

# Freeware & Open Access Databases

*Towards a Wiki Drug Discovery?*

# Is it possible to use “free” accessible information to run a drug design project?



# WHAT DO WE NEED?

1. What
2. Where
3. How
4. Example
5. Reference



# WHAT DO WE NEED?

1. Drug Design

2. Information

3. Mining

4. Zika

5. Reference

*FREE VS COMMERCIAL*



# FREE vs COMMERCIAL

## Free / Open Source Systems



## Commercial Systems



# WHAT DO WE NEED?

1. What

2. Where

3. How

4. Example

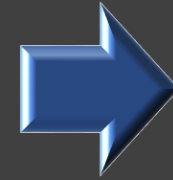
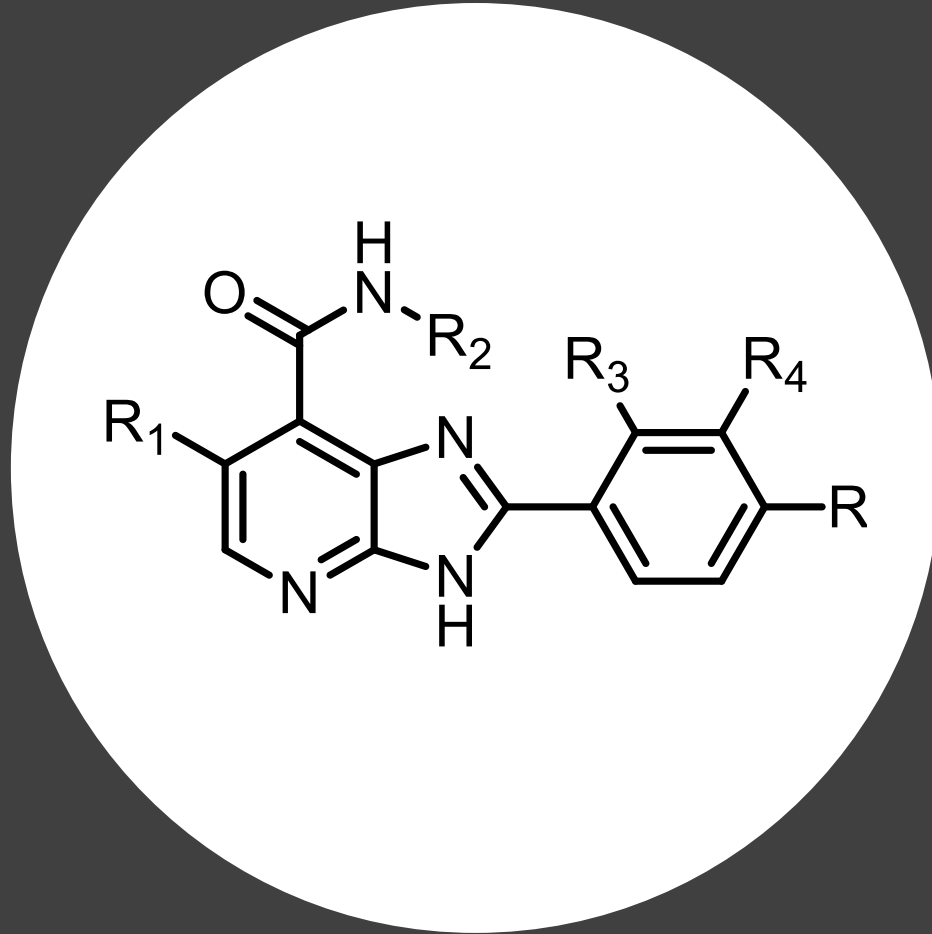
5. Reference



