monthly webinar series

Fragment Based Ligand Discovery: X-ray crystallography as a screening tool

Smaller is Better

Laurent Maveyraud, 03/03/21

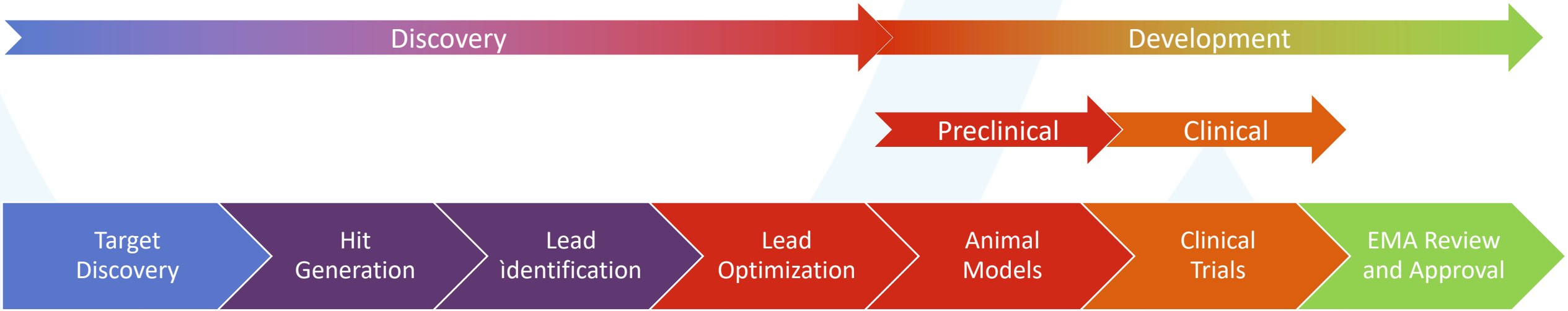
Overview

- The drug discovery process
- How can we find our first hit?
- FBLD: smaller is better
- FBLD: screening methods
- X-ray crystallography as a screening tool
- Practical considerations
  - Optimizing crystallization
  - Efficient soaking/co-crystallization
  - High throughput data collection and processing
  - Looking for binders
  - Refinement
  - PDB deposition
- Does it really work?

Overview

- The drug discovery process
- How can we find our first hit?
- FBLD: smaller is better
- FBLD: screening methods
- X-ray crystallography as a screening tool
- Practical considerations
  - Optimizing crystallization
  - Efficient soaking/co-crystallization
  - High throughput data collection and processing
  - Looking for binders
  - Refinement
  - PDB deposition
- Does it really work?

The drug development process



- Program Selection
- Target identification
- Target validation

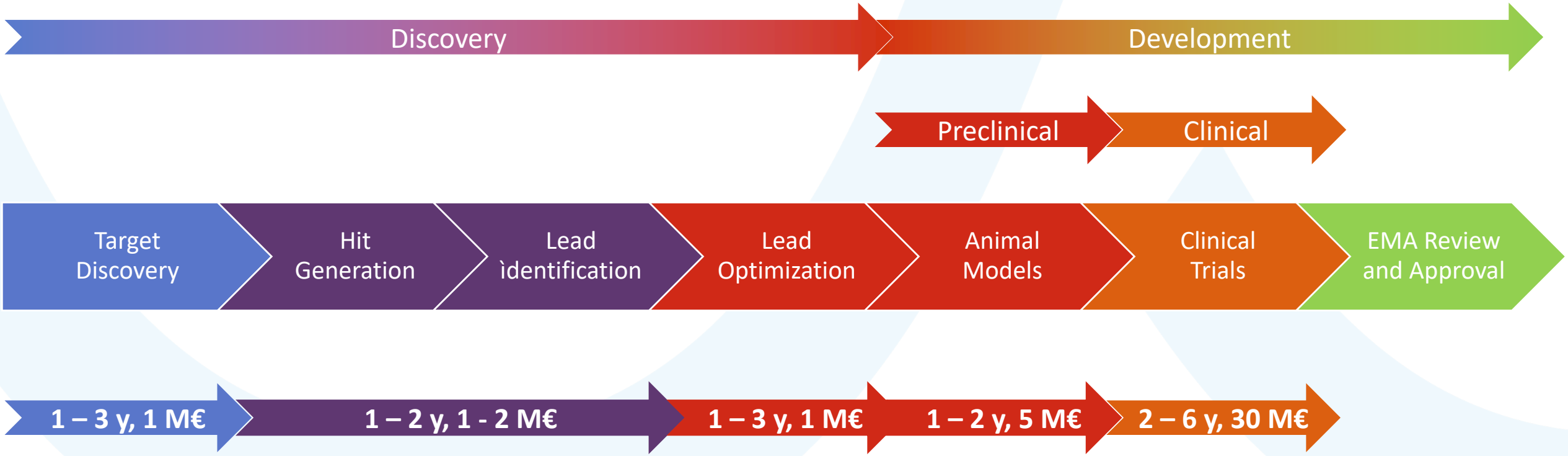
- Assay Development
- High Throughput Screening
- Virtual screening
- Structure Based Drug Design
- Hit to Lead

- Lead to Drug
- ADMETox

- ADMETox
- Efficacy
- Safety

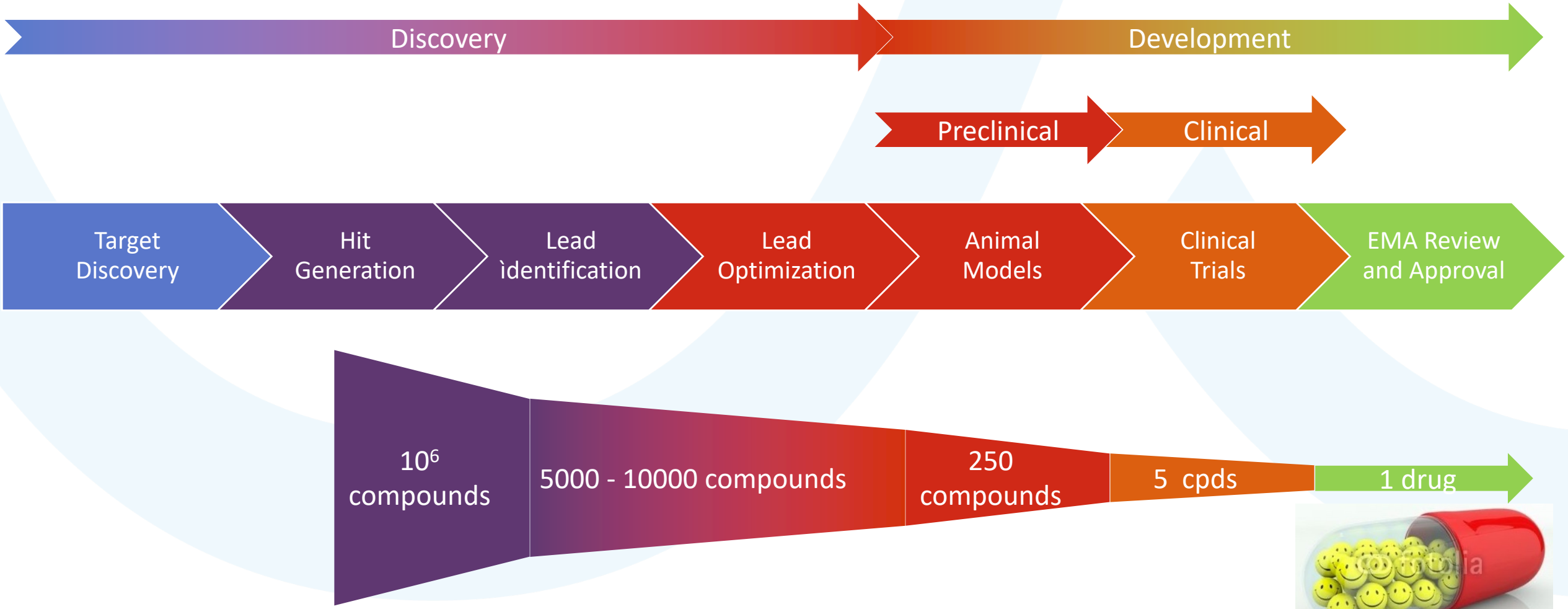
- Phase I
- Phase II
- Phase III

A long and expensive process

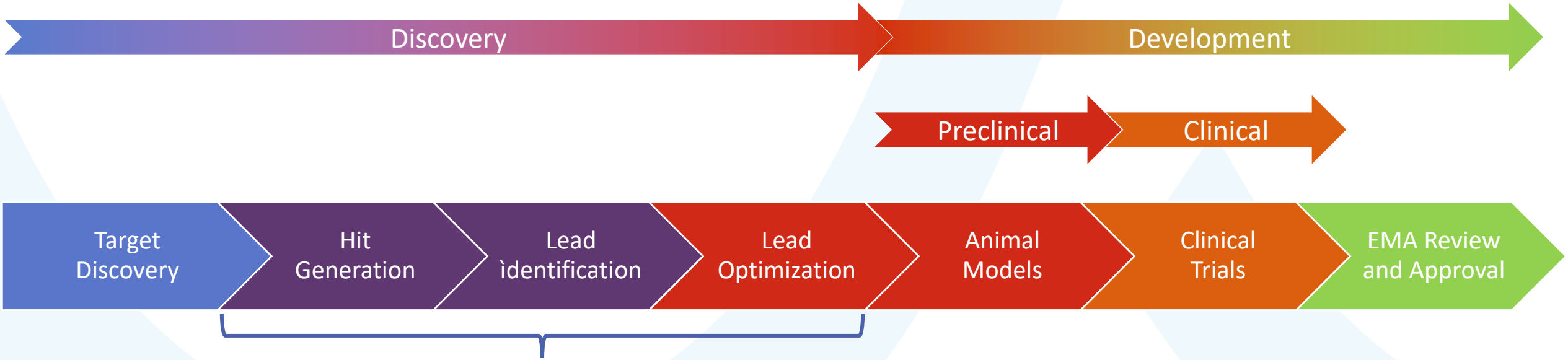


6 – 15 years
25 – 40 M€

A high attrition rate



Where can we help?



Biophysics (SPR, BLI, DSF, NMR, ITC…)

Structural Biology (NMR, RX, CryoEM…)

Overview

- The drug discovery process
- How can we find our first hit?
- FBLD: smaller is better
- FBLD: screening methods
- X-ray crystallography as a screening tool
- Practical considerations
  - Optimizing crystallization
  - Efficient soaking/co-crystallization
  - High throughput data collection and processing
  - Looking for binders
  - Refinement
  - PDB deposition
- Does it really work?

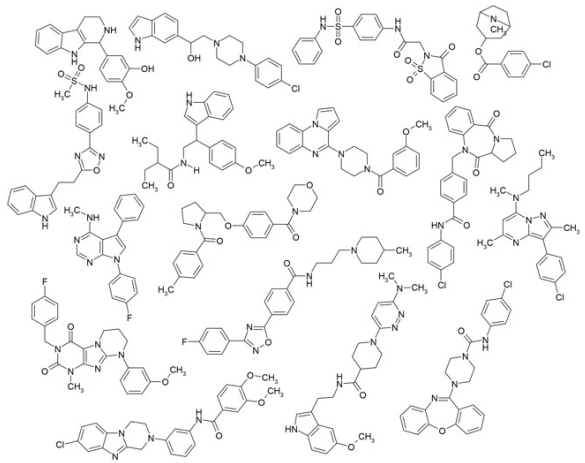
At the beginning, is the hit

A hit is any chemical compound that binds at the target site.

It provides a scaffold upon which medecinal chemists can elaborate more potent compouds.

- Prior knowledge
- HTS
- Virtual screening
- …

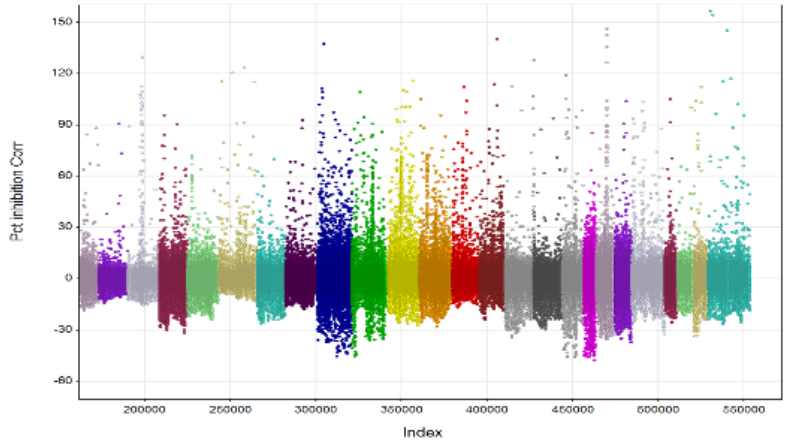
HTS: huge libraries, large compounds



Chemical library
(10^5 – 10^6 compounds)

Screening
(enzymatic, cellular, phenotypic,…)

Single point



Primary hits (1 %)

Confirmation assay (tripicate)

Counter assay

Dose response

A few compounds left

HTS: what do we have in the libraries?

- A **detectable** biological activity, ie an affinity for the target in the (high-) micromolar range
- No highly reactive species (known as PAINS)
- « easily » transformed in drugs (biodisponibility, solubility…)

The rule of five (RO5, aka Lipinsksi's rules) :

- < 500 Da
- ClogP < 5
- < 5 H-bond donors
- < 10 H-bond acceptors

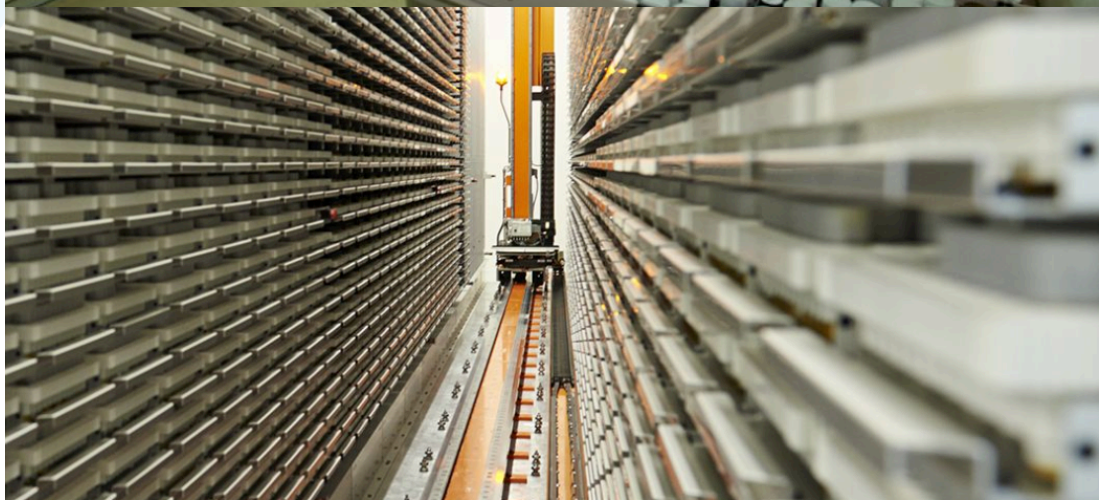
logP = log(partition coefficient octanol/water)

Lipinski, 1997

HTS: what do we have in the libraries?

Sanofi's chemical library (2015) :

- >1,5 M compounds
- 22 FTE for compound management
- 4500 m^2
- 30 M€ running costs

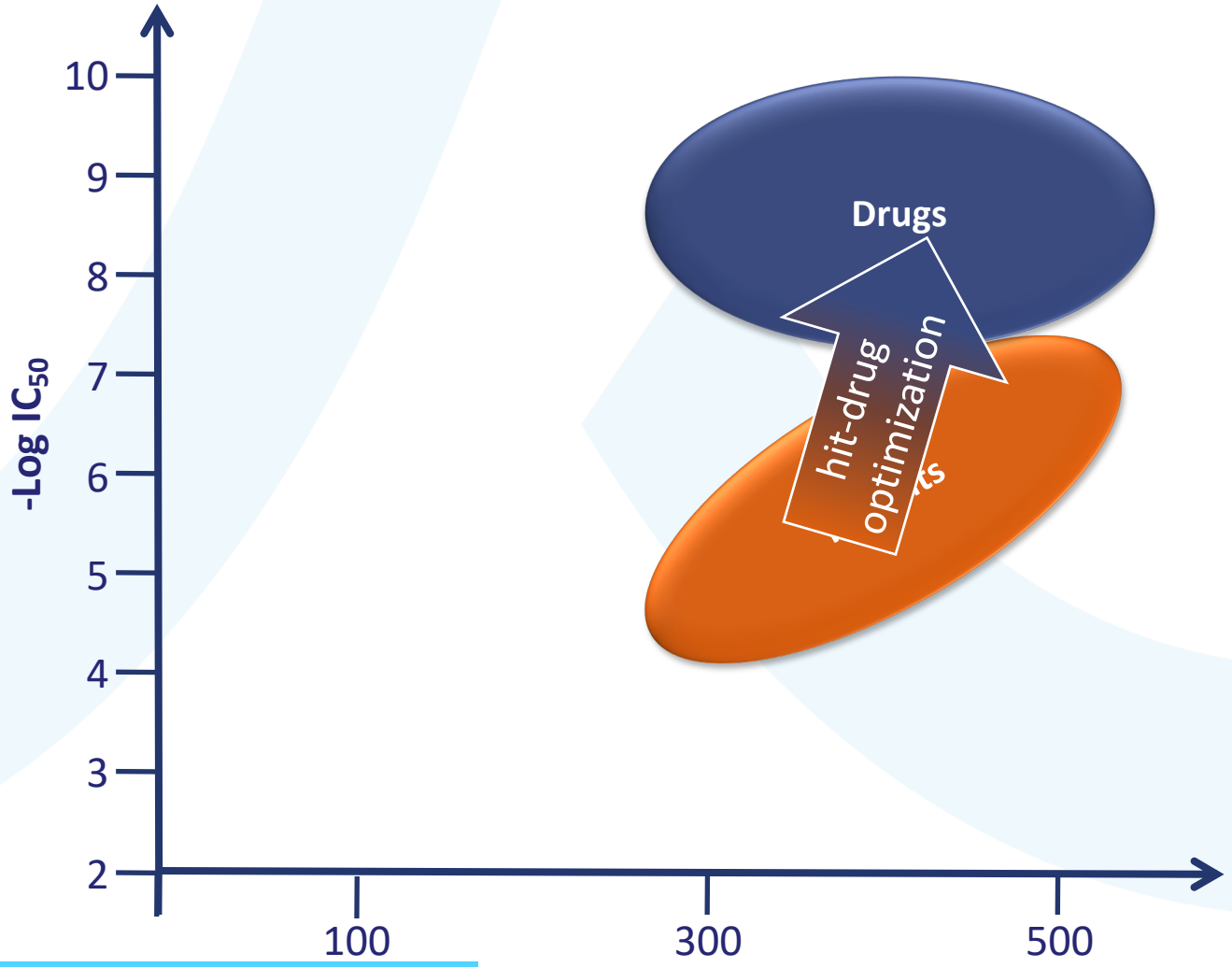


Only possible for big pharma

HTS: what about chemical diversity?

- 500 Da : about 30 non-H atoms
- 10^{60} possible compounds
- 10^{8} know compounds
- 10^{6} in the library

Accessible chemical diversity is a tiny portion of the real chemical diversity



Hits have to be optimized at a constant molecular weight

Bohacek, 1996
Hoffmann, 2019
Scott, 2012

If you are not a big pharma, keep it small

Rules of 5

- MW < 500 Da
- < 5 H-bond donnors
- < 10 H-bonds acceptors
- ClogP < 5

DRUG LIKE

- 10^{60} possible compounds
- 10^{8} known compounds
- 10^{6} in library

Rules of 3

- MW < 300 Da
- < 3 H-bond donnors
- < 3 H-bonds acceptors
- ClogP < 3

FRAGMENT LIKE

- 10^{9} possible compounds (180 Da)
- 10^{5} known compounds
- 10^{3} in library

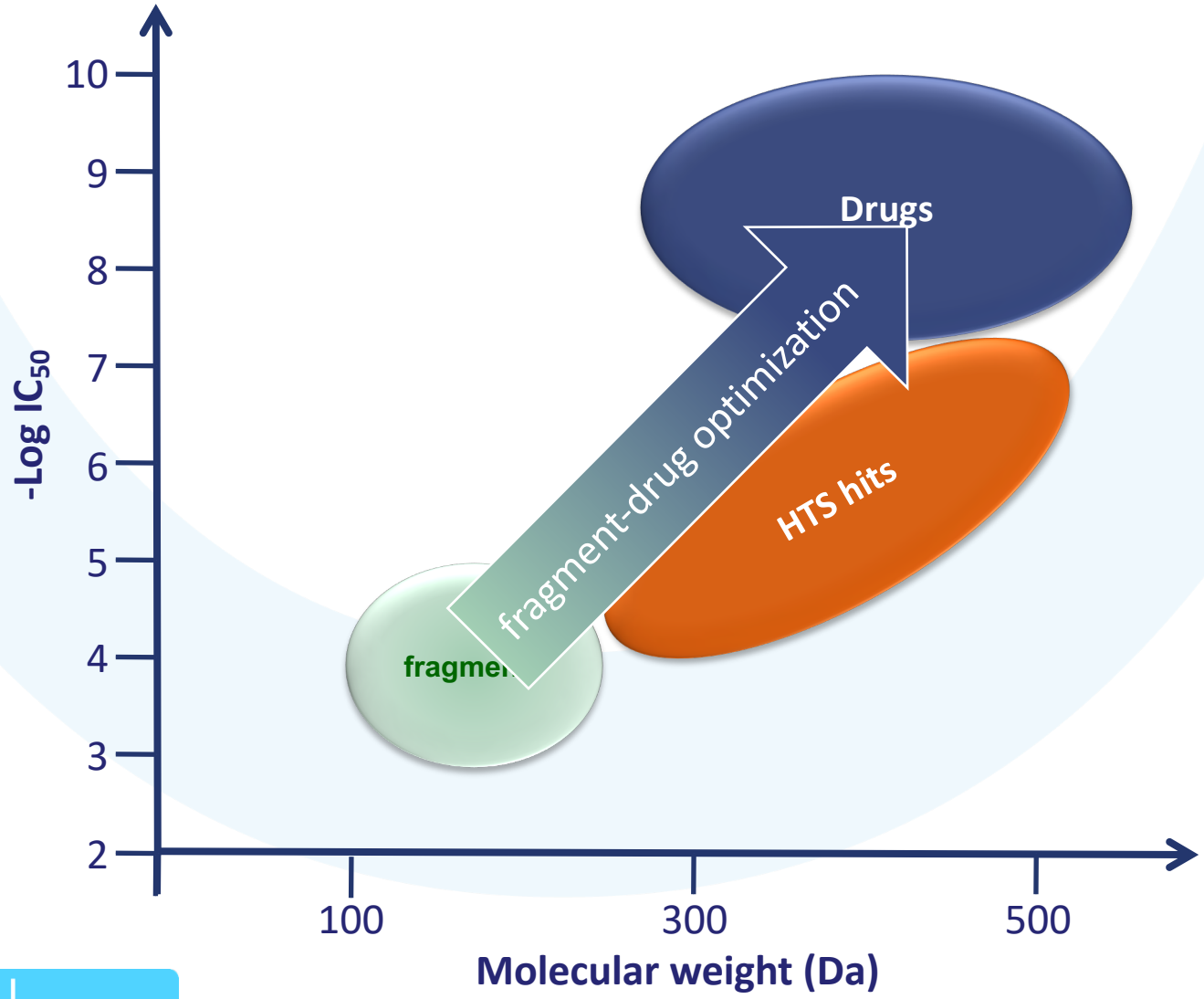
Fragments:
better at exploring the available chemical diversity

Fink, 2007
Reymond, 2012

Overview

- The drug discovery process
- How can we find our first hit?
- FBLD: smaller is better
- FBLD: screening methods
- X-ray crystallography as a screening tool
- Practical considerations
  - Optimizing crystallization
  - Efficient soaking/co-crystallization
  - High throughput data collection and processing
  - Looking for binders
  - Refinement
  - PDB deposition
- Does it really work?

