The medicinal chemistry construction set

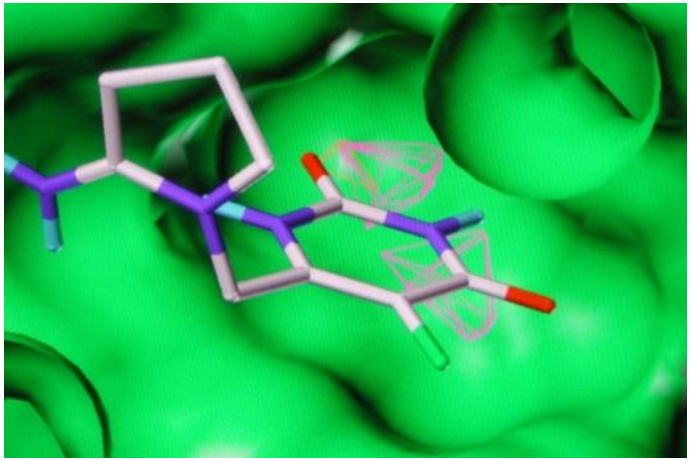
5. receptor-based *de novo* design

5. De novo design

- the cinderella method

De novo design methods

- Active site analysis
- Whole molecule
- Connection methods
  - Fragment reconnection
  - Site point
  - Sequential build-up
  - Fragment connection



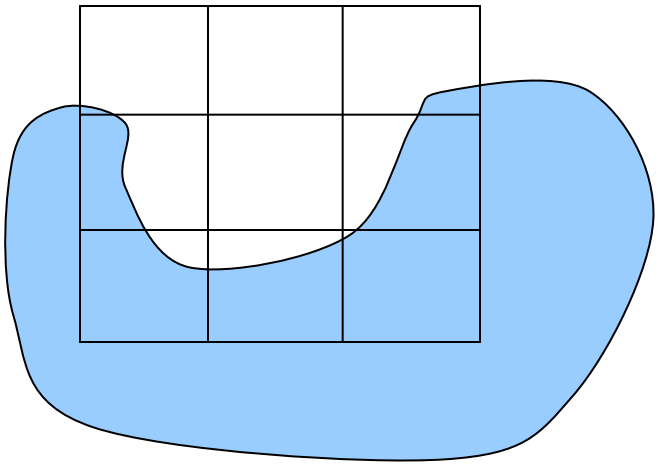
5.1. Active site analysis: Fragment placement methods

5.1.1. Grid-based functional group mapping

- probe potential interactions with active site of protein using selected small molecules to represent functional groups

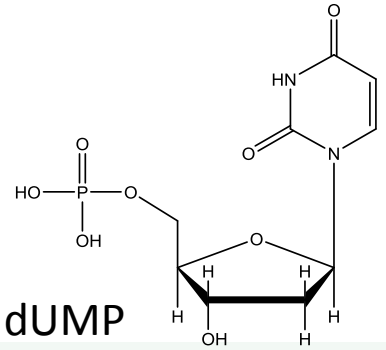
e.g. GRID program

- uses forcefield to calculate ΔU_{inter} of chemical probes



Chemical probe	**Ligand functional group/interaction type**
	hydrophobic
	amino
	ammonium
	hydroxyl
	aromatic hydroxyl

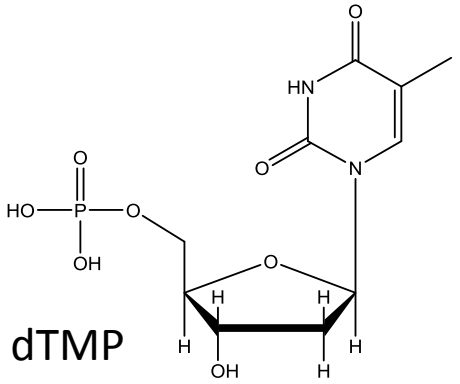
Example: thymidylate synthase



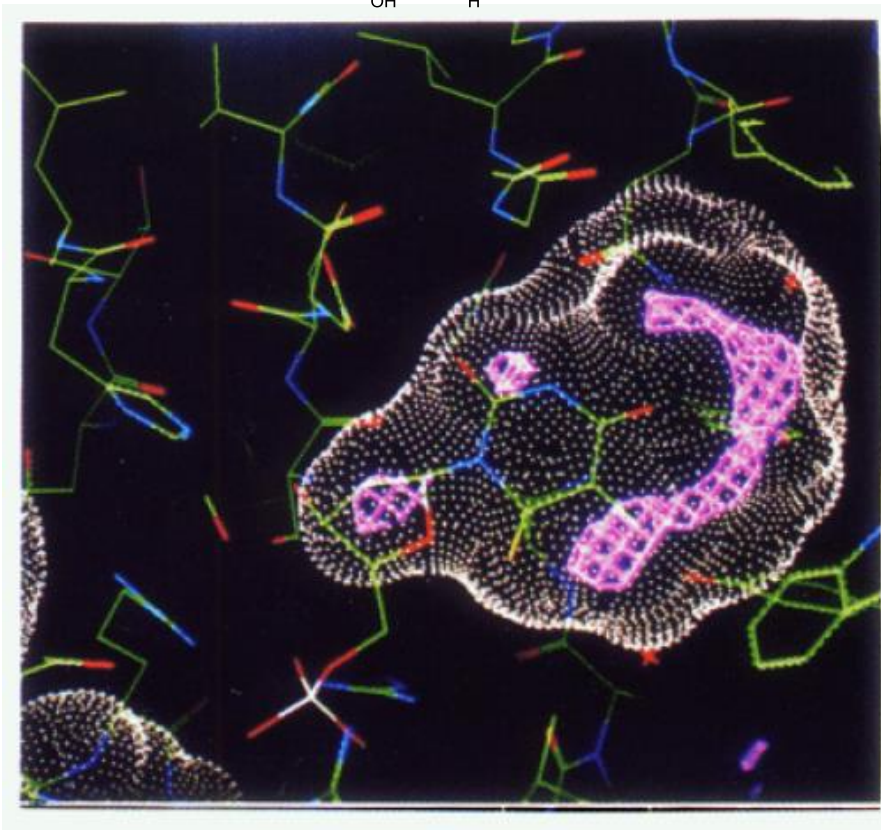
dUMP

thymidylate synthase

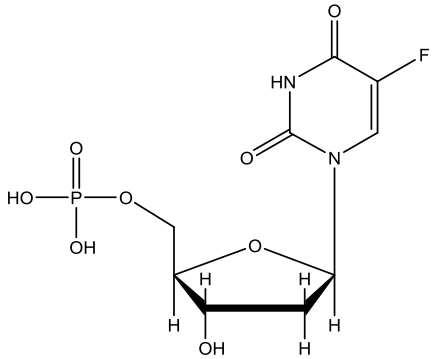
N^5,N^{10}-methylene THF → DHF



dTMP

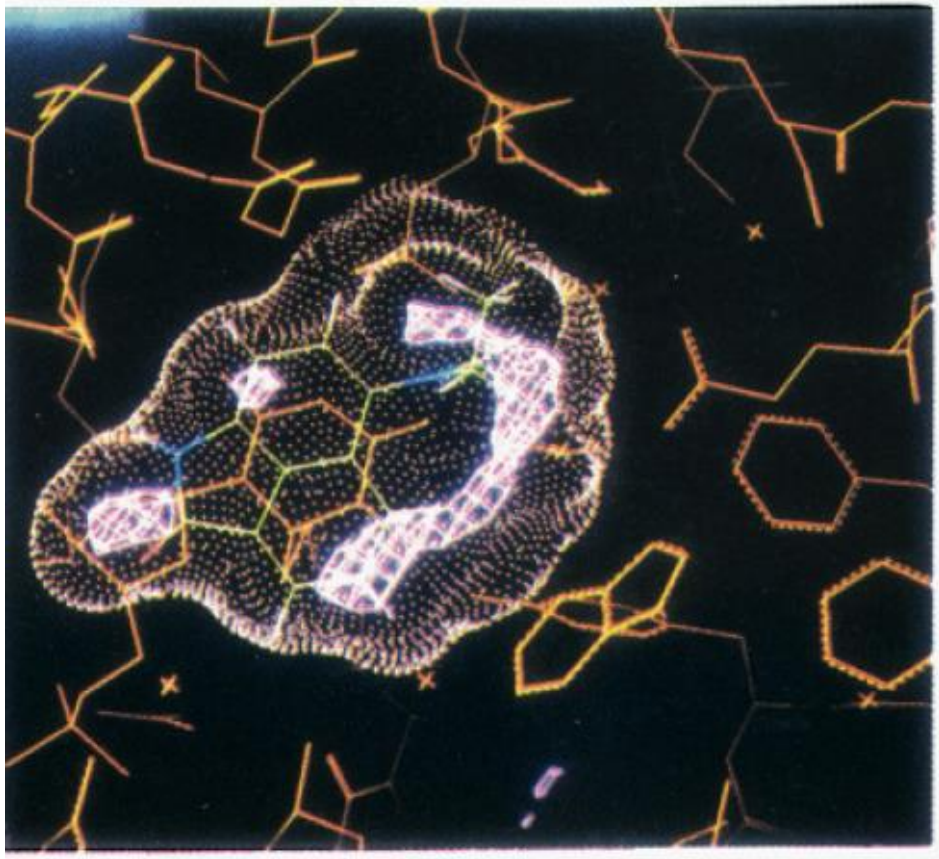


X-ray structure of the enzyme with 5F-dUMP:

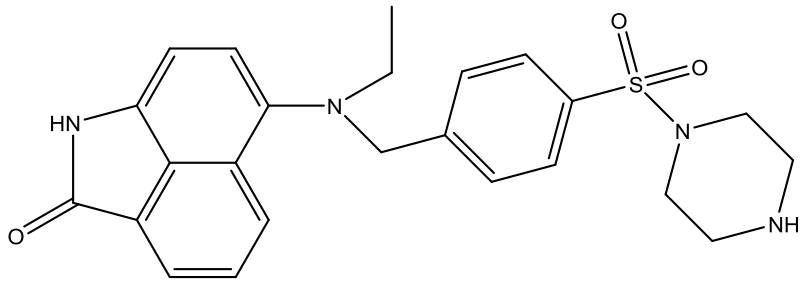


purple = Me probe of GRID (ie. hydrophobic)

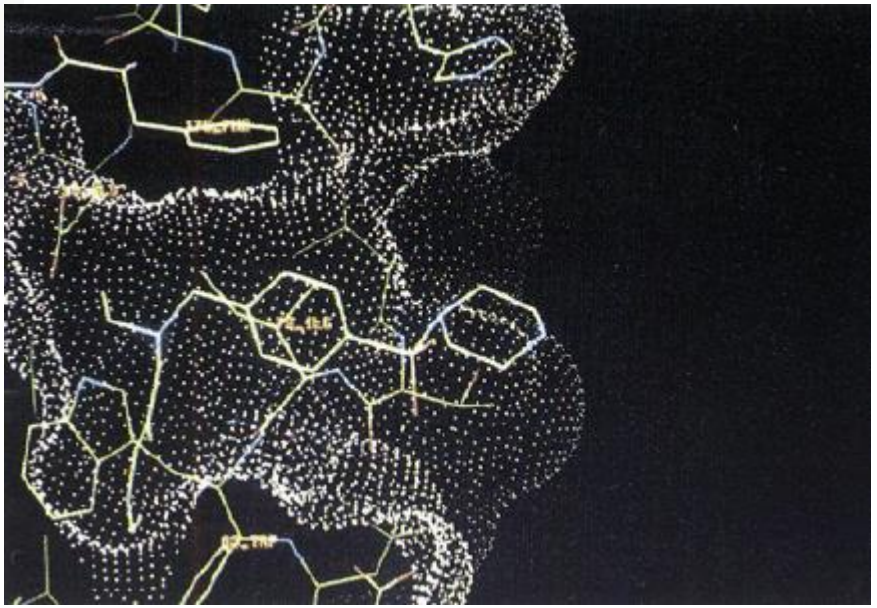
Journal of Medicinal Chemistry, **1994**, 37, 1036



(tricyclic portion of inhibitor)

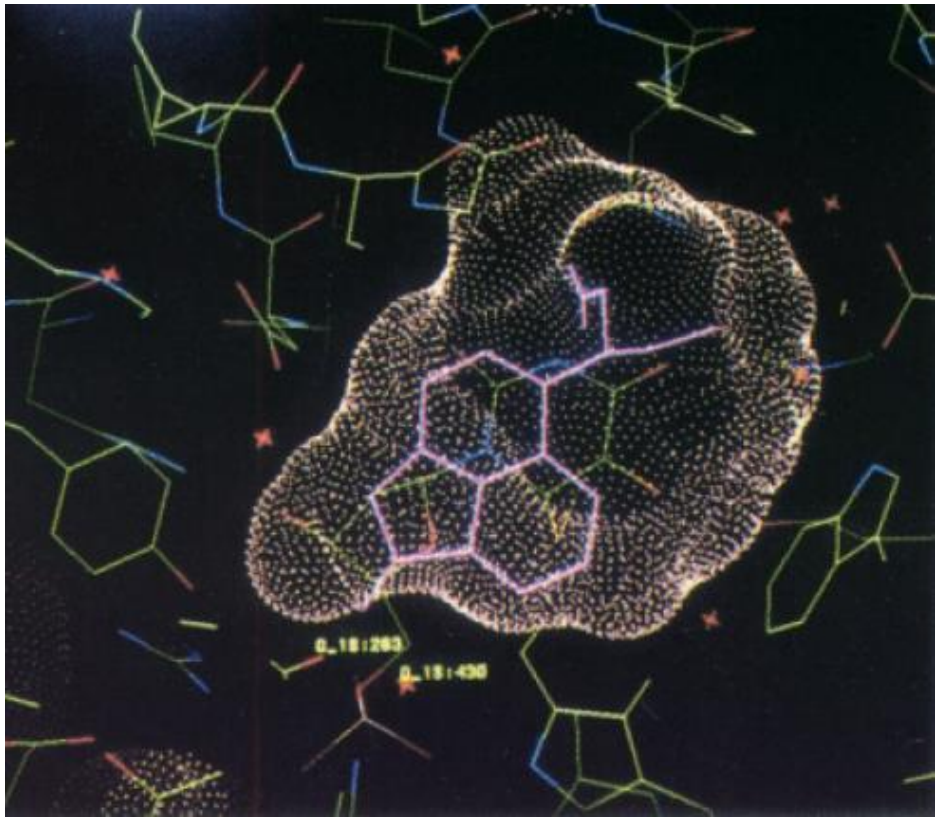


benzindole scaffold

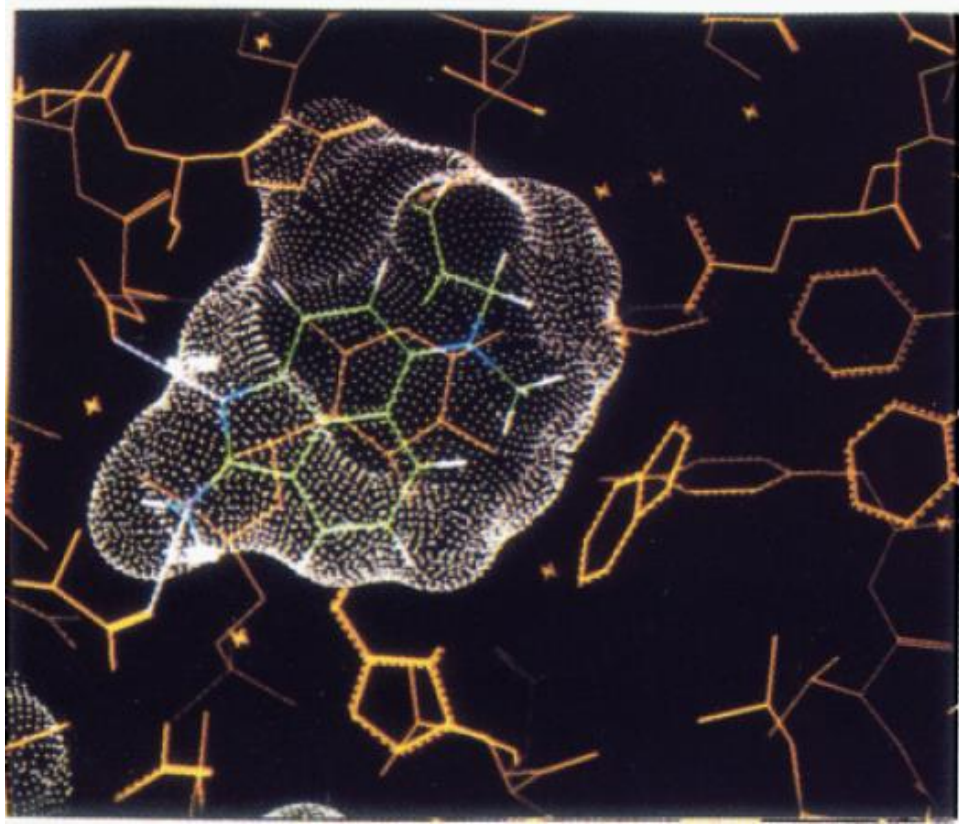


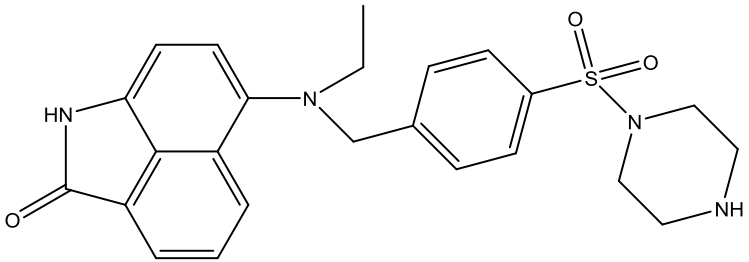
(benzylic portion of inhibitor)

Predicted (Ki = 1.6 μM)



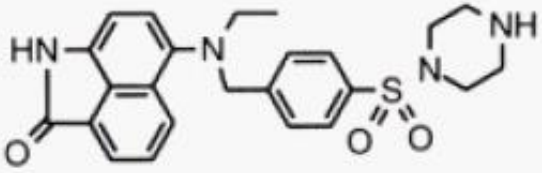
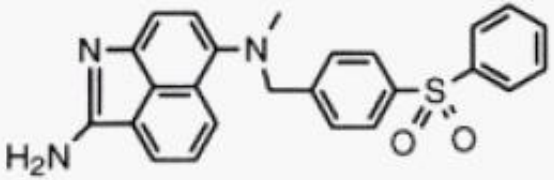
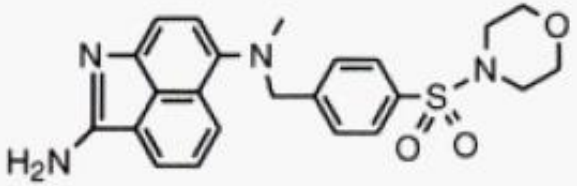
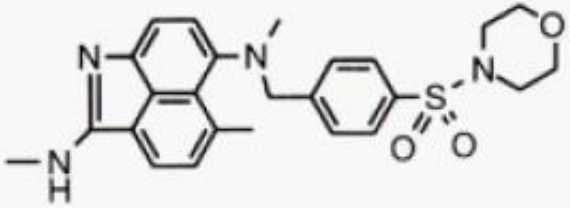
X-ray structure

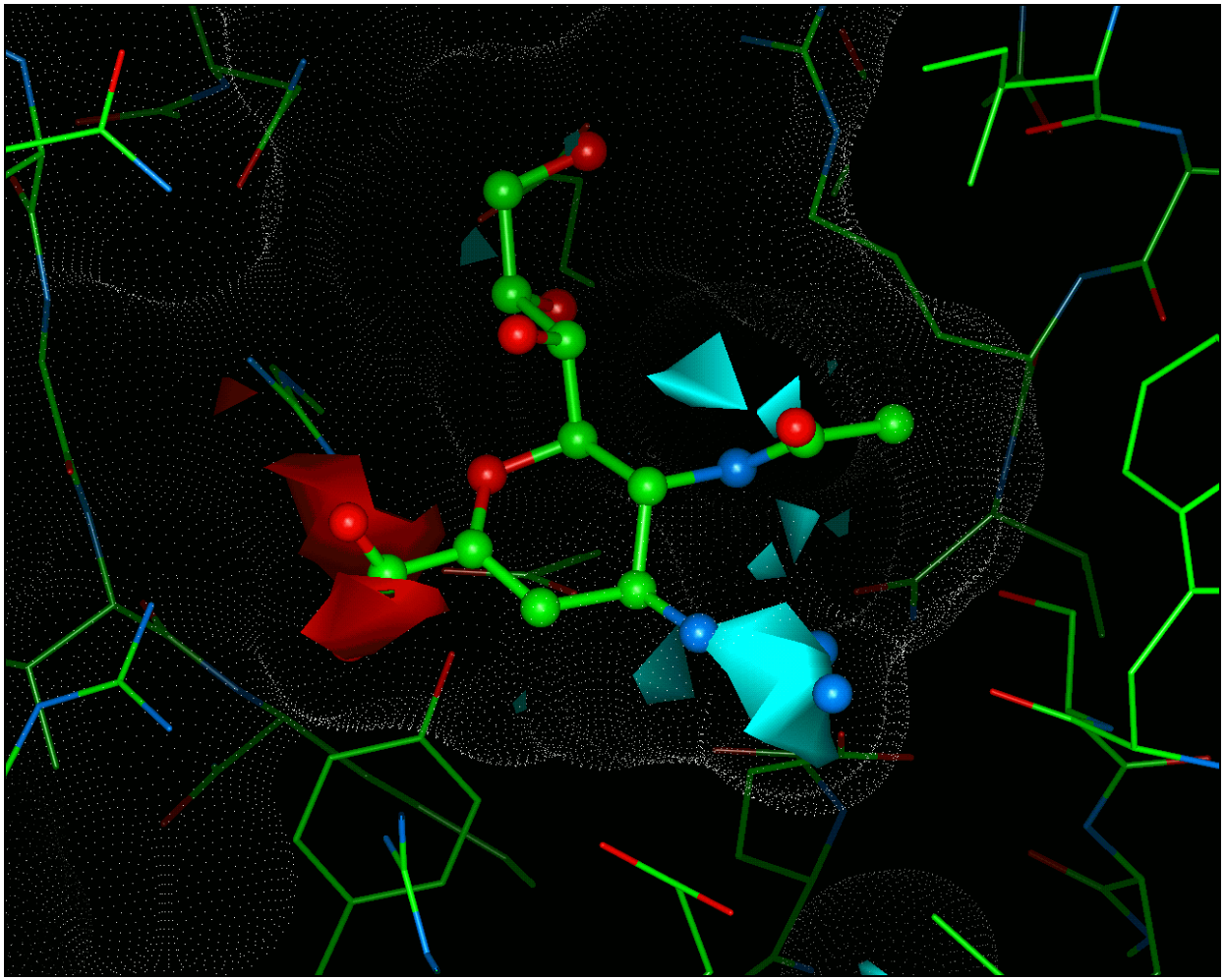




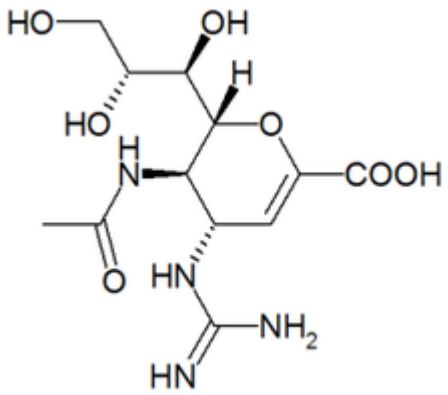
- Optimisation

Table 8. Rapid Optimization of a Novel Lead

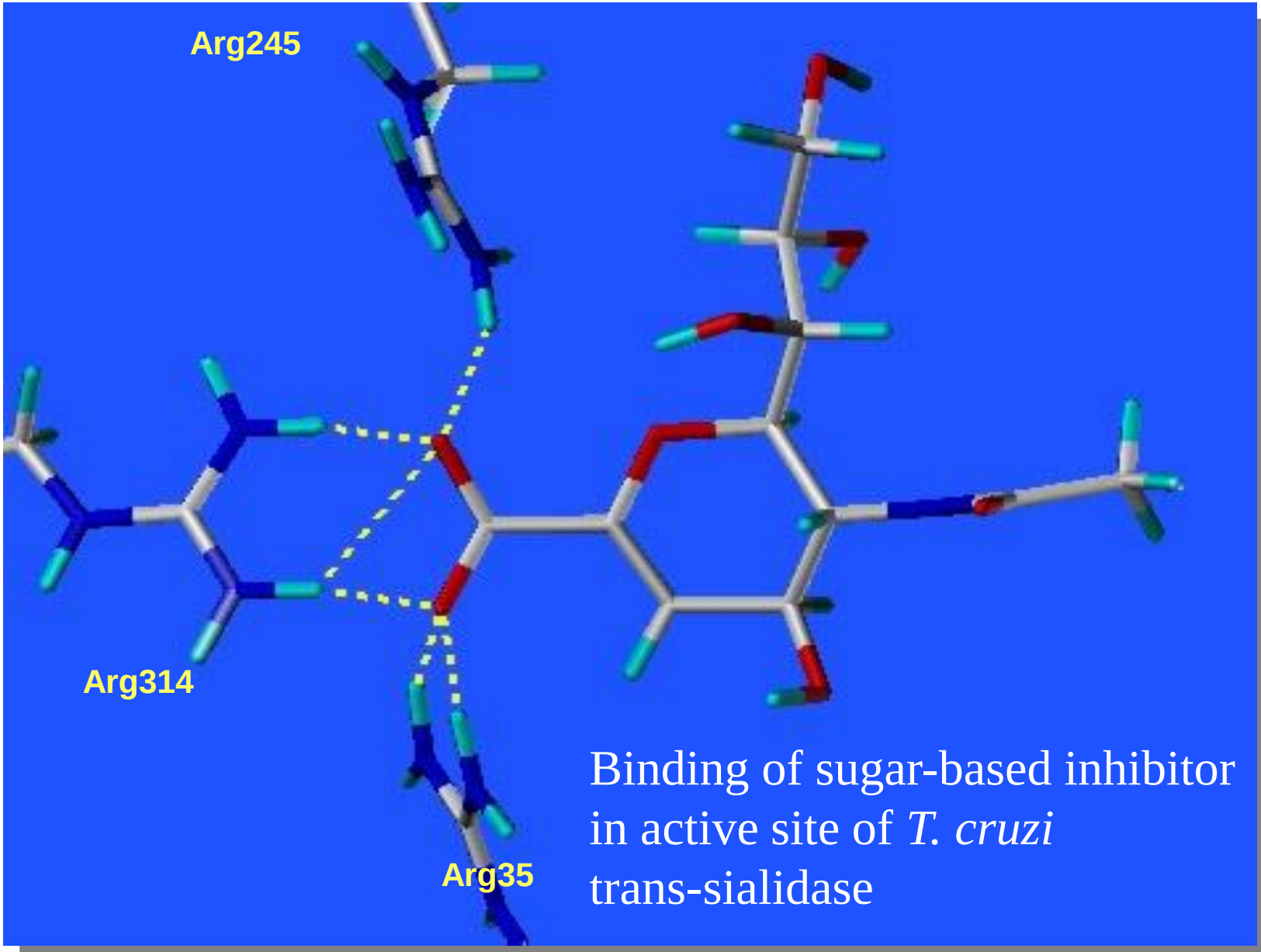
no.	structure	K_i (μM) human TS	IC_{50} (μM) L1210 cells	thymidine reversal
22		1.6	6.0	1.1
23		0.034	0.38	2.3
24		0.002	0.3	8.5
25		0.002	0.15	12.7