# The Big Picture

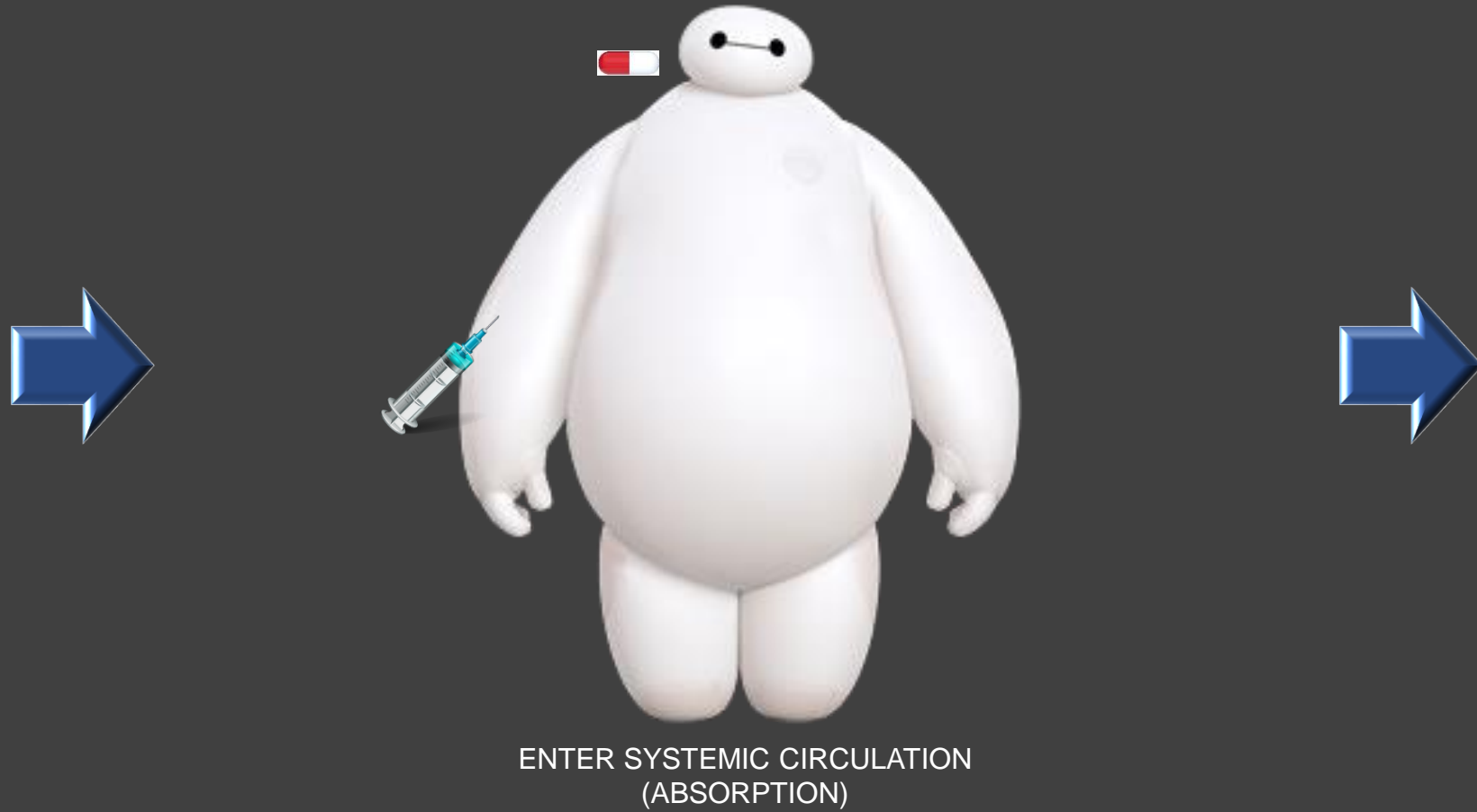


# DRUG DISCOVERY & MOLECULAR PROPERTIES

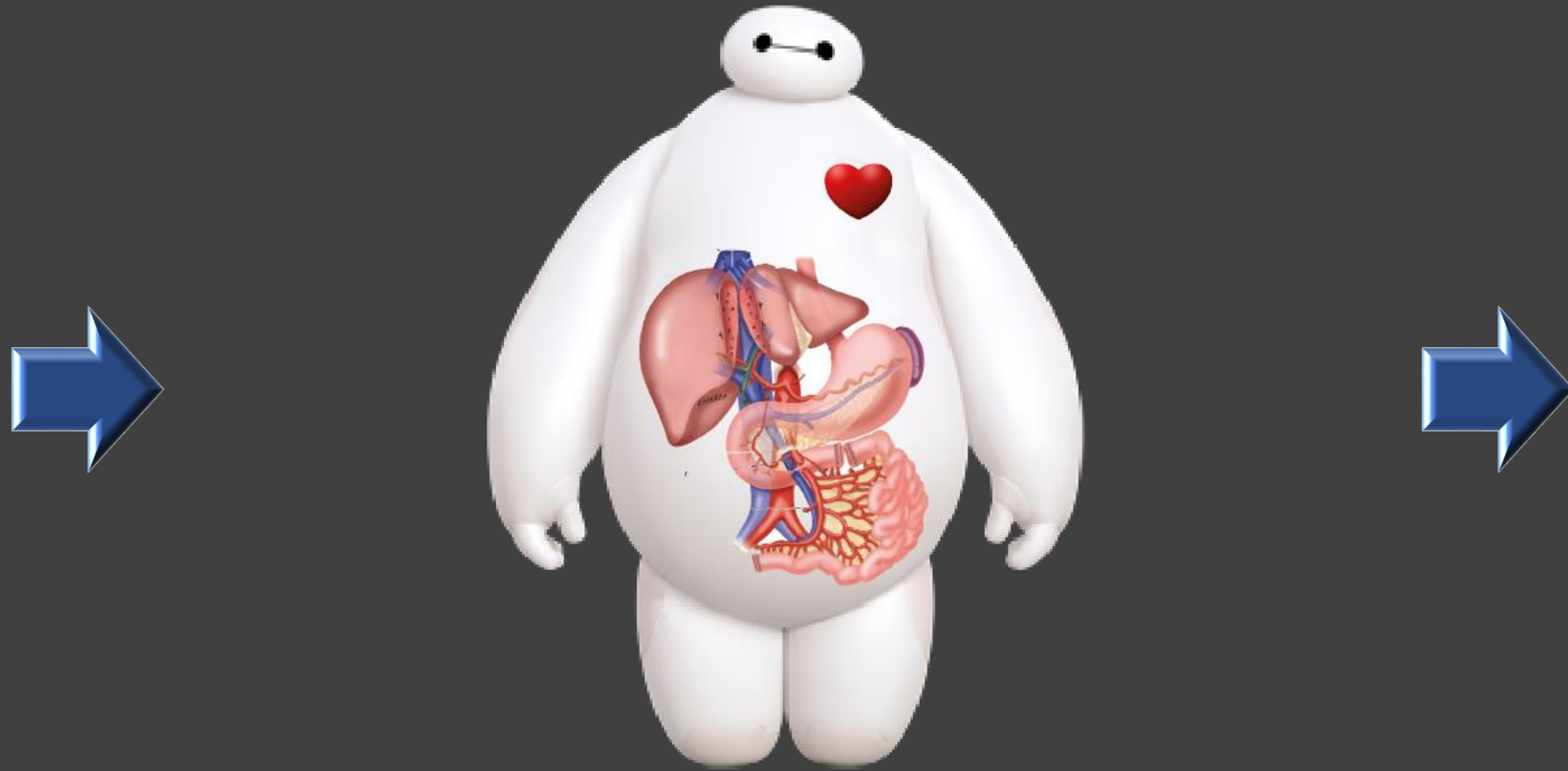




# DRUG DISCOVERY & MOLECULAR PROPERTIES

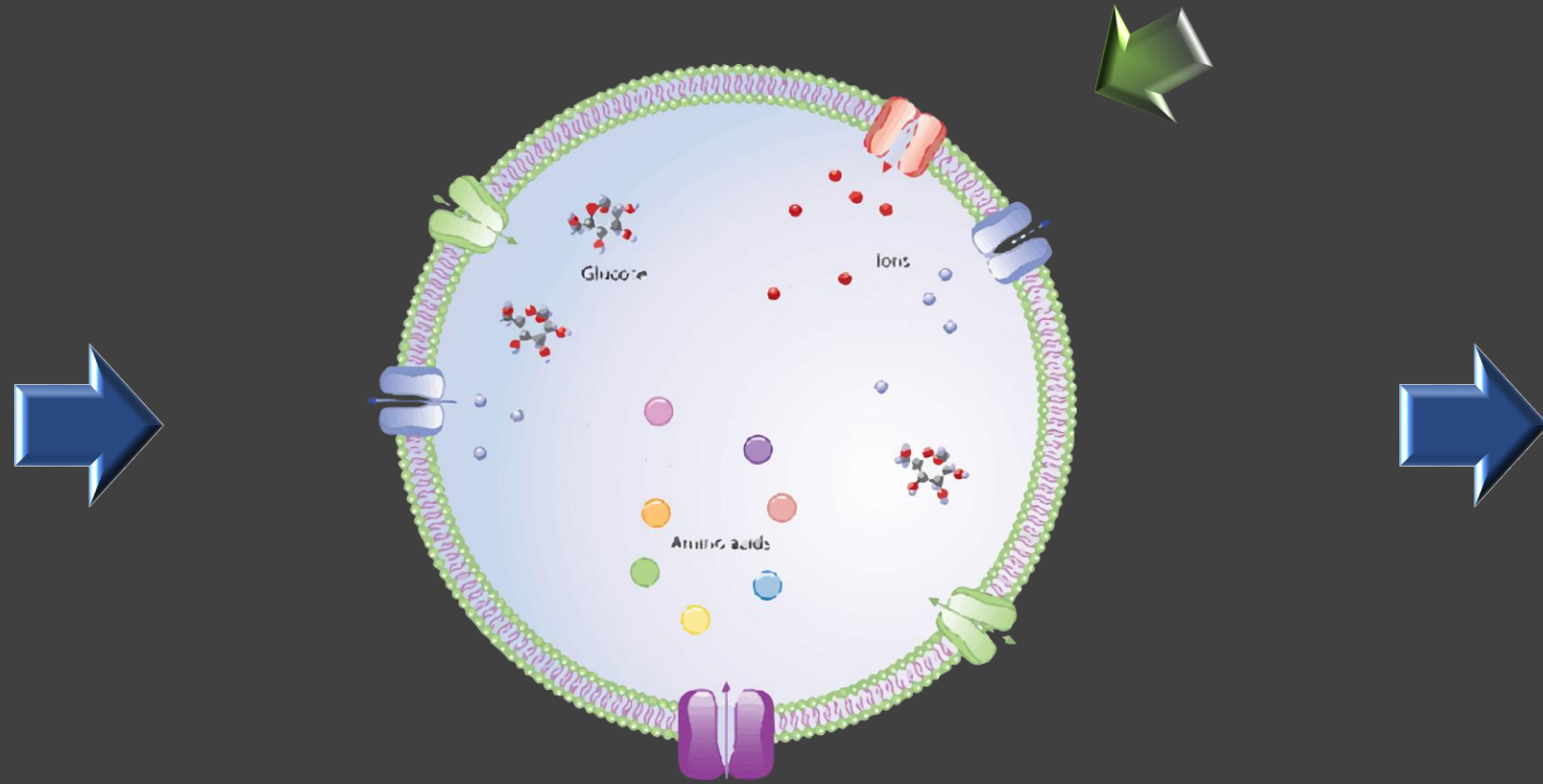


# DRUG DISCOVERY & MOLECULAR PROPERTIES



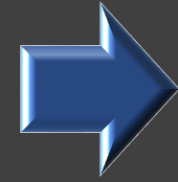
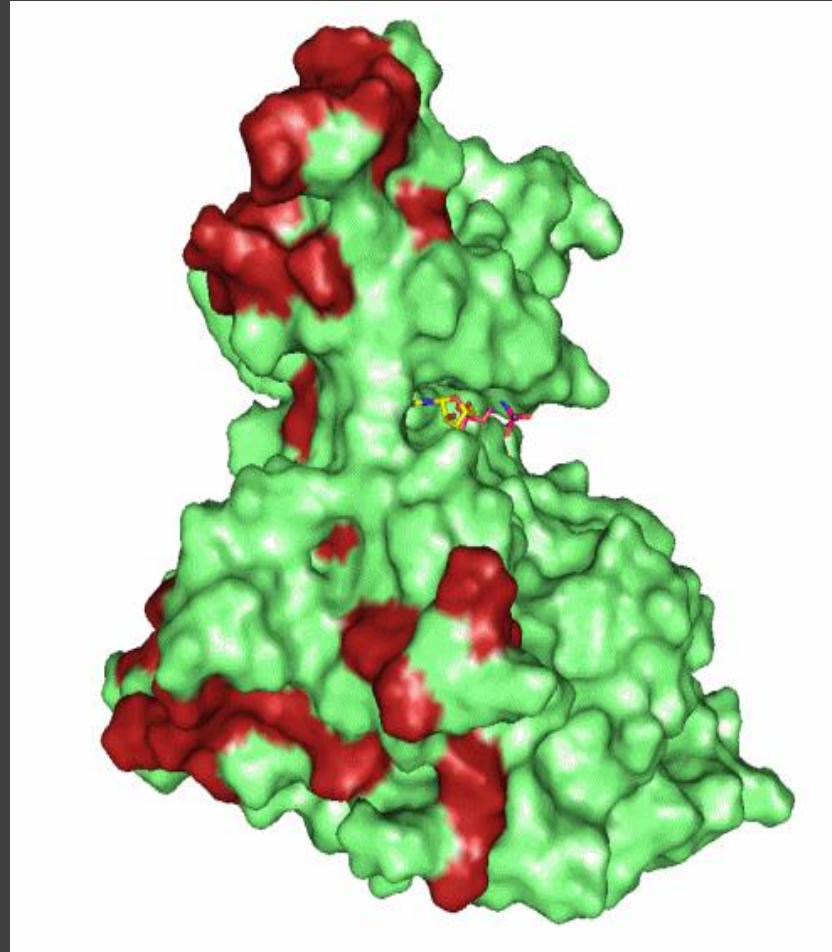
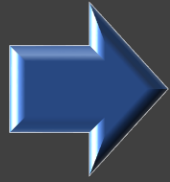
TISSUE/ORGAN DISTRIBUTION

# DRUG DISCOVERY & MOLECULAR PROPERTIES



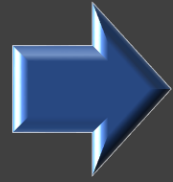
CELLULAR PERMEABILITY

# DRUG DISCOVERY & MOLECULAR PROPERTIES



BIND TO RECEPTOR

