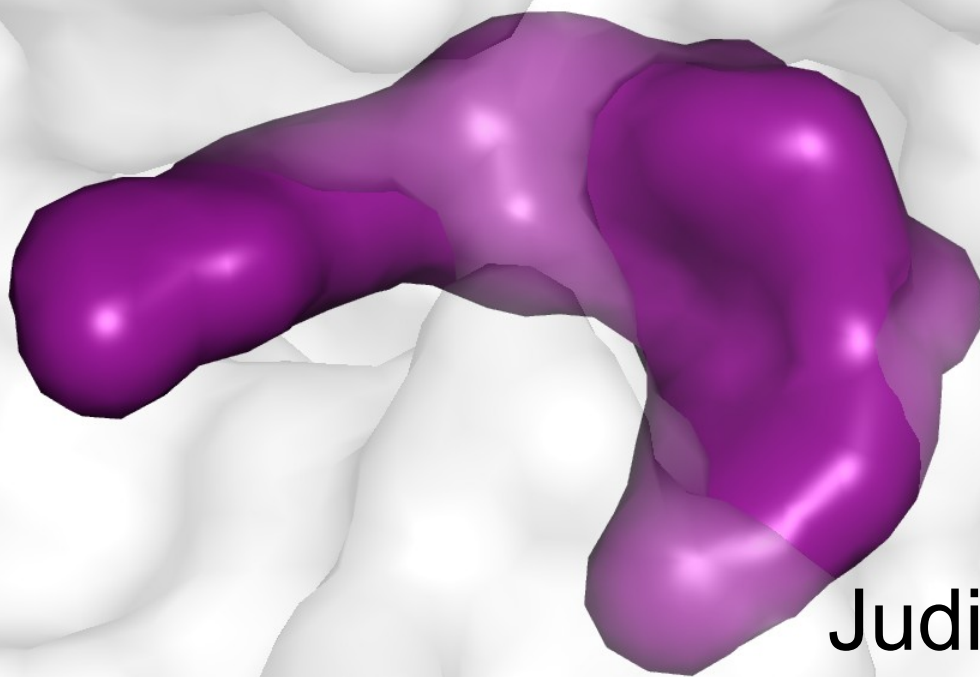


Crystallography in industry



Judit Debreczeni

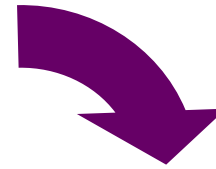
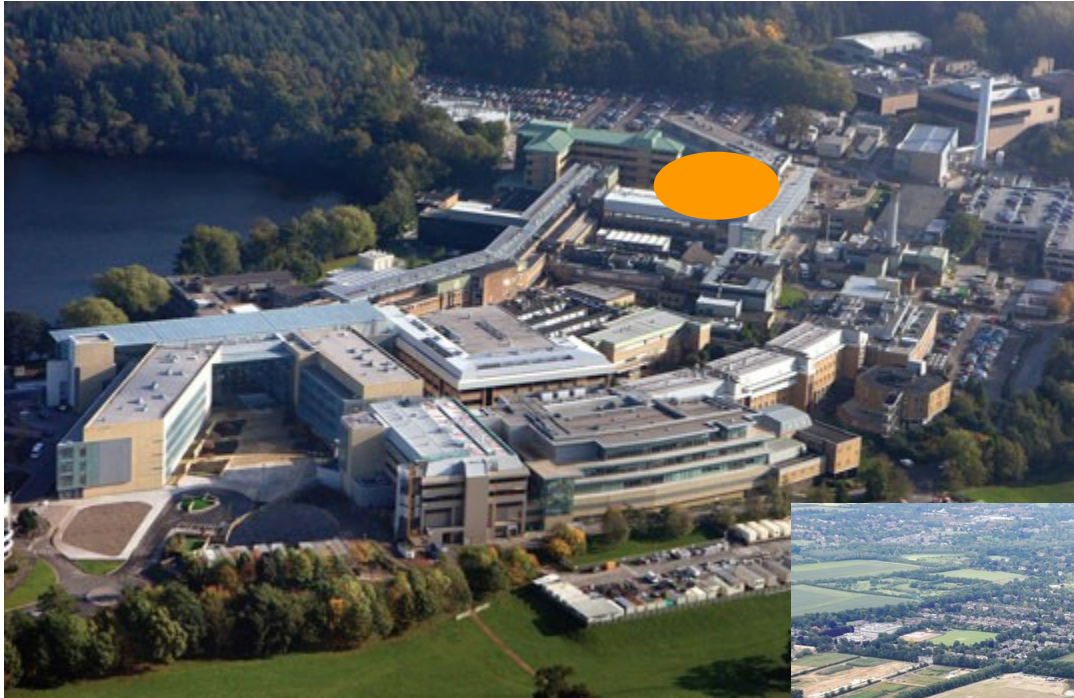
AstraZeneca

Diamond/CCP4 workshop 2014

Acknowledgements



- Chris Phillips
- Claire Brassington
- Jason Breed
- David Hargreaves
- Tina Howard
- Richard Norman
- Derek Ogg
- Jon Read
- Steve StGallay
- Martin Packer
- Willem Nissink
- Richard Ward
- Sebastien Degorce
- Cliff Jones
- Jason Kettle
- Richard Pauptit
- Julie Tucker
- Joe Patel
- Stefan Gerhardt
- Helen Gingell
- Tony Mete
- Choham Camaldeep



- Alderley Park, Cheshire



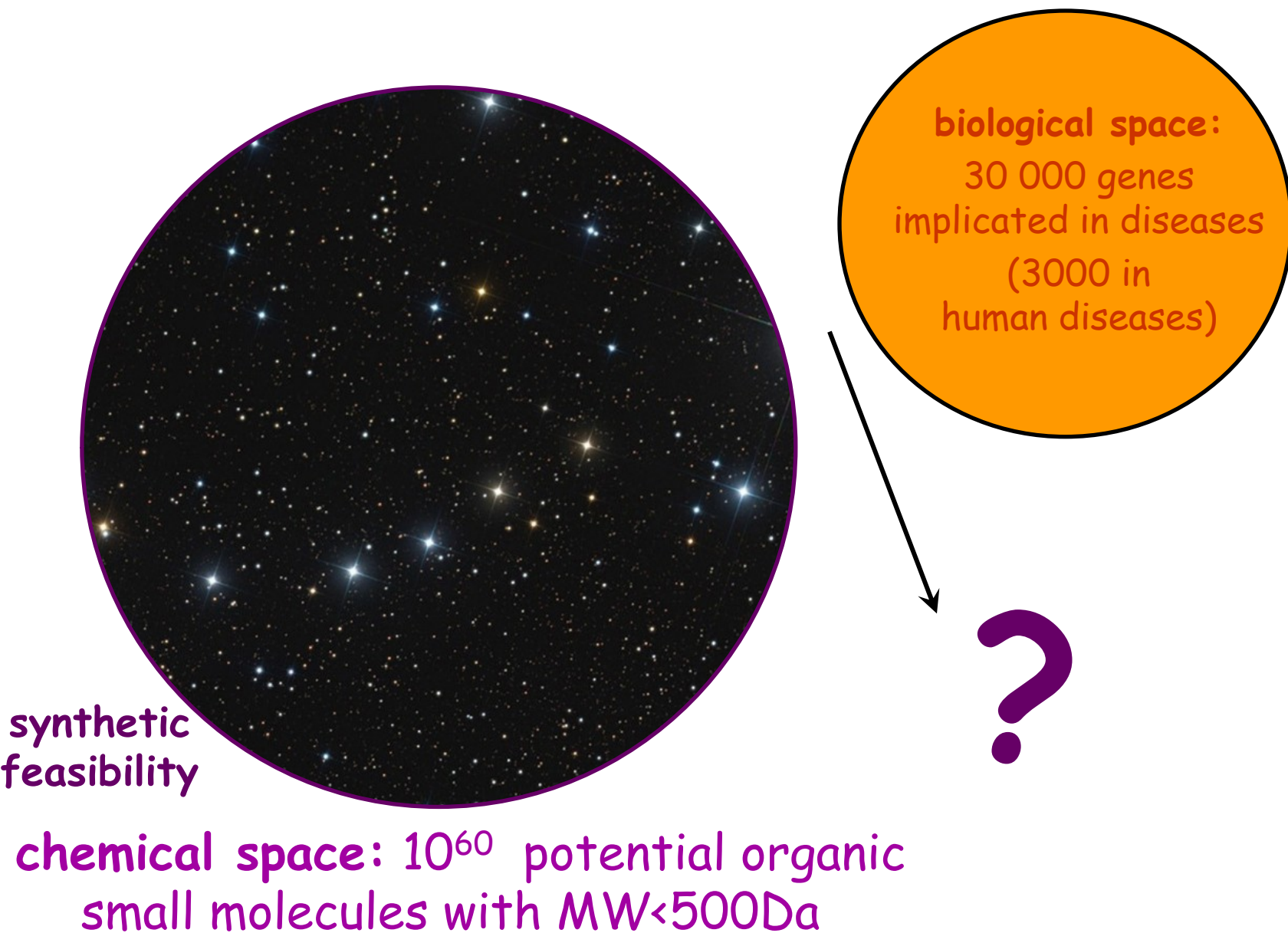
- Cambridge, CBC

Outline



- big picture - Big Pharma
- industrial crystallography - fit for purpose
- pipelines
- examples
- outlook

Chemical vs biological space



Chemical vs biological space



- undruggable biological space?
- unexplored bioactive space
- legacy compound collections
- unprecedented/challenging chemistry
- inadequate or unexplored assay techniques

