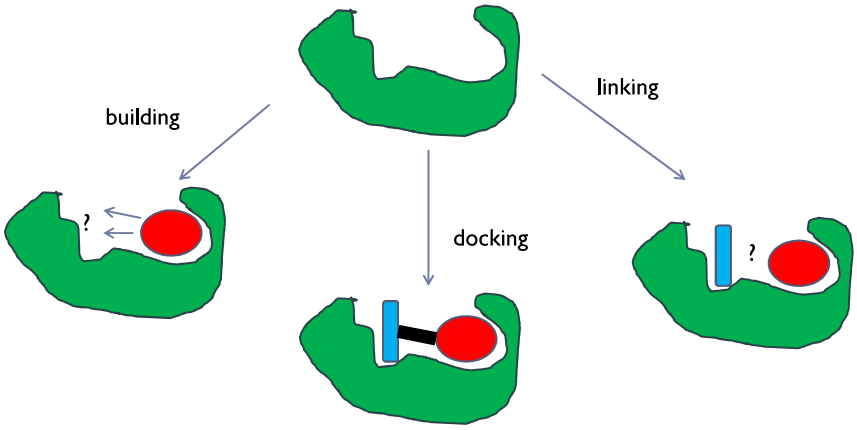
Rational Drug Design lecture 13

Łukasz Berlicki

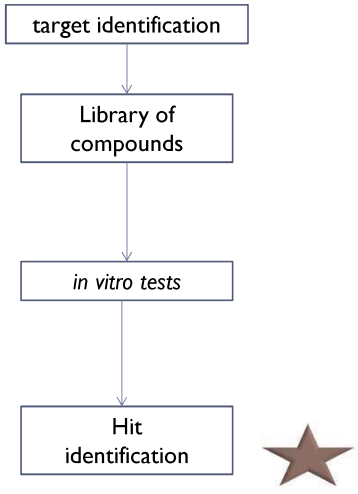
Computer-aided design

- Structure-based design of medicines often uses computer methods to:
  - Analyze a large number of structures
  - Evaluate ligand binding energy
  - Design new compounds

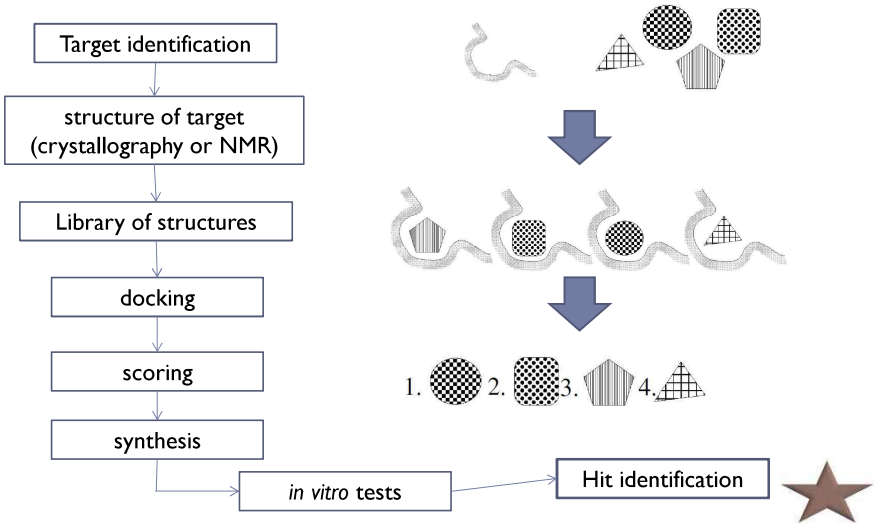
Strategies for searching for ligands



Classic approach



Computer-aided design

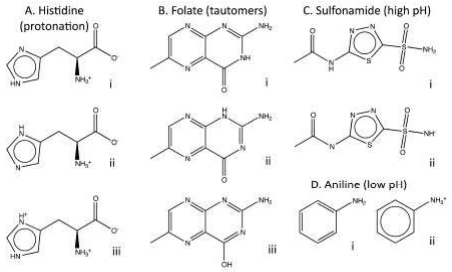


Compounds databases

- Free public databases:
  - NCI database (*National Cancer Institute*), ca. 400 000 structures
  - ZINC database (University of California, San Francisco), ca. **35 000 000** structures of compounds
- Databases of compounds in pharmaceutical companies

ZINC database

- http://zinc.docking.org/
- The sum of catalogs (295) of companies providing chemical compounds
- Divided into partial databases
- Describes various states of the protonation
- Format prepared for docking