# DRUG DISCOVERY & MOLECULAR PROPERTIES



ELIMINATION

# DRUG DISCOVERY & MOLECULAR PROPERTIES

## Pharmaco-kinetics & Pharmaco-dynamics

ENTER SYSTEMIC CIRCULATION  
(ABSORPTION)

TISSUE/ORGAN DISTRIBUTION

CELLULAR PERMEABILITY

BIND TO RECEPTOR

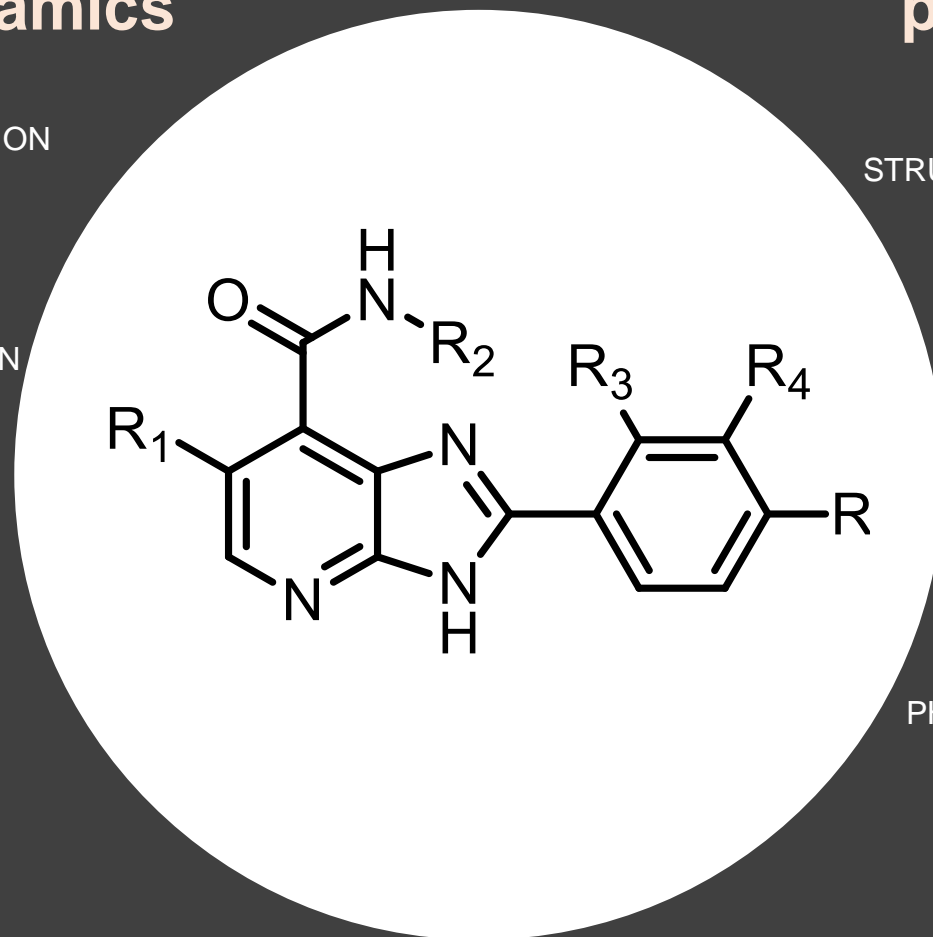
ELIMINATION

## Chemical & Biological properties

STRUCTURAL PROPERTIES

BIOLOGICAL PROPERTIES  
(measured)

PHYSICO CHEMICAL PROPERTIES  
(calculated)



# DRUG DISCOVERY & MOLECULAR PROPERTIES

## Pharmaco-kinetics & Pharmaco-dynamics

ENTER SYSTEMIC CIRCULATION  
(ABSORPTION)

