X-RAY CRYSTALLOGRAPY AND STRUCTURAL BIOLOGY IN THE PROCESS OF DISCOVERY AND DEVELOPMENT OF NEW DRUGS

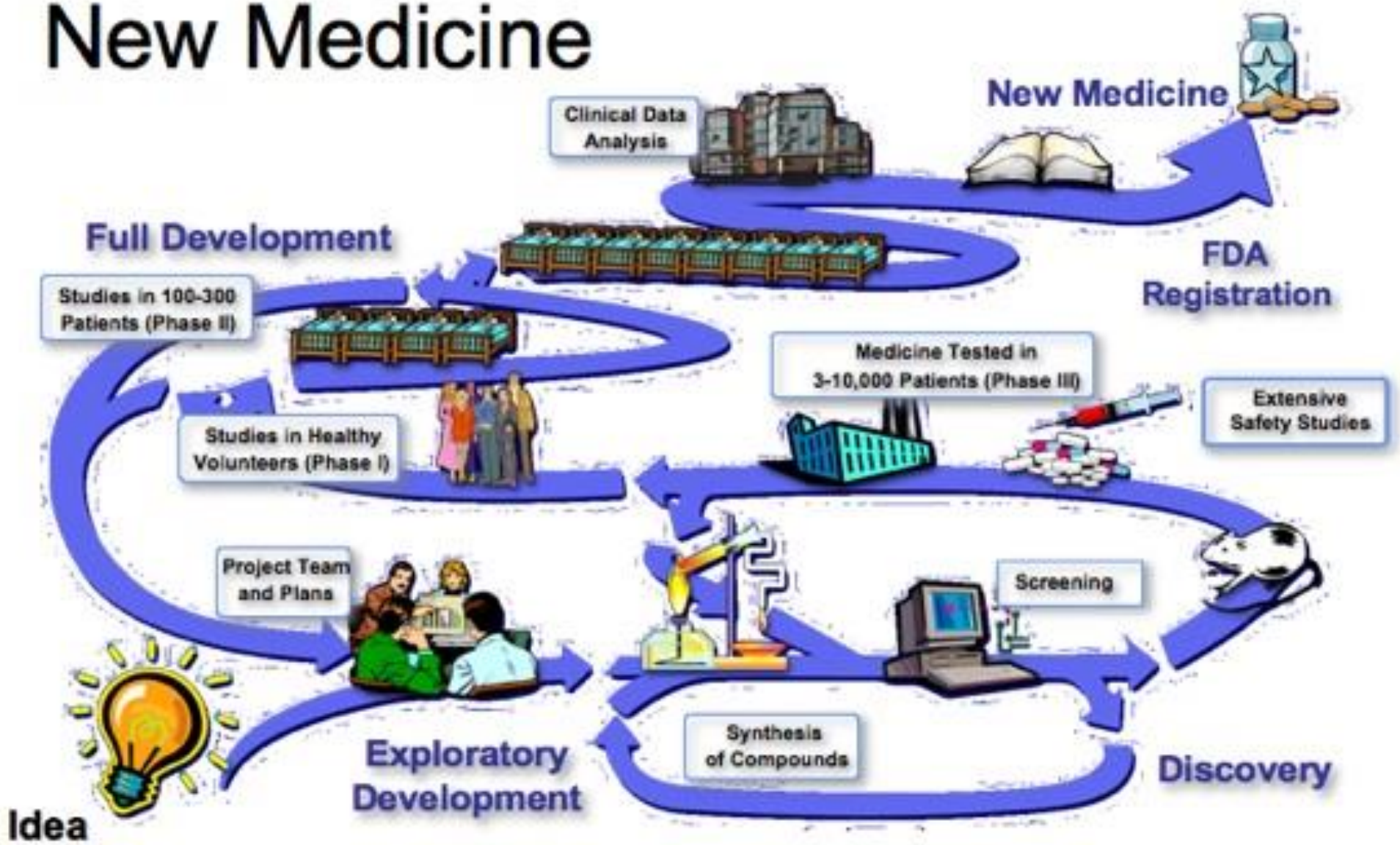
Matthias Frech
Lisbon, July 2022

MERCK

Agenda for today

- Introductory remarks on the Drug Discovery Process, with special focus on the early phase
- Where to place Crystallography / Structural Biology
- The Organization and Automation to get Structural Information
- Selected Examples
- Outlook

The Long Road to a New Medicine



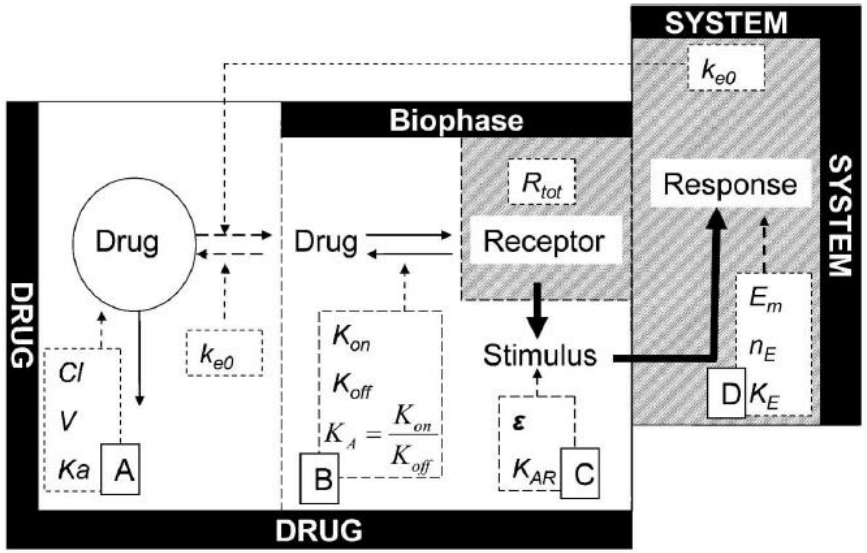
How do medicines work?

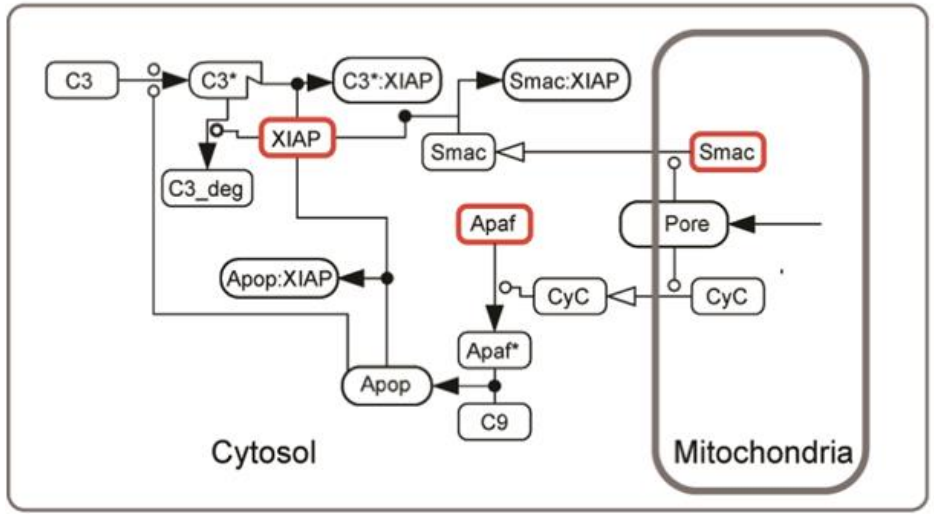
Drug action begins with binding

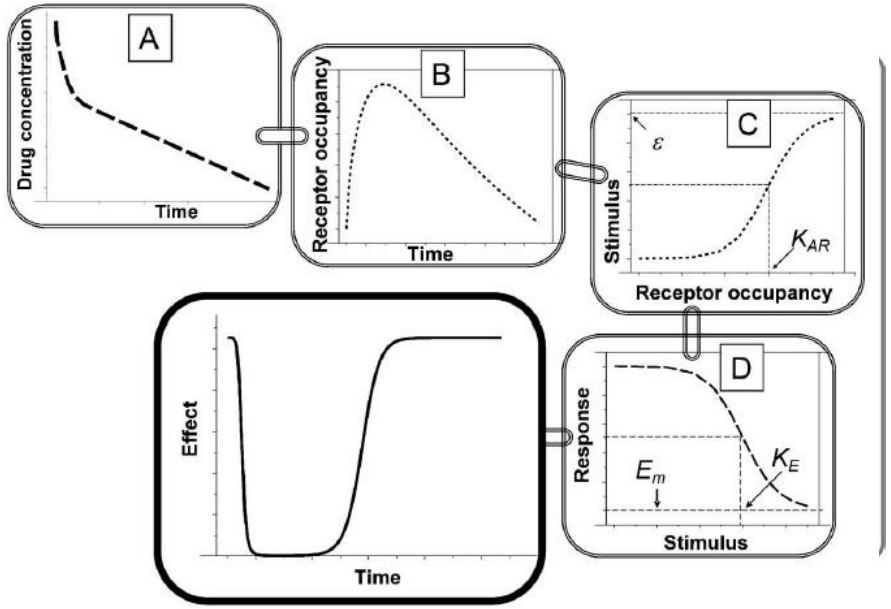
'Corpora non agunte nisi fixata'

A substance will not work unless it is bound

-Paul Ehrlich, 1913

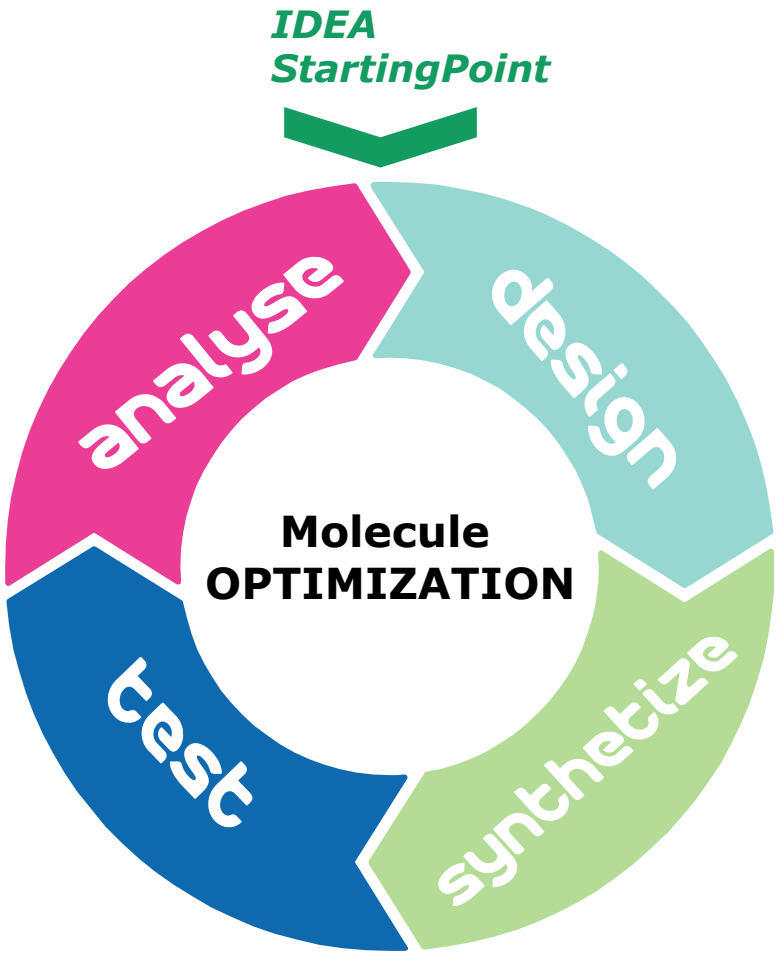


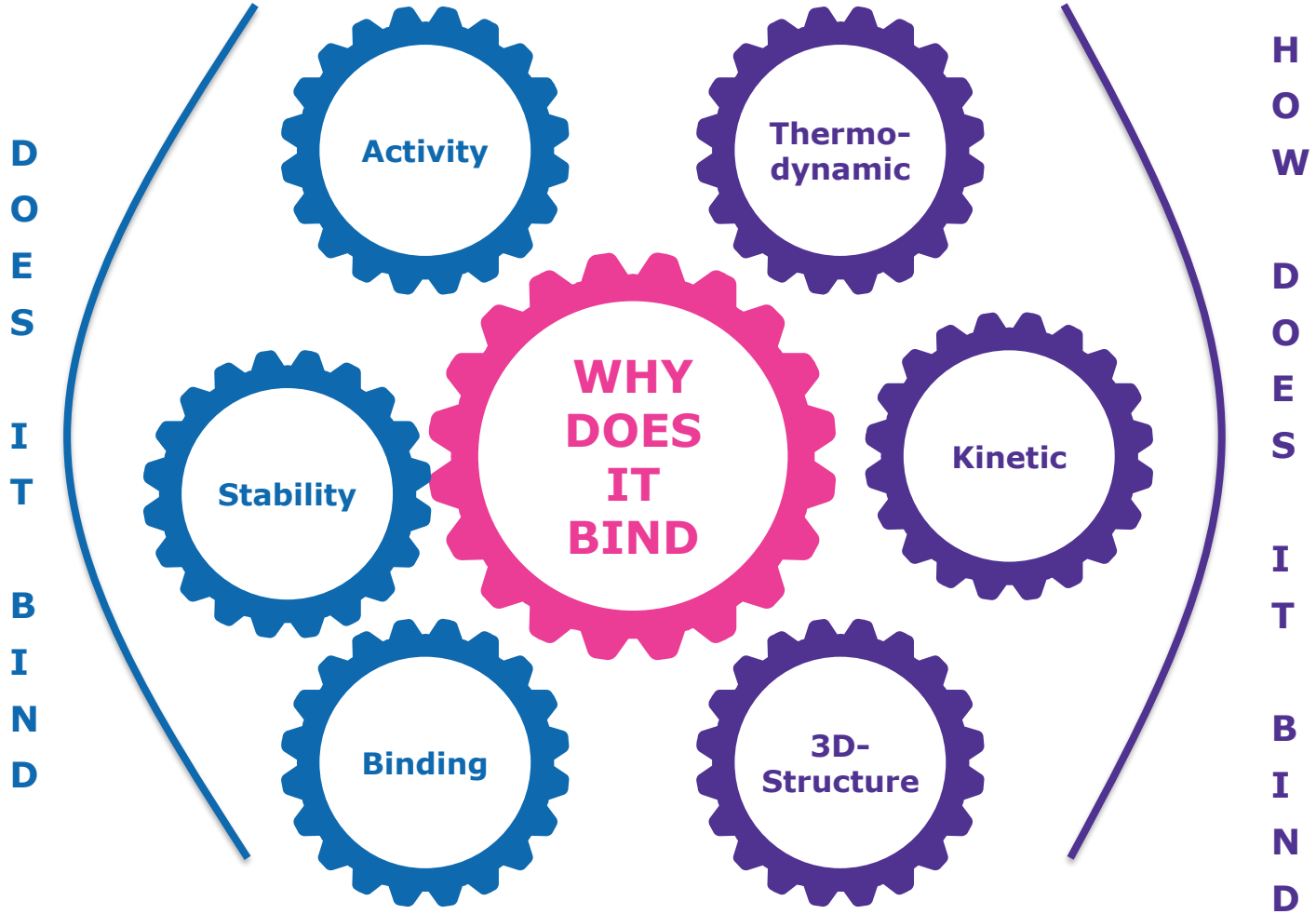




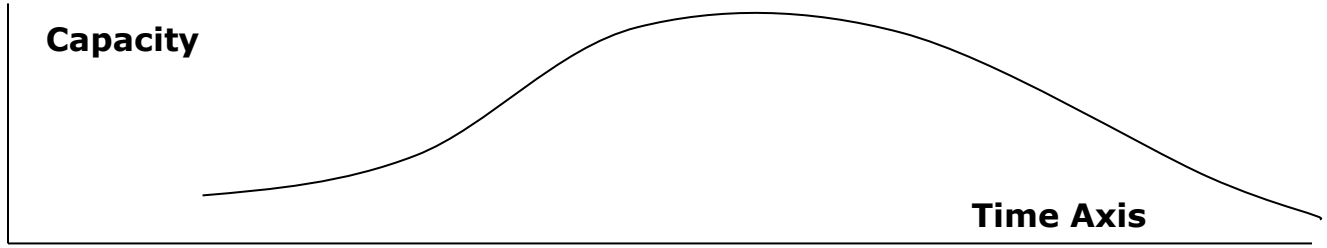
Ploeger BA, van der Graaf PH, Danhof M. Drug Metab Pharmacokinet. 2009;24:3-15.

Dissection of Molecular Interaction





Structural Biology is an essential Part in the early Phase of the Drug Discovery Process



Target research

Verification of target-ligand, target-substrate or target-small molecule interactions

Hit discovery

- Validation and ranking of HTS hits
- Identification of promiscuous binders
- Hit series prioritization
- Fragment screening

Hit optimization

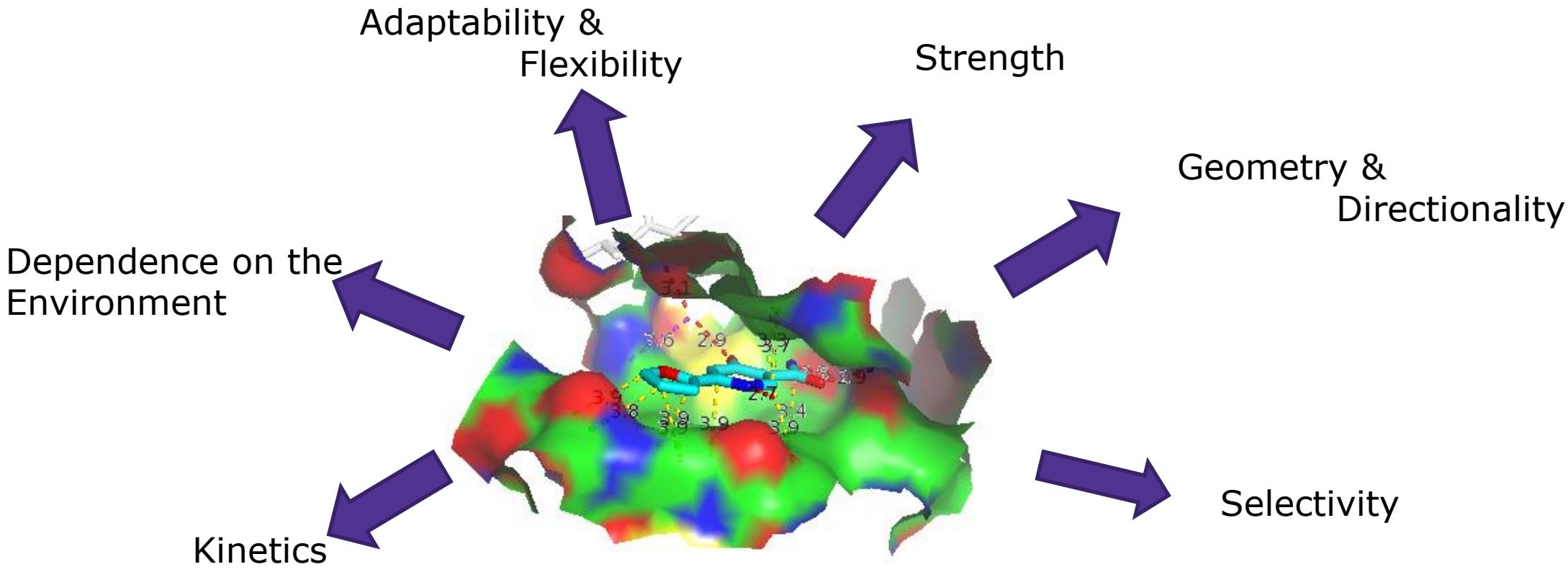
Lead optimization

- Protein Co-crystal structures to deliver insight into binding modes
- Support of chemical synthesis by delivering thermodynamics and kinetics as additional decision criteria
- Rank very high affinity binders
- Gaining insight into mode of mechanism
- Profiling of drug candidates and backup series regarding their binding to isoforms or homologs of the target protein

further development

Molecular Interactions for New Chemical Entities

Can be regarded as „bridge“ between the world of chemistry and biology.