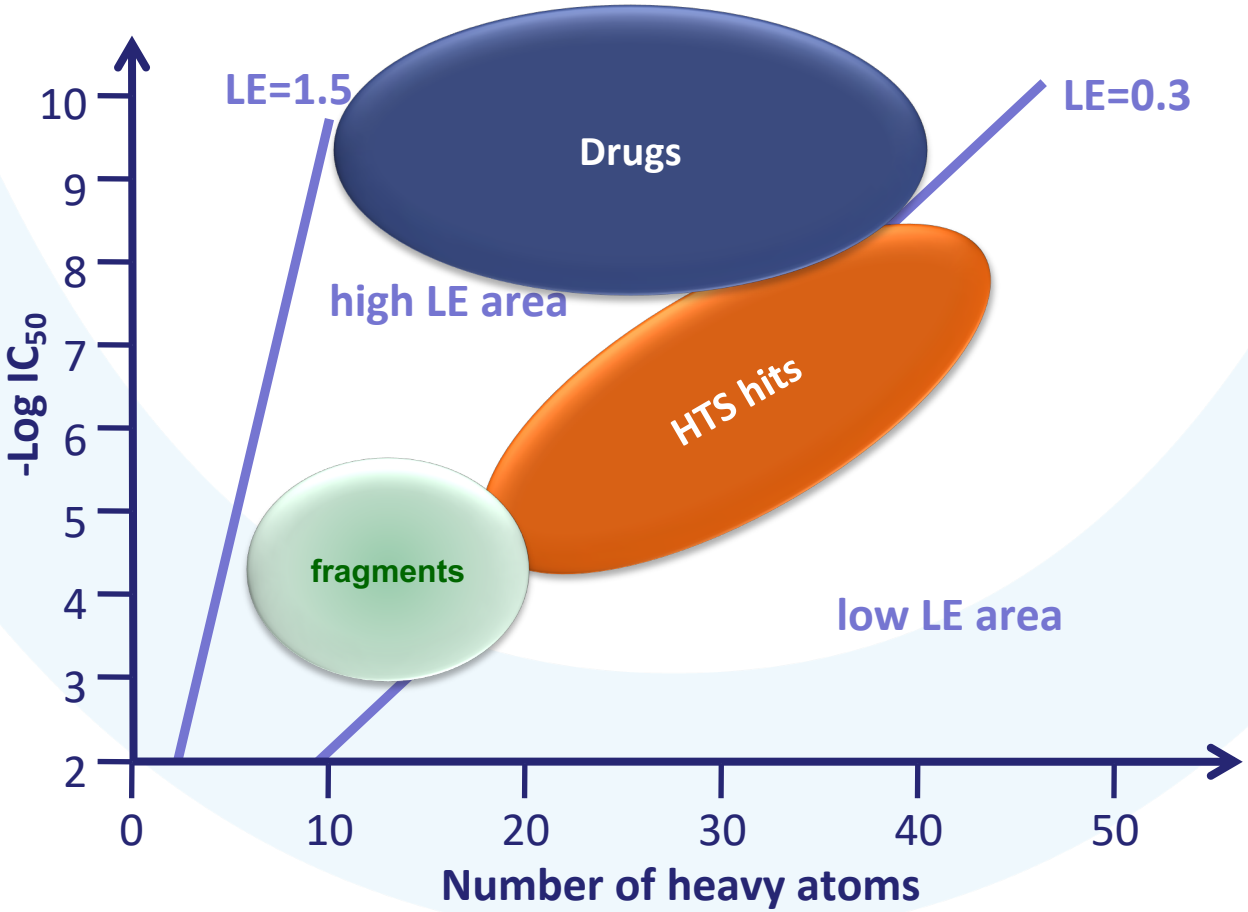
Smaller is better…



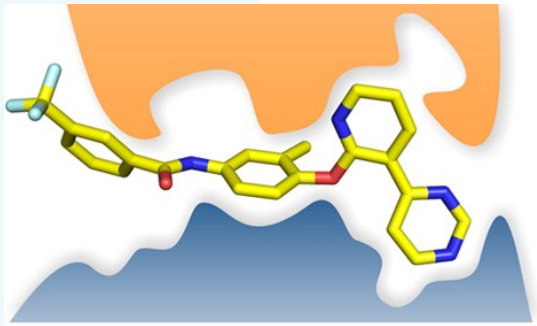
Scott, 2012

Fragments:
« Easier » optimization

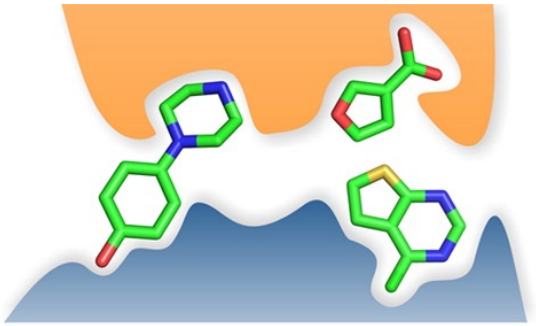
Smaller is better…



$$LE = \frac{-\Delta G\ (kJ/mol)}{nb.of\ non\ hydrogen\ atoms}$$



HTS hit: multiple suboptimal interactions

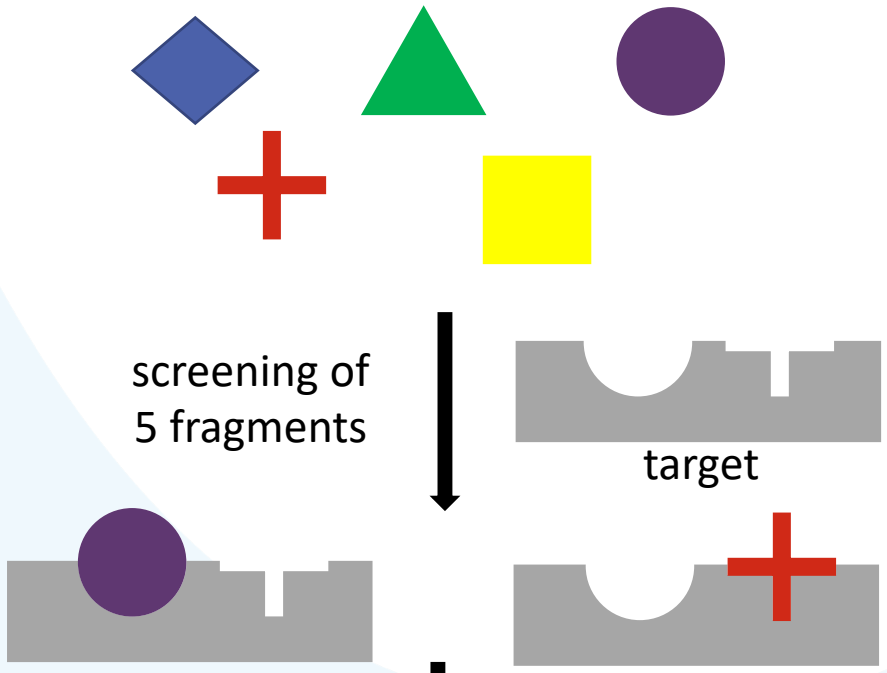


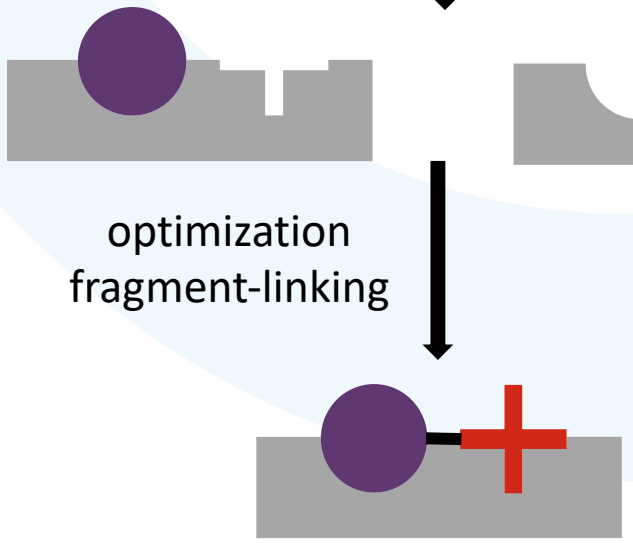
fragment hit: few optimized interactions

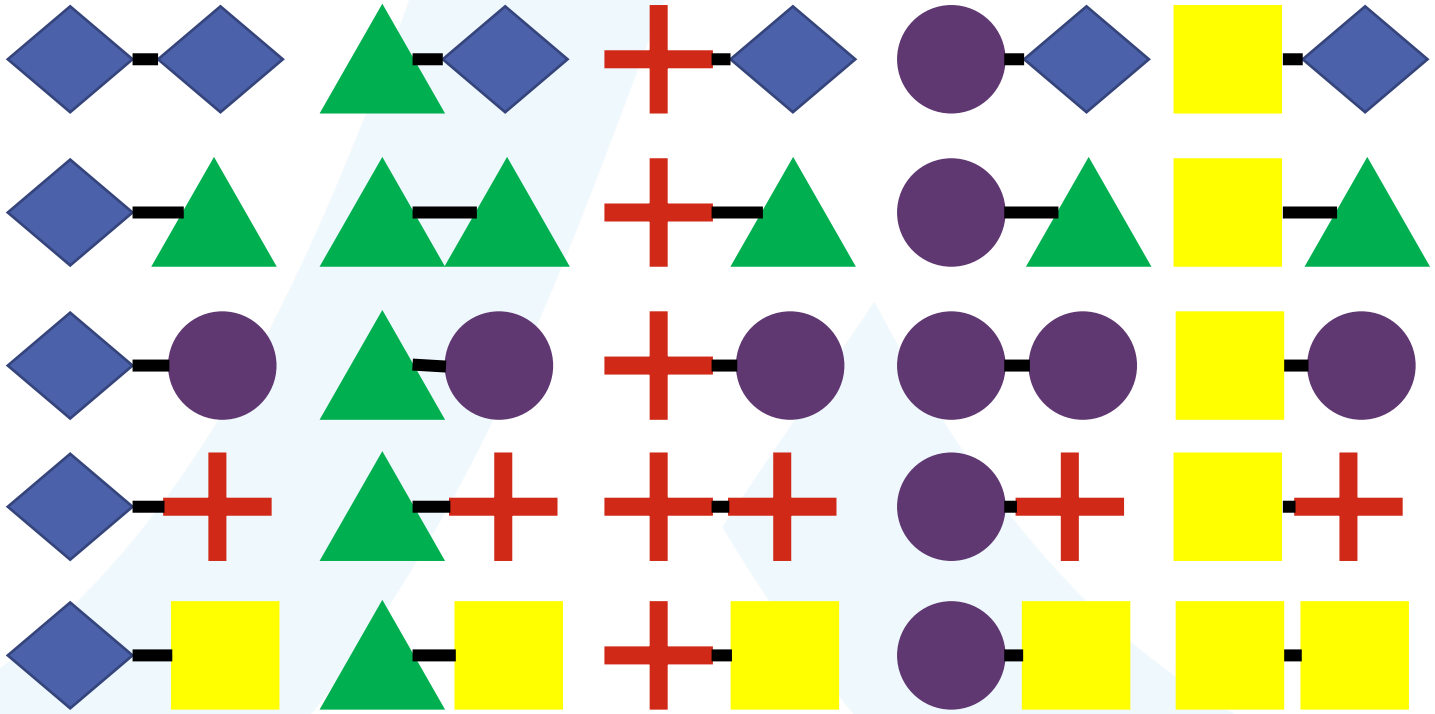
Fragments: better ligand efficiency

Siegal, 2008
Abad Zapatero, 2005
Scott,2012

Smaller is better…



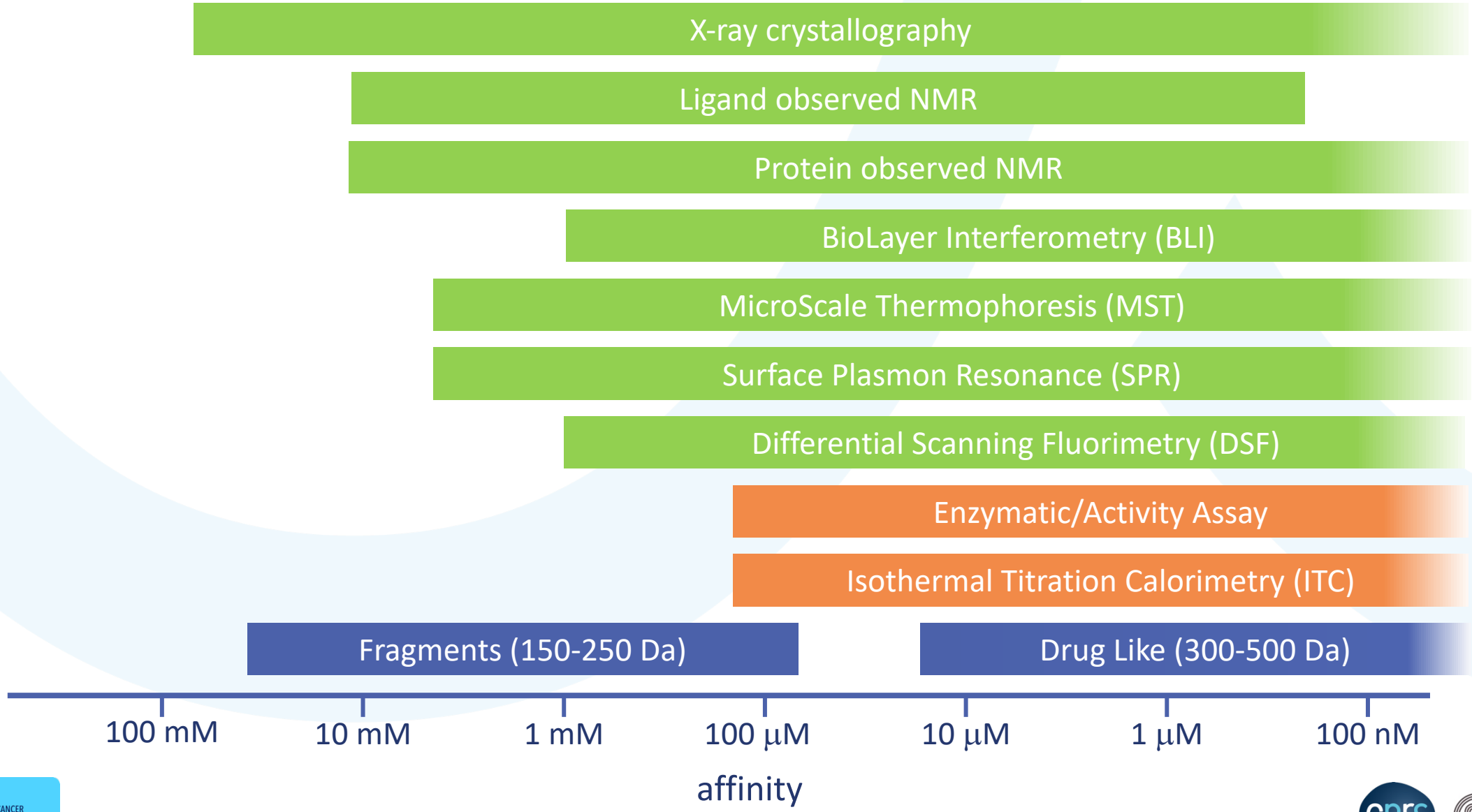




Fragments:
a higher hit rate

Galy, 2012

…but binding is weaker

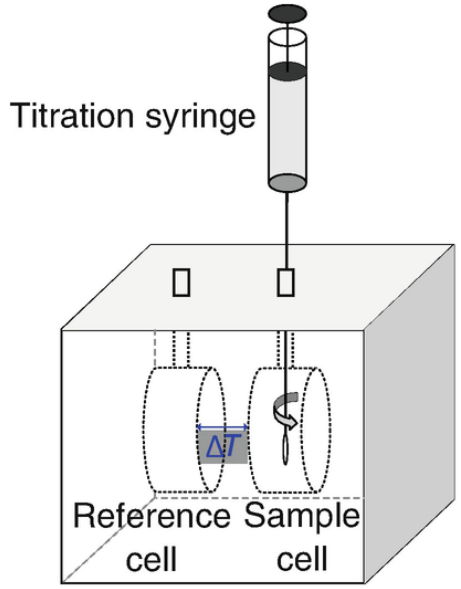


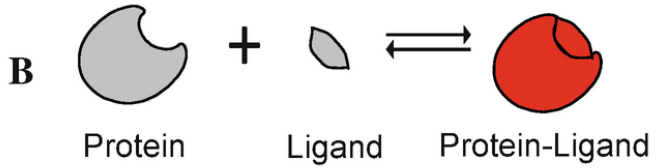
Adapted from Hubbard, 2011

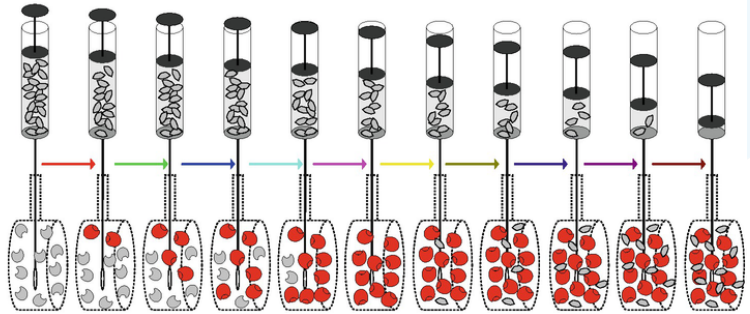
Overview

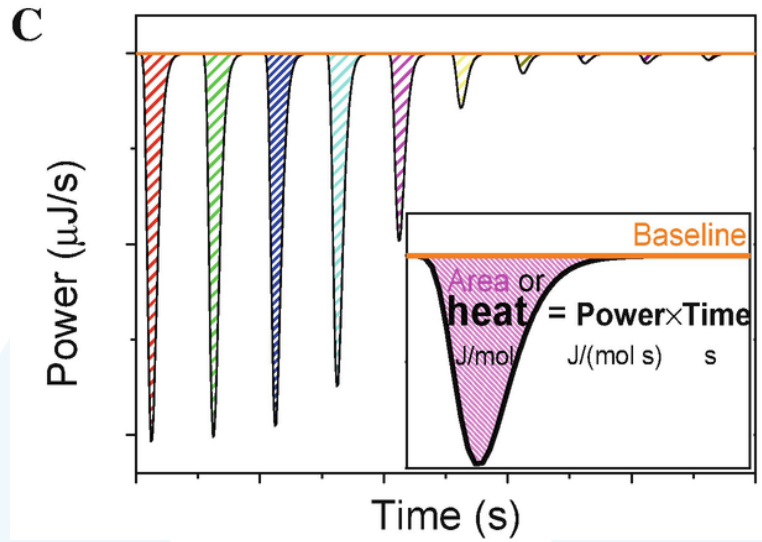
- The drug discovery process
- How can we find our first hit?
- FBLD: smaller is better
- FBLD: screening methods
- X-ray crystallography as a screening tool
- Practical considerations
  - Optimizing crystallization
  - Efficient soaking/co-crystallization
  - High throughput data collection and processing
  - Looking for binders
  - Refinement
  - PDB deposition
- Does it really work?

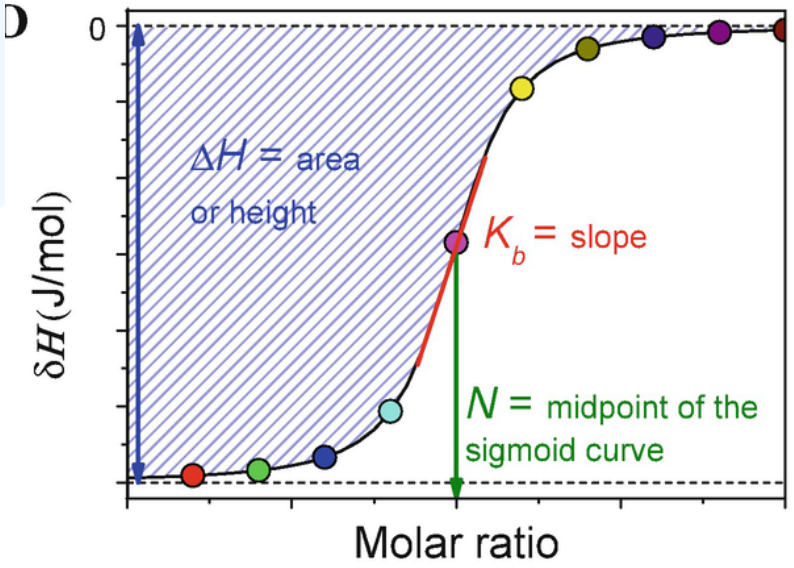
Isothermal Titration Calorimetry (ITC)





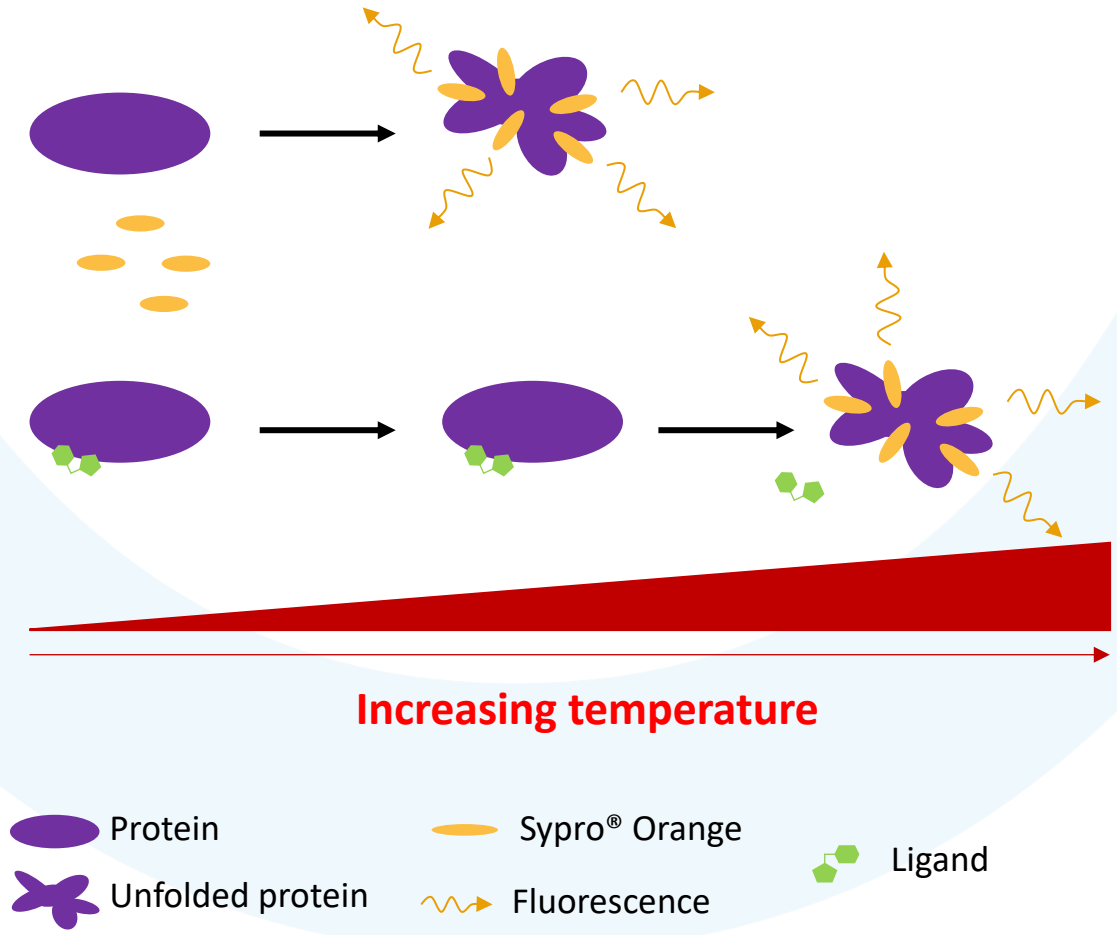


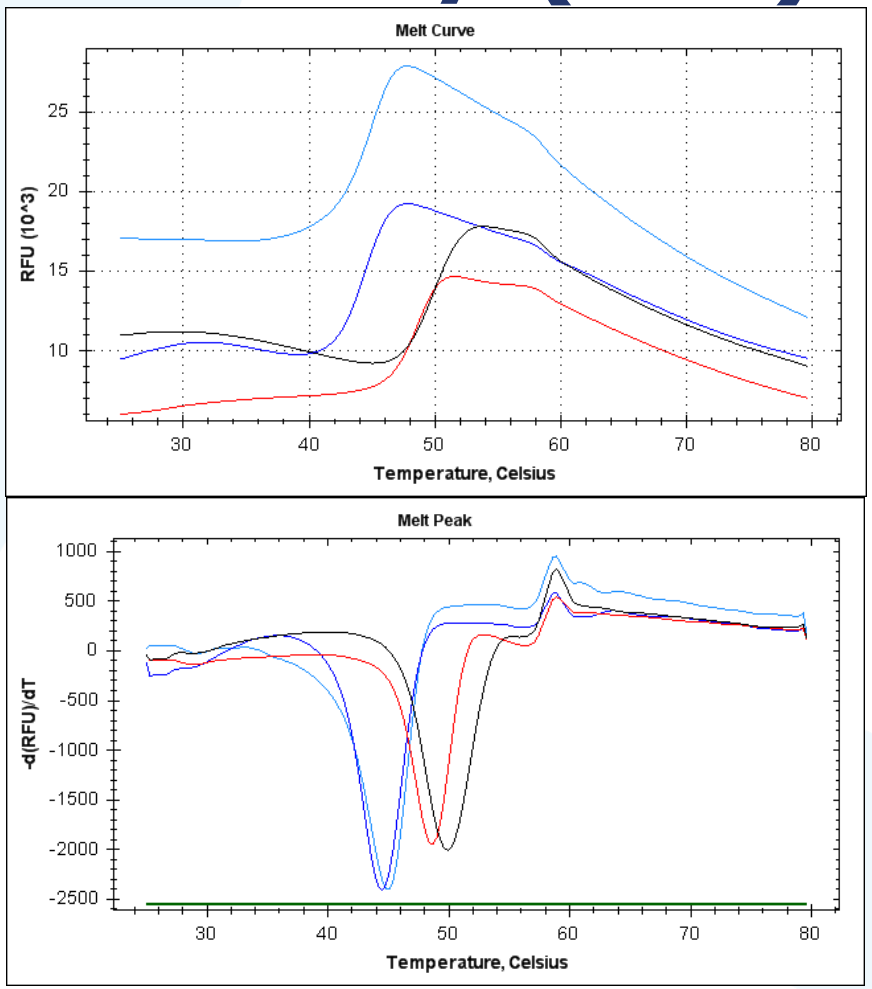




Paketuryté, 2018

Differential Scanning Fluorimetry (DSF)