TISSUE/ORGAN DISTRIBUTION

CELLULAR PERMEABILITY

BIND TO RECEPTOR

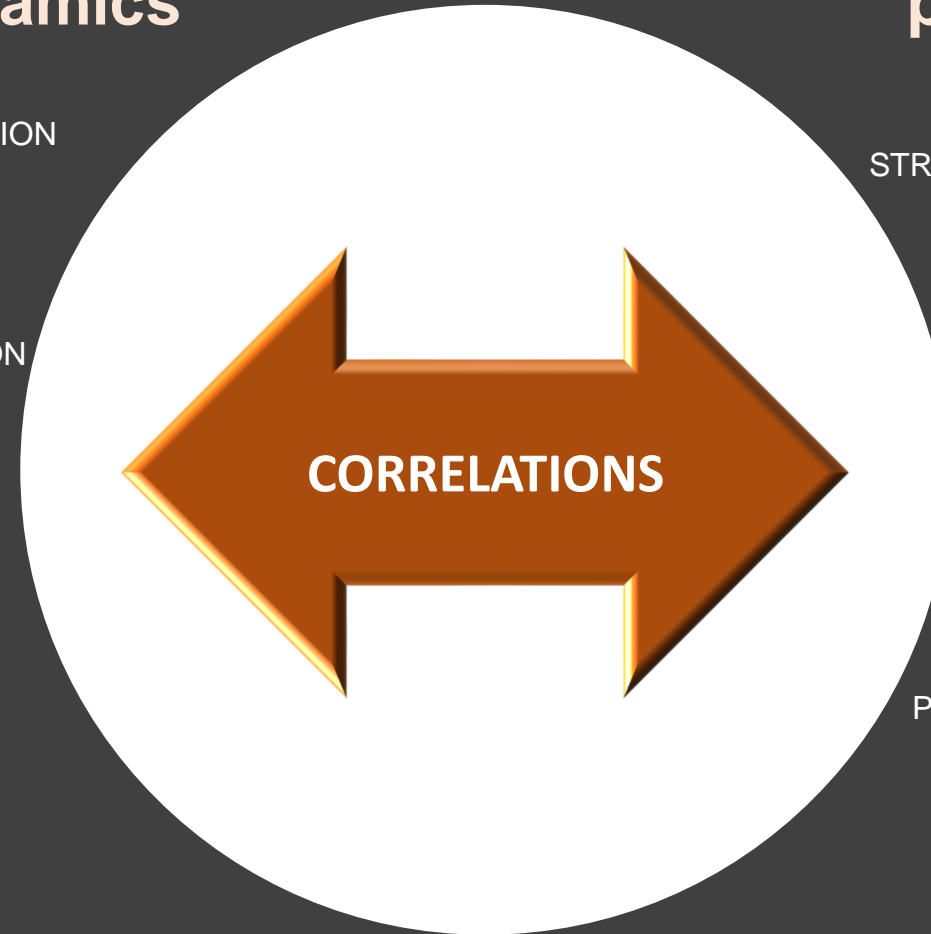
ELIMINATION

## Chemical & Biological properties

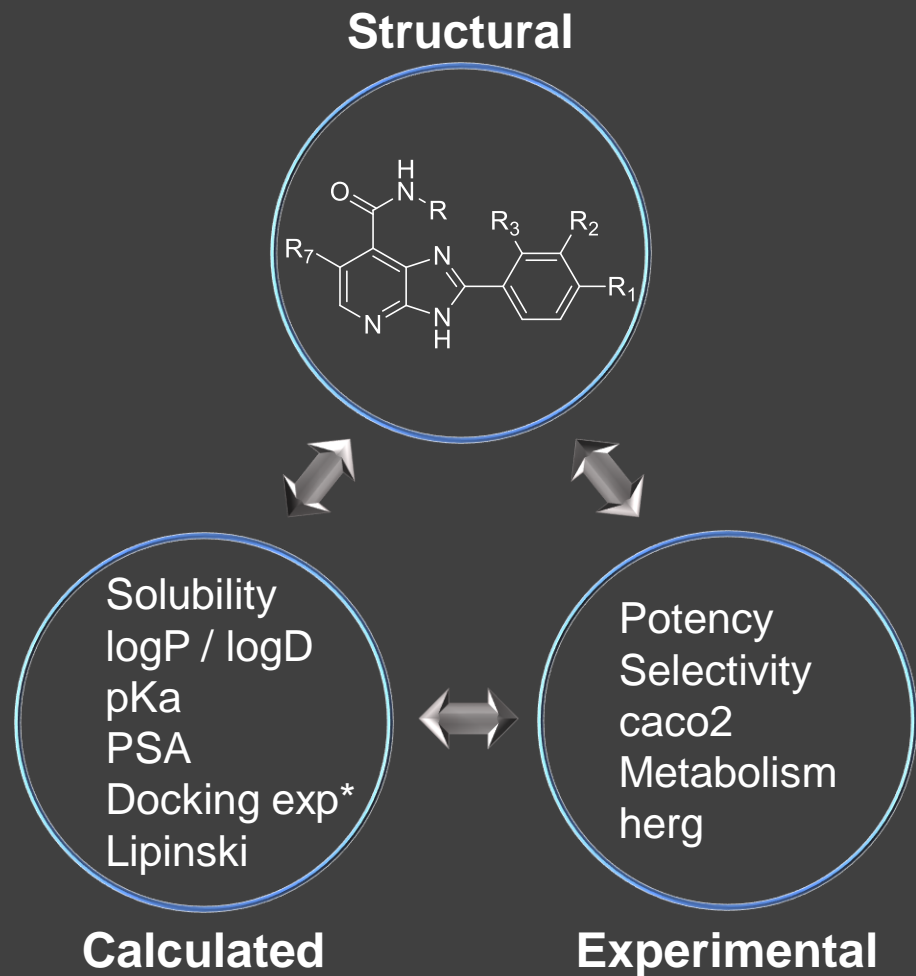
STRUCTURAL PROPERTIES

BIOLOGICAL PROPERTIES  
(measured)

PHYSICO CHEMICAL PROPERTIES  
(calculated)

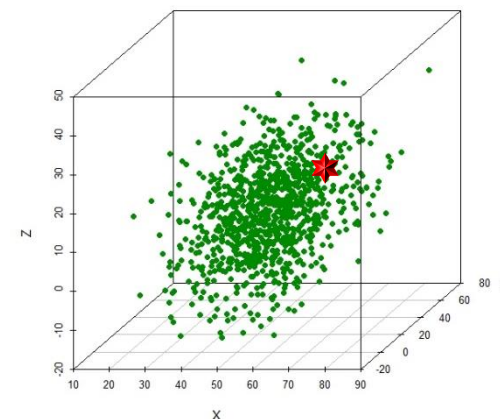
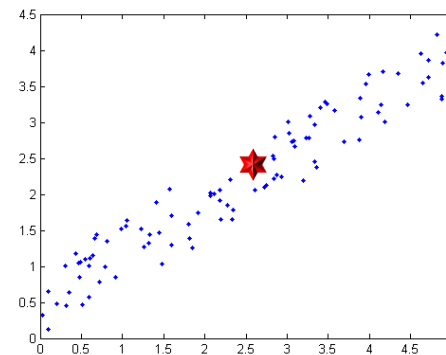


## Med Chem Properties

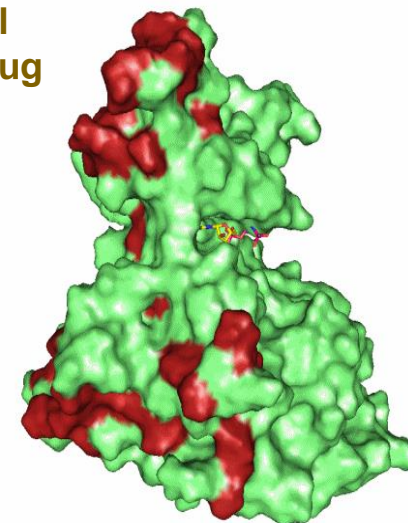


## Data Mining

Available information (building correlations)



★ **New Potential Candidate Drug**





**one more thing...**



one more thing...



