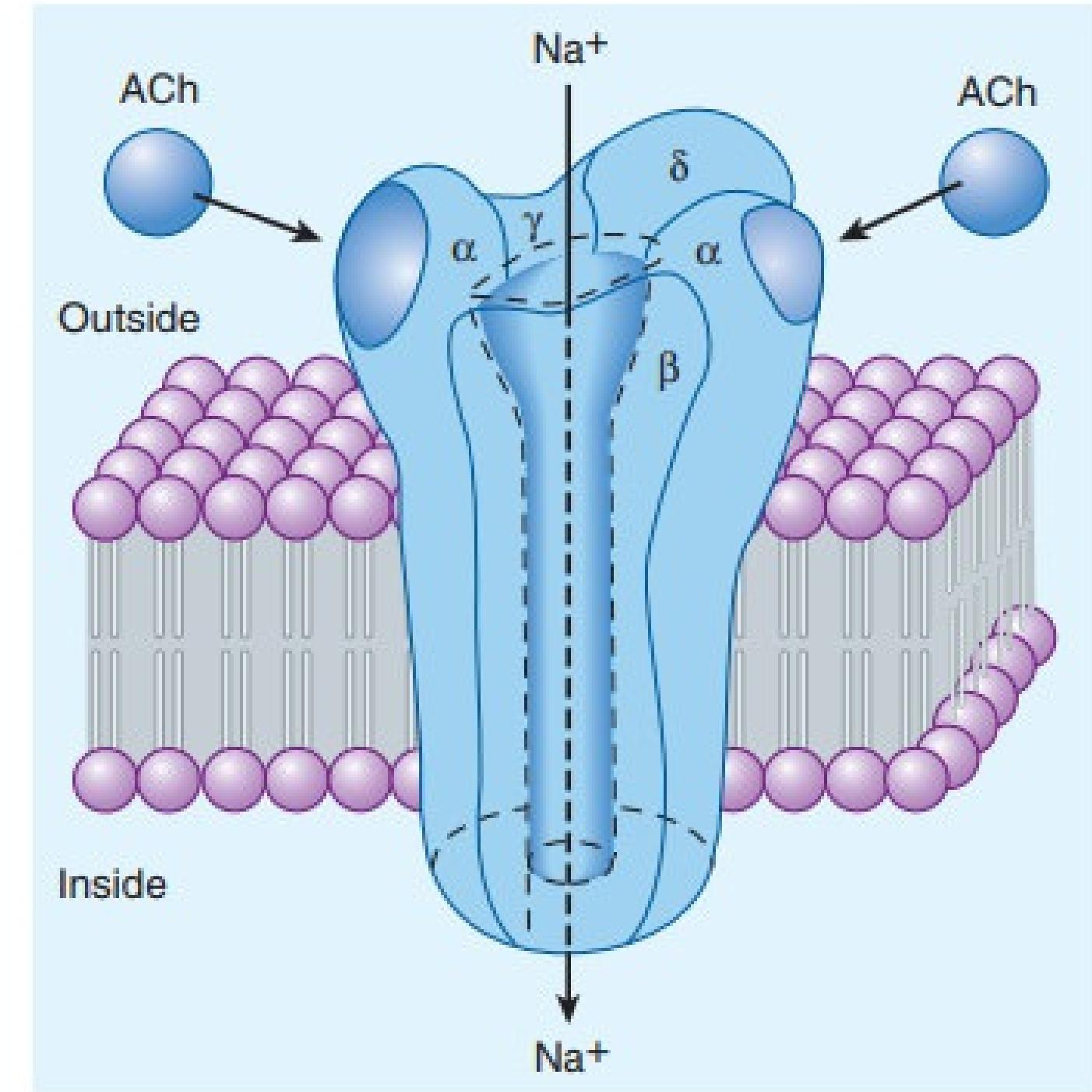
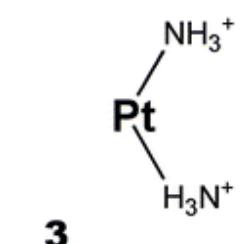
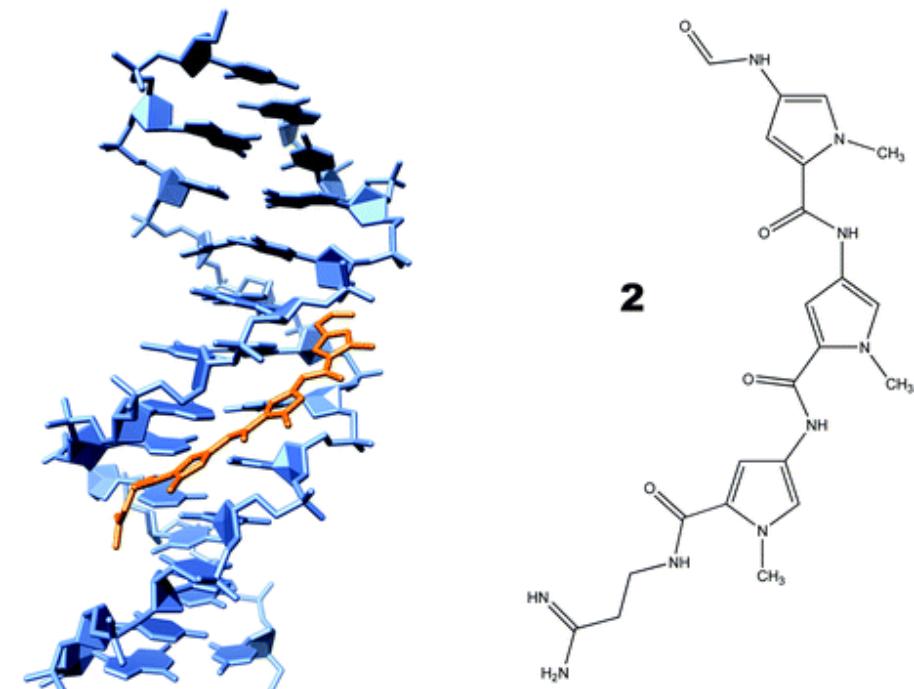
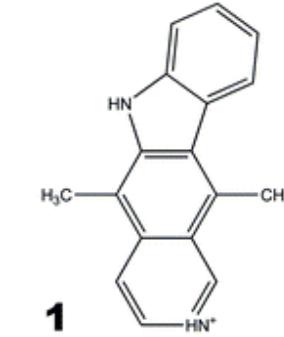
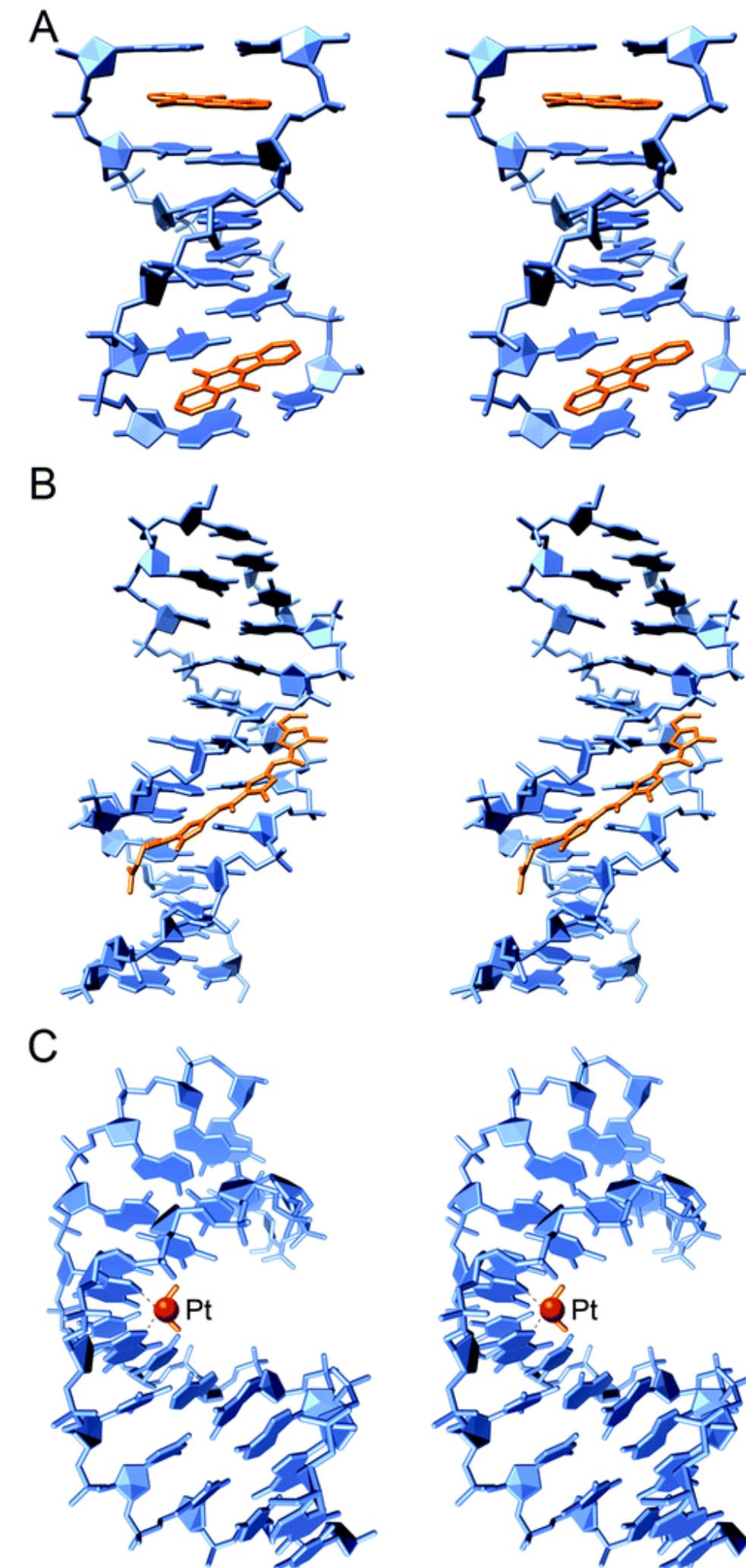


FIGURE 2–9 The nicotinic acetylcholine (ACh) receptor, a ligand-gated ion channel. The receptor molecule is depicted as embedded in a rectangular piece of plasma membrane, with extracellular fluid above and cytoplasm below. Composed of five subunits (two α , one β , one γ , and one δ), the receptor opens a central transmembrane ion channel when ACh binds to sites on the extracellular domain of its α subunits.