- Valuable and instrumental to manipulate biologial functions with molecuar precision to eventually generate a benefit for patients.

Structure Biology in Drug Discovery at Merck

Methods we use to get Three Dimensional Structural Information

X-Ray Protein Crystallography

Cryo Electron Microscopy

Protein based Nucler Magnetic Resonance Spectroscopy

More for analytical questions:

Negative Stain Electron Microscopy

Small angle X-Ray Scattering

Dynamic Light Scattering

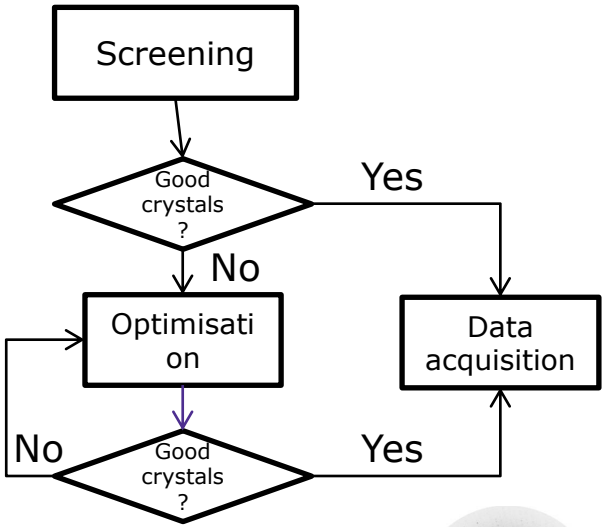
Protein Crystallisation – Still tedious

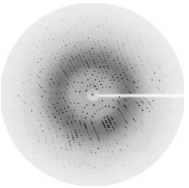
- The buffer composition for optimal crystallisation need
  to be discovered empirically

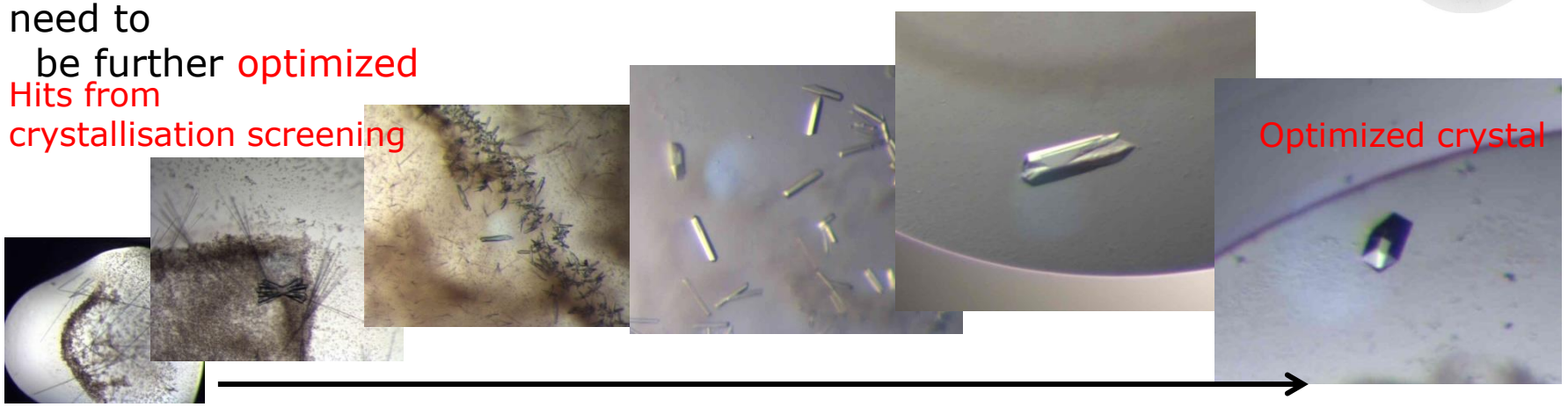
- The buffer solution contains precipitants like PEG or
  amoniumsulfate to form crystals

- Different variations of buffer solutions are used for
  crystallisation screening

- Initial crystals from crystallisation screening need to
  be further optimized





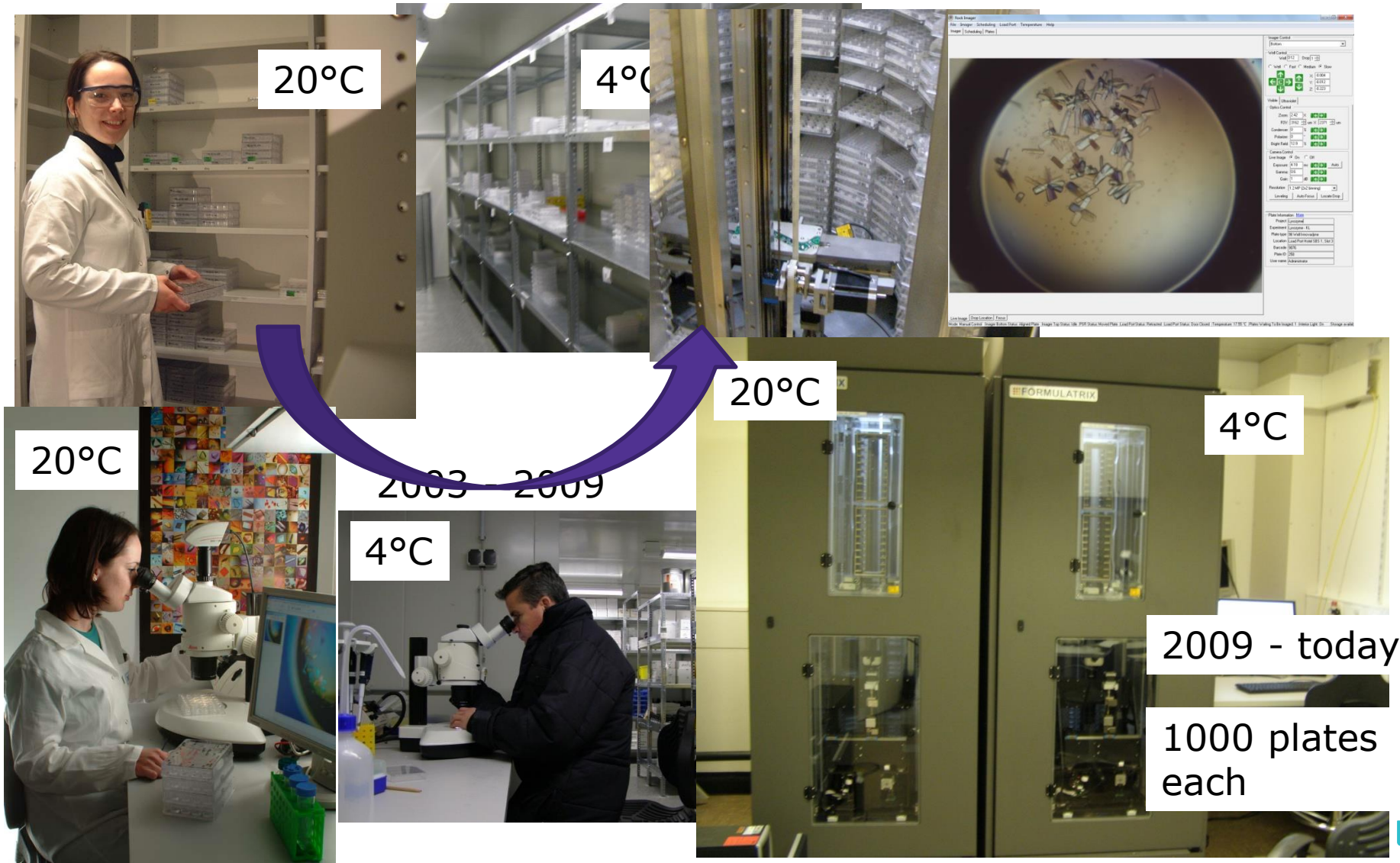


Automation in the crystallisation lab

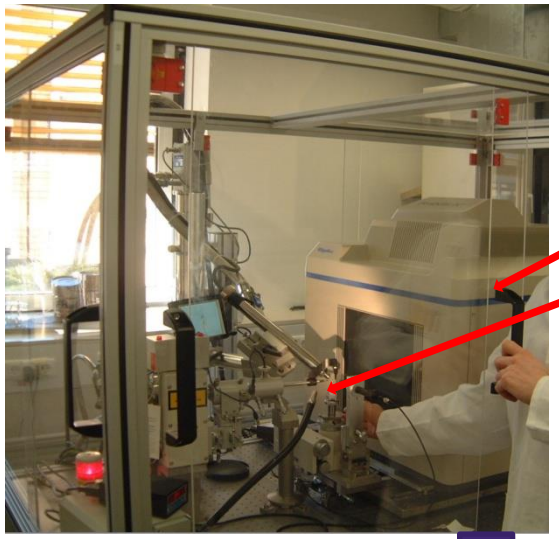
- **2003: Manually using 1µl protein droplet**
  1 Screening plate : 90 min 384 screening conditions
  1 Opt. Plate: 25 min
- **2005: Tecan Evo using 0,5µl Protein droplet**
  1 Screening plate : 90 min **576** screening conditions
  1 Opt. Plate: 25 min
- **2011: ARI Phoenix using 0,1µl Protein droplet**
  1 Screening plate : **5** min **1728** screening conditions
- **2013: Formulator using 0,1µl Protein droplet**
  1 Opt. Plate : **5** min 24 conditions / plate

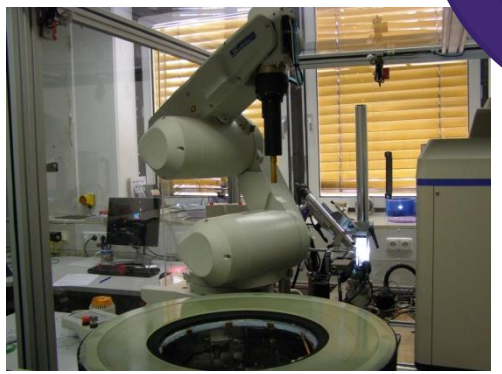
With the same protein amount, number of crystallisation conditions is increased by a factor of 4.5

Automation of crystallisation process

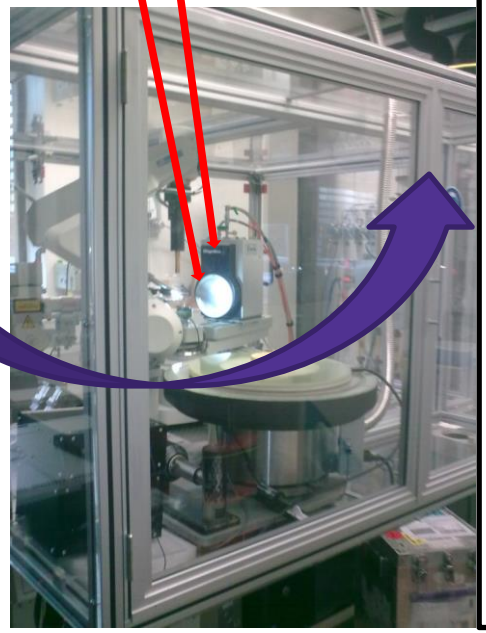


Characterization of protein crystals in-house



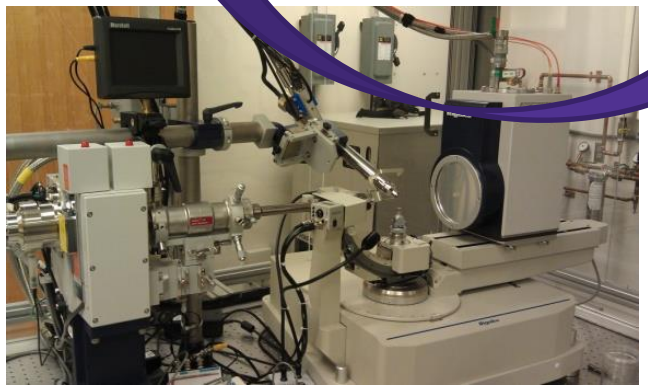


Sample changer



2004:	Diffractometer setup
2004-2009:	1 dataset 2-4 days
2008:	New 40 Sample changer robot in place
2009:	New CCD detector in place 1 dataset 2-3 hours
2020:	New X-Ray source installed

Measurement of protein crystals at the Swiss Light Source / Villigen



1 dataset less than 2 minutes

In-house tests with sample changer:
preselection of best diffracting crystals

Automation in Structural Biology

- **The process to deliver a protein structure can be divided into a few basic working packages**

- **Crystallization** – a partly automated step. To titrate one plate with hundreds of different conditions it takes about 3 min. Plate is transferred manualy to the imager system. Continious observation of protein crystallization is done in an automated way. Inspection and interpretation is done by human intervention.

- **Data Aquisition** - nowadays at the Synchrotron aquiring of a data set takes less than 2 min per protein crystal

- **Data Processing** – process the synchrotron data until a first structure is automated can work with 80 data sets over night. It delivers the information: structure can be solved and refined further, the ligand can be detected

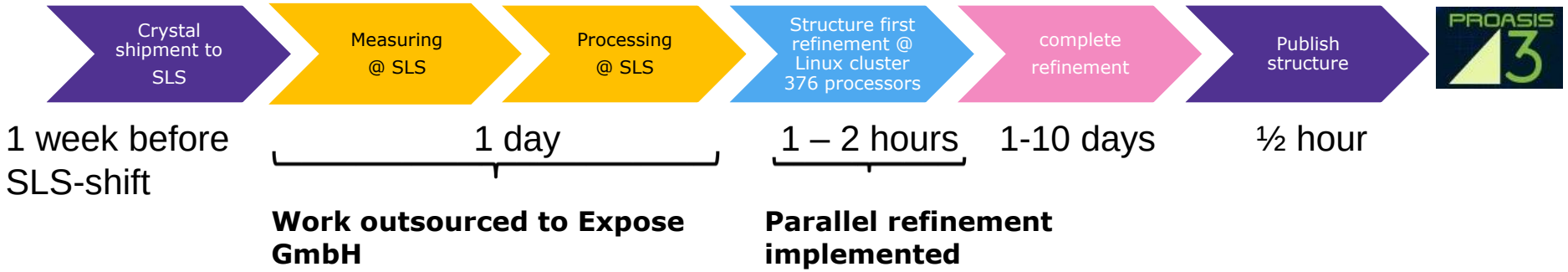
- **Structure Refinement** – the final refinement to a high resolution can take only a day but also weeks of time depending on data complexity, quality, available capacity, experience of involved crystallographer

Improvements in timelines for data processing, structure refinement and publishing 2003 - today

Until 2008: Timeline 4.5 – 8.5 weeks

- Travelling to SLS (once per month) and all steps sequentiell

Today: Timeline 1.5 – 3 weeks



- Global coordination of synchrotron logistics between Da, Billerica and IBET/Portugal
- Crystals are shipped to synchrotron (every 3 weeks)
- 1st refinement of up to 98 datasets in parallel

Number of structural requests and solved structures

2011 – 2013

requested structures	766	
Structures delivered	395	
In-house	245	
In collaborative refinement efforts		
Extern	150	

2014 – 2016

requested Structures	1050	
Structures delivered	583	
In-house	417	
In collaborative refinement efforts		-
Extern		166

2017 (till Nov.)

requested Structures	386	
Structures delivered	275	
In-house	36	
In collaborative refinement efforts		56
Extern		183 (Vertex 116)

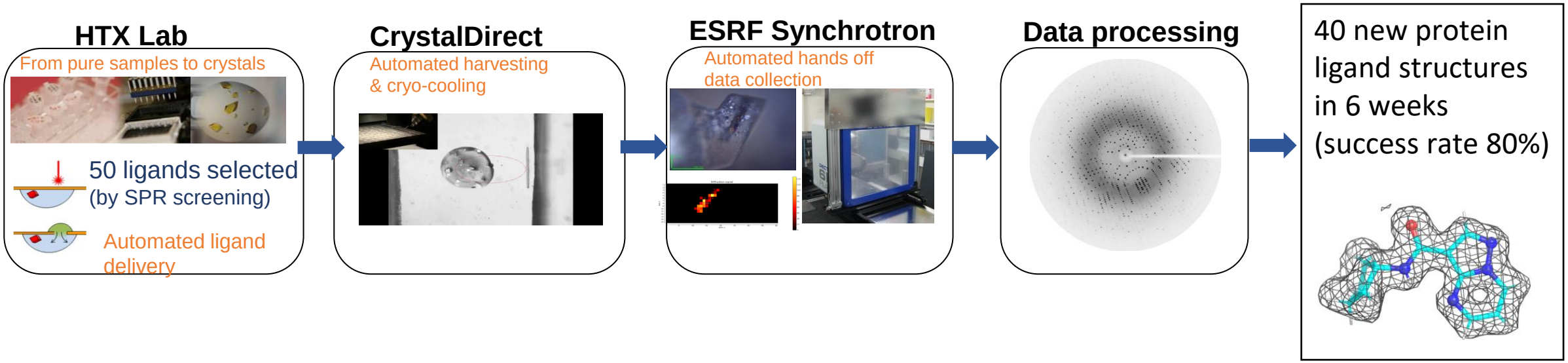
We have an attrition in our structure solution process

- Within one year 665 crystals were sent the synchrotron.
- This resulted in 360 datasets for processing and refinement
- Which delivered 159 solved structures

- Bundles of co-crystals with defined priorities are shipped to the SLS in Switzerland
- The inhouse produced protein crystals are prechecked with the inhouse Xray source to ensure a basic quality and reduce drop out at the synchrotron.
- The first refinement is done over night. Giving the information: ligand is bound , structure can be refined