ZINC database

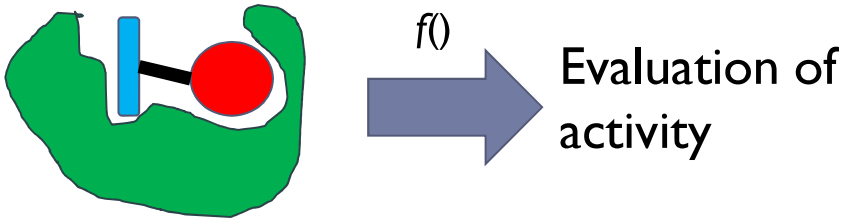
	Lead-Like	Fragment-Like	Drug-Like	All
Standard Size Updated	**Lead-Like** 6,053,287 2014-09-29	**Fragment-Like** 2,098,733 2014-11-24	**Drug-Like** 17,900,742 2014-11-24	**All Purchasable** 22,724,825 2014-11-28
Clean Size Updated	**Clean Leads** 4,591,276 2014-09-25	**Clean Fragments** 1,611,889 2014-09-24	**Clean Drug-Like** 13,195,609 2013-11-05	**All Clean** 16,403,865 2013-12-18
In Stock Size Updated	**Leads Now** 3,687,621 2014-06-25	**Frags Now** 1,768,827 2014-09-17	**Drugs Now** 10,639,555 2014-11-24	**All Now** 12,782,590 2014-05-01
Boutique Size Updated	**Boutique Leads** 5,114,169 2012-12-24	**Boutique Frags** 2,755,555 2013-11-08	**Boutique Drugs** 10,292,210 2012-11-27	**All Boutique** 12,217,845 2012-11-27
Comments/Citation	Teague, Davis, Leeson, Oprea, Angew Chem Int Ed Engl. 1999 Dec 16;38(24):3743-3748.	Carr RA, Congreve M, Murray CW, Rees DC, Drug Discov Today. 2005 Jul 15;10(14):987	Lipinski, J Pharmacol Toxicol Methods. 2000 Jul-Aug;44(1):235-49.	Purchasable chemical space
Filtering Critieria	p.mwt <= 350 and p.mwt >= 250 and p.xlogp <= 3.5 and p.rb <= 7	p.xlogp <=3.5 and p.mwt <=250 and p.rb <= 5	p.mwt <= 500 and p.mwt >= 150 and p.xlogp <= 5 and p.rb <=7 and p.psa < 150 and p.n_h_donors <= 5 and p.n_h_acceptors	

Docking programs

- Assumptions for docking:
  - Solvent is omitted
  - Protein in a fixed conformation (rigid)
  - The ligand can take different conformations
- Types of algorithms:
  - Systematic search
  - Stochastic (Monte Carlo and genetic)
  - Molecular dynamics

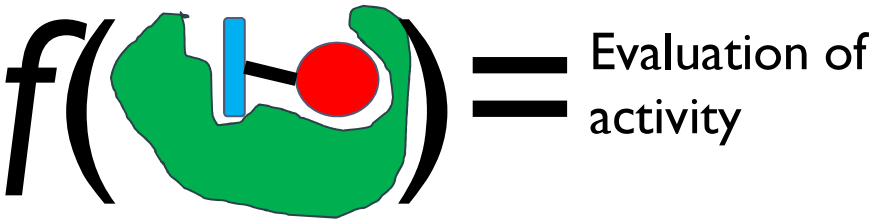
scoring function

- **Scoring function** – an algorithm for assessing the ligand-to-protein binding constant for a given ligand-protein complex



scoring function

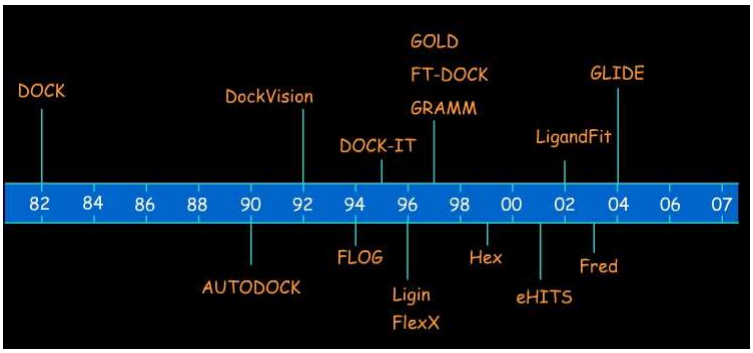
- **Scoring function** – an algorithm for assessing the ligand-to-protein binding constant for a given ligand-protein complex

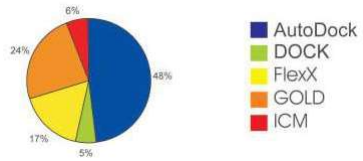


scoring function

- Typs of scoring functions:
  - **force field-based -** calculated on the basis of equations describing molecular mechanics (using force field).
  - **empirical** – equal to the specific interactions developed on the basis of known ligand-protein complexes
  - **knowledge-based** – equations correlating known inhibitory activities with complex structures based on the distance of individual atom pairs.
  - **consensus scoring** – combination of several evaluation functions.