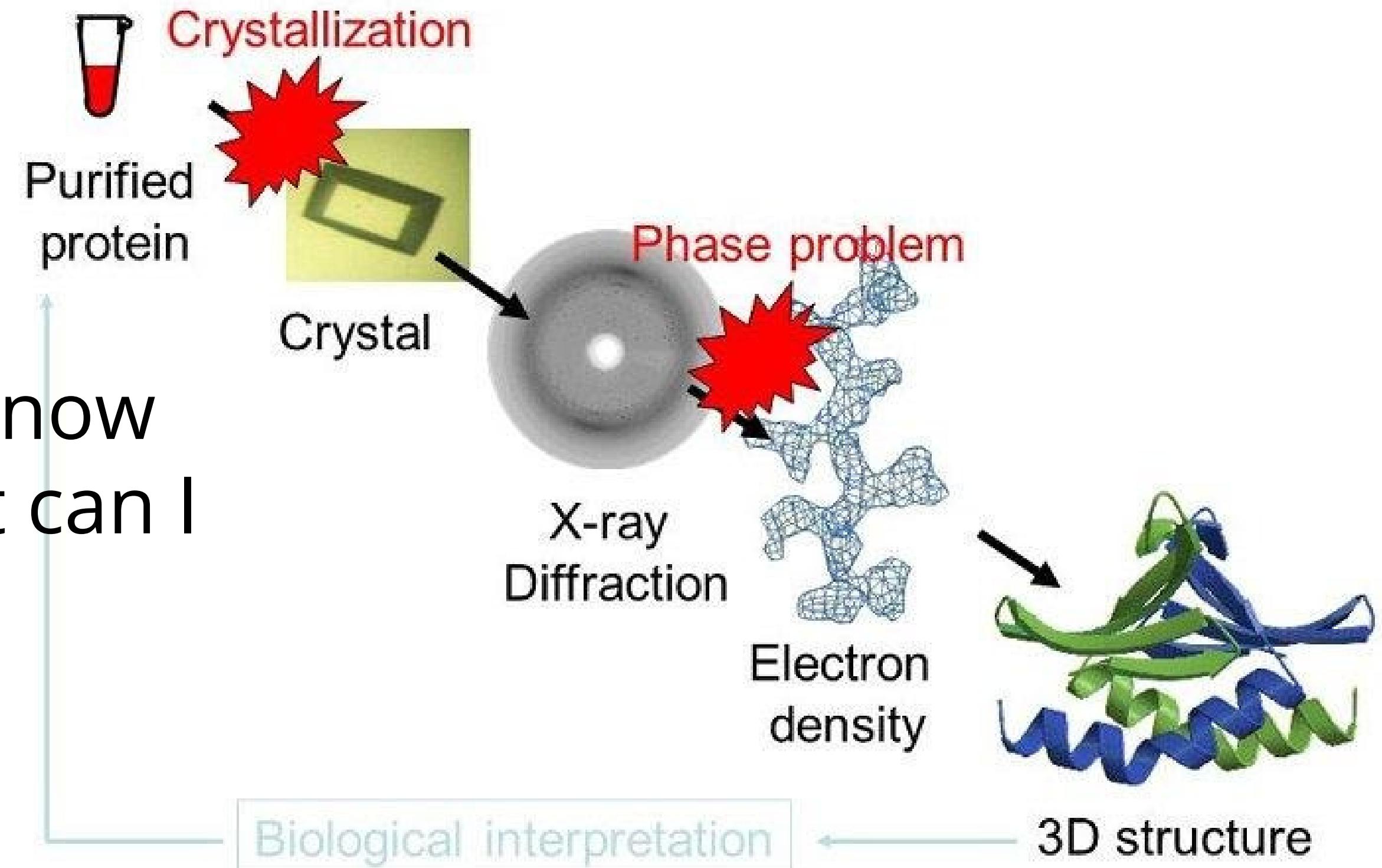
Drug target IV

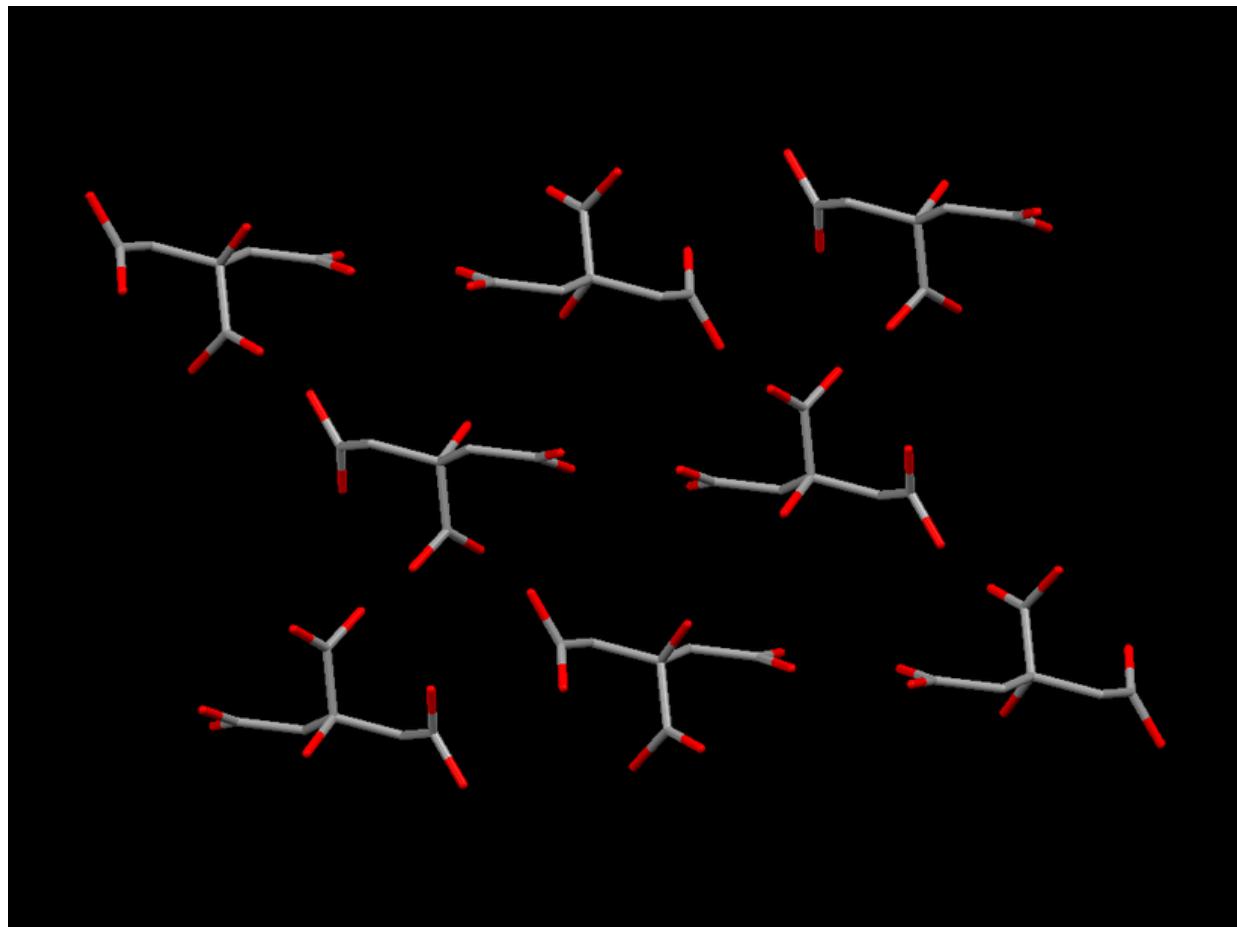
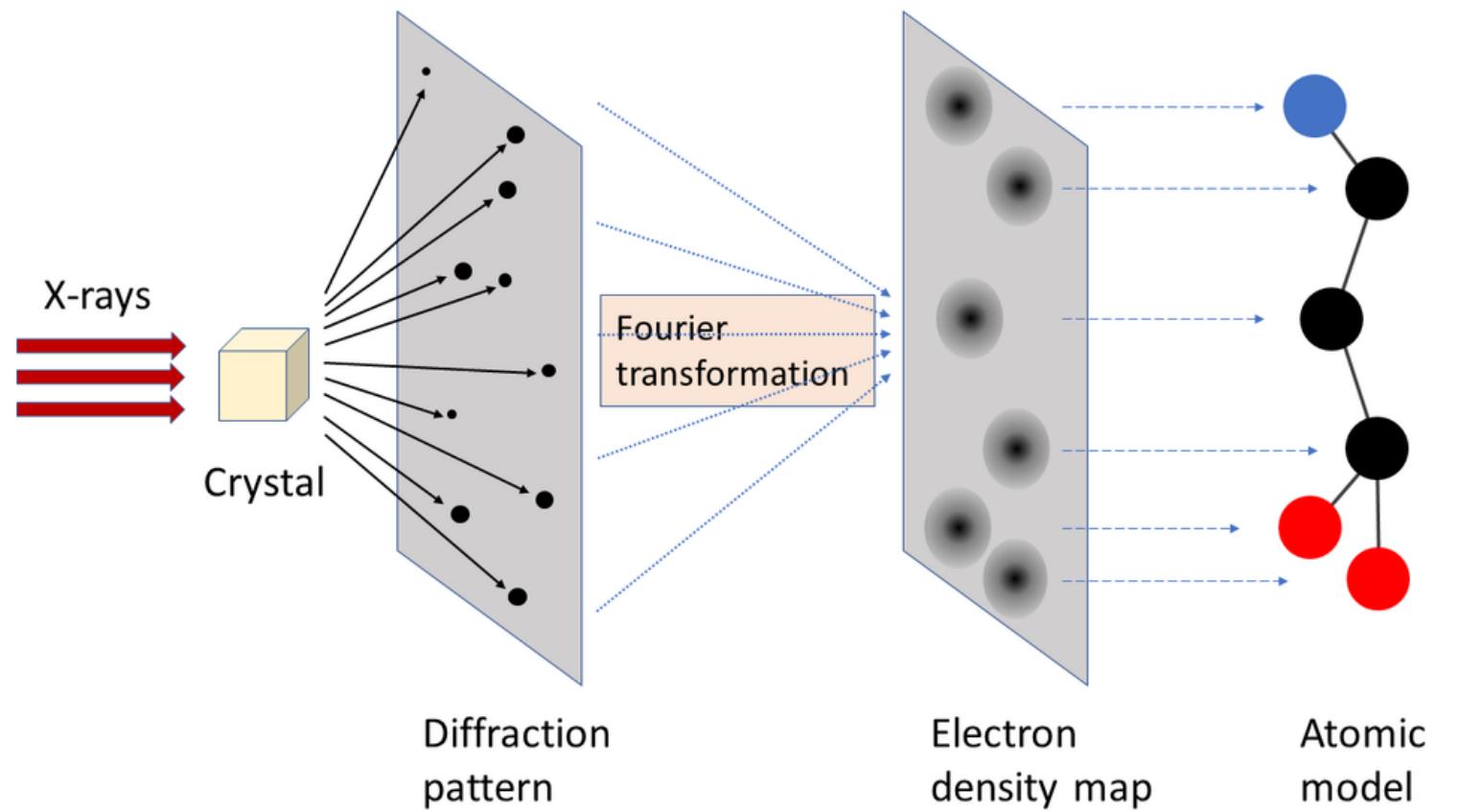
Nucleic Acids (DNA, RNA)