High Throughput Protein Crystallization in collaboration with the EMBL Core Facilites (Grenoble)

- High throughput crystallography to complement HTS-Eval generating more diverse structural data set in the early HO phase to support the selection of promising chemical scaffolds.
- Enhanced throughput of selected HO / LO projects to generate protein ligand structures
- Use protein crystallography as a screening method for fragments



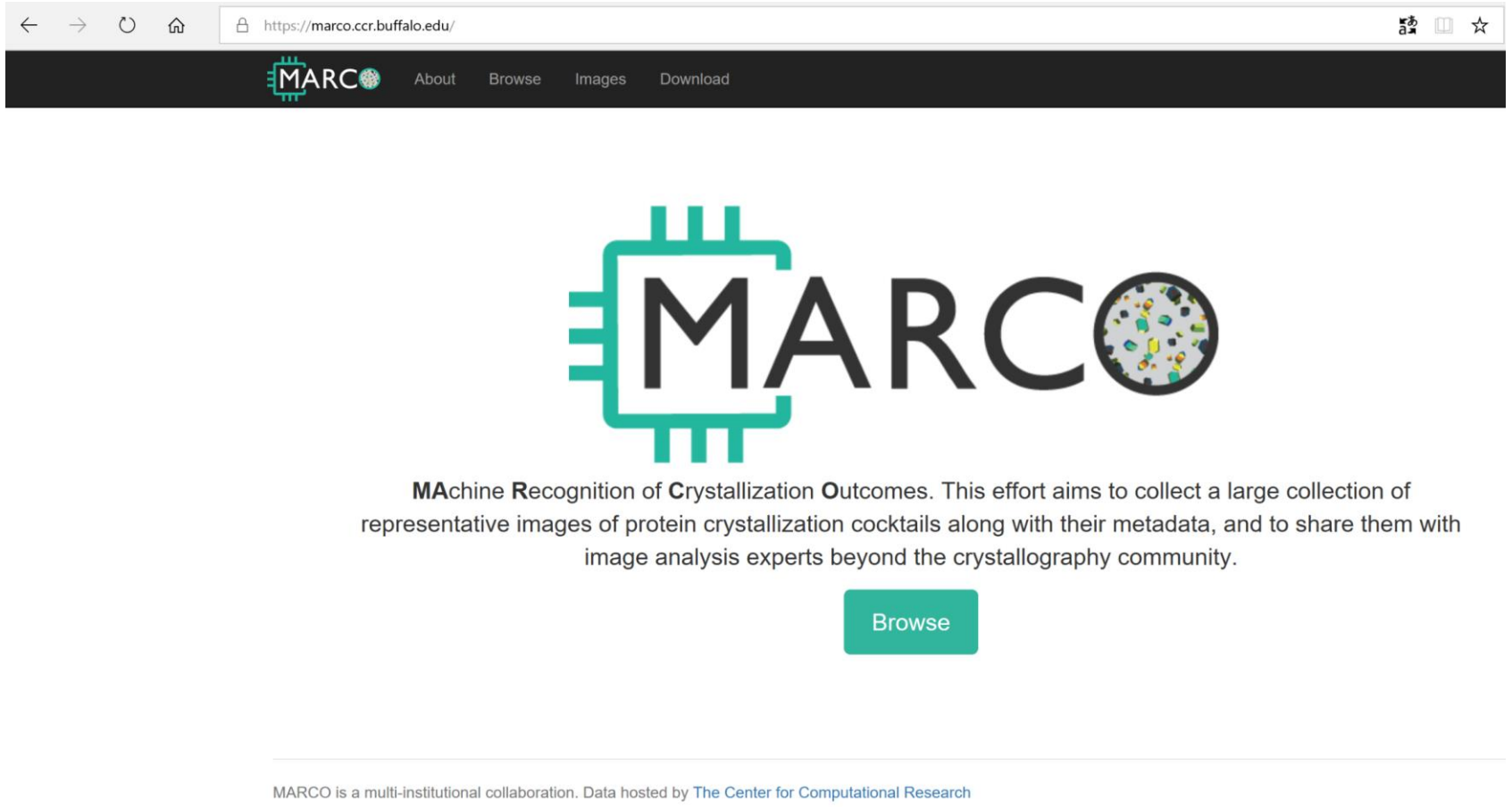
- Improved (current) throughput: 50 compounds soaked, measured and bound hits partially refined in 3 weeks
- Appropiate selection of project is mandatory. Soakable system is needed, cystallization system needs be ready at the time of the HTS to facilitate the seamless transfer to Grenoble, Refinement resources needs preparation in advance

Automated Protein Crystal Detection
MARCO Initiative: pre-competitive initiative

Collaborative initiative to develop an automated computer vision-based system to identify protein crystals with machine learning.

Contributors:

- GlaxoSmithKline
- Merck & Co.
- Hauptman-Woodward Medical Research Institute
- Collaborative Crystallization Centre
- Bristol-Myers Squibb
-



Automated Protein Crystal Detection
Project Description

Objective: Develop an automated system to acquire and identify images of protein crystals

Background: Crystallographers look through individual pictures of crystallization drops to identify protein crystals. Manual review is time consuming and sometimes difficult to identify a protein crystal from other contaminants such as precipitate, debris, etc. An automated system could save time and potentially identify 'hard to find' crystals

In the future, use other biophysical measurements like UV/VIS or SONICC system from Formulatrix to accelerate learning and identification of crystals

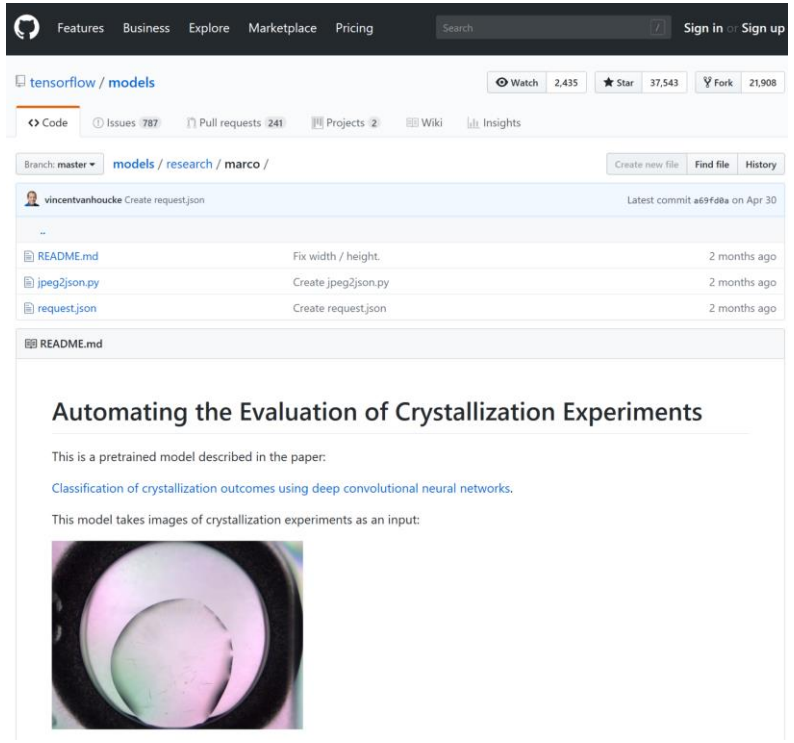
Assets available: database of protein crystallization experiments and images (not annotated)

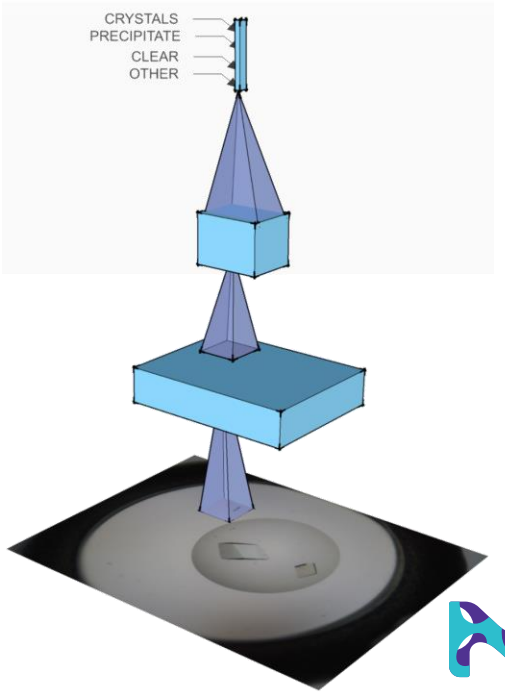
Automated Protein Crystal Detection
MARCO Initiative: publicly available assets

Images including training and test sets (~6GB)

Institution	Technical Setup	# of Images
Bristol-Myers Squibb	Formulatrix Rock Imager (FRI)	8719
CSIRO	Sitting drop, FRI, Rigaku Minstrel [23, 24]	15933
HWMRI	Under oil, Home system [15]	79632
GlaxoSmithKline	Sitting drop, FRI	83126
Merck	Sitting drop, FRI	305804

Github repository for machine learning algorithms (TensorFlow)

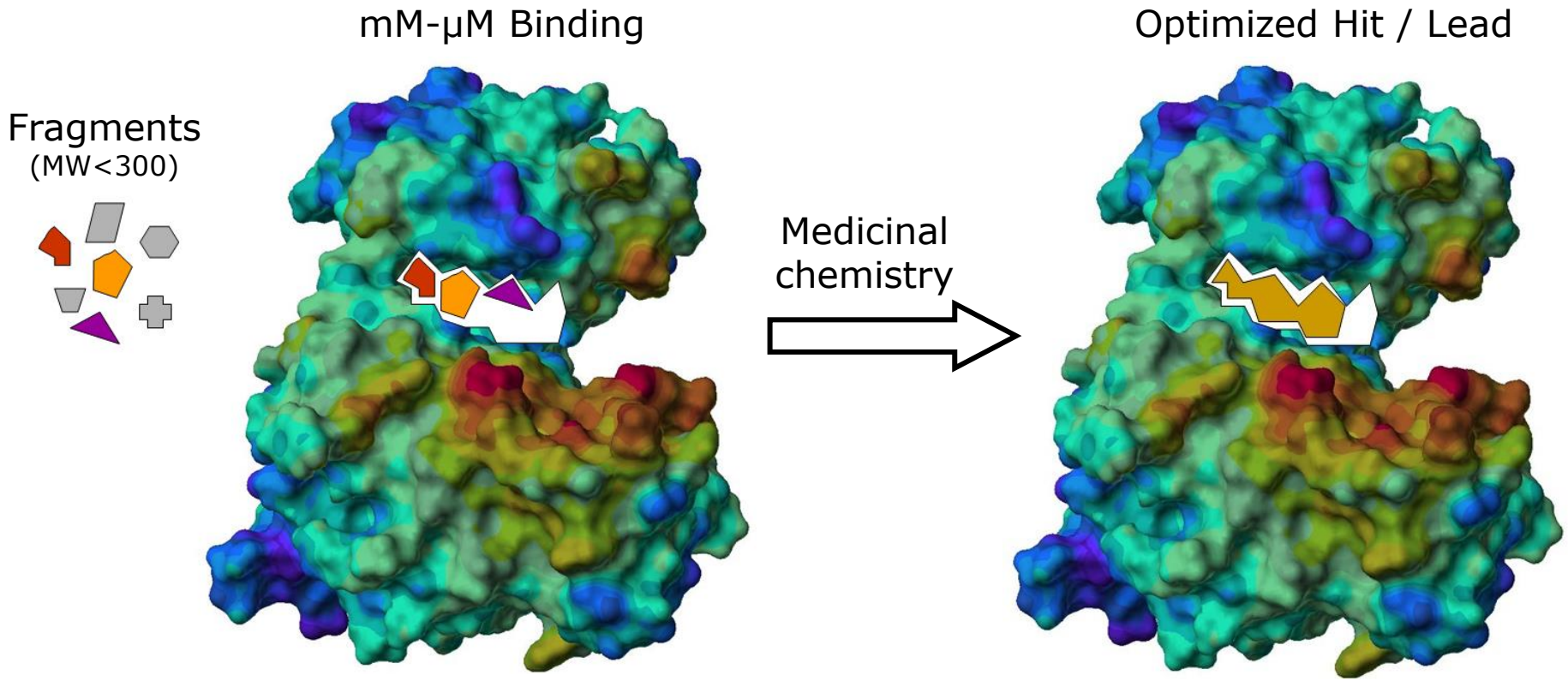




Collection of Drug Discovery Examples

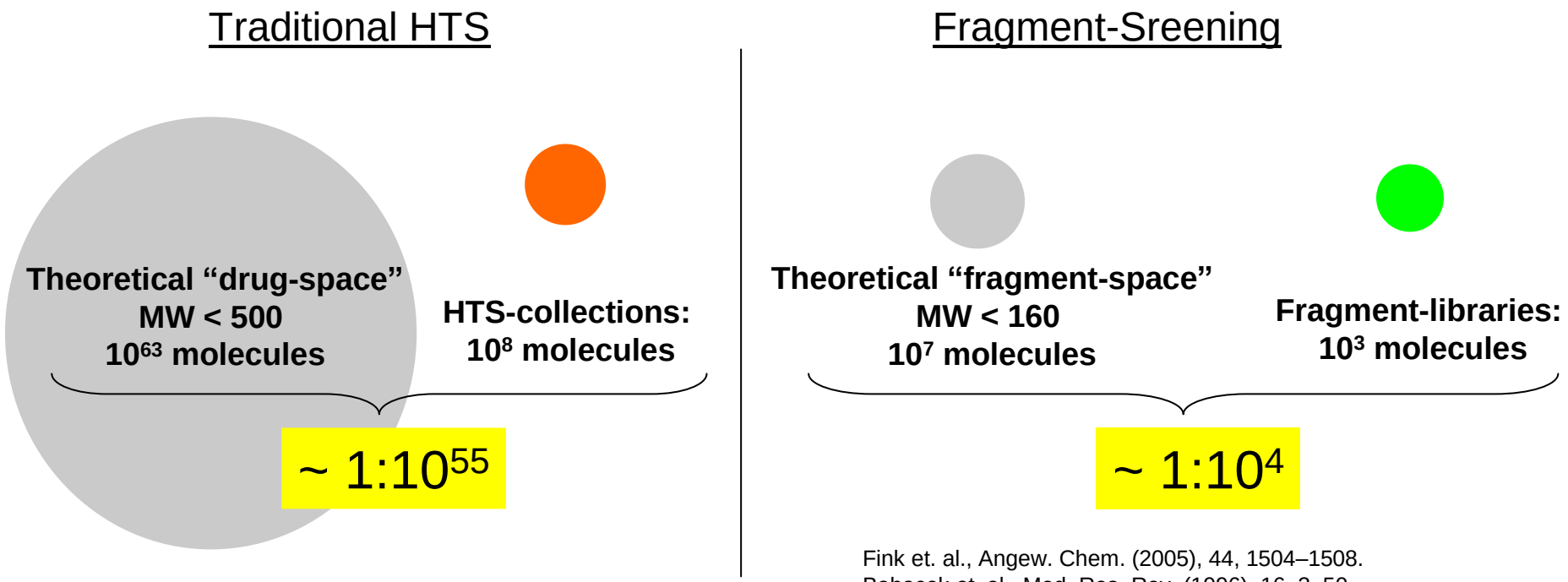
- Fragment Screening and Crystallization
- Tyrosine Spleen Kinase to look for Water Molecules
- Fexibility assessed in the structural model HSP90

Fragment-based Lead Discovery (FBLD) Concept



Low complexity fragments with higher probability to match binding site

Diversity of Fragments



Fink et. al., Angew. Chem. (2005), 44, 1504–1508.
Bohacek et. al., Med. Res. Rev. (1996), 16, 3–50.

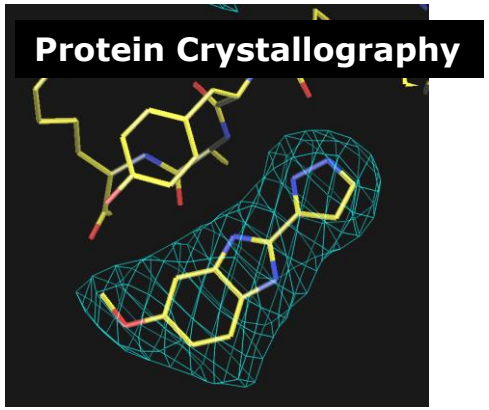
Fragment libraries sample higher chemical diversity

=> Significantly fewer compounds to be screened at higher concentrations

MERCK

Fragement Based Lead Discovery-Process

Experimental screening of 2000+ fragments using SPR & NMR



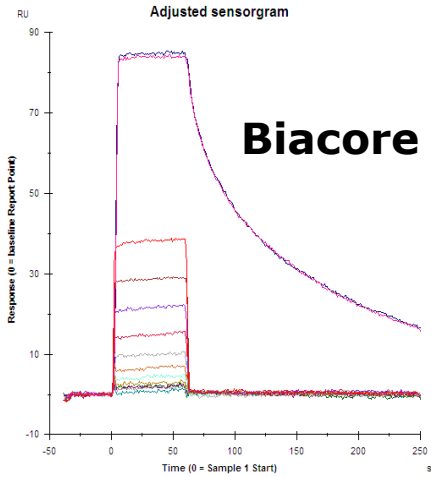
confirmed Hits

Crystallographic soaking

Co-Crystal Structure

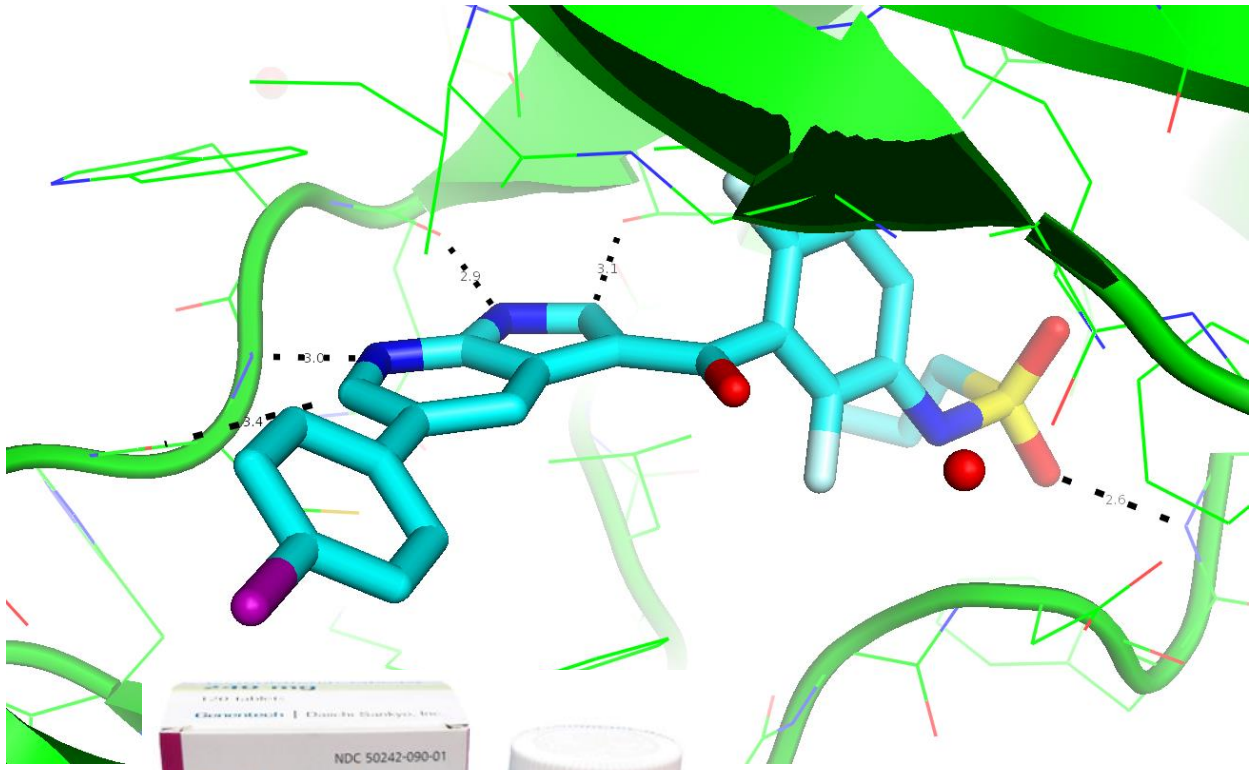
Selection of close analogues (information on SAR)