protein

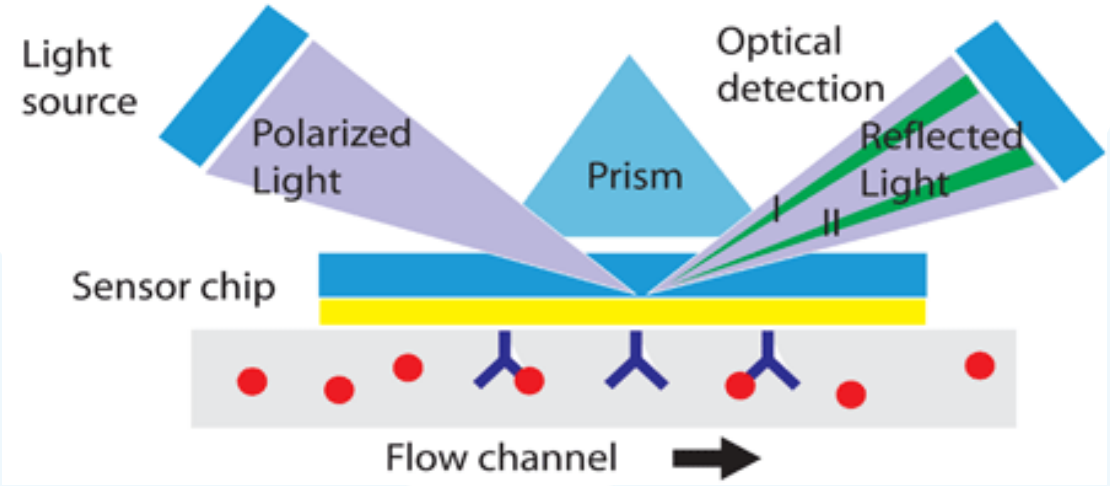
protein + fragment

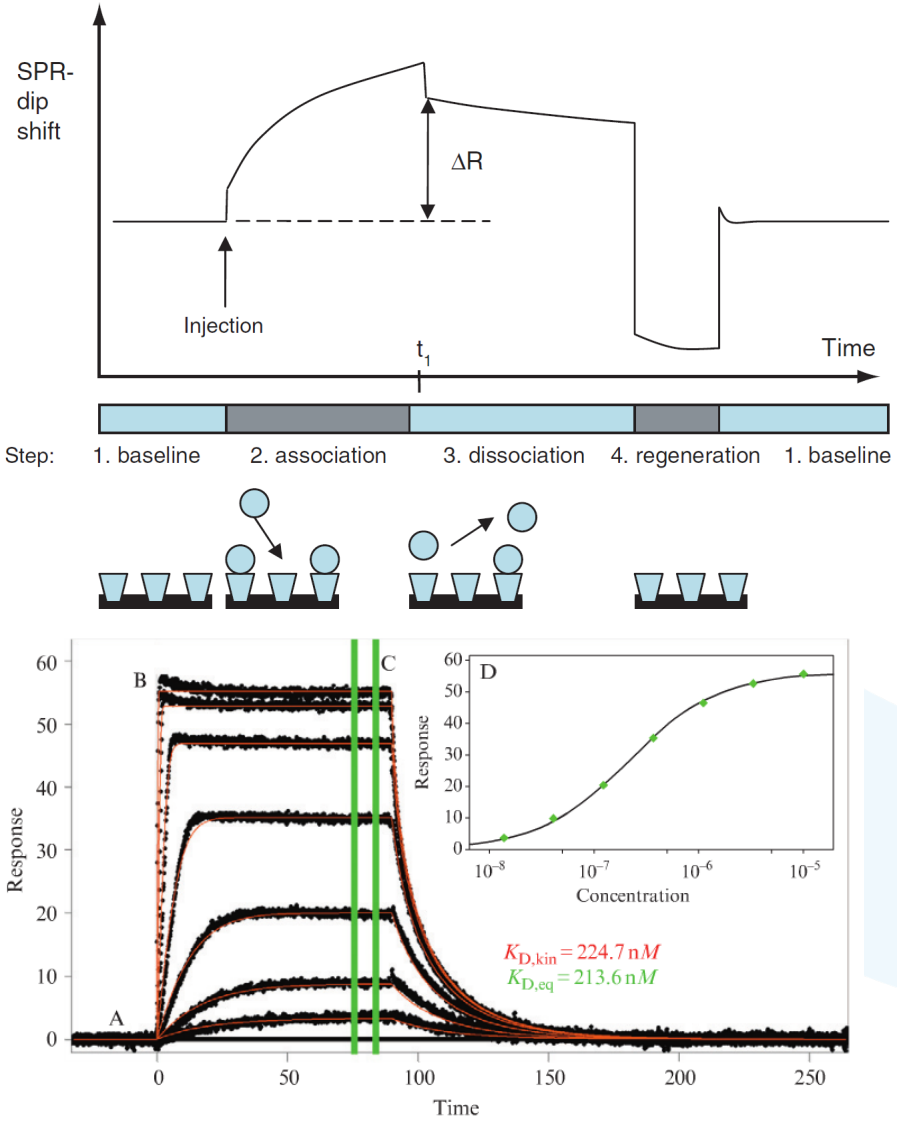
protein + ligand 1

protein + ligand 2

Galy, 2012

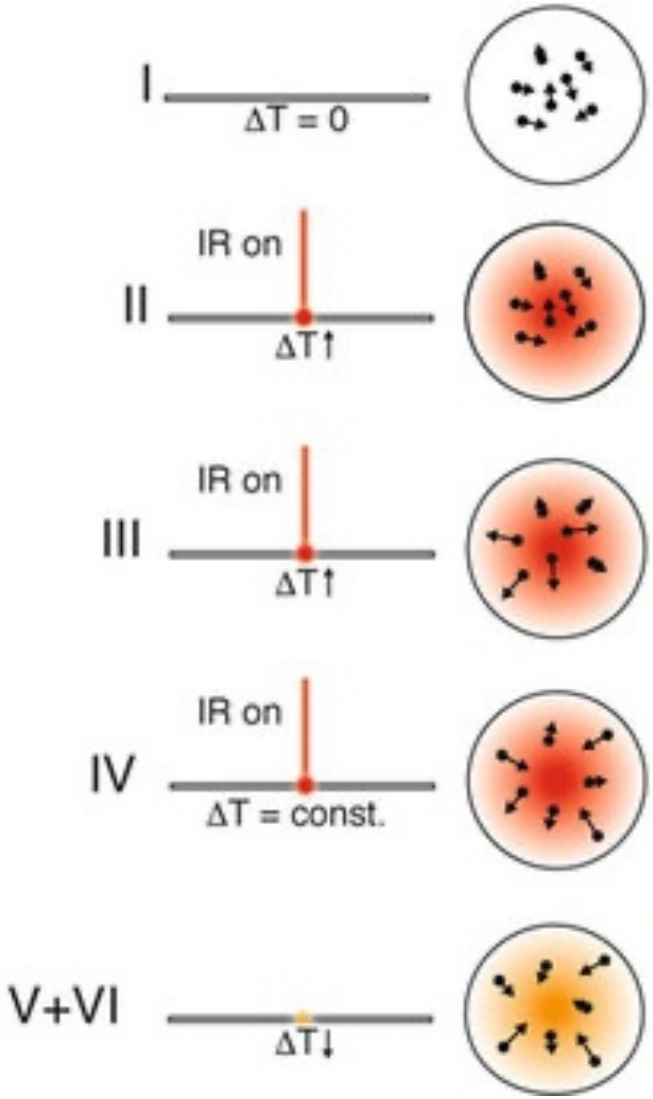
Surface Plasmon Resonance (SPR)

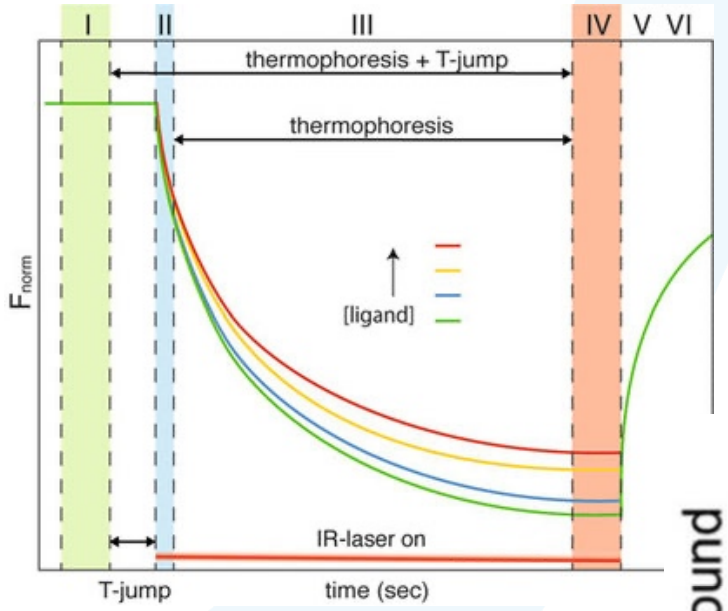


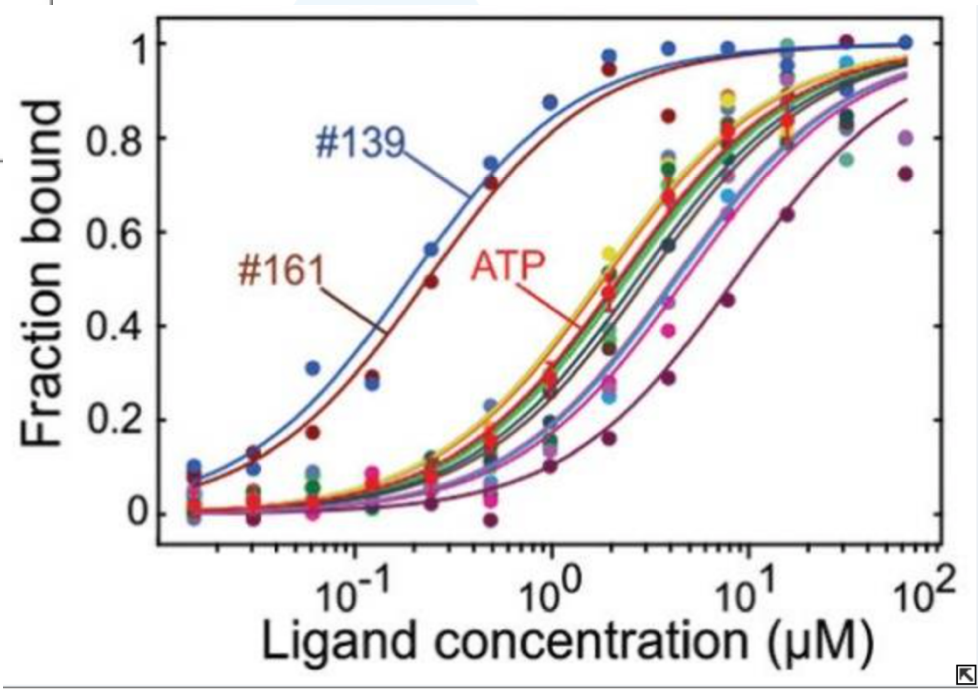


Tudos, 2008
Gianetti, 2011
http://nfs.unipv.it/nfs/minf/dispense/immunology/agabint.html

MicroScale Thermophoresis (MST)

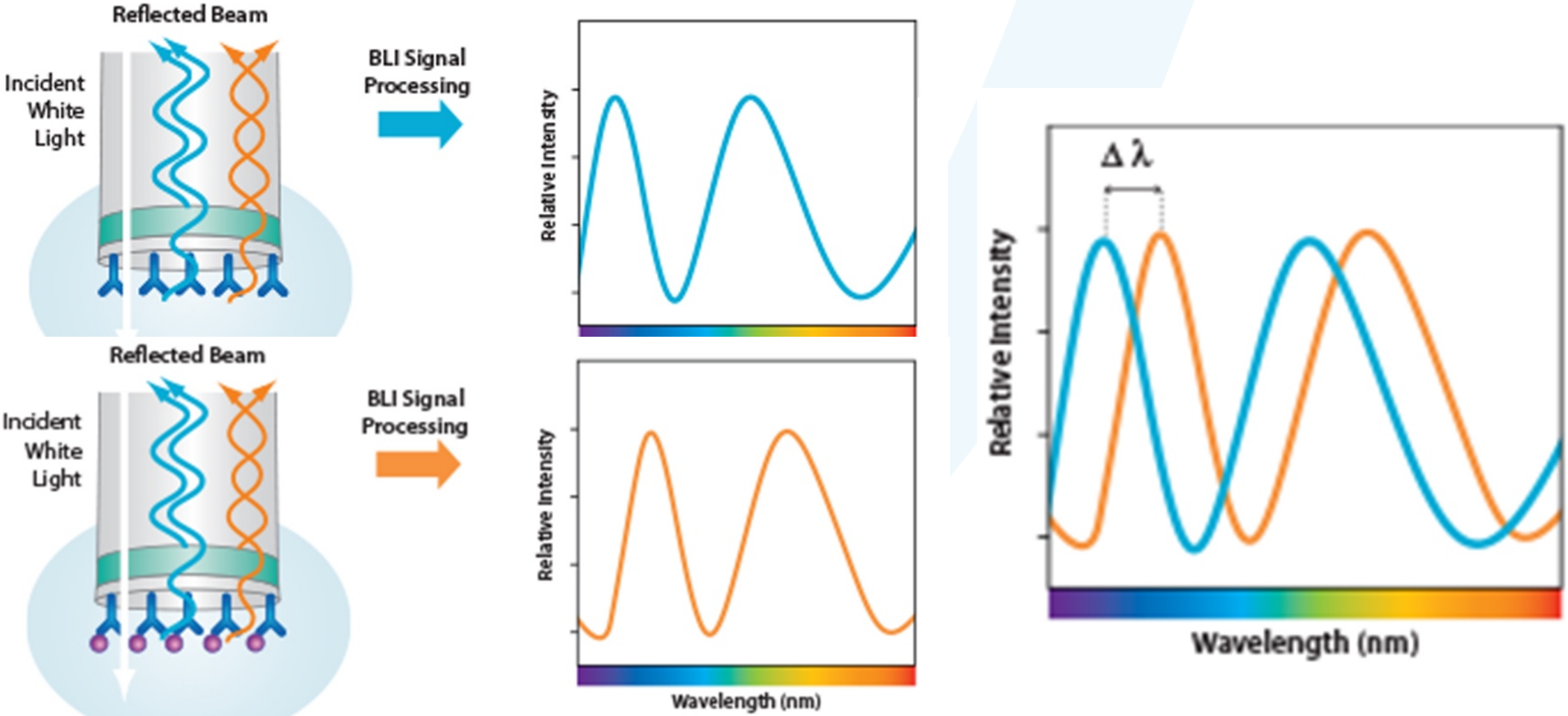






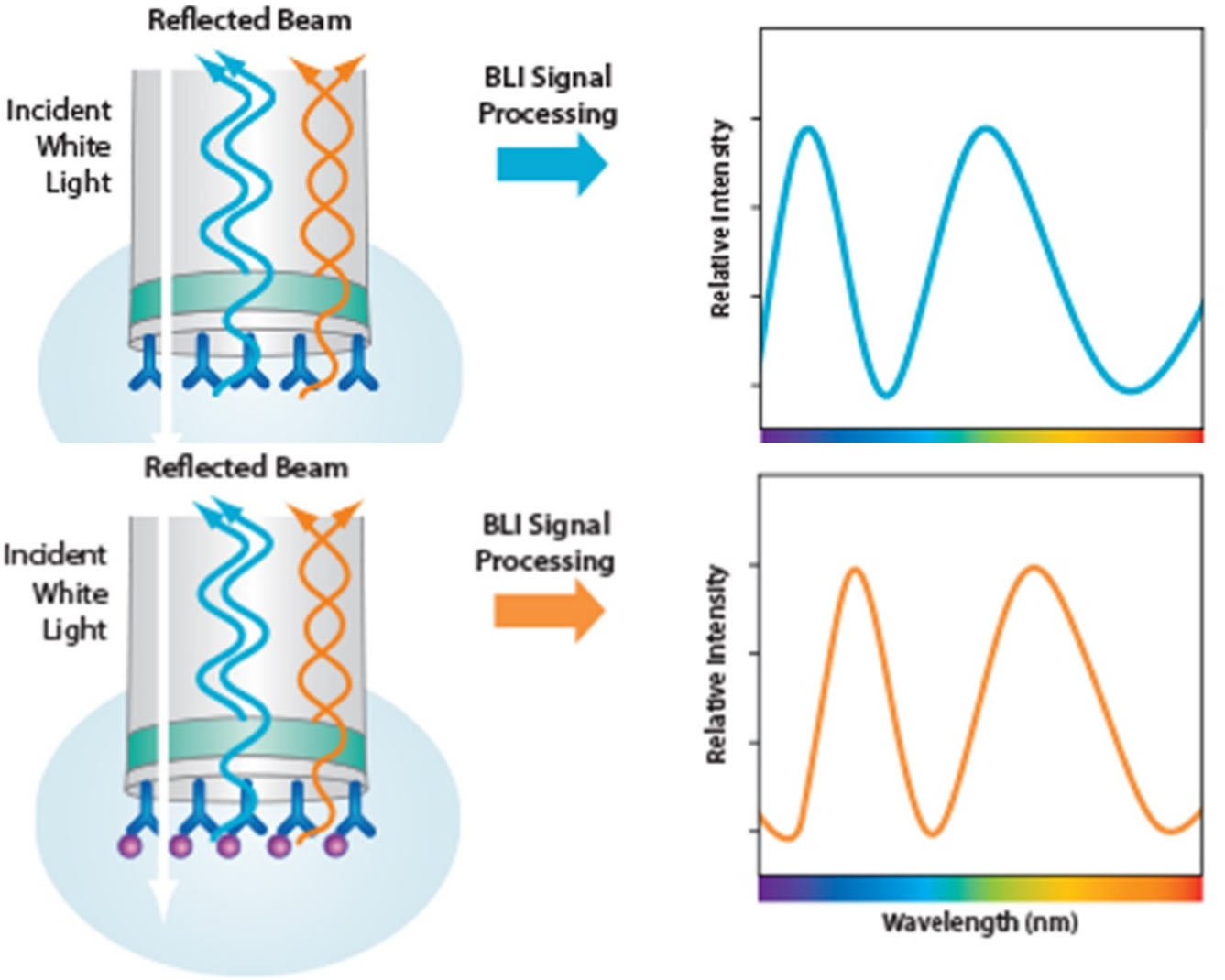
Corberski, 2018
Linke, 2016

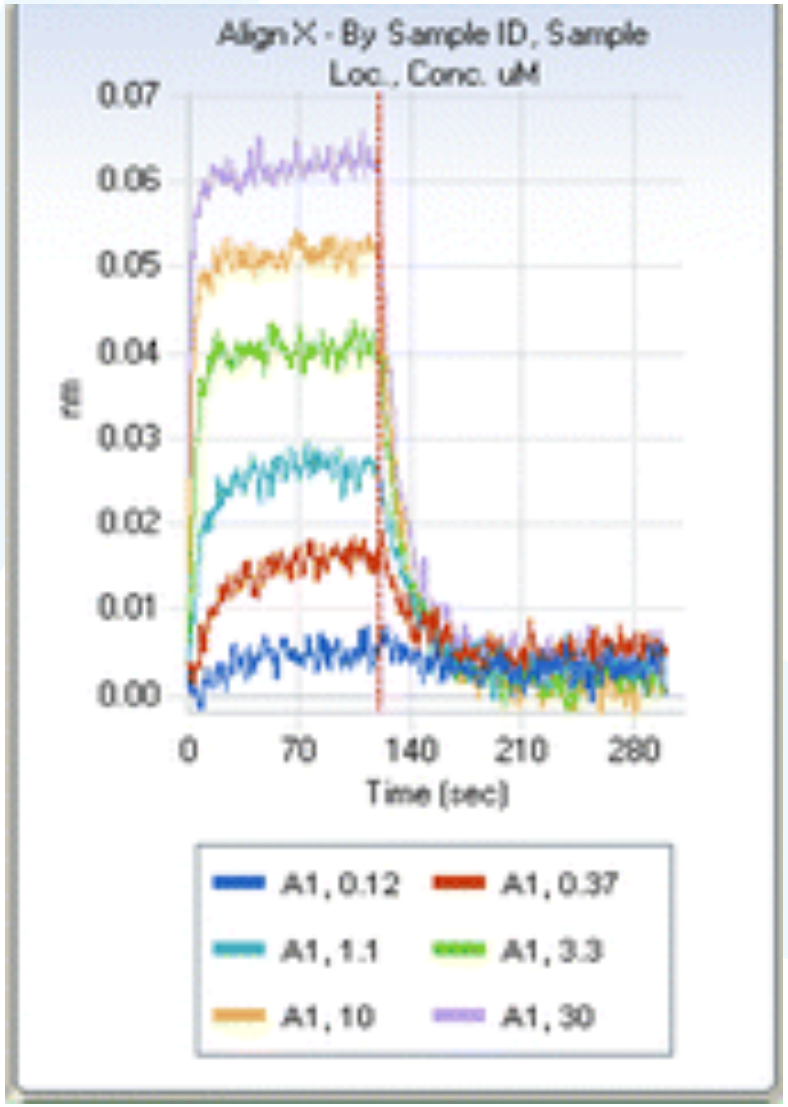
BioLayer Interferometry (BLI)



https://www.creative-biostructure.com/
Wartchowv, 2011

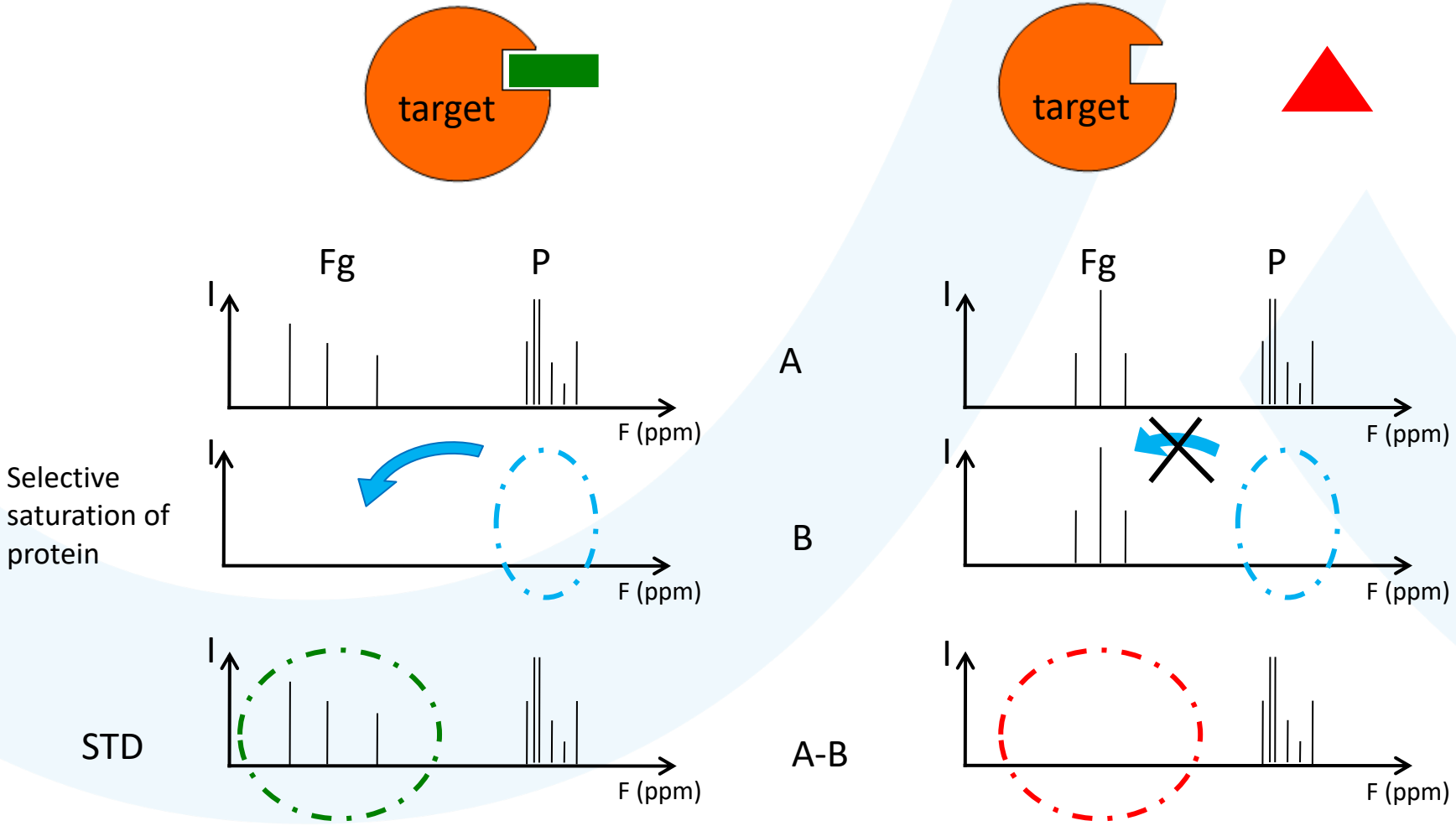
BioLayer Interferometry (BLI)





https://www.creative-biostructure.com/
Wartchowv, 2011

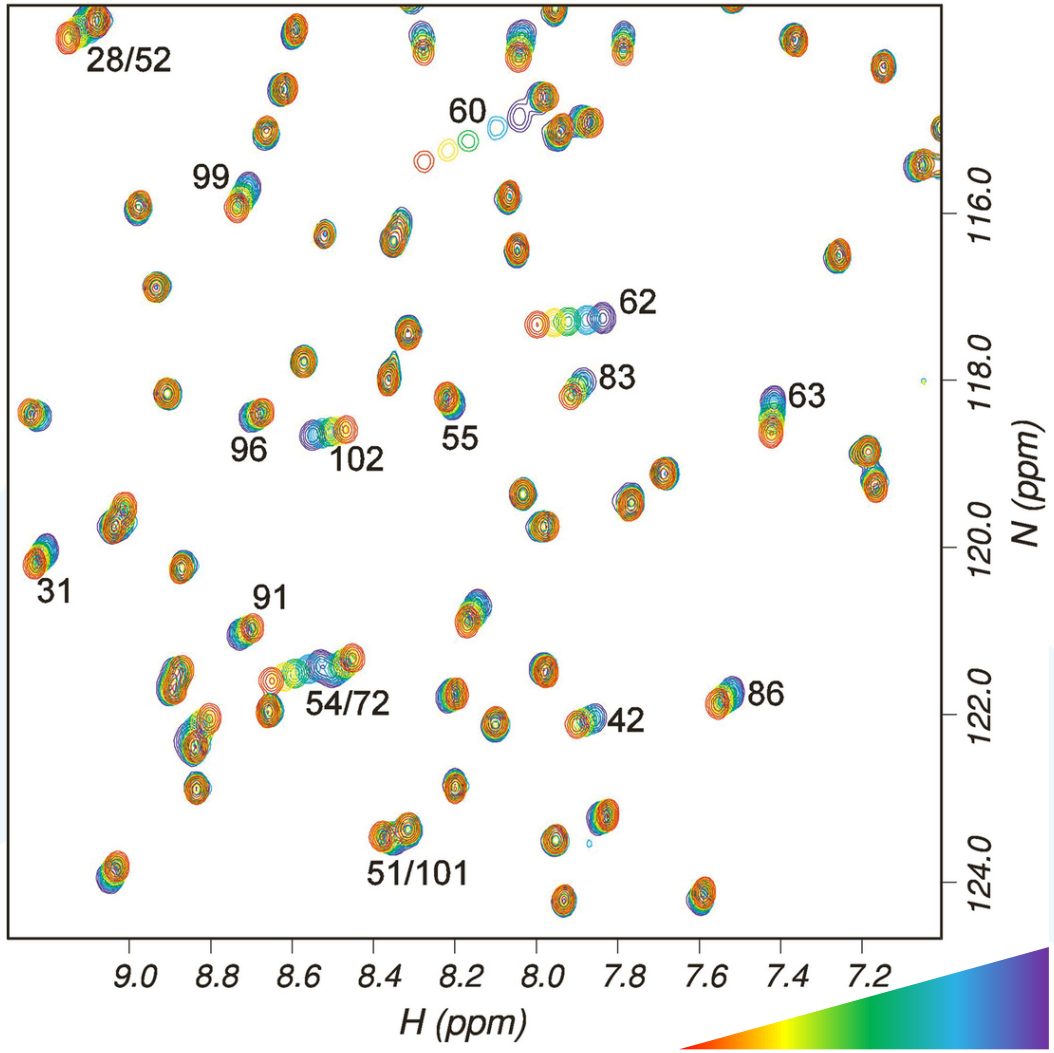
NMR: ligand observed NMR (STD)