Drug discovery strategies



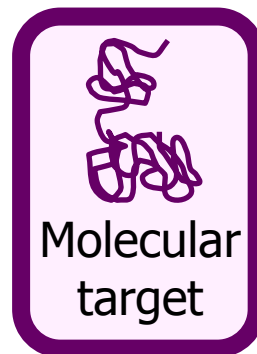
Disease



In vivo
model



In vitro
model



Molecular
target

PHENOTYPIC DRUG DISCOVERY



- Disease linkage
- Efficacy
- Toxicity



- SAR deconvolution and rational design
- Mode of action
- Screening speed

TARGETED DRUG DISCOVERY

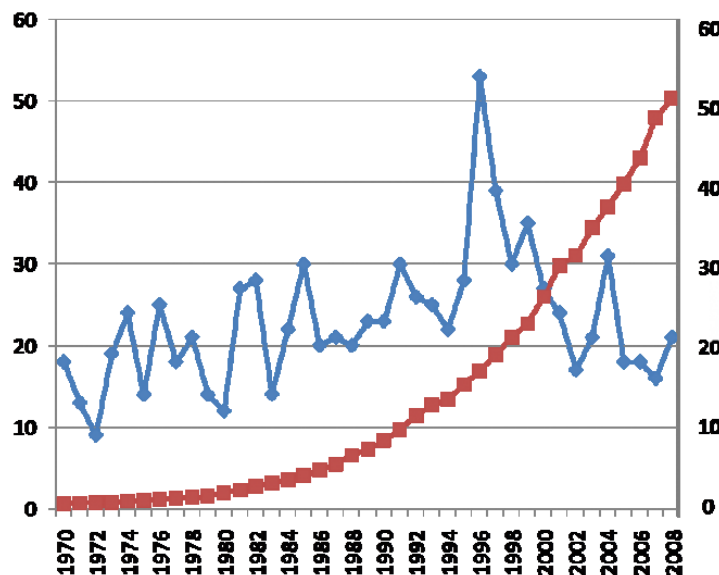


- Rational design
- SAR
- Screening speed



- Off-target effects
- Efficacy
- Target validation

R&D spend vs New Molecular Entities

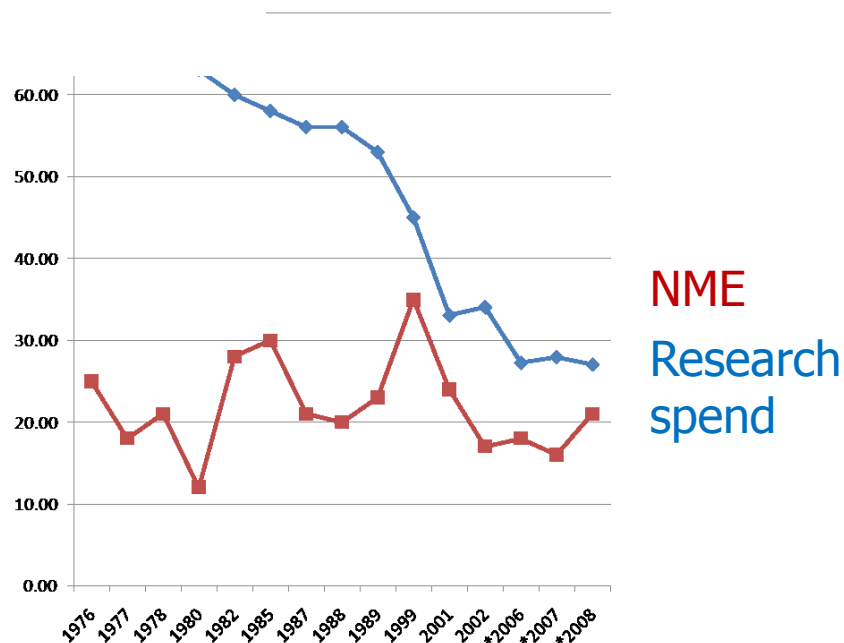


R&D cost

New Molecular Entities per year

Pressure to

- reduce timelines and cost
- early attrition



NME

Research
spend

Drug discovery/development value chain



24 months



Structural biology
impact

Lead generation strategies



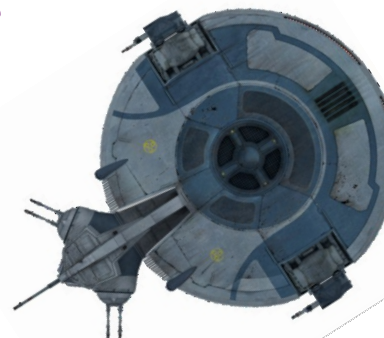
- The ultimate weapon:
high throughput screening (HTS)



- test all compounds against new targets
 $10^6 >$ compounds
- very fast (~ 1 week)
- Assay quality: compromise between speed & accuracy
- Compound conc. $\sim 10\mu\text{M}$; only measure $\text{IC}_{50} < 10\mu\text{M}$
- Compound collection of legacy chemistry
- $\sim 25\%$ of HTS screens deliver useful hits

- Powerful predator:
subset screening:

- 10-100K compounds
- even faster
- Focused on a target family or chemistry
- Biased towards prior knowledge



Lead generation strategies

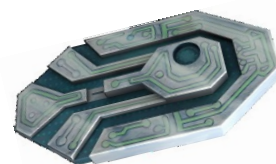


- Dim witted infantry:
fast follower strategy



- Targeted chemistry, library designs
- E.g. scaffold hopping, minor changes to tweak properties and introduce "novelty"

- An enigmatic coding device: ...
virtual screening



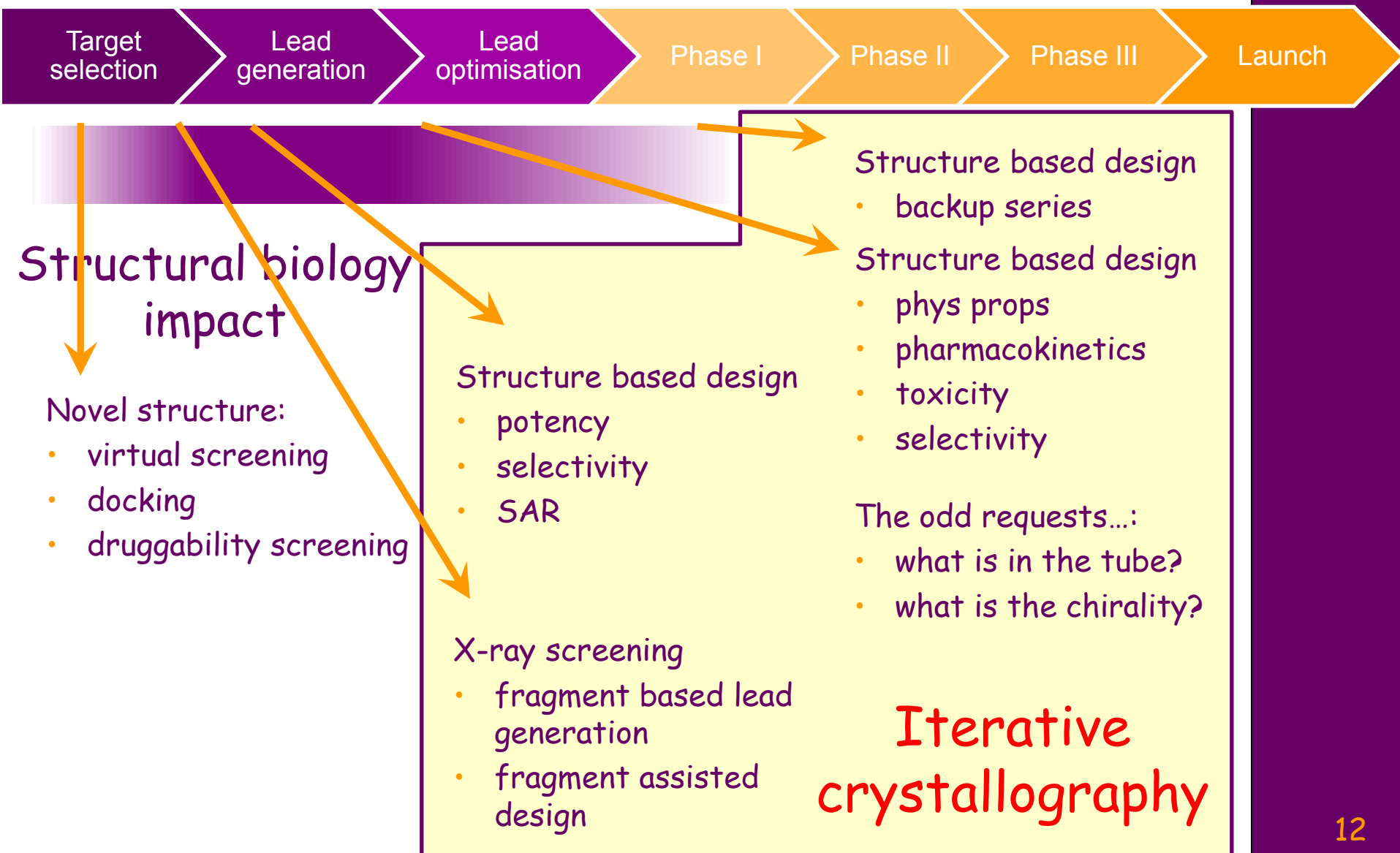
- pre-filtering on chemical properties and predictability
- topological searches (2D, 3D), pharmacophore filtering
- docking, similarity searching

- An elegant weapon ...
fragment based drug design

- Small library of small but diverse compounds screened (1-10K)
- Biophysical methods or high concentration screening to detect weak binders
- X-ray crystallography can be a primary screening technique requires a robust system
- Initial hits are weak and not drug-like!
- Limited chemistry with small conservative changes to follow up initial hits