Docking programs



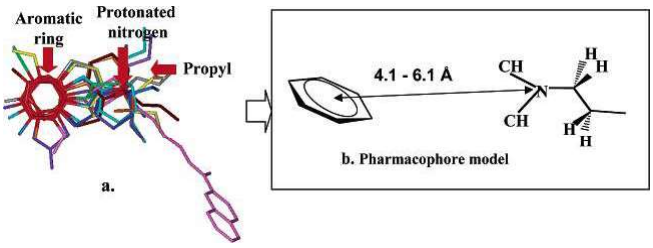


Virtual screening vs HTS

- Search for inhibitors of protein tyrosine phosphatase (PTP 1B)
  - High-throughput screening (HTS)
    - 400 000 compounds,
    - 300 compounds with activity < 300 μM,
    - 85 tested compounds IC_{50} < 100 μM,
    - **hit rate** = 0.021 %
  - virtual screening (program DOCK v. 3.5)
    - 235 000 compounds,
    - 365 structures with high scores,
    - 127 tested compounds IC_{50} < 100 μM,
    - **hit rate** = 34.8 %

Ligands of D_3 receptor

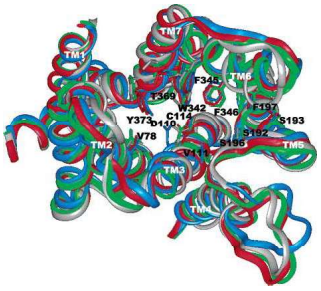
- Virtual screening
  - **250 251** compounds (NCI database)
  - Fitting to pharmacophore

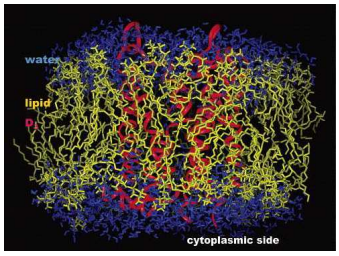


  - **6 727** compound fitted to pharmacophore

Ligands of D_3 receptor

- Docking to 4 conformations of receptor

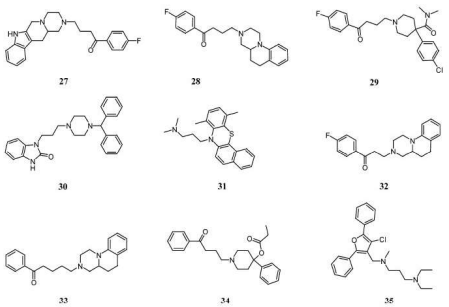




- 20 compounds were selected that represented new chemotypes and had high scores for docking for at least 2 receptor conformations

Ligands of D_3 receptor

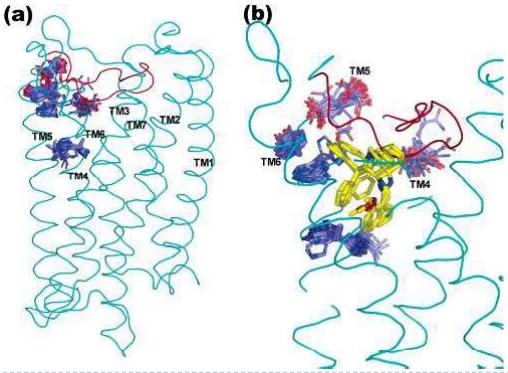
- 8 compounds showed activity
  < 0.5 μM
- Most active compound showed
  K_i = 11 nM