Chemical and Chemoinformatic Analysis



Information on Ligand/Target interaction gives opportunities for chemical optimisation

First approved drug from Fragment screening



FDA approval 2011

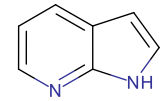
metastatic melanoma

January 2006
PLX4032
(Zelboraf)

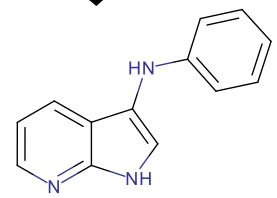
February 2005
(Plexxikon)

B-Raf V600E

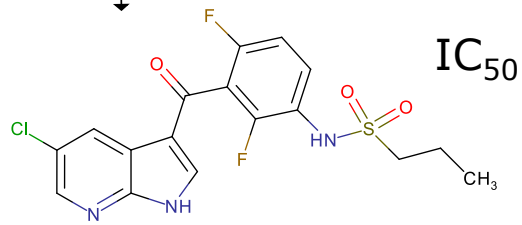
IC_{50}>200 μM



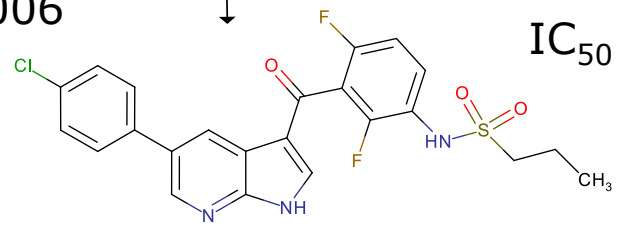
IC_{50}~100 μM



IC_{50}=0.013 μM



IC_{50}=0.031 μM



Previous results of in-house FBLD: Example in TGF-beta

MSRDB

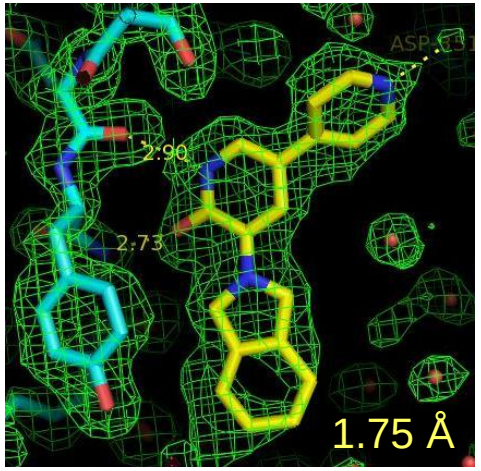
Selection by docking

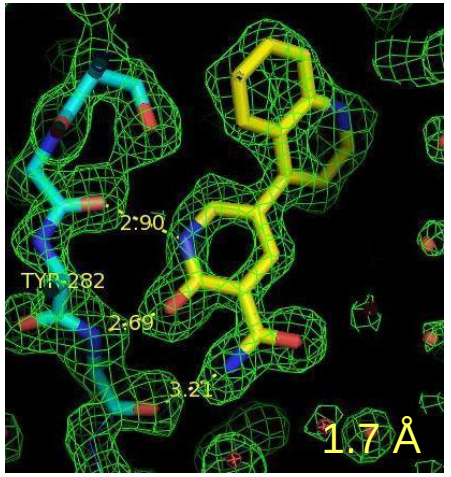
164 fragments

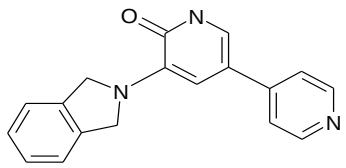
Biochemical assay @ 100 µM
(CHK1, PDK1, TGF-beta)

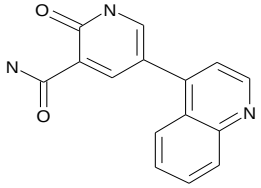
27 Hits
(IC_{50}<100µM)

X-ray structures of fragments in TGF-beta





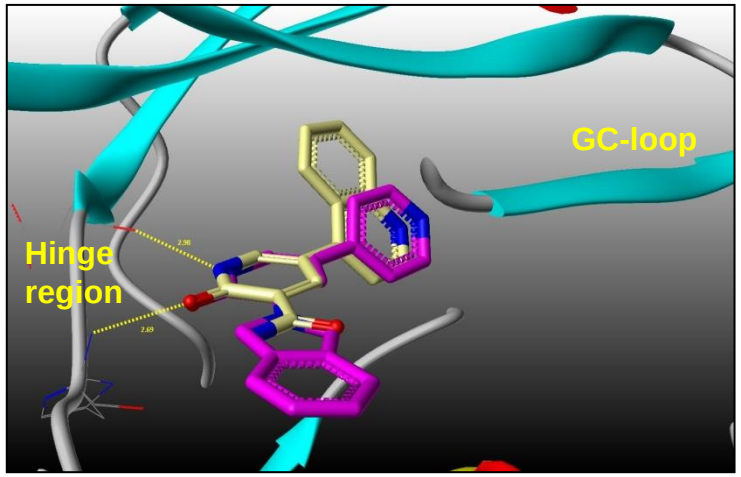


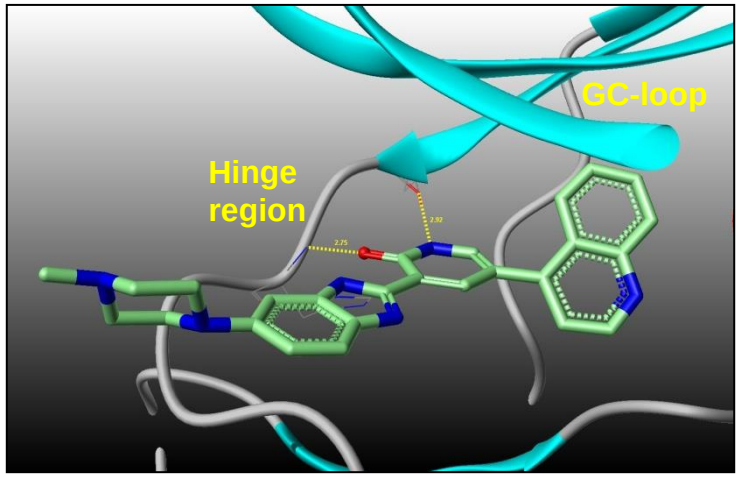


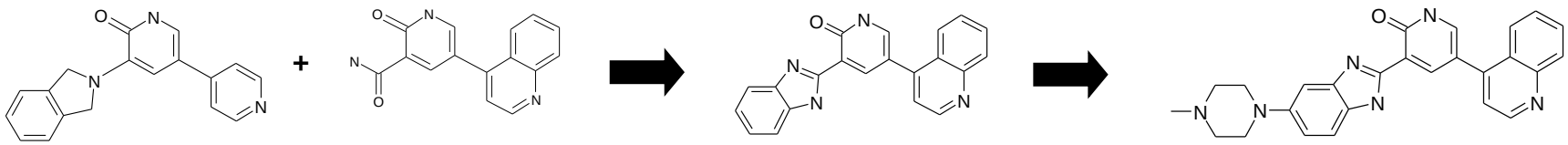
MSC1043364A
TGF-beta: IC_{50}=49 µM
CHK1: IC_{50}=6.7µM
PDK1: IC_{50}=35 µM

MSC1044909A
TGF-beta: IC_{50}=48µM
CHK1: 92%cntrl@100µM
PDK1: 86%cntrl@100µM

Optimization of fragments for TGF-beta





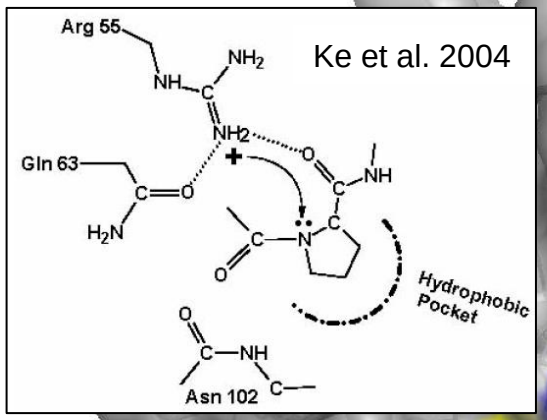


IC_{50}= 49 µM + IC_{50}= 48 µM → IC_{50}= 370 nM → IC_{50}= 37 nM

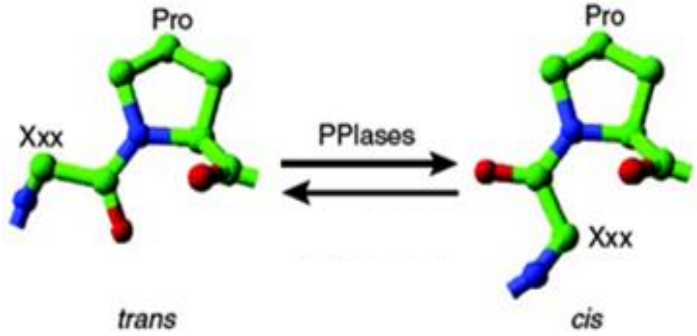
Structure-based optimization

- Favourable IP-situation
- Promising selectivity profile
- Weak cellular activity

Cyclophilin D enzymatic mechanism



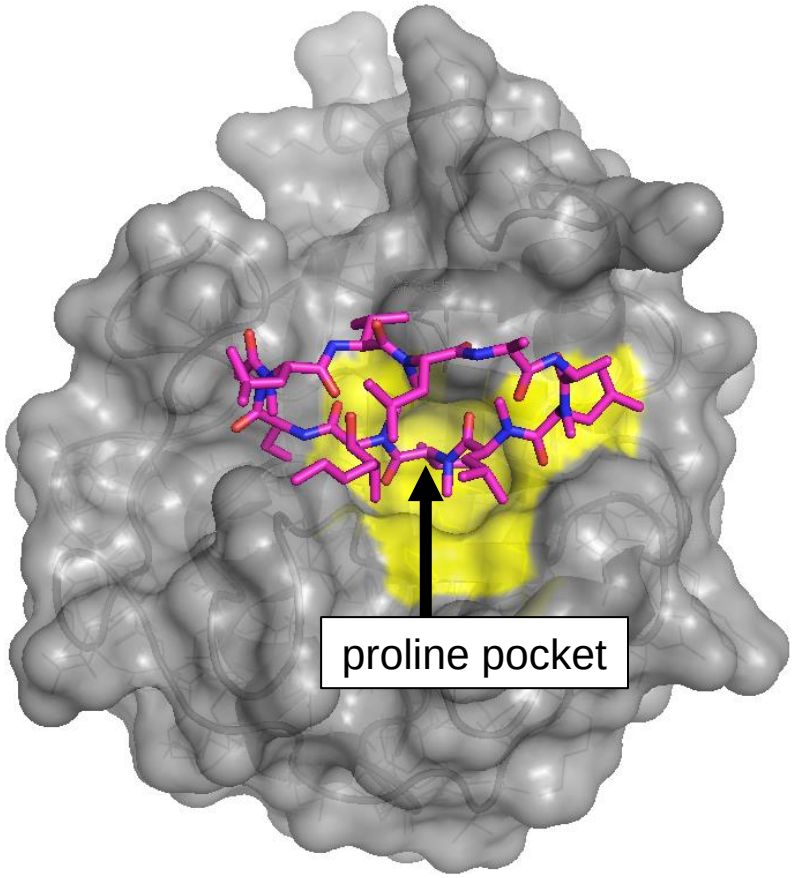
Ke et al. 2004



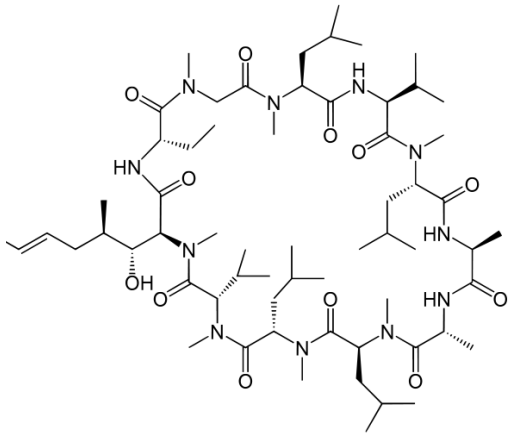
- Cyclophilins catalyze cis-trans isomerization of prolines
- Proline bound in small lipophilic pocket
- Conserved arginine involved in isomerization mechanism

Cyclophilin D inhibiton by Ciclosporin A

Crystal structure of Cyclophilin D in complex with Ciclosporin A



Kajitani et al. 2008



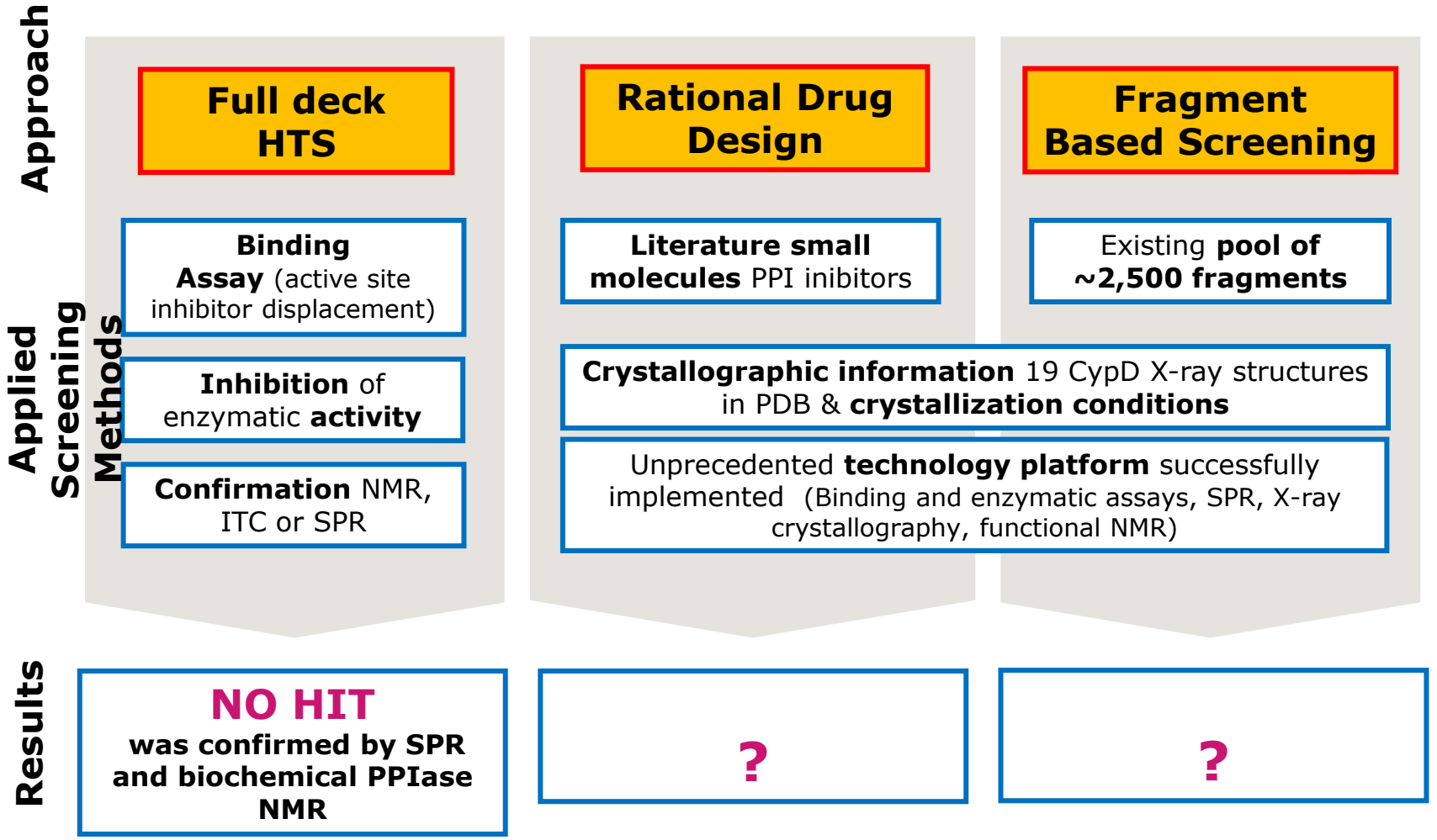
Ciclosporin A

SPR : K_D = 30 nM

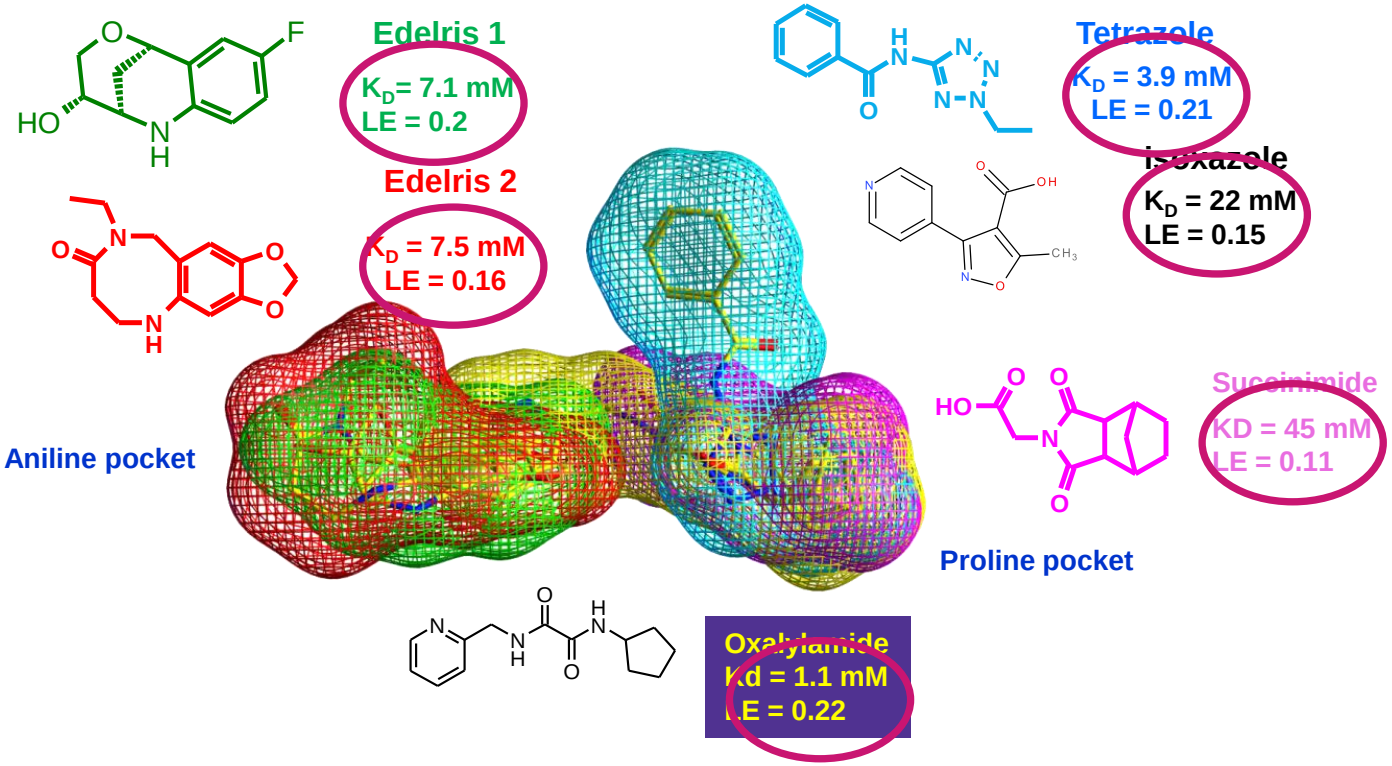
Binding : IC_{50} = 20 nM

- Immunosuppressant drug Ciclosporin A (CsA) inhibits cyclophilin family members
- CsA binding encompasses large contact area (~150 Å^2) including proline pocket
- CsA binding site highly conserved among cyclophilins