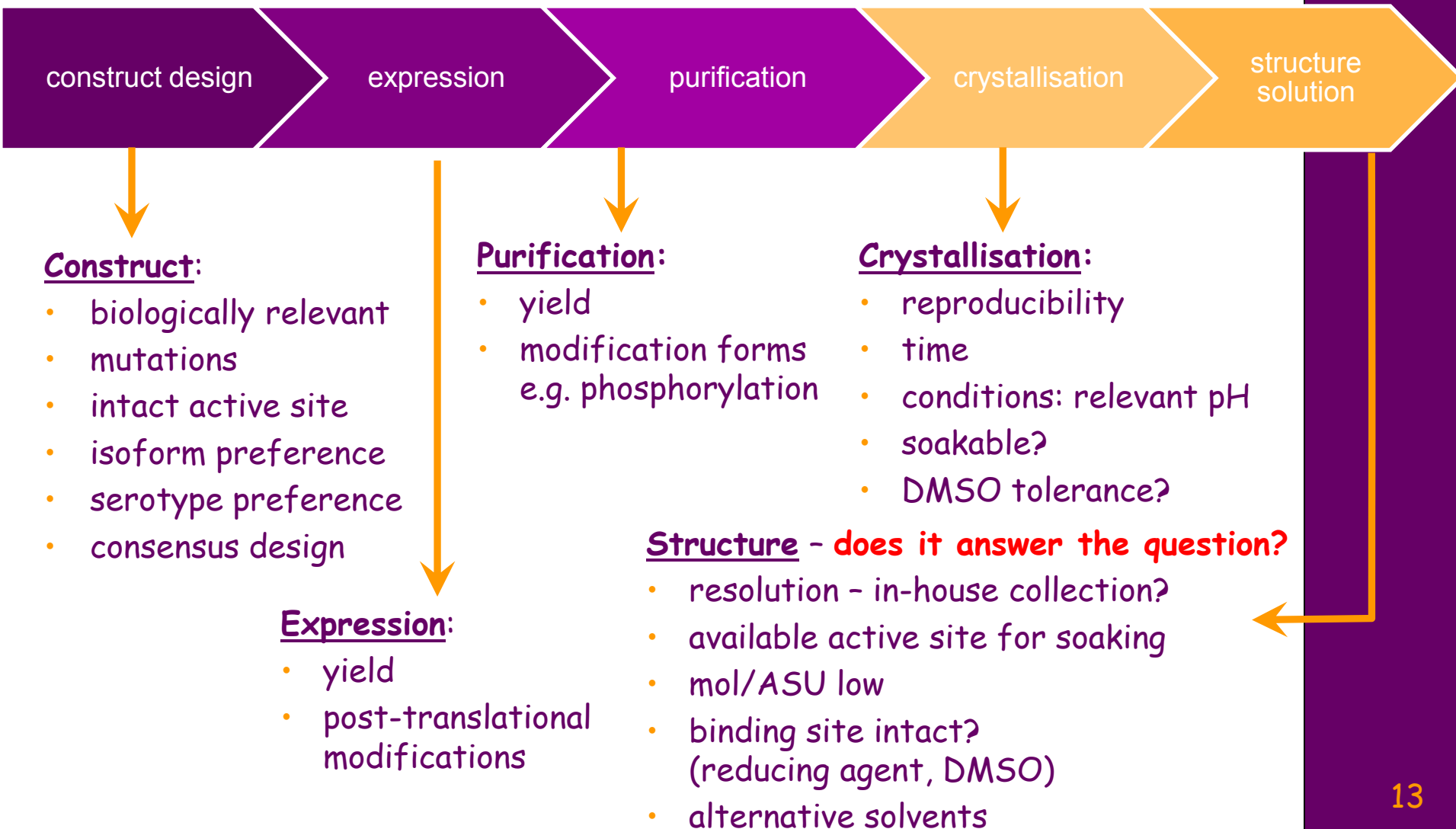
Drug discovery/development value chain



Industrial crystallography – “fit for purpose”



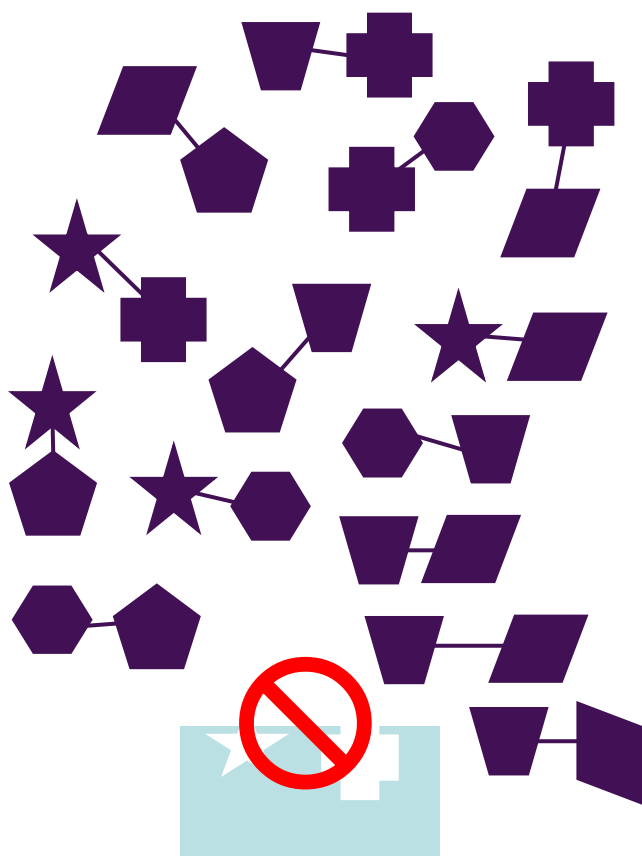
- purpose: to be able to answer specific drug discovery questions



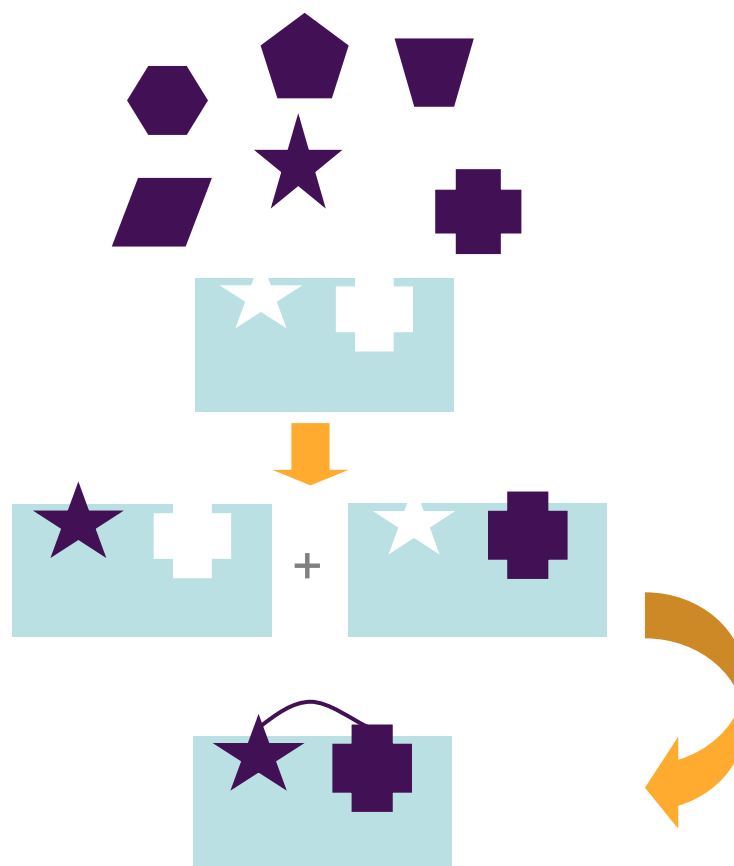
X-ray screening



Biological assay-based HTS



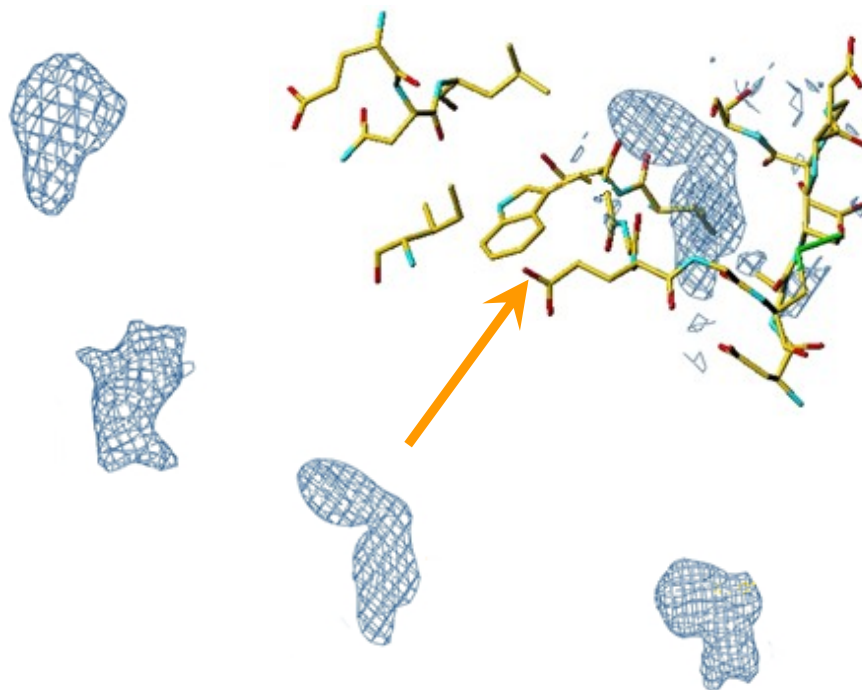
Fragment screen



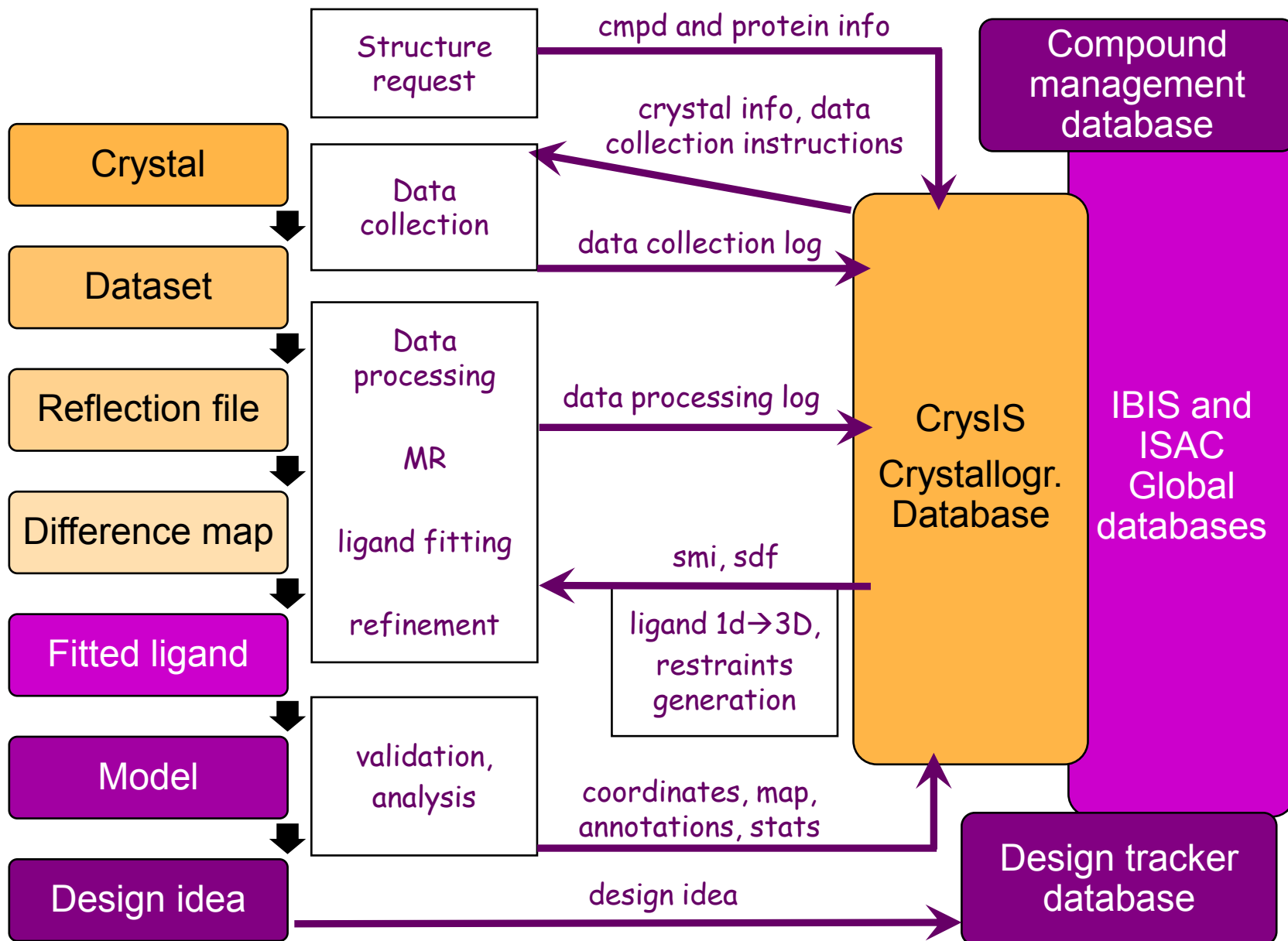
X-ray screening



- primary screening in fragment based design
 - or part of a cascade including biophysical methods or high concentration screening
 - low-ish throughput
 - cocktail screening - bespoke X-ray fragment library based on shape diversity
 - heavy use of automation and databases
-
- 8h shift at the synchrotron:
~100 datasets
 - automated data processing and ligand fitting
 - manual evaluation
-
- if multiple compounds bind:
deconvolution
(single experiments)