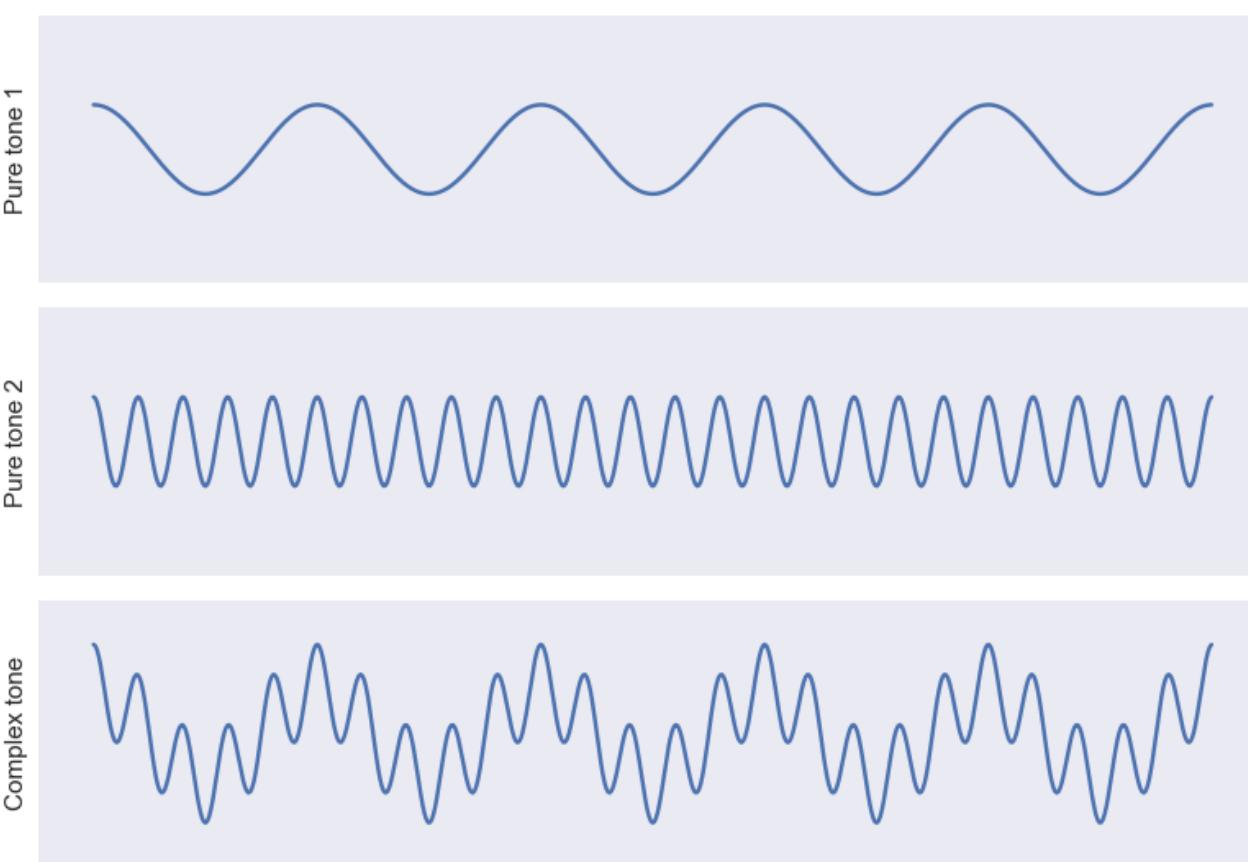
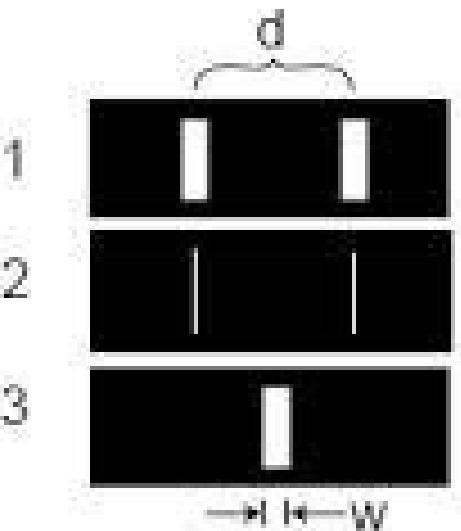
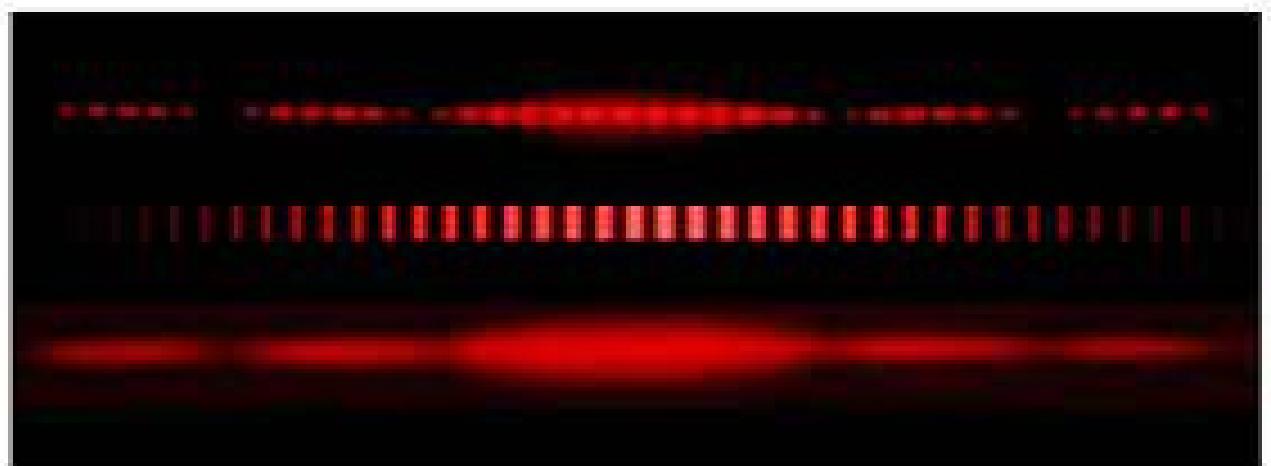
Structure determination method X-ray crystallography

OK, now that I know my target. What can I do with it?

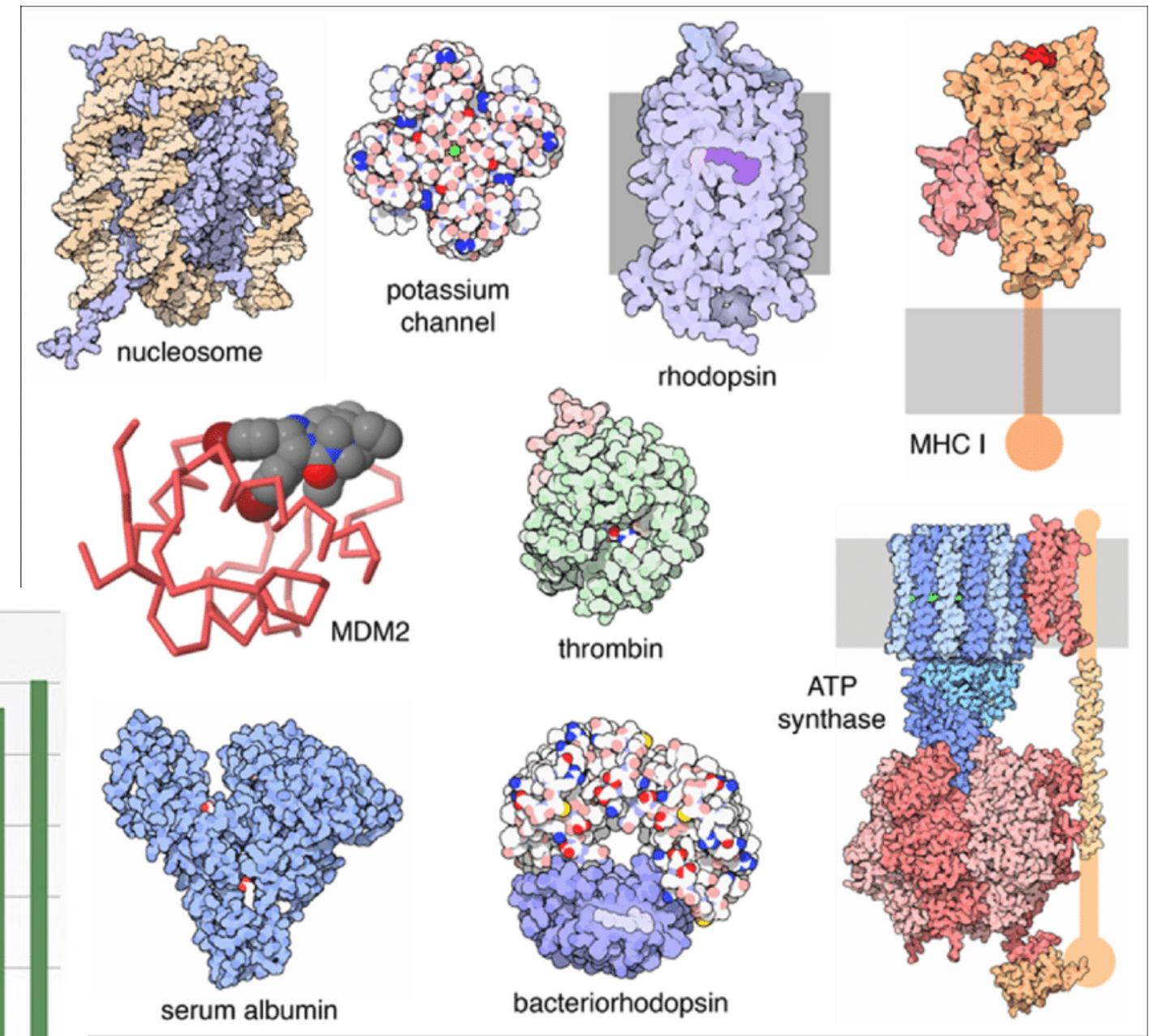
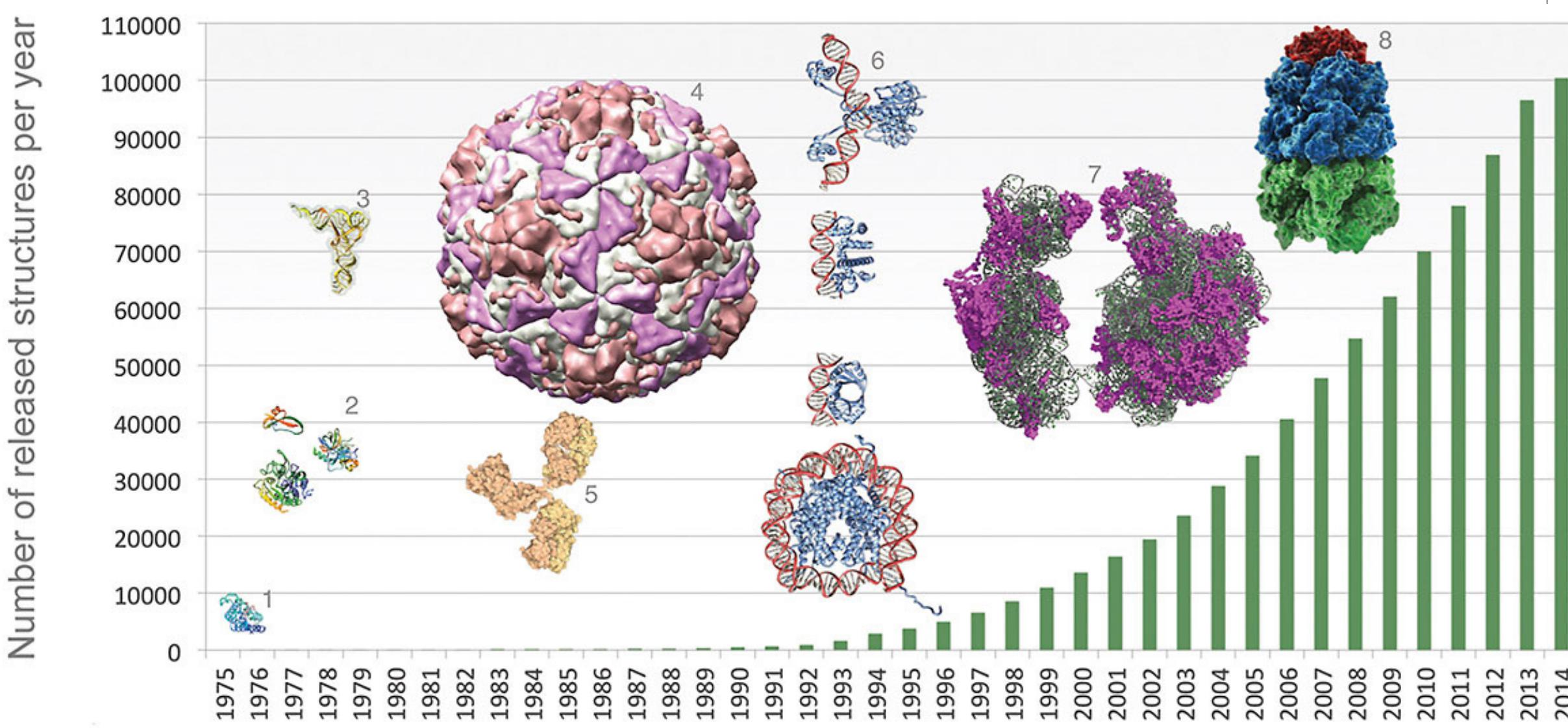




X-ray crystallography?



Voilà!