	NCI no.	rank at each D_3 conformer I (%)	II (%)	III (%)	IV (%)	rank (%)	$K_i \pm$ SD (nM)[b]
27	143691	1.5	0.8	3.1	0.5	0.8	11.0 ± 0.6
28	131405	8.6	39.1	17.4	10.0	10.0	83.5 ± 7.3
29	170979	30.3	63.8	16.0	59.4	30.3	43.2 ± 16.7
30	309710	6.1	9.7	4.3	23.1	6.1	62.7 ± 22.4
31	24116	71.1	15.7	12.0	10.7	12.0	442 ± 62
32	147865	15.7	34.4	26.9	21.0	21.0	465 ± 83
33	147980	19.0	41.0	18.1	22.1	19.0	297 ± 90
34	167762	46.9	19.9	23.3	48.4	23.3	429 ± 42
35	402703	24.9	5.1	21.3	77.4	21.3	1381 ± 466
36	349646	2.0	10.5	10.8	1.3	2.0	2412 ± 233
37	186753	51.8	27.7	28.1	32.8	28.1	2615 ± 358
38	13636	15.5	6.8	2.5	17.4	6.8	>10000
39	22808	6.7	9.5	9.1	6.1	6.7	>10000
40	201722	4.3	5.5	2.1	21.8	4.3	>10000
41	202072	29.9	56.2	55.2	20.6	29.9	>10000
42	246981	7.0	24.0	34.0	1.3	7.0	>10000
43	298248	1.6	0.7	2.8	3.9	1.6	>10000
44	300859	37.5	24.0	12.3	38.4	24.0	>10000
45	349645	2.5	10.1	2.4	8.1	2.5	>10000
46	330803	12.6	24.1	4.4	33.0	12.6	>1000[c]

Antagonists of neurokinine 1 receptor

- The neurokinin 1 receptor is a GPCR
- He is responsible for binding neurotransmitters and transmitting pain.
- The crystal structure is unknown
- For the needs of virtual screening a model based on homology was developed



Antagonists of neurokinine 1 receptor

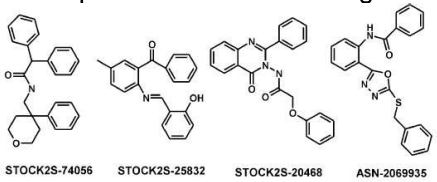
- 8 structural databases were used - **826 952** compounds
- Filters: molar mass and number of rotatable bonds - **419 747** compounds
- Filters: hydrophobic features and hydrogen donor / acceptor – **131 667** compounds
- Adjustment to the pharmacophore – **11 109** compounds

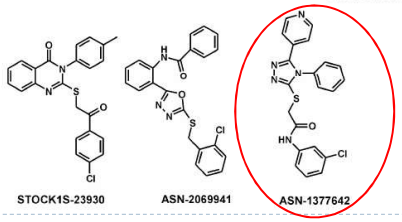
	ACD		AMBINTER		AEGC		AEPC	
filter step	no. of compds[a]	%	no. of compds[a]	%	no. of compds[a]	%	no. of compds[a]	%
	215212	100.00	115815	100.00	182485	100.00	44549	100.00
1. rotatable bonds/MW	135502	62.96	59877	51.70	91677	50.24	9417	21.14
2. requested no. of hydrophobic, donor, and acceptor properties	30878	19.34	19764	17.07	36302	19.89	2740	6.15
3. pharmacophore hypothesis	8645	4.02	5353	4.62	10534	5.77	1018	2.29
4. excluded volumes	3084	1.43	1510	1.30	2998	1.64	334	0.75

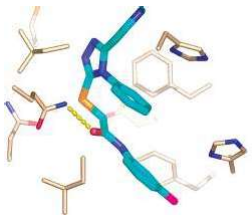
	ChemStar		IBS		LEADQUEST		Σ	
filter step	no. of compds[a]	%	no. of compds[a]	%	no. of compds[a]	%	no. of compds[a]	%
	57927	100.00	158942	100.00	52002	100.00	826952	100.00
1. rotatable bonds/MW	28712	49.57	76321	48.02	18231	35.04	419747	50.76
2. requested no. of hydrophobic, donor, and acceptor properties	11229	19.38	24571	15.46	6483	12.47	131967	15.95
3. pharmacophore hypothesis	3547	6.12	5463	3.44	2144	4.12	36704	4.44
4. excluded volumes	1226	2.12	1362	0.86	595	1.14	11109	1.34

Antagonists of neurokinine 1 receptor

- Docking to the receptor model (FlexX program) and evaluation (Drug Score)
- **1 000** compounds with the highest evaluation of manually tested
- **7** compounds selected for testing