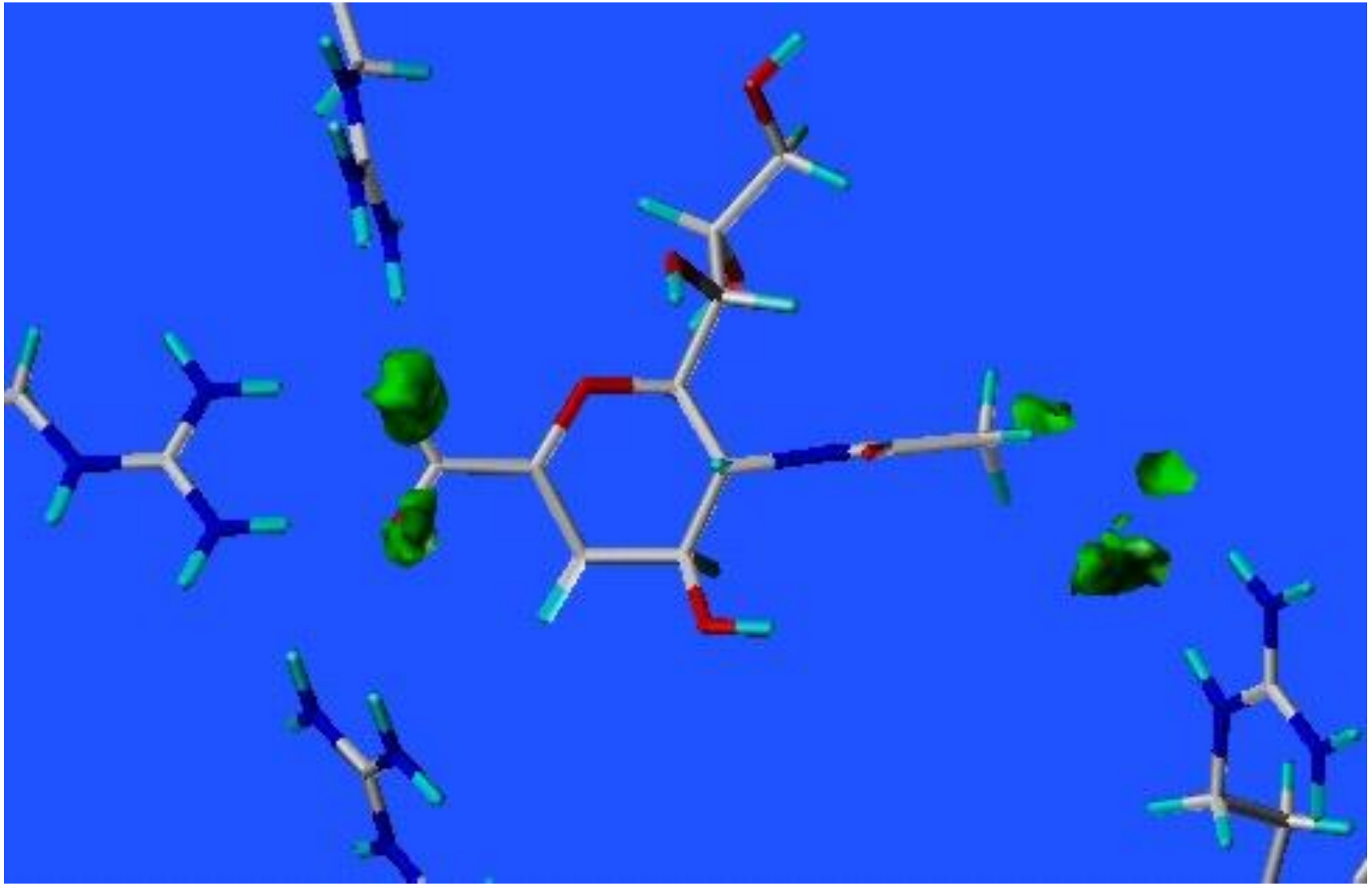
GRID analysis of neuraminidase active site



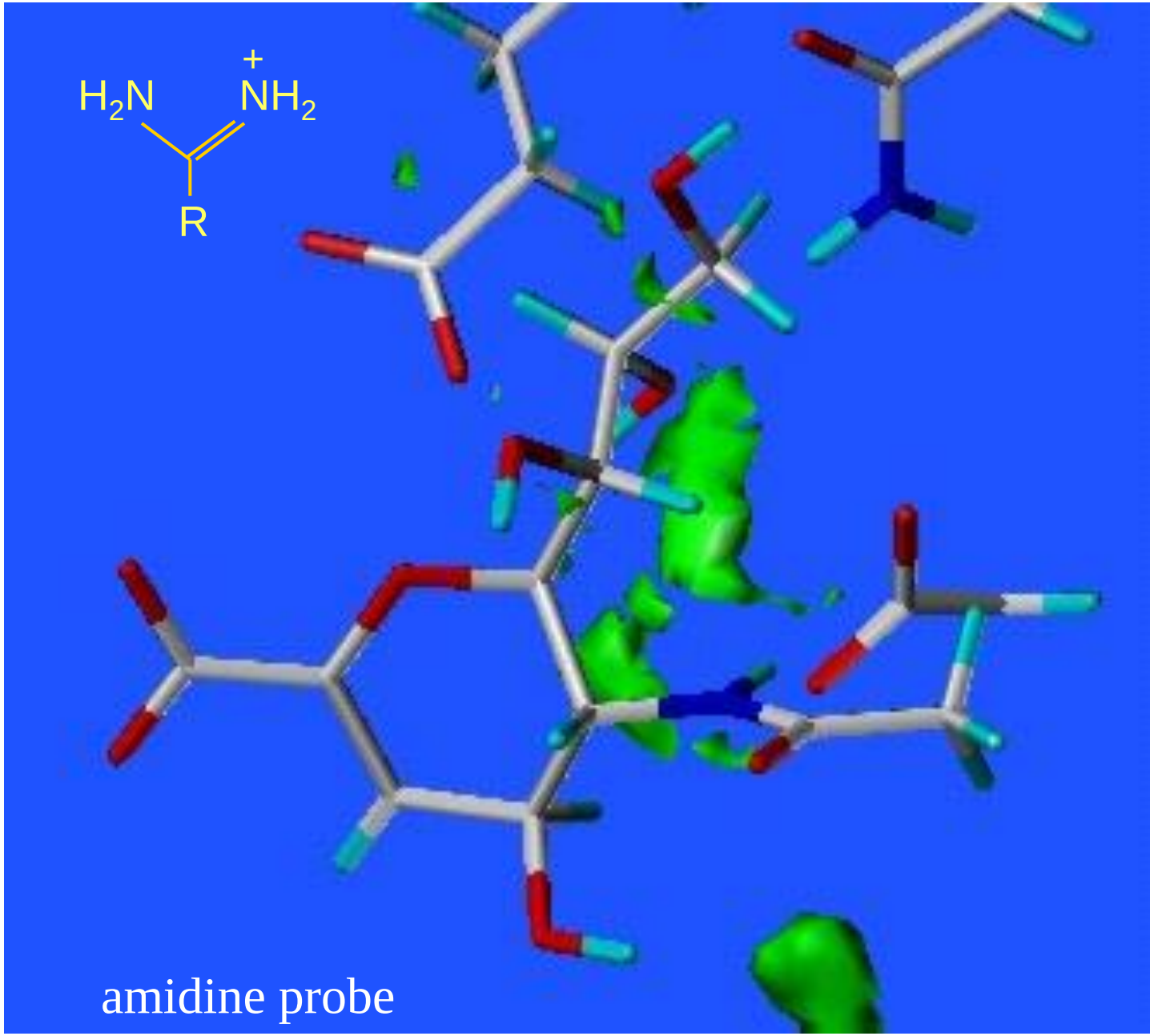
zanamavir (Relenza)



Binding of sugar-based inhibitor in active site of *T. cruzi* trans-sialidase



GRID analysis of TcTS site: green = (green: most favourable regions of carboxylate probe

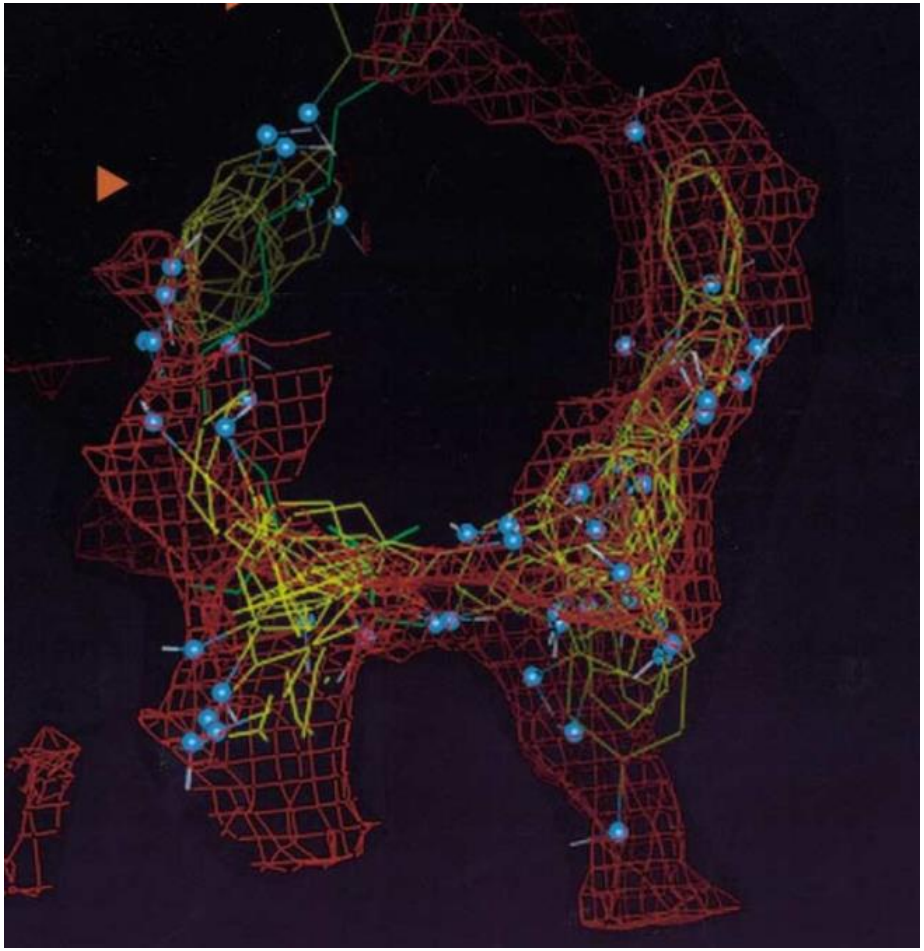


Other active site analysis methods – MCSS, GREEN, HINT, BUCKETS, SiteMap, SiteFinder, fpocket

5.1.2. MD-based functional group mapping

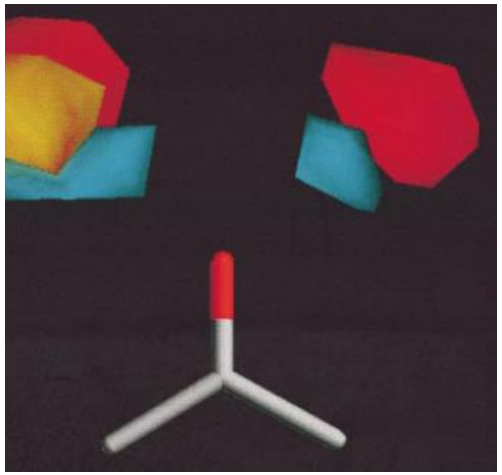
MCSS (Multiple Copy Simultaneous Search) **method**

- run multiple MD simulations of the same probe
  - gradually cool system
  - see where phenol probe molecule prefers to locate (blue spheres = phenol oxygen sites)
  - compares well with GRID's phenol density

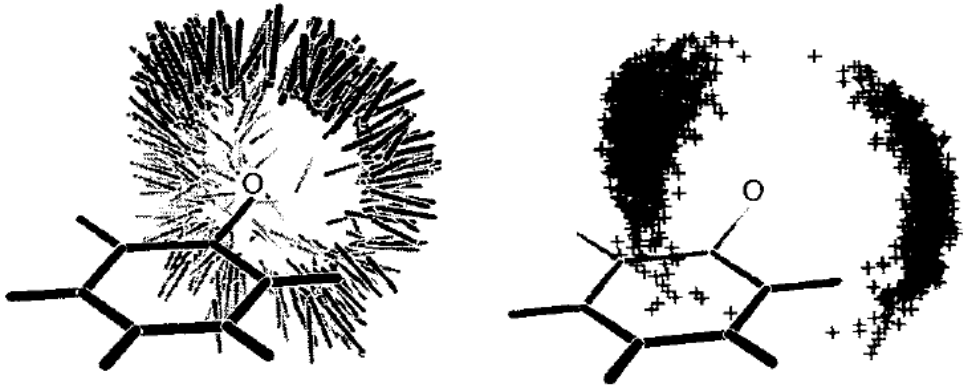


5.1.3. Crystal structure-based functional group mapping

- Predict where atoms prefer to locate in an active site based on their preferred distribution in known X-ray crystal structures of protein-ligand complexes
  - SuperStar program to analyse database of X-ray structures in the PDB



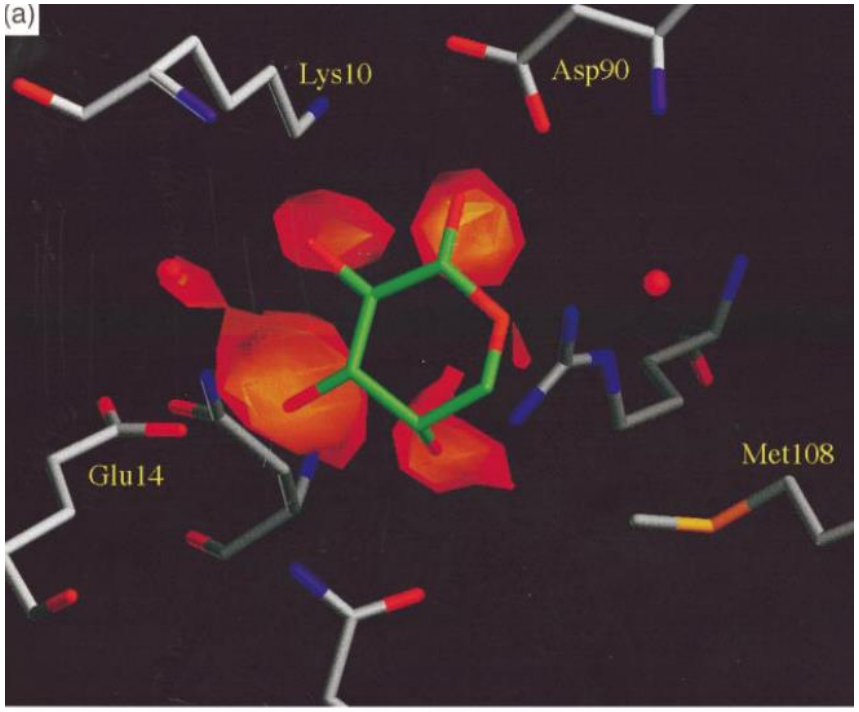
eg. where ligand H (blue) and O (red) atoms prefer to locate around a protein's carbonyl group



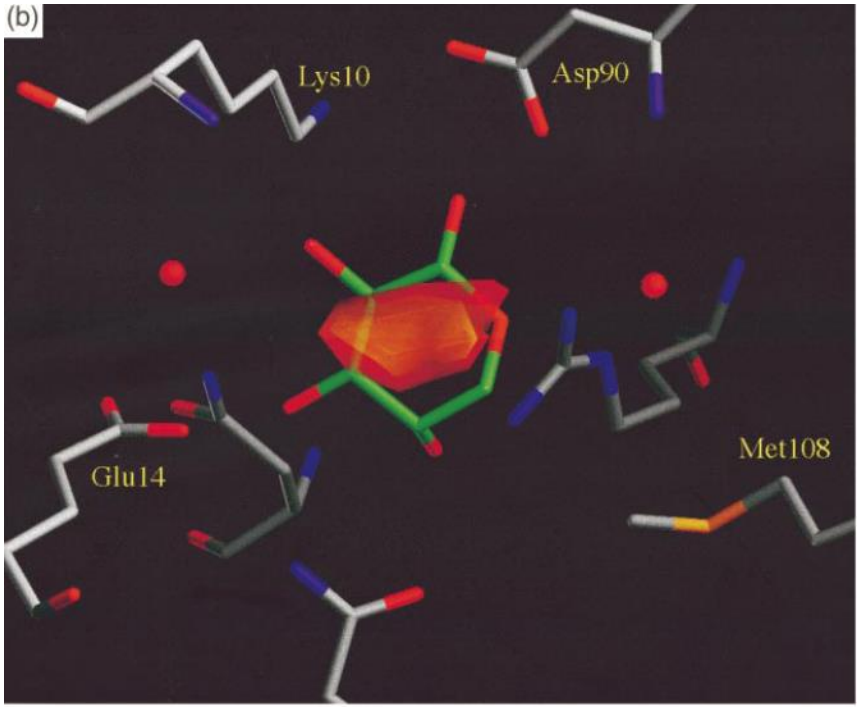
phenolic oxygen HBA/HBD

Angew. Chim. Int. Ed. 1996, 35, 2588
J. Mol. Biol. (1999) 289, 1093-1108

- eg. Arabinose Binding Protein active site
  - SuperStar's predicted atom densities superimpose onto where ligand actually observed to sit from X-ray

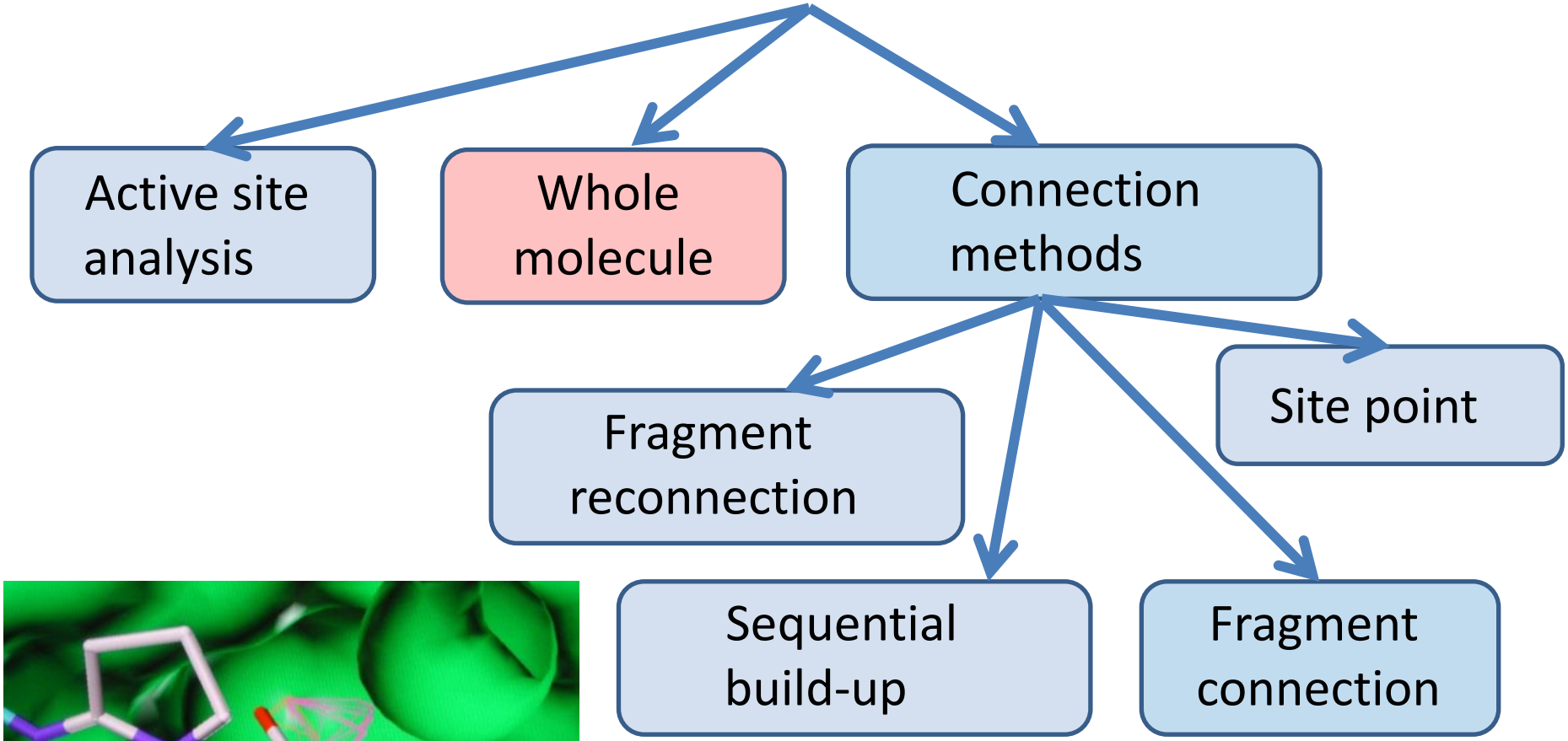


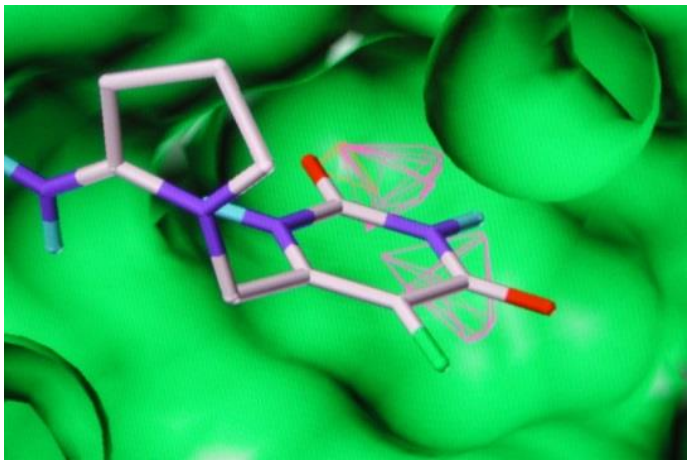
Hydroxyl contacts



Carbon atom contacts

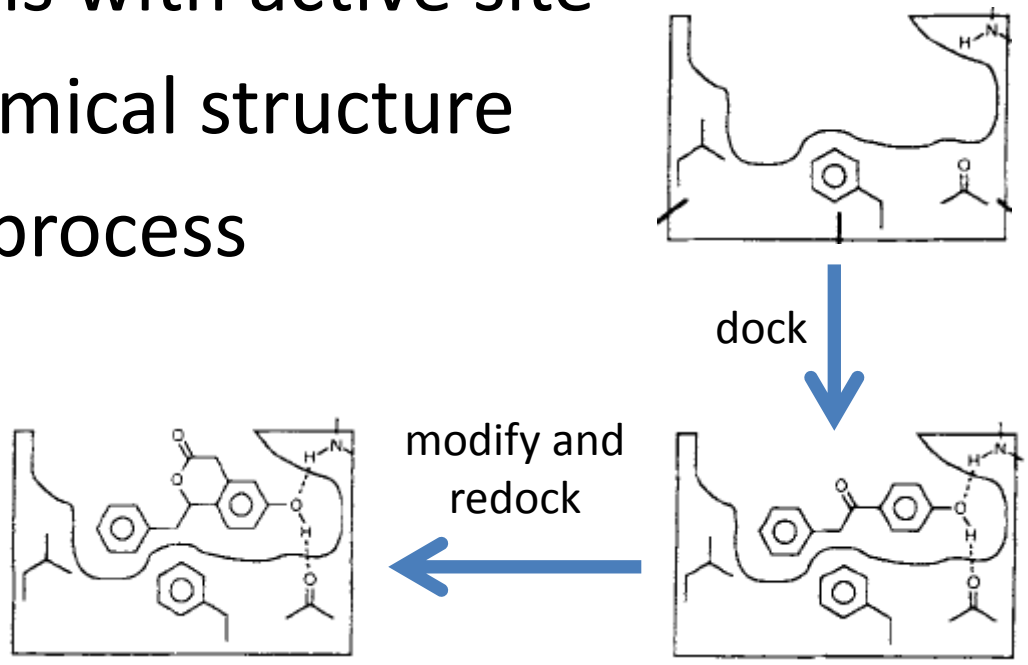
De novo design methods





5.2. Whole Molecule De Novo Design

1. “manually” build individual ligands on computer
2. use docking to predict their bound pose
3. examine interactions with active site
4. modify ligand’s chemical structure
5. redock and repeat process



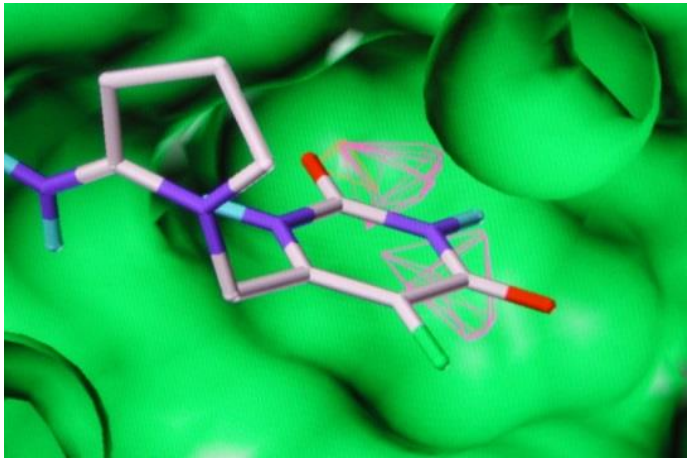
eg. using AutoDock software

Lacks diversity – you end up with something similar to what you started with

Angew. Chim. Int. Ed. 1996, 35, 2588

De novo design methods

- Active site analysis
- Whole molecule
- Connection methods
  - Fragment reconnection
  - Site point
  - Sequential build-up
  - Fragment connection