Left sidebar (blue background):

- RCSB PDB
- Deposit ▾
- Search ▾
- Visualize ▾
- Analyze ▾
- Download ▾
- Learn ▾
- More ▾
- Documentation ▾
- Careers

Right sidebar (white background):

- MyPDB ▾
- Contact us

Top right corner icons:

- Search icon
- Download icon
- Upload icon
- Print icon
- Help icon
- Logout icon

Main content area:

RCSB PDB PROTEIN DATA BANK

199,803 Structures from the PDB
1,000,361 Computed Structure Models (CSM)

3D Structures [Advanced Search](#) | [Browse Annotations](#)

Enter search term(s), Entry ID(s), or sequence

Include CSM [Search](#)

[Help](#)

Logos: RCSB PDB-101, wwPDB, EMDataResource, Nucleic Acid Database, wwPDB Foundation

Social media icons: Facebook, Twitter, YouTube, GitHub

Announcement banner: NEW! Computed Structure Models (CSM) [Learn more](#)

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

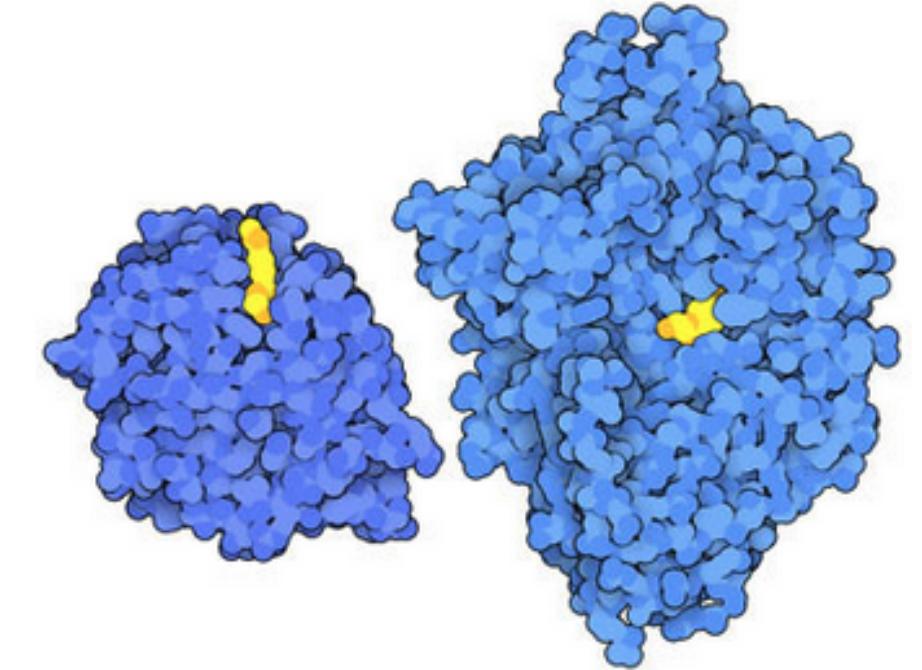
- Experimentally-determined 3D structures from the **Protein Data Bank (PDB)** archive
- Computed Structure Models (CSM)** from AlphaFold DB and ModelArchive

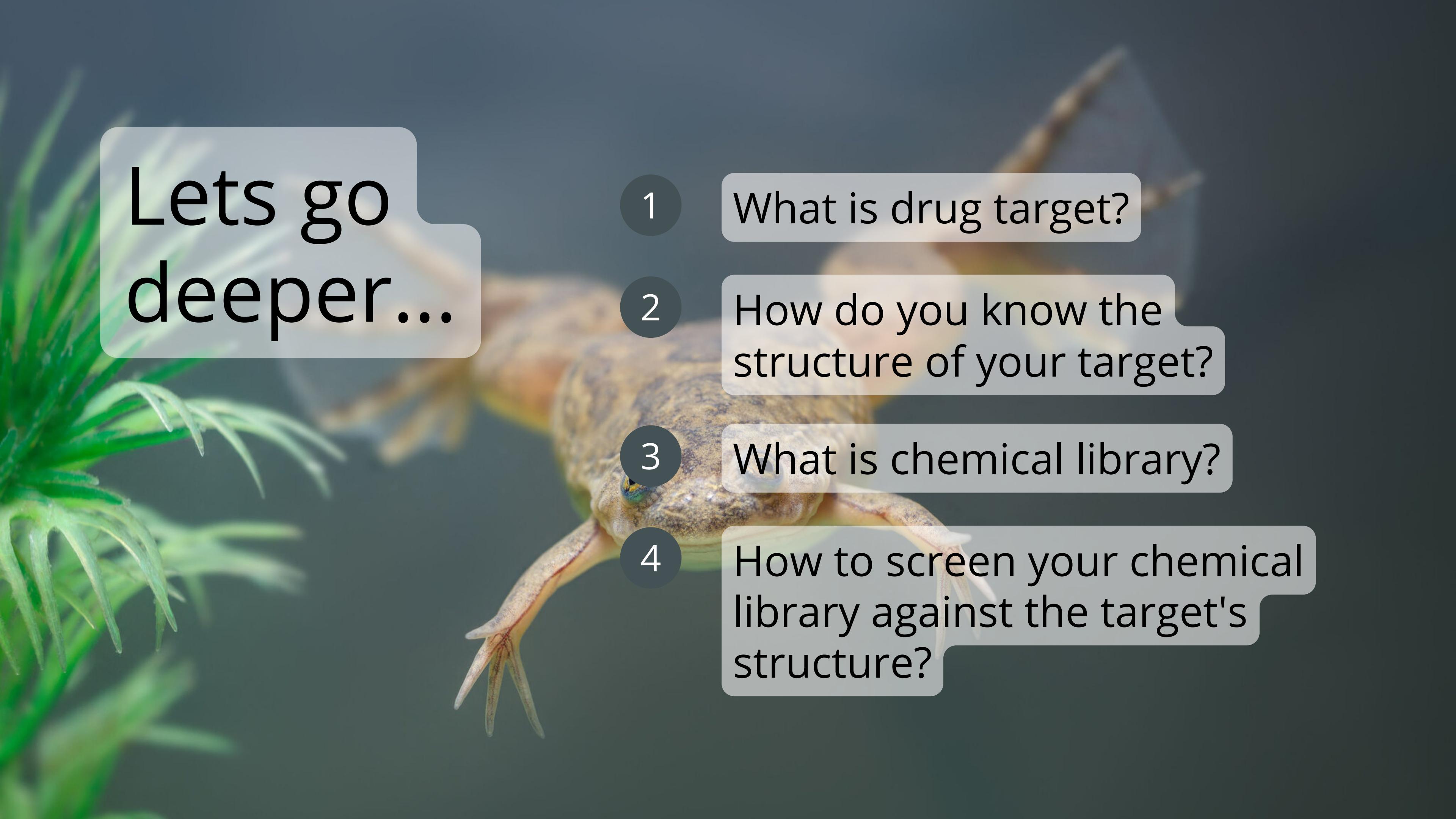
These data can be explored in context of external annotations providing a structural view of biology.

COVID-19 CORONAVIRUS Resources

MOLECULAR LANDSCAPES FOR 2023

Plastic-eating Enzymes





Lets go
deeper...

1

What is drug target?

2

How do you know the
structure of your target?

3

What is chemical library?

4

How to screen your chemical
library against the target's
structure?

What is a chemical library?

A database of chemical compound (may contain thousands or even millions compound)

Curated with its properties, and structure 1D, 2D, or even 3D.

A

ZINC1590 (Isoniazid)
In: in-stock metabolites for-sale bb fda

Google Wikipedia PubMed

Added	Seen	Purchasability	Since	Mwt	logP	Heavy Atoms	Tranche	Download
2015-08-07	2015-09-04	Premier	2015-08-07	137.142	-0.315	10	ABDA	Download

SMILES: NNC(=O)c1ccncc1
InChI: InChI=1S/C6H7N3O/c7-9-6(10)5-1-3-8-4-2-5/h1-4H,7H2,(H,9,10)
InChI Key: QRXWMOHMRWLFEY-UHFFFAOYSA-N

Available 3D Representations

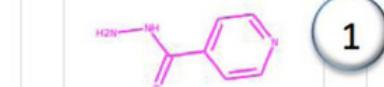
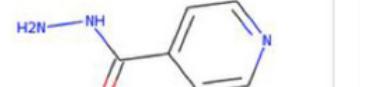
pH range	Net charge	H-bond donors	H-bond acceptors	tPSA	Rotatable bonds	Apolar desolvation	Polar desolvation	Download
Reference	0	2	3	68	1	-0.55	-11.63	Download

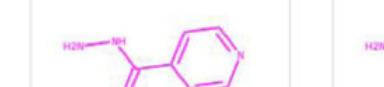
Vendors (62 Total):
ChemDiv 0272-0055
Frontier Scientific Services 500012803

Annotated Catalogs (50 Total):
MicroSource Spectrum 01500355
MicroSource US 01500355

B

Interesting Analogs

Endogenous	Metabolites	Natural Products	Aggregators	ZINC1590
None Found Similar Endogenous	ZINC1590 Isoniazid  Identity	ZINC1590 Isoniazid  Identity	None Found Similar Aggregators	 3