Hit Discovery Strategy : Three Tracks in Parallel



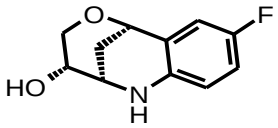
Solved PPIase Fragment Structures Occupy optimal Space for Growíng and Linking Approaches



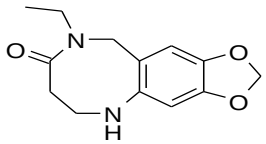
Cyclophilin D: in-house FBLD application

From millimolar fragments to 3 nanomolar series

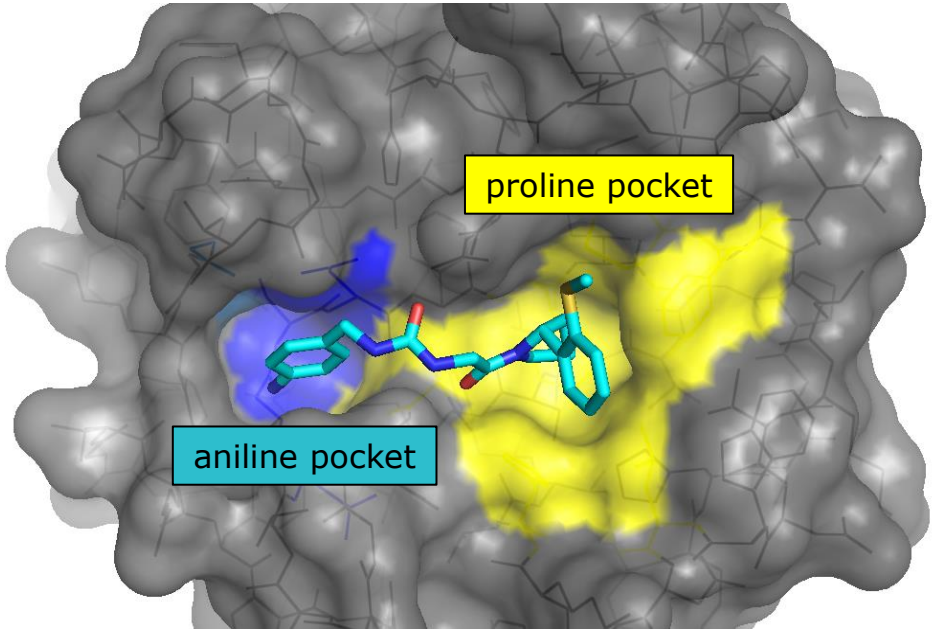
fragments in aniline pocket



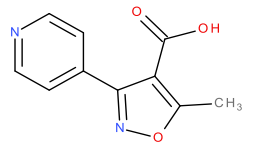
K_D= 7.1 mM
LE = 0.2



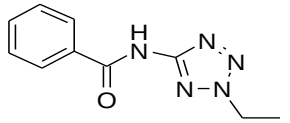
K_D = 7.5 mM
LE = 0.16



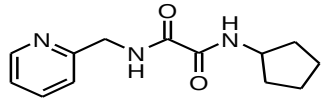
fragments in proline pocket



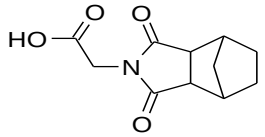
K_D = 22 mM
LE = 0.15



K_D = 3.9 mM
LE = 0.21



K_D = 1.1 mM
LE = 0.22

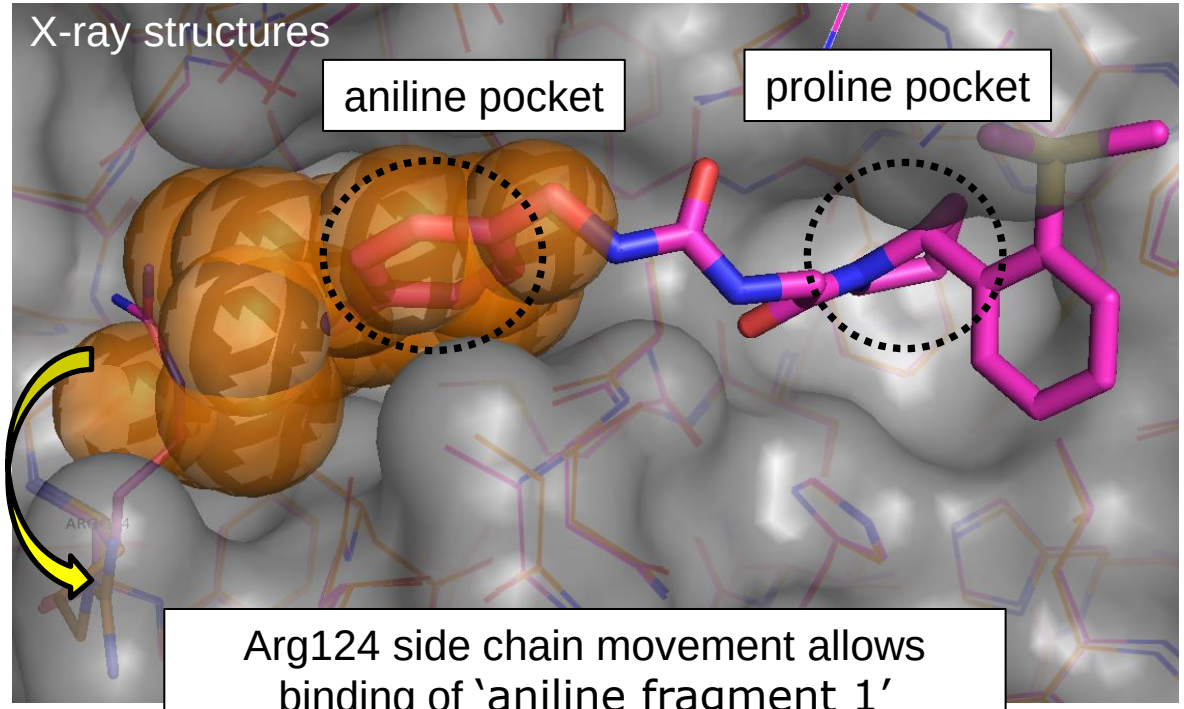


K_D = 45 mM
LE = 0.11

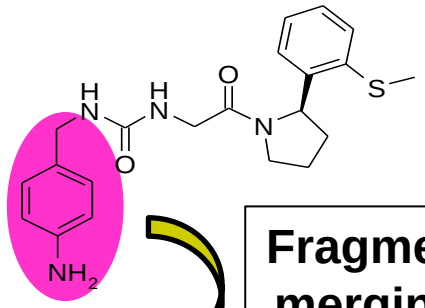
- HTS yielded no confirmed hit
- Fragment-screen by SPR identified 58 hits, but only with millimolar potencies (KDs 1-58 mM)
- X-ray structures of 6 fragments served as basis for merging and linking to yield 3 nanomolar series

The Power of Biophysical Methods:

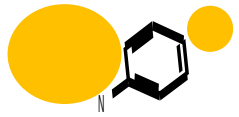
Fragment merging with ‚Aniline fragment 1‘



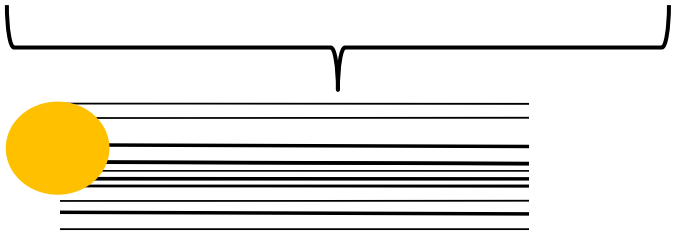
CypD reference inhibitor
SPR : K_D = 204 nM
Binding : IC_{50} = 520 nM



Fragment merging



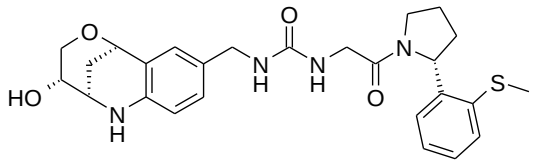
‘Aniline fragment 1’
K_D= 7.1 mM
LE = 0.2



SPR : K_D = 7.7 nM
Binding : IC_{50} = 3.7 nM

- 26-fold potency enhancement (K_D = 204 nM -> 7.7 nM) achieved through replacement of ‘simple’ aniline by ‘aniline fragment 1’

Medicinal Chemistry – Series 1
General profile

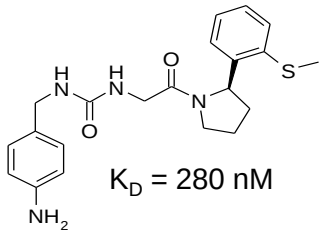


CsA	PhysChem/ *in vitro* Pharmacology		DMPK / Safety	
1202/ 3.8	MW / logP	496 / 1.8	Stability in PBS / SGF	Yes / No
235 / 5	tPSA / HBD	103 / 4	Stability in plasma (h, r, m)	Yes
	CNS-MPO score	3	u-somes stab. (h, m, r)	78/ 579 / 16
0.02	Binding (competition) IC_{50} uM	0.031	SPB (h, m, h)	og
0.03	Binding (SPR) KD uM	**0.0077**	Caco-2 (Papp / E)	0.02 / 564
0.04	PPI ase activity IC_{50} uM	**0.0037**	MDR1/MDCK	0.23 / 5
Yes	PPIase activiy (NMR)	Yes	CYP inhibition	> 10 uM
NA	Kinetic Solubility PBS uM	> 200	HTS-MNT	pNegative
EC_{50} = 3.5 uM	Rat primary cortical neuron viability assay	40% effect @ 100 uM	hERG (binding)	> 30 uM

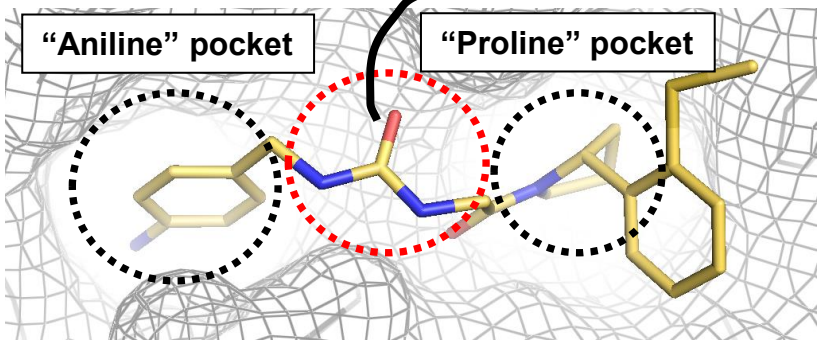
- ~100 molecules synthesized, **nM potency** achieved
- Optimisation of **permeability** needed by reducing Mw/ PSA and HBD, no safety flag

6 CypD fragment X-ray structures

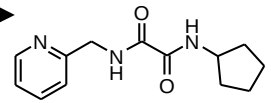
Reference inhibitor



K_D = 280 nM

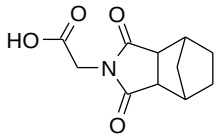


Oxalylamide



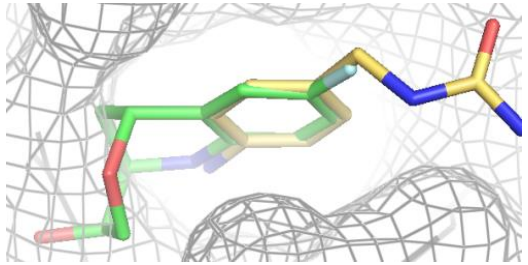
K_D = 1.1 mM
LE = 0.22

Succinimide

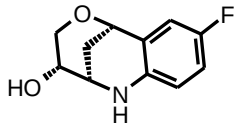


K_D = 45 mM
LE = 0.11

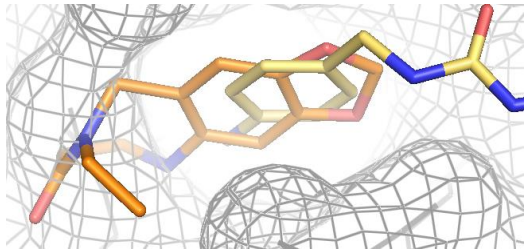
Fragments binding in **"Aniline" pocket**



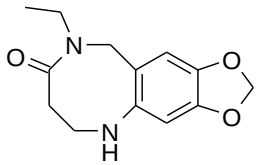
"Aniline fragment" 1



K_D= 7.1 mM
LE = 0.2

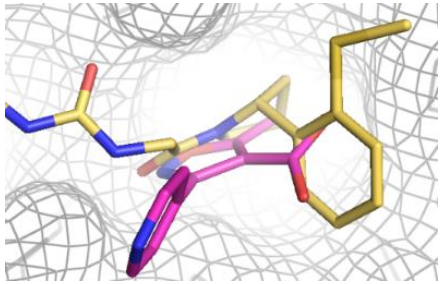


"Aniline fragment" 2

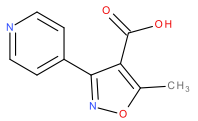


K_D = 7.5 mM
LE = 0.16

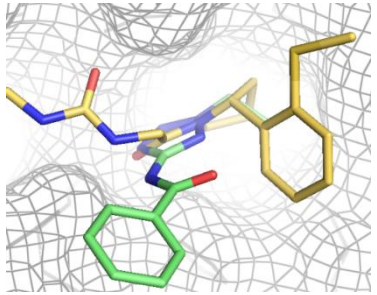
Fragments binding in **"Proline" pocket**



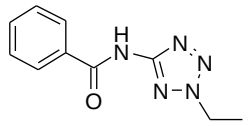
Isoxazoles



K_D = 22 mM
LE = 0.15



Tetrazoles



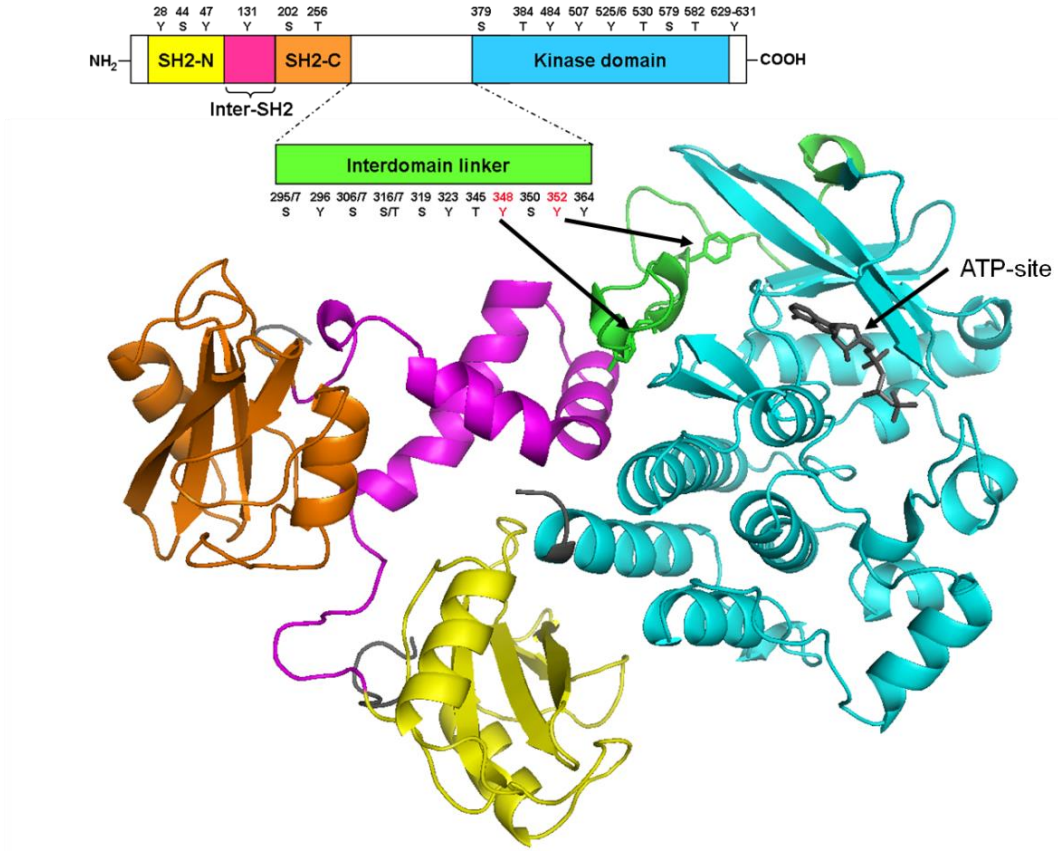
K_D = 3.9 mM
LE = 0.21

Protein Crystallography: Spleen Tyrosine Kinase SYK

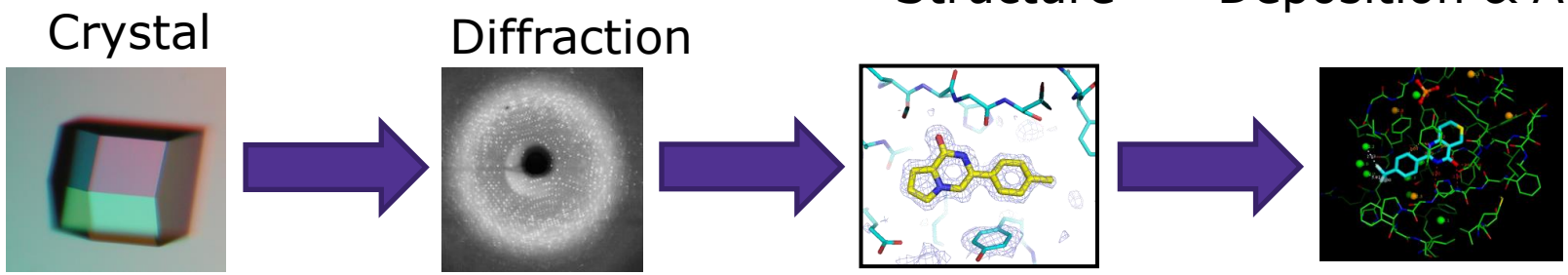
Domain-model and proposed phosphorylation sites of SYK