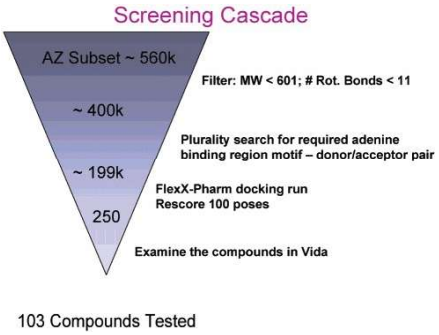




IC_{50} = 250 nM

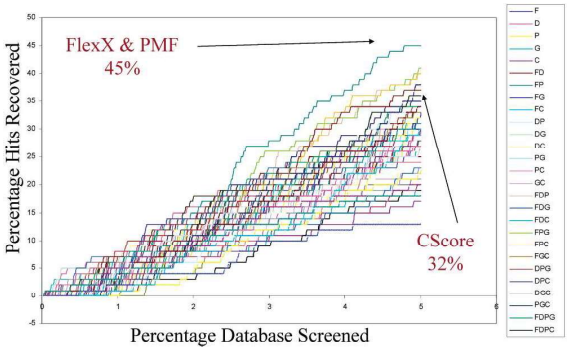
Chk1 kinase inhibitors

- **560 000** compounds from the AstraZeneca database
- Application of filters:
  - molar mass <601
  - number of rotable bonds <11
- **400 000** compounds,
- Presence of hydrogen donor pair / acceptor (binding motif)
- **199 000** compounds
- Docking to the ATP binding site



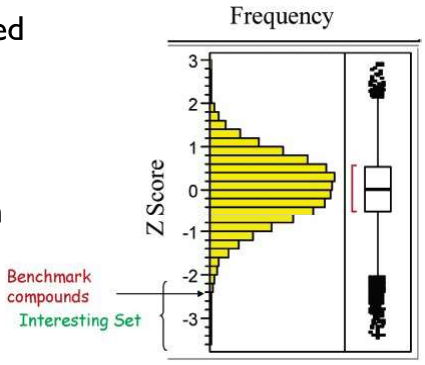
Chk1 kinase inhibitors

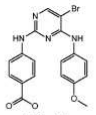
- Evaluation of docking based on the **consensus function**
- Development of a consensus function based on the evaluation of CDk2 kinase inhibitors (100 inhibitors among 8 000 compounds)
- Checking different combinations of evaluation functions allowed to indicate the one that indicated the most inhibitors.
- Combination of functions FlexX and PMF pointed 45% of inhibitors in 5% the best results.

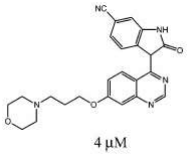


Chk1 kinase inhibitors

- **250** compounds were selected based on the consensus function
- **103** compounds tested
- **36** compounds had activity in the range 0.11 do 68 μM.

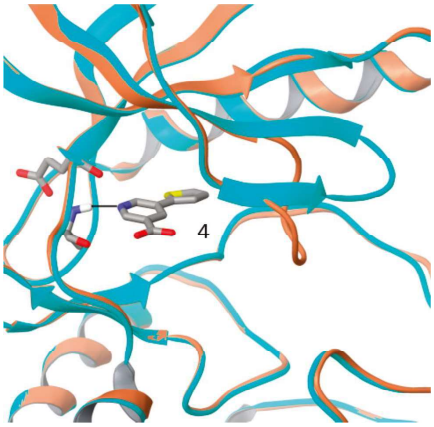






FGFR Kinase inhibitors

- FGFR Kinase (*Fibroblast Growth Factor Receptor 1 Kinase*) plays a role in the development of various cancers.
- The active site can take two conformations
- ZINC base with 2.2 million compounds was used.
- Docking was performed on both conformations using the Glide program.

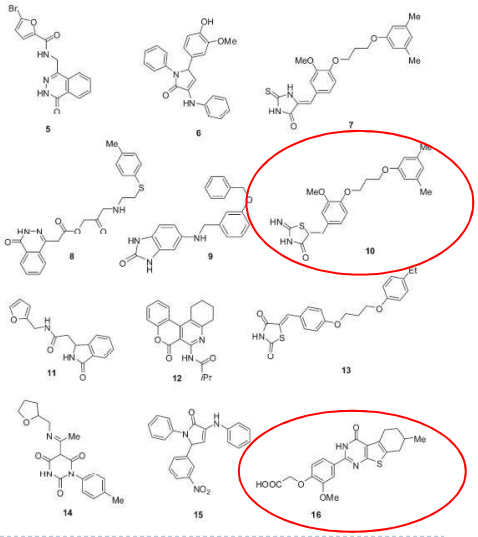


FGFR Kinase inhibitors

- All compounds (**2.2 million**) from the ZINC database were docked in standard precision mode to both conformations.
- The **40 000** best-rated compounds are docked again in high-precision mode for both conformations.
- **1 000** of the best-assessed compounds are docked to other kinases (EGFR, InsR, VEGFR2, Src and MEK).
- The **100** best-assessed compounds that were not well evaluated for other kinases.
- **24** compounds were tested.
- **2** compounds showed activity (IC_{50} 23 and 50 μM).