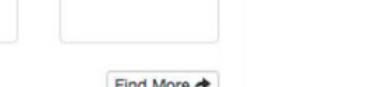
Drugs: ZINC1590 Isoniazid

2

In Man: ZINC1590 Isoniazid

2

Bioactives: ZINC1590 Isoniazid

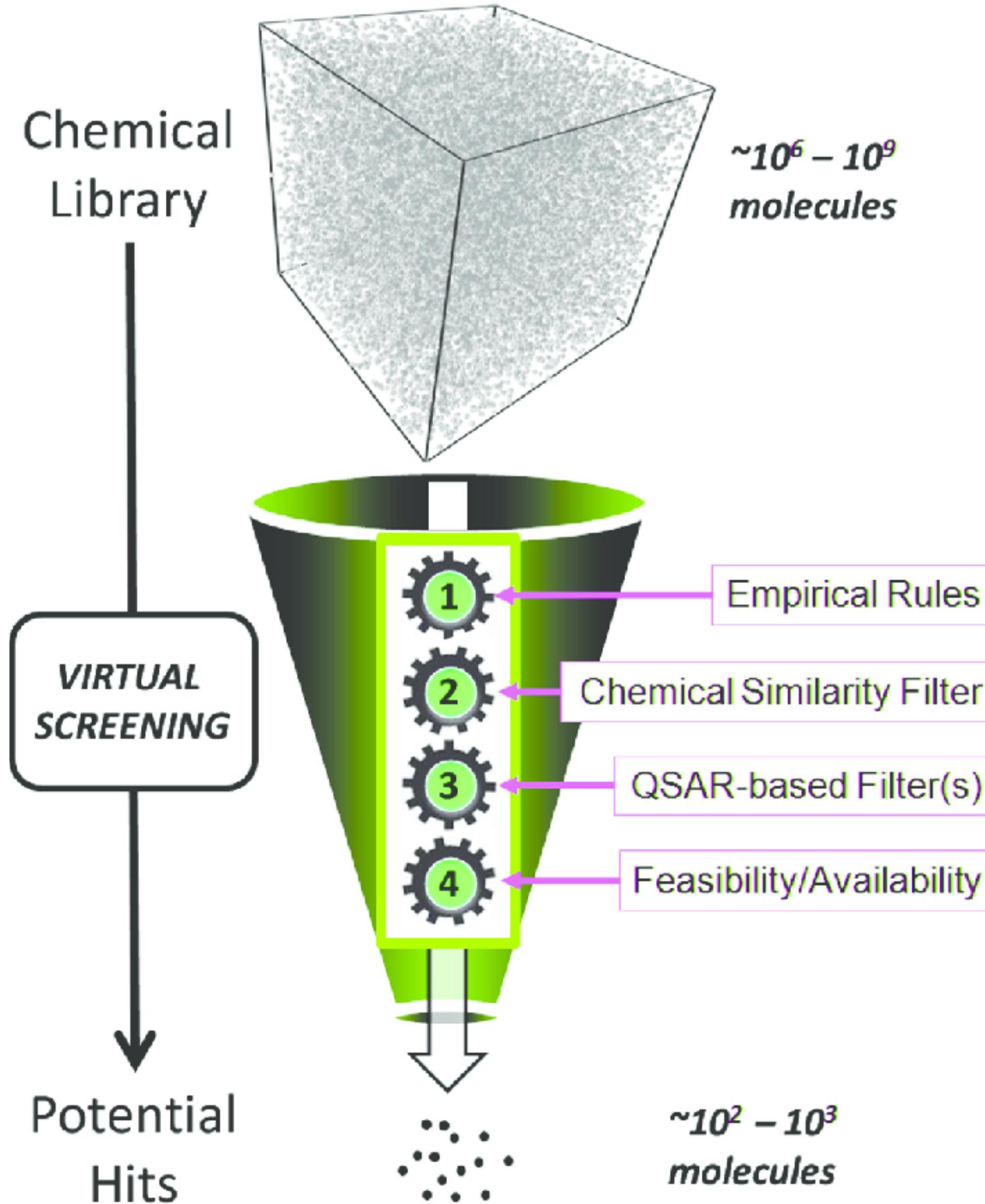
5

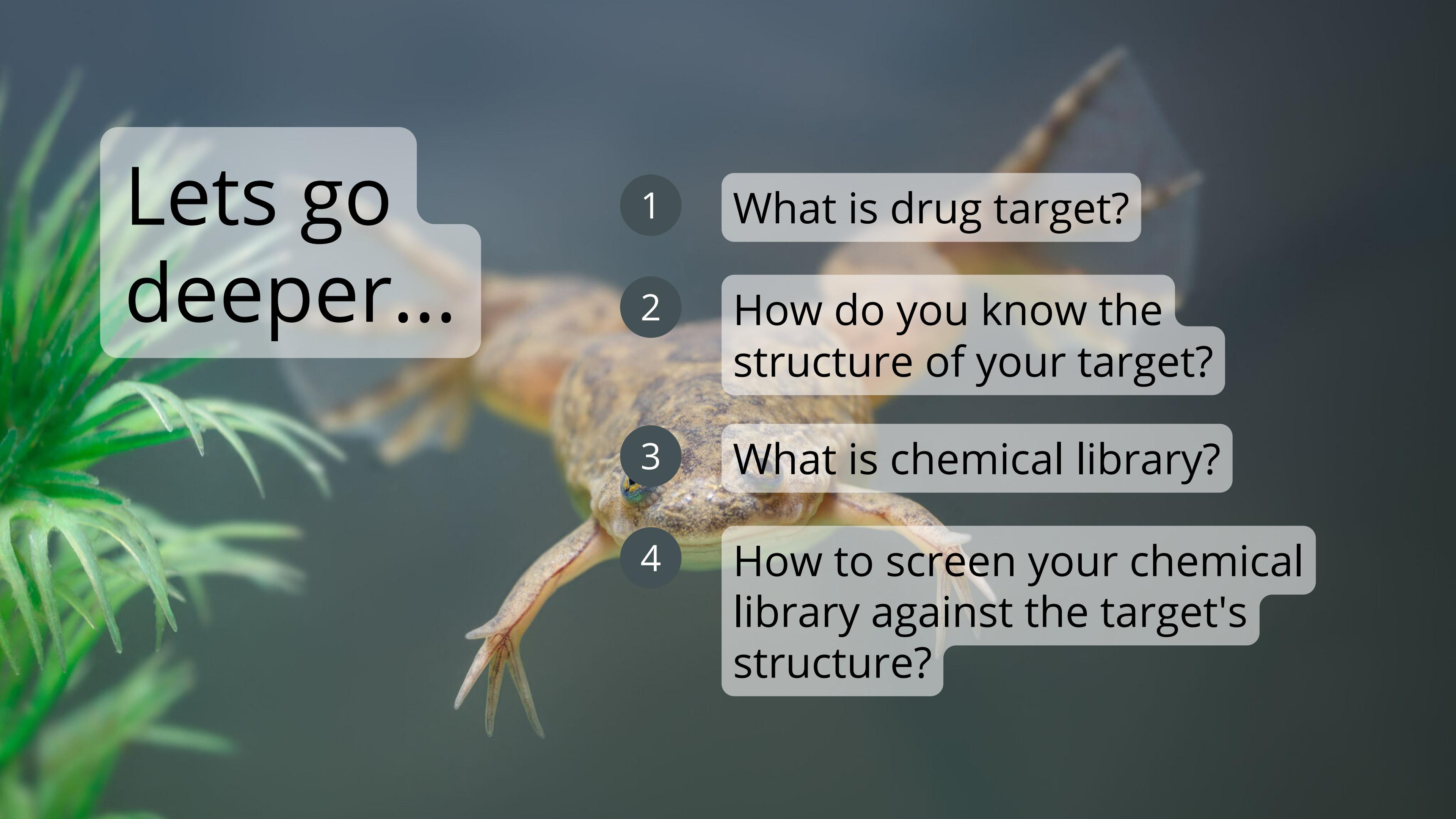
Purchasable: ZINC1590 Isoniazid

137786322
137786322

Scaffold of this compound

What is a chemical library?

Can be filtered through screening process
--> To generate hits/leads





Lets go
deeper...

1

What is drug target?

2

How do you know the
structure of your target?

3

What is chemical library?

4

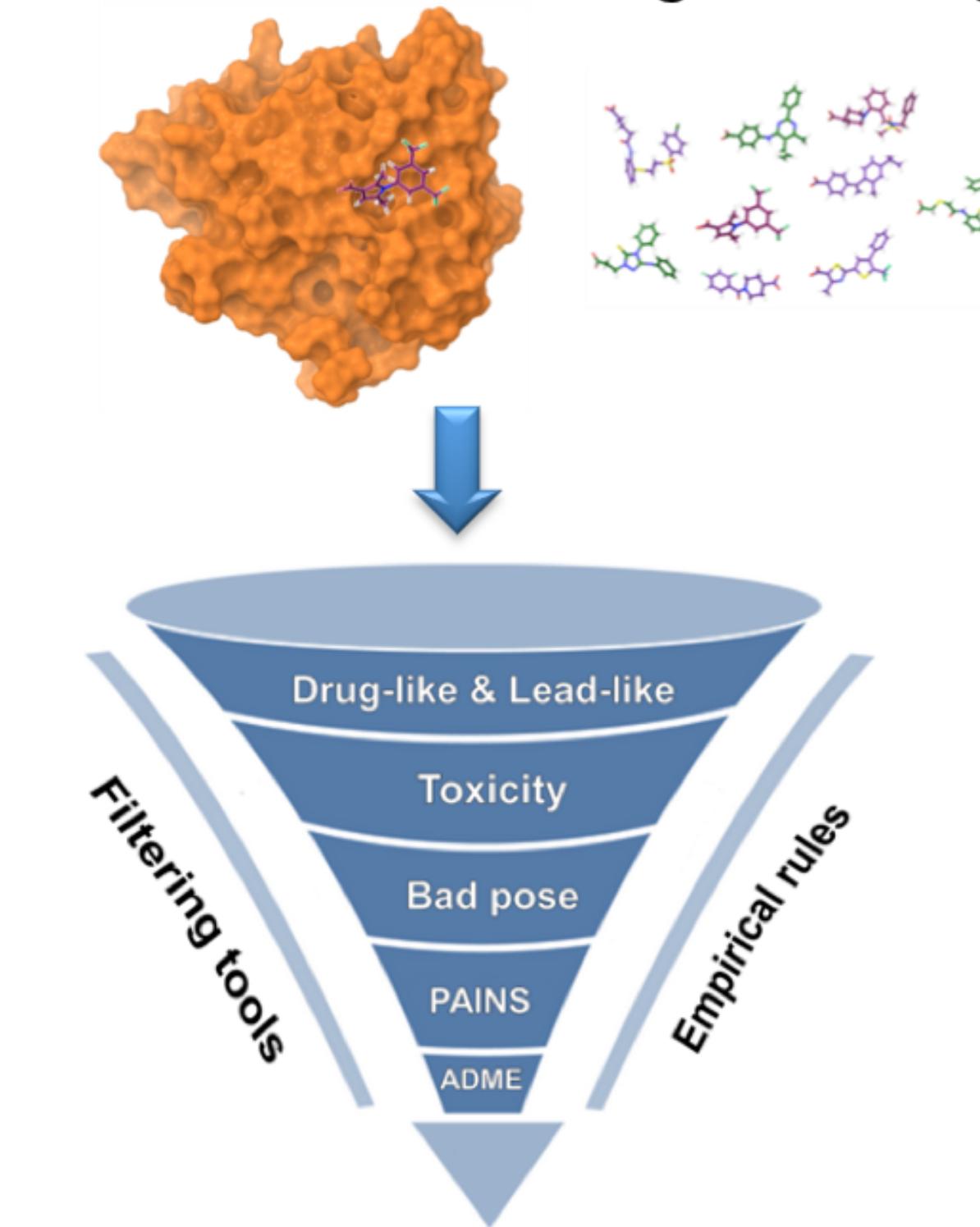
How to screen your chemical
library against the target's
structure?

Virtual Screening & Scoring

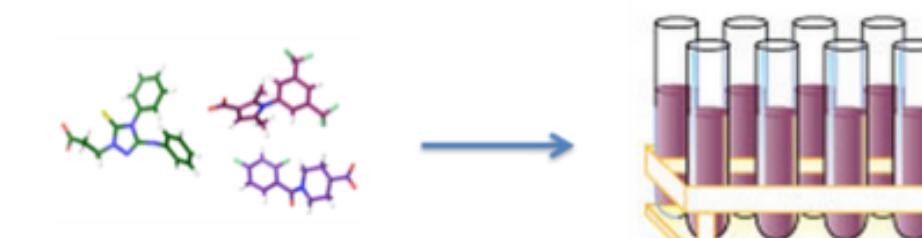
How to screen your **chemical library** against the **target's structure?**

We can group screening into two category:

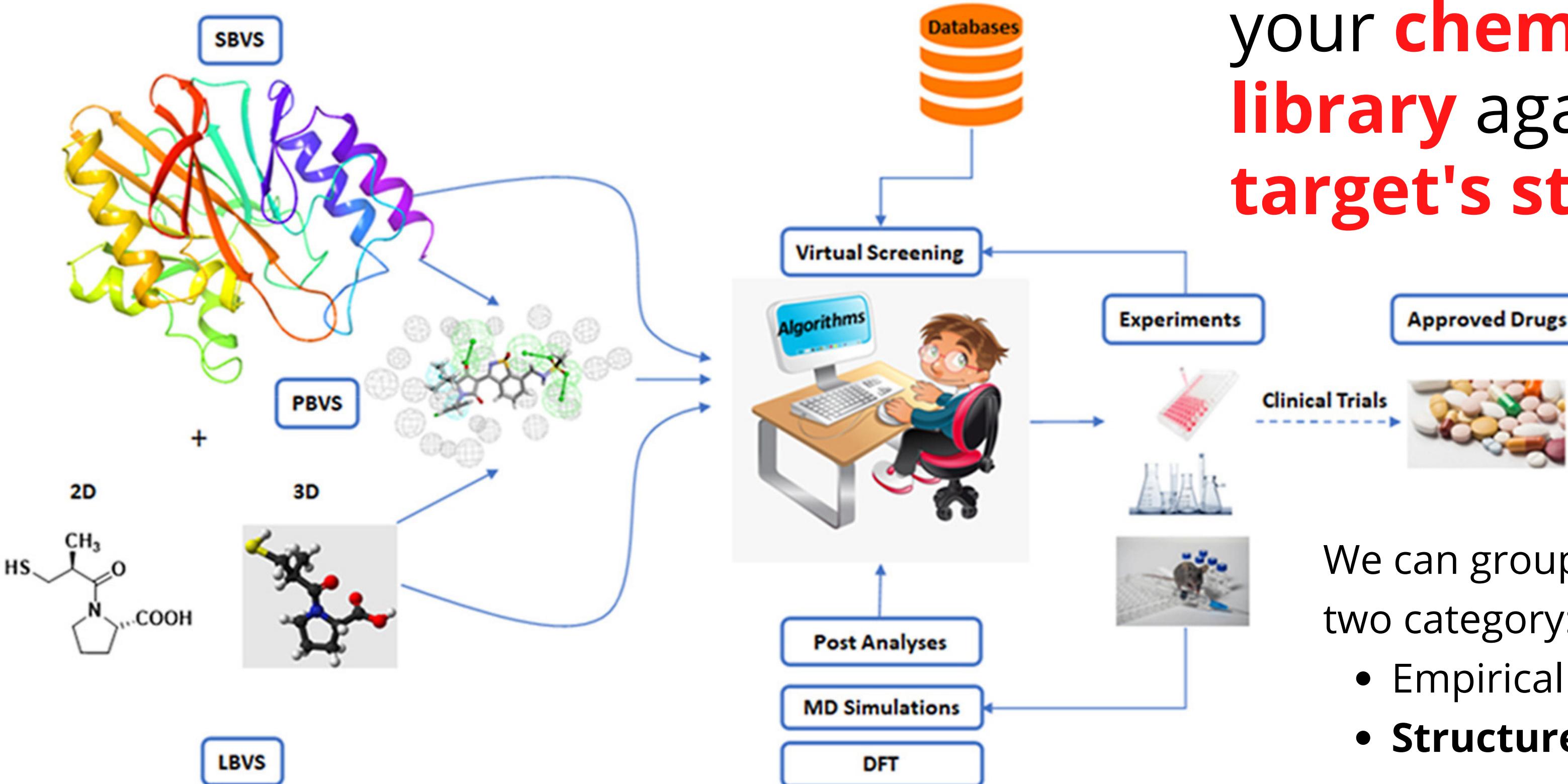
- **Empirical rules**
- Structure-based



Re-ranking Virtual Screening results
Compound selection & *in vitro* assays



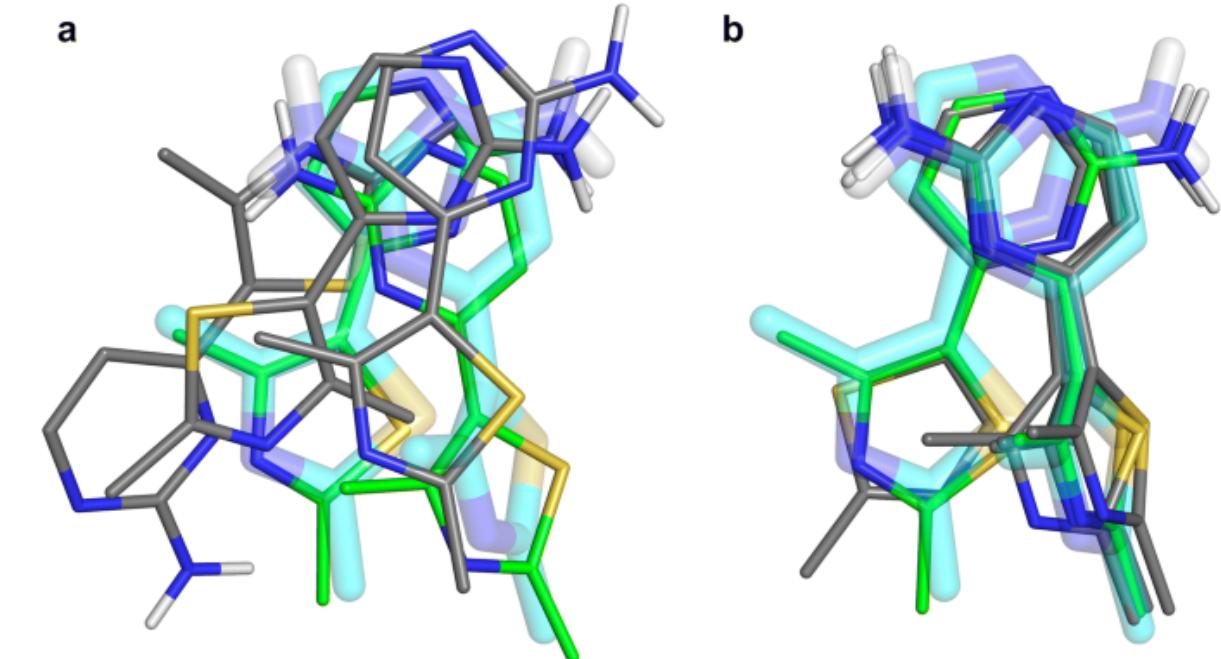
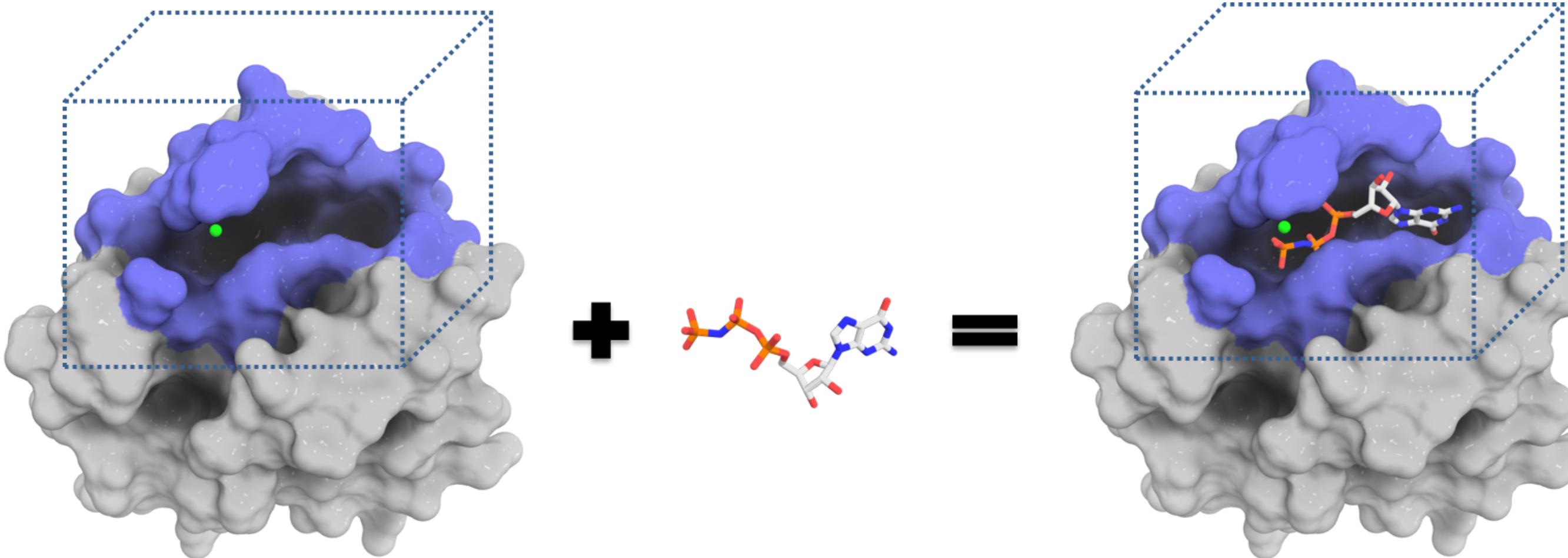
How to screen your **chemical library** against the target's structure?



We can group screening into two category:

- Empirical rules
- **Structure-based**

How to do structure-based virtual screening?



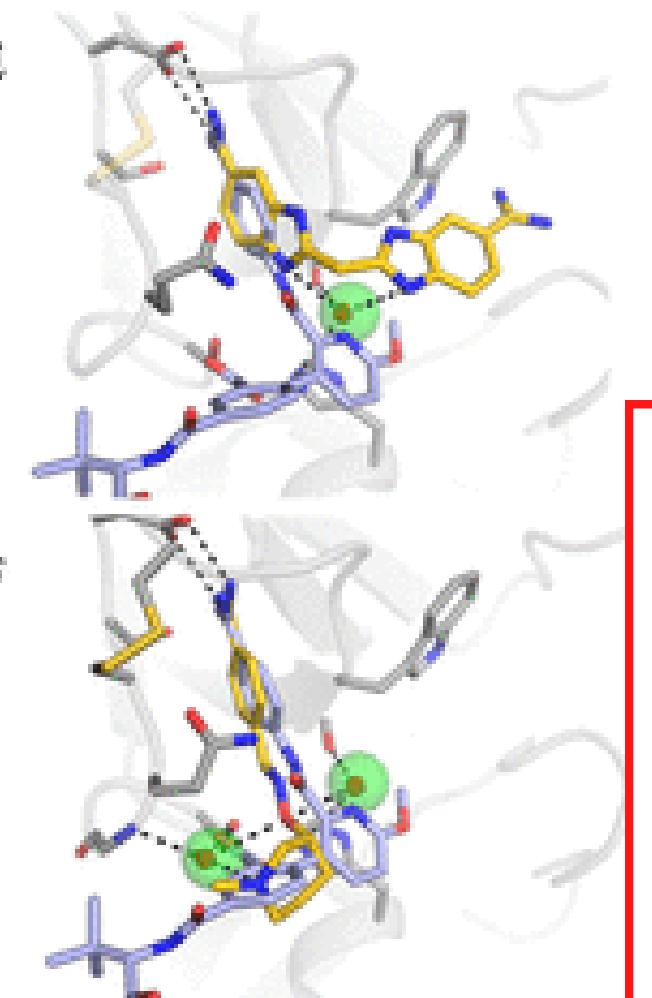
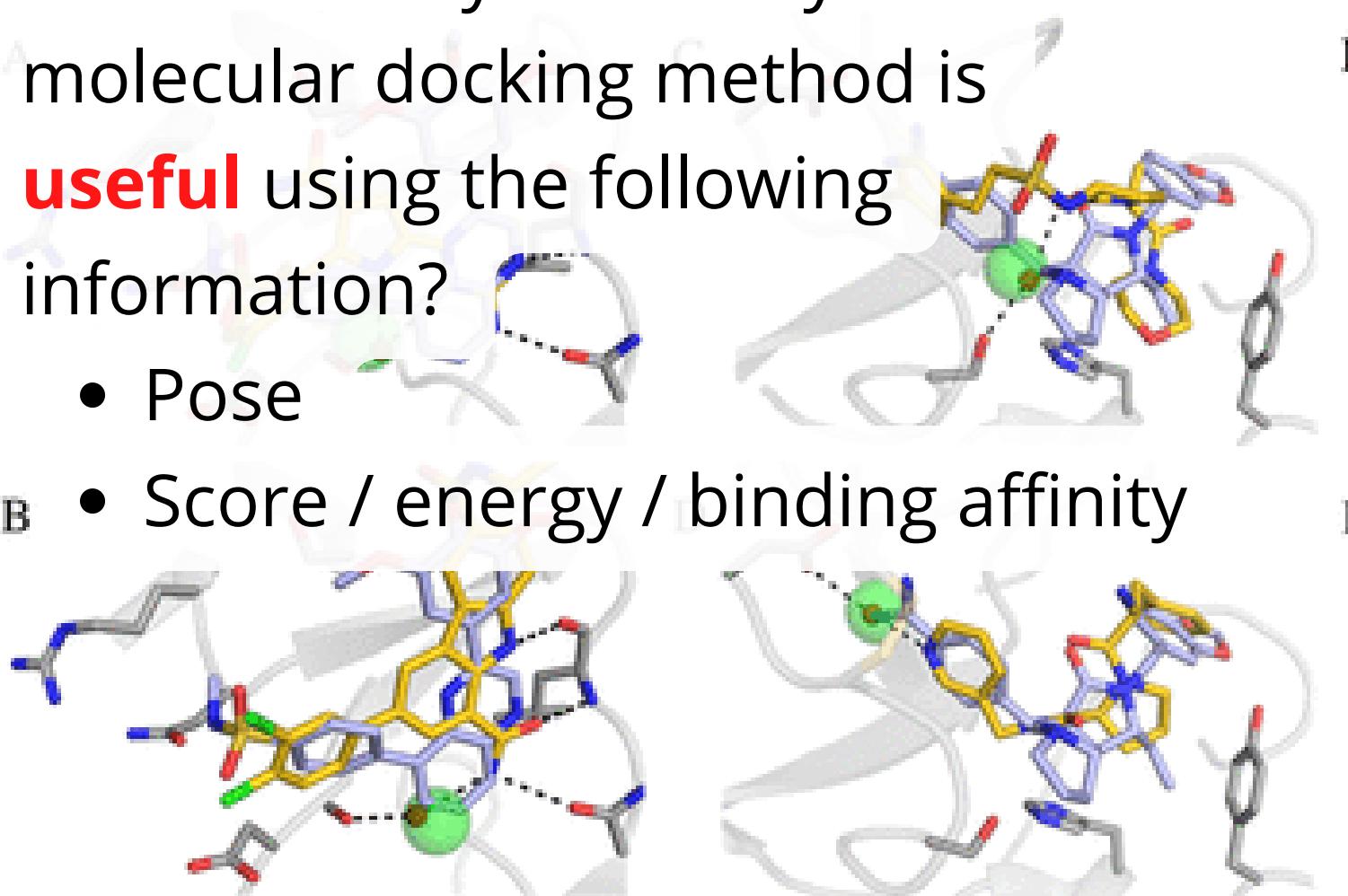
molecular docking
will generate:

- Pose
- Score / energy /
binding affinity

How to do structure-based virtual screening?

then how do you know your molecular docking method is **useful** using the following information?