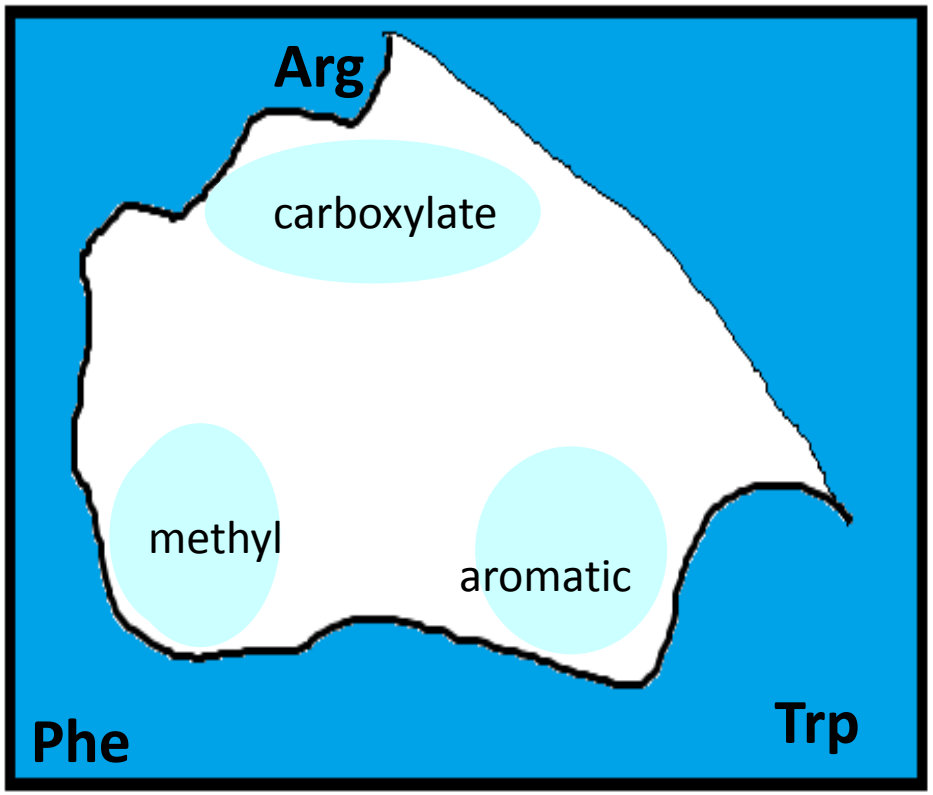


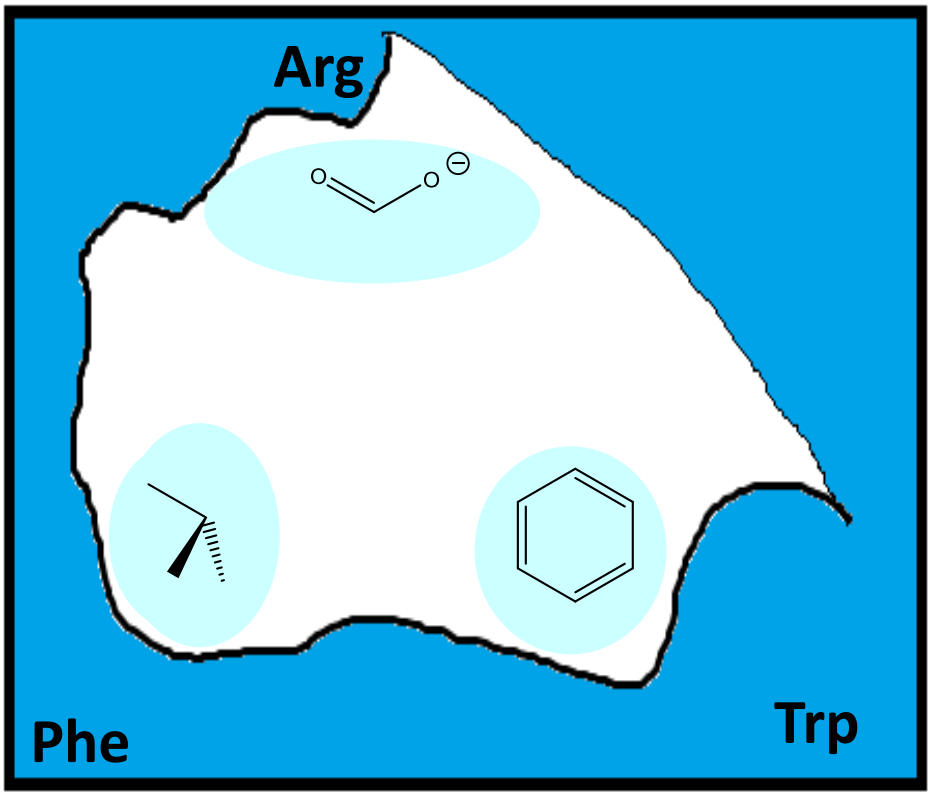
5.3. Connection methods

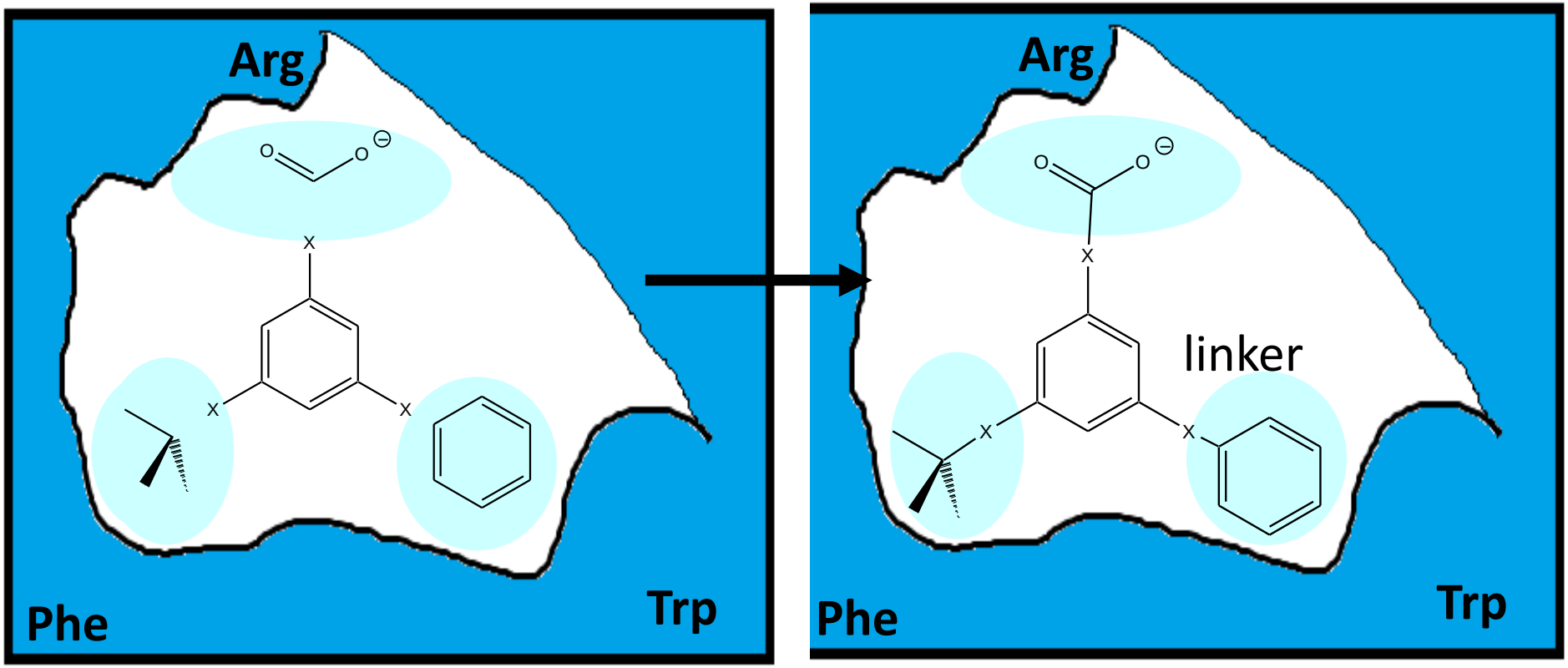
5.3.1. Fragment connection

- join favourably-bound functional groups together with chemically sensible linkers (called “outside-in” approach)
  - rely on fragment placement methods
  - *e.g.* CAVEAT program

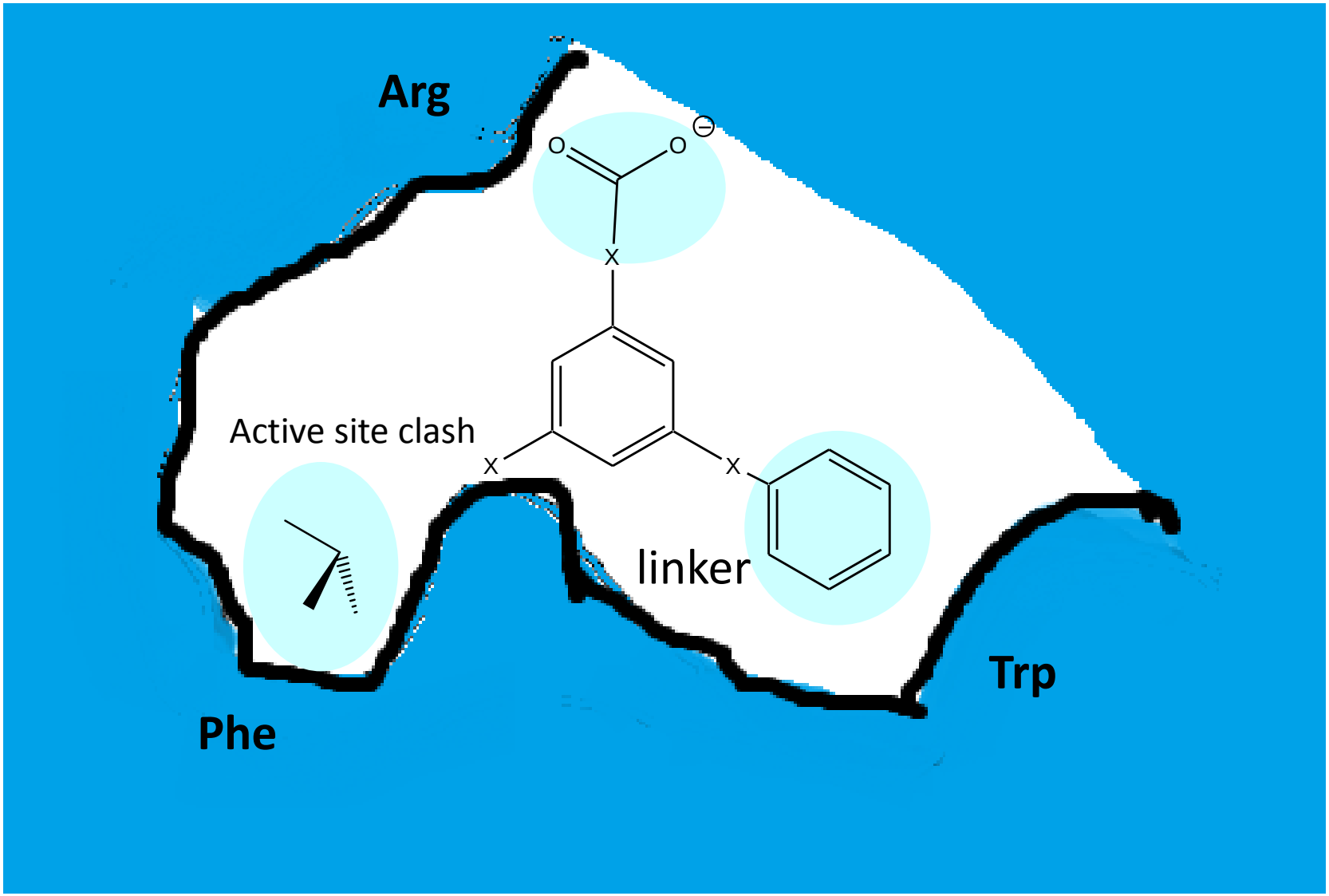
- Run probes, place fragments





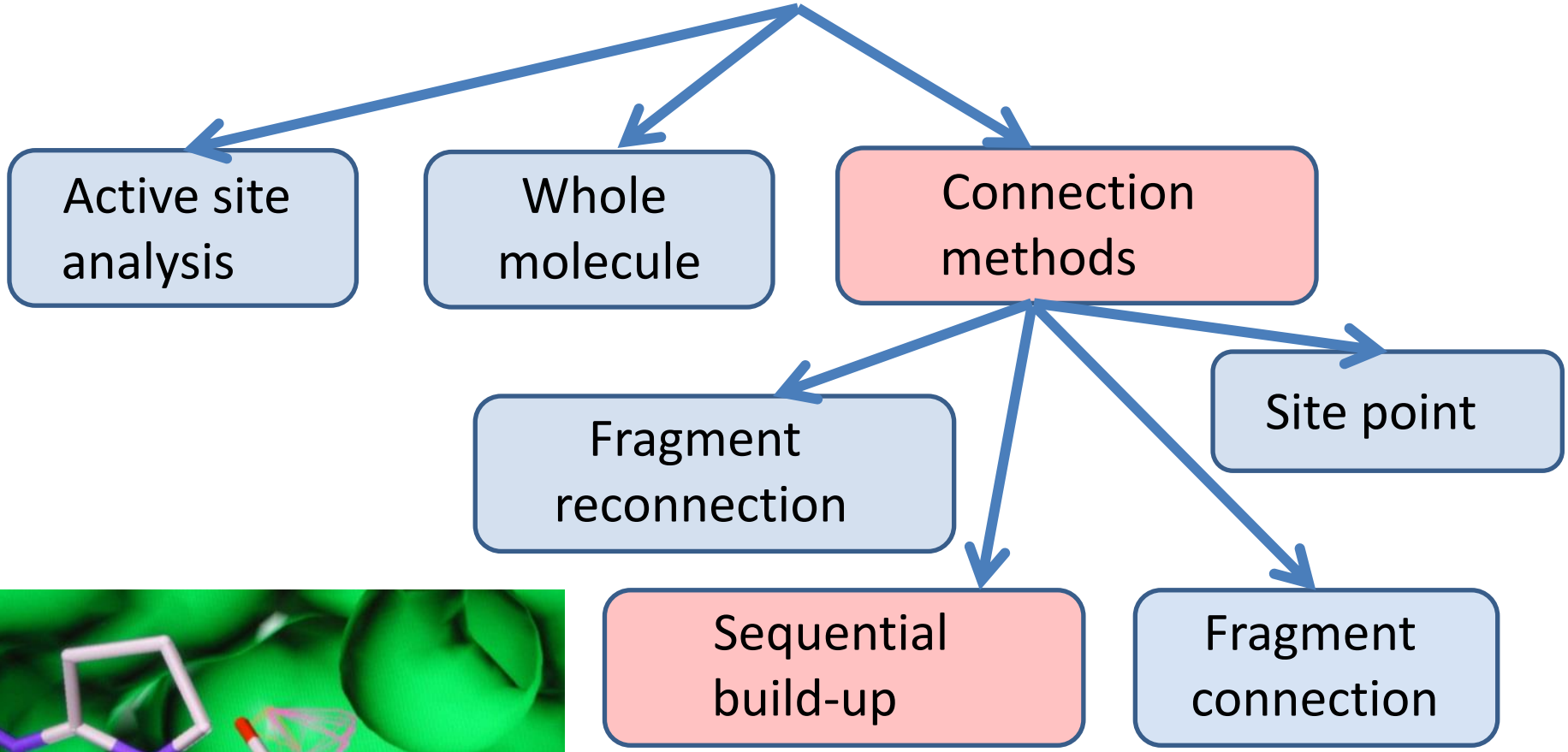

Arg
Phe
Trp
Arg
linker
Phe
Trp

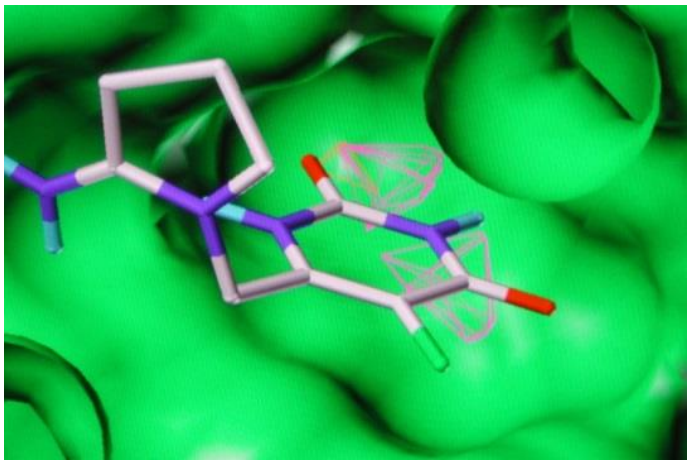
Account for active site sterics



Need to sample linker conformations (eg. by Monte Carlo)

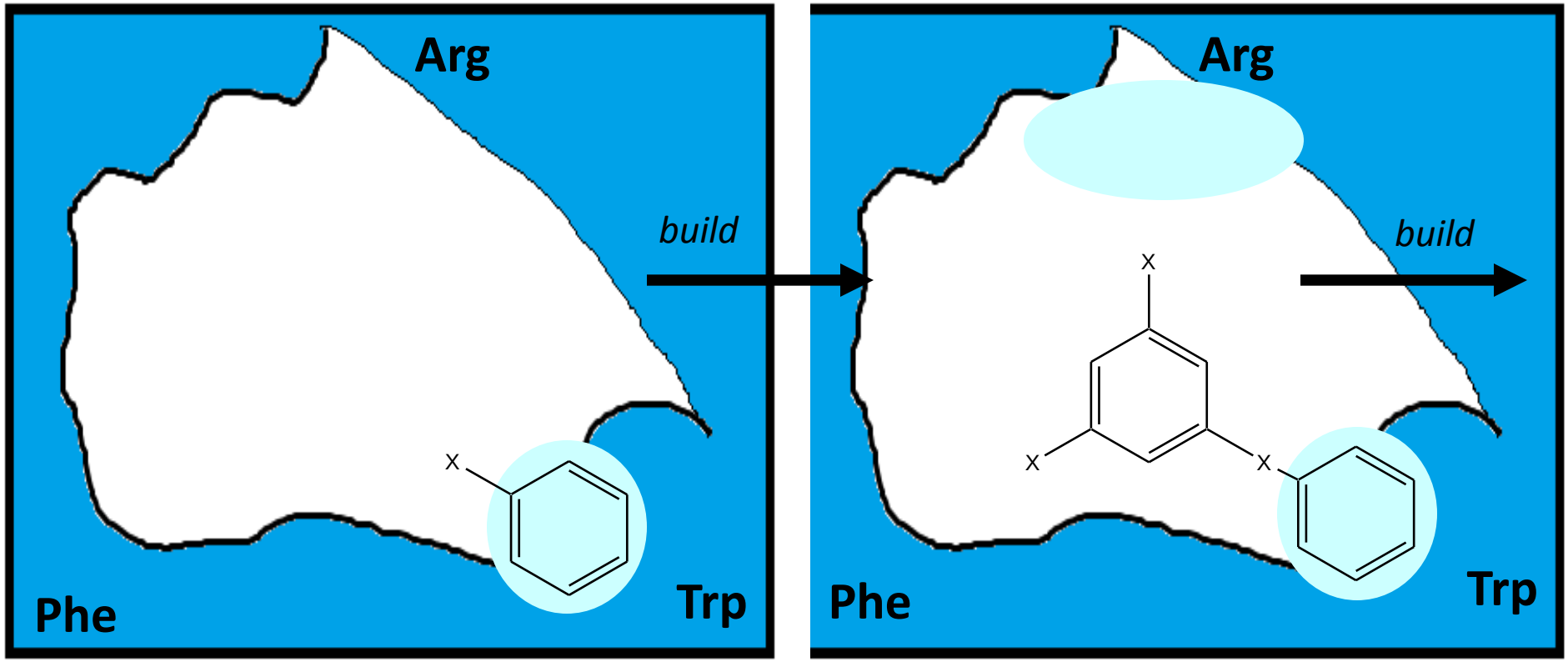
De novo design methods





5.3.2. Sequential build-up

- inside-out (or *aufbau*) approach, eg. GROWMOL, SmoG

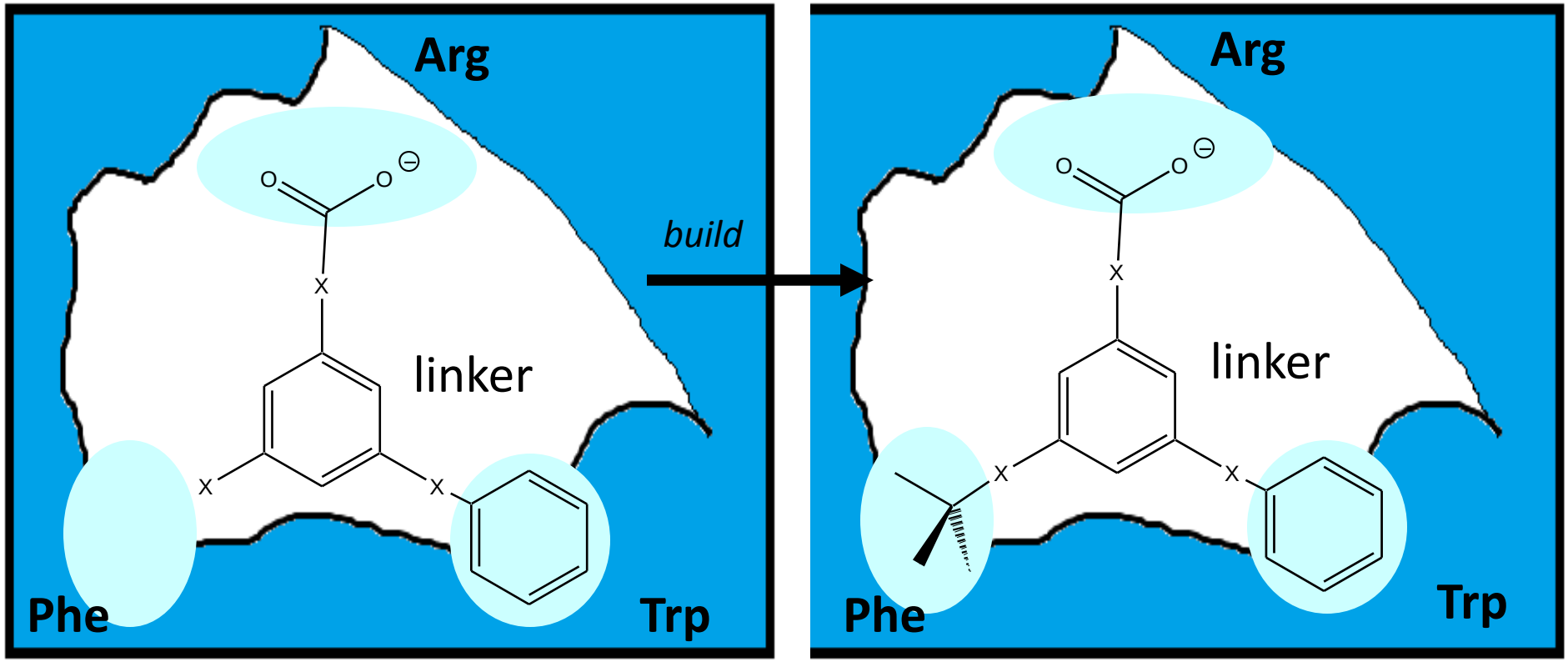


Use Monte Carlo to search space of additional parts of growing molecule

Sequential build-up

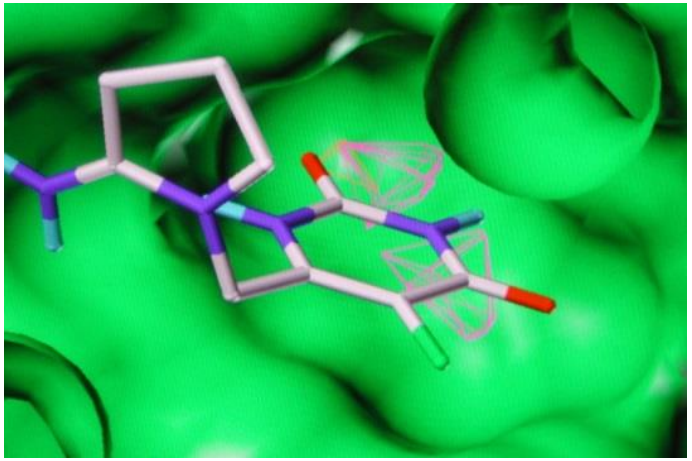
risk of combinatorial explosion

- a lot of conformations to explore as molecule grows



De novo design methods

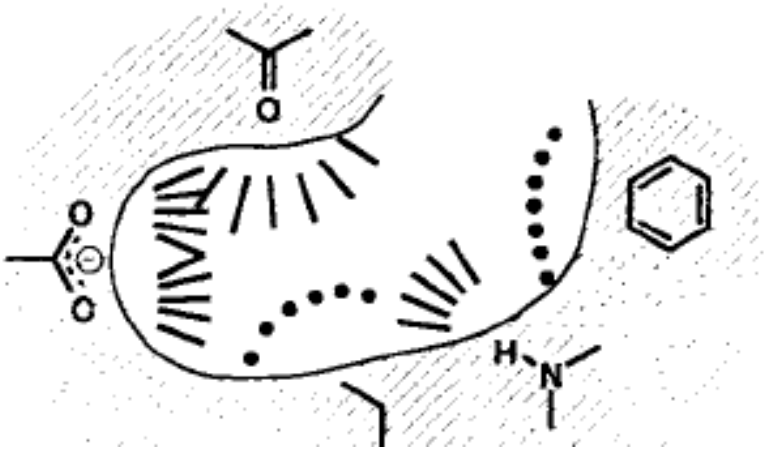
- Active site analysis
- Whole molecule
- Connection methods
  - Fragment reconnection
  - Site point
  - Sequential build-up
  - Fragment connection



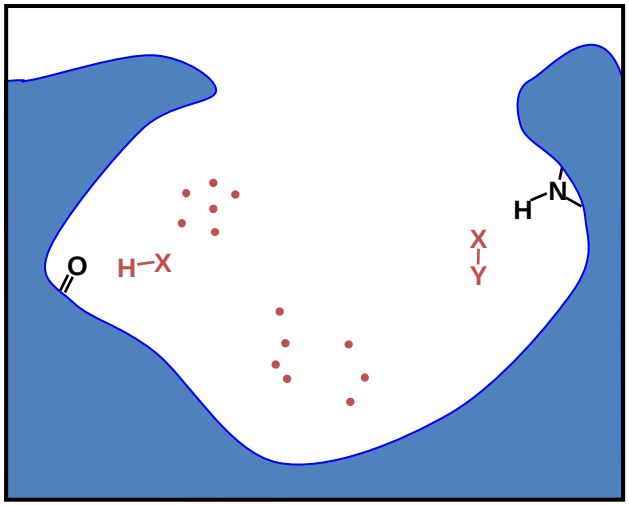
5.3.3. Site-point connection

- identifies possible interaction sites for binding regions within the binding site
- fits molecular fragments to different regions of the binding site
- links fragments

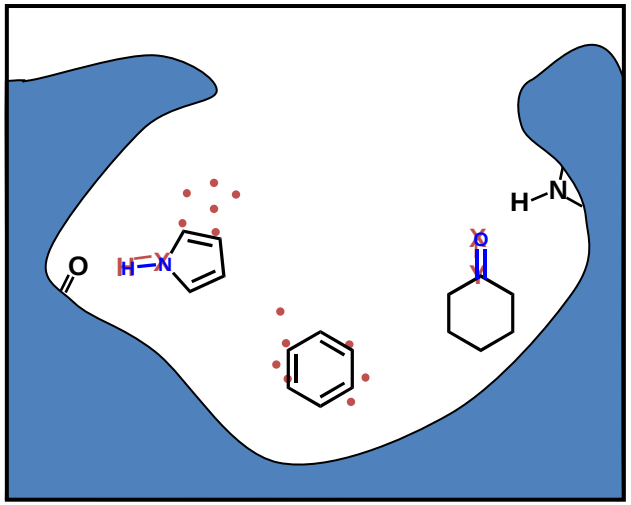
Eg. LUDI, SPROUT programs



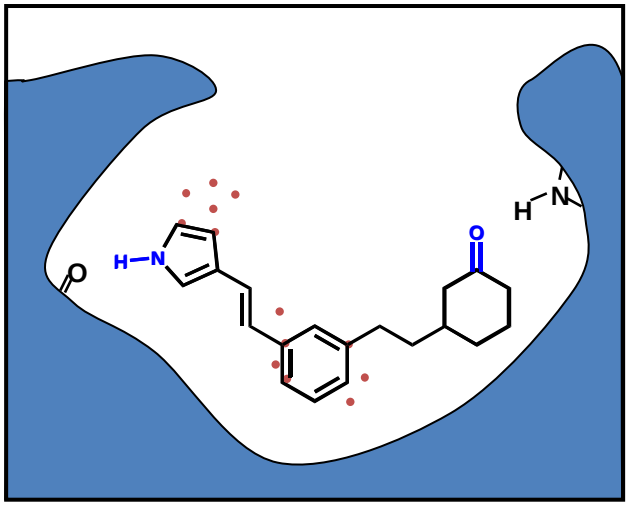
5.3.3.1. LUDI: Site-point connection



Interaction sites

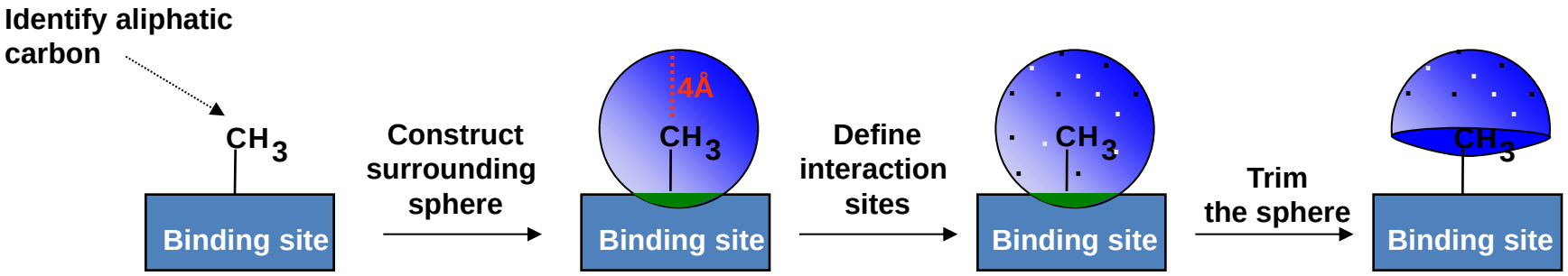


Fragment fitting



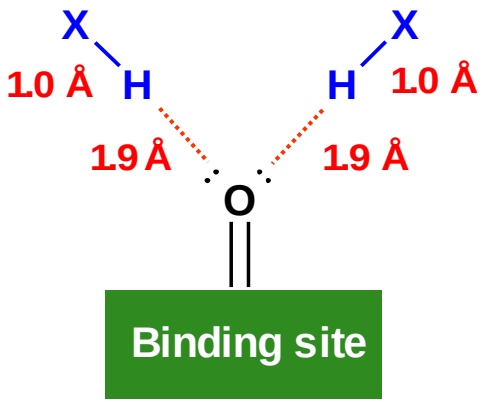
Bridging

Van der Waals site points

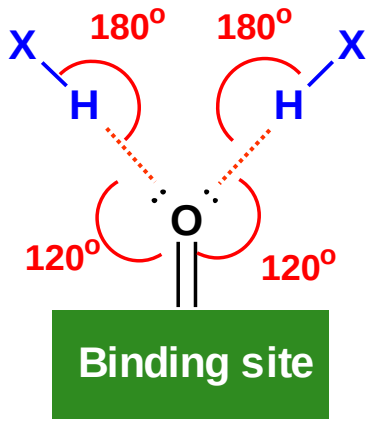


- Non-directional interaction
- Radius of sphere is optimum distance for interaction
- Interaction sites evenly distributed points on surface of sphere
- Trimming sphere removes points that are too close to binding site surface