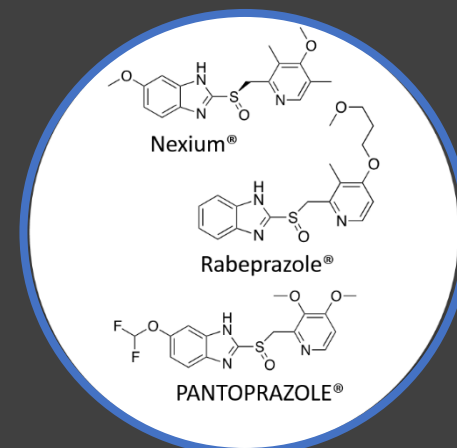
Nature



Analogy



HTS



Chemical Evolution



Structures and associated data

# Information Sources (open access)

- easy integration with analytical platforms (free and commercial)
- most popular (and up to date)

1.What

2.Where

3.How

4.Example

5.Reference

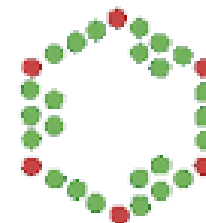
 DRUGBANK

Pubchem

**ChemSpider**  
Search and share chemistry

**CDD**  
COLLABORATIVE DRUG DISCOVERY

ChEMBL



UniProt

**ZINC**

... and many more

# WHERE TO FIND THE INFORMATION

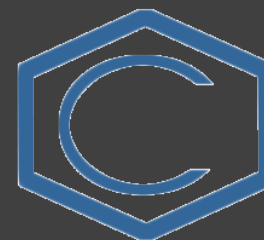
1.What

2.Where

3.How

4.Example

5.Reference



- Free
- NCBI
- FTP / API
- BioAssay / Compounds

# WHAT DO WE NEED?

## Data Mining & Visualization (*open access*)

[www.predictiveanalyticstoday.com](http://www.predictiveanalyticstoday.com)

1.What

2.Where

3.How

4.Example

5.Reference



## WHAT DO WE NEED?

### Data Mining & Visualization (*open access*)



1.What

2.Where

**3.How**

4.Example

5.Reference

- Popular
- User community
- Integration
- Compounds & reactions
- Flexible (applications)

## WHAT DO WE NEED?

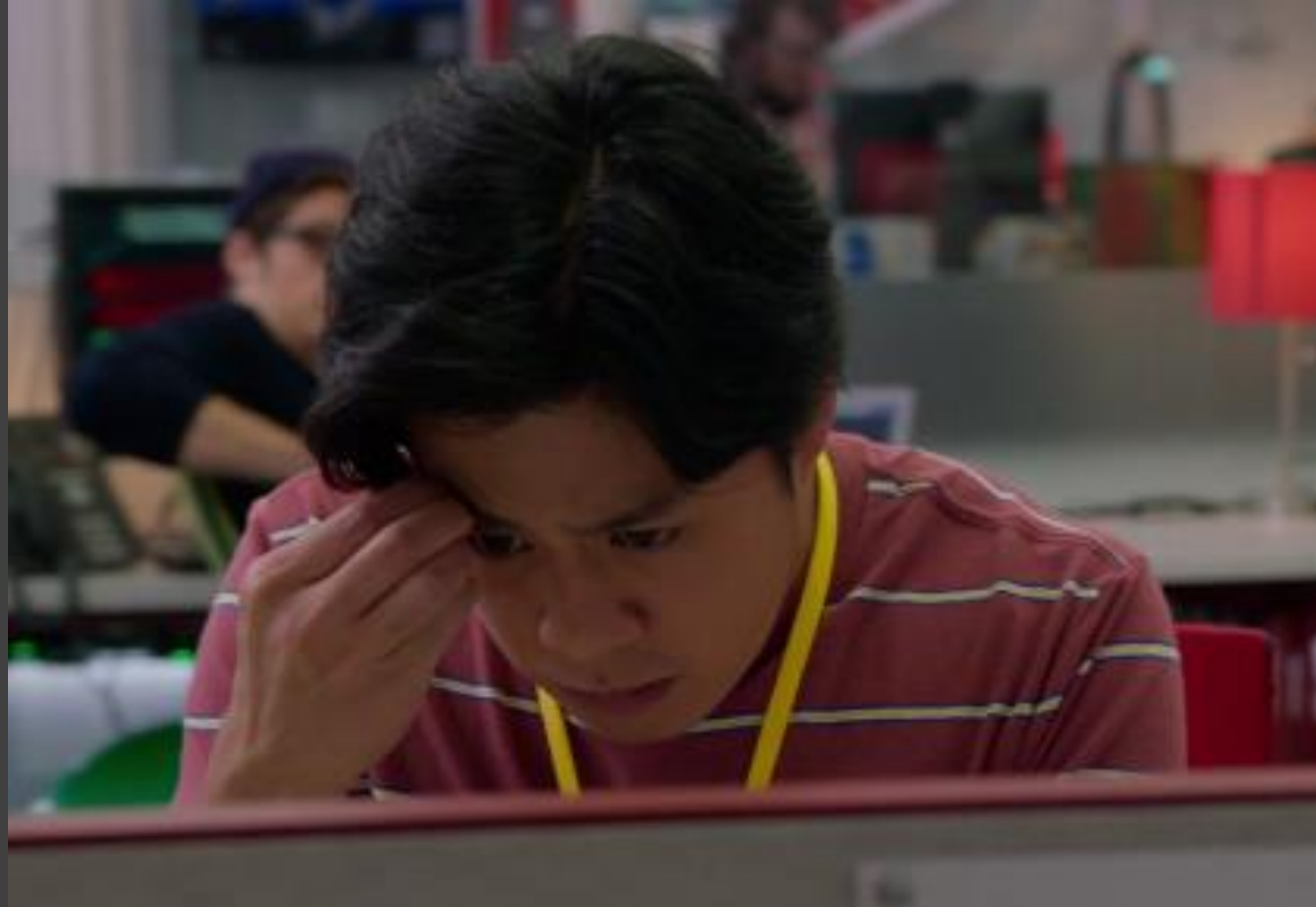
1. What

2. Where

**3. How**

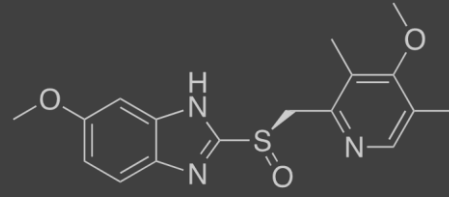
4. Example

5. Reference

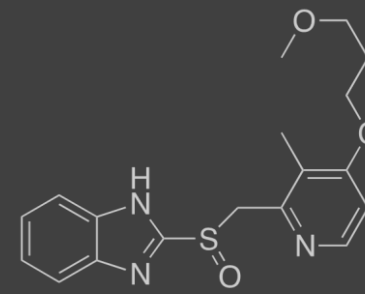


... we do not have own data to compare with...

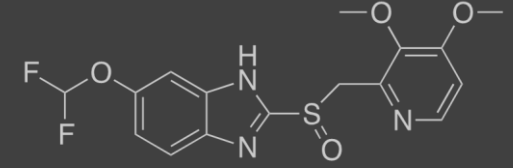
# WHAT DO WE NEED?



Nexium®



Rabeprazole®



Pantoprazole®

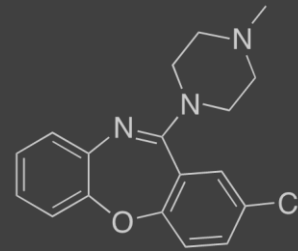
1. What

2. Where

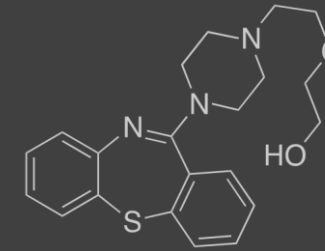
3. How

4. Example

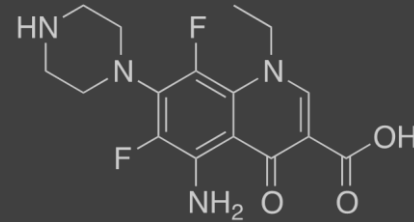
5. Reference



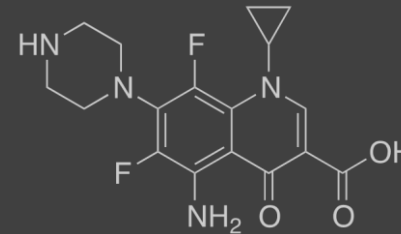
Loxapine®



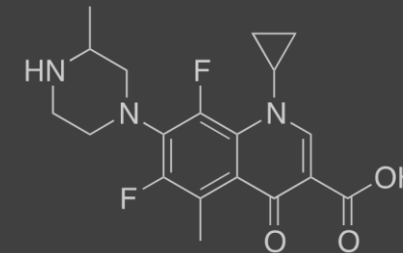
Seroquel®



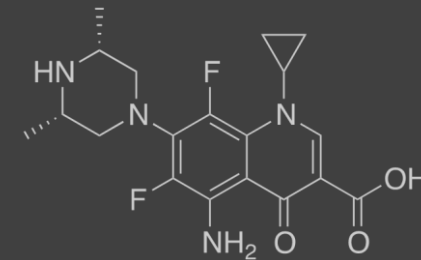
Norfloxacin  
(Noroxin™)