- Pose
- Score / energy / binding affinity



```
#####
# If you used AutoDock Vina in your work, please cite:      #
#                                                       #
# J. Eberhardt, D. Santos-Martins, A. F. Tillack, and S. Forli   #
# AutoDock Vina 1.2.0: New Docking Methods, Expanded Force     #
# Field, and Python Bindings, J. Chem. Inf. Model. (2021)      #
# DOI 10.1021/acs.jcim.1c00203                                #
#                                                       #
# O. Trott, A. J. Olson,                                         #
# AutoDock Vina: improving the speed and accuracy of docking    #
# with a new scoring function, efficient optimization and       #
# multithreading, J. Comp. Chem. (2010)                          #
# DOI 10.1002/jcc.21334                                         #
#                                                       #
# Please see https://github.com/ccsb-scripps/AutoDock-Vina for          #
# more information.                                              #
#####
```

```
Scoring function : vina
Rigid receptor: 6cox_receptor.pdbqt
Ligand: 6cox_ligand.pdbqt
Grid center: X 23.665 Y 23.312 Z 47.826
Grid size : X 20 Y 20 Z 20
Grid space : 0.375
Exhaustiveness: 32
CPU: 0
Verbosity: 1
```

```
Computing Vina grid ... done.
Performing docking (random seed: -1699093488) ...
0% 10 20 30 40 50 60 70 80 90 100%
|-----|-----|-----|-----|-----|-----|-----|-----|
*****
```

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	dist from best mode rmsd u.b.
1	-10.65	0	0
2	-8.537	4.05	5.611
3	-8.239	4.966	6.407
4	-7.556	5.185	7.071
5	-6.793	5.38	7.191
6	-6.419	3.884	5.149
7	-4.388	1.73	2.389
8	-2.38	8.768	10.31
9	-1.181	3.663	4.893

How to do structure-based virtual screening?

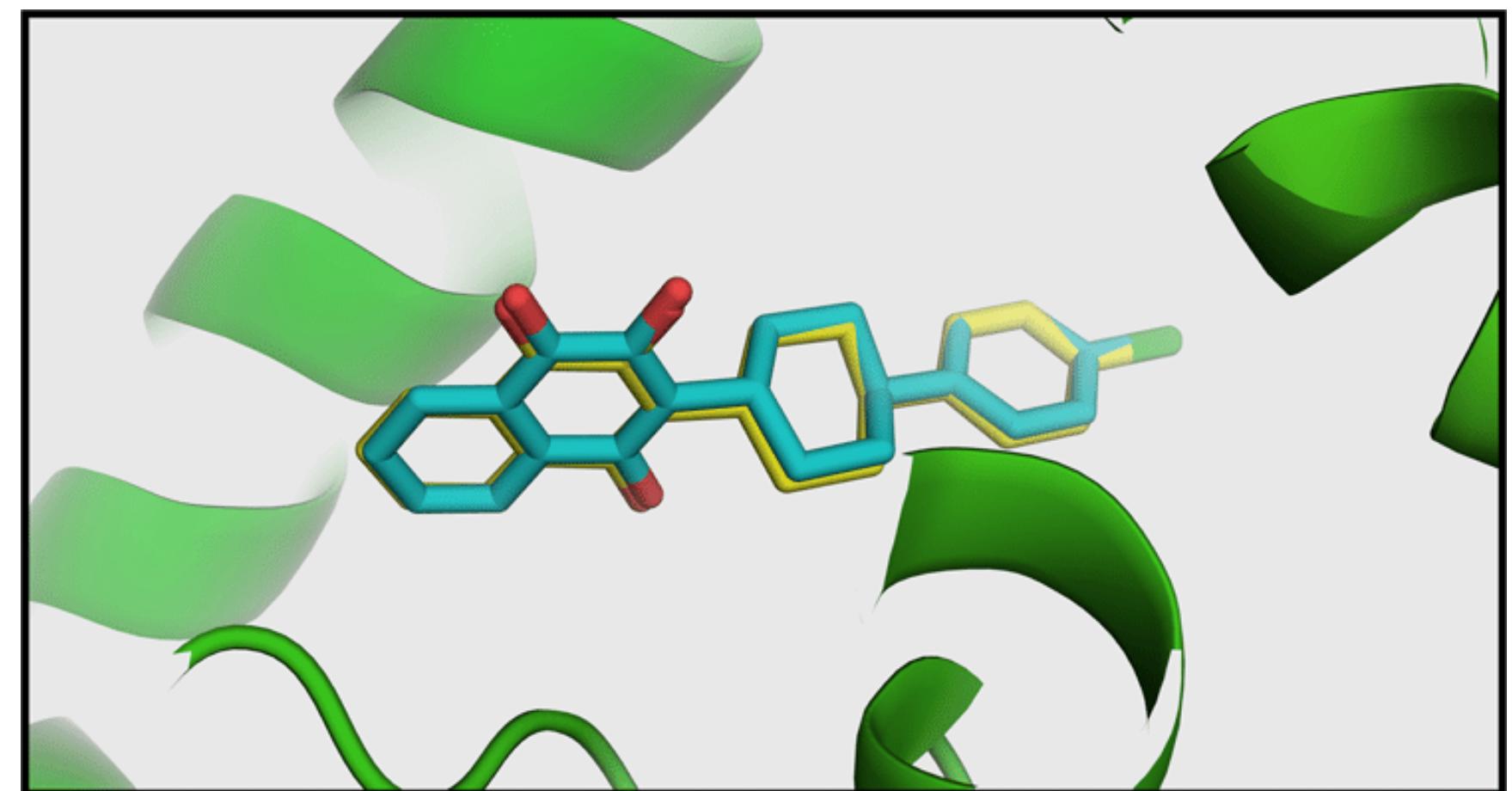
then how do you know your molecular docking method is **useful** using the following information?

- Pose
- Score / energy / binding affinity



Easiest way: Redock

Rule of thumb: RMSD < 2 Angstrom



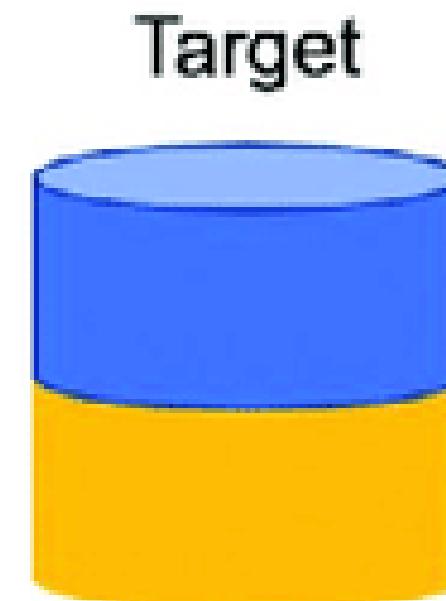
How to do structure-based virtual screening?

then how do you know your molecular docking method is **useful** using the following information?

- Pose
- Score / energy / binding affinity

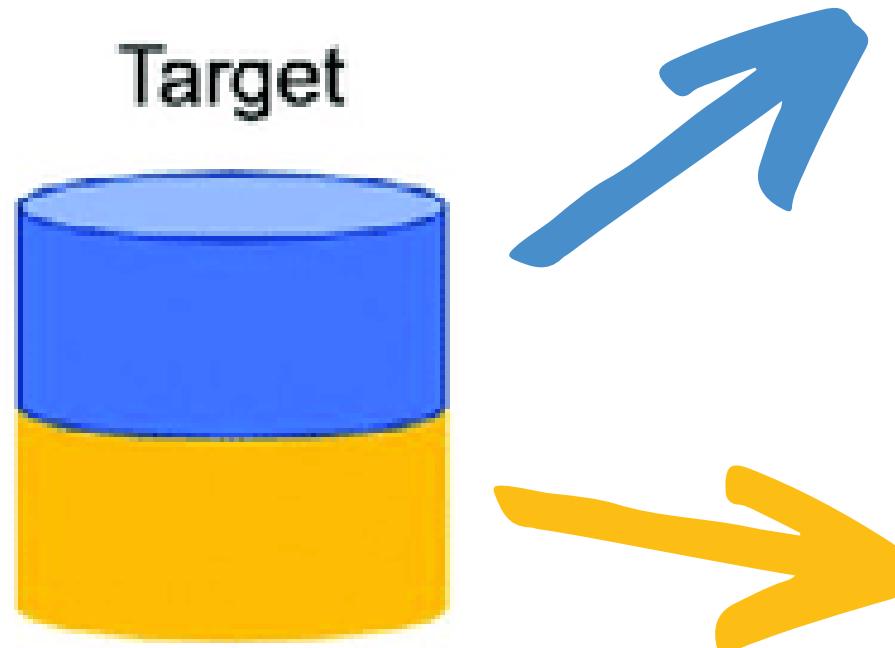


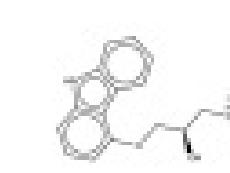
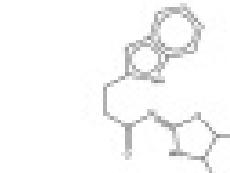
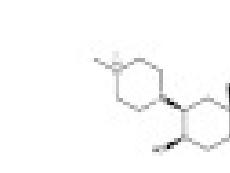
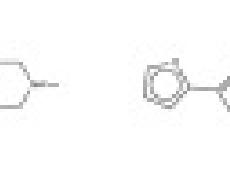
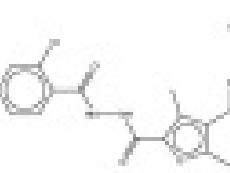
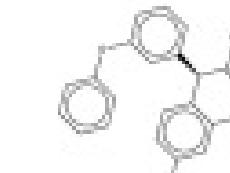
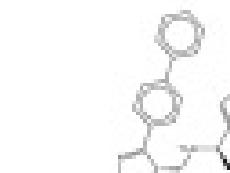
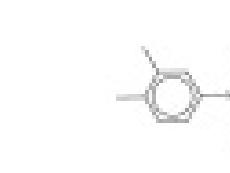
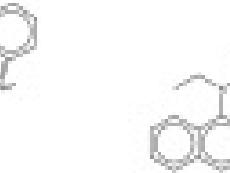
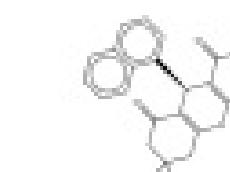
We can do better:
Retrospective validation
Use **decoy** and **known active** to test our model



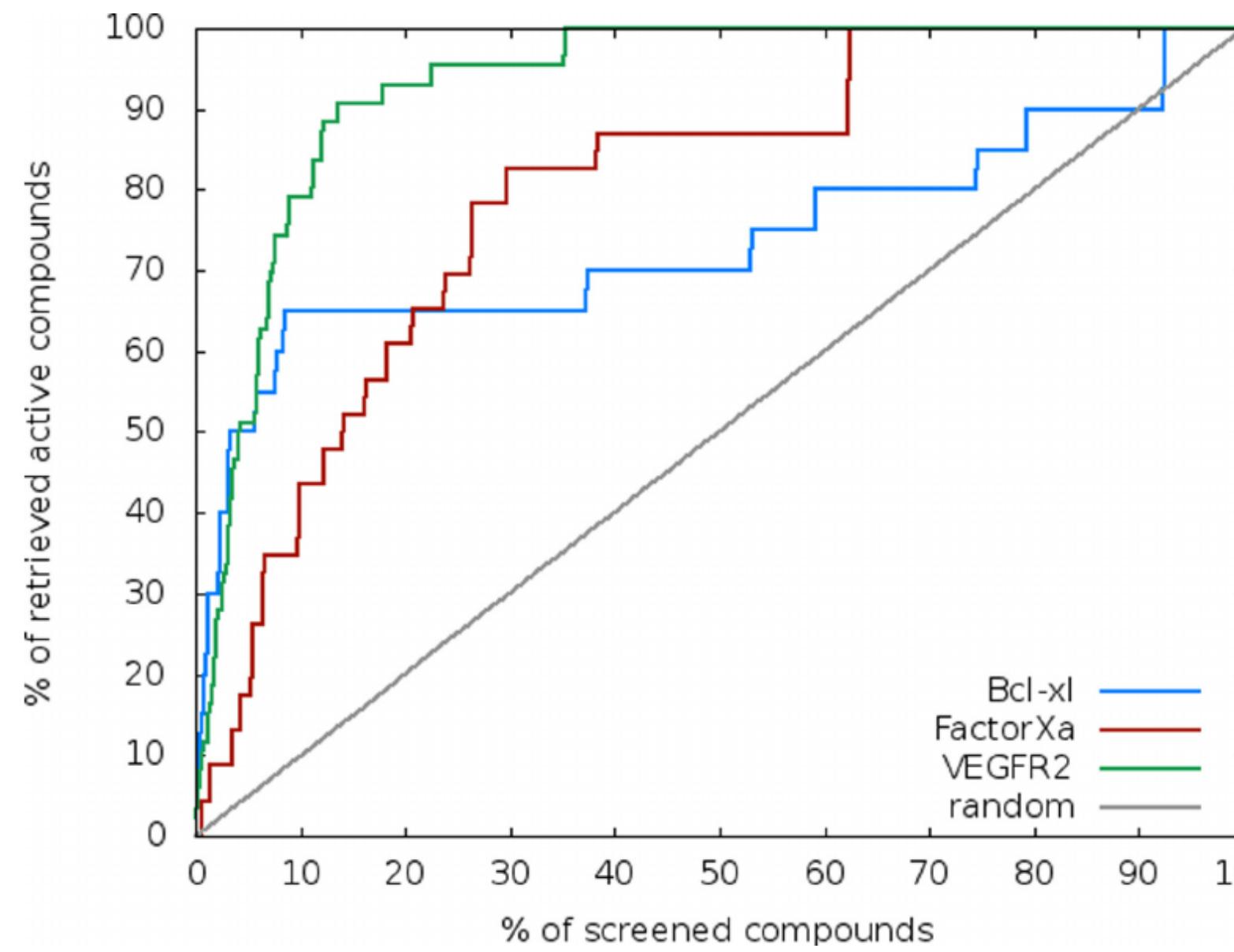
Decoys?