Hydrogen bond site points



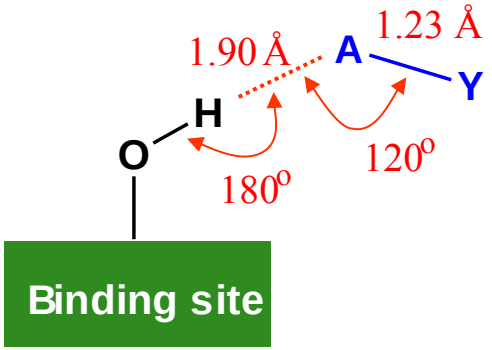
Required distances



Required bond angles

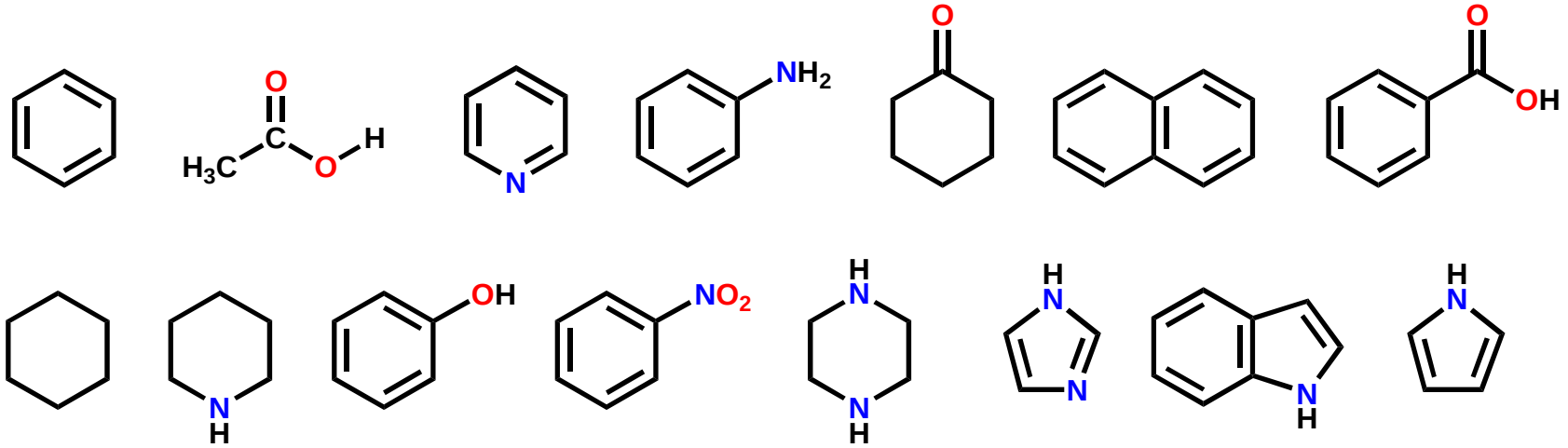
- Directional interactions involving an optimum distance and angle
- Identify amino acid HBAs
- Define site points as a vector involving two atoms (H-X)

Site-point connection



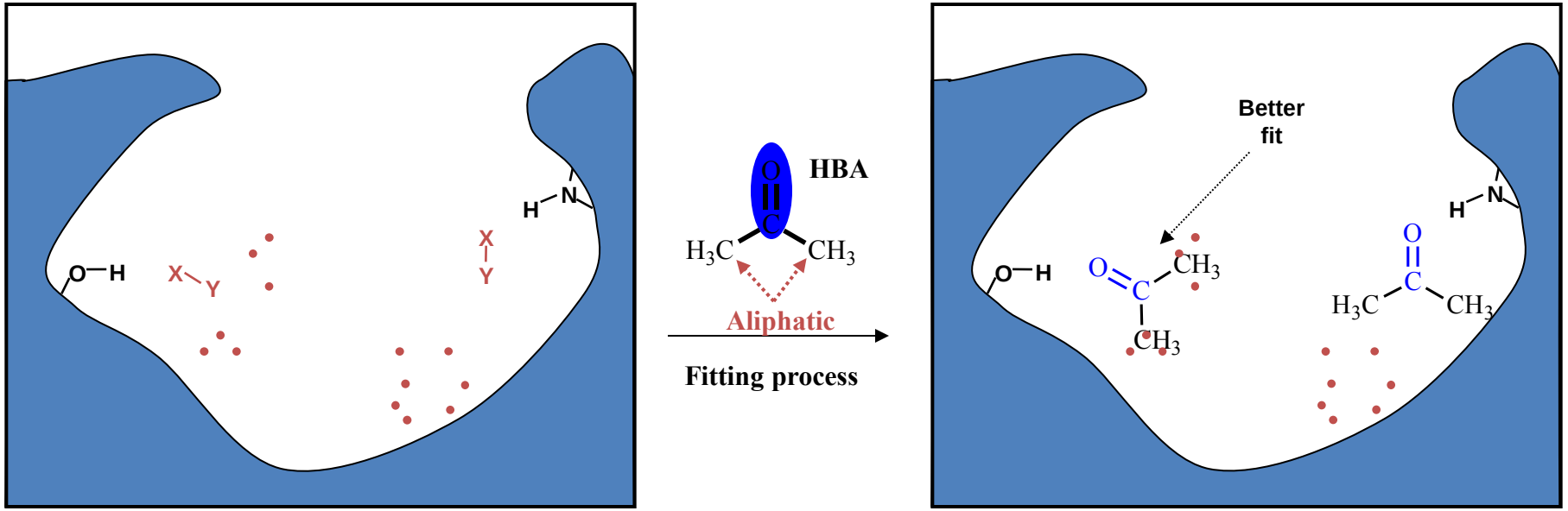
- Directional interactions involving an optimum distance and angle
- Identify amino acid HBDs
- Define site points as a vector involving two atoms (A-Y)

Fragment library



- Flexible fragments present as several conformations
- Atoms of fragments defined for fitting process
- Aliphatic atoms fitted onto van der Waals interaction points
- HBAs and HBDs fitted onto H-bond interaction points

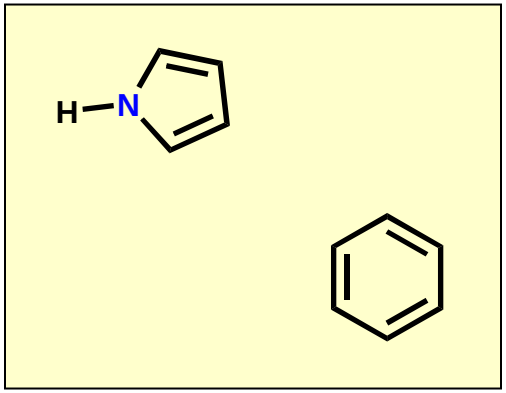
Fit fragments to maximise interactions



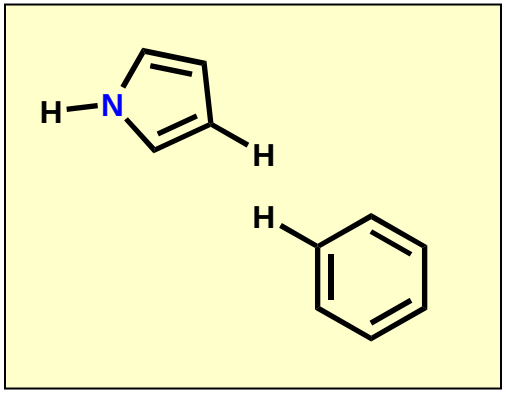
Interaction sites

Fitting fragments

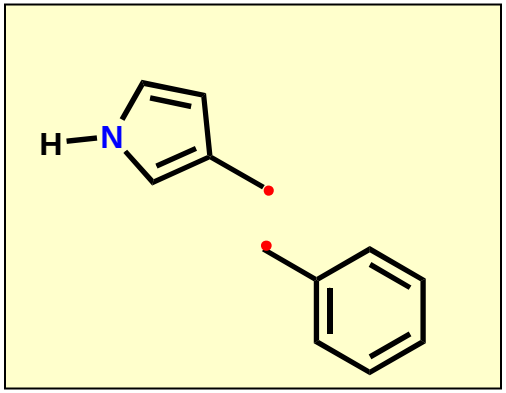
Join fragments



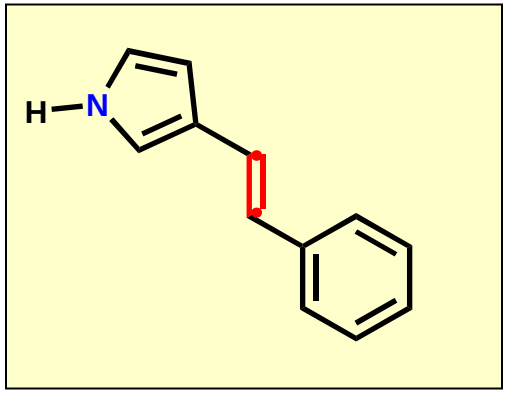
Identify closest fragments



Identify closest hydrogens

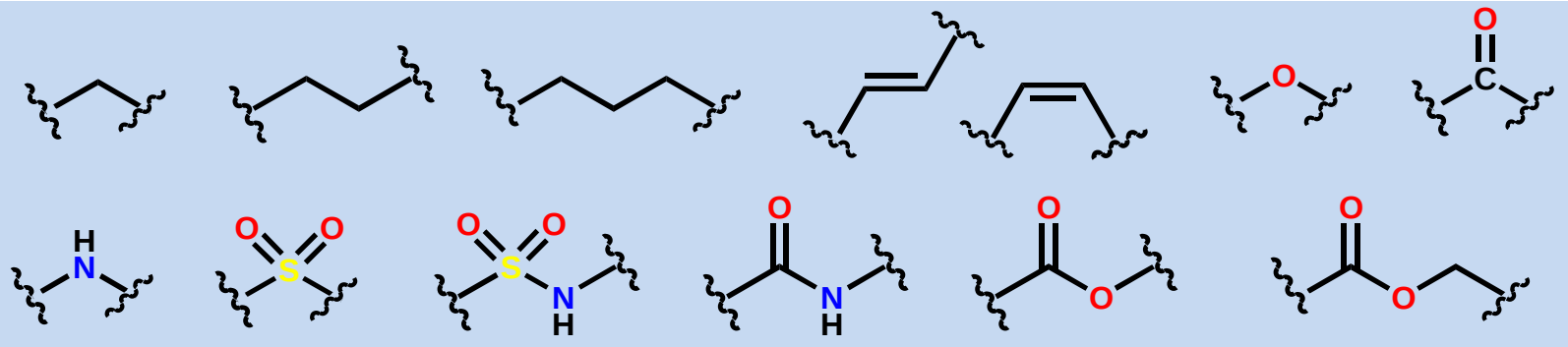


Link points

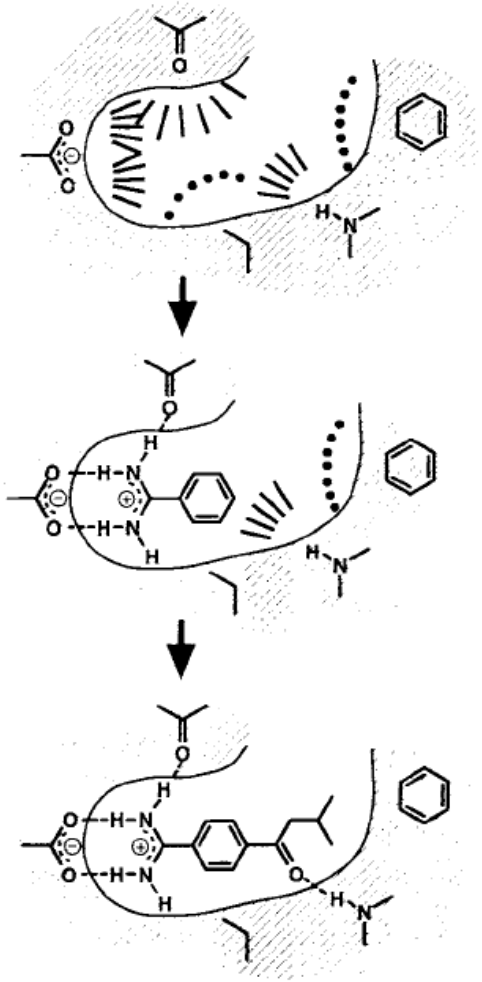


Fit bridging linker

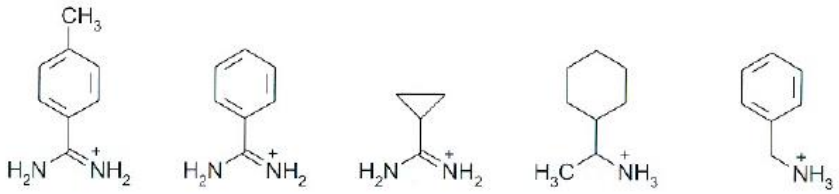
- Bridging linkers



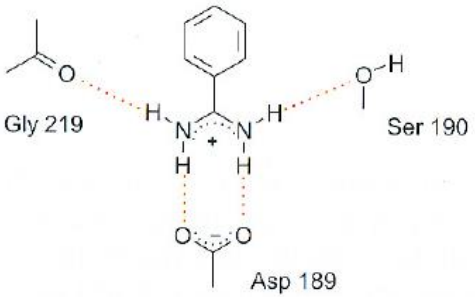
Example: de novo design of Trypsin inhibitors with LUDI

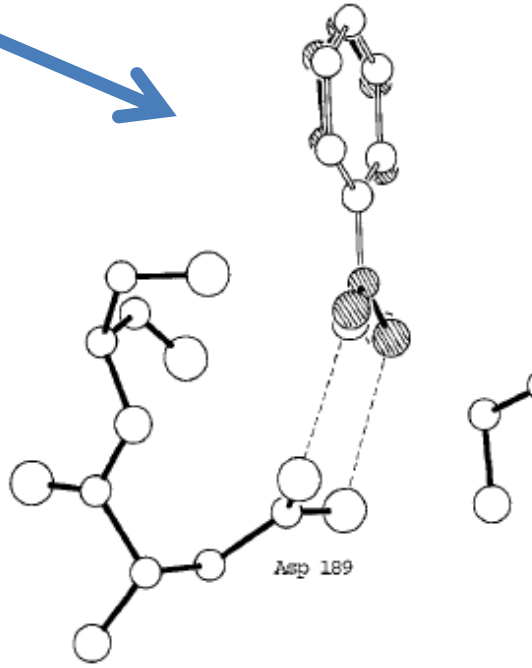


trypsin = serine protease



Fragments derived from 30k known small molecules





Top-ranking was known inhibitor, benzamidine

Angew. Chim. Int. Ed. 1996, 35, 2588-2614

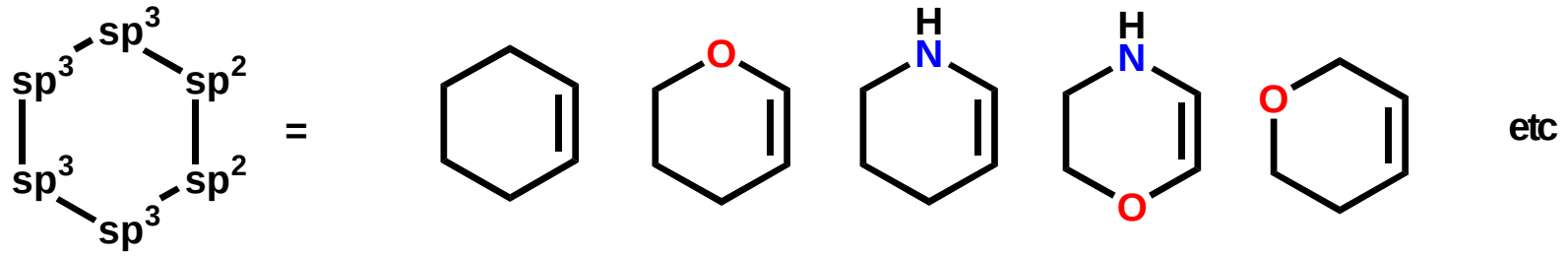
5.3.3.2. SPROUT: Site-point connection

- Sites can be:
  - from fragments placed in active site
  - pharmacophore points, derived from receptor
  - pharmacophore points, derived from active ligands

- then fit “fragment templates” to sites…

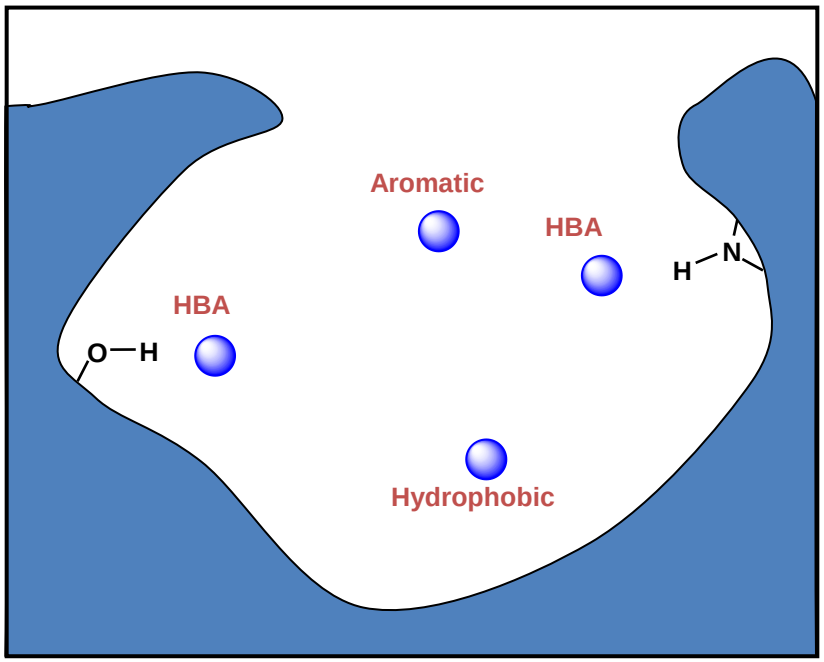
Fragment templates

- Each “fragment template” represents several different real fragments
  - atoms represents a generalised sp, sp^2 or sp^3 hybridised atom
  - bonds can be a single, double or triple bond

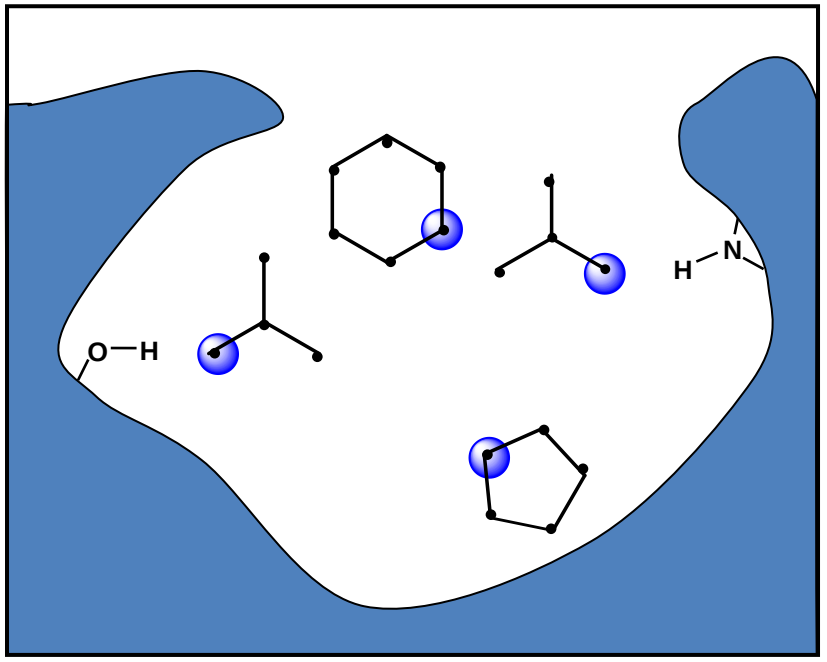


Fragment template **matching molecular fragments**

- Fragment templates selected randomly
- Atoms of fragment template chosen randomly and fitted to target sites



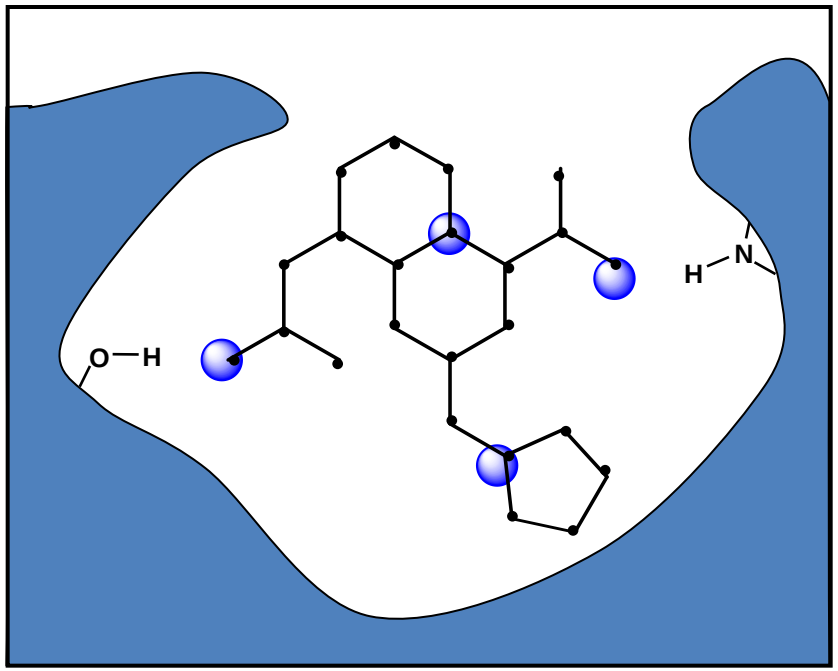
Interaction sites



Fitting “fragment templates”

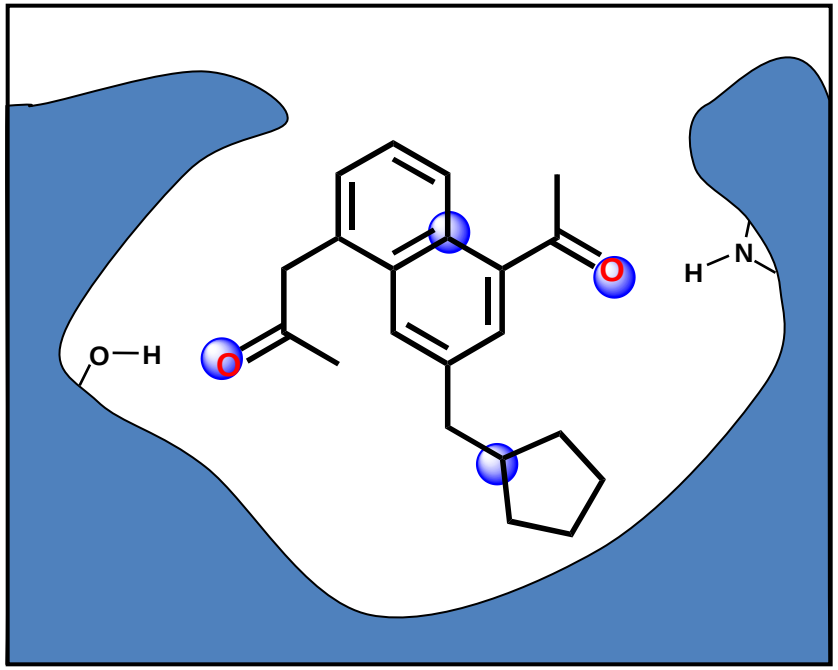
- Fragment templates grown inwards and linked
- Resulting combined template (“molecular skeleton”) converted to a range of real molecules
  - atoms added to allow required interactions
  - large number of molecules possible for each skeleton

Growth and linkage



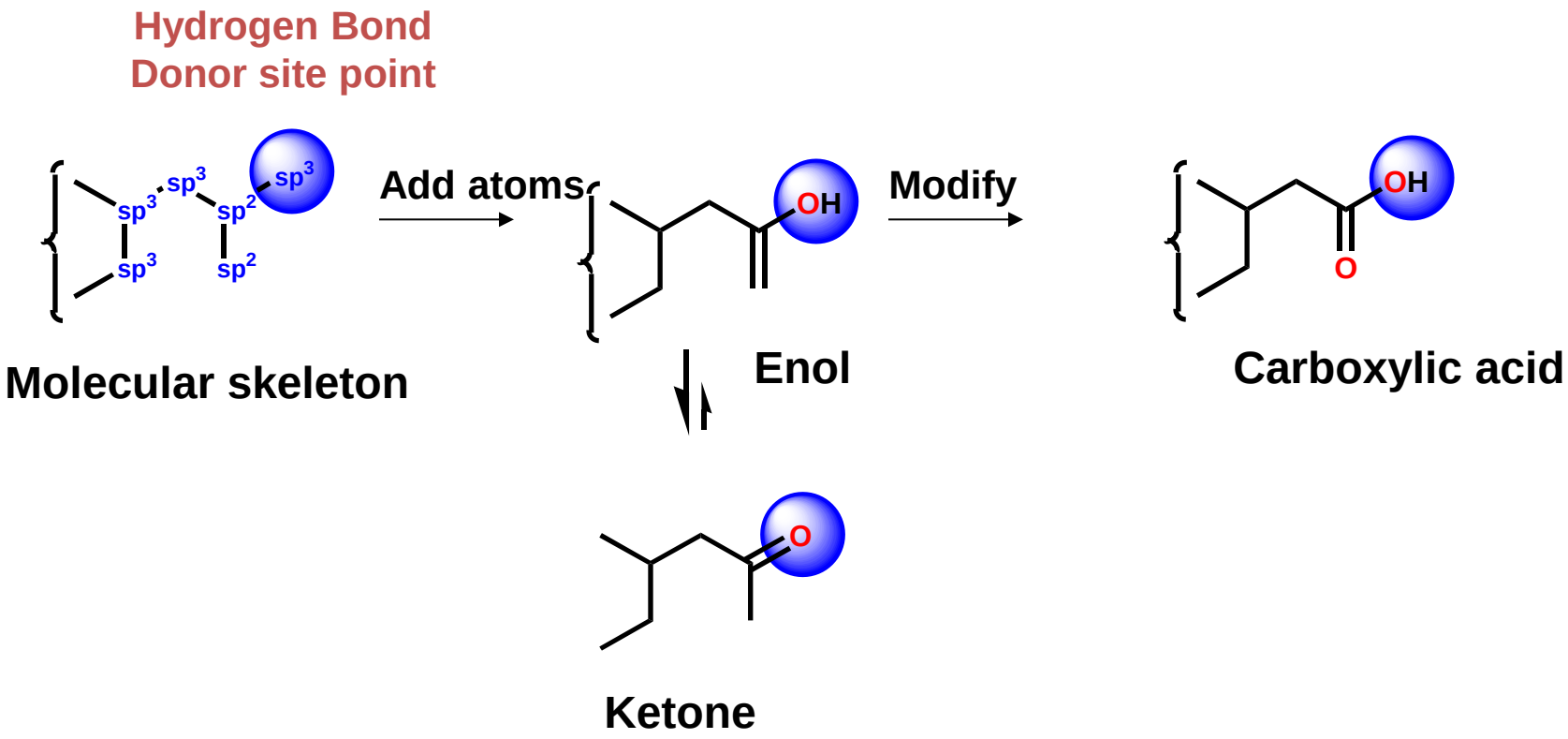
“Molecular skeleton”