Ciprofloxacin  
(Cipro™)



Grepafloxacin  
(Raxar™)



Sparfloxacin  
(Zagam™)



# WHAT DO WE NEED?

1.What

2.Where

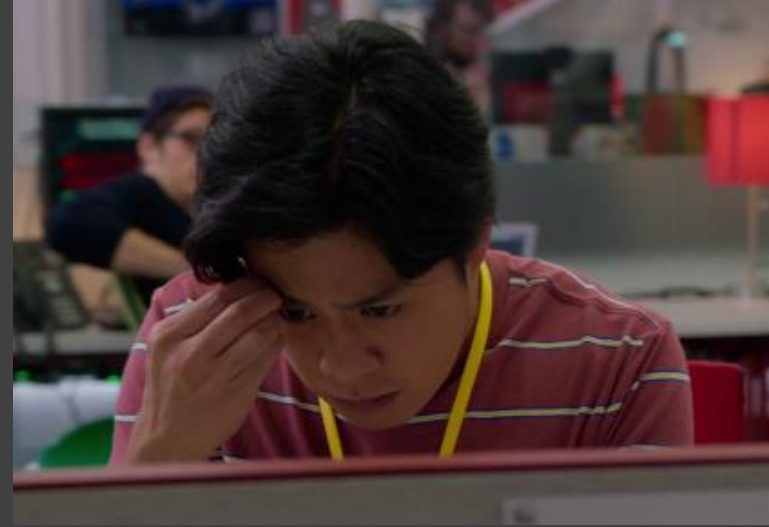
3.How

4.Example

5.Reference

## STRUCTURAL ANALYSIS

*+ ASSOCIATED DATA*

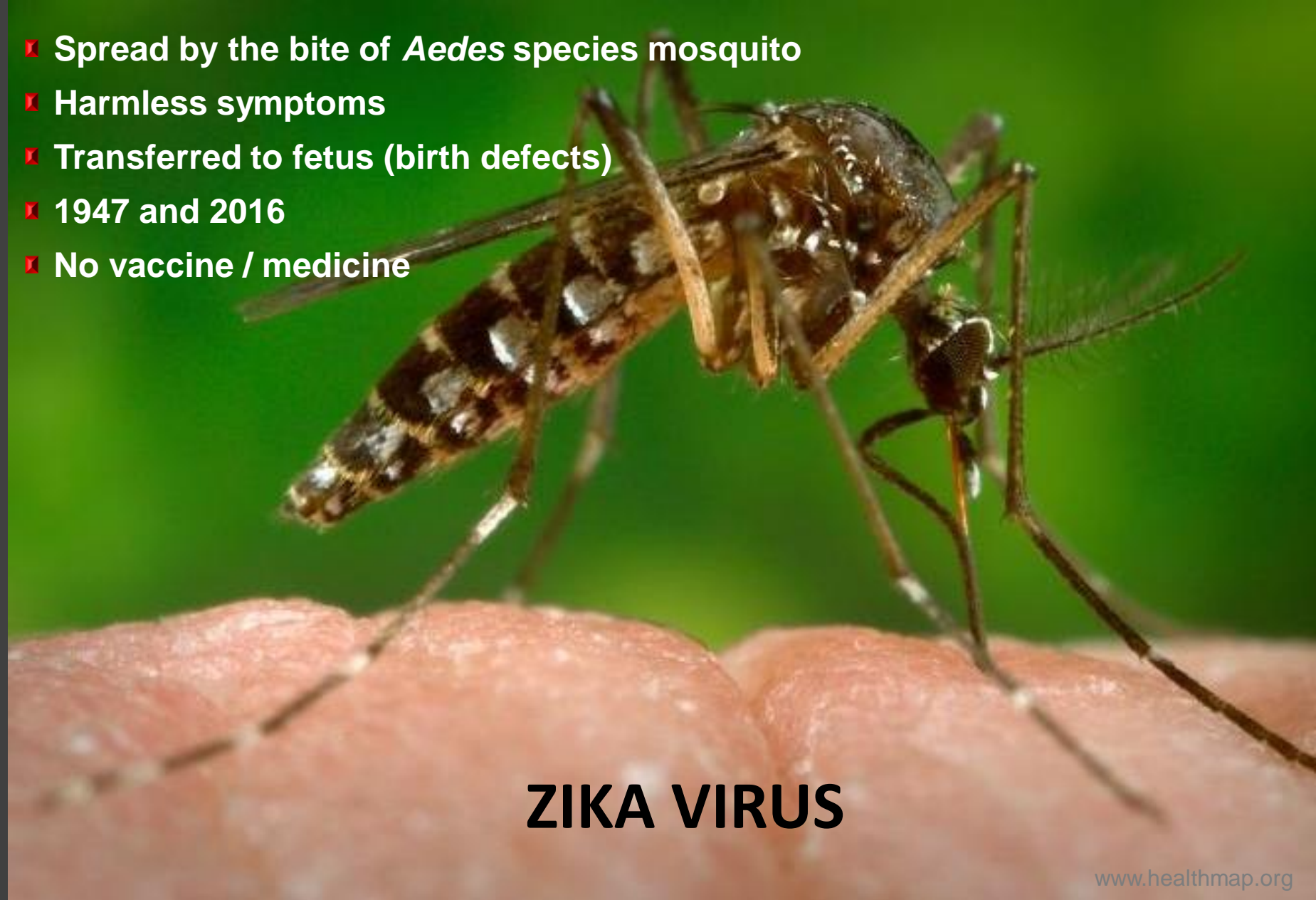


... we do not have own data to compare with...

## WHAT DO WE NEED?

1. What
2. Where
3. How
4. Example
5. Reference

- ❑ Spread by the bite of *Aedes species* mosquito
- ❑ Harmless symptoms
- ❑ Transferred to fetus (birth defects)
- ❑ 1947 and 2016
- ❑ No vaccine / medicine

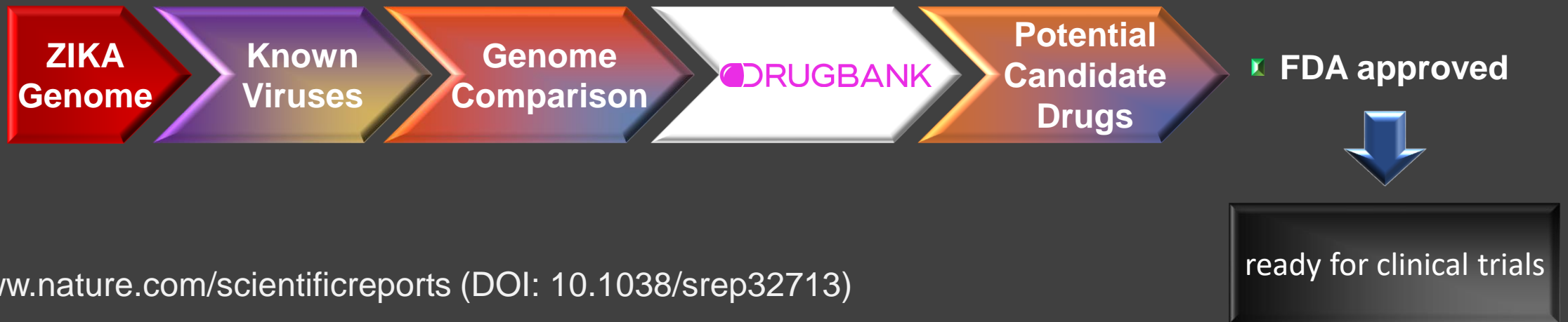


**ZIKA VIRUS**

[www.healthmap.org](http://www.healthmap.org)