Pipelines, software, databases



Pipelines, software, databases



Data collection:

- In-house: sample changer
- Synchrotron: queuing system

Software pipelines:

- in-house - interface corporate databases
 - shell script wrapper for CCP4 (**molrep**, **phaser**, **refmac**) and Global Phasing tools (**autoPROC**, **autoBUSTER**, **grade**, **rhofit**)
 - python wrapper for **pipedream**
- other:
 - **xia2** at synchrotrons
 - **Dimple** at Diamond

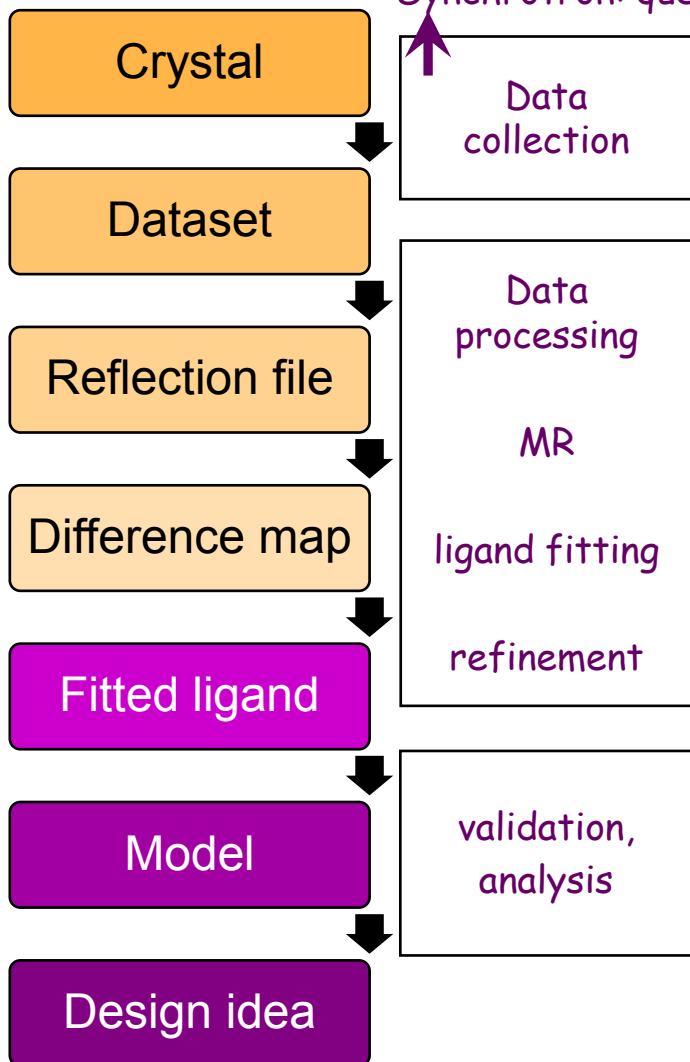
Ligand tools: 1D,2D→3D and fitting

- **grade**, **rhofit** - Global Phasing
- **afitt**, **flynn**, **writedict** - OpenEye
- **corina**
- **acedrg**, **pyrogen**, **libcheck**, **JLigand**, **cprodr**, **Coot** - CCP4

- ligand fitting
- restraints manipulation
- 2D editing
- customisations

Refinement:

- **refmac**
 - **autoBuster** (GPhL)
 - **primeX** (Schrodinger)
- } forcefield for ligands (MMFF)



Pipelines, software, databases



Crystal

Dataset

Reflection file

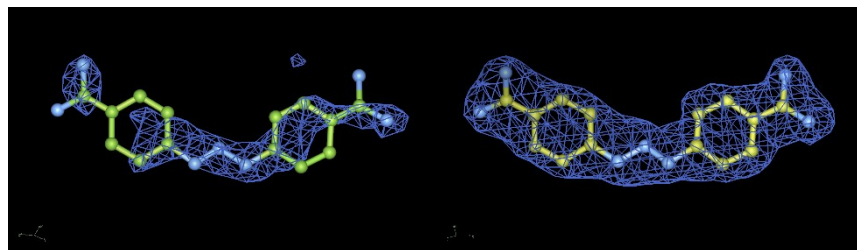
Difference map

Fitted ligand

Model

Design idea

- Is it there? - density fit



- 268D, 2.0 Å
- R=16%, 99% complete
- B = 29 Å³
- RSR: 0.44
- 1D63, 2.0 Å
- R=17%, 73% complete
- B = 31 Å³
- RSR: 0.11

- Is it the right compound?

ISAC-PDB SMILES Check

AZ11596624 Binding site 'INH I 1' of 'lvizp'

Comments

☒ Valid ☐ Invalid

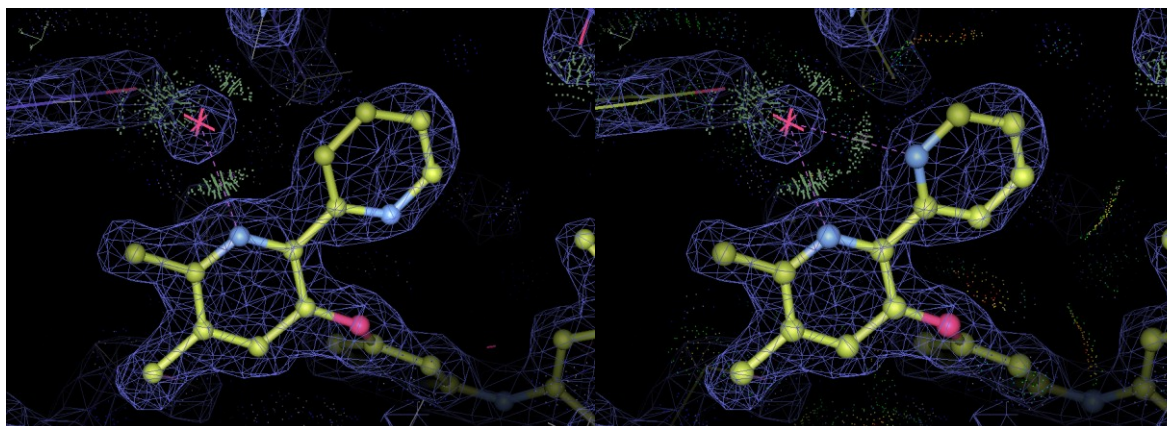
Update Cancel

validation,
analysis

Pipelines, software, databases



- Binding mode? - atom typing, interactions, bonds probe and reduce in *Coot*, comp chem tools



Crystal

Dataset

Reflection file

Difference map

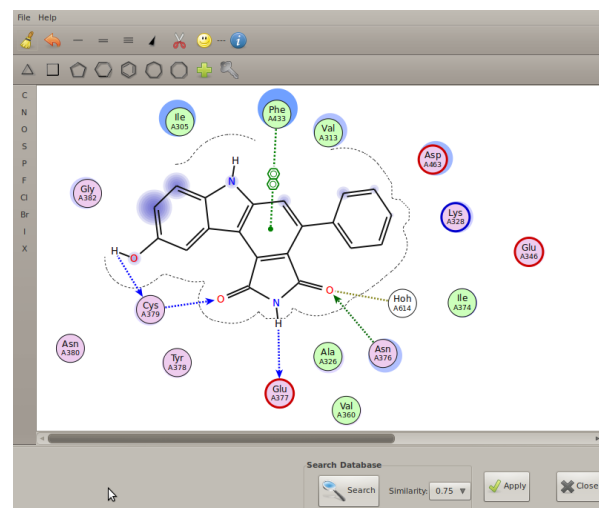
Fitted ligand

Model

Design idea

validation,
analysis

- interactions
Coot,
MOE,
Maestro,
Ligplot
etc



Pipelines, software, databases



Crystal

Dataset

Reflection file

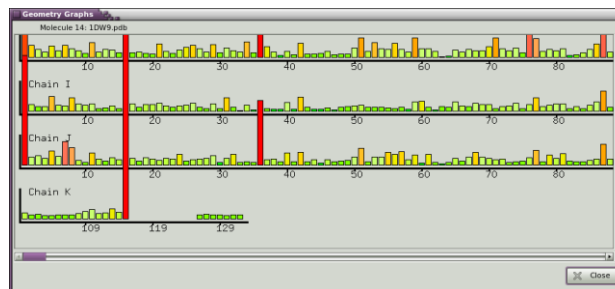
Difference map

Fitted ligand

Model

Design idea

- Geometry - fit to restraints
geometry analysis in *Coot*



- Geometry - compared to small molecule X-ray structures

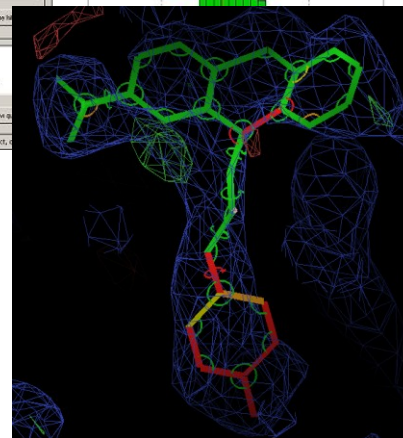
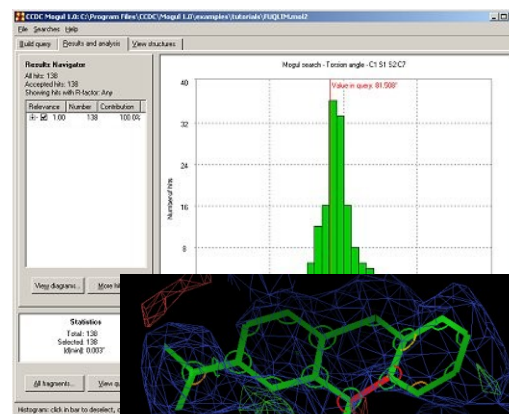
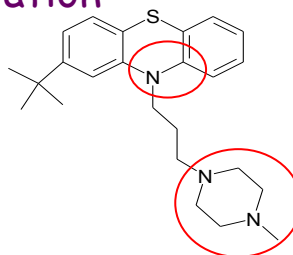
Mogul (CSD)
mean, Z-score, # of hits etc.