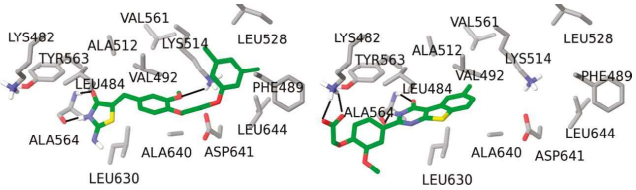
FGFR Kinase inhibitors

compd	XP rank	XP score	SP rank	IC_{50} (μM)[a]
5	1	−16.18	112	na
6	7	−14.84	11628	na
7	16	−14.29	33929	na
8	19	−14.20	22167	na
9	25	−13.98	550	na
10	38	−13.79	61627	23
11	46	−13.71	8409	na
12	51	−13.66	13883	na
13	52	−13.64	6377	na
14	74	−13.49	36663	na
15	77	−13.48	396	na
16	84	−13.45	20052	50
17	2	−18.53	429	na
18	10	−17.38	10443	na
19	29	−17.01	12782	na
20	37	−16.96	7676	na
21	41	−16.93	14137	na
22	45	−16.89	19473	na
23	55	−16.84	9711	na
24	56	−16.84	15905	na
25	64	−16.78	6525	na
26	72	−16.75	4238	na
27	93	−16.62	27217	na
28	97	−16.60	22135	na



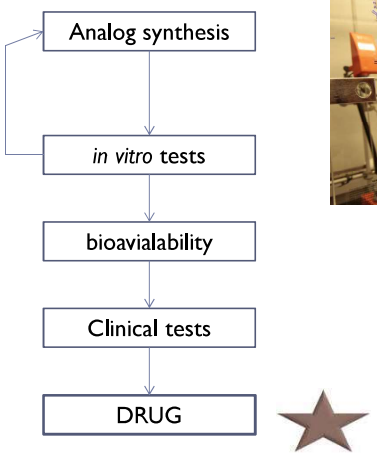
FGFR Kinase inhibitors

- Despite the procedure to increase selectivity, the compounds obtained did not show significant selectivity.