Match to real molecules

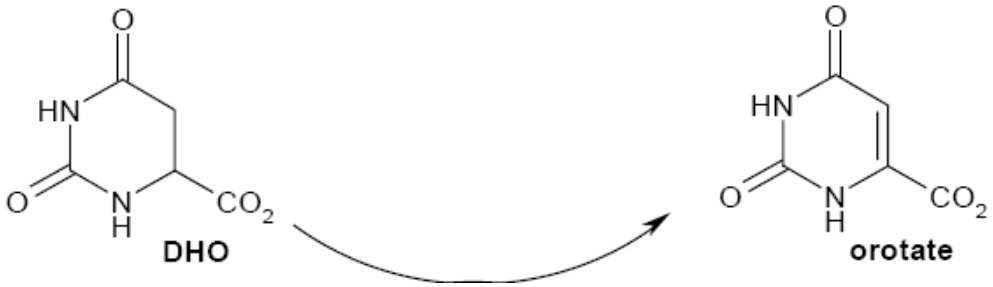


A molecular structure “solution”

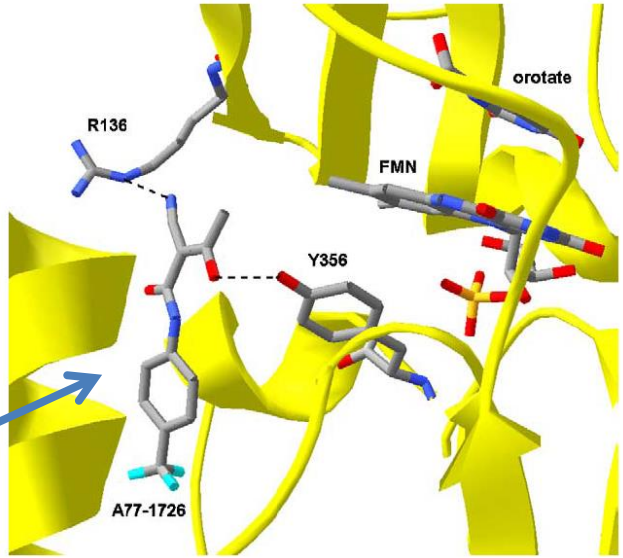
In-built rules to modify chemically unrealistic features



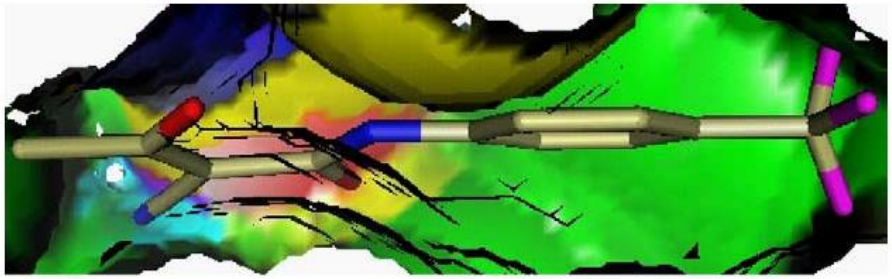
Example: de novo design of DHODH inhibitors with SPROUT



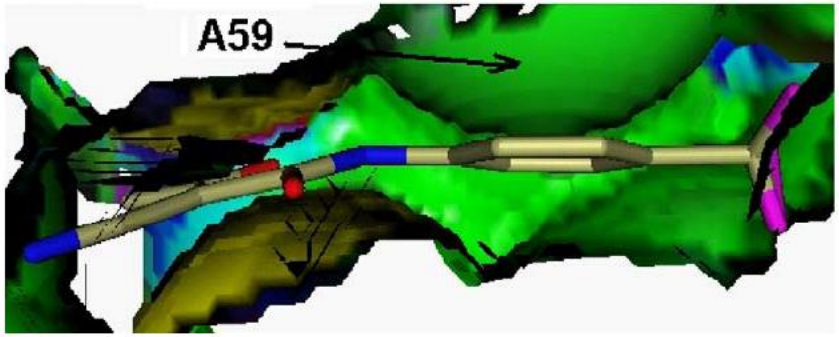
dihydroorate dehydrogenase (DHODH)



ubiquinone channel



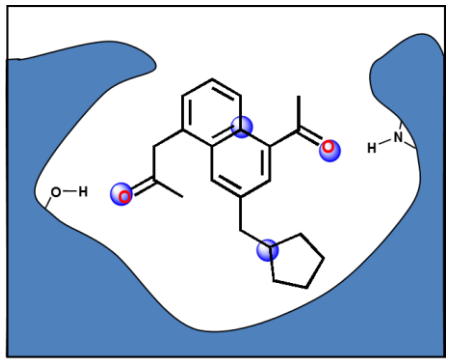
P. falciparum DHODH



Human DHODH

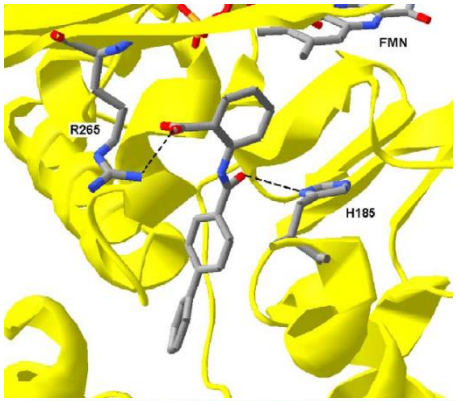
Bioorg. Med. Chem. Lett. 16 (2006) 88–92

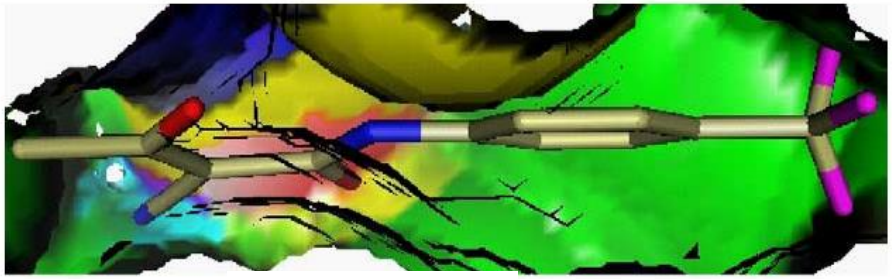
Example: de novo design of DHODH inhibitors with SPROUT



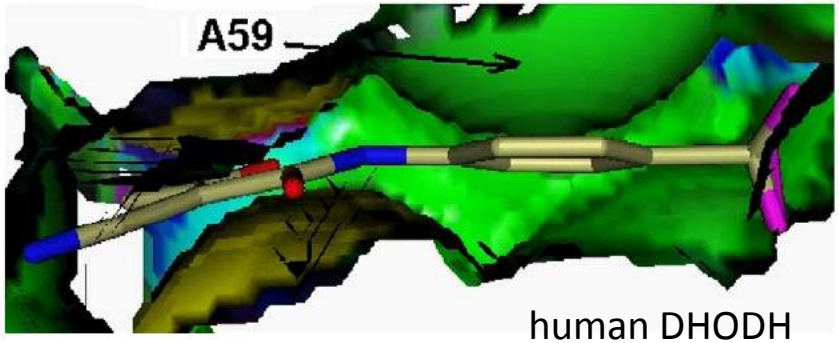
Molecular structure

Of 20 templates made by SPROUT, best was amides of anthranilic acid

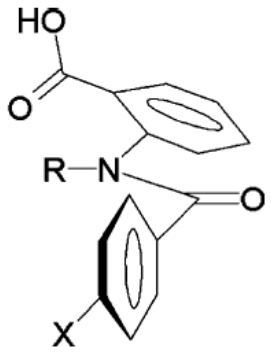




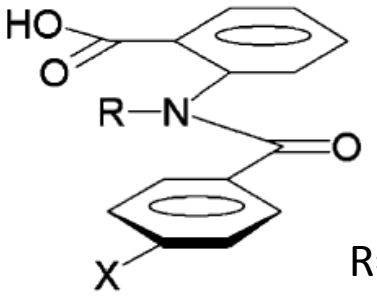
PfDHODH



human DHODH

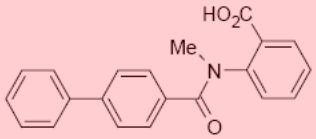
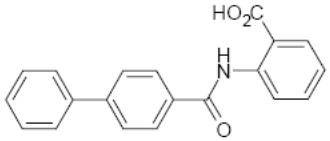
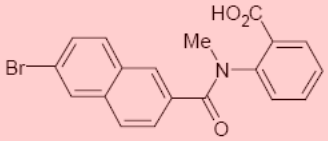
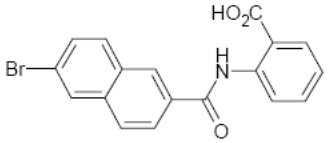
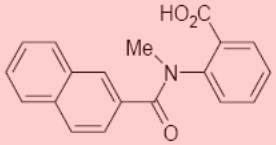
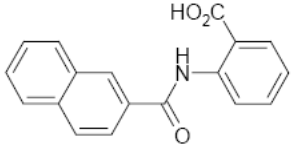


R=Me, twisted



R=H, flat

Example: de novo design of DHODH inhibitors with SPROUT

Entry	Structure	IC_{50} (Pf)	IC_{50} (h)	K_i^{app} (Pf) μM	K_i^{app} (h)
1		42.6 (4.6)	>200	4.9	n.a.
2		153.5 (13.2)	5.0 (1.6)	17.7	0.7
3		93.4 (6.4)	>200	10.8	n.a.
4		142.6 (14.2)	8.4 (2.7)	16.4	1.1
5		>200	>200	n.a.	n.a.
6		>200	13.8 (3.3)	n.a.	1.8

Example de novo design programs

Program	Method	year
GRID	Active Site Analysis	1985
MCSS	Active Site Analysis	1991
LUDI	Site Point	1992
SPROUT	Site Point	1993
CAVEAT	Fragment Connection	1994
SmoG	Sequential Build-up	1996
SkelGen	Fragment Connection	1997
BREED	Fragment Reconnection	2004
LIQUID	Site Point	2007

5.3. De novo design

Advantages

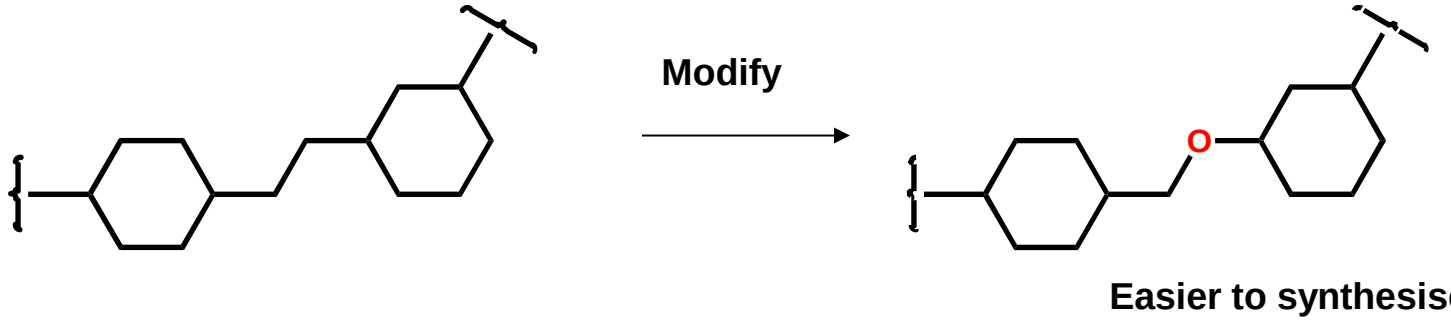
- Can produce completely new ideas for design
- Some methods do not need a receptor structure
- Growing track record, especially in “lead hopping”

Disadvantages

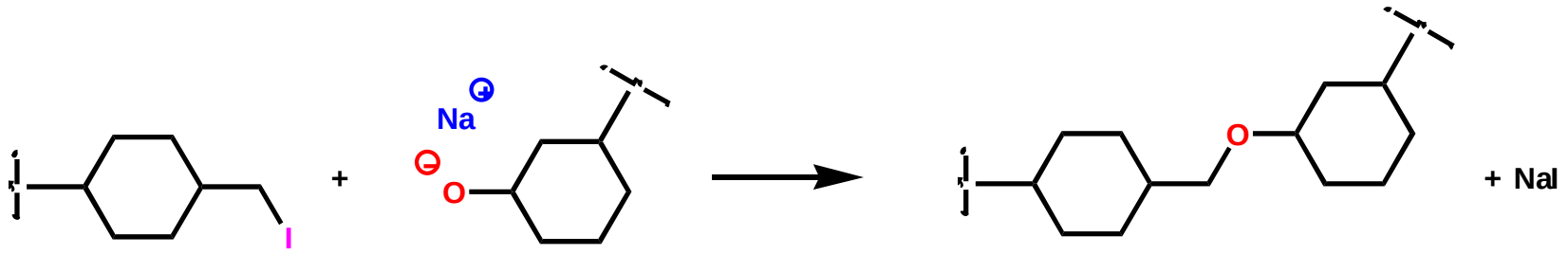
- “Wrong“ placement of start fragment(s)
- May be built into active site in an unreasonably high energy conformation
- Lack synthetic feasibility
  - avoid chemical instability, multiple chiral centres, large fused-ring systems
- Often requires a receptor structure

Problem #1: synthetic feasibility

- Solution: Some programs can modify structure to improve synthetic feasibility (eg. CAESA)

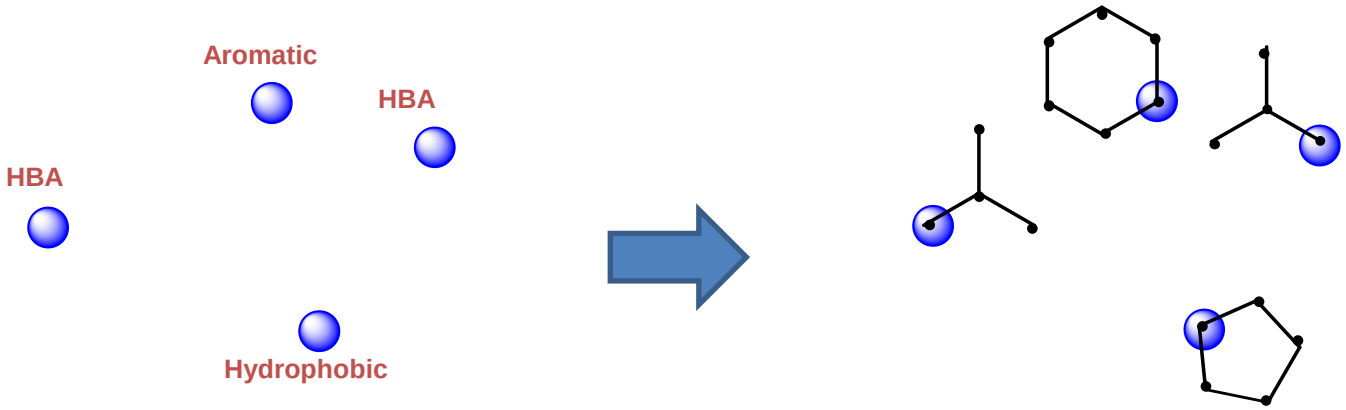


Synthesis



Problem #2: No receptor structure

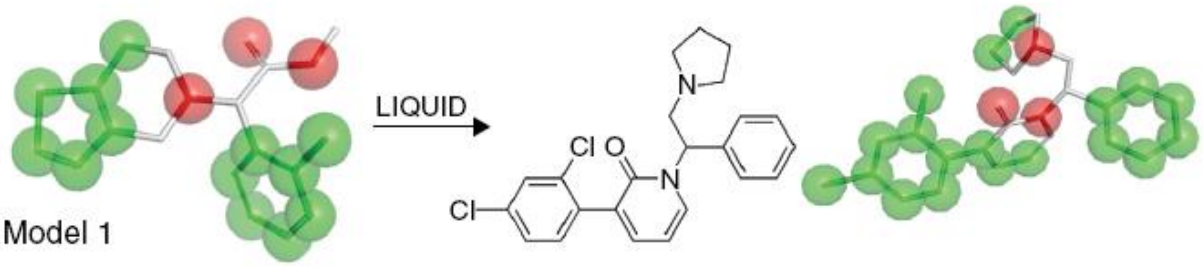
- Solution #1: Use ligands to generate pharmacophore
  - can use as input sites for some site connection methods, eg. SPROUT



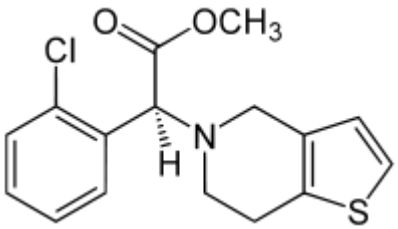
  - new program LIQUID uses “fuzzy” pharmacophore…

LIQUID

- ligand-only de novo design method
  - uses “fuzzy” pharmacophore to generate different molecules

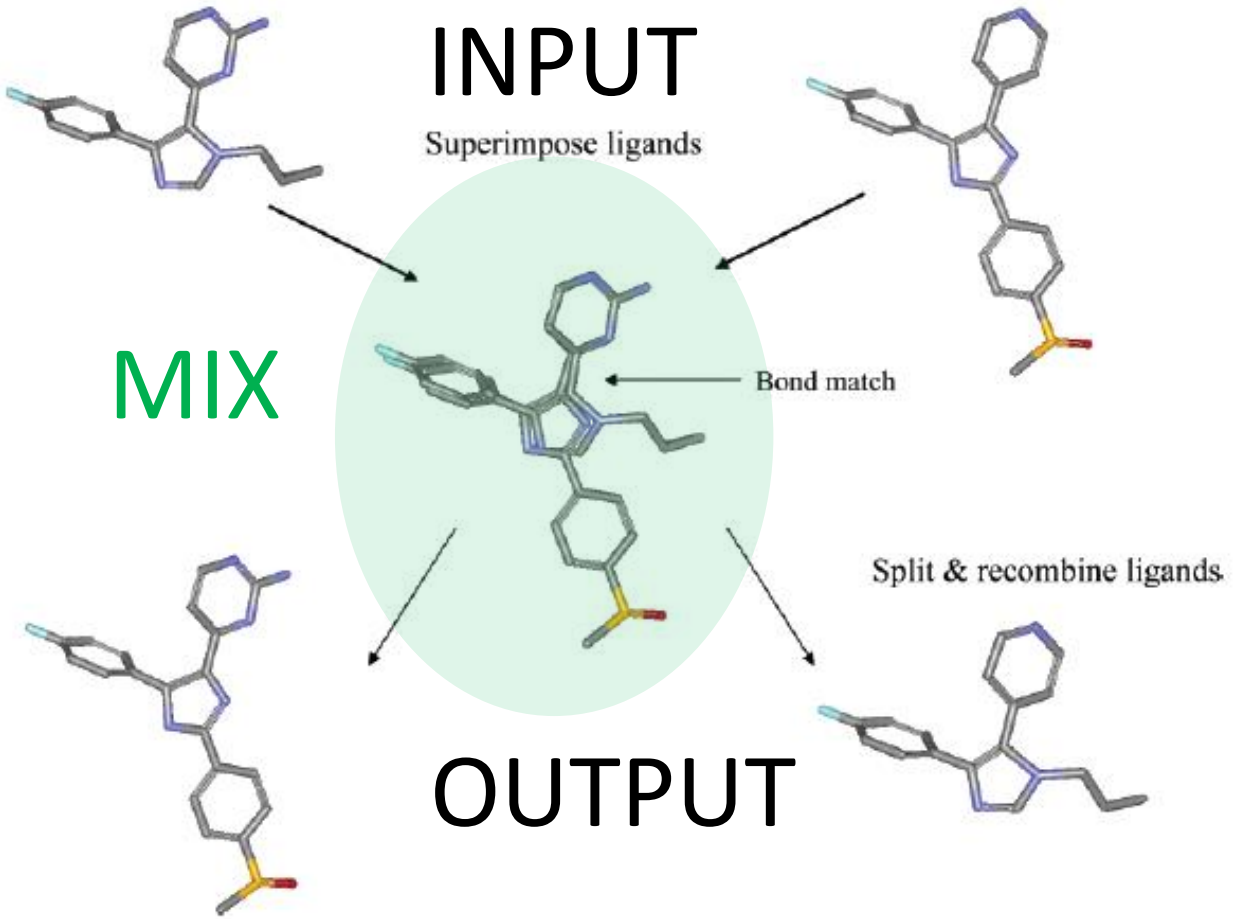


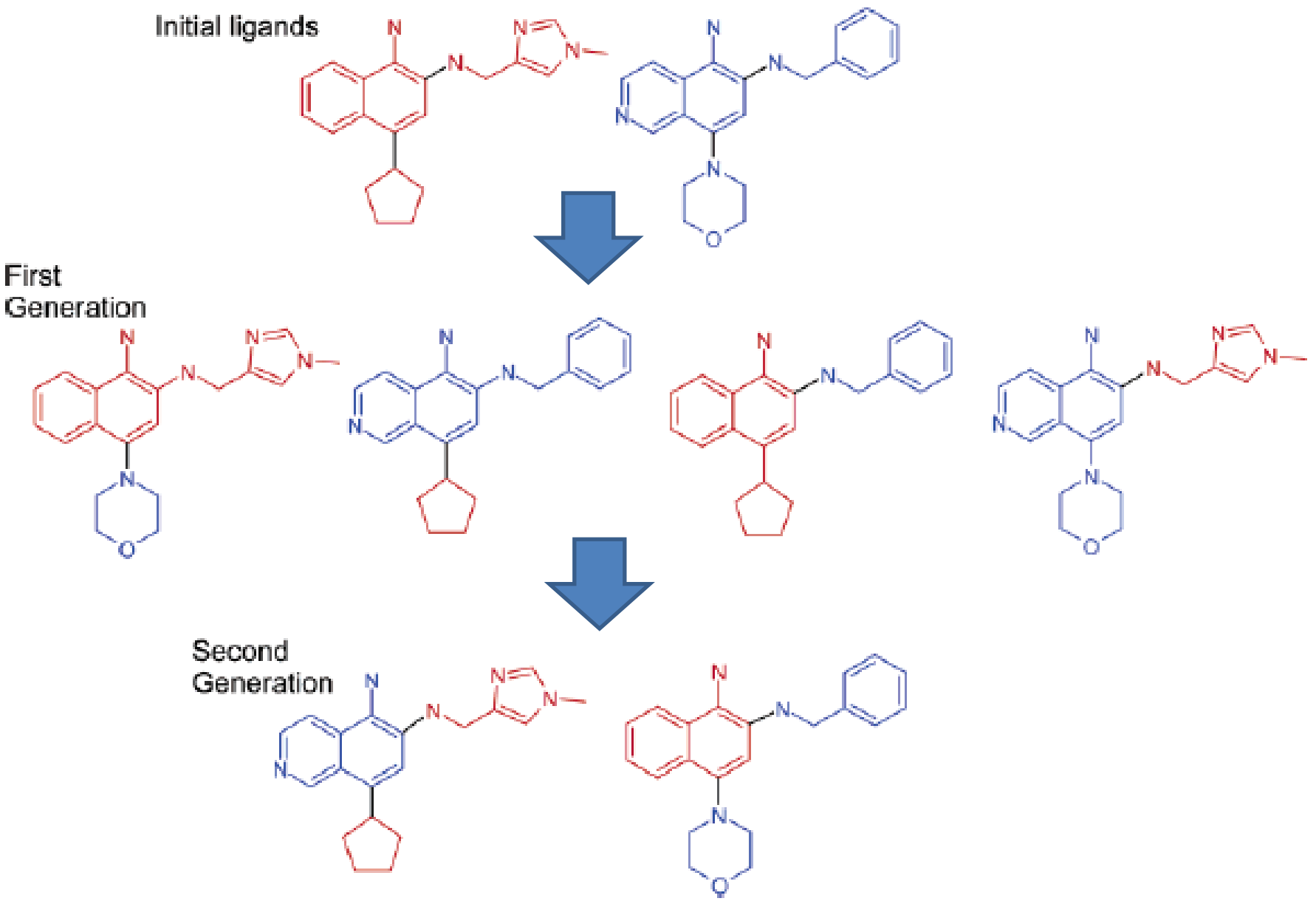
clopidogrel



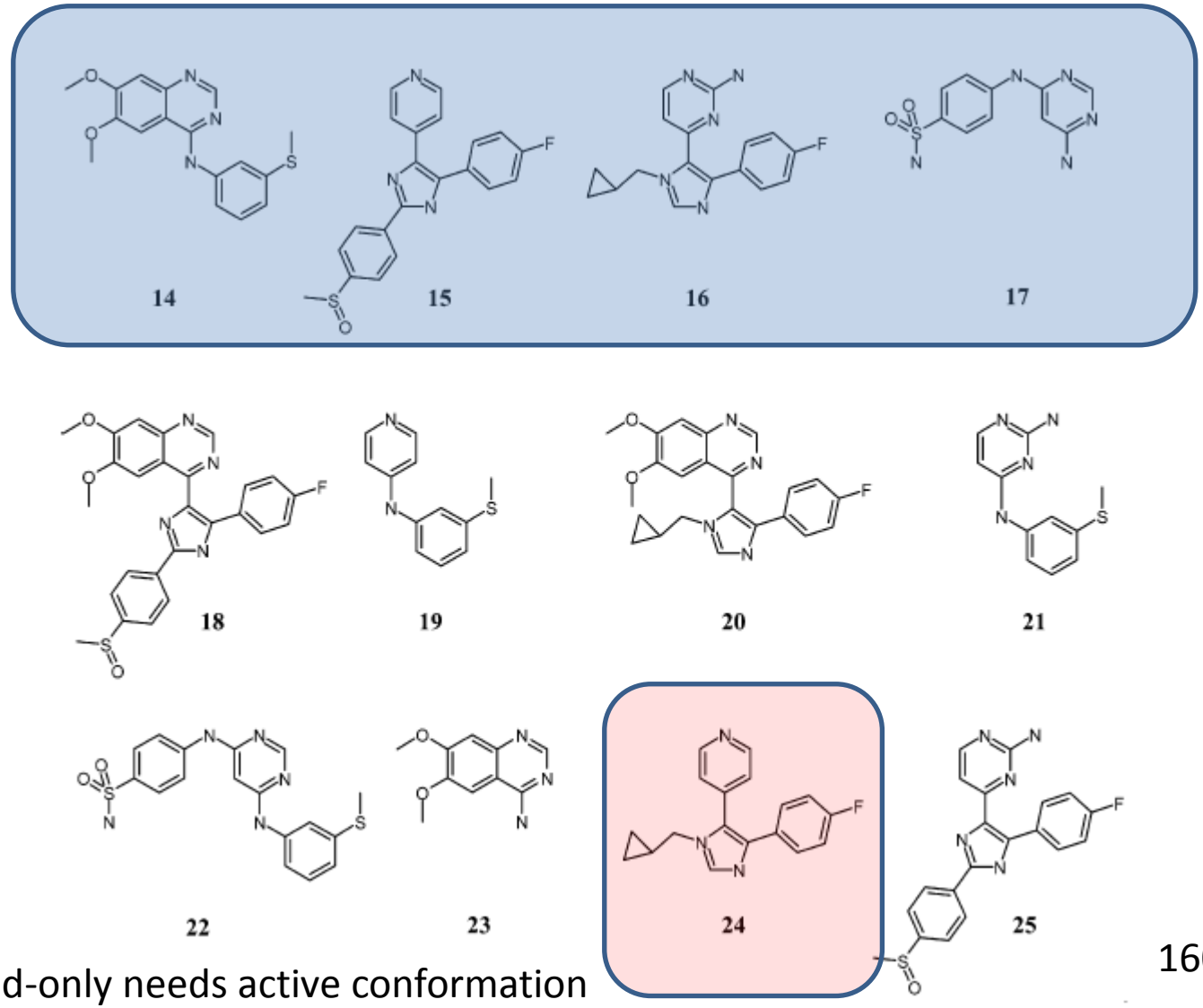
Solution #2:

- BREED program: breed together known active ligands to produce new ligands, “cut-and-paste”




Initial ligands
First Generation
Second Generation

Kinase inhibitors



Recall: Ligand-only needs active conformation

160 nM (JNK3)

De novo design today

- No longer Cinderella
  - growing track record
  - synthetic feasibility of solutions improved
  - can take account of active site flexibility
  - success in scaffold-hopping (with or without receptor structure)
    - but need to ensure we have the active ligand conformation

Target	Ligand affinity	De novo program(s)	Reference
thrombin	10 nM	LUDI	Bohm et al. JCAMD 1999, 13:51
CDK-4	<1 μM	LEGEND, LUDI, LeapFrog	Honma et al. J Med Chem 44:4628
hCB-1 cannaboid receptor	300 nM	TOPAS	Roger-Evans et al. QSAR Comb Sci 23:426
lansoterol 14a-demethylase	40 μg/mL	MCSS + LUDI	Ji et al. 2003, J Med Chem 46:474
HIV-1 RT	10 nM	BOMB	Jorg 2006
HIV-1 RT	4 μM	SYNOPSIS	Vinkers et al. J Med Chem 2003, 46: 2765
HIV-1 protease	42 nM	BREED	Pierce et al. J Med Chem 2004, 47: 2768
Kv1.5 potassium channel	< 1 μM	TOPAS	Ang. Chem. Int. Ed. 2000, 39:4130

5.3.2. De novo design “tips”

- don’t bind a ligand too tightly/fill binding site
  - may be experimental error in crystal structure
  - different binding modes from predicted
  - space needed for drug optimisation eg. PK
- differing opinions on whether rigid or flexible molecules are better
  - flexible ones allow for error in prediction
  - rigid ones bind better due to pre-paid entropy

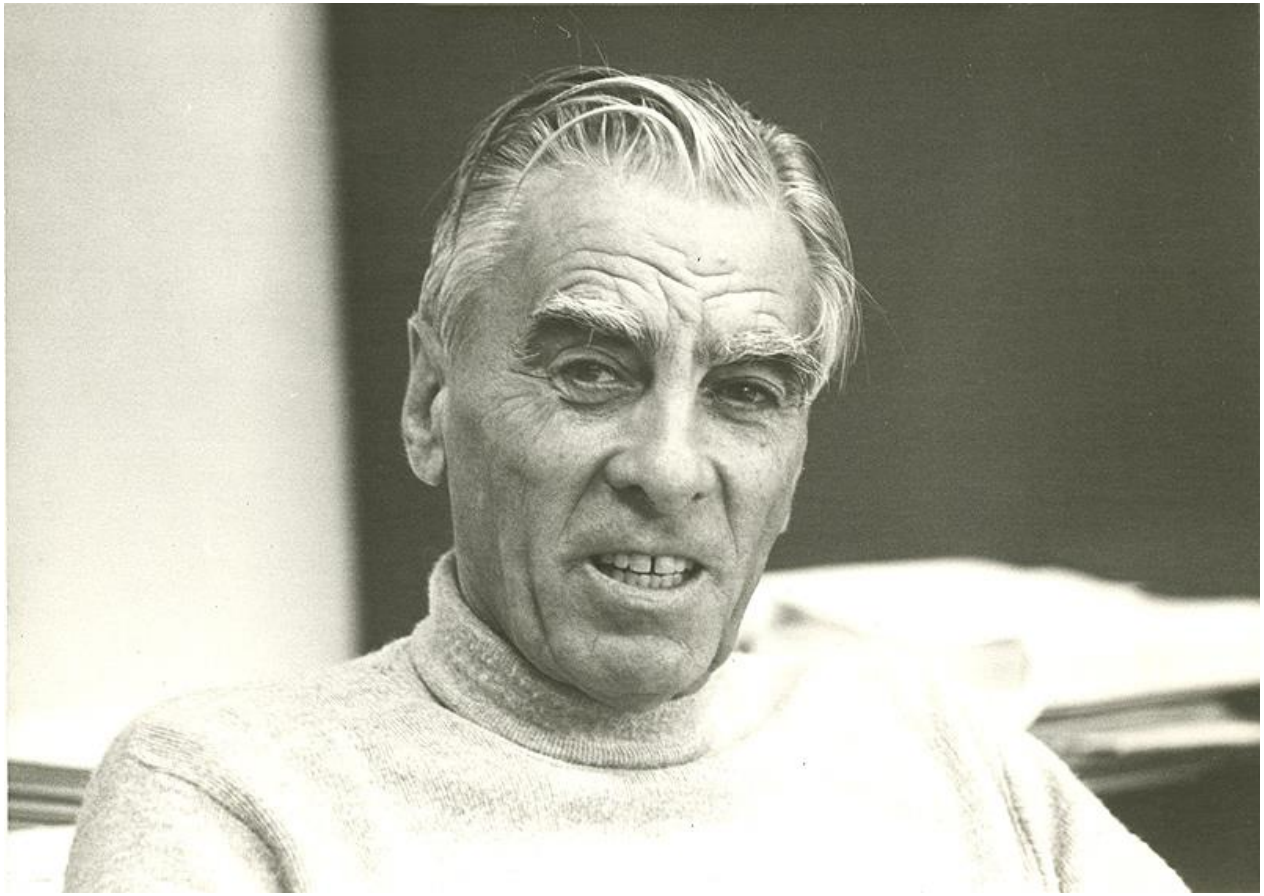
Go deeper: **Patrick, Introduction to Medicinal Chemistry, Ch. 17. Section 17.15**

Design in the dark

6. 3D QSAR

6. Quantitative Structure-Activity Relationships (QSAR)

- also known as Hansch Analysis



6.1. QSAR

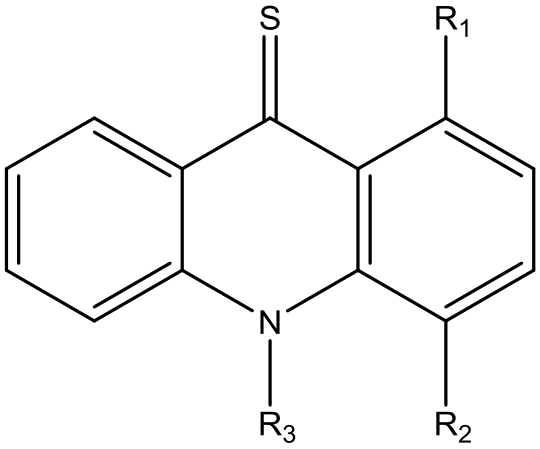
= *mathematical* relationship between:

and

- use (*linear or nonlinear) regression analysis* to derive this mathematical relationship

- DNA binding (K_d) of thioacridone derivatives
  - potential anticancer compounds

R_1	R_3	R_2	K_d	$\log(1/K_d)$	logP
$NH(CH_2)_2N(CH_3)_2$	CH_3	Cl	0.011	1.96	3.51
$NH(CH_2)_2N(CH_3)_2$	H	Cl	0.012	4.42	3.27
$N(CH_2CH_2Cl)_2$	H	H	0.005	5.30	4.92
$N(CH_2CH_2Cl)_2$	H	Cl	0.003	5.81	5.48
$N(CH_2CH_2Cl)_2$	H	CH_3	0.006	5.12	5.41



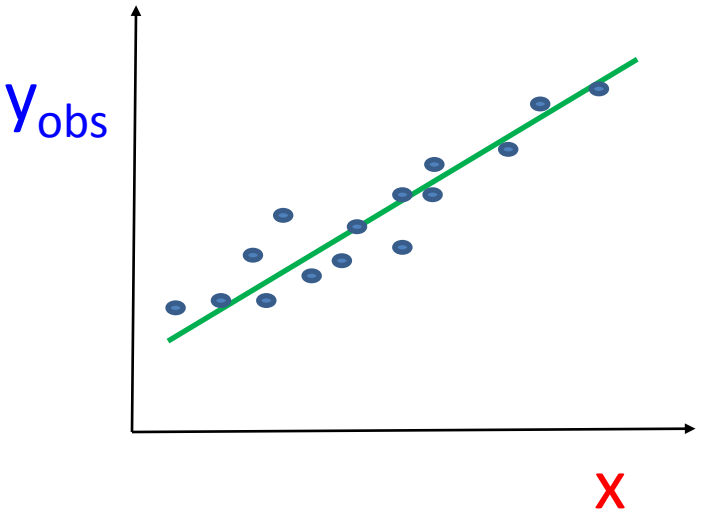
$\log(1/K_d)$ = potency of compounds

Bioorganic & Medicinal Chemistry 13 (2005) 689–698

Linear regression analysis

- For one variable (descriptor) = *linear*:

$y_{calc} = m\,x + c$

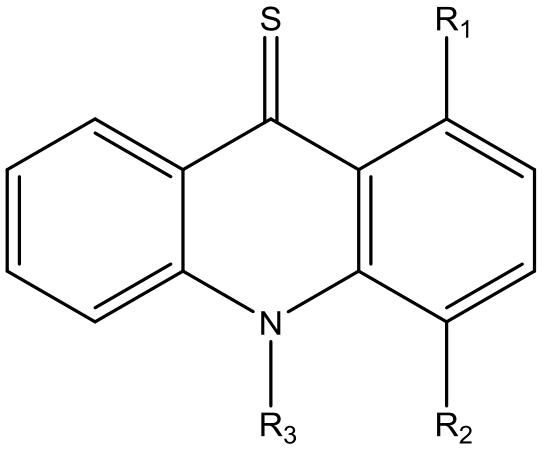


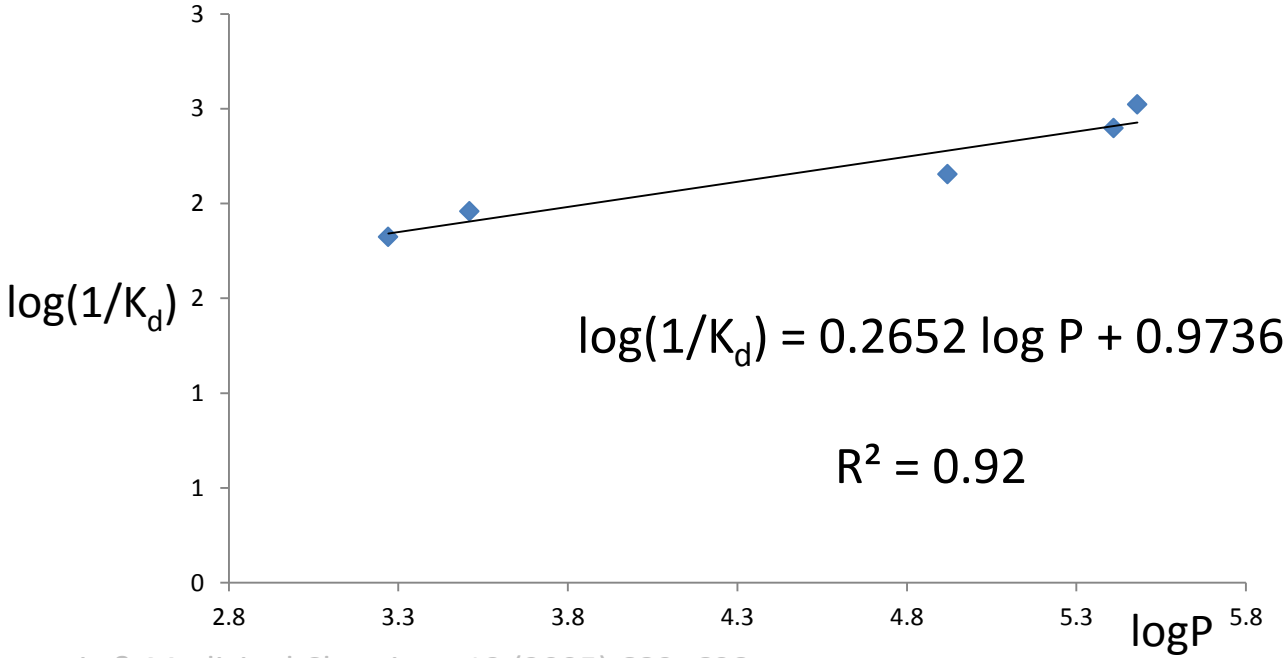
fit {*m*,*c*} such that y_{calc} gives best estimate of y_{obs}

cf. MS Excel "trendline"
(if all data fits perfectly on the line,
then square of correlation coefficient, $R^2 = 1$)

DNA binding (K_d) of thioacridone derivatives

R_1	R_3	R_2	K_d	$\log(1/K_d)$	logP
$NH(CH_2)_2N(CH_3)_2$	CH_3	Cl	0.011	1.96	3.51
$NH(CH_2)_2N(CH_3)_2$	H	Cl	0.015	1.82	3.27
$N(CH_2CH_2Cl)_2$	H	H	0.007	2.15	4.92
$N(CH_2CH_2Cl)_2$	H	Cl	0.003	2.52	5.48
$N(CH_2CH_2Cl)_2$	H	CH_3	0.004	2.40	5.41



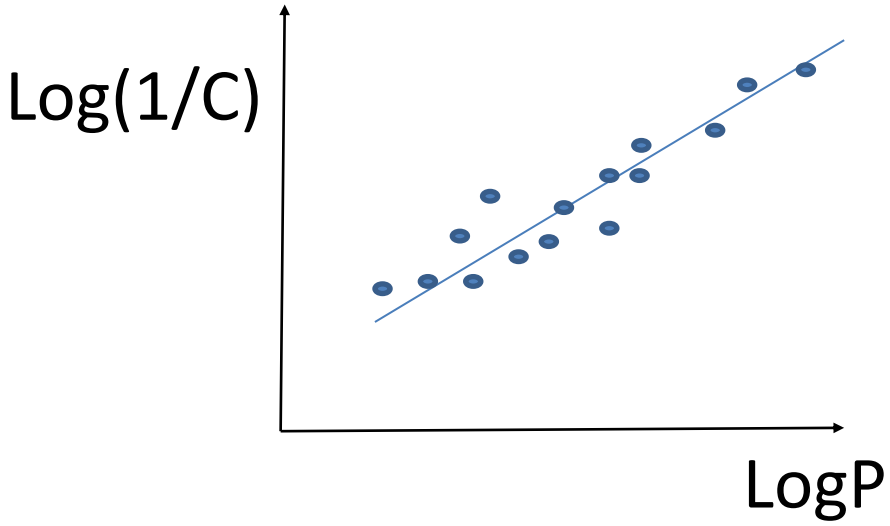


Bioorganic & Medicinal Chemistry 13 (2005) 689–698

Linear relationship

- Toxicity of alcohols to red spiders (Hansch, 1971):

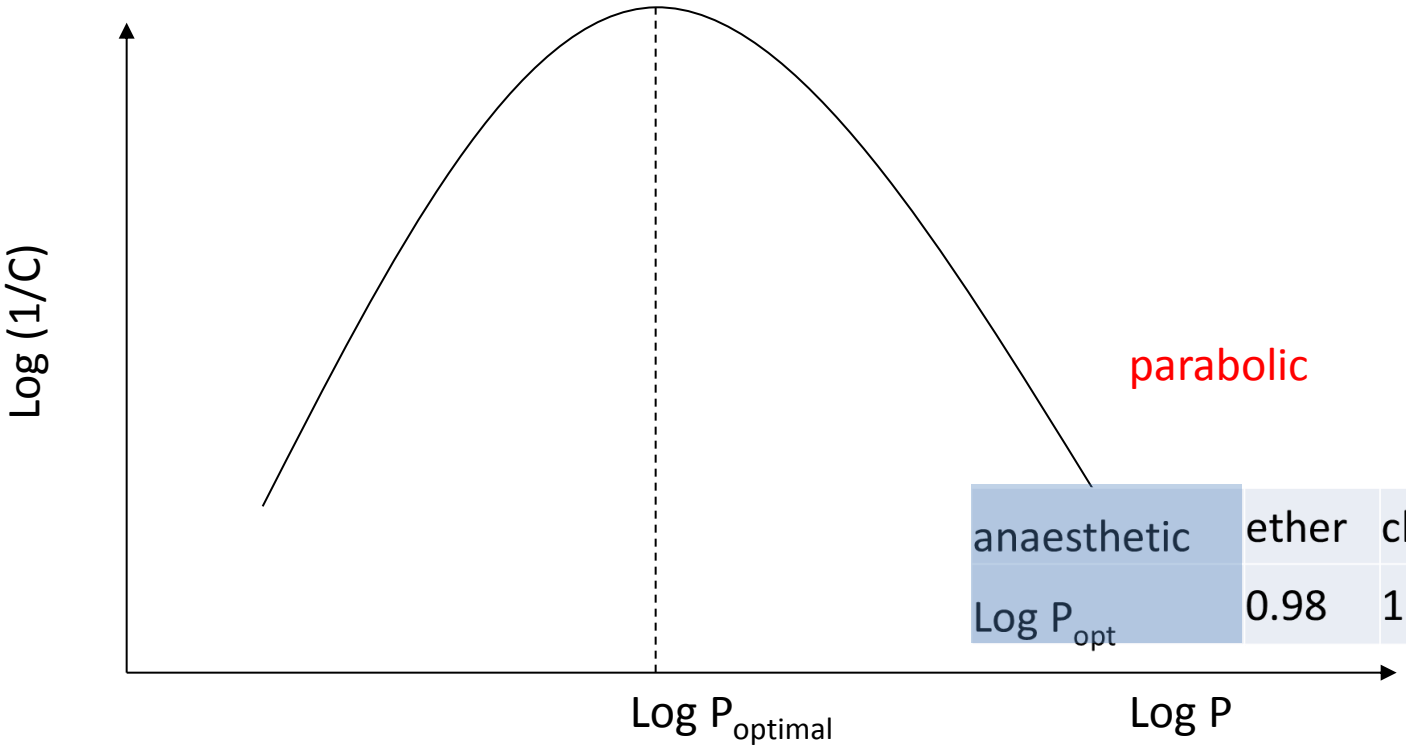
$\text{Log}(1/C) = 0.69\ \text{logP} + 0.16 \qquad R^2 = 0.958,\ n = 14$



Non-linear relationship

- eg. anaesthetic ethers

$$\text{Log}(1/C) = -0.22(\log P)^2 + 1.04 \log P + 2.16$$

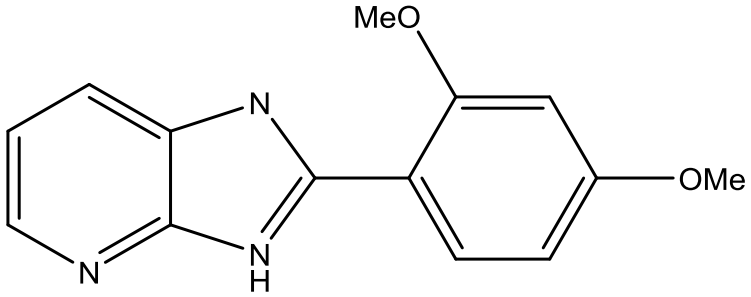


parabolic

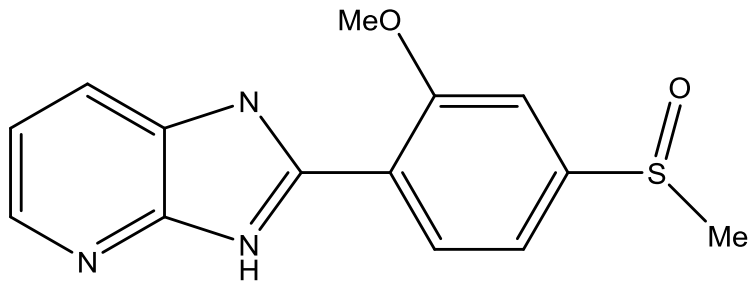
anaesthetic	ether	chloroform	halothane
Log P_{opt}	0.98	1.97	2.30

Optimal logP values

- Drugs with Log P values ~2
  - enter CNS efficiently
    - *e.g*. barbiturate sedatives have log P values ~2
  - can avoid CNS side effects by lowering Log P



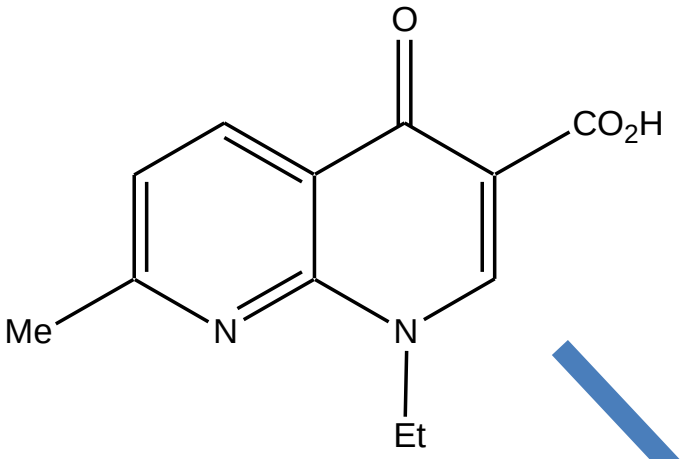
Log P = 2.5



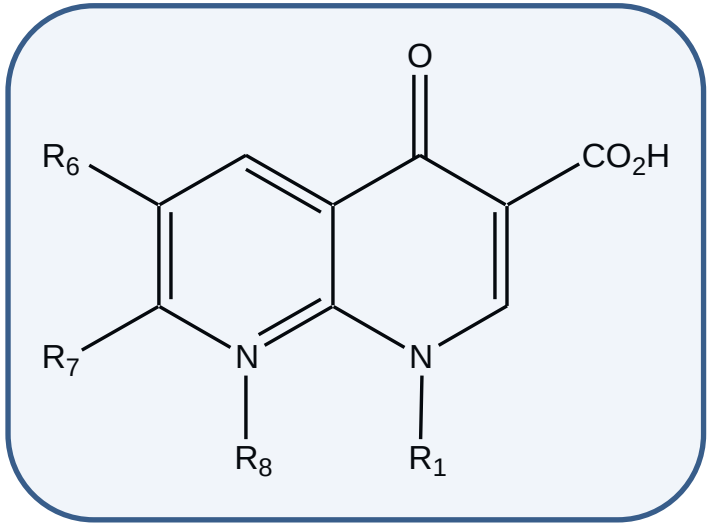
Sulfazole, Log P = 1.2

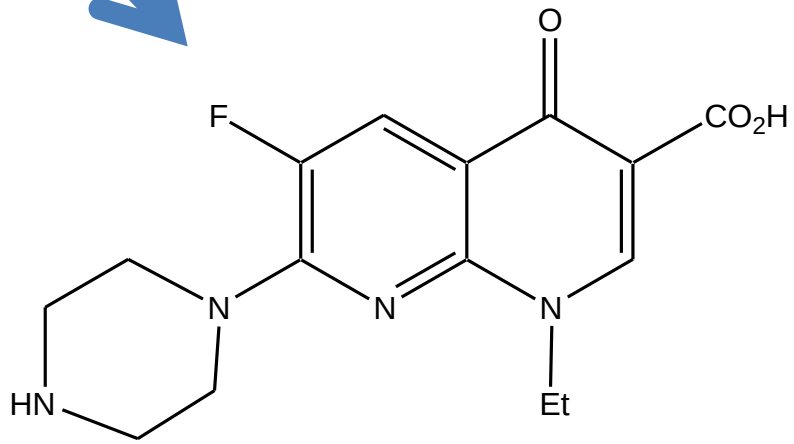
Go deeper: **Patrick, Introduction to Medicinal Chemistry, Ch. 18.**

QSAR works: Norfloxacin



nalidixic acid





norfloxacin

…but limited by congeneric series

6.2. Three-dimensional (3D) QSAR

- derive mathematical relationship between a molecule's biological/pharmacological activity and its 3D structure

- properties calculated for whole molecule not just for substituents (unlike 2D QSAR)
  - properties known as *molecular fields:*
    (i) **steric field**: defines shape of molecule
    (ii) **electrostatic field**:
    defines electronic character of molecule
- assumes non-covalent steric and electrostatic interactions with the enzyme

see G. L. Patrick, Ch. 18

6.2.1. Comparative Molecular Field Analysis (CoMFA)

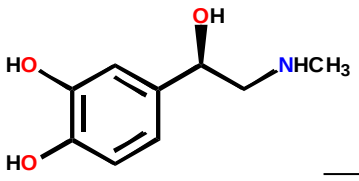
- CoMFA = most common 3D QSAR method
- need a set of molecules with desired pharmacological activity

1. each molecule built on computer using molecular modelling software
2. active conformation identified by energy minimisation
3. active pharmacophore defined
4. each molecule fitted in turn into a lattice of grid points in the same relative position and orientation

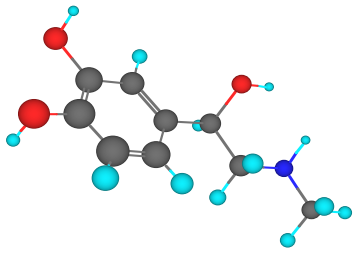
6.2.1. Comparative Molecular Field Analysis (CoMFA)

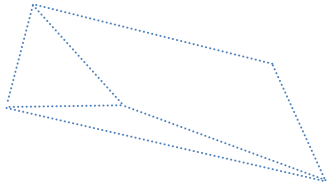
- CoMFA = most common 3D QSAR method
- need a set of molecules with desired pharmacological activity

1. each molecule built on computer using molecular modelling software
2. active conformation identified by energy minimisation
3. active pharmacophore defined
4. each molecule fitted into a lattice of grid points in the same relative position and orientation



Build 3D model



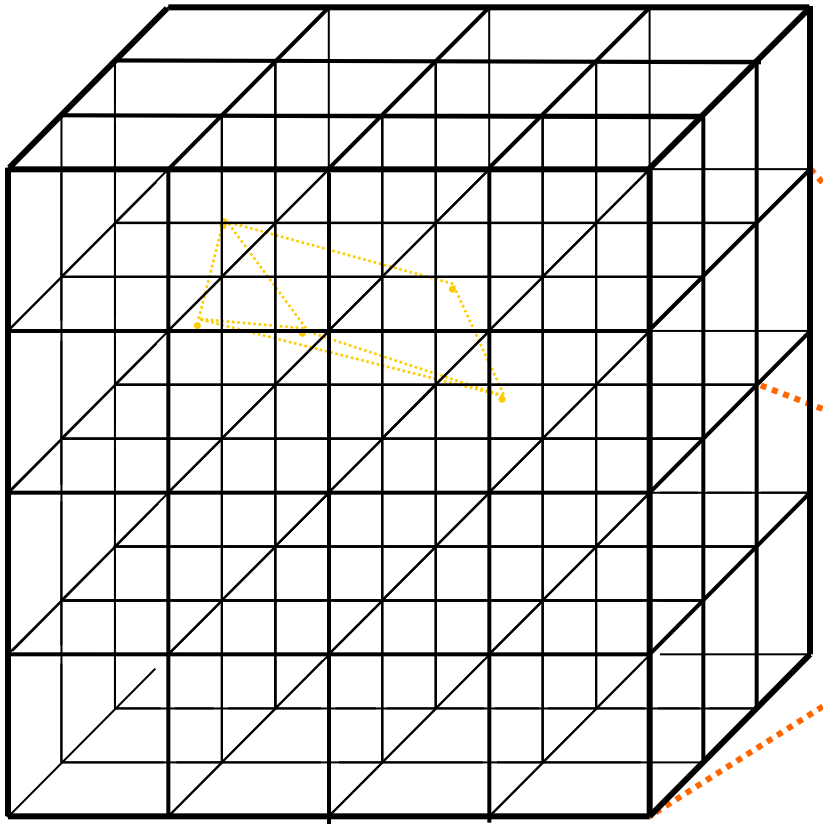


Patrick/OUP

Active conformation

Define pharmacophore

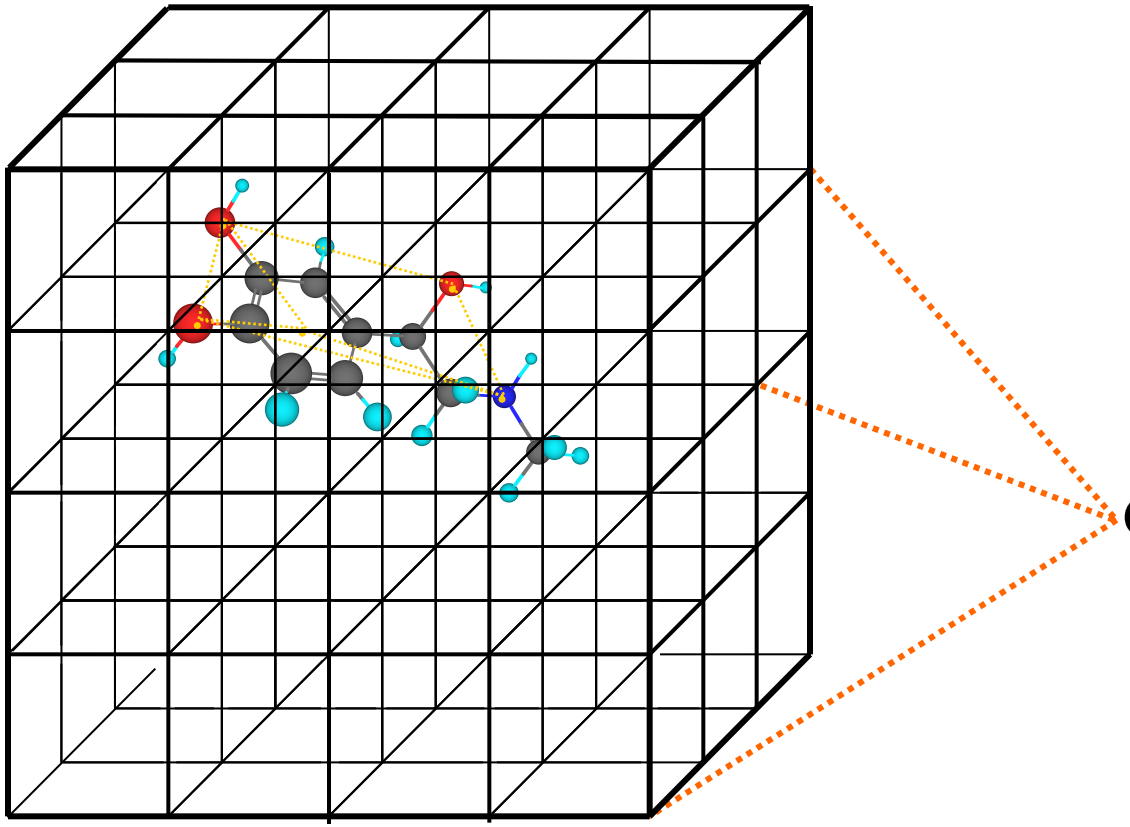
•Place the pharmacophore into a lattice of grid points