Galy, 2012

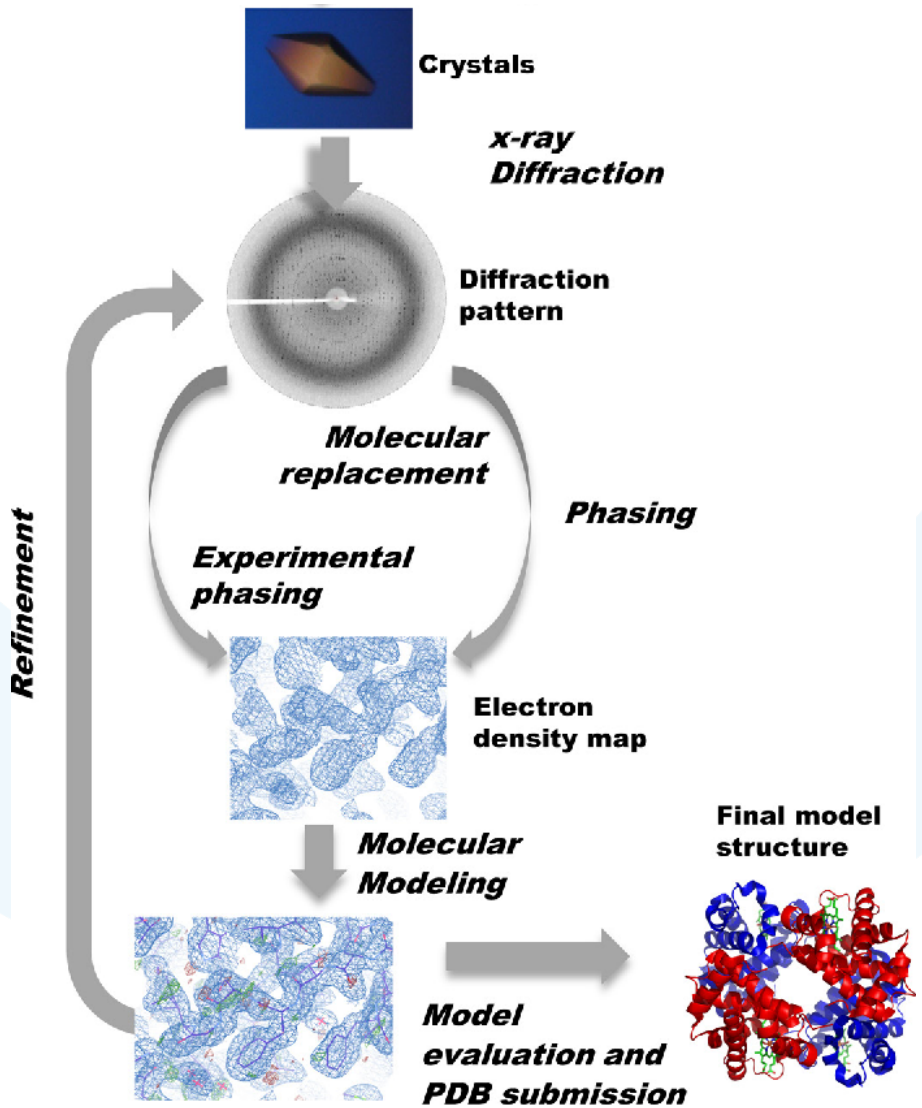
NMR: protein observed (CSP)

[^{15}N, ^{1}H]-HSQC spectrum of a ^{15}N labelled protein in absence and in presence of (increasing concentrations of) an unlabelled ligand.



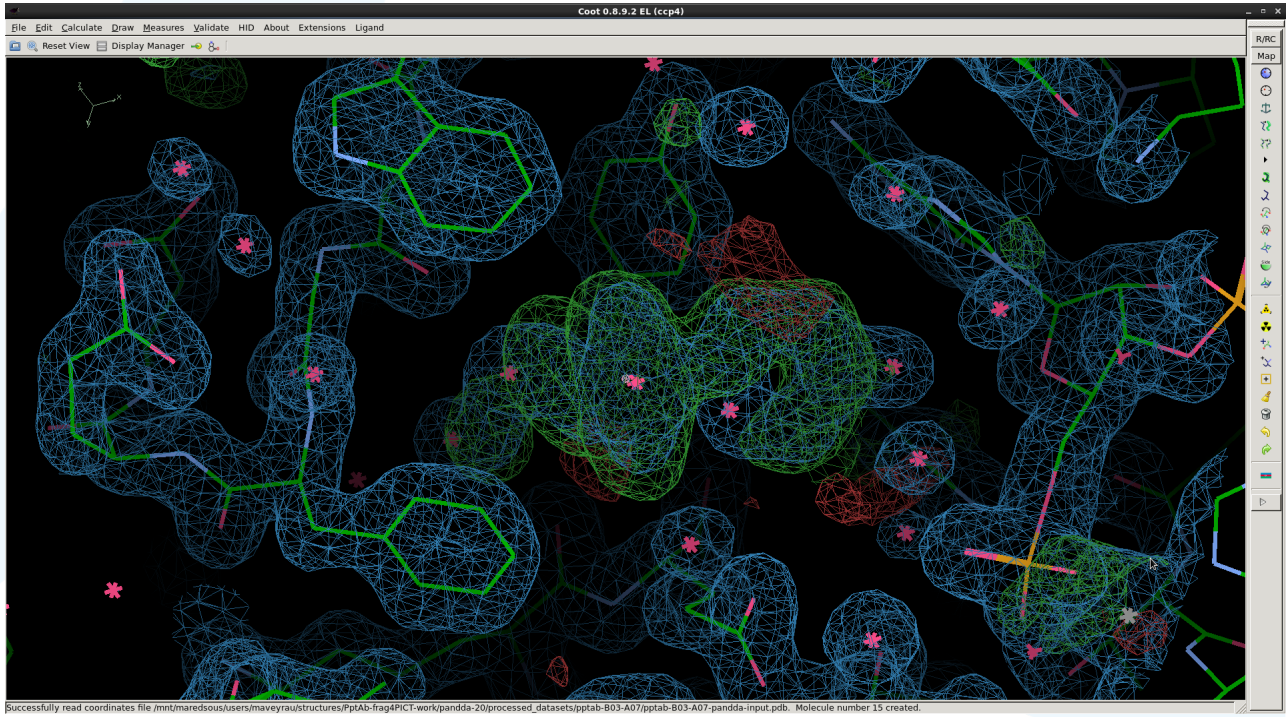
Adapted from Williamson, 2018

Xray crystallography



Büttner, 2015

Fapo-Fcomplex difference map after crystal soaking:



Biophysical techniques: what do we see?

Method	Label free	Immobilized	Kd range	Binding	Site	Structure
ITC			10^{-3} - 10^{-9}			
DSF			10^{-3} - 10^{-12}			
SPR			10^{-3} - 10^{-11}			
MST			10^{-3} - 10^{-12}			
BLI			10^{-3} - 10^{-11}			
STD			10^{-2} - 10^{-10}			
CSP	^{15}N		10^{-3} - 10^{-9}			
X-ray			10^{-1} - 10^{-12}			

Adapted from Kairys, 2019

Overview

- The drug discovery process
- How can we find our first hit?
- FBLD: smaller is better
- FBLD: screening methods
- X-ray crystallography as a screening tool
- Practical considerations
  - Optimizing crystallization
  - Efficient soaking/co-crystallization
  - High throughput data collection and processing
  - Looking for binders
  - Refinement
  - PDB deposition
- Does it really work?

X-ray crystallography provides the most comprehensive information

- No false positive (but false negative are possible)
- Binding site is known
- Structure of the complexe is known, ready for hit optimisation
- But…
  - It is slow (but not that slow)
  - It needs carefull thinking beforehand
  - It requires a lot of beam time…

Overview

- The drug discovery process
- How can we find our first hit?
- FBLD: smaller is better
- FBLD: screening methods
- X-ray crystallography as a screening tool
- Practical considerations
  - Optimizing crystallization
  - Efficient soaking/co-crystallization
  - High throughput data collection and processing
  - Looking for binders
  - Refinement
  - PDB deposition
- Does it really work?

FBLD by X-ray crystallography

A typical fragment library includes 1000 compounds

- Optimize you crystallisation conditions:
  - Reproducible, cryo-ready, high-resolution (better than 2.2 A)
- Preparation of complexes:
  - Soaking, co-crystallization
- Crystal cooling
- Data-collection and processing:
  - Unattended automatic data collection & processing
- Detect structures with bound ligand ASAP:
  - PANDDA, pipedream…
- Focus only on complexes (expect 50-100 structures!)

Optimization of crystallization

Many crystals will be needed (at best 1 crystal/fragment, but likely more):

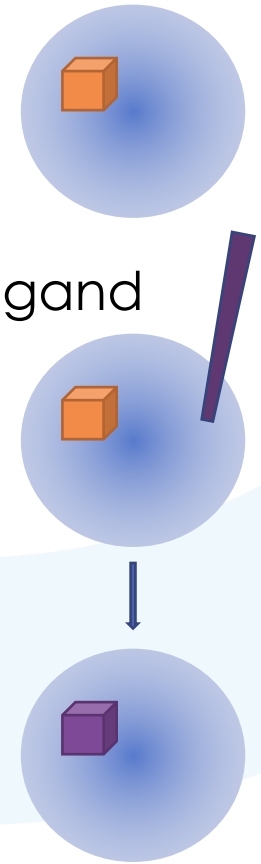
- Aim at > 90% success rate (eg about 80-90 drops with crystals in a 96-well plate)
- If possible, use a cryo compatible crystallization condition (you don't want to cryoprotect each crystal!)
- Check that the crystals diffract to high resolution
- Check that crystals tolerate a few % DMSO
- Check that crystals allow for complex formation
- Higher symmetry spacegroup preferred (not P1!)

Soaking vs. Co-crystallization

Soaking: get the crystal and add the ligand

1) Crystallize protein

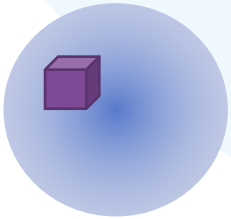
2) Add droplet with ligand



Co-crystallization: add the ligand and get the crystal

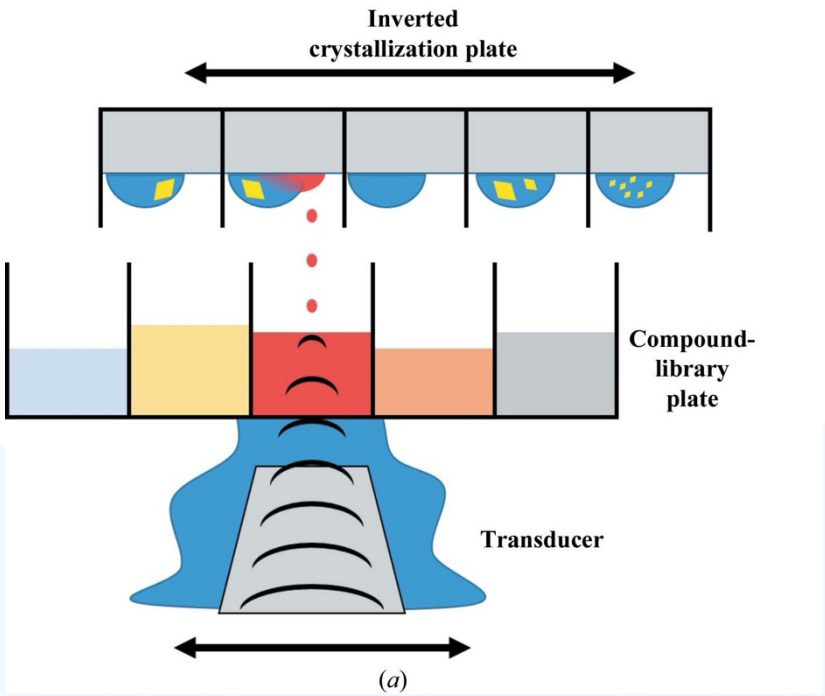
1) Add ligand to protein solution

2) Crystallize the complex

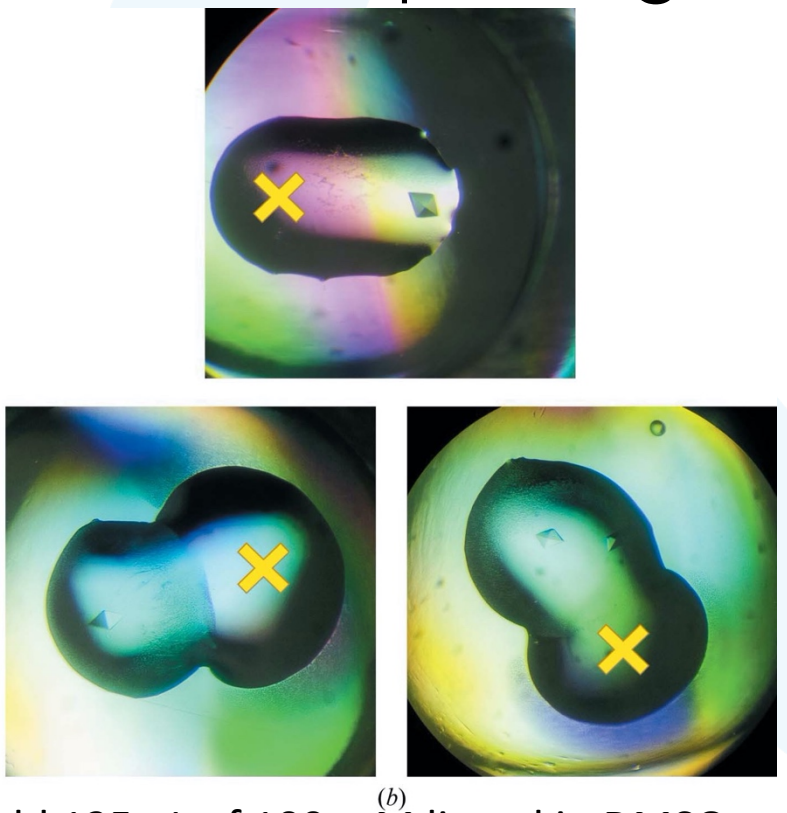


Soaking vs. Co-crystallization

- High throughput soaking: use acoustic dispensing



Allow for accurate dispensing of nanovolumes to defined positions



Add 135 nL of 100 mM ligand in DMSO to a 200 nL crystal-containing droplet

Collins, 2017

Soaking vs. Co-crystallization

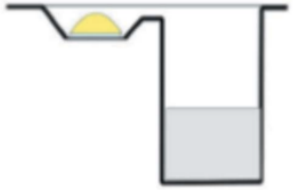
- High-throughput co-crystallization: pre-coating of crystallization plates

Dispensing ligand solution

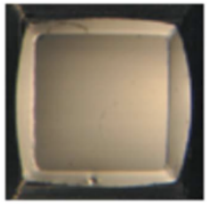
1 µL, 10 mM
in 100 % DMSO

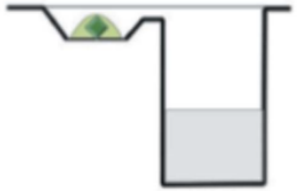
Solvent evaporation

1-2 days at RT



Dispensing protein and mother liquor





Protein crystal growth

Gelin, 2015

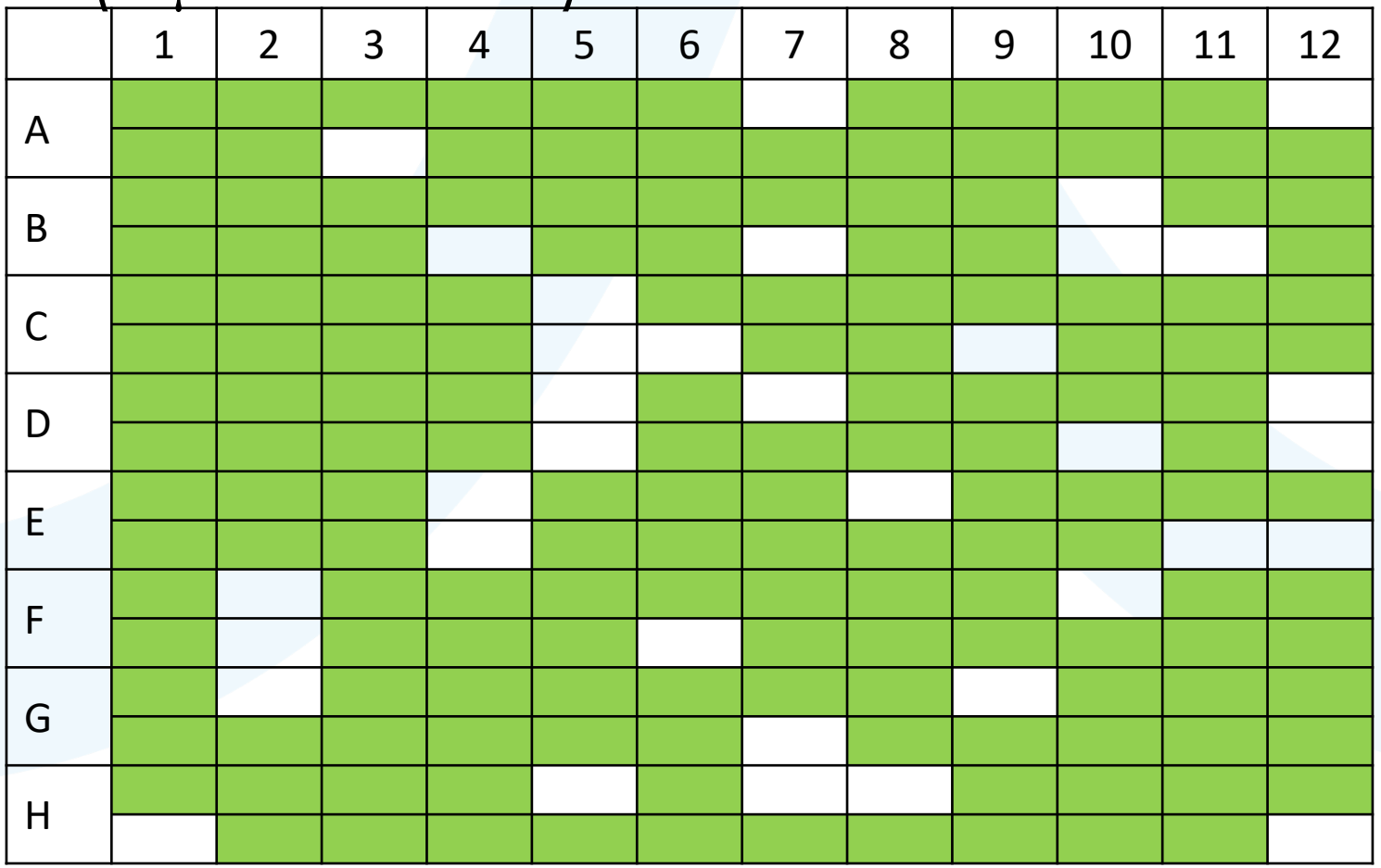
Soaking vs. Co-crystallization

- High-throughput co-crystallization: pre-coating of crystallization plates (1 µl at 10 mM)

192 crystallization drops in the presence of dried fragment

82 % success rate

(95 % in absence of dried compound) .



Data collection and processing

You will need **a lot of** beamtime (about 7 days cumulated beamtime for our 940-fragments library).

Unattended automatic data collection is the way to go:

MASSIF-1 beamline at ESRF (about 150 dataset/day).

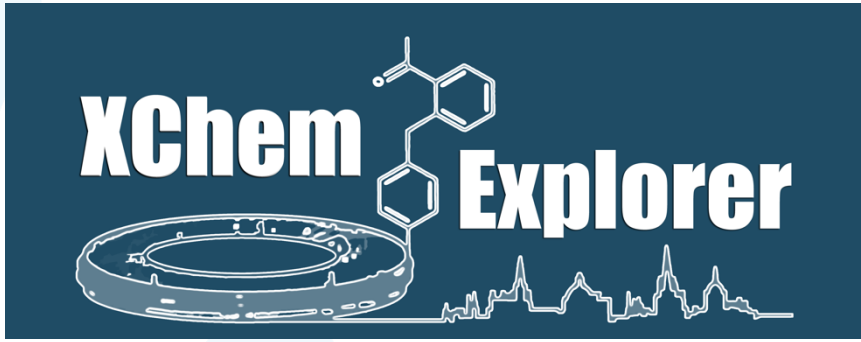
Data processing is performed automatically with various software (XDS, XDSAPP, AutoProc, GRENADES…).

Selecting the best data processing is manual…

Data collection and processing

How to keep track of everything?

Hundreds of data collection and data sets are expected….



XChemExplorer is developped at the Diamond Light Source

Usable with data from any sources… with some tweaking…

Krojer, 2017

XChemExplorer

1 directory per fragment :

```
PICT-B01A01
PICT-B01A02
PICT-B01A03
…
…
…
PICT-B10H09
PICT-B10H10
PICT-B10H11
PICT-B10H12
reference
```

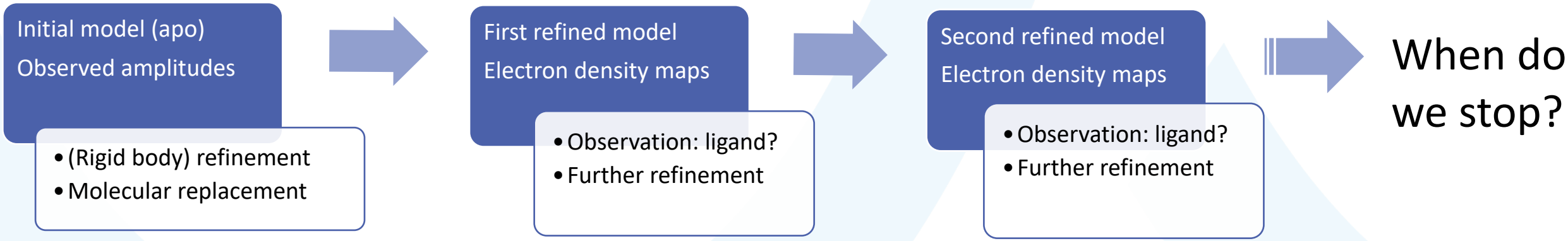
In each directory:

`PICT-B01A01.log` : aimless logfile
`PICT-B01A01.mtz` : MTZ with processed data
`fragment.cif` : fragment CIF file

In reference directory:

`reference.pdb` : refined apo protein structure
`reference.mtz` : apo protein dataset

Is a ligand present in the structure?



if (ligand) = yes then keep on refining
If (ligand) = no then

1. keep refining, but how long?
2. discard the structure… but what if?

dimple (CCP4)

pipedream (GlobalPhasing)

- Data processing (pipedream)
- Consistent indexing with reference, rigid-body/molecular replacement, restrained refinement (both)
- Analyse electron density maps (both)
- Build ligand and further refine (pipedream)

Is a ligand present in the structure?