# ZIKA: OPEN ACCESS DRUG DISCOVERY PROJECT

- Collection of biological & chemical data
- 464 potential drug targets
- Direct mining possible (not straightforward)
- Download data for external mining



[www.nature.com/scientificreports](http://www.nature.com/scientificreports) (DOI: 10.1038/srep32713)

[www.openzika.ufg.br](http://www.openzika.ufg.br)

<http://bioinfo.imtech.res.in/manojk/zikavr/index.php>

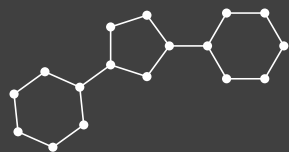
# ZIKA: OPEN ACCESS DRUG DISCOVERY PROJECT

## 3 main biological targets:

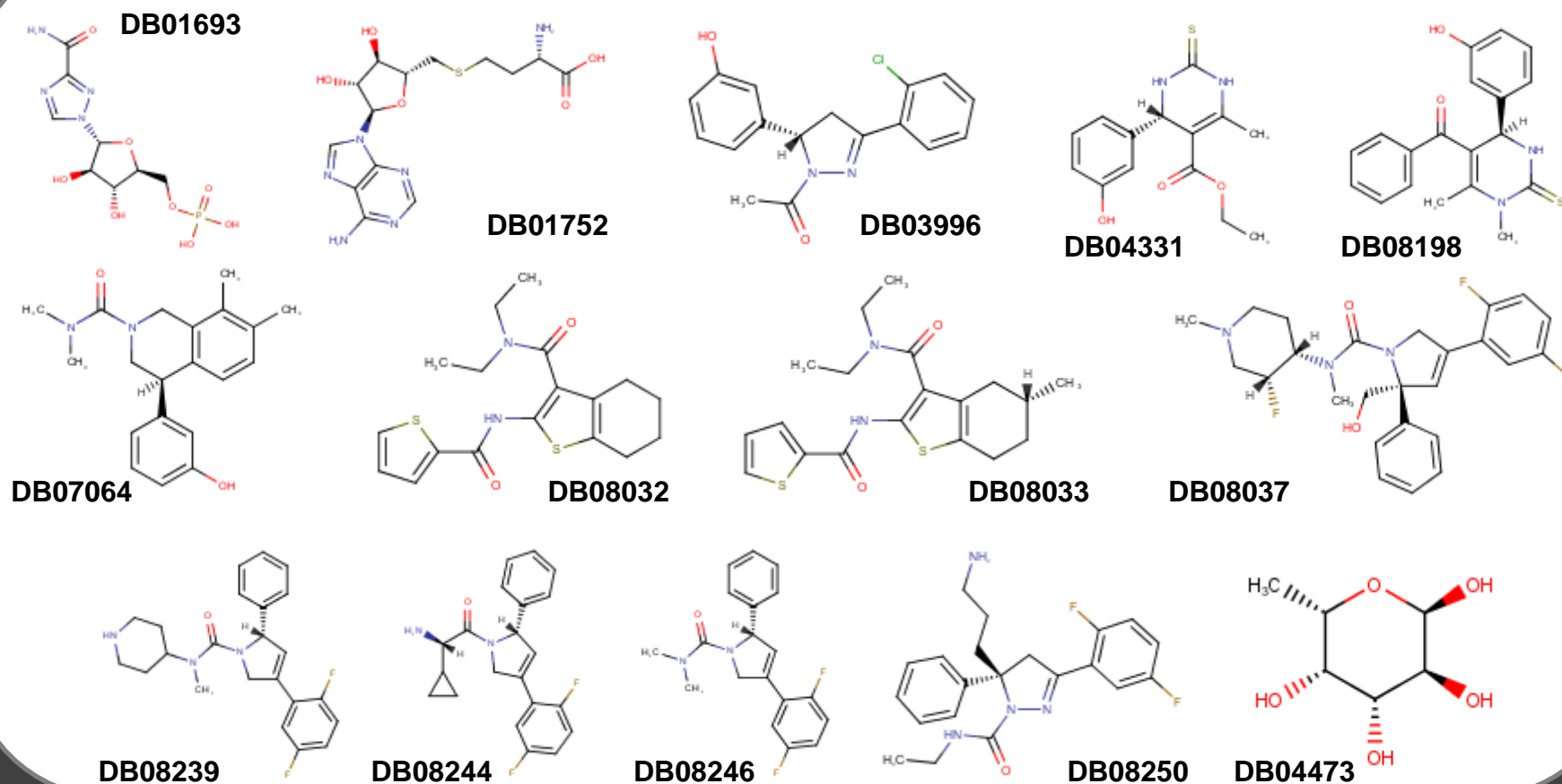
- Interleukin-4-receptor subunit alpha
- Genome polyprotein
- Kinesin-like protein KIF11

## 1 structural "coincidence"

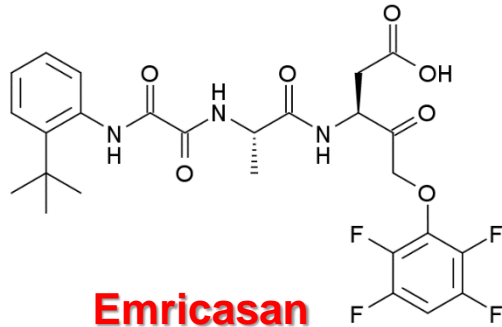
- 6 Molecules share the same subgraph



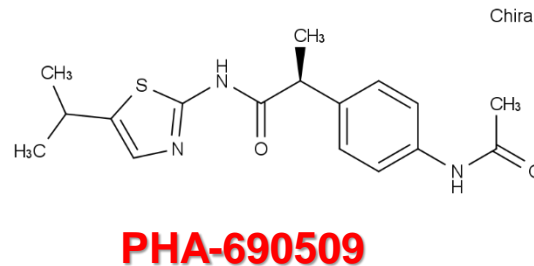
DRUGBANK



# ZIKA: OPEN ACCESS DRUG DISCOVERY PROJECT

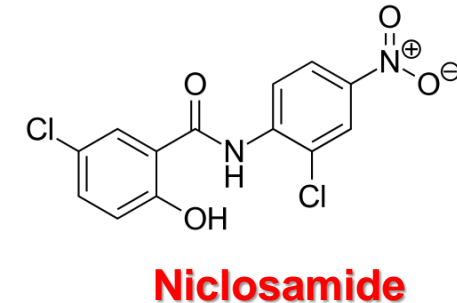


- pan-caspase inhibitor
- neuroprotective activity
- not suppress ZIKV replication



- cyclic-dependent kinase inhibitor (CDKi)\*
- antiviral activity

\**Ten structurally unrelated CDKis inhibit ZIKA replication*



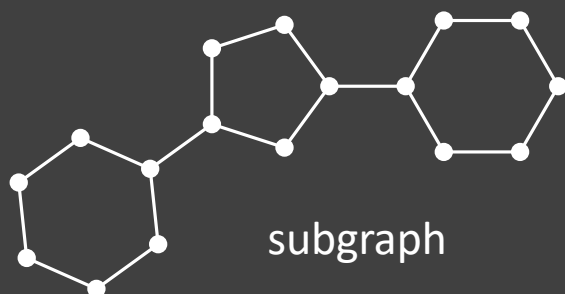
- FDA-approved drug
- treatment of worm infections

## Combination of Drugs

## Taken together ...

<http://bioinfo.imtech.res.in/manojk/zikavr/index.php>

### 1 structural “coincidence”



- Interleukin-4-receptor subunit alpha
- Genome polyprotein
- Kinesin-like protein KIF11