Grid points

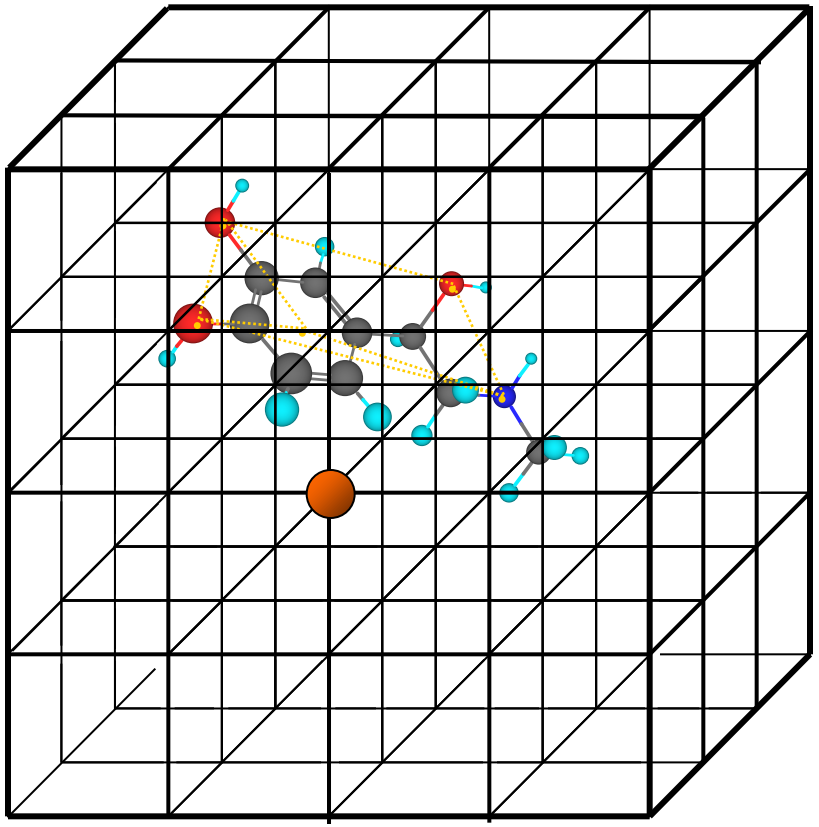
•Each grid point defines a point in space

•Position molecule to match the pharmacophore



•Each grid point defines a point in space

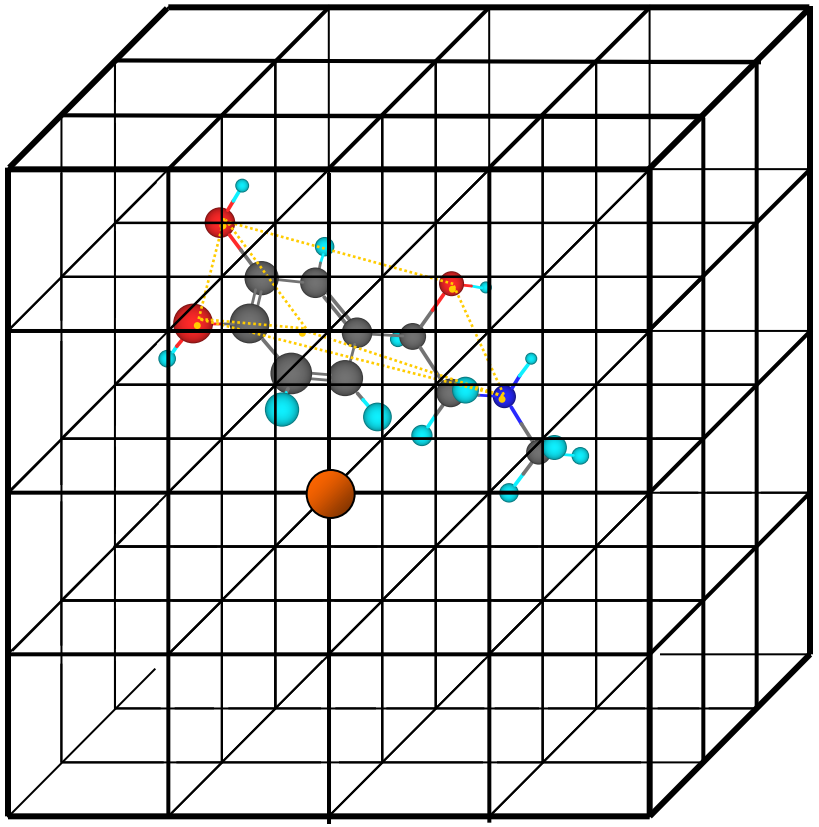
•A probe atom is placed at each grid point in turn



Probe atom

•Probe atom = a proton or sp^3 hybridised carbocation

•A probe atom is placed at each grid point in turn



Probe atom

•Measure the steric or electrostatic interaction of the probe atom with the molecule at each grid point

5. Steric and electrostatic fields around molecule measured and defined

- probe atom (H^+ or sp^3 hybridised carbocation) placed at each grid point in turn

- calculate steric and electrostatic interactions between probe and each atom on molecule

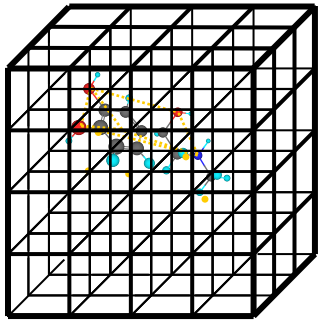
- closer probe atom to molecule, the higher the steric (or electrostatic) interaction

- grid points with equal interaction energy connected by contour lines to define a field

-quantitatively related to biological activities as in traditional QSAR

  - 1000s of field values generated as potential parameters to use in obtaining a best fit
  - to reduce the number of parameters in the fit, analyse using a “partial least squares” (PLS) approach

- identify steric and electrostatic effects on biological activity at specific points in 3D space around the molecule; visualise mathematical relationship:



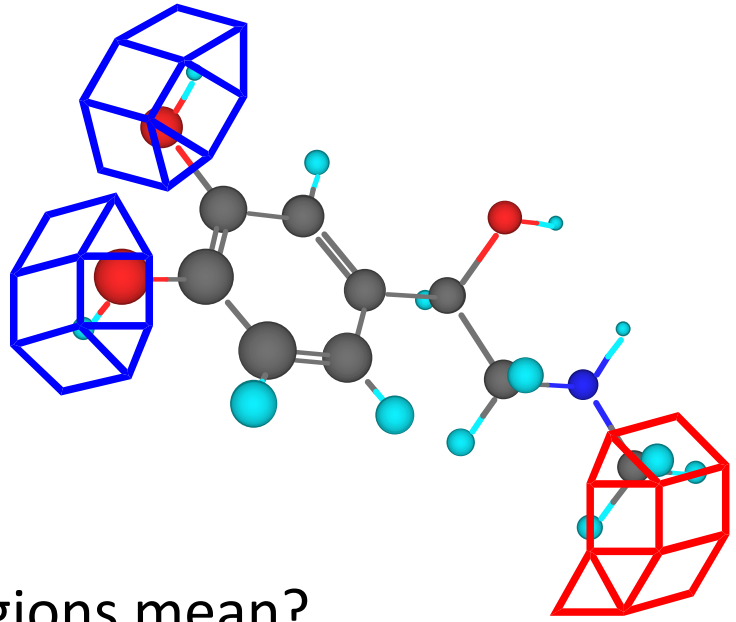
Tabulate fields for each compound at each grid point

Compound	Biological activity	Steric fields (S) at grid points (001-998)					Electrostatic fields (E) at grid points (001-998)				
		S001	S002	S003	S004	S005 etc	E001	E002	E003	E004	E005 etc
1	5.1										
2	6.8										
3	5.3										
4	6.4										
5	6.1										

Partial least squares analysis (PLS)

QSAR equation Activity = aS001 + bS002 +……..mS998 + nE001 +…….+yE998 + z

- identify steric and electrostatic effects on biological activity at specific points in 3D space around the molecule
- visualise mathematical relationship:

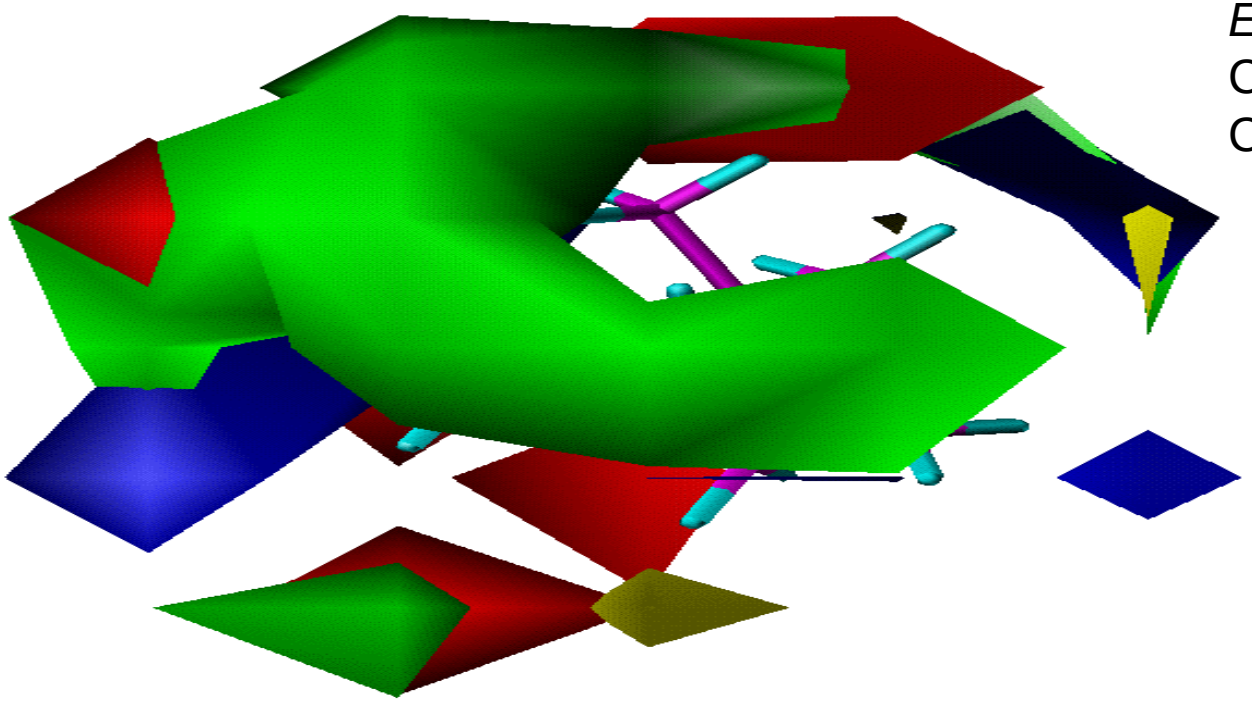


- What do regions mean?

6.2.2. Understanding a 3D QSAR analysis: coloured representation

Maps of electrostatic fields:
BLUE - positively charged groups favourable
RED - negatively charges groups favourable

Maps of steric fields:
GREEN – space-filling areas for best binding
YELLOW – space-conflicting areas



Example: 3D-QSAR of $CYP450_{cam}$ ligands with CoMFA

Understanding a 3D QSAR analysis: black and white representation

(a) steric fields

(i) *solid lines* $\Rightarrow$ bulky groups favourable for activity

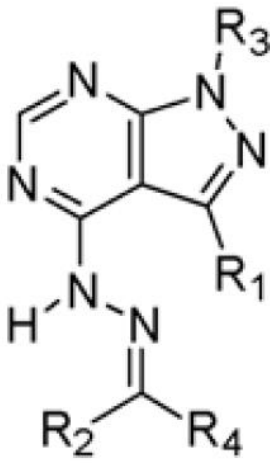
(ii) *dashed lines* $\Rightarrow$ bulky groups unfavourable for activity

(b) electrostatic fields

(i) *solid lines* represents regions where +ve groups would improve activity

(ii) *dashed lines* where –ve groups would improve activity

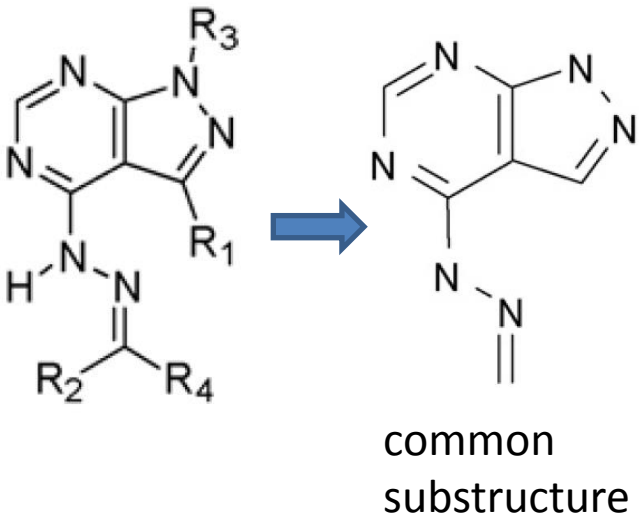
6.2.3. Case study: glycogen synthase kinase (GSK-3β) inhibitors



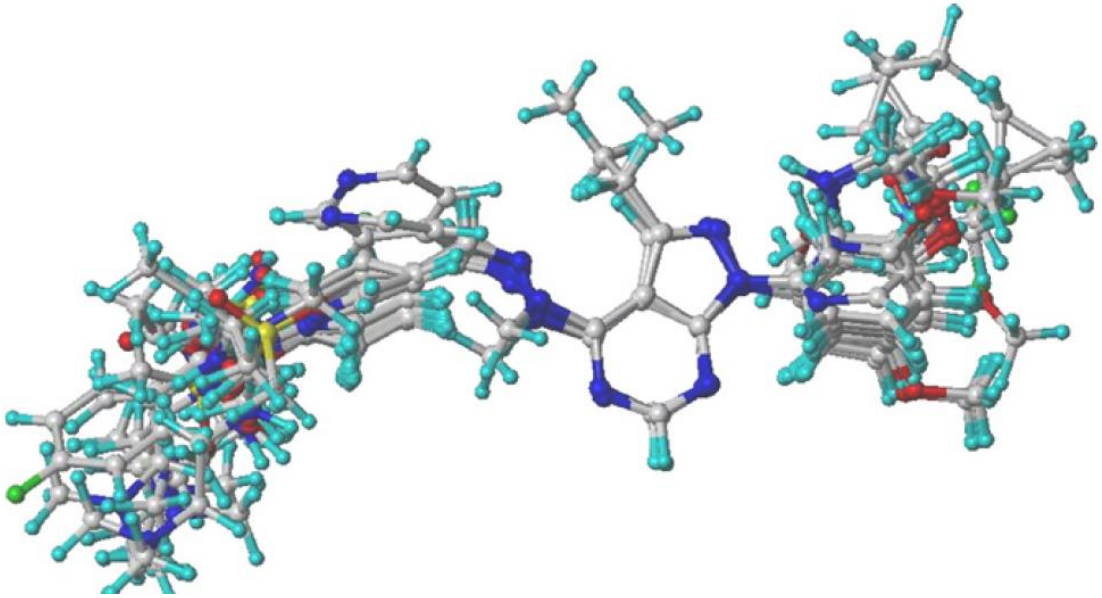
Cpd	R1	R2	R3	R4	Actual pIC_{50}
1	H	H	–Ph-2-OMe	–4-Pyridyl	5.60
2	H	H	–Ph-4-OMe	–4-Pyridyl	6.00
3	H	H	–Ph-2-OEt	–4-Pyridyl	7.00
4	H	H	–Ph-2-OCF3	–4-Pyridyl	6.50
5	H	H	–Ph-2-NHnPr	–4-Pyridyl	5.40
6	H	H	–Ph-2-NH(CH_2)cyclopropyl	–4-Pyridyl	5.70
7	H	H	–Ph-2-NHAc	–4-Pyridyl	6.80
8	H	H	–Ph-3-OMe	–Ph-4-F	8.10
9	H	H	–Ph-3-OMe	–Ph-4-SO_2Me	8.60
10	H	H	–Ph-3-OMe	–Ph-4-CO_2H	7.50
11	H	H	–Ph	–Ph-(3-OMe,4-OH)	7.00
12	H	H	–Ph-2-NH(CO)*n*Pr	–4-Pyridyl	6.20
13	H	H	–Ph-2-F	–4-Pyridyl	6.50
14	H	H	–3-pyridyl	–4-Pyridyl	7.50
15	H	H	–4-Pyridyl	–4-Pyridyl	8.50
16	H	H	–2-pyridyl-3-OMe	–4-Pyridyl	7.40
17	H	H	–2-thiazole	–4-Pyridyl	7.40

Journal of Molecular Graphics and Modelling 25 (2007) 885–895

- Align common pharmacophore/substructure

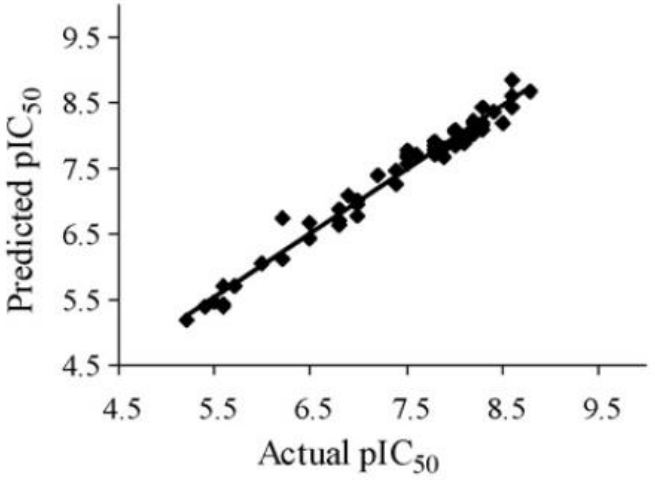


common substructure

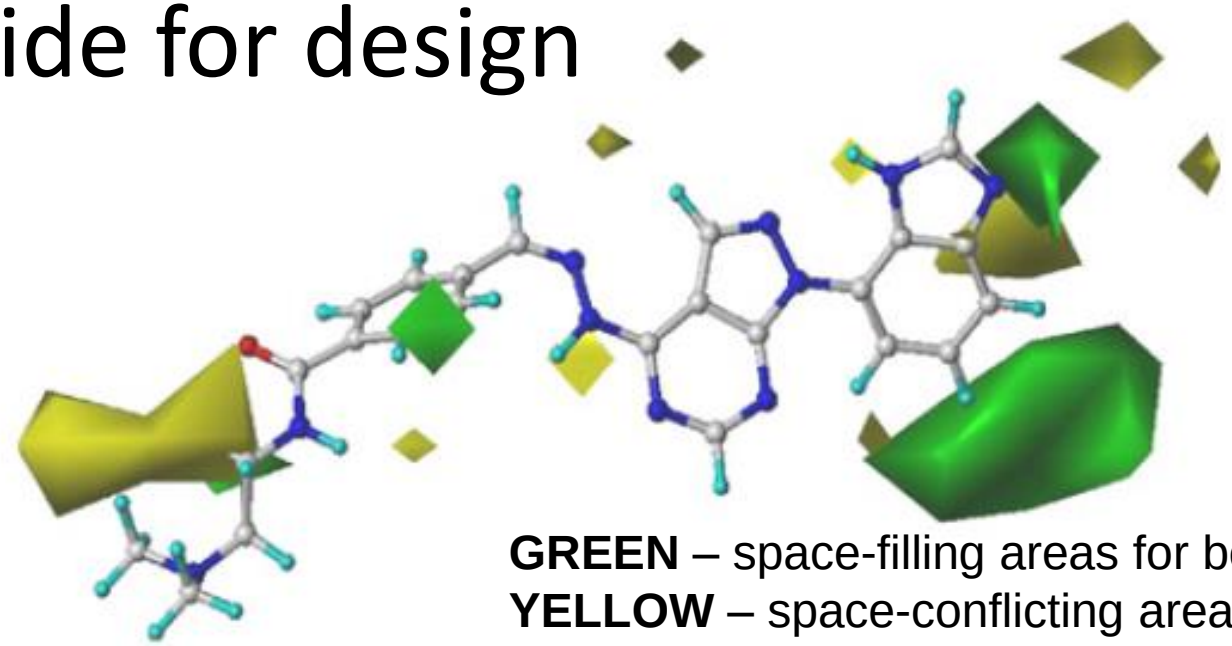


Perform fitting

Cpd	R1	R2	R3	R4	Actual pIC_{50}	Predicted pIC_{50}		Residuals	
						CoMFA	CoMSIA	CoMFA	CoMSIA
1	H	H	–Ph-2-OMe	–4-Pyridyl	5.60	5.69	5.85	–0.09	–0.25
2	H	H	–Ph-4-OMe	–4-Pyridyl	6.00	6.06	6.44	–0.06	–0.44
3	H	H	–Ph-2-OEt	–4-Pyridyl	7.00	7.02	6.88	–0.02	0.12
4	H	H	–Ph-2-OCF3	–4-Pyridyl	6.50	6.68	6.50	–0.18	0.00
5	H	H	–Ph-2-NHnPr	–4-Pyridyl	5.40	5.39	5.66	0.01	–0.26
6	H	H	–Ph-2-NH(CH_2)cyclopropyl	–4-Pyridyl	5.70	5.72	5.68	–0.02	0.02
7	H	H	–Ph-2-NHAc	–4-Pyridyl	6.80	6.72	6.95	0.08	–0.15
8	H	H	–Ph-3-OMe	–Ph-4-F	8.10	7.97	7.82	0.13	0.28
9	H	H	–Ph-3-OMe	–Ph-4-SO_2Me	8.60	8.87	8.74	–0.27	–0.14
10	H	H	–Ph-3-OMe	–Ph-4-CO_2H	7.50	7.79	7.67	–0.29	–0.17
11	H	H	–Ph	–Ph-(3-OMe,4-OH)	7.00	6.95	6.72	0.05	0.28
12	H	H	–Ph-2-NH(CO)*n*Pr	–4-Pyridyl	6.20	6.12	6.19	0.08	0.01
13	H	H	–Ph-2-F	–4-Pyridyl	6.50	6.44	6.57	0.06	–0.07
14	H	H	–3-pyridyl	–4-Pyridyl	7.50	7.73	7.76	–0.23	–0.26
15	H	H	–4-Pyridyl	–4-Pyridyl	8.50	8.19	8.30	0.31	0.20
16	H	H	–2-pyridyl-3-OMe	–4-Pyridyl	7.40	7.27	7.45	0.13	–0.05
17	H	H	–2-thiazole	–4-Pyridyl	7.40	7.48	7.39	–0.08	0.01

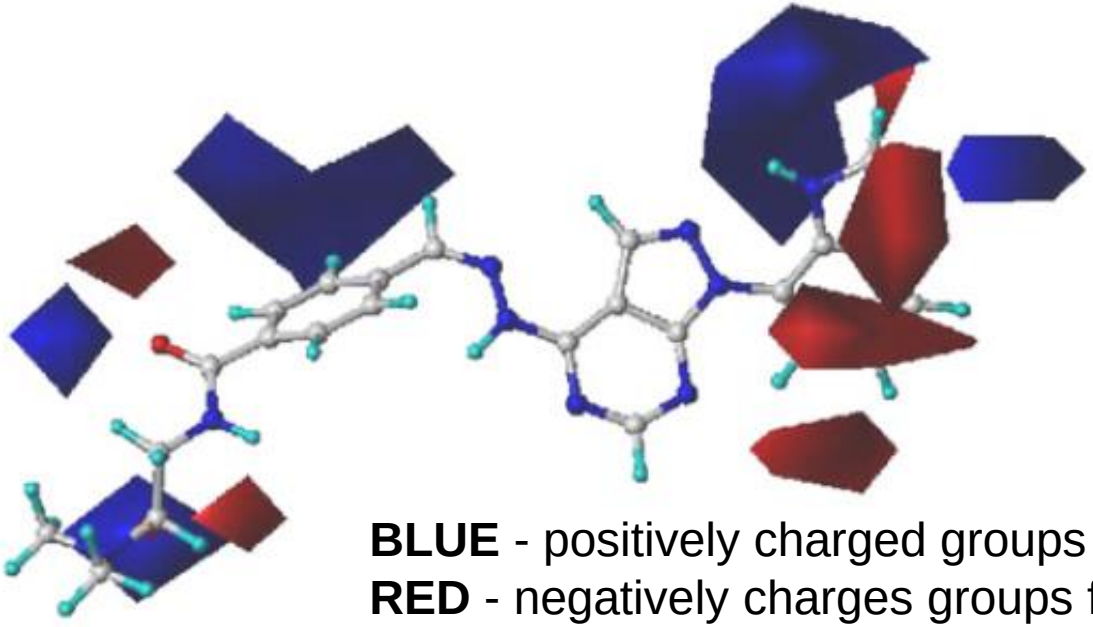


Guide for design



steric

GREEN – space-filling areas for best binding
YELLOW – space-conflicting areas



electrostatic

BLUE - positively charged groups favourable
RED - negatively charges groups favourable