Hopefully, XCE can automate the dimple process…

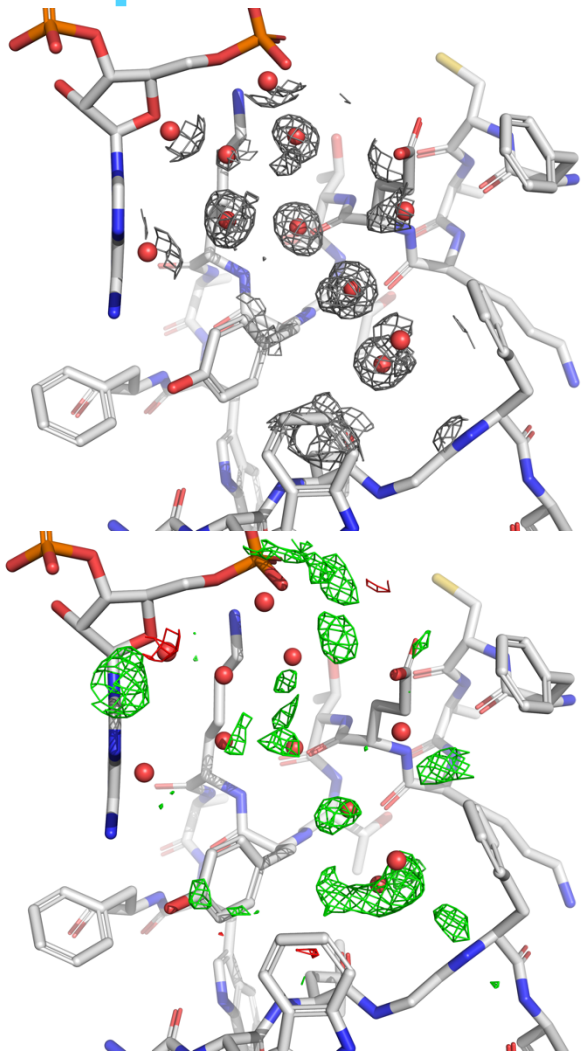
For each folder, if MTZ file present:

- Check indexing (`reference.mtz`)
- Rigib-Body refine (MolRep if needed)
- Perform restrained refinement (refmac)
- Compute maps and search for large peaks
- Score the peak (if any)

```
PICT-B01A01
PICT-B01A02
PICT-B01A03
…
…
…
PICT-B10H09
PICT-B10H10
PICT-B10H11
PICT-B10H12
reference
```

Latest version of XCE also handles pipedream!

Is a ligand present in the structure?



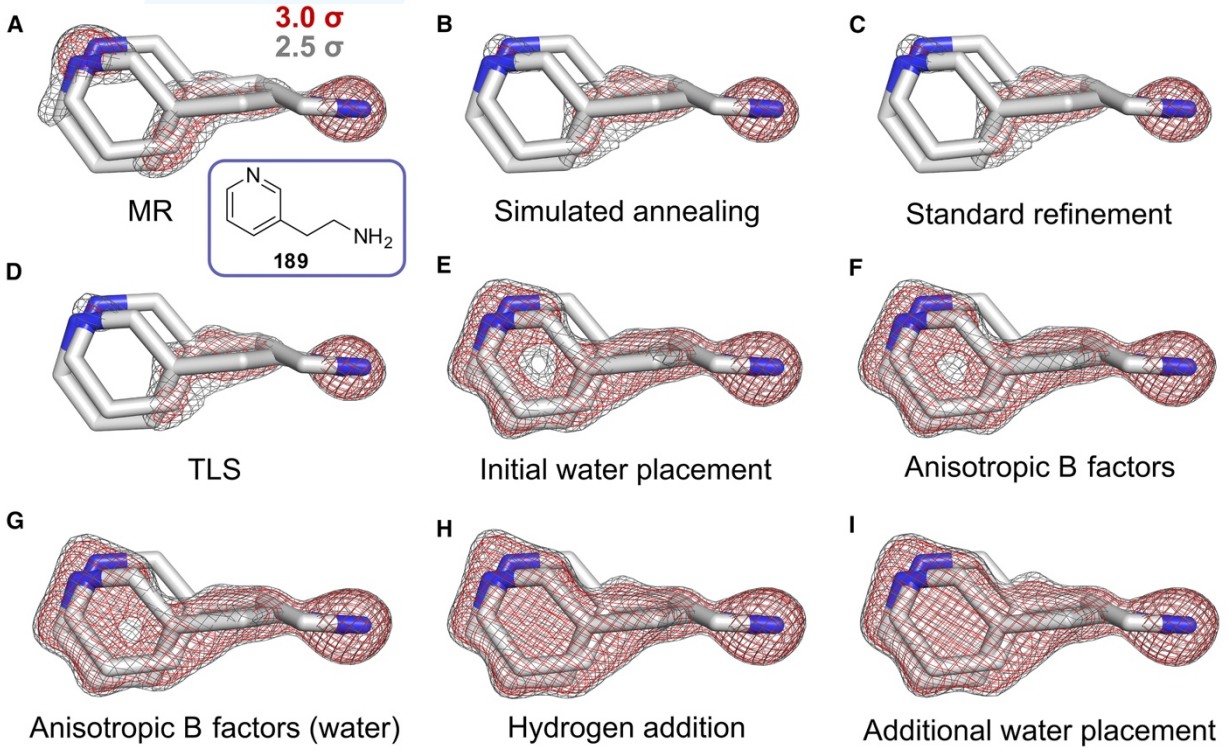
2Fo-Fc,
1sig

Nothing to be seen…

Discard?

Fo-Fc,
3sig

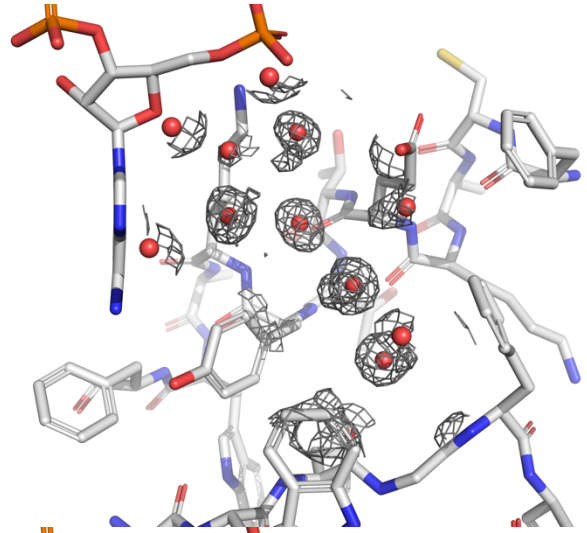
Difference map for ligand at various refinement stage…



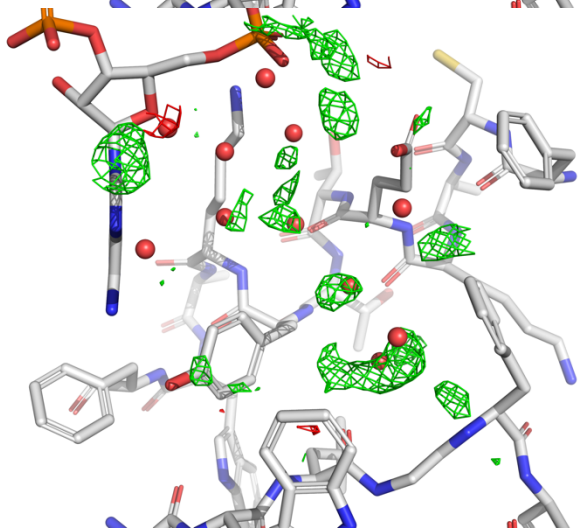
Maveyraud & Mourey, 2020
Schiebel, 2016

Is a ligand present in the structure?

What you see…

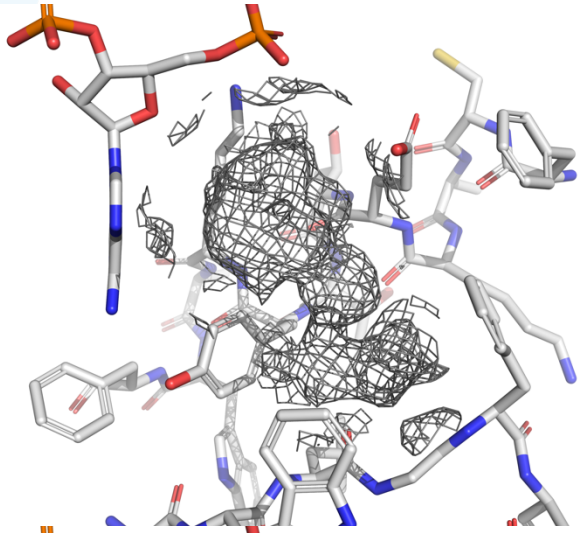


2Fo-Fc, 1sig

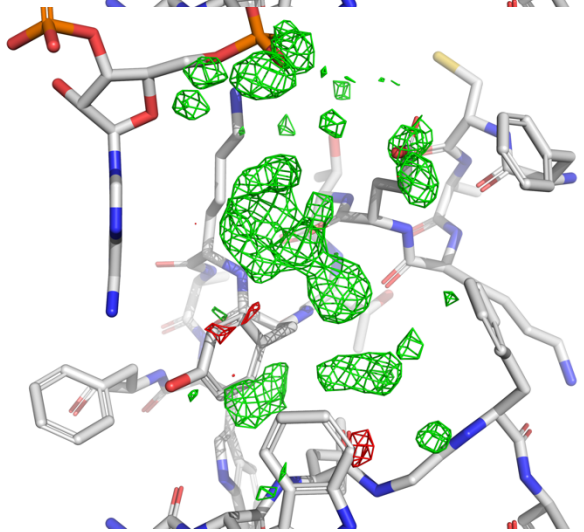


Fo-Fc, 3sig

What a Pandda sees…



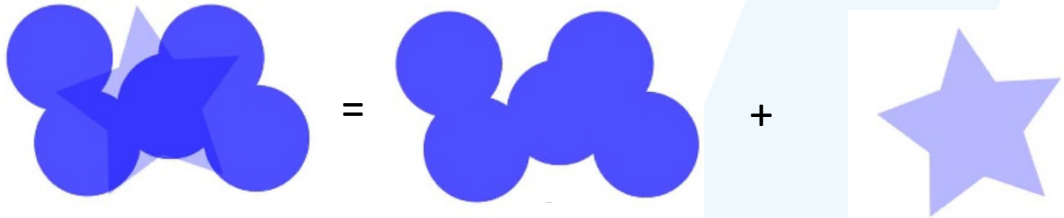
Event map



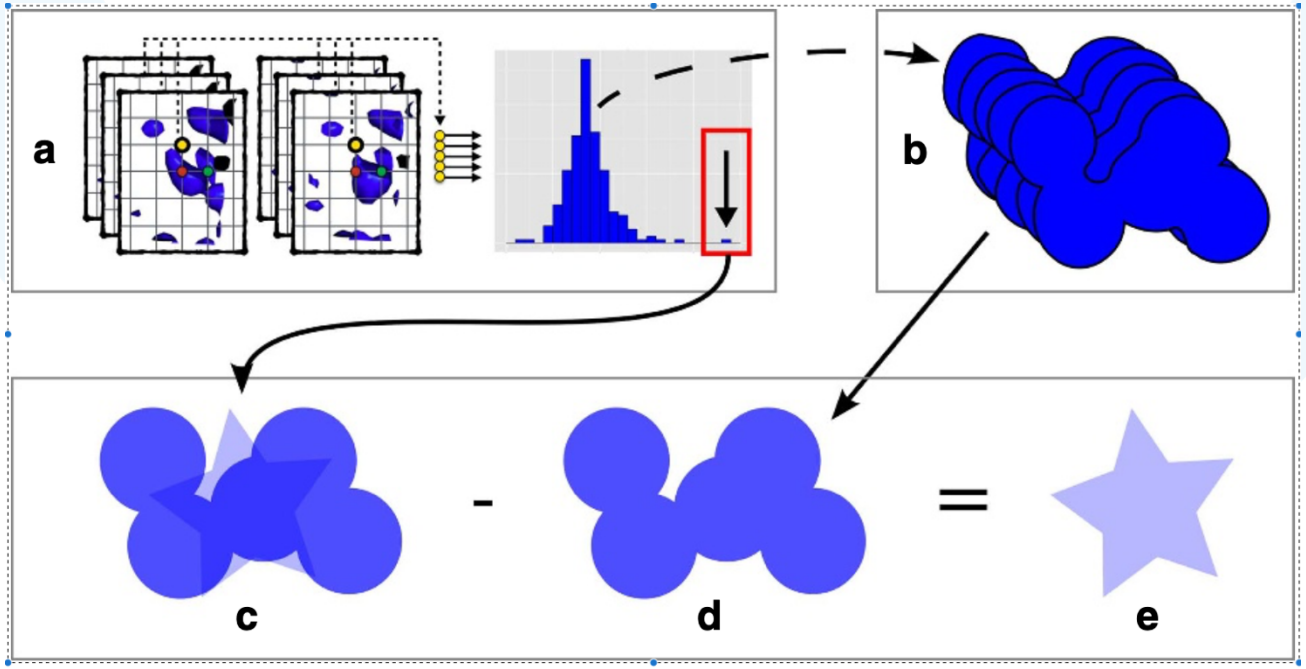
Z-map

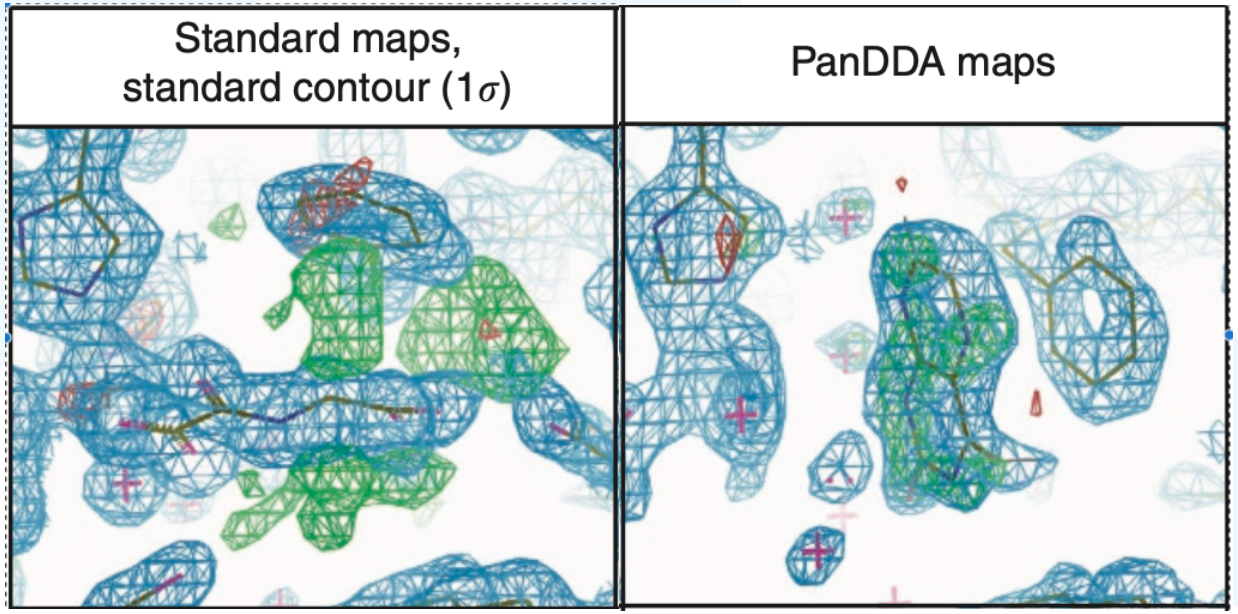
Maveyraud & Mourey, 2020
Pearce, 2017

How does Pandda work?



structure of complex = structure of apo + structure with bound ligand



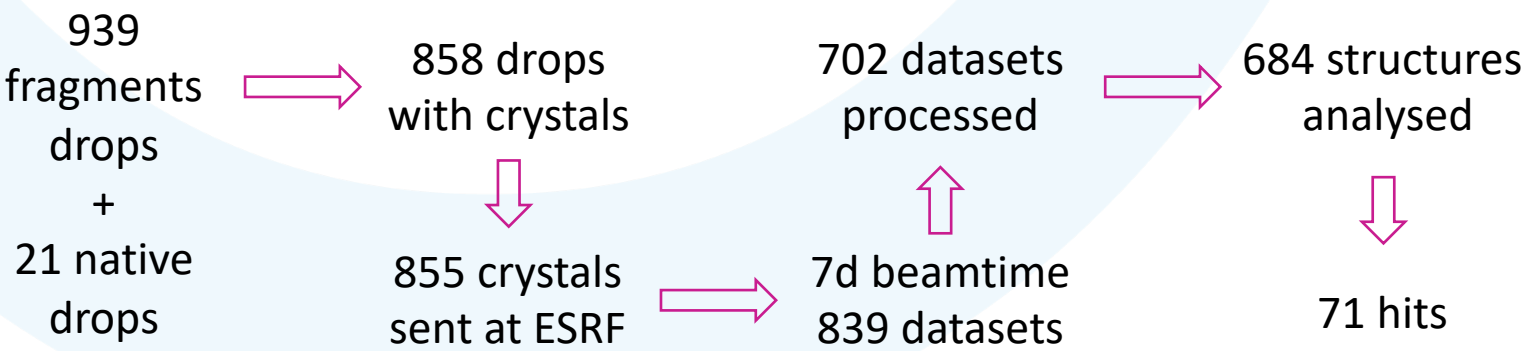


Pearce, 2017

Pandda – real life example

The project:

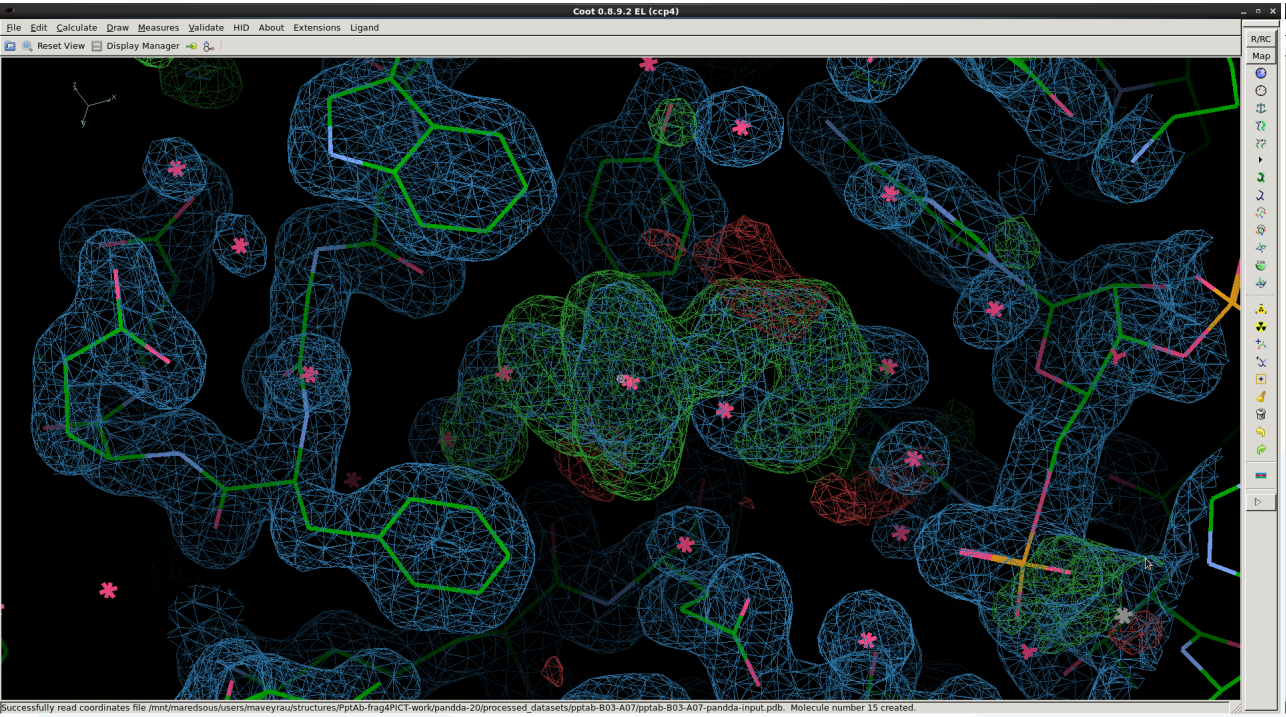
- Target: essential enzyme from *Mycobacterium abscessus*, *apo* structure includes CoA, 2 Mn2+ ions
- 939 fragments (PICT fragment library)



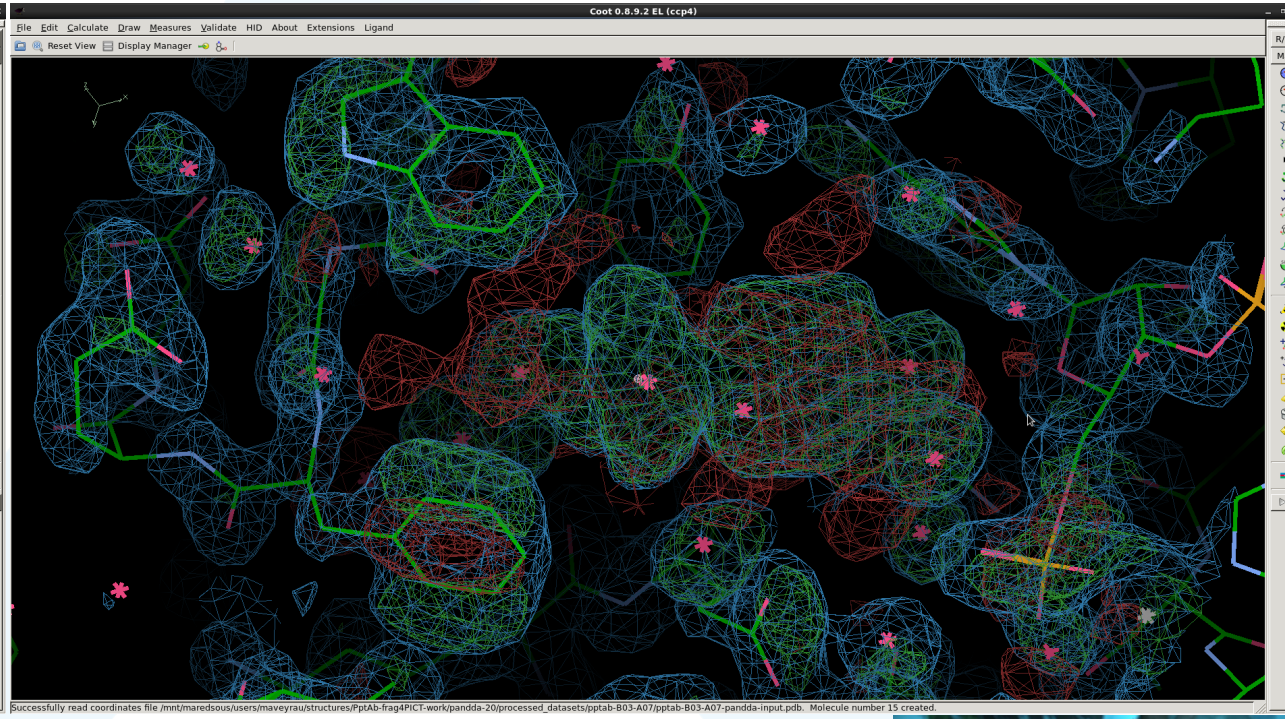
71 hits for 939 fragments (7.6 % hit rate)