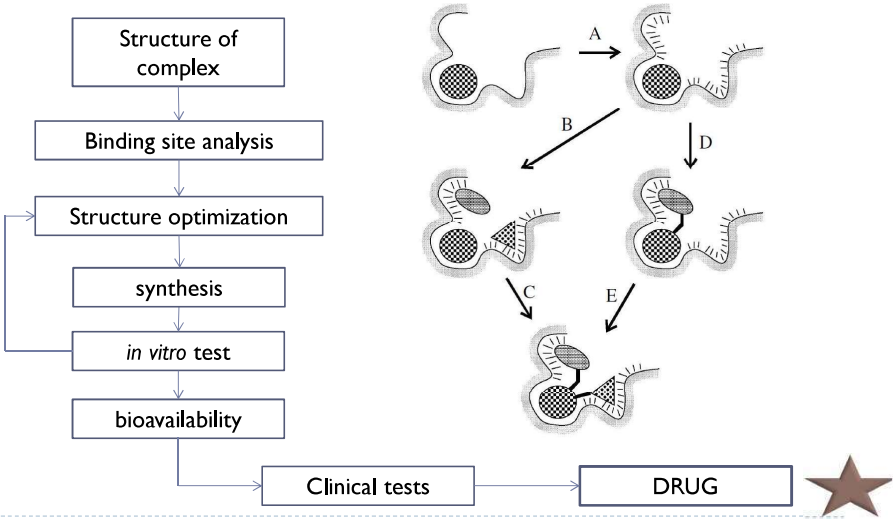


	IC_{50} (μM)			
compd	FGFR1	EGFR	Src	InsR
10	23	56	10	47
40	1.9	2.4	1.9	na

Optimization - classic approach

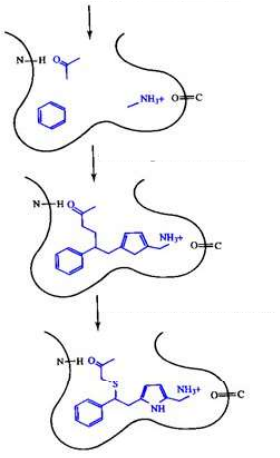


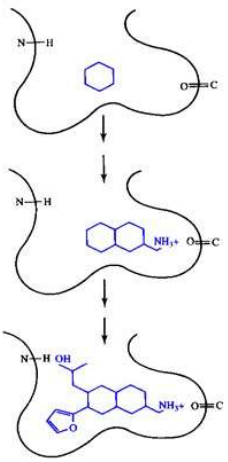
Computer-aided optimization



Optimization strategies

- Fragment linking
- building





Optimization strategies

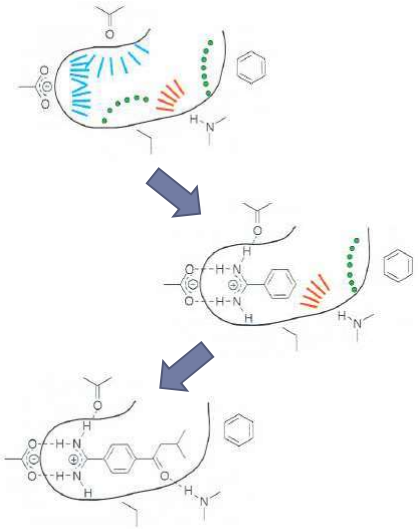
- Connecting atoms
  - Advantages:
    - Large structural diversity of compounds
    - Effective use of all possible interactions
    - No need to use the fragment database
  - Disadvantages:
    - Designed compounds can be difficult to synthesize or unstable
- Connecting fragments
  - Advantages:
    - The compounds consist of fragments that are synthesizable
  - Disadvantages:
    - Structural diversity of compounds depends on the base used

Programs for design

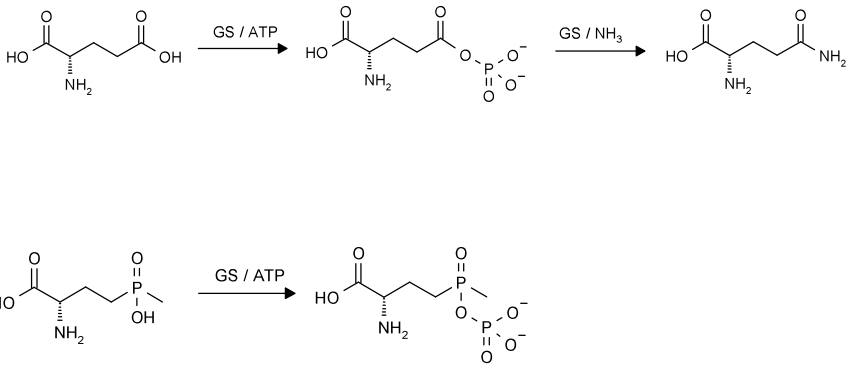
GENSTAR	atoms	building up molecules
GROUPBUILD	fragments	sequential growth
GROW	aminoacids	sequential growth
GROWMOL	fragments	sequential growth stochastic search
HOOK	fragments	fragment linking
LEGEND	atoms	stochastic search
LUDI	fragments	combinatorial search
MCSS	fragments	stochastic search
PRO-LIGAND	fragments	sequential growth
PRO-SELECT	fragments	scaffold-linker approach
SKELGEN	small fragments	Monte-Carlo search
SPROUT	fragments	sequential growth combinatorial search

LUDI program

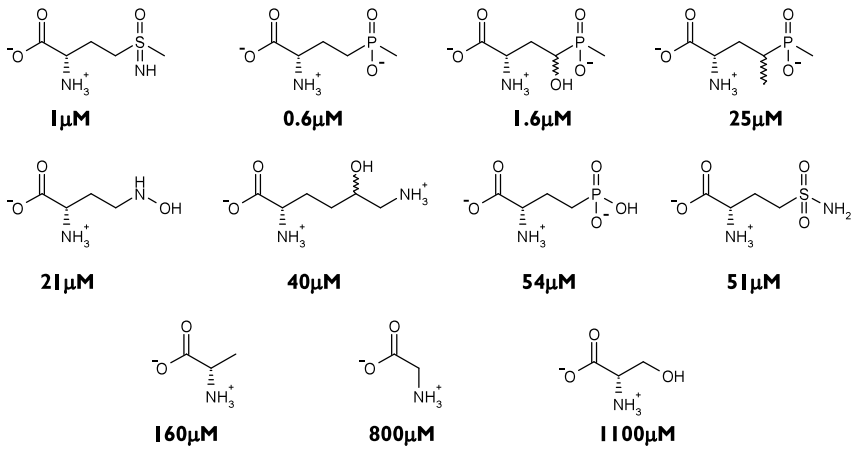
- Finding potential interactions:
  - Lipophilic (green)
  - Hydrogen bond donor (red)
  - Hydrogen acceptor (blue)
- Superimposing fragments to interaction sites
- Attachment of subsequent fragments (overlapping with the base molecule and interaction sites)