Coot

Grade, *acerdg*, *pyrogen*

- restraint
generation

validation,
analysis



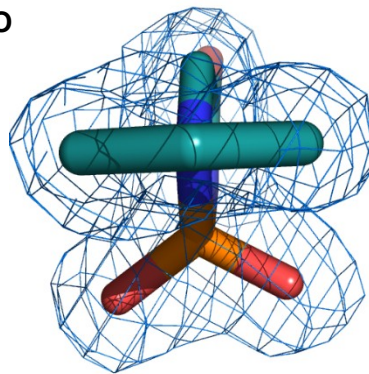
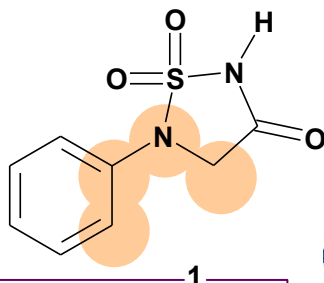
Pipelines, software, databases



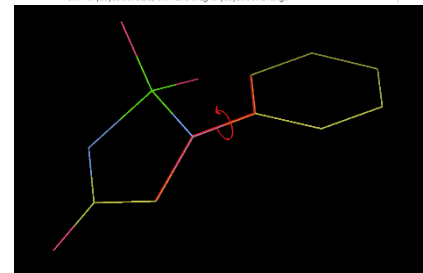
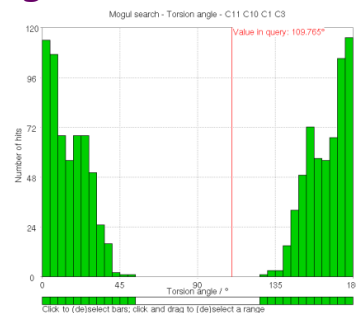
• PDB validation report:

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
7	INH	X	1	-	24,?,?	0.90	1 (4%)	31,?,?	1.33	5 (16%)
7	INH	X	2	-	24,?,?	1.04	1 (4%)	31,?,?	1.76	5 (16%)
7	INH	X	3	-	24,?,?	0.89	1 (4%)	31,?,?	1.04	3 (9%)
7	INH	X	4	-	24,?,?	1.07	1 (4%)	31,?,?	1.46	5 (16%)

- Torsions: not restrained (typically)
- detect energy penalty paid upon binding



validation,
analysis



Design teams – DMTA cycle



- **Crystallographers:**

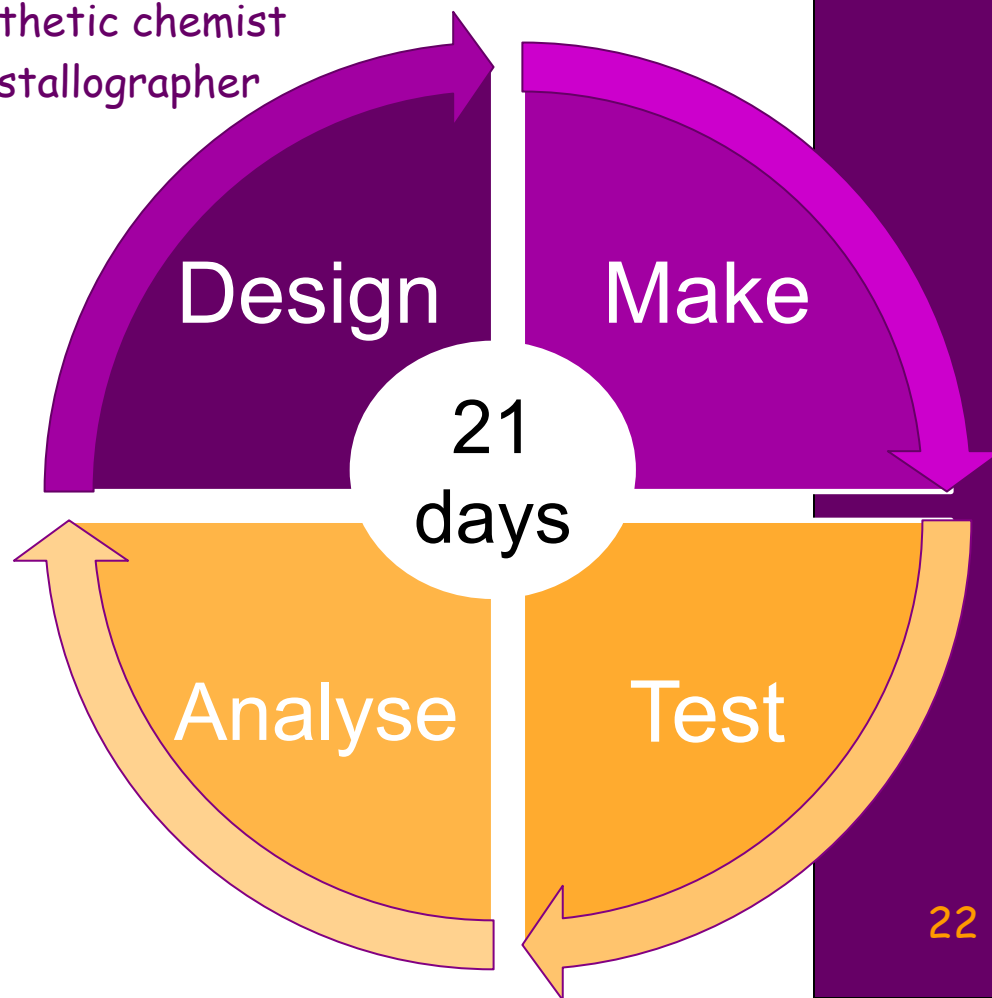
- provide structure
- provide analysis
 - binding site, binding mode
 - interactions → potency
selectivity
 - buried/accessible regions
 - context - comparisons, overlays
 - mode of action
- influence design
- influence synthesis

- **Challenges:**

- 21 day challenge
- education for chemists
- library design is cheaper than bespoke synthesis!

- **Design team:**

- medicinal chemist
- computational chemist
- synthetic chemist
- crystallographer

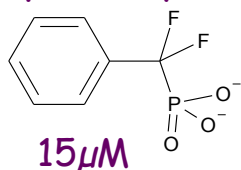


example 1: PTP1B – fragment based design

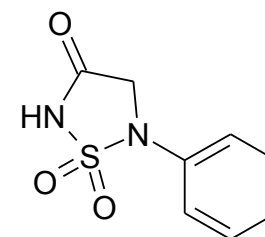


- Diabetes target: negative regulator of insulin signalling (dephosphorylates insulin receptor)
- Obesity target: potentiation of leptin signalling

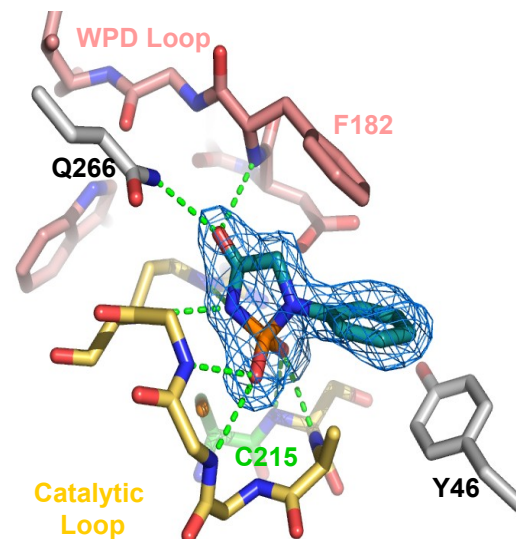
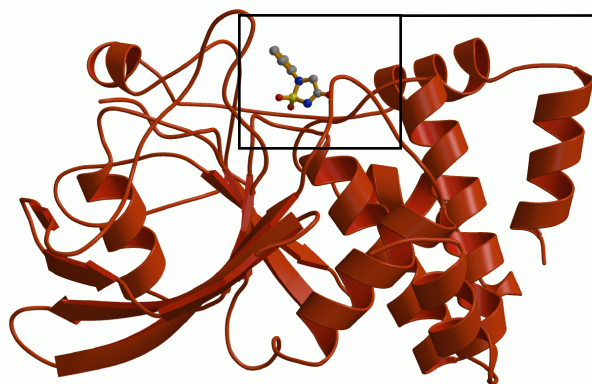
- Full HTS identified 20 000 hits
- all false positives!
- fragment based design starting point: small library: phosphor-Tyr mimetics, e.g.:



15N 2D NMR
screening
→ X-ray



Compound 1
3mM



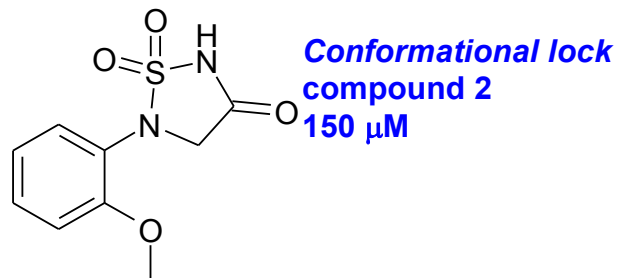
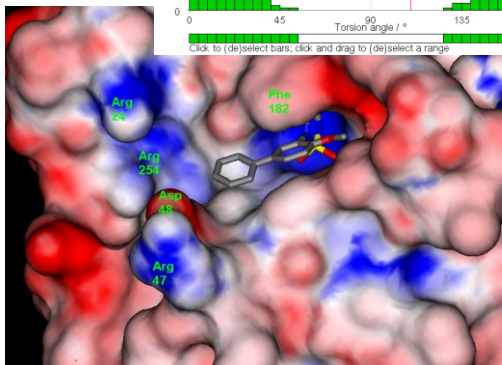
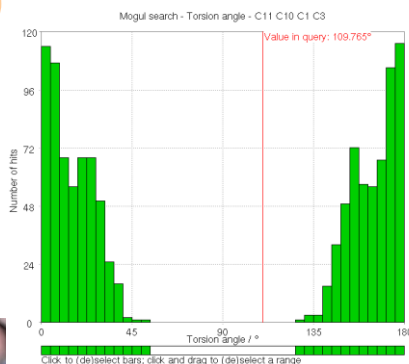
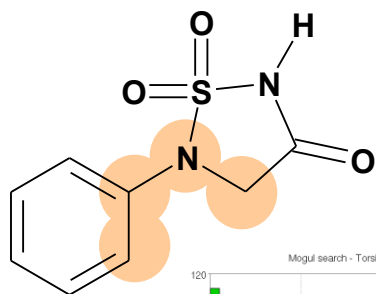
example 1: PTP1B – fragment based design