Nguyen, 2020
Unpublished results

Pandda in real life: the good…



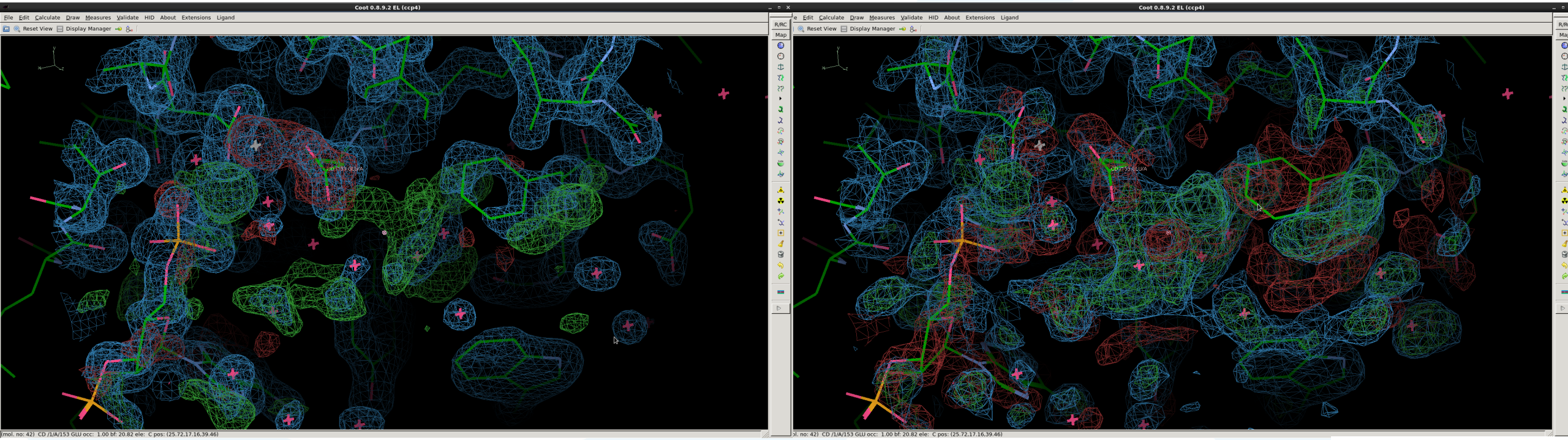
The dimple **Fo-Fc map**: presence of the ligand obvious, despite minute refinement.
=> Binding with high occupancy



The Pandda **event map**: presence of the ligand is also obvious

Unpublished results

Pandda in real life: the bad…

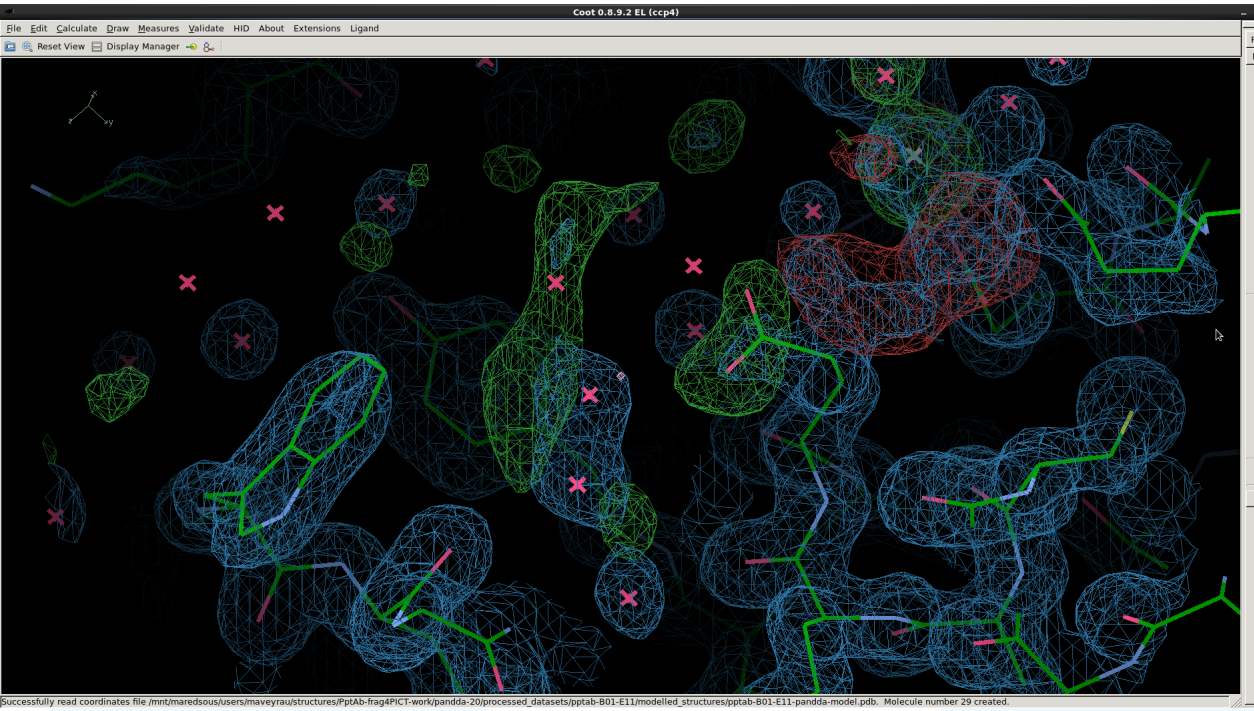


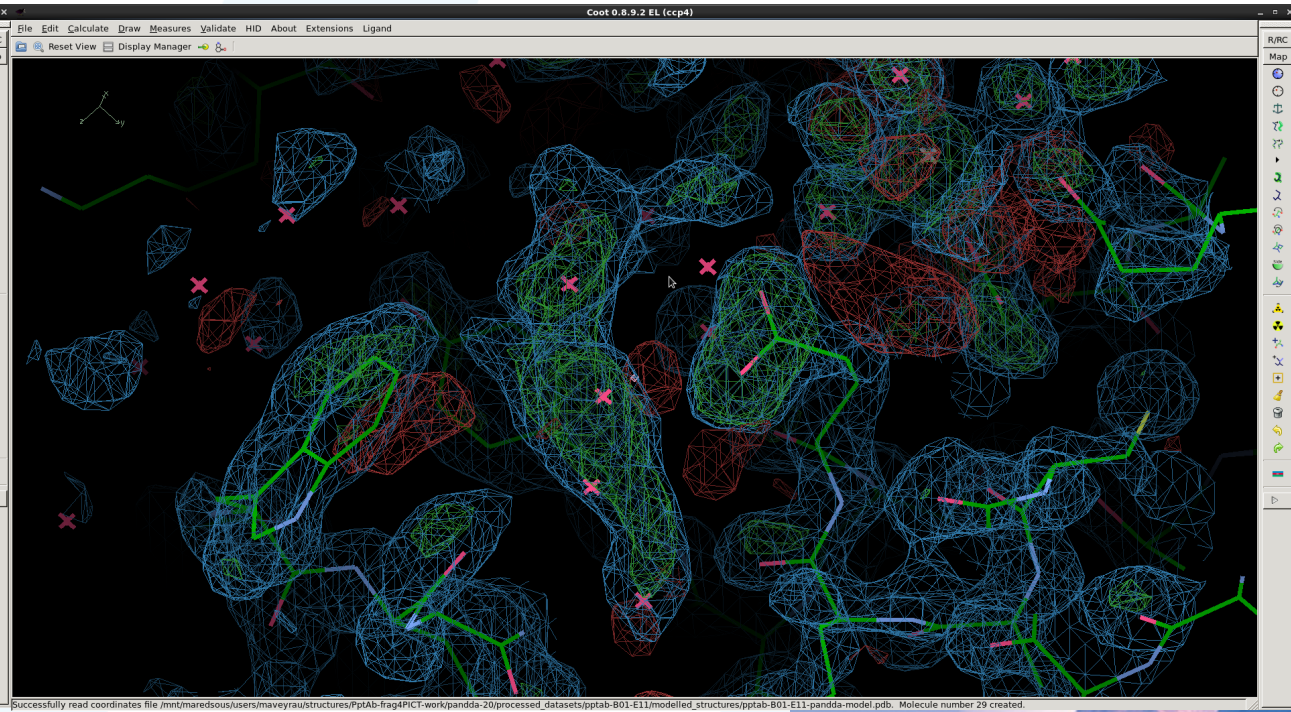
The dimple **Fo-Fc map**: presence of the ligand is possible, but impossible to build

The Pandda **event map**: presence of the ligand is obvious, easily built
=> Binding with low occupancy

Unpublished results

Pandda in real life: and the ugly.





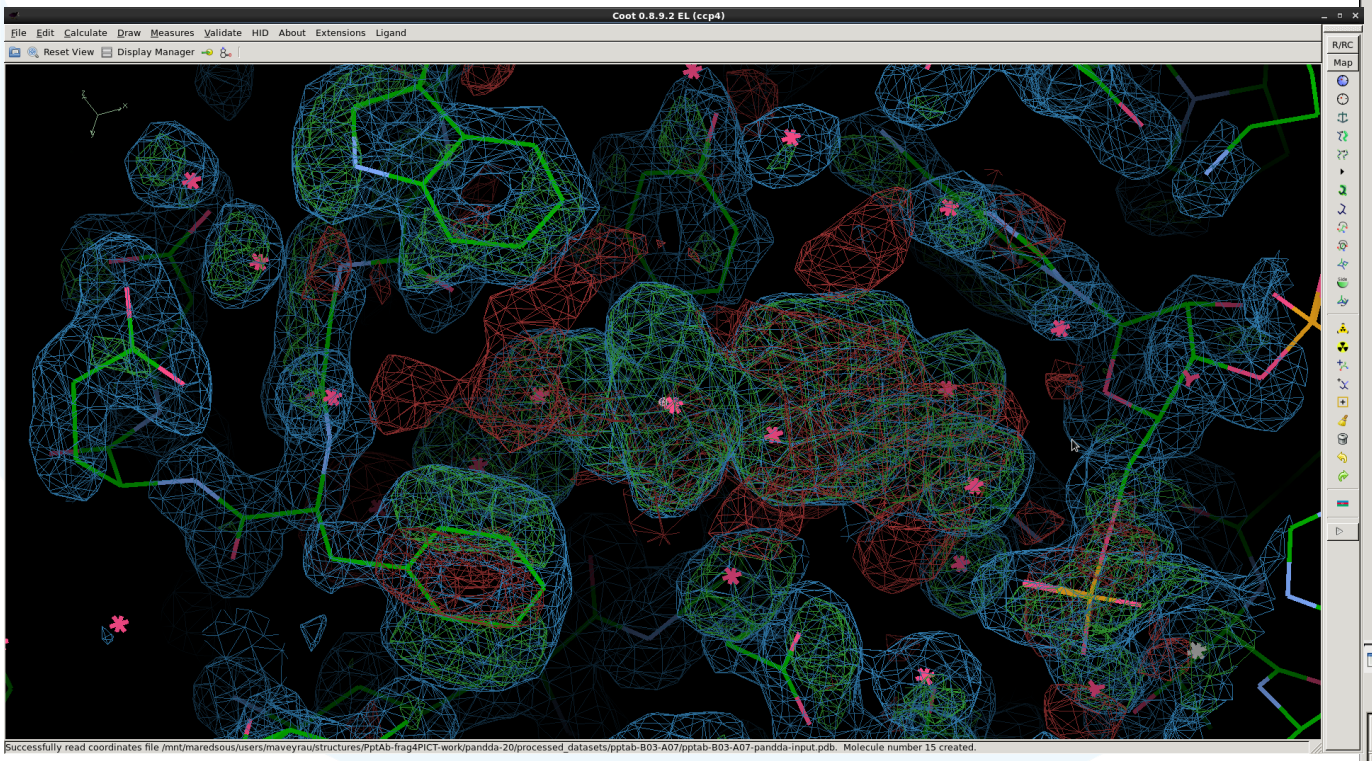
The dimple **Fo-Fc map**: seeing a ligand there would be highly optimistic!

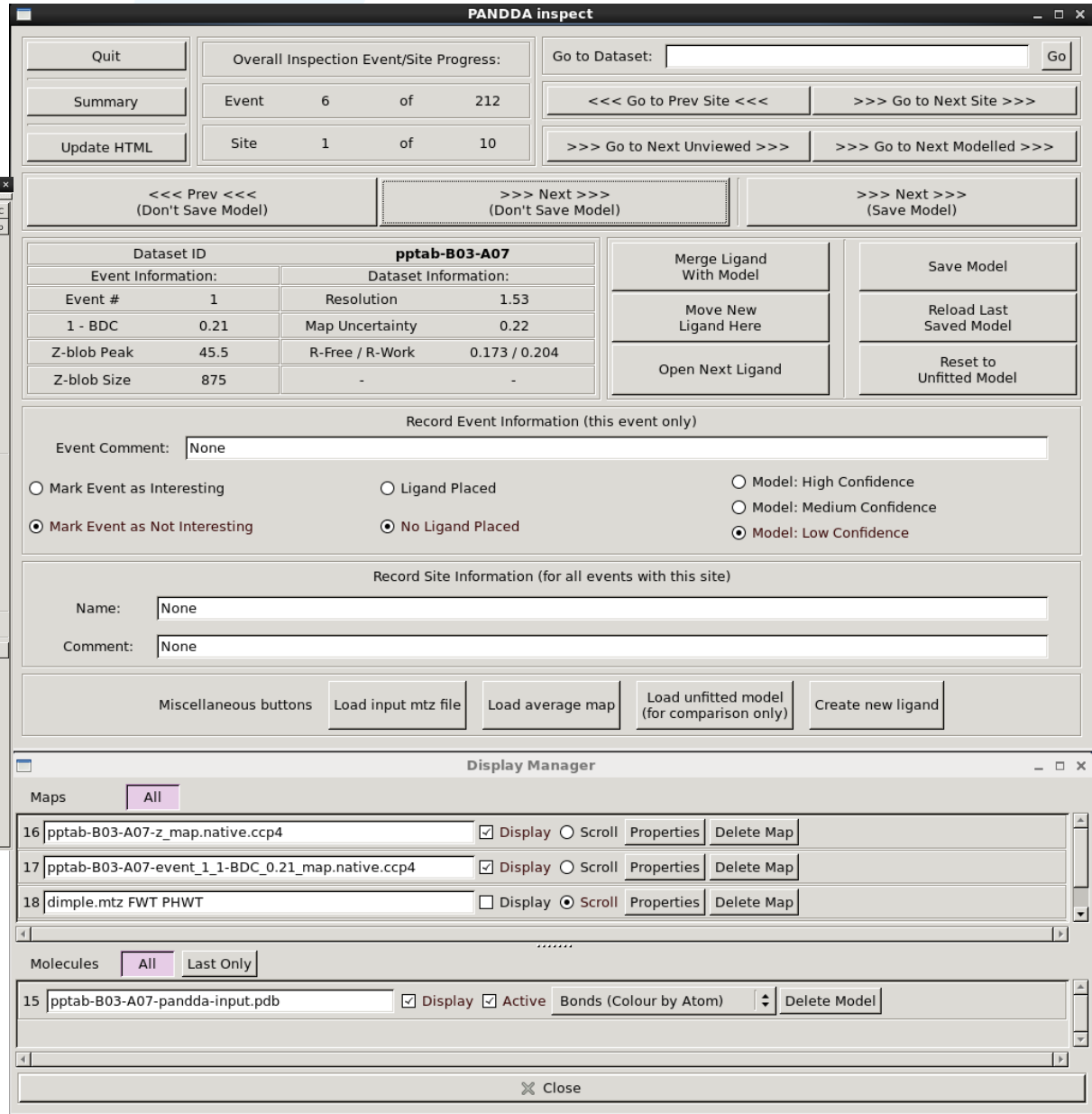
The Pandda **event map**: something is there… But impossible to build the ligand with confidence!

Unpublished results

Pandda in real life

Th Pandda / Coot / XCE interface



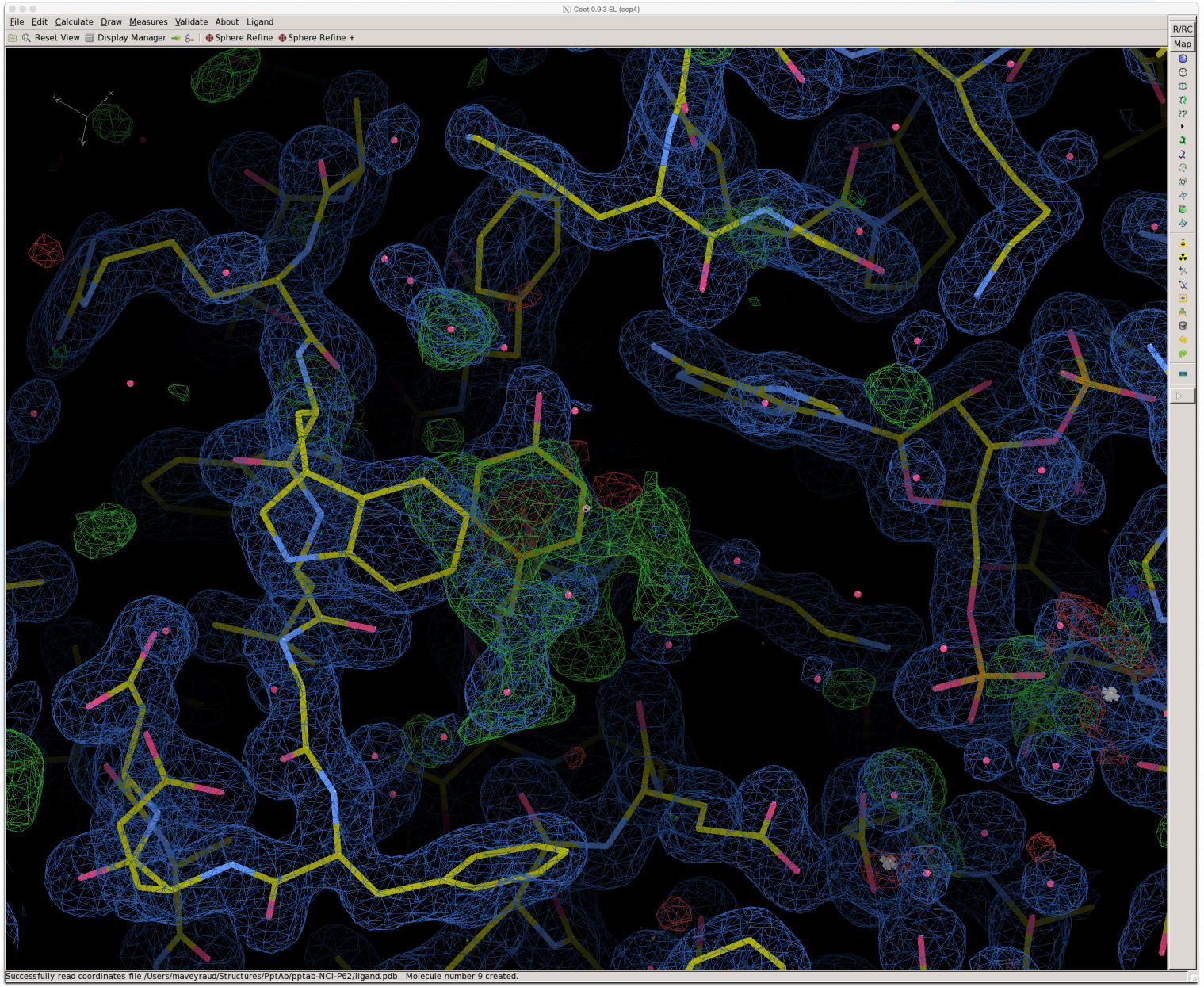


Pipedream

Peform refinement as thoroughly as possible before checking for presence of a ligand… :

- In the penultimate refinement cycle, place « dummy waters » in unmodelled density
- Groups of connected dummy water suggest possible ligand binding site
- Perform last refinement cycle without these waters: should improve difference maps

Pipedream/Pandda



Dimple:
Resolution 1.3
R/Free : 0.1753 / 0.1960

Unpublished results

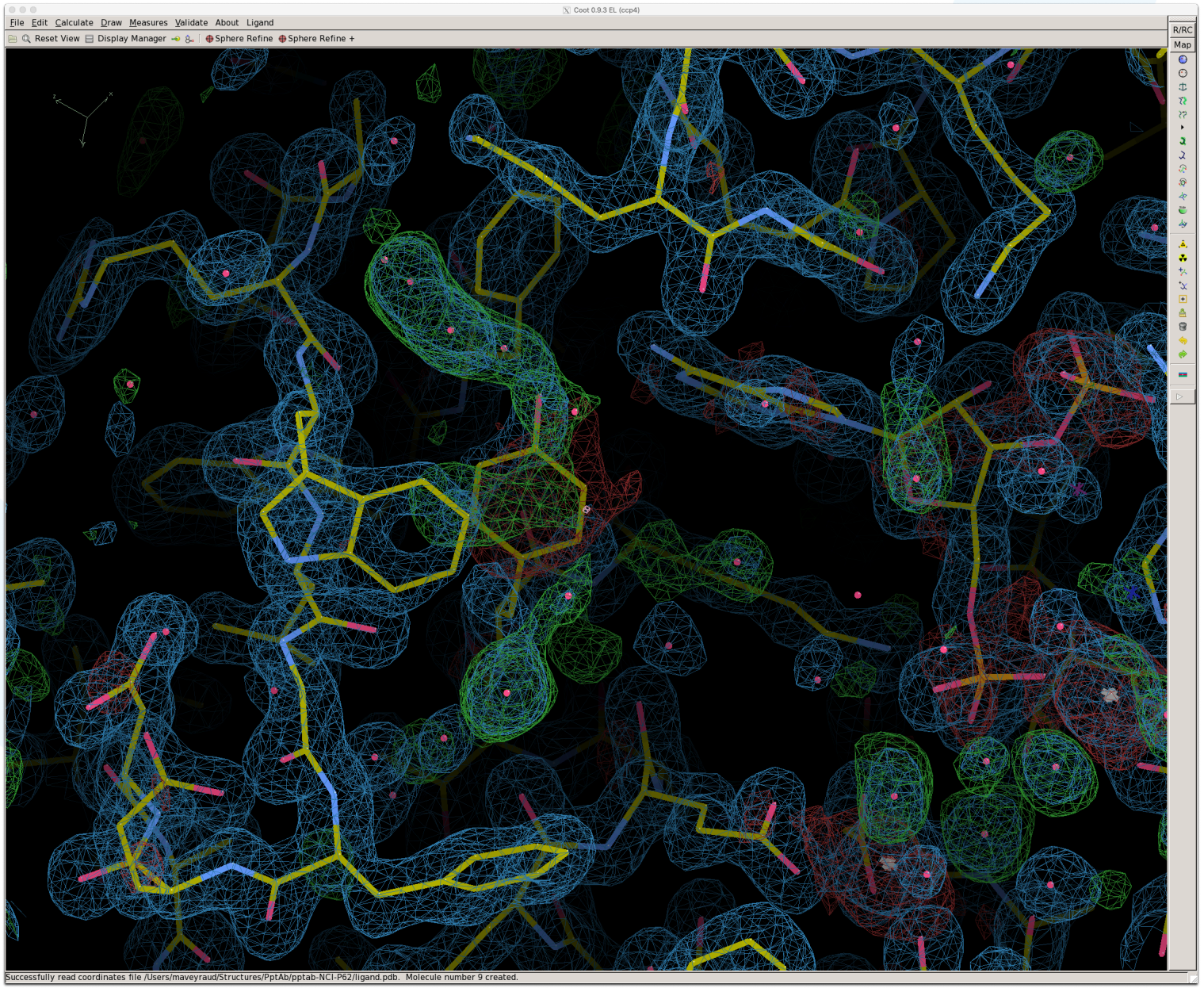
ipbs TOULOUSE CANCER INFECTION INFLAMMATION

GENOTOUL Pict

cnrs

UNIVERSITÉ TOULOUSE III PAUL SABATIER

Pipedream/Pandda



Pipedream:
Resolution 1.3
R/Free : 0.2099 / 0.2313

Unpublished results

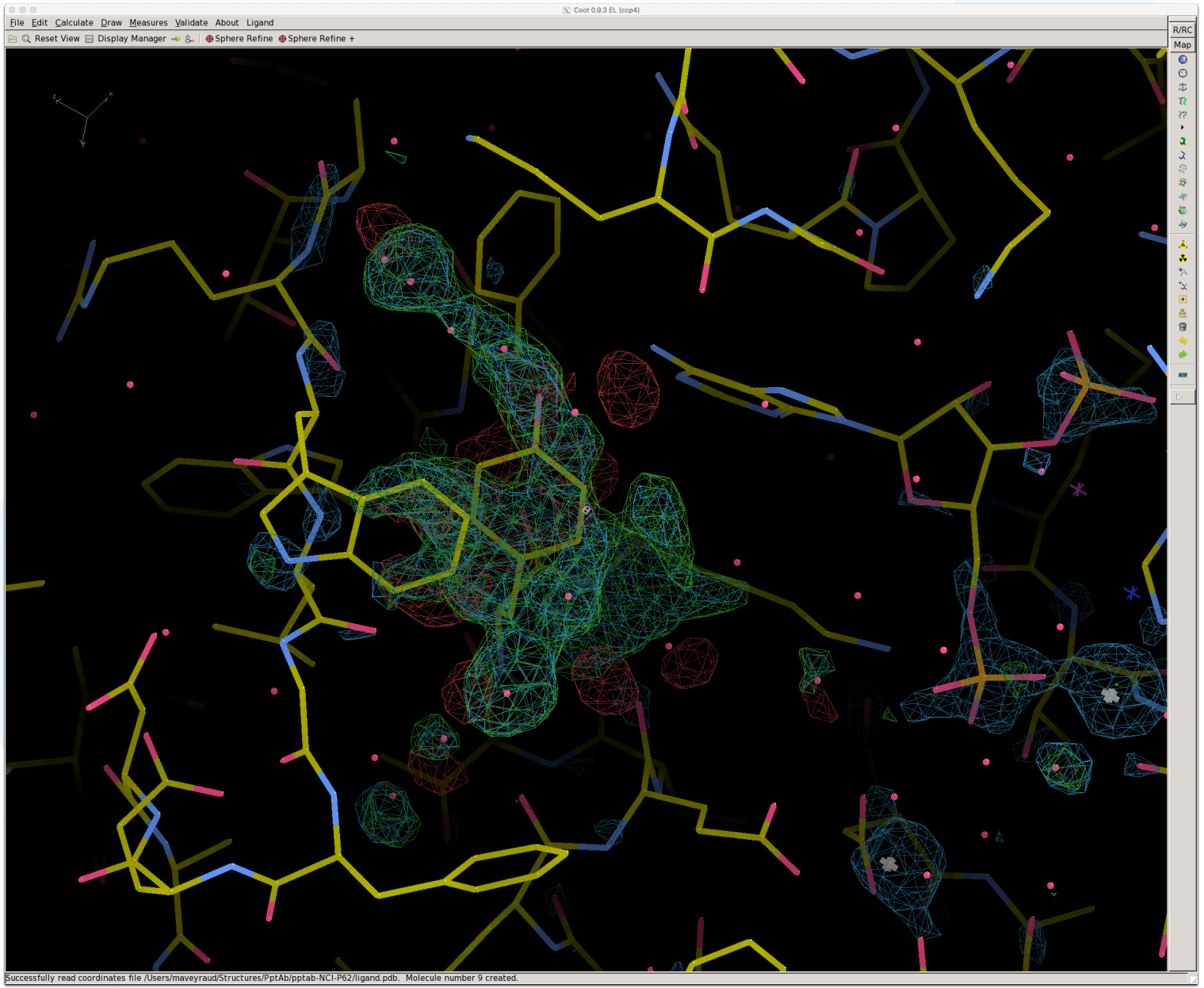
ipbs TOULOUSE CANCER INFECTION INFLAMMATION