Any chemicals retrieved from chemical libraries that have the **similar properties** with known actives.



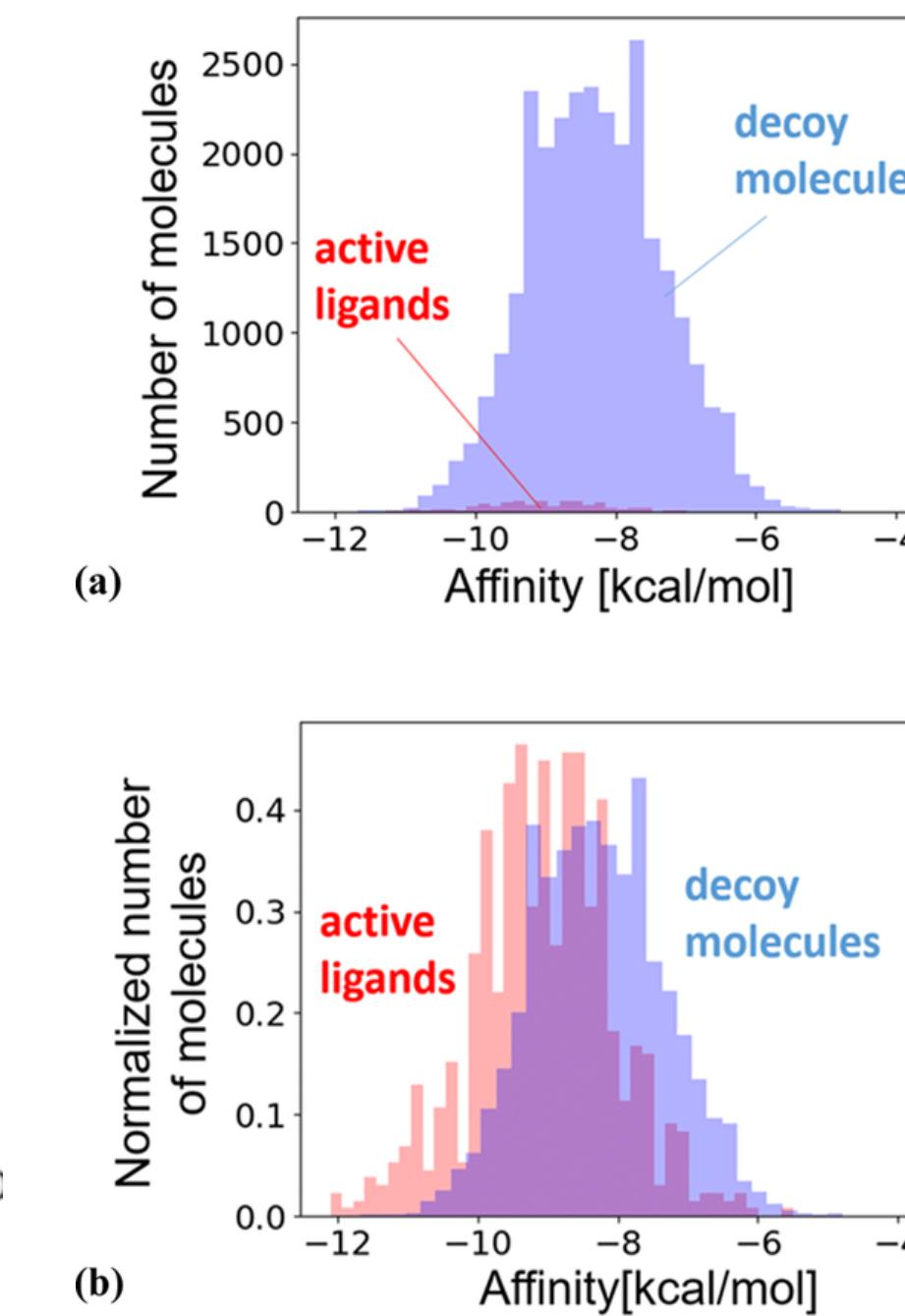
	AA2AR	ADRB1	ADRB2	CXCR4	DRD3
Actives	 P91159500	 P91036100	 P91101100	 P9100100	 P92006600
Decoys	 P97081816	 P99400250	 P74036011	 P53898657	 P54577091
	 P69477083	 P197024550	 P59891568	 P97463531	 P54100299
	 P59108886	 P89156581	 P13915327	 P13498808	 P99103717

How to use decoys in retrospective validation?



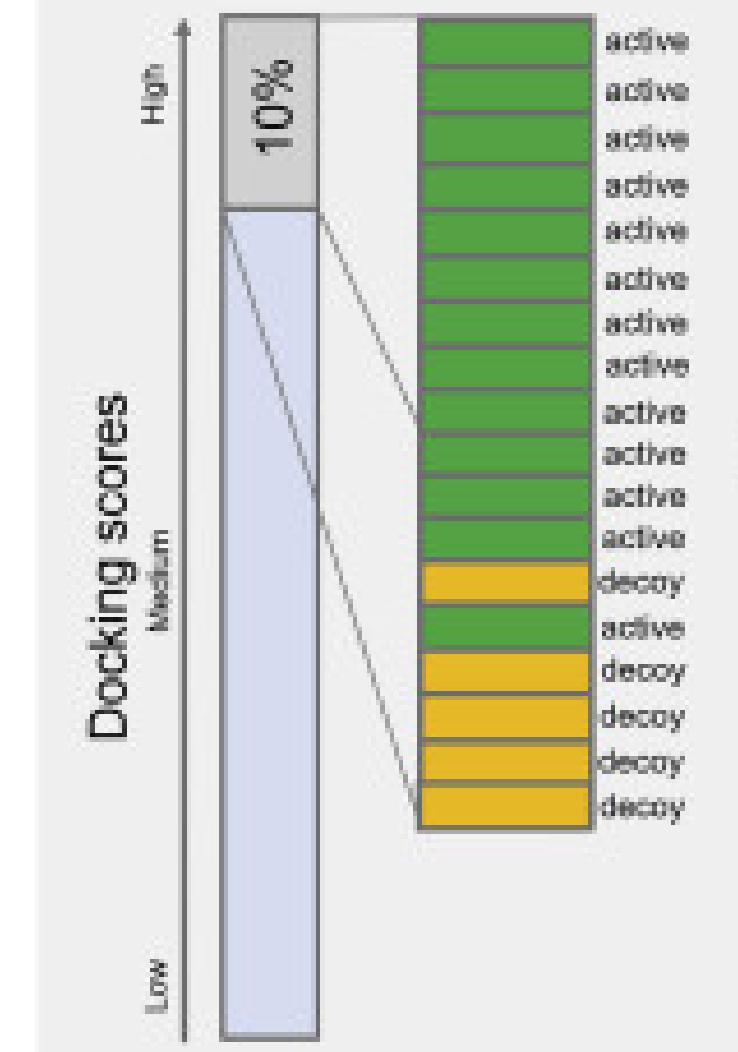
Usually 40-400 decoys per known active.

E.g. using 300 actives and 15000 decoys in a test set.



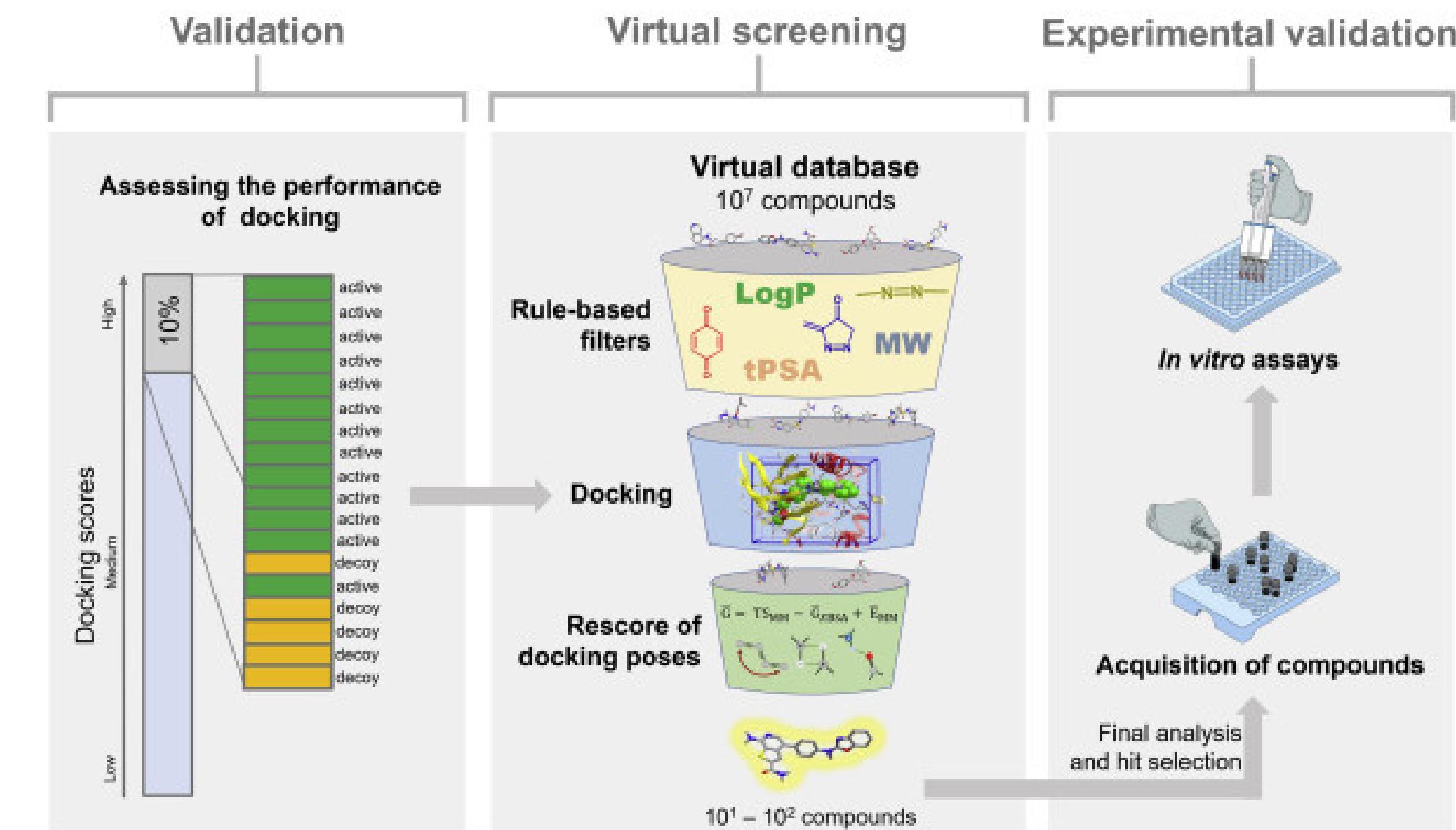
Validation

Assessing the performance of docking



My Virtual Screening protocol seems valid, what next?

- Use it on Prospective Application
 - Combine with empirical rule-based screening
 - Validate the result with in vitro assays



What you will learn next week



- How molecule **represented** in computer
- How molecule got **visualized**
- The nook and cranny of Docking tools
- The nook and cranny of Interaction Fingerprinting
- **How to do docking** from preparation to analysis
- **Best practices** in silico research

Thank you

My lecture could be wrong,
but I hope that it will be useful for all of us!

Muhammad Radifar, 2023

github.com/radifar/open-lecture