Glutamine synthetase inhibitors

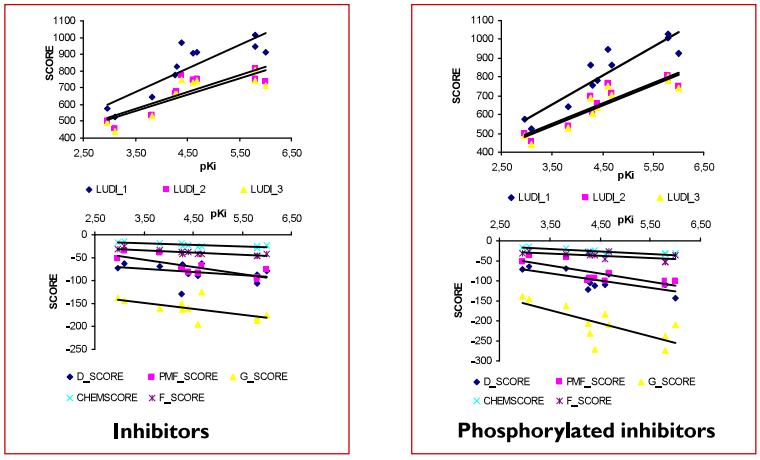


Glutamine synthetase inhibitors



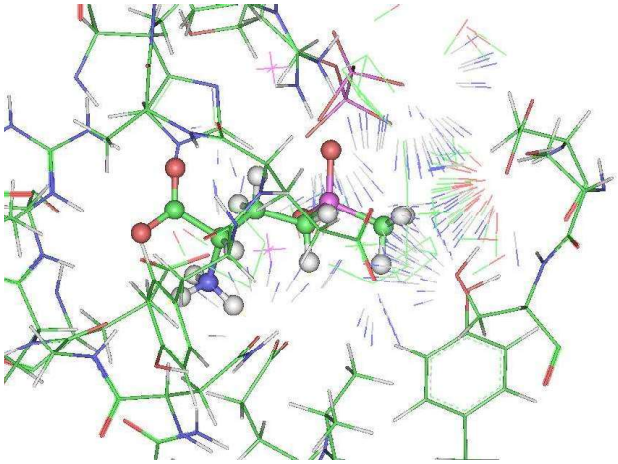
Glutamine synthetase inhibitors

- Scoring of inhibitor-enzyme complexes



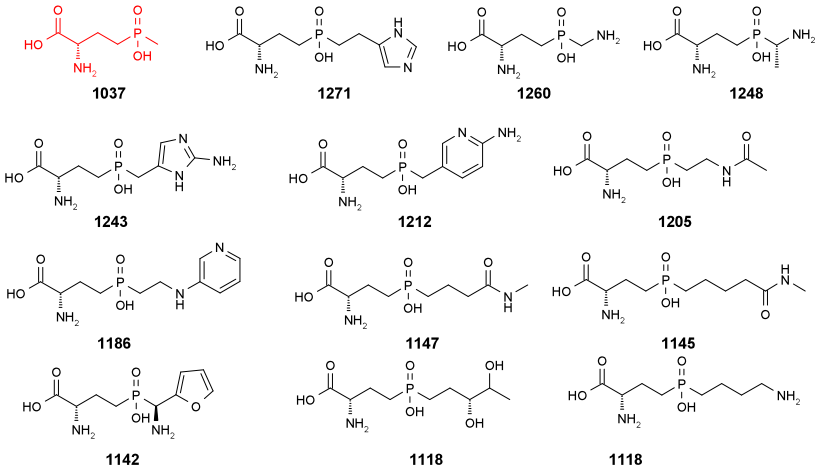
Glutamine synthetase inhibitors

- Interaction sites (LUDI program)



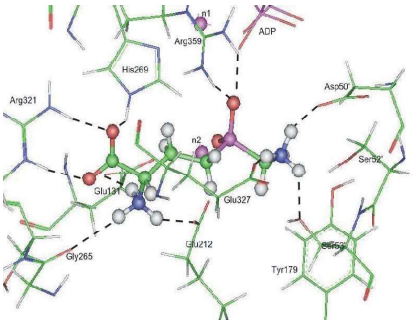
Glutamine synthetase inhibitors

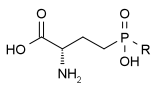
- Structures of designed compounds



Glutamine synthetase inhibitors

- The structure of the enzyme-inhibitor complex





	K_i [μM]		
	E. coli	*M. tuberculosis*	kukurydza
$-CH_3$	0.6	ND	1.1
$-CH_2NH_2$	0.59	0.8	1.8
$-CH_2OH$	2.1	5.7	8.5
$-CH(CH_3)NH_2$	33	3.4	21.6
$-CH_2CH_2NH_2$	55	ND	320
$-CH_2CH_2COOH$	970	ND	ND

Summary

- Computer-aided search for new active molecules is significantly more efficient than classical HTS techniques.
- Designing new molecules leads to structures difficult to obtain by other methods.