GENOTOUL Pict

cnrs

UNIVERSITÉ TOULOUSE III PAUL SABATIER

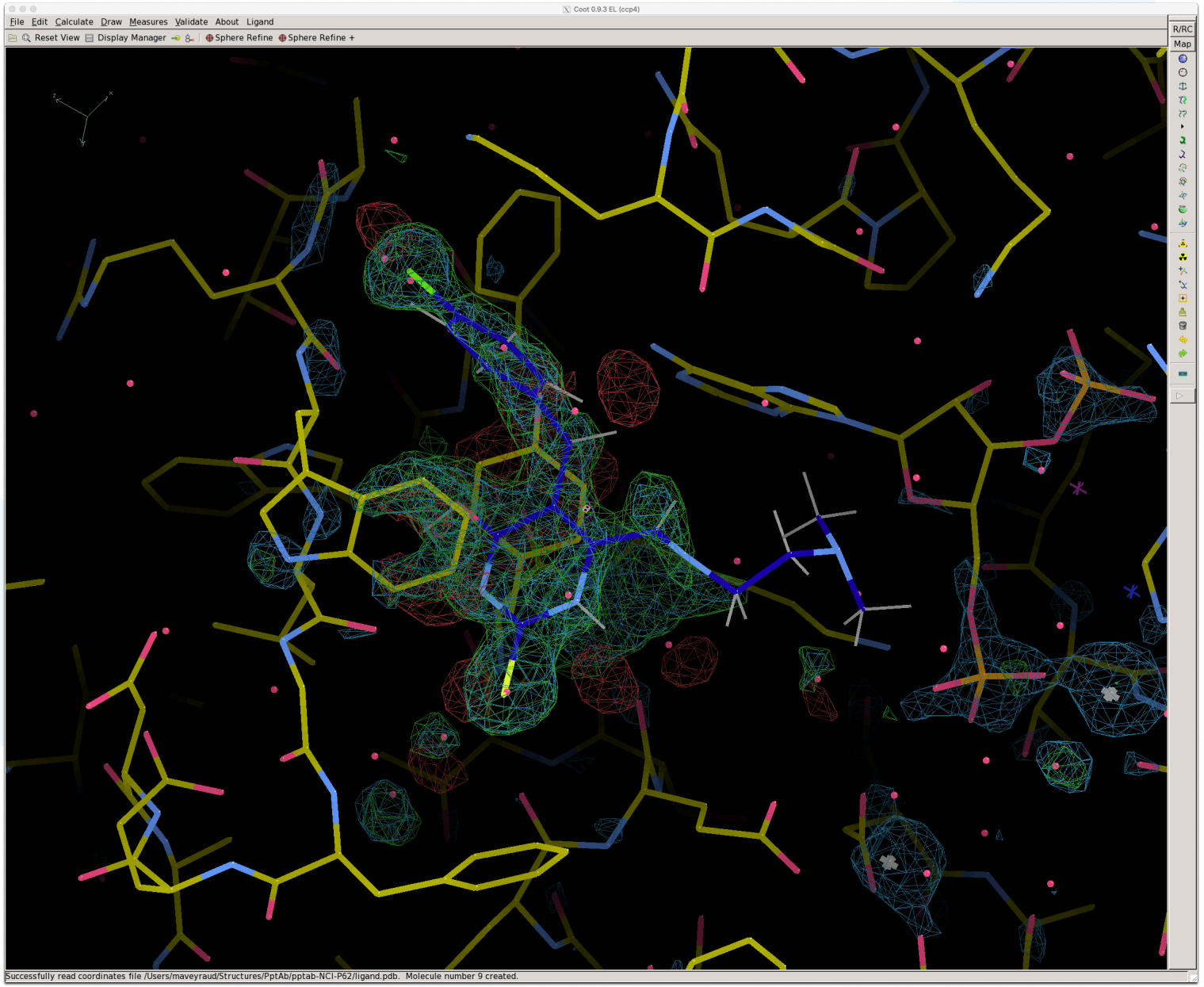
Pipedream/Pandda



Dimple + Pandda:
Resolution 1.3
R/Free : 0.2099/0.2313

Unpublished results

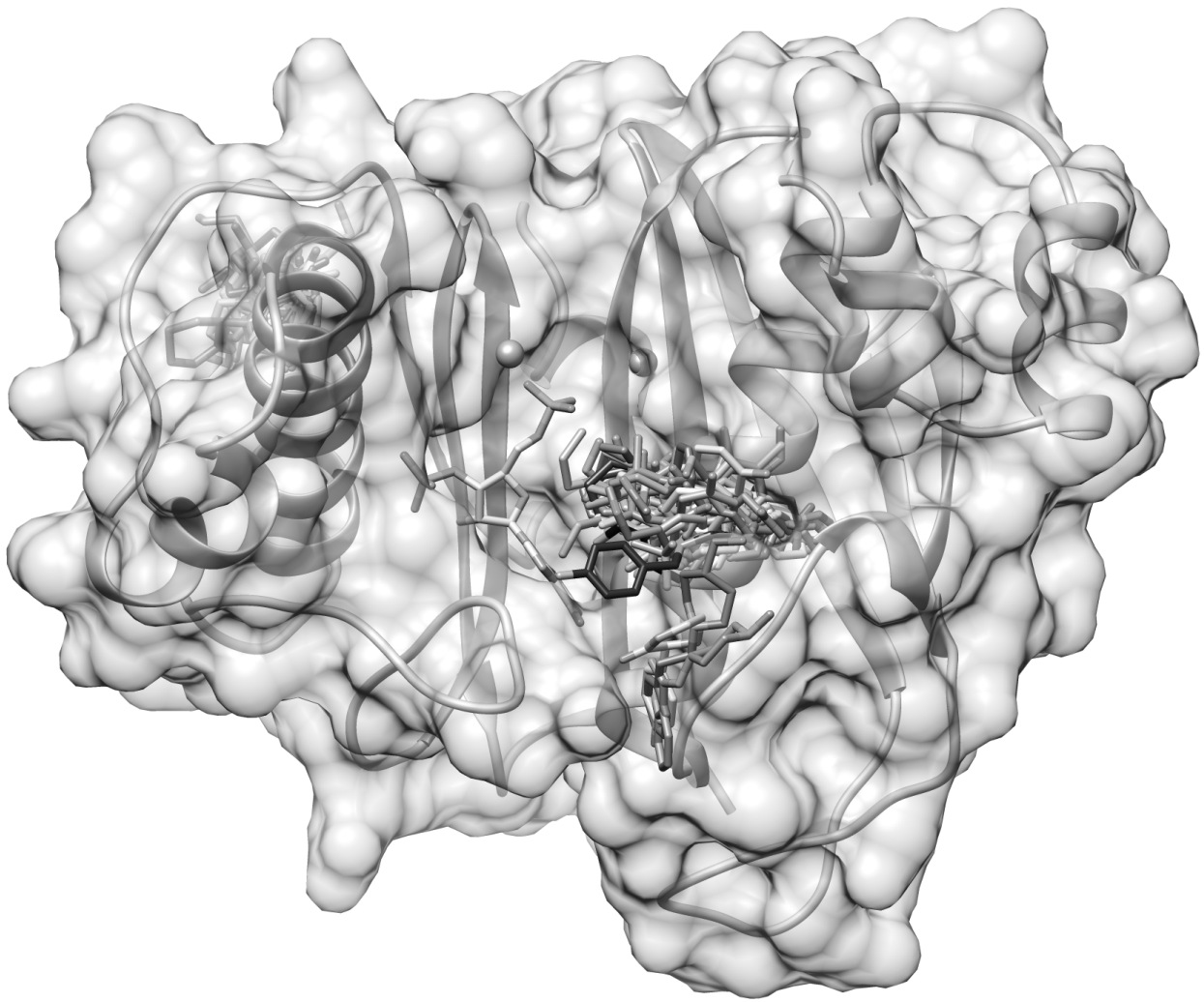
Pipedream/Pandda

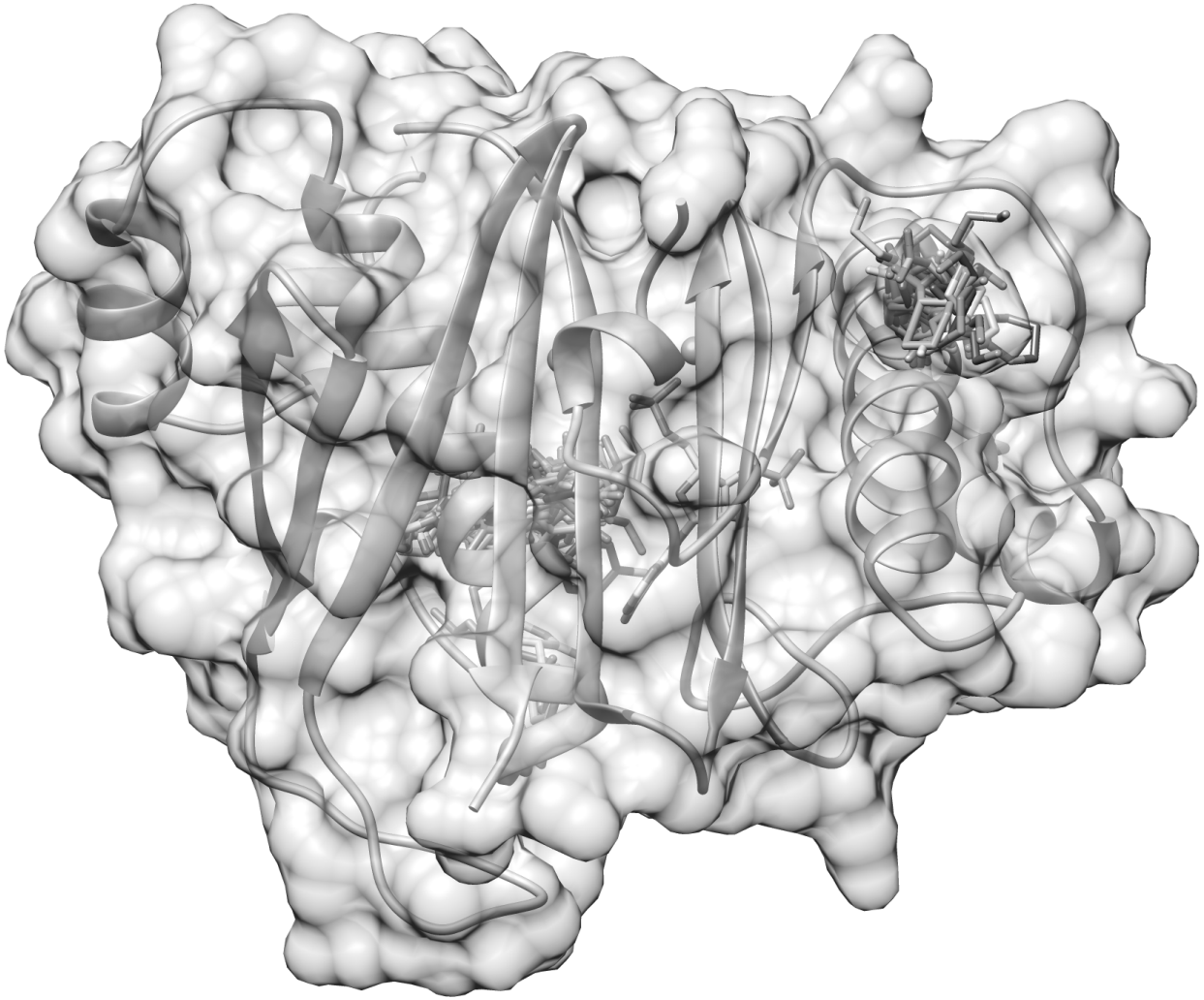


Refined ligand:
Resolution 1.3
R/Free : 0.1685 / 0.1948
Refined occupancy: 0.58

Unpublished results

939 fragments screened, 71 hits



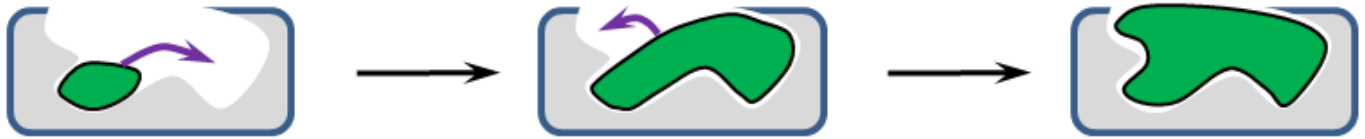


Unpublished results

Optimizing fragments

Fragments linking

Fragments growing



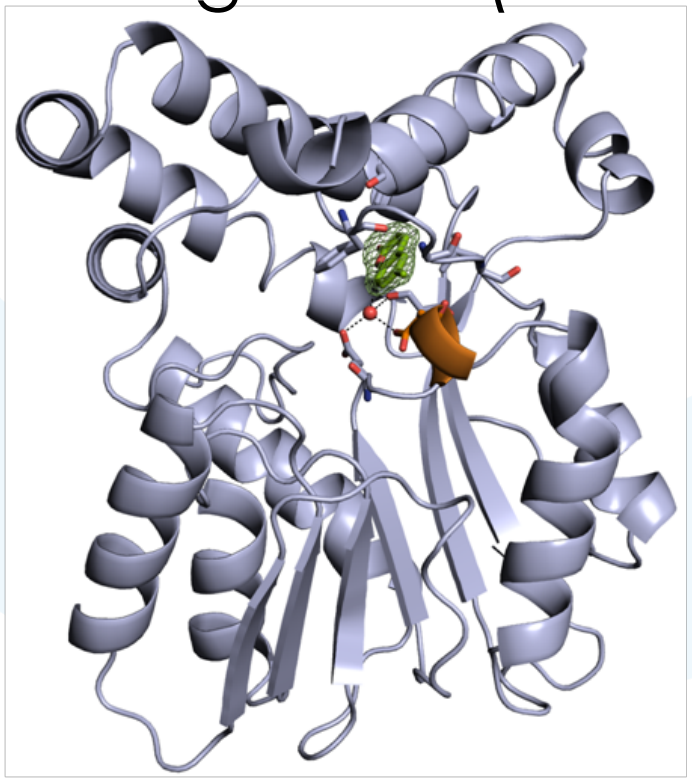
Fragments merging

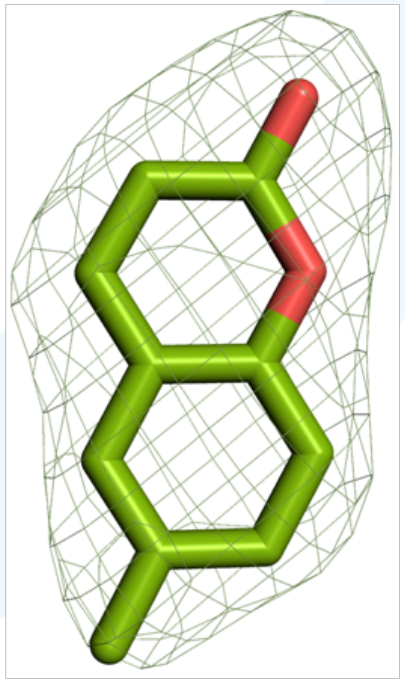
Galy, 2012

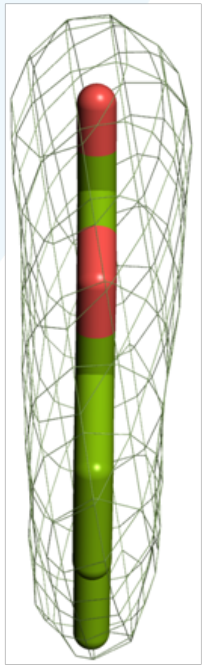
Optimizing fragments

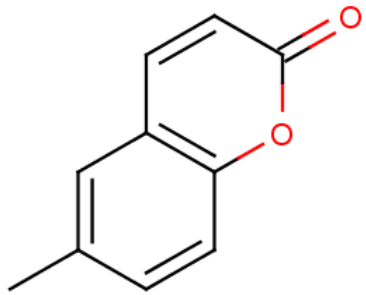
The project:

- Target: methyltransferase from *Mycobacterium tuberculosis*
- 160 fragments (commercial library)









A

-20.9 kJ.mol^{-1}

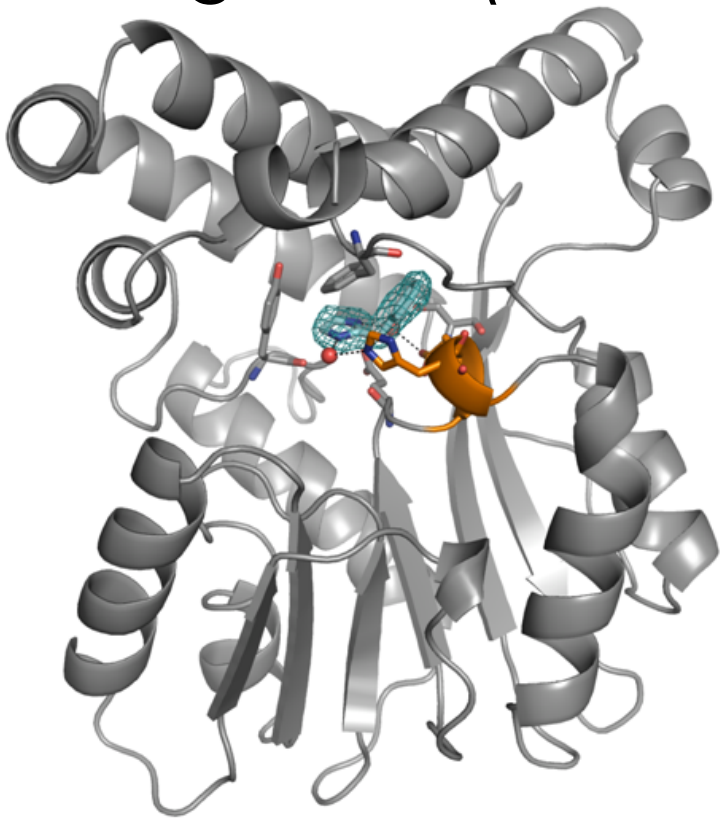
LE = 1.74

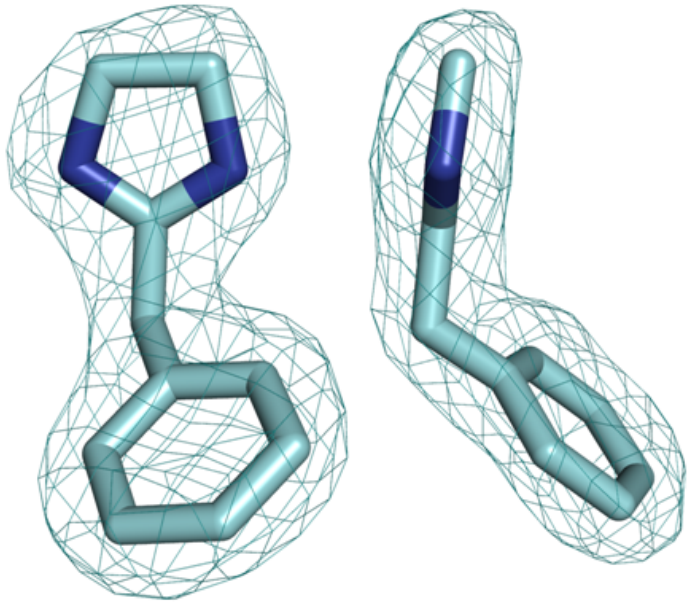
Galy, unpublished

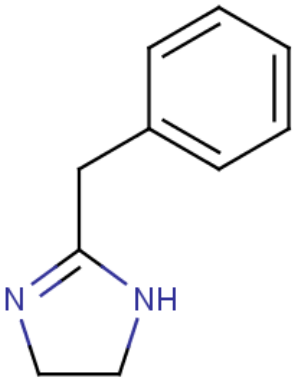
Optimizing fragments

The project:

- Target: methyltransferase from *Mycobacterium tuberculosis*
- 160 fragments (commercial library)