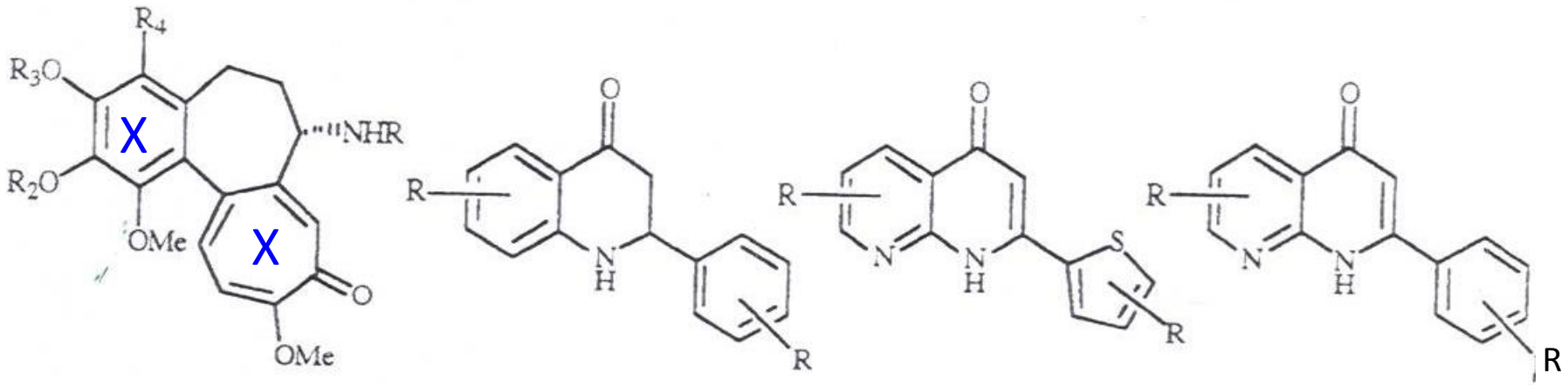
6.2.4. Case study: colchicine

- alkaloid from Morning Crocus
- inhibitor of tubulin polymerisation
- useful in treatment of arthritis

colchicine:

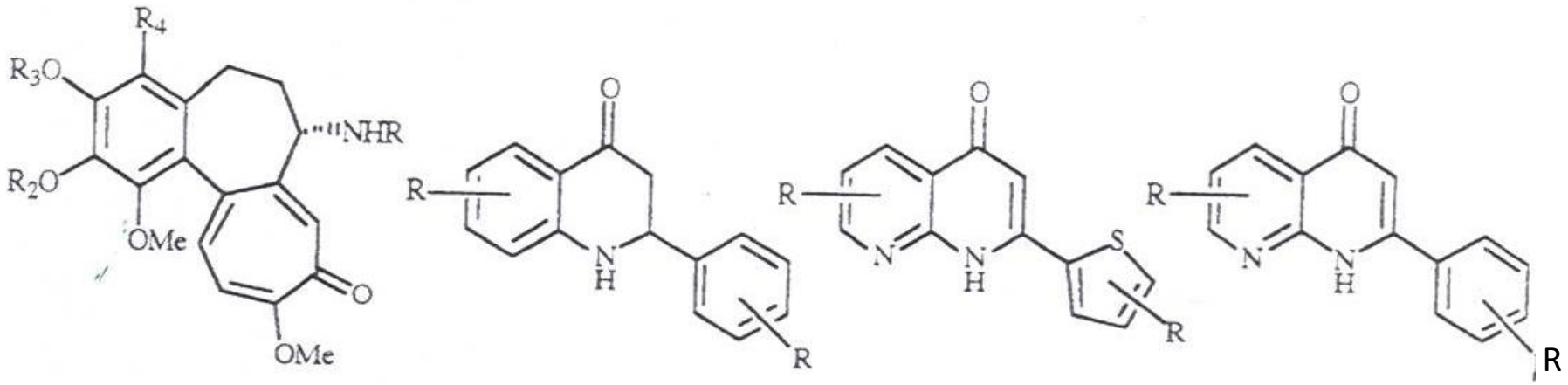
- lead molecule in a 3D QSAR study of 104 molecules of various structural types
- rigid
- high affinity for tubulin

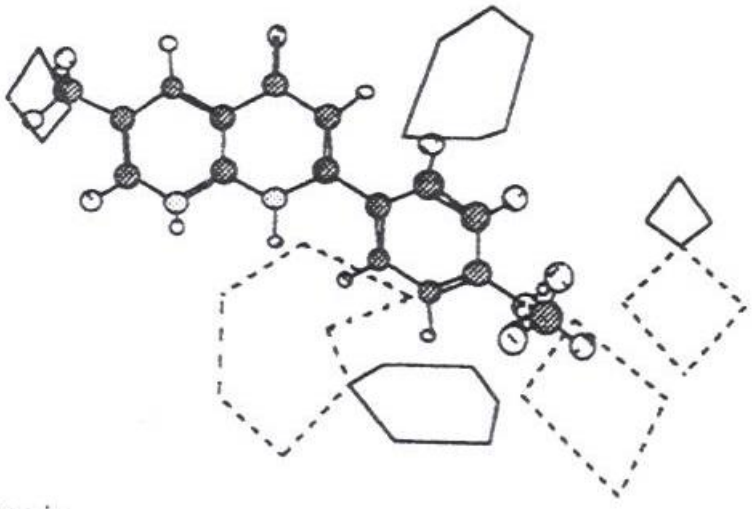
- pharmacophore identified as the two aromatic rings



Colchicine, R=Ac,R_1=R_2=R_3=Me,R_4=H

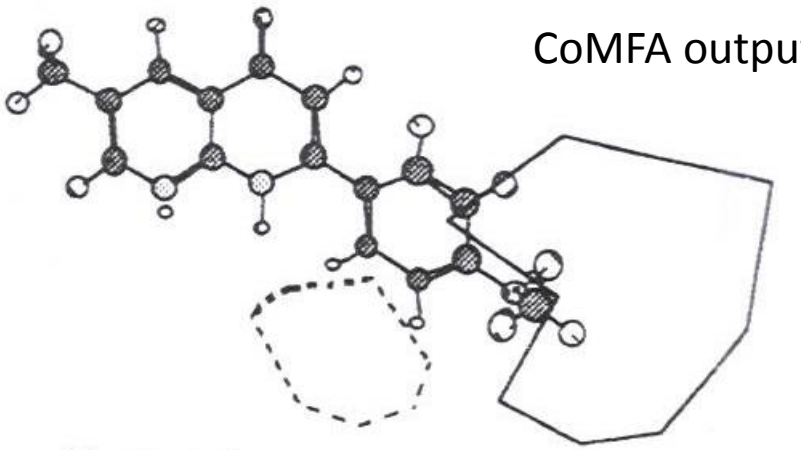
- pharmacophore identified as the two aromatic rings





Steric

CoMFA output



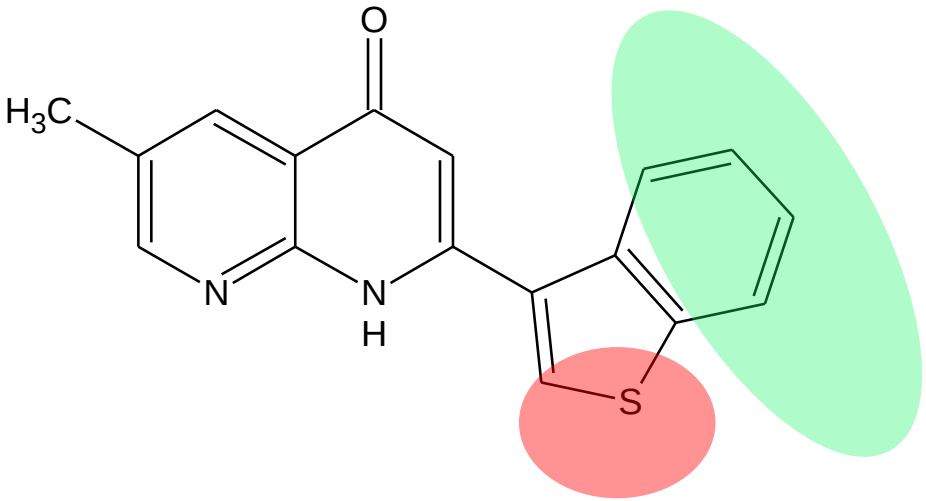
Electrostatic

analysis of results:

1. introduction of steric bulk around the single aromatic ring more effect on activity than introducing steric bulk around the bicyclic system

2. negative electrostatic area near to aromatic ring
⇒ introduce electronegative groups in this area

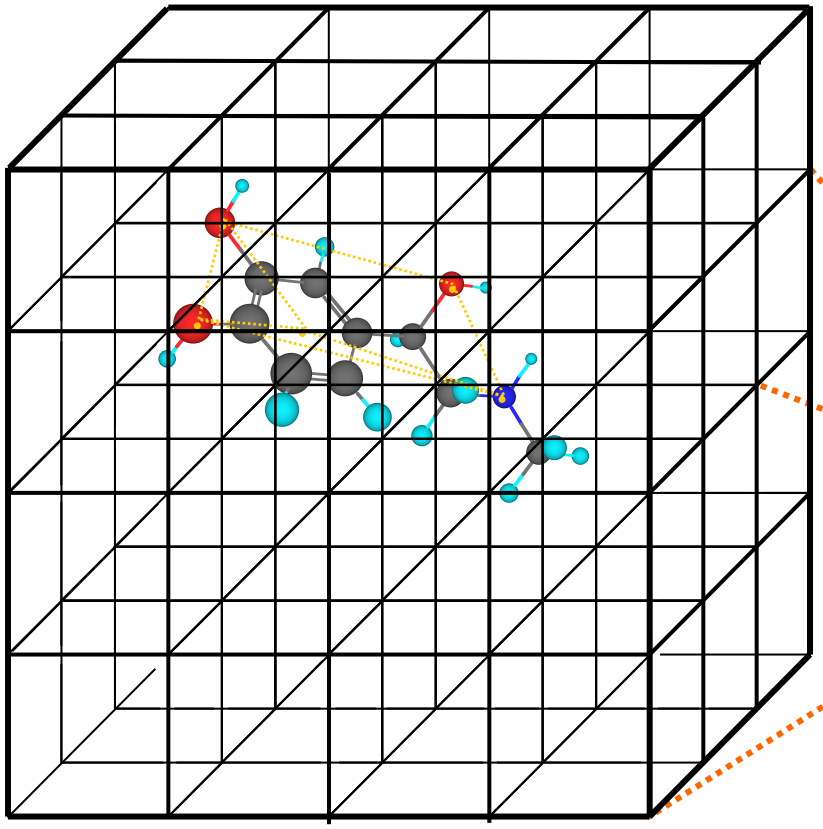
3. novel compound with high activity synthesised based on this evidence:



Comparative Molecular Similarity Indices Analysis (CoMSIA)

- calculate a wider range of properties than CoMFA
- CoMSIA probes:
  - steric
  - electrostatic
  - hydrophobic
  - hydrogen bond donor
  - hydrogen bond acceptor
- assign value for each property to each atom in ligand molecule

•Position molecule to match the pharmacophore

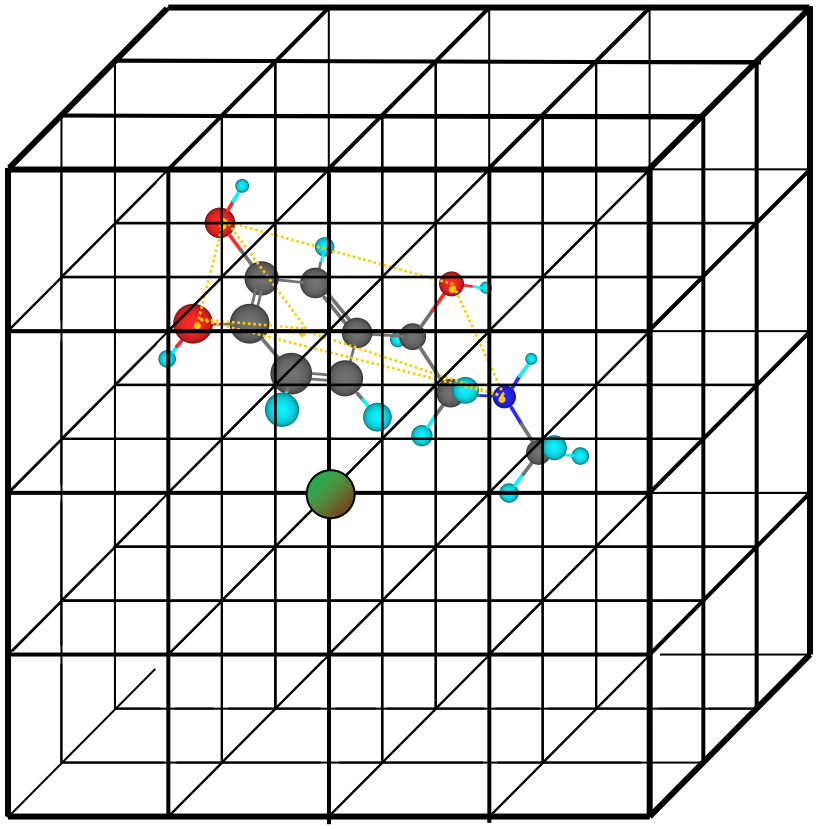


Grid points

•Each grid point defines a point in space

Patrick/OUP

•A probe *point* is placed at each grid point in turn



Probe *point*

calculate score of selected property (eg. hydrophobicity) at probe point

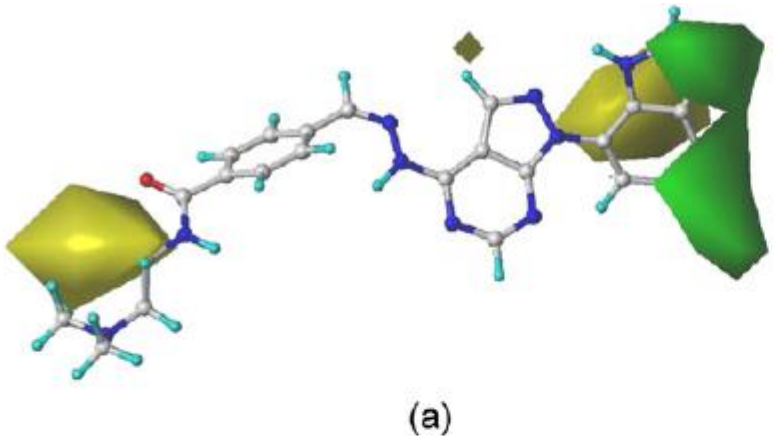
- strength depends on (1) distance of ligand atoms from probe point and (2) nature of ligand atoms

  - eg. if probe point close to a group of ligand hydrophobic atoms, we will have a large hydrophobic score at that probe point

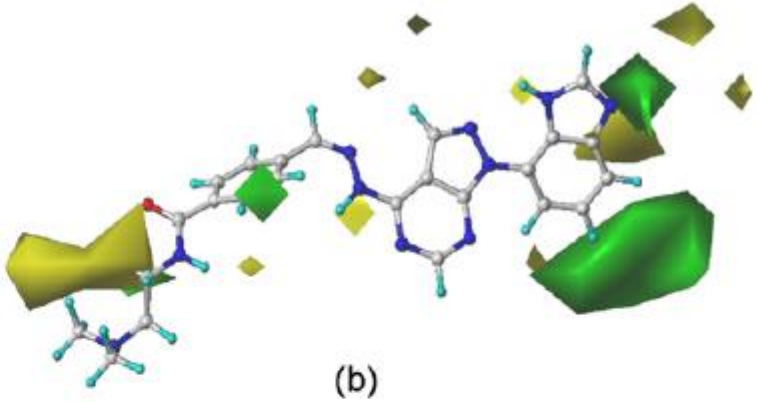
CoMSIA

CoMFA

Steric

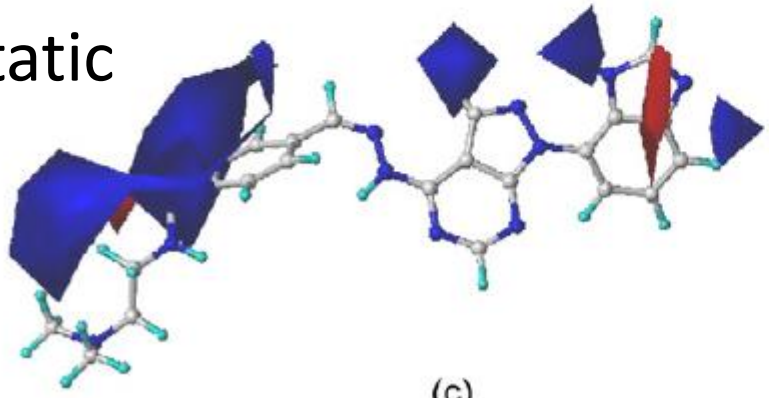


(a)

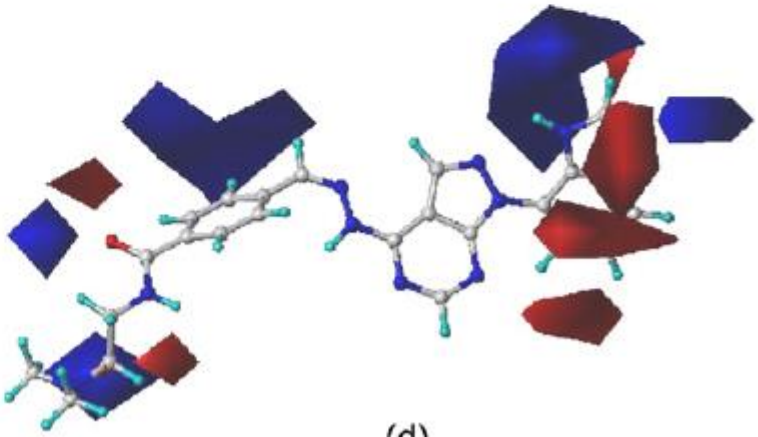


(b)

Electrostatic

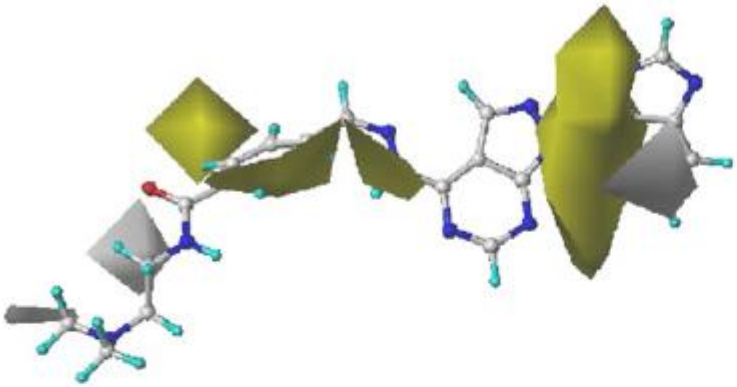


(c)

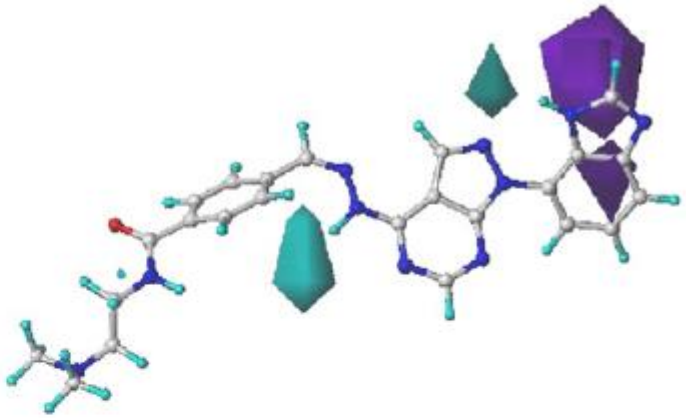


(d)

CoMSIA



hydrophobic



HB donor

HB acceptor

6.2.5. Potential problems of 3D QSAR

1. molecule must be in its bioactive conformation

- often, energy minimisation is used to produce a stable (local) conformation which is assumed to be same as most active conformation

- best if got a rigid molecule in set (difficult if they are very flexible)

2. each molecule should be aligned properly with the others to ensure their pharmacophores match

- sometimes it is difficult to identify the pharmacophore in molecules

- overlay combinations of molecule conformations: combinatorial explosion

  eg. 2 ligands, 5 rotatable bonds, 10° step

  ~4×10^{15} possible combinations of overlay

3. problematic if one compound in a series binds to the receptor in a different conformation from the others

6.2.6. Benefits of 3D QSAR

1. properties calculated for each individual molecule on computer

2. no reliance on experimental parameters

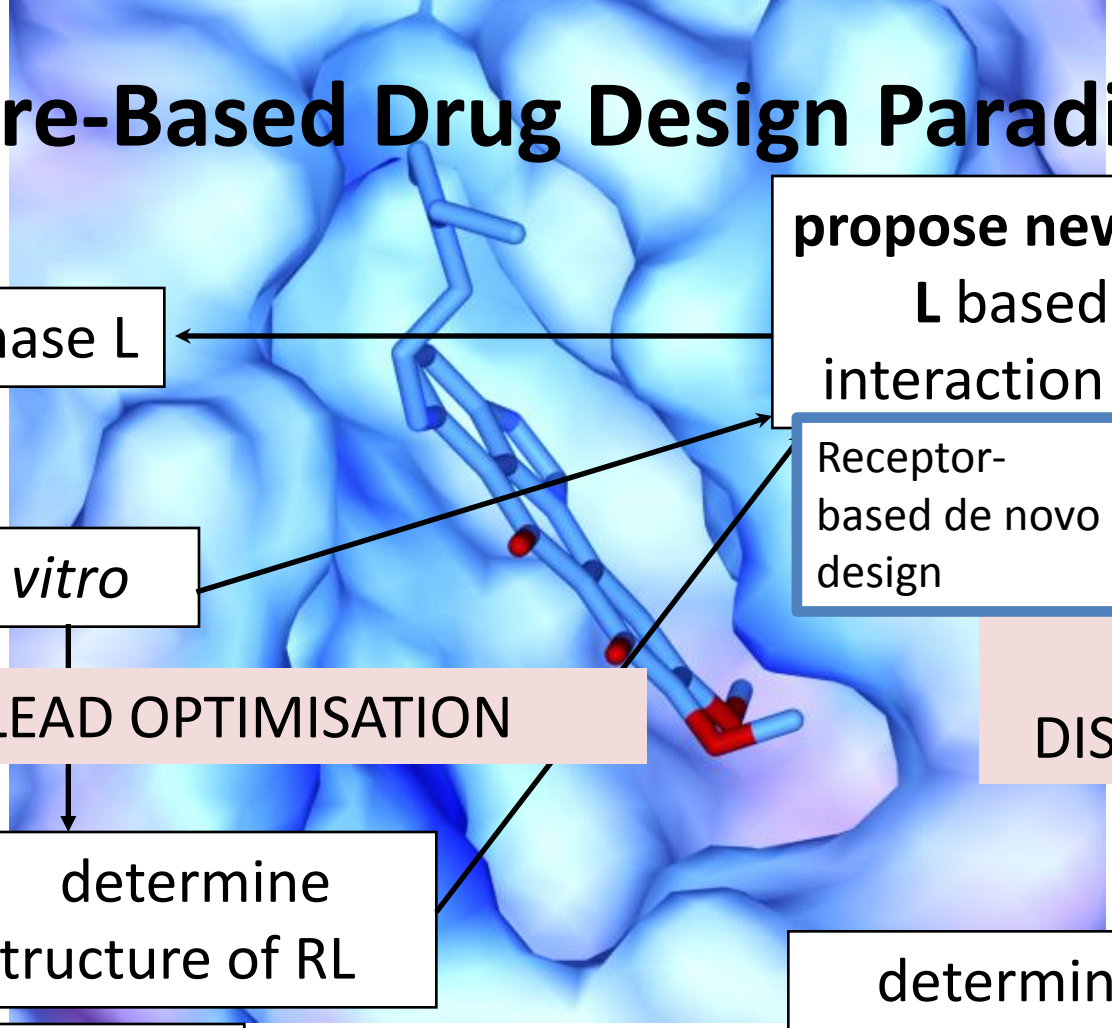
3. no need for congeneric series provided that a similar pharmacophore can be defined

4. mathematics complex but graphical representation of beneficial and non- beneficial interactions allows easier design of new leads

BUT REMEMBER: predictions only as good as the mathematical model used which may have made many approximations

Go deeper: **Patrick, Introduction to Medicinal Chemistry, Ch. 18. Section 18.10**

The Structure-Based Drug Design Paradigm



propose new ligand L based on interaction with R

synthesise/purchase L

assay L *in vitro*

LEAD OPTIMISATION

determine structure of RL

assay L *in vivo (cell, tissue, etc. models)*

drug candidate

Receptor-based de novo design

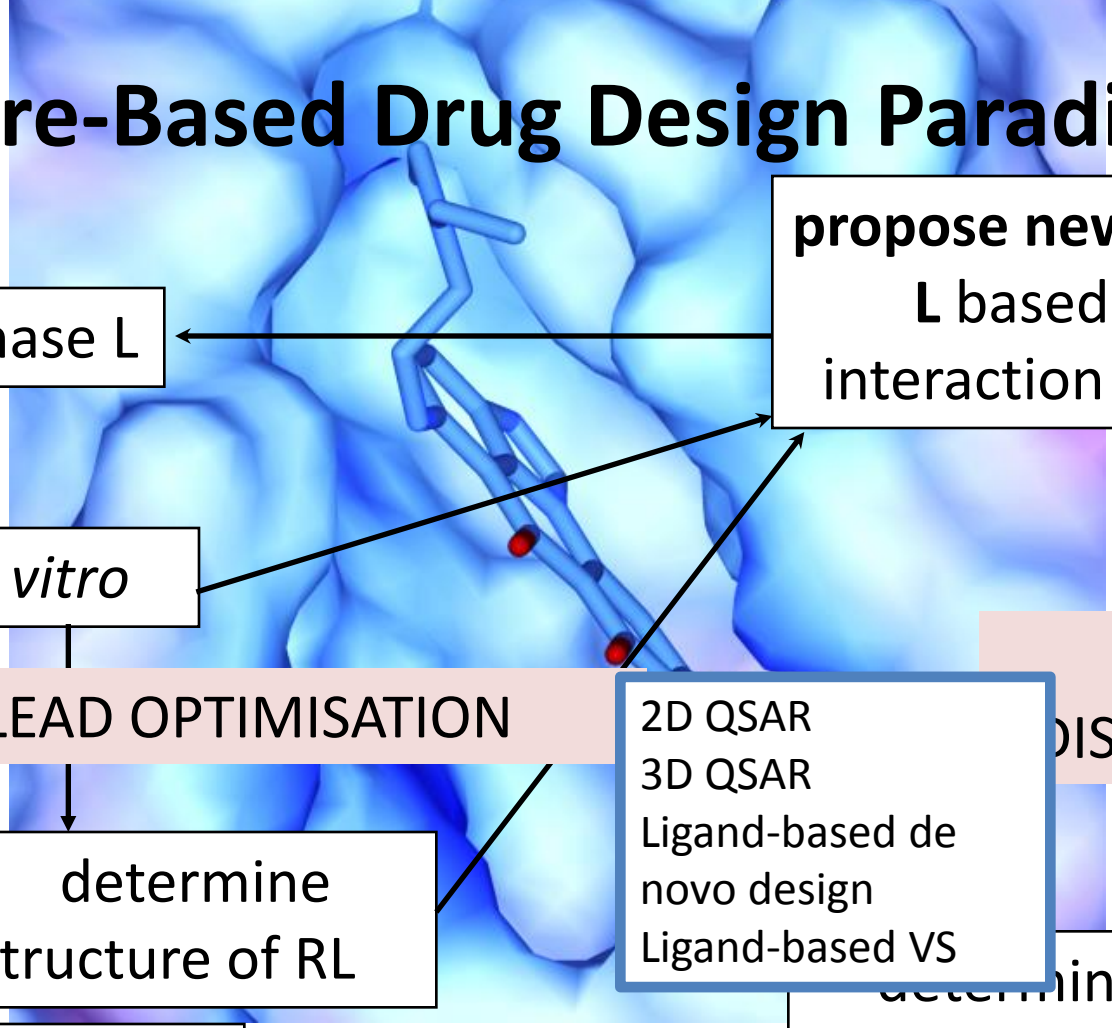
Receptor-based VS

LEAD DISCOVERY

determine **structure of R**

select appropriate target receptor R

The Structure-Based Drug Design Paradigm



propose new ligand L based on interaction with R

synthesise/purchase L

assay L *in vitro*

LEAD OPTIMISATION

determine structure of RL

2D QSAR
3D QSAR
Ligand-based de novo design
Ligand-based VS

LEAD DISCOVERY

determine structure of R

assay L *in vivo (cell, tissue, etc. models)*

drug candidate

select appropriate target receptor R