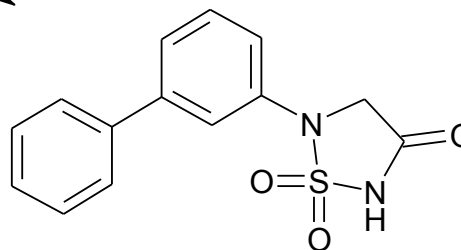
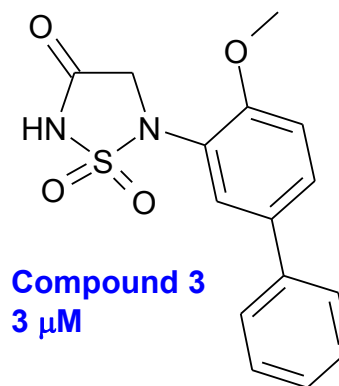
- structure based improvement of fragment hit:

- Starting point

- 1. conformationally strained:



hybridisation



Hydrophobic m-subst
130 μM

example 2: “M” – HTS follow-up

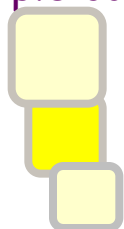


- “M”: novel metabolic enzyme
- oncology target with very few known inhibitors

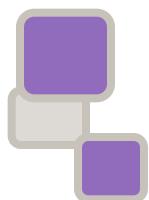
- HTS with full compound collection (2 orthogonal enzyme assays, single shot)
- technology hitter assay
- hit confirmation and profiling with ITC



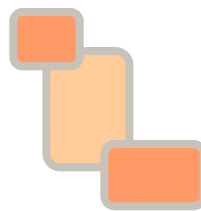
Multiple compound clusters (5 front runners and 4 lesser series):



Series	110
Cluster Size	53
pIC50 M3 (Ave)	6.3 (4.2)
cLogP (Ave)	0.5 (2.5)
pIC50 M1 (Ave)	4.1 (4.1)
ITC binding	Yes

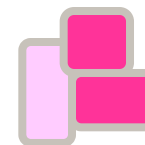


Series	130
Cluster Size	120
pIC50 M3 (Ave)	6.2 (4.8)
cLogP (Ave)	3.9 (3.2)
pIC50 M1 (Ave)	- (5.7)
ITC binding	No



Series	134
Cluster Size	6
pIC50 M3 (Ave)	6.1 (5.3)
cLogP (Ave)	2.7 (1.7)
pIC50 M1 (Ave)	- (5.7)
ITC binding	No

Slope factor > 9



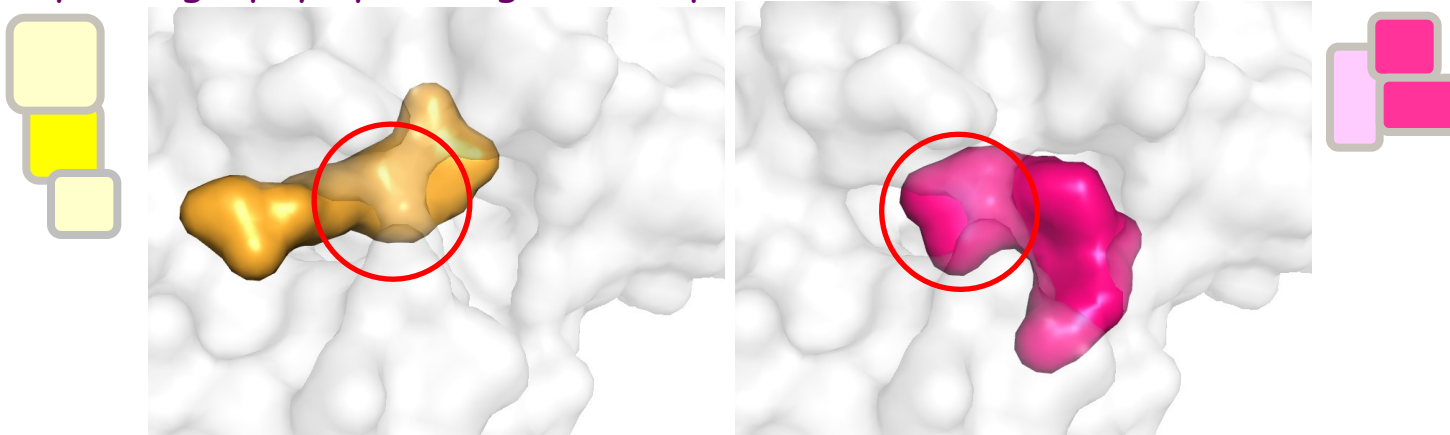
Series	27
Cluster Size	16
pIC50 M3 (Ave)	6.5 (4.8)
cLogP (Ave)	4.6 (4.5)
pIC50 M1 (Ave)	7.4 (7.0)
ITC binding	Yes

example 2: “M” – HTS follow-up

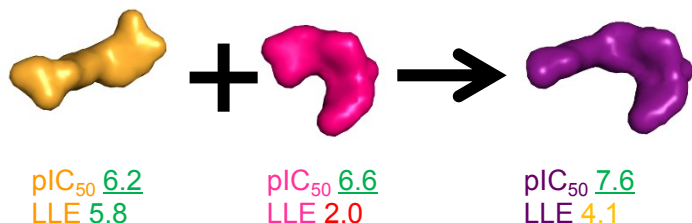


- structure based approach -
towards improved potency and better phys-chem props

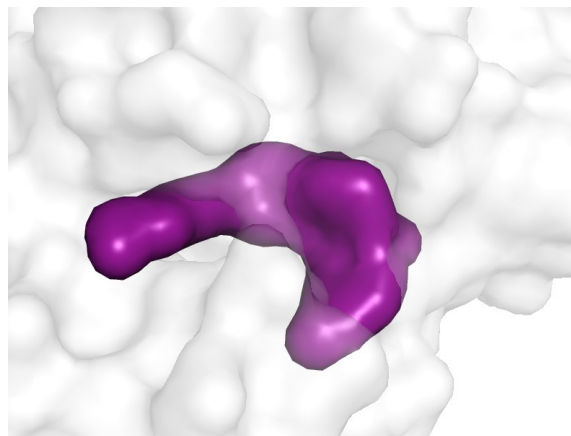
Crystallographic profiling of multiple series:



hybridisation



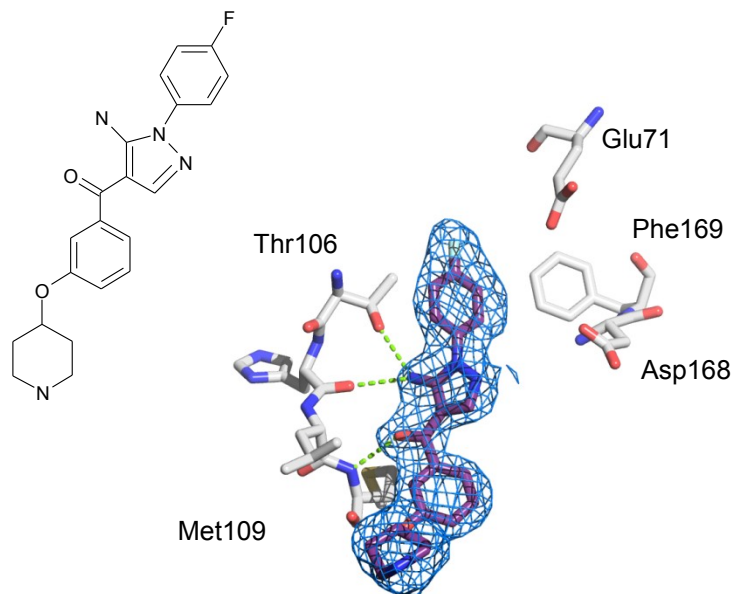
Novel, more potent compound series:



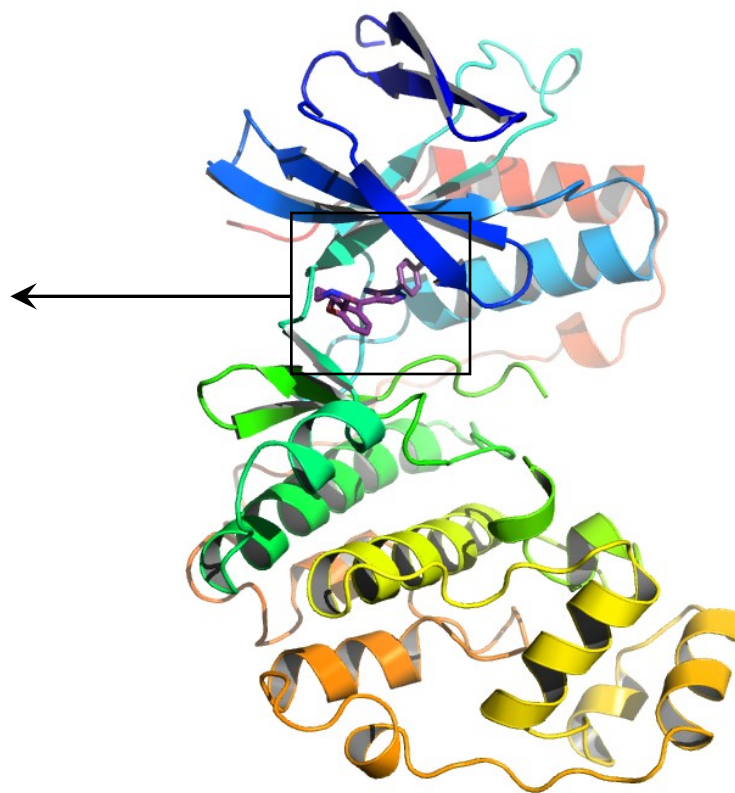
example 3: P38 – unclear SAR



- pyrazolamines



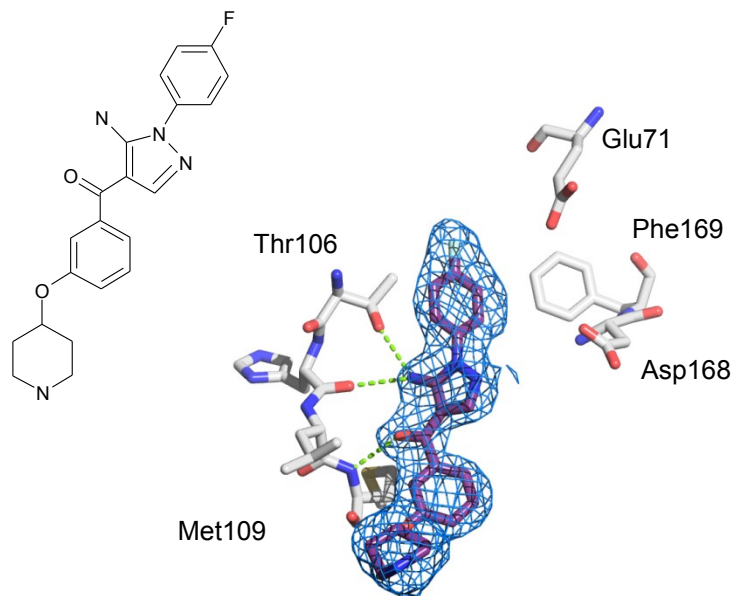
- p38 MAP kinase:
 - activated by MKK3 and 6
 - mediates the release of $\text{TNF-}\alpha$
 - rheumatoid arthritis target