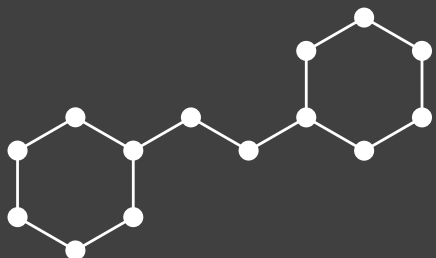
*Nature Medicine* 22, 10, 2016, 1101

**cyclic-dependent kinase inhibitor (CDKi)**

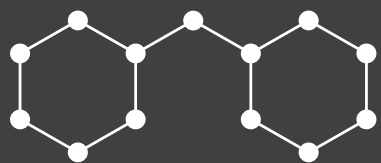
*Ten structurally unrelated CDKis inhibit ZIKA replication*

1. Search bioactives related to the subgraph
2. Search for KIF11 inhibitors (Structural comparison)
3. Search for CDK inhibitors (Structural comparison)

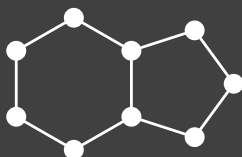
## CDKi's (27 molecules)



4/27



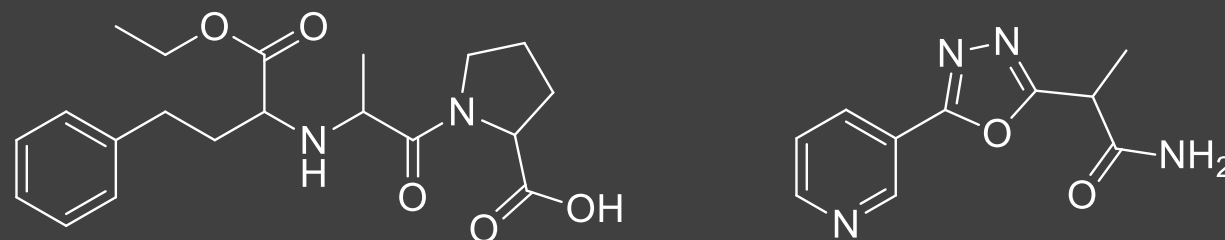
11/27



13/27

*Ten structurally unrelated CDKis inhibit ZIKA replication*

**257 structurally unrelated CDKi's  
(based on subgraph)**



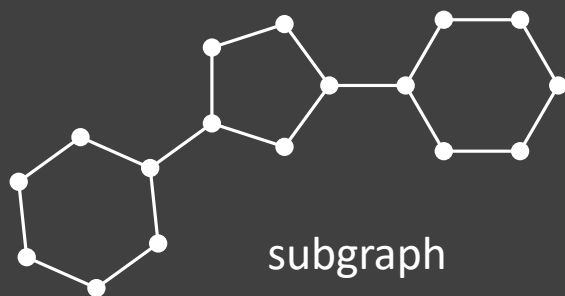
**...structures and bio-assay data available**

**1452 CDKi's found (CDK1-CDK20)**

# ZIKA: OPEN ACCESS DRUG DISCOVERY PROJECT

14 Molecules in [www.drugbank.ca](http://www.drugbank.ca)

1 structural “coincidence”

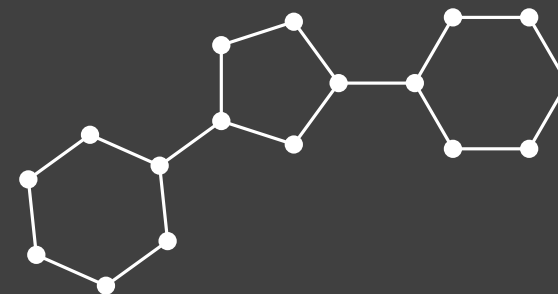


6/14

- Interleukin-4-receptor subunit alpha
- Genome polyprotein
- Kinesin-like protein KIF11

7/14

KIF11 inhibitors found:



0/5776

Graph analogues found: 1100

- 18 reported as Kinesin like protein 1
- Kinesin superfamily (non-standard names)
- >1000 compounds (Tanimoto >0.7)



## Summary

- **Analogues with same subgraph (Tanimoto >0.7)**
- **Non related structures active on same targets**
- **Bioassay data**
- **Other analysis possible**
  - ❑ Lipinski
  - ❑ Pharmacophore
  - ❑ Match Pairs
  - ❑ SAR
  - ❑ R-Group Decomposition

# WHAT DO WE NEED?

1. What

2. Where

3. How

4. Example

5. Reference

REAXYS API

KNIME Analytics Platform

File Edit View Node Help

KNIME Explorer

- EXAMPLES (knime-guest@http://public)
- LOCAL (Local Workspace)
  - \_20\_BindingDB\_large\_download
  - Alex
  - Chemnotia
  - ICIC\_Heidelberg
  - IOS
  - Mining (reaction databases)
  - Modules FF
  - RISE\_PROJECTS
  - Starting Materials Mining
  - Tools Downloaded
  - Tools FF
  - Work\_in\_progress

Node Repository

- IO
- Manipulation
- Views
- Analytics
- Database
- Other Data Types
- Structured Data
- Scripting
- Tool Integration
- Community Nodes
- KNIME Labs
- Workflow Control
- Social Media
- Reporting
- Chemistry
- Chemistry/Infocom
- Elsevier
  - Reaxys Bioactivities
  - Reaxys Citations
  - Reaxys Reactions
  - Reaxys Substances
- InfoChem
- ICEditKnime

2: ios17\_s2\_Zika\_June17

\*0: PK17002\_May20

Welcome to KNIME Analytics Platform

0:1404:0 - Cluster Preparation

2:1933 - kinesin-like (Node 1933)

Node 1909: List of CDK inhibitors used in Nature Medicine paper

Node 1911

Node 1913

Node 1914

SEARCH ON PUBCHEM: ANALOGUES OF THE 28 CDKI's REPORTED ON NATURE PAPER!!

Table Creator

String Manipulation

GET Request

Row Splitter

Cell Splitter

Ungroup

Target CDKL1-5 manual listing the only way? can't find a way to do CDKx Possible at all or just me?

make url for CDKLi targets

get data (requires 30ms delay)

Keep status 200 (else error)

Node 1943

19 assays

String Manipulation

GET Request

Row Splitter

Cell Splitter

Ungroup

make url from AID

get data (requires 30ms delay)

Keep status 200 (else error)

Node 1940

>95 cmpds

Column Filter

Column Rename

GroupBy

Get Structures

Cluster Preparation

Row Filter

Molecule Type Cast

DB vs Rxys Kinesin cores

common skeletons

Node 1931

Node 1933

filtered duplicates: 17 cmps

CDKLi structures

Node 1953

Node 1954

Node 1934

PUBCHEM CPDS clusters + graph

CDKLi structures

19 assays

>95 cmpds

filtered duplicates: 17 cmps

CDKLi structures

PUBCHEM CPDS clusters + graph

Node 1953

Node 1954

Node 1934

common skeletons

Node 1316

Simple analysis of 27CDKI structures "common skeletons"

Target CDKL1-2 (not CDKL) manual listing the only way? can't find a way to do CDKx Possible at all or just me?

Out port1: common cor

Out port2: core not in pub

## WHAT DO WE NEED?

### REAXYS API

1.What

2.Where

3.How

4.Example

5.Reference

- CDKi's ca 40000 structures (25x more)
- KIF11 inhibitors ca 4000 structures (1000 less)
- Equal amount of data
- Same conclusions

# Is it possible to use “free” accessible information to run a drug design project?

YES

but ...

- ▣ Quality
- ▣ Accessibility
- ▣ Up to Date



## OUR ANSWERS

# Towards a Wiki Drug Discovery?

## YES

- ✦ Neglected diseases (Malaria)
- ✦ Orphan diseases
- ✦ Unmet medical needs

## NO

- ✦ All others



FREE vs COMMERCIAL

## Free / Open Source Systems



## Commercial Systems



YOUR TASTE!  
YOUR NEEDS!