→Y348 and Y352 play a crucial role

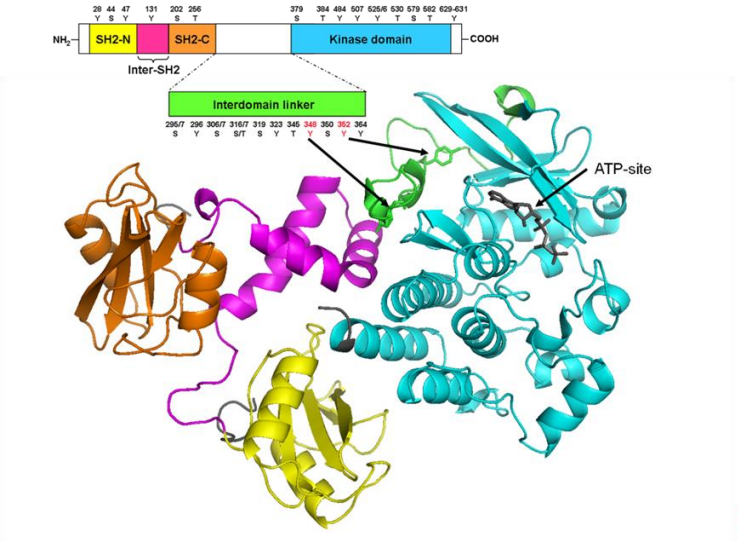
→in the activation mechanism of SYK

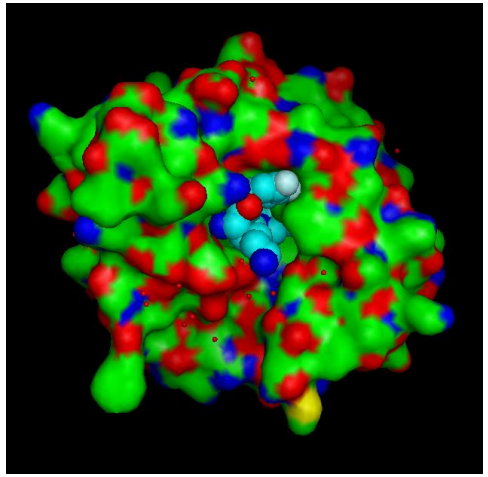


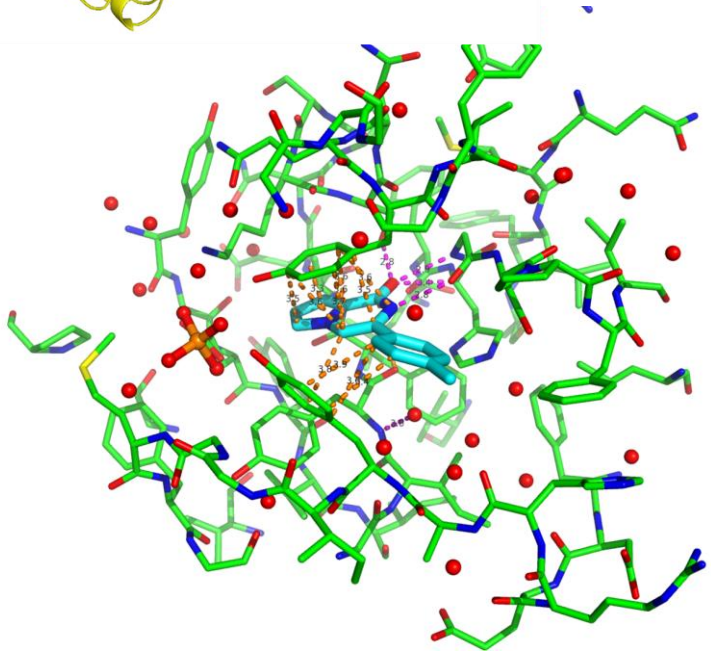
Structure Deposition & Analysis

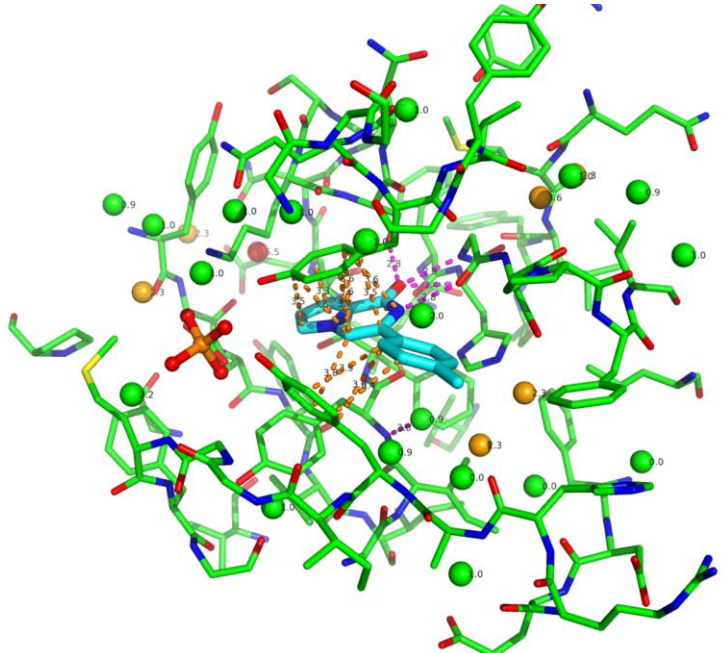


„Classification of Water Molecules “



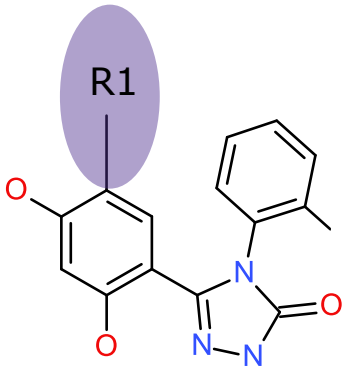


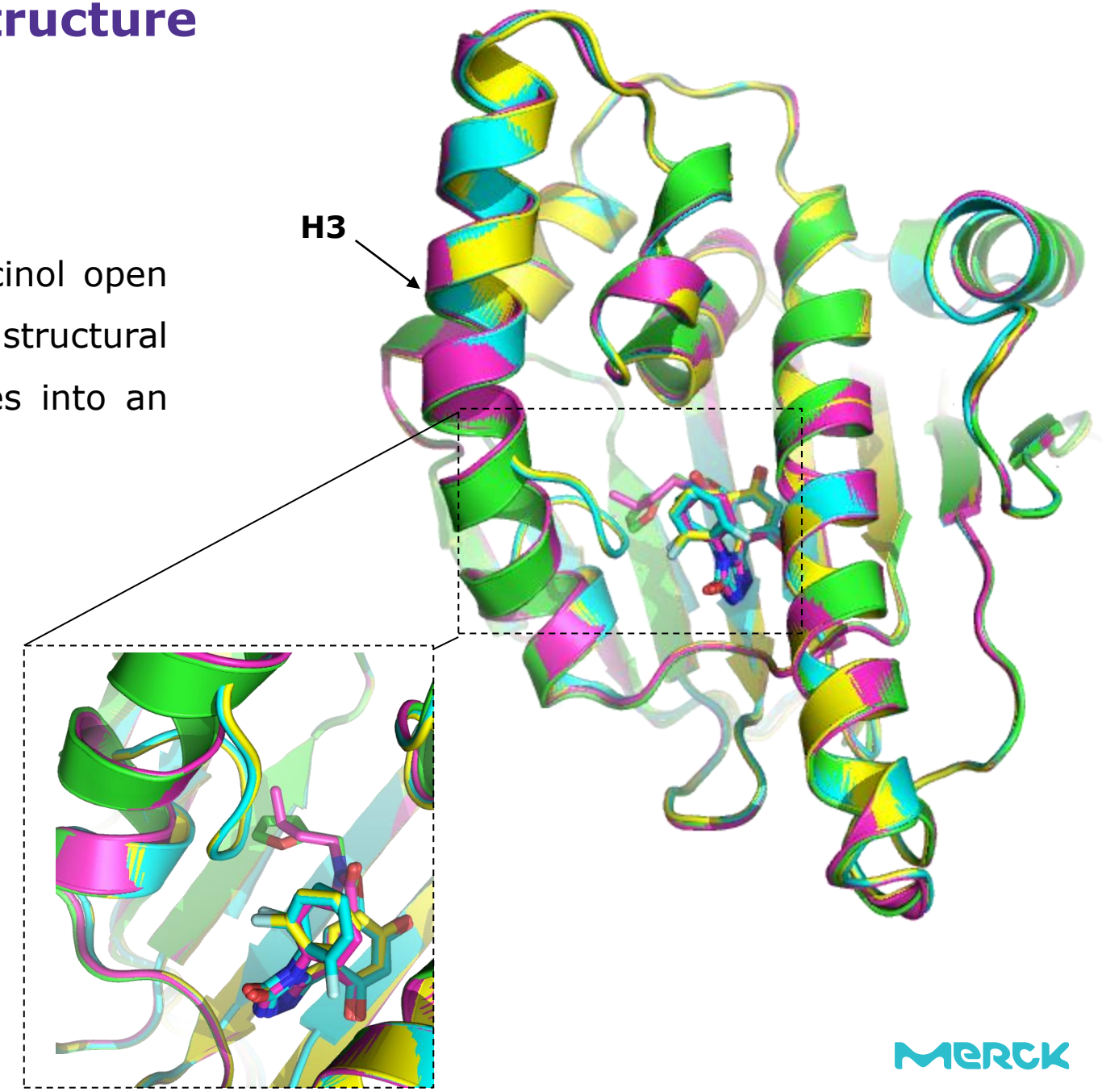




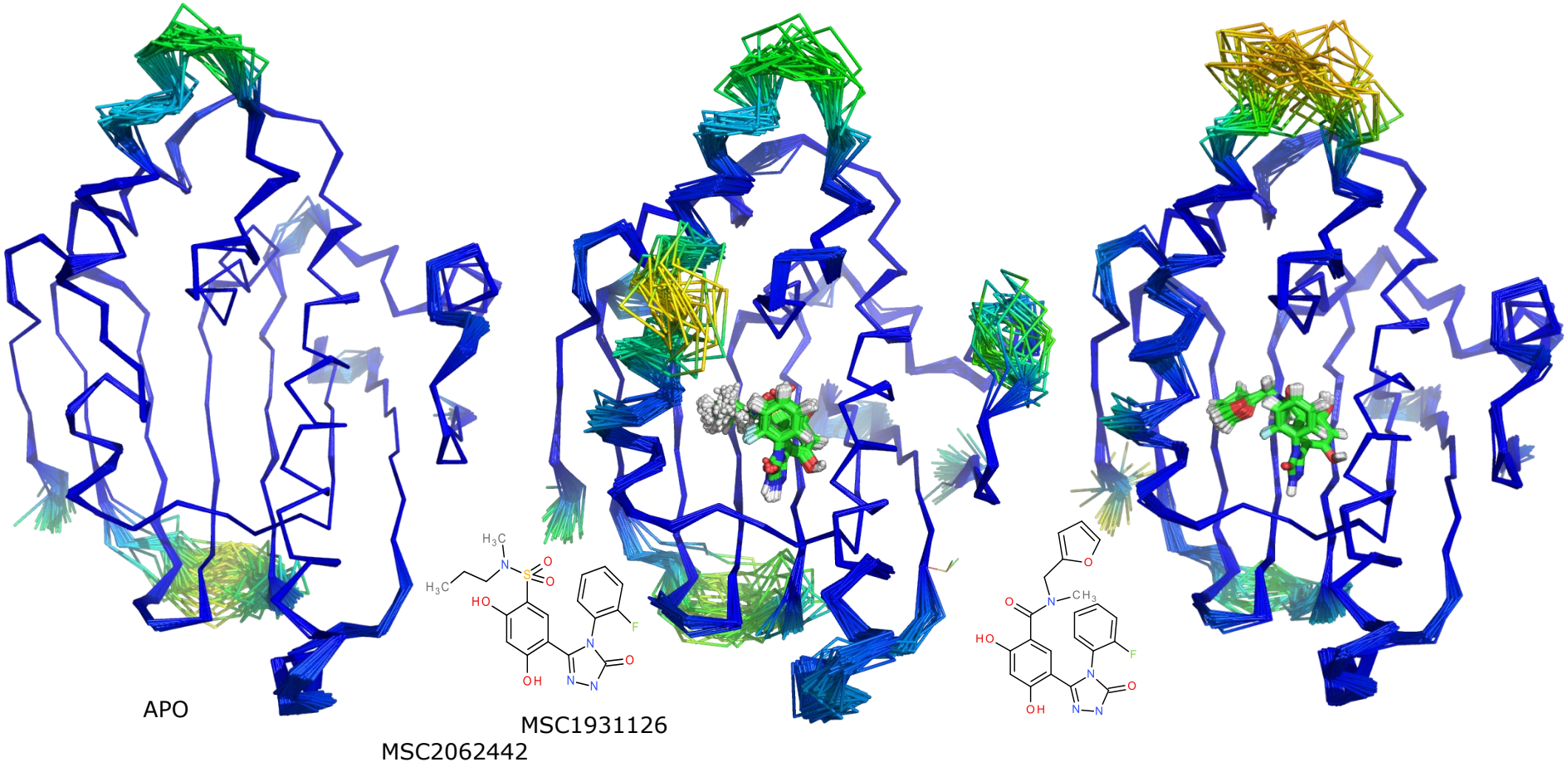
Structural analysis of HSP90 structure complexes

Substituents in the 5’ position of the resorcinol open up a large lipophilic pocket by enforcing a structural change so that segment 103-110 rearranges into an helical conformation





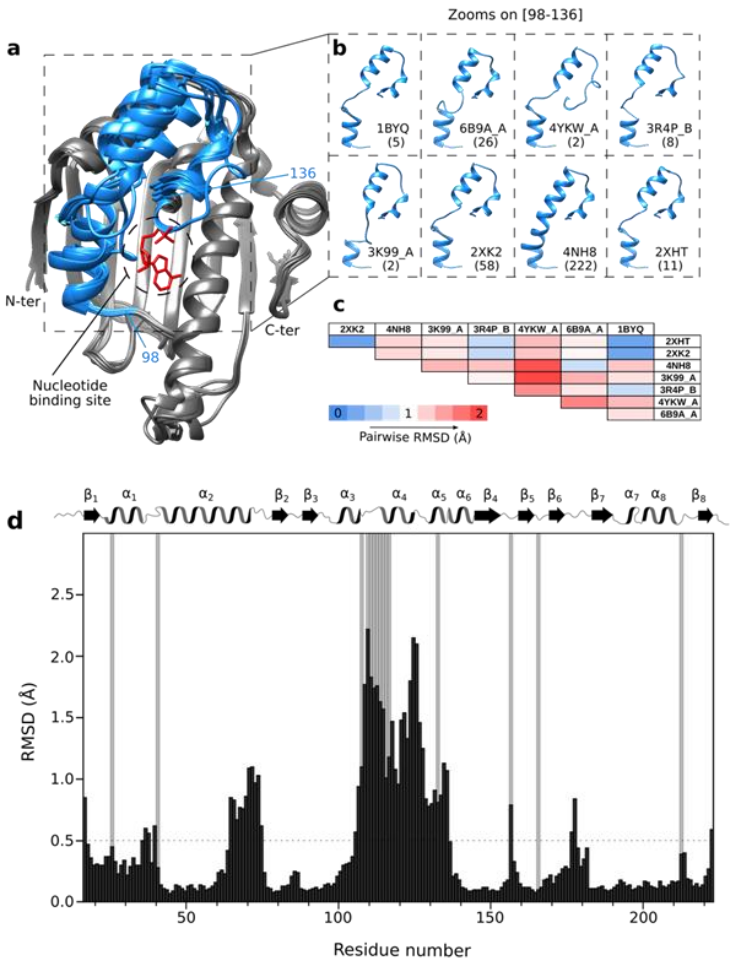
Modelling Dynamics in Protein Crystal Structures by Ensemble Refinement



- Residues 103-110 in the Apo structure are significantly more stable when compared to complex strucures that induce helix formation.