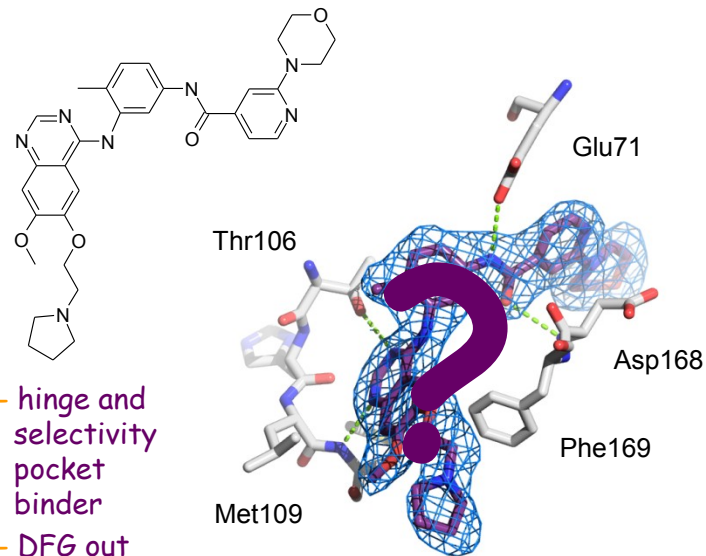
example 3: P38 – unclear SAR



- pyrazolamines

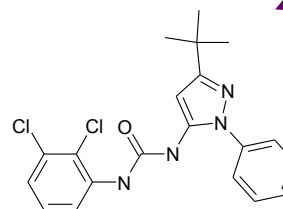
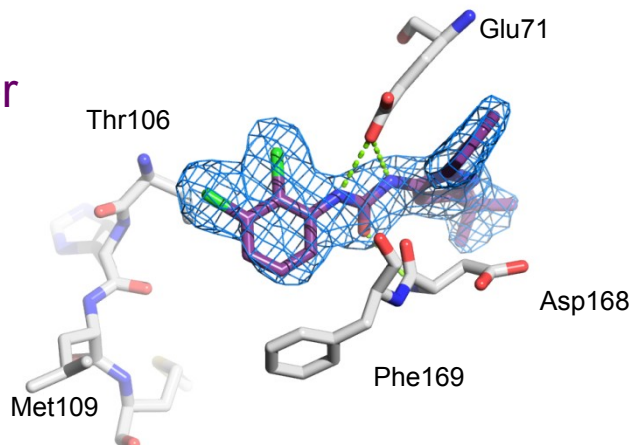


- MPAQ series: - SAR unclear
- modelling unsuccessful



pyrazoloureas:

- no hinge binder component
- DFG out



example 4: ALK5 – improving phys-chem props

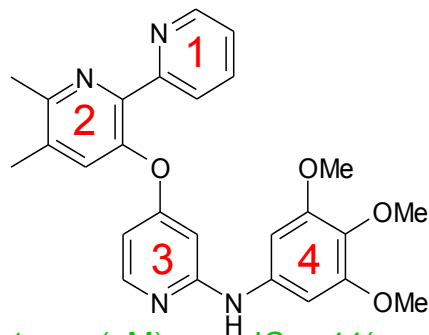


- ALK5: kinase involved in TGF- β signalling, phosphorylated smad2 and 3
- ALK5 inhibitors: against TGF- β driven tumours.

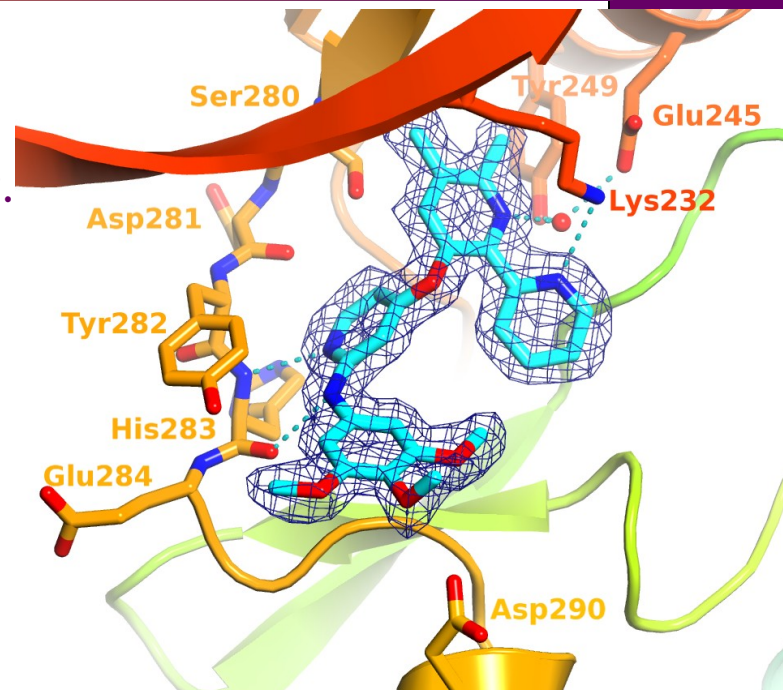
- Structure used for docking: kinase binders and their fragments
- library design: combinations of hinge binder, solvent channel and selectivity pocket groups
- Conserved interactions retained

library synthesis: novelty, tractability screening

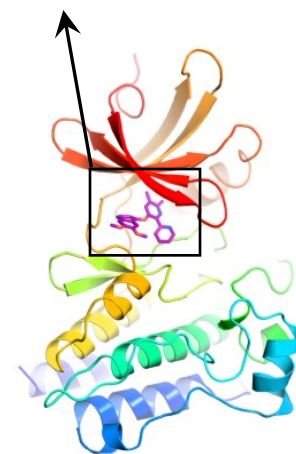
Hit:



- Potency (nM): IC_{50} : 44(enzyme), 55(cell)
- Bioavailability: F: 4%
- Lipophilicity: $\log D_{7.4}$: 3.2



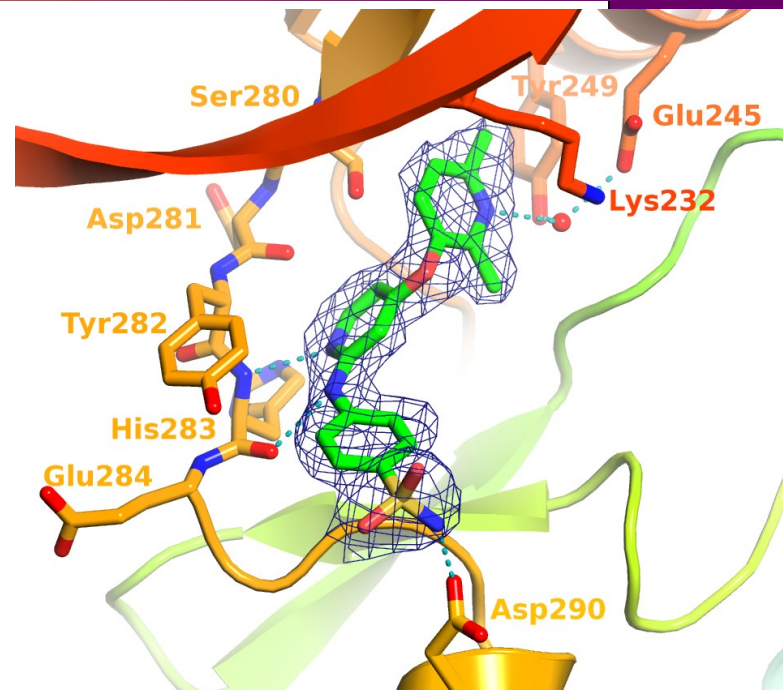
what drives potency?



example 4: ALK5 – improving phys-chem props

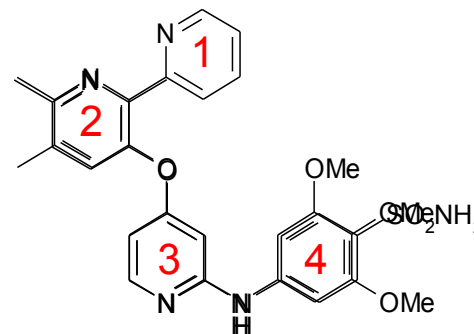


- crystallography of fragments of known binders to assess individual binding modes
- **fragment** assisted approach for optimisation: strained binding?
- systematic variation of N atoms in selectivity pocket group
- library screening of solvent channel group



SAR:

- ring **1** interaction not important
- large, electron withdrawing group in solvent channel