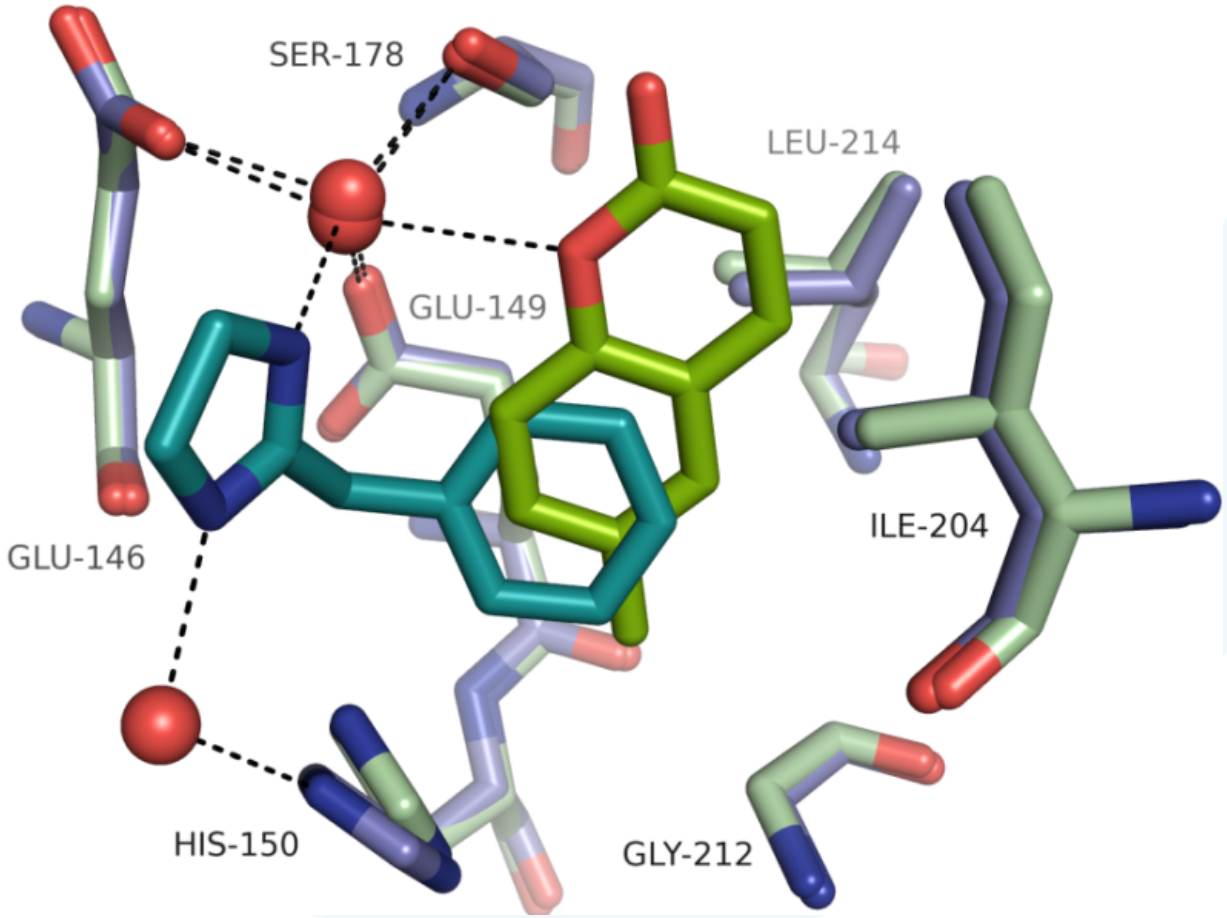
B

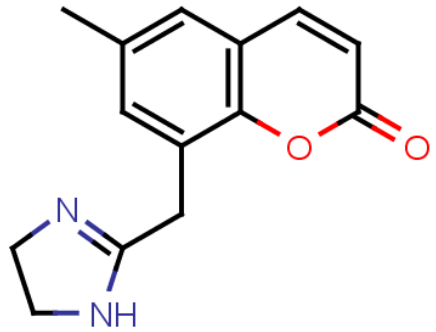
-20.7 kJ.mol^{-1}

LE = 1.73

Galy, unpublished

Optimizing fragments

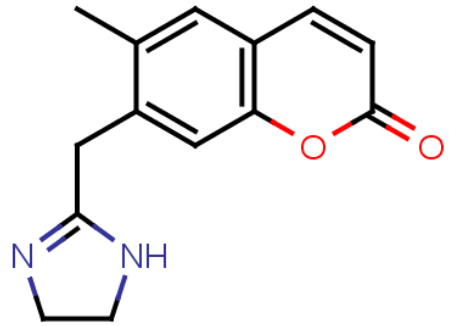




C

-26.3 kJ.mol^{-1}

LE = 1.5



D

-34.5 kJ.mol^{-1}

LE = 1.92

Galy, unpublished

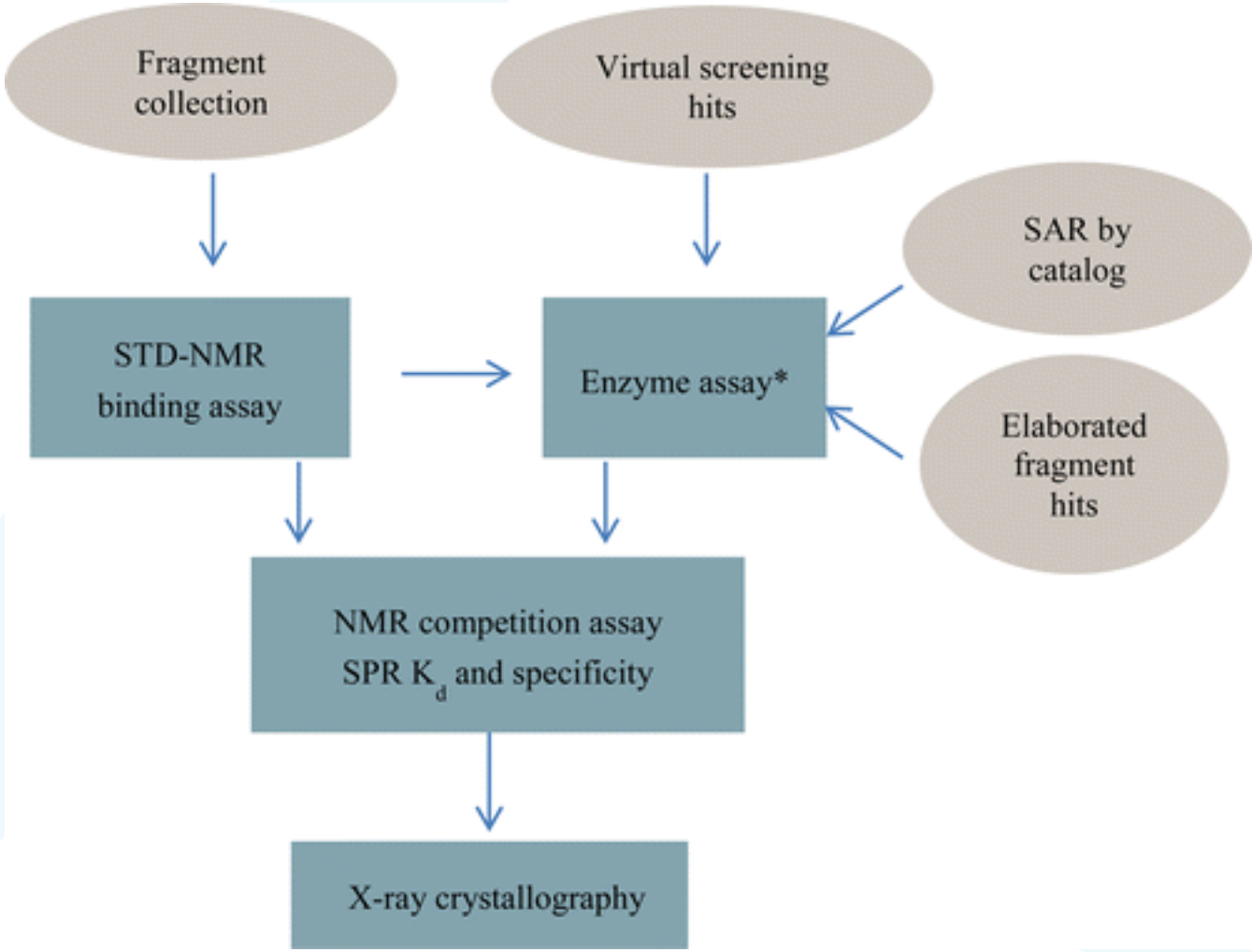
Fragments in the clinic

Approved!		
Erdafitinib (2019)	**Astex/J&J**	**FGFR1-4**
Pexidartinib	**Plexxikon**	**CSF1R, KIT**
Vemurafenib (2011)	**Plexxikon**	**B-RAF^{V600E}**
Venetoclax (2016)	**AbbVie/Genentech**	**Selective BCL-2**
Phase 3		
Asciminib	**Novartis**	**BCR-ABL**
Lanabecestat	Astex/AstraZeneca/Lilly	BACE1
Verubecestat	Merck	BACE1
Phase 2 : 19 compounds in the pipeline		
Phase 1 : 21 compounds in the pipeline		

Nanomolar inhibitors of LDH

- Lactate dehydrogenase: last enzyme in the glycolytic cycle, converts lactate into pyruvate
- In cancer cells, ATP is preferentially derived from glycolysis rather than oxydative phosphorylation

LDH is a target in anticancer therapies



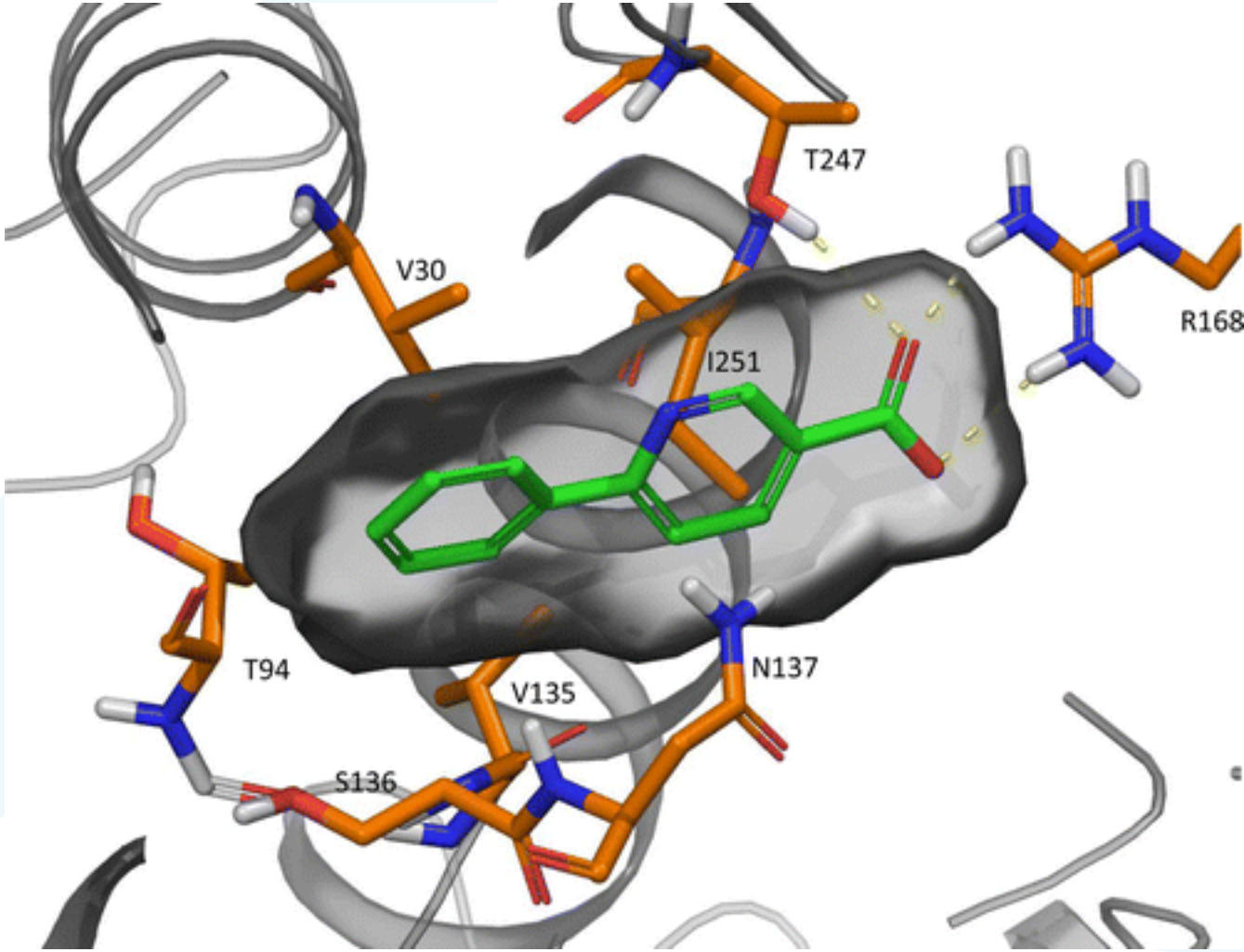
Kohlmann, 2013

Nanomolar inhibitors of LDH

Validation of hits by SPR…

Too low Kd (1-5 mM) for SAR

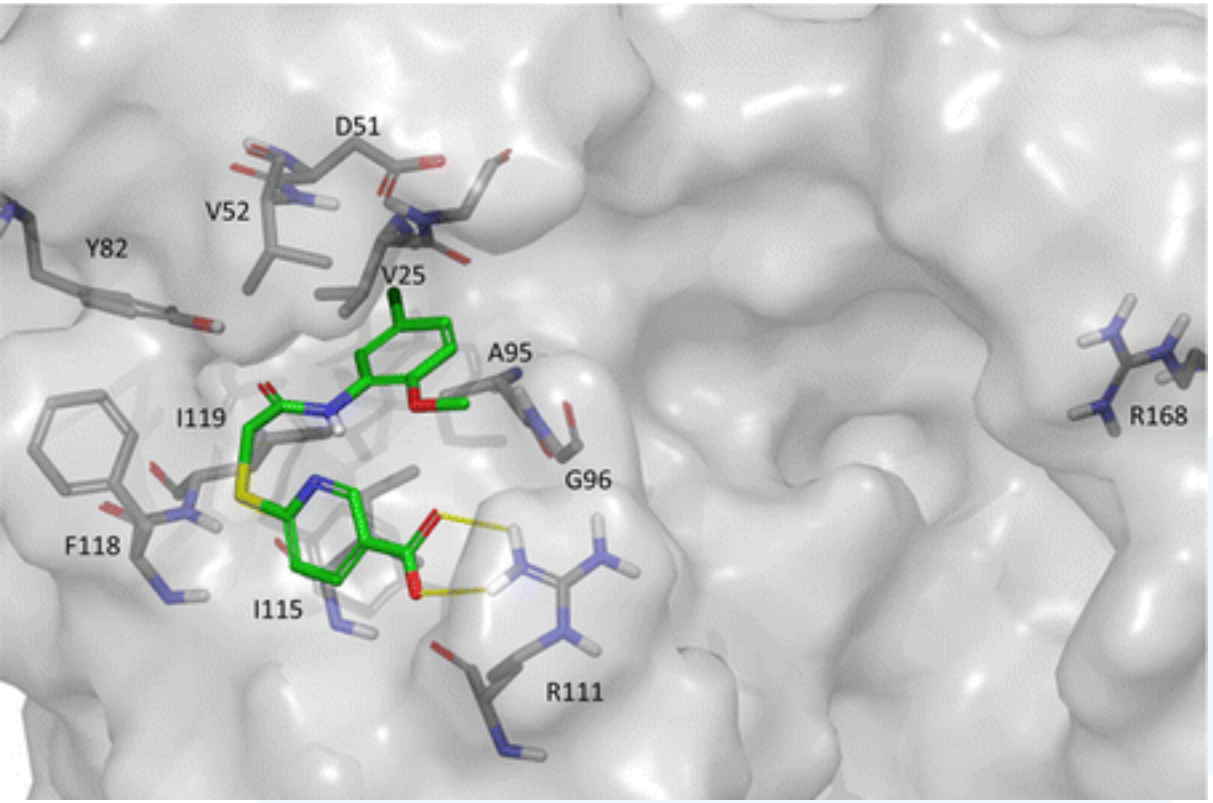
X-ray structure



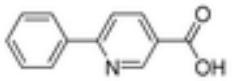
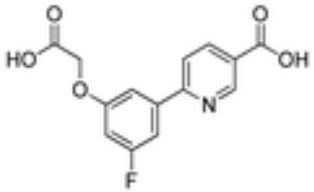
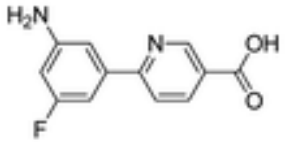
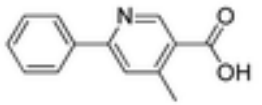
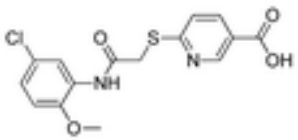
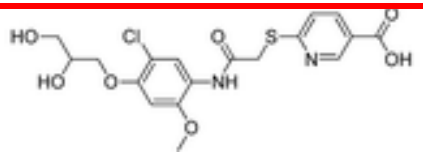
Kohlmann, 2013

Nanomolar inhibitors of LDH

SAR by catalog



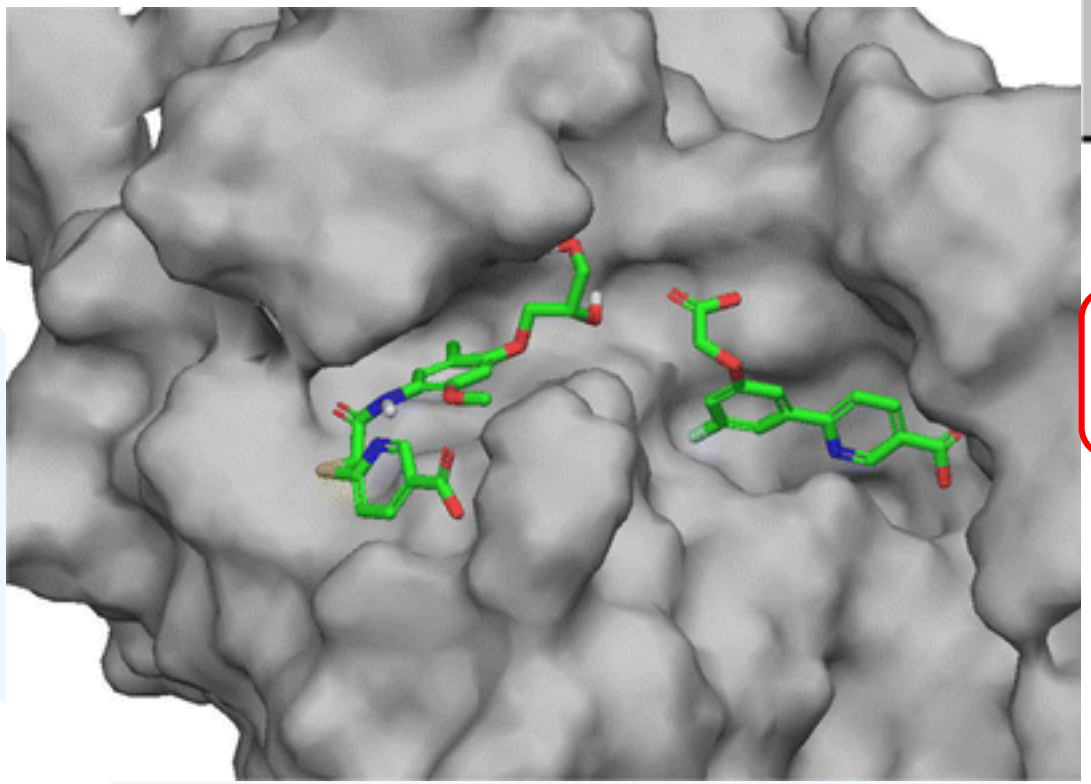
Does not bind as expected!

Compound	Structure	IC_{50},µM	K_d,µM	LE
1		> 2000	2300	0.18
2		2200	1300	0.14
3		1300	na	
4		na	1900	0.17
5		770	137	0.17
6		342	360	0.12

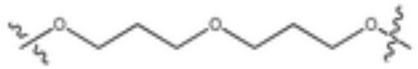
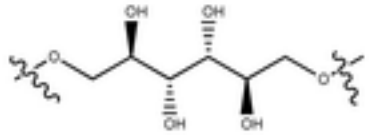
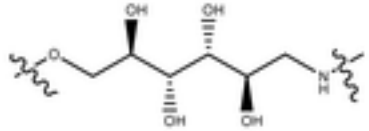
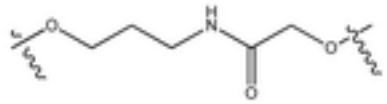
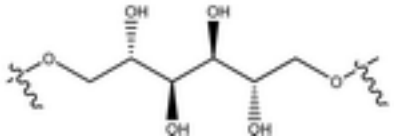
Kohlmann, 2013

Nanomolar inhibitors of LDH

Two fragments at distinct site: let's link'em



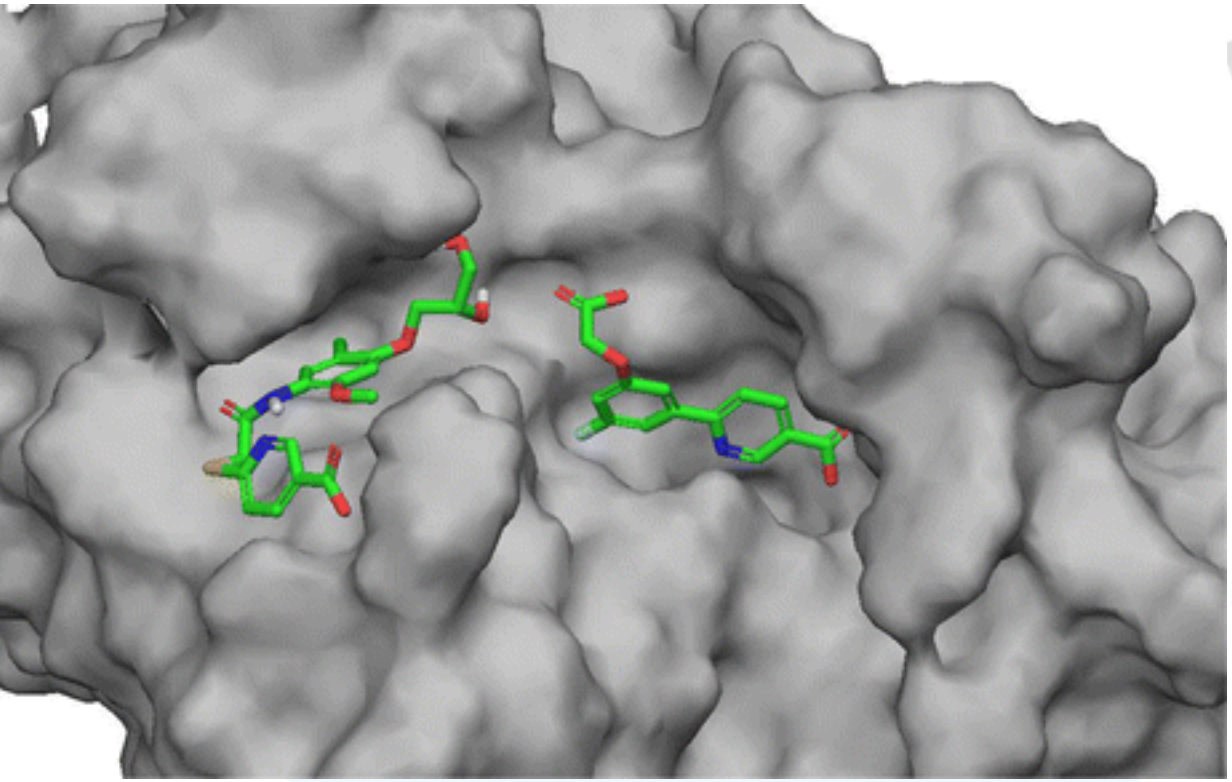
Co-soaked structure

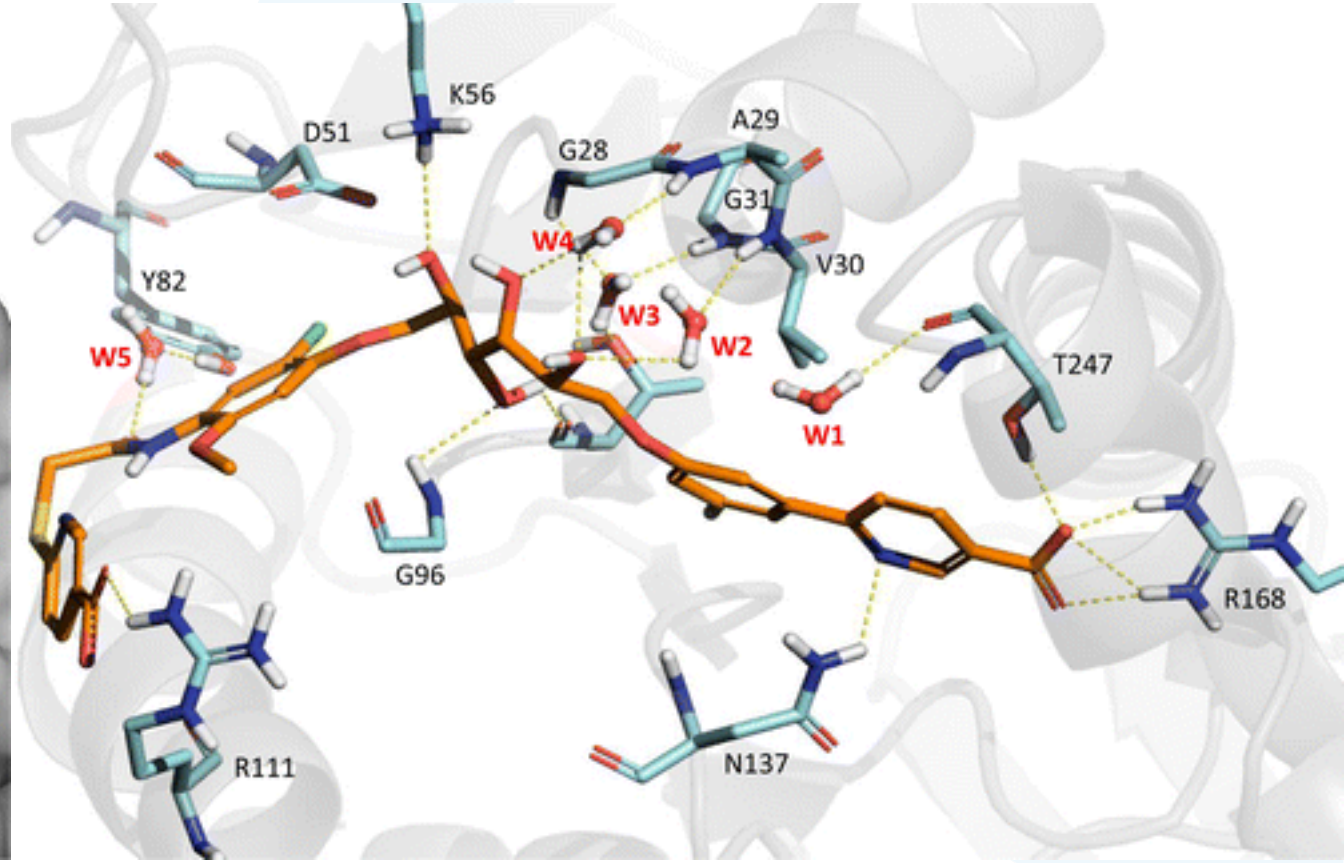
Compound	Linker	IC_{50},µM	K_d,µM	LE
7		59.0	0.850	0.13
8		0.440	0.068	0.14
9		0.120	0.019	0.15
10		2.40	0.175	0.14
11		15.2	12.0	0.10

SAR of linkers

Kohlmann, 2013

Nanomolar inhibitors of LDH





Linked fragment binds as expected

Kohlmann, 2013

Thank you!

PICT :

Plateforme Intégrée de Criblage de Toulouse

Lionel Mourey (IPBS)
Olivier Saurel (IPBS)
Yves Génisson (SPCMIB)
Magali-Rémaud Simeon (TBI)

Hma

Romain Galy (IPBS)
Sylviane Julien (IPBS)
Stéphanie Ballereau (SPCMIB)
Yves Génisson (SPCMIB)
Jean Christophe Plaquevent (SPCMIB)
Olivier Saurel (IPBS)
Alain Milon (IPBS)
Georges Czaplicki (IPBS)
Lionel Mourey (IPBS)

PPTase

Coralie Carivenc (IPBS)
Minh Chau Nguyen (IPBS)
Jean-Denis Pedelacq (IPBS)
Christian Chalut (IPBS)
Christophe Guilhot (IPBS)
Claire Blanger (SPCMIB)
Yves Génisson (SPCMIB)
Stéphanie Ballereau (SPCMIB)
Lionel Mourey (IPBS)

PROJET FINANCÉ PAR L'ANR - PROJECT FUNDED BY THE ANR
ANR