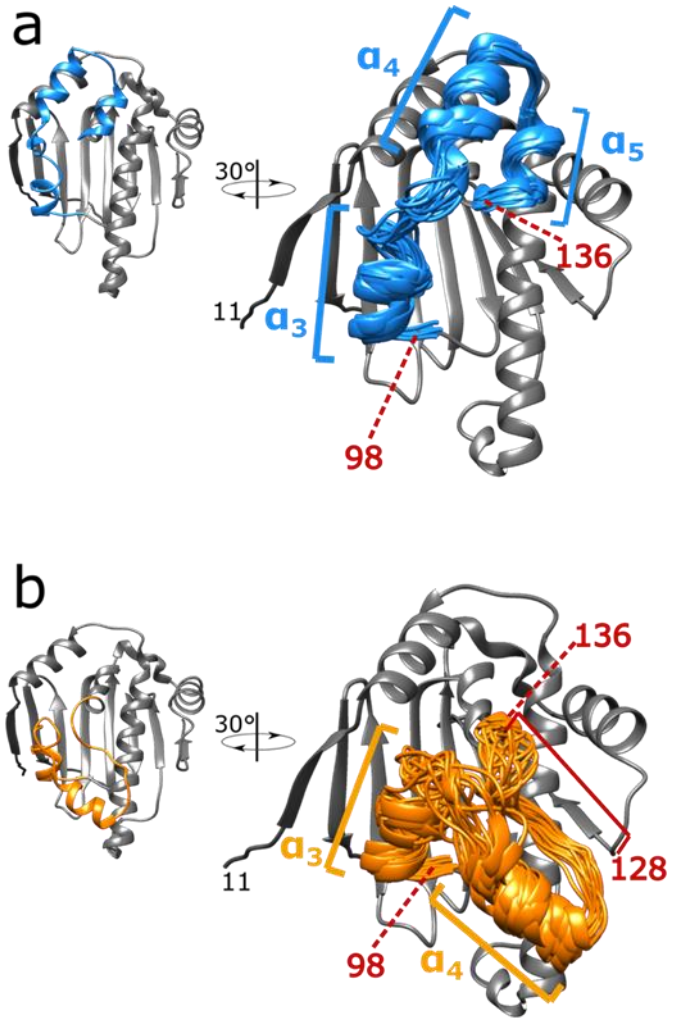
- Helix stability is affected by the chemistry of the ligand.

Analysis of available X-ray structures of human HSP90 N-terminal domain.



Superimposition of 8 centroids representing the 8 clusters describing the 334 structures of isolated HSP90-NTD available in the Protein Data Bank (on January the 5th of 2021).

NMR Solution structure ensembles of HSP90α-NTD ATP-lid open and closed states.



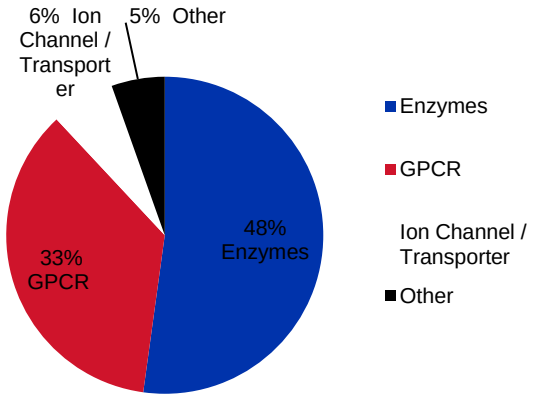
The HSP90 ATP-lid populates a closed conformation distant by up to 30 Å from the previously known open ATP-lid state.

Both structures could be solved and kinetic and thermodynamic parameters could be derived.

Biochemical mechanism of new molecular entities approved during 2001 and 2004

During 2001 and 2004 the FDA approved 85 new molecular entities.

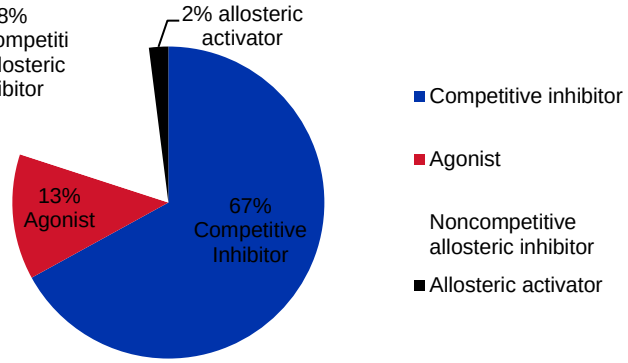
Majority targeted enzymes (48%) or GPCRs (33%),



80% interacted with the same site as endogenous effectors

Three biochemical operations defined the modes of action

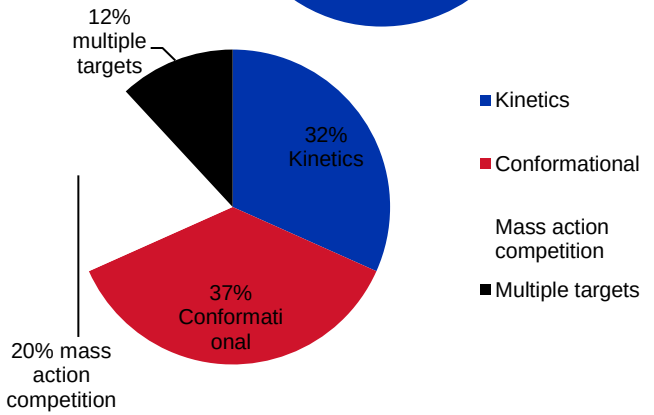
1) Mass action competition
2) Drug stabilized conformational change important to the response
3) Drug action is less-responsive to mass action competition with effectors due to non-equilibrium kinetics



About 80% of the NMEs elicit a response

Utilizing conformational

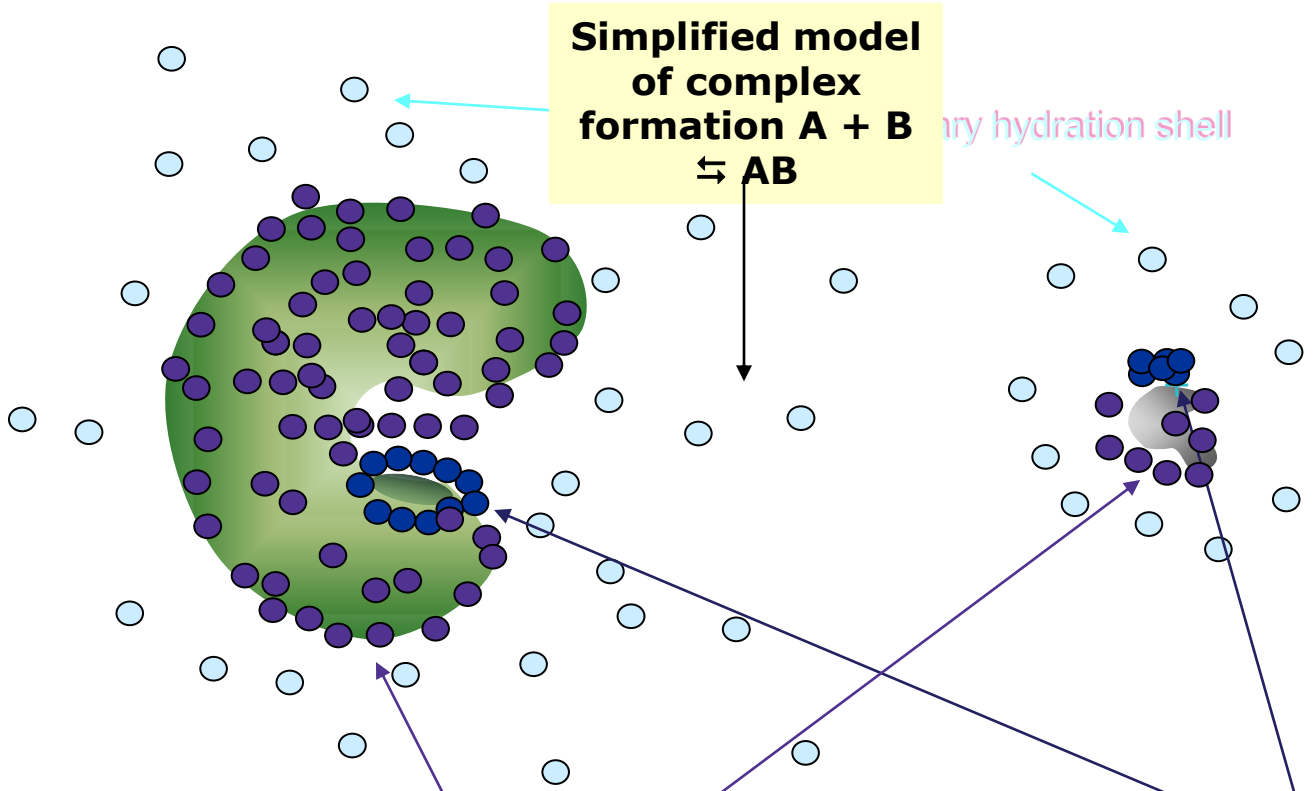
and / or non-equilibrium kinetics mechanisms



Swinney Current Topics in Med. Chem. 2006, 6, 461-478, more recent Swinney focus on mechanism of new drugs

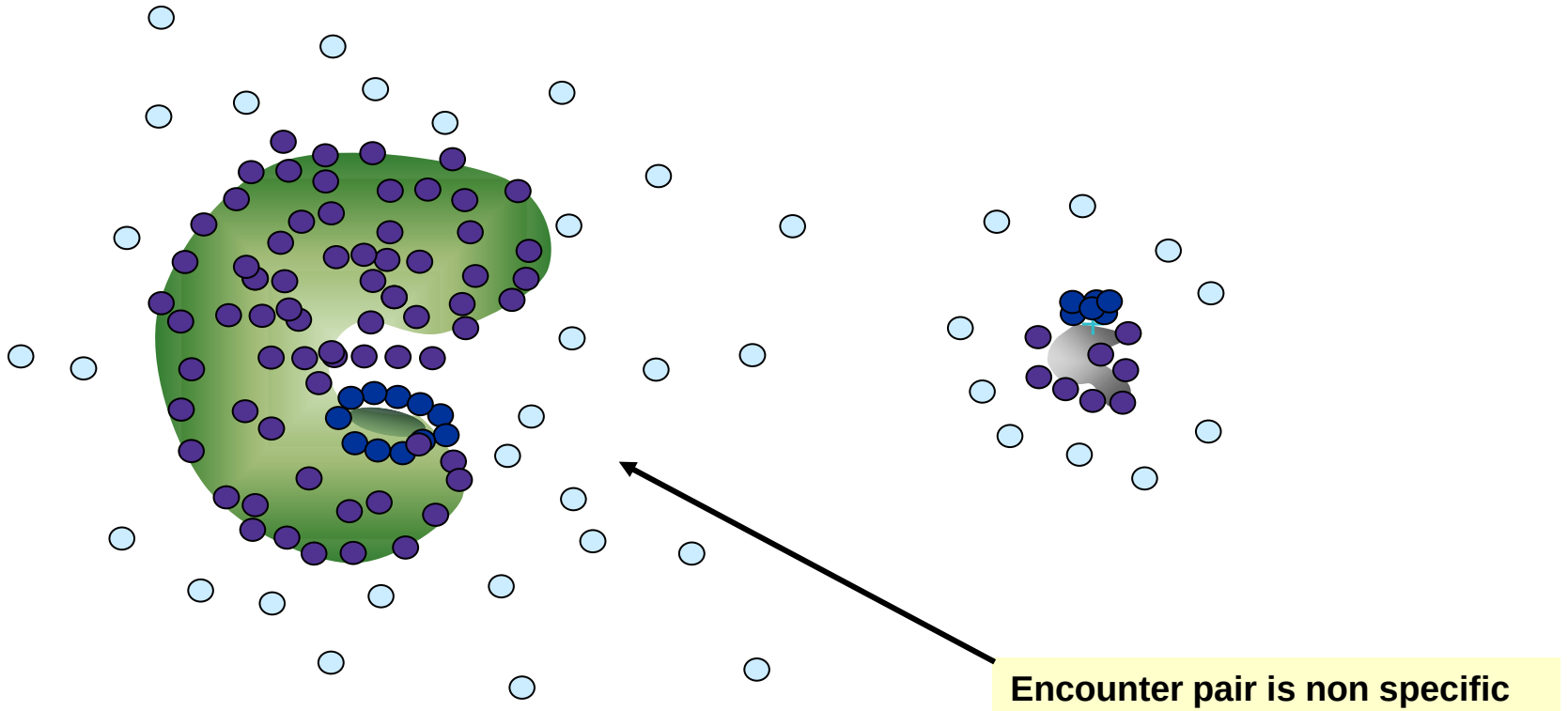
MERCK

…binding is not only a two players game



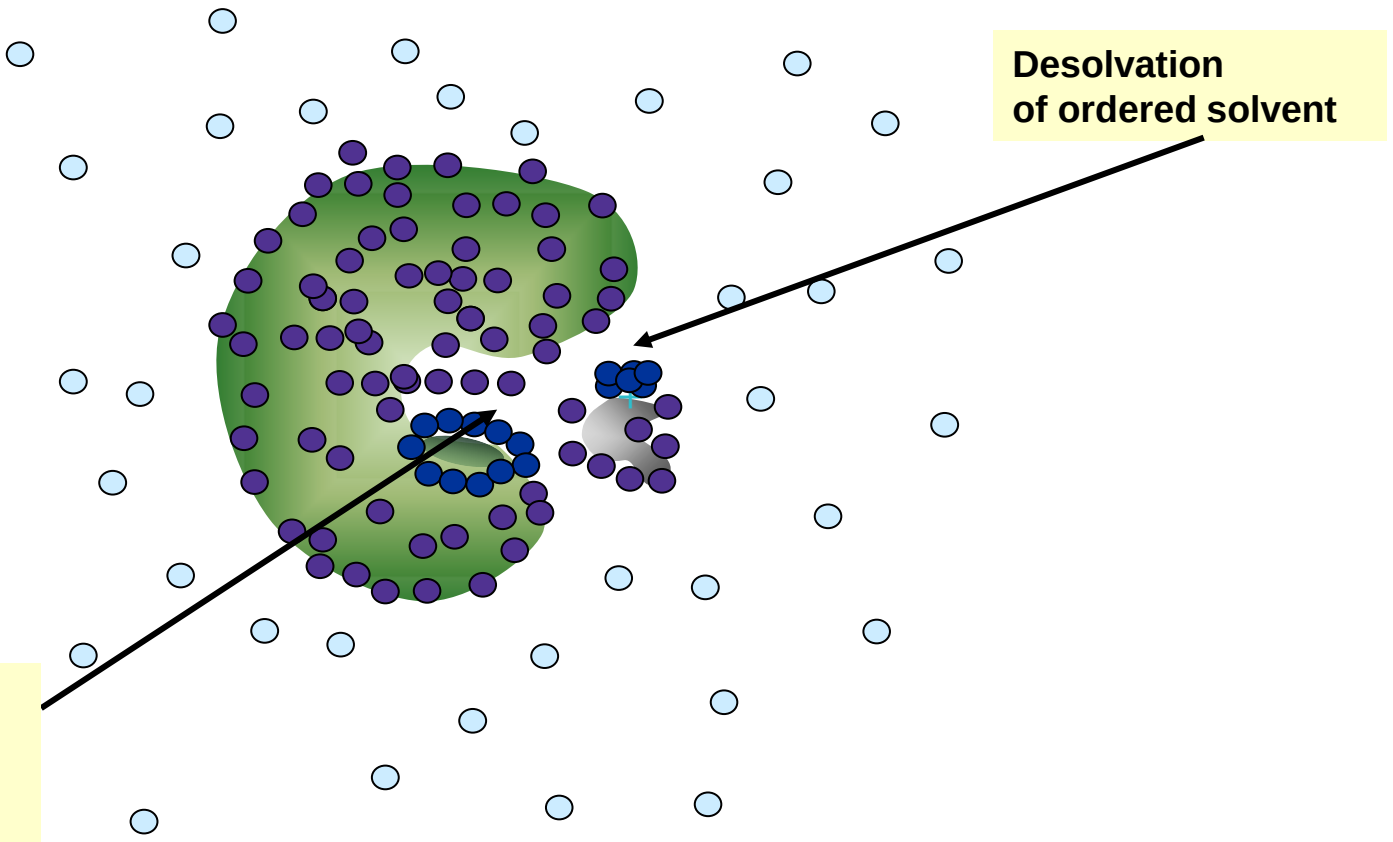
$$A + B \leftrightarrows A\cdots B \leftrightarrows A—B \leftrightarrows AB$$

…it involves numerous molecules



$$A + B \leftrightarrows A\cdots B \leftrightarrows A{-}B \leftrightarrows AB$$

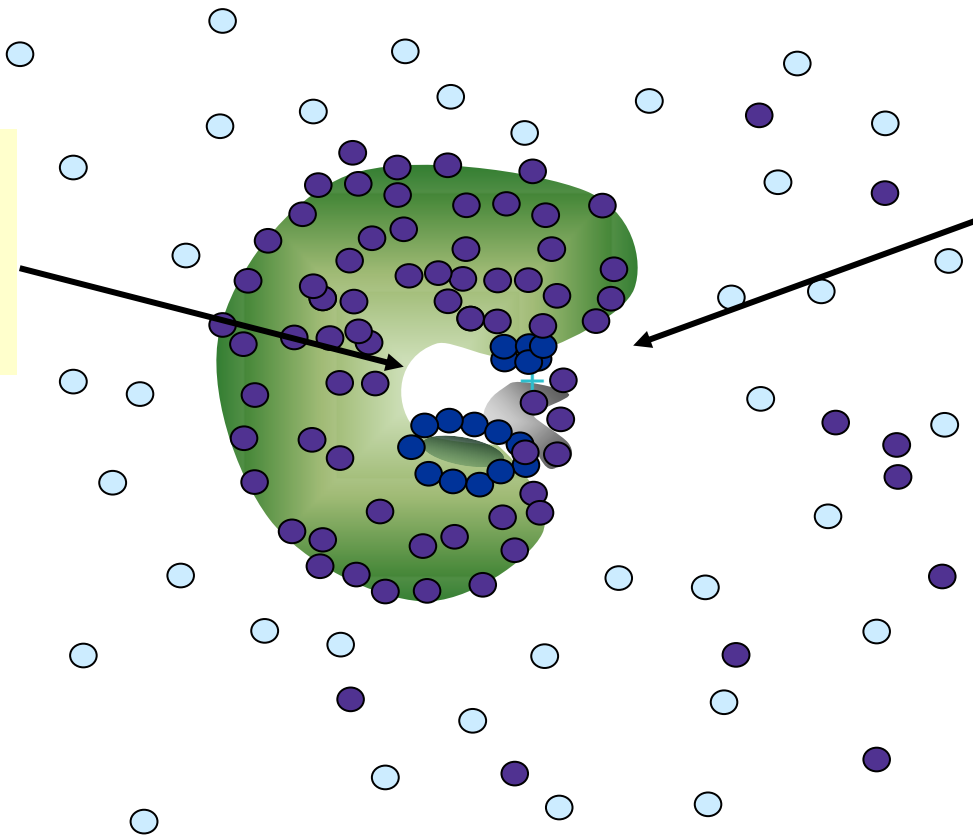
....and is a very complex multi step process