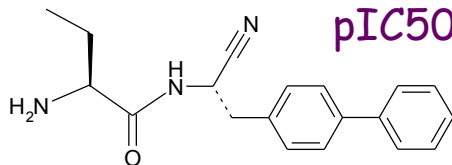
- Potency (nm): IC_{50} : 22(cell)
- Bioavailability: F: 75%
- Lipophilicity: logD7.4: 2.5

example 5: CatC – improving stability and PK

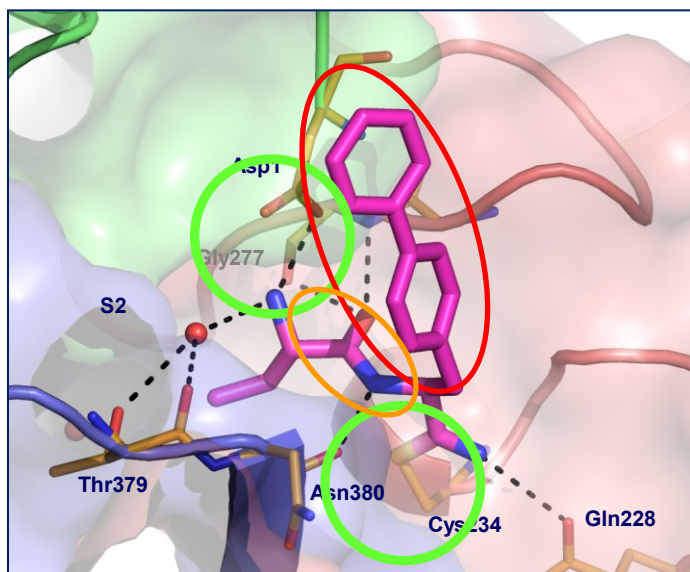


- CatC (DPP1): Cys protease
- activates pro-inflammatory proteins (NE, CatG, PR-3)
- COPD target

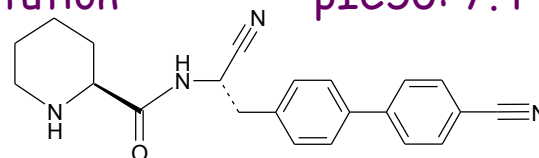
Known inhibitors: covalent warheads
pIC50: 8.4



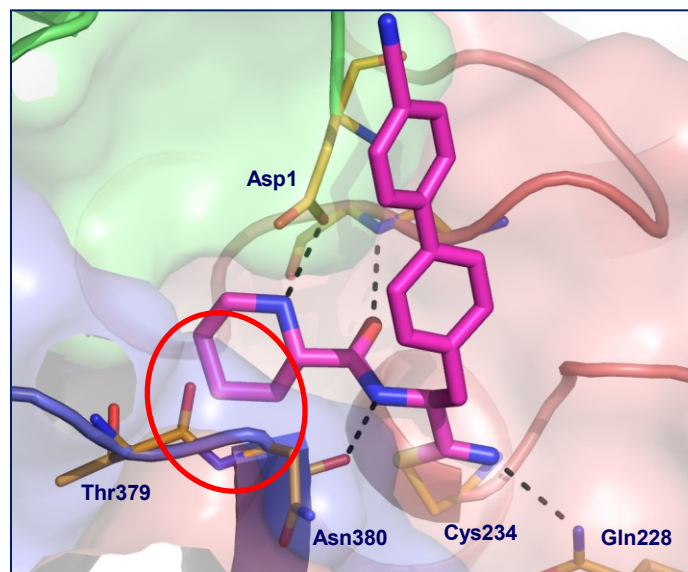
Issues: Plasma instability
Metabolic clearance



Cyclisation and electron withdrawing
substitution
pIC50: 7.4



Improved stability and clearance
But: potency loss



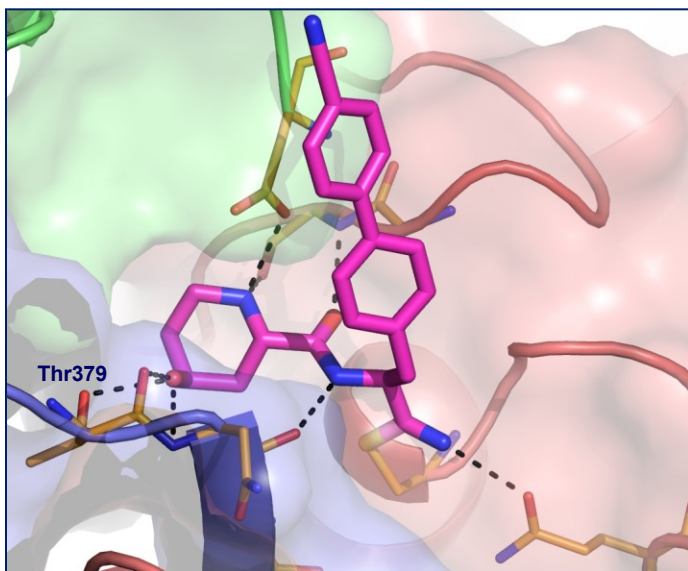
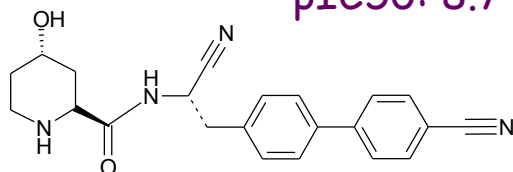
example 5: CatC – improving stability and PK



- regaining potency
- strategy: keep cyclisation, incorporate H-bonds
- two approaches:

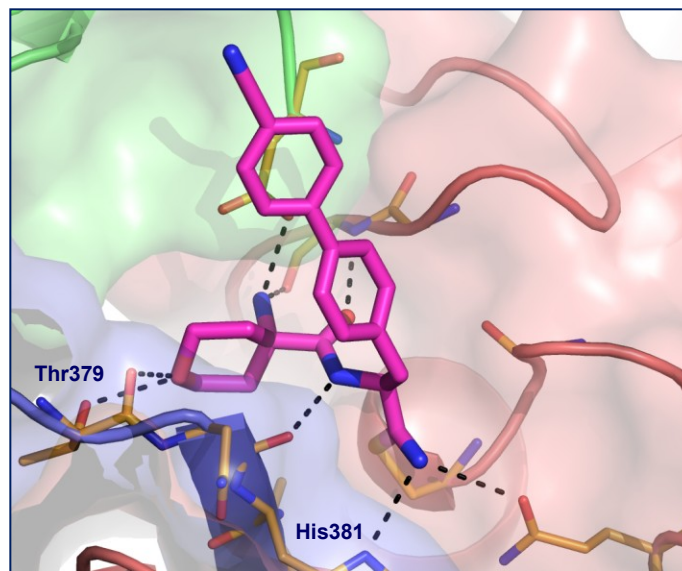
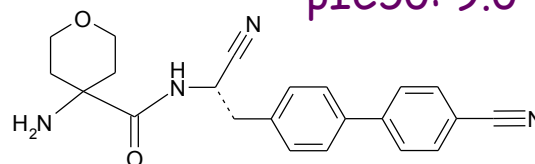
1. 4(S)-hydroxyl

pIC₅₀: 8.7



2. Pyran ring

pIC₅₀: 9.0



example 6: IL15/17 – biologics

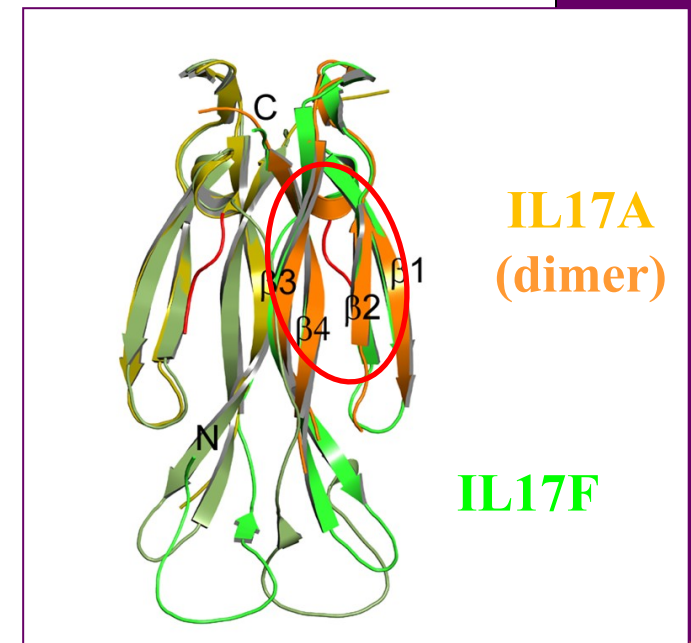
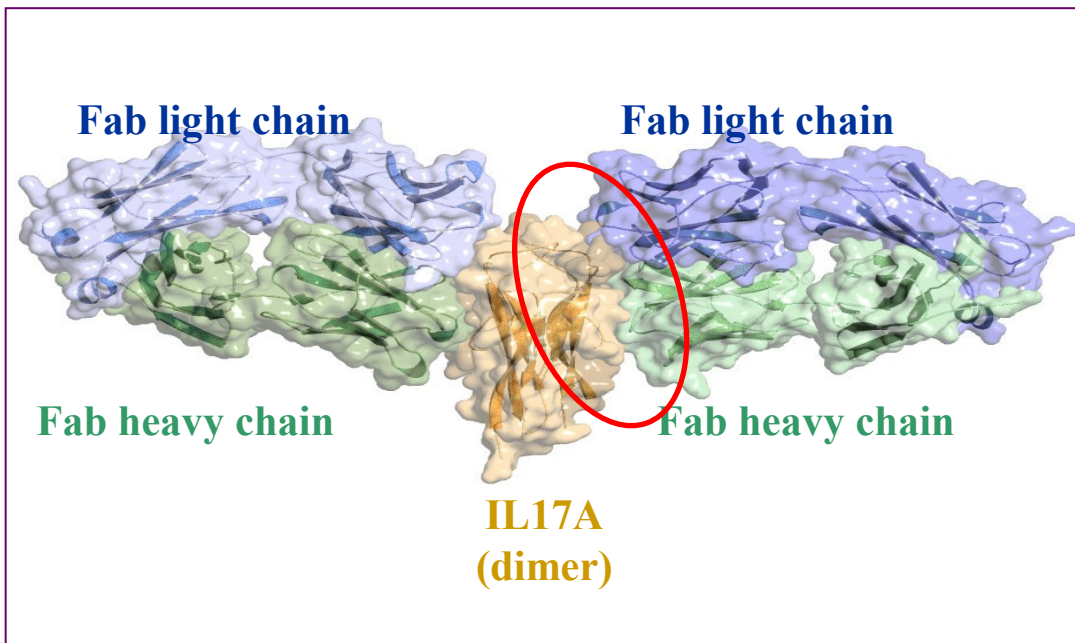


Crystallography in biopharmaceuticals: antibodies,

- providing definitive epitope mapping for patenting
- providing a structural context for understanding optimisation

IL17A

- dimeric cytokine
- role in adaptive immune response and maintaining inflammatory responses





- fuzzy comp.chem-crystallography boundaries
 - force fields in refinement or parameterless refinement
 - CSD and COD in protein crystallography
 - X-ray terms in comp chem minimisation
- integrated structure and biophys groups
- new screening techniques, e.g. DNA encoded libraries