$$A + B \leftrightarrows A\cdots B \leftrightarrows A{-}B \leftrightarrows AB$$

MERCK

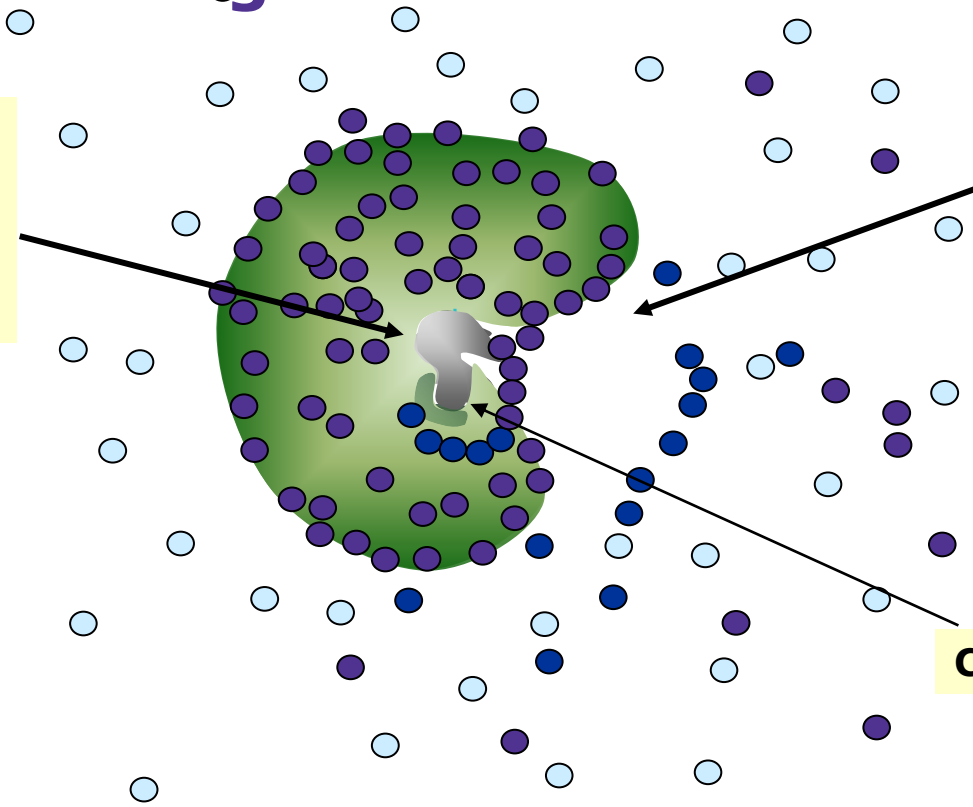
…docking of the ligand to form the complex



$$A + B \leftrightarrows A\cdots B \leftrightarrows A{-}B \leftrightarrows AB$$

MERCK

…eventually with induced fit or structural rearrangements



Formation of Ionic / hydrogen bonds v. d. Waals interaction

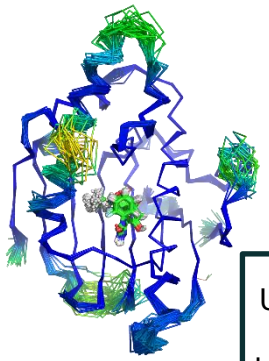
Full desolvation of residual solvent

Conformational change

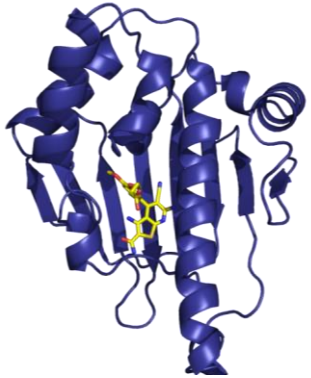
$$A + B \leftrightarrows A\cdots B \leftrightarrows A{-}B \leftrightarrows AB$$

MERCK

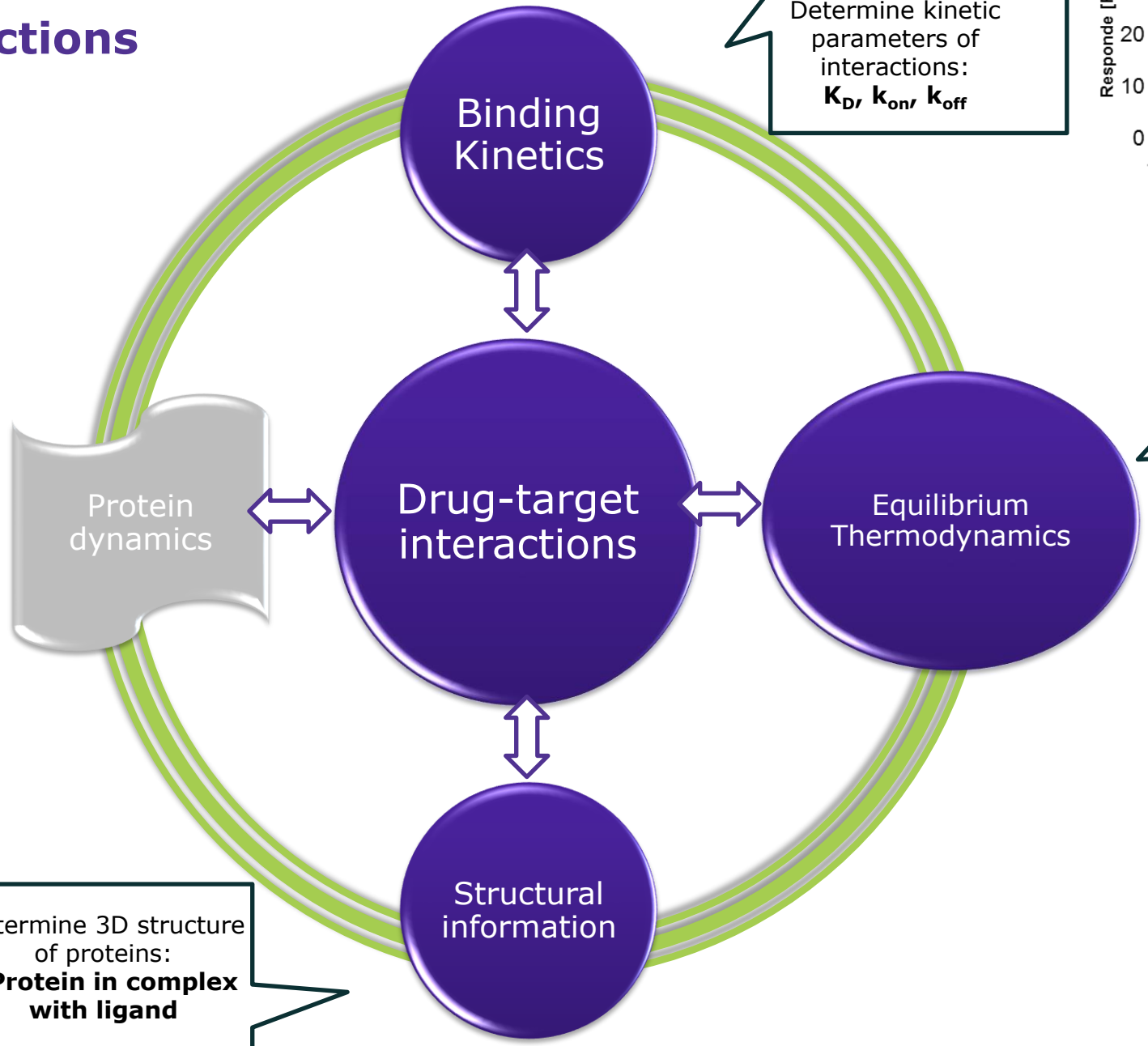
Drug-Target interactions



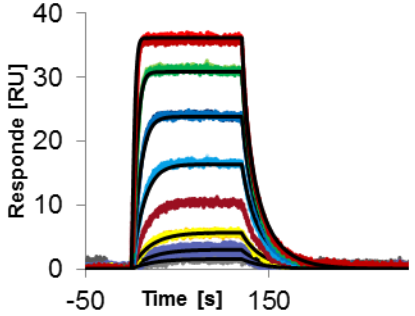
Understand the role of protein dynamics in drug target interactions: **NMR and MD**



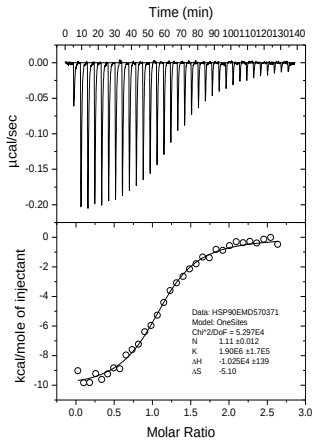
Determine 3D structure of proteins:
- **Protein in complex with ligand**



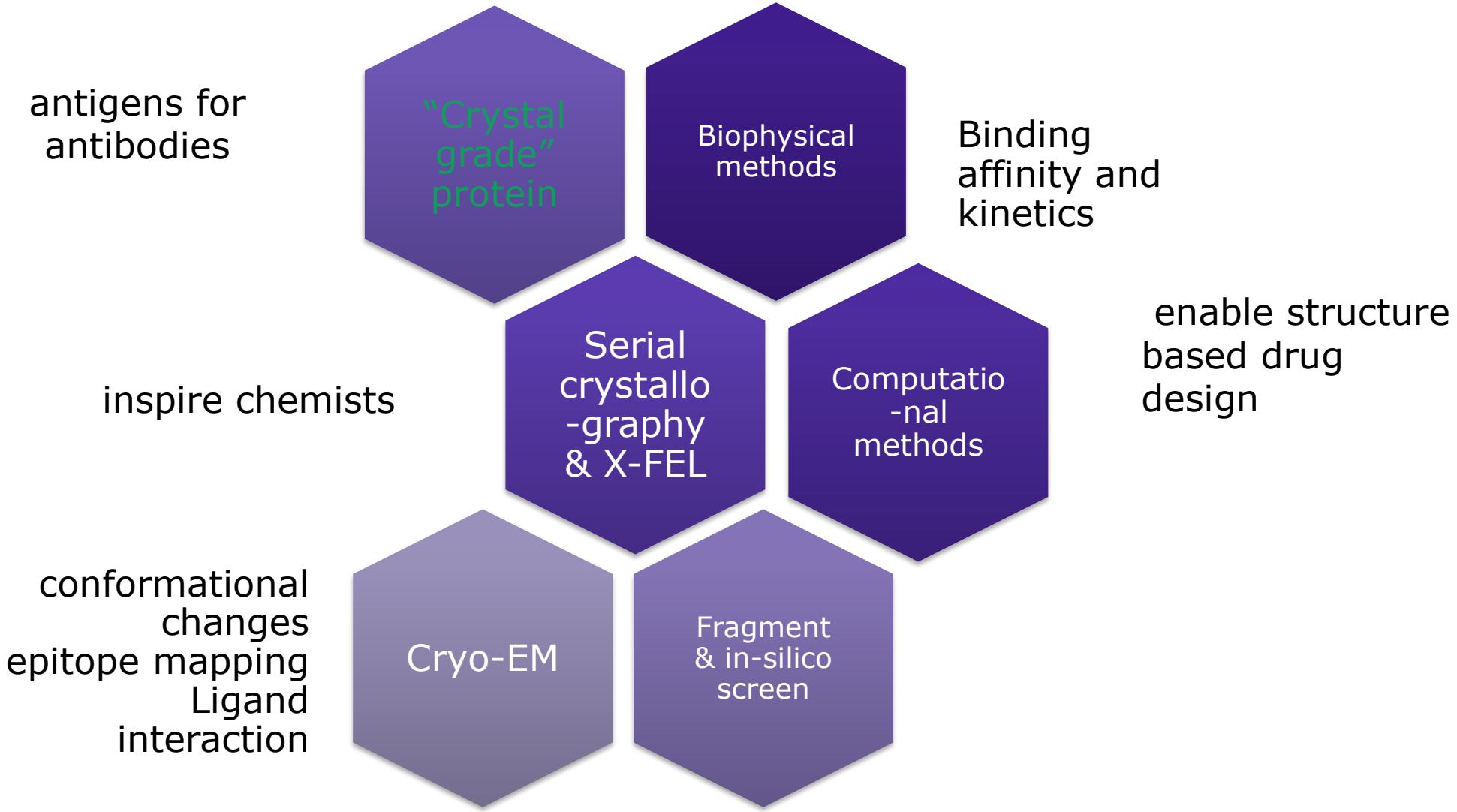
Determine kinetic parameters of interactions: **K_D, k_{on}, k_{off}**



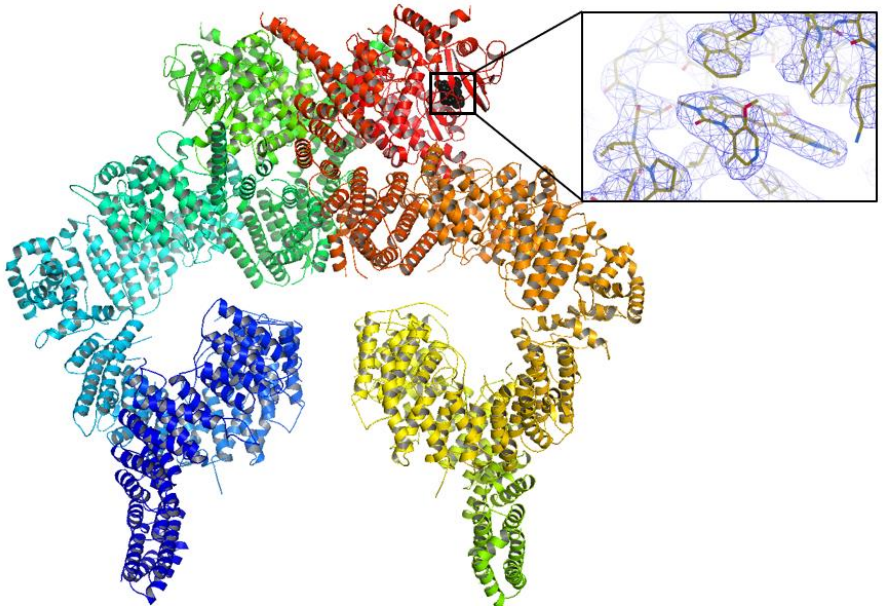
Determine thermodynamic parameters of interactions in solution: **K_D, ΔG, ΔH, ΔS**



The proper integration of the methods into the organization and projects enables an efficient drug discovery

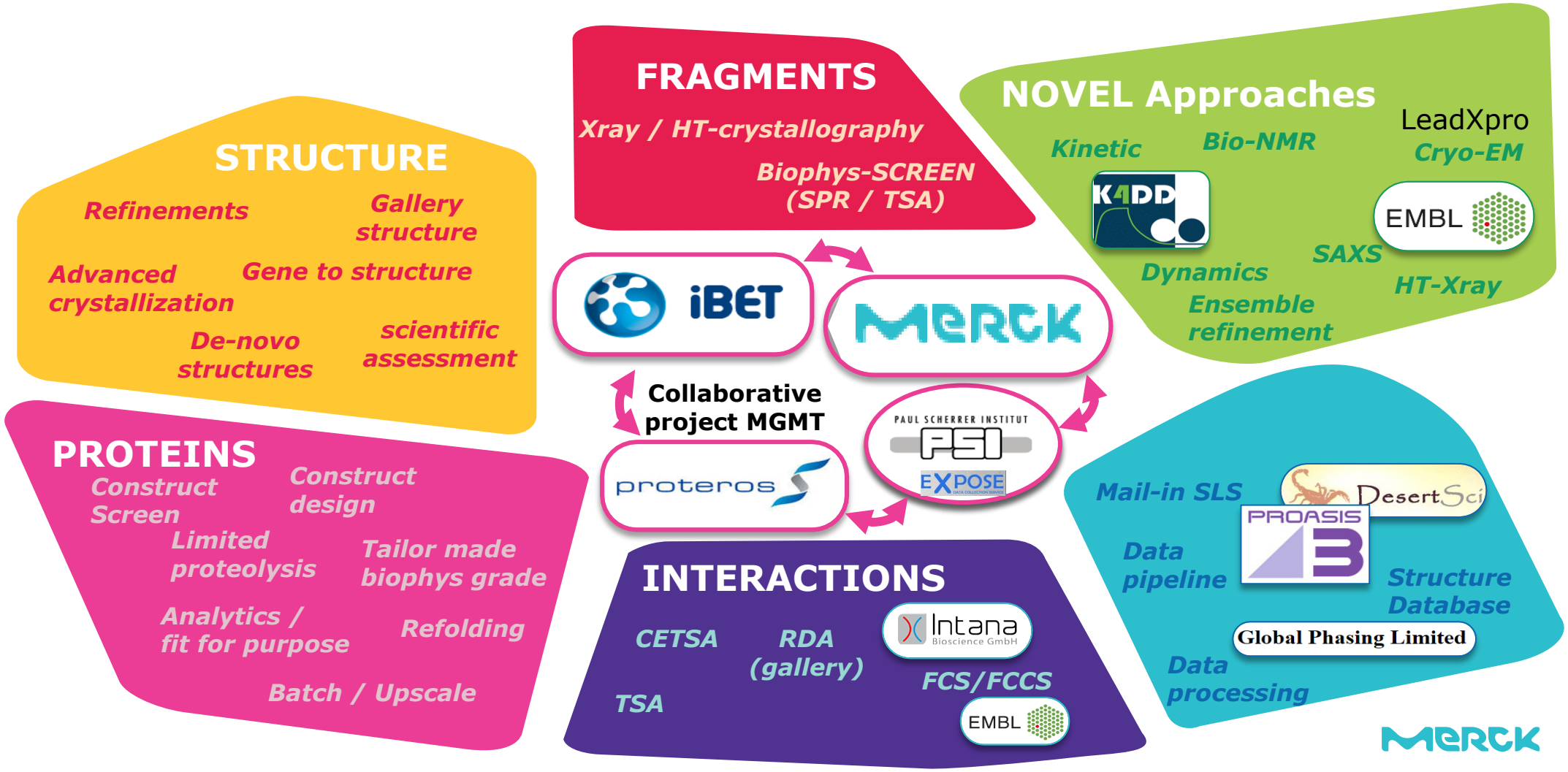


Cryo EM can help to analyze difficult targets and will be a highly valuable Methods to further advance our target understanding



Cryo EM structure of human full-length ATM with inhibitor M4076 solved at 3.0 Å. Published in Nature Struct. & Mol. Biology. This would be the first published ATM structure with an inhibitor and at the highest resolution.

Internal and External Scientific Collaboration Network to support NBE and NCE Projects



Thank you for your attention

Marta Amaral, Ulrich Graedler, Djordje Musil, Daniel Schwarz Jörg Bomke, Ansgar Wegener, Hans-Peter Buchstaller

Rebecca Wade, Daria Kokh

Pedro Matias
Tiago Bandeiras
Filipe Freire
Diana Silva
Carolina Cassona
Manuel Freitas

Doris Schuetz
Lars Richter
Gerhard Ecker

Matthias Dreyer
Jérémie Topin
Marc Bianciotto

Jörn Güldenhaupt

- NMR-Bio SAS and IBS / Grenoble
- Pavel Macek, Rime Kerfah,
- Faustine Henot, Jerome Boisbouvier