

## Projektowanie związków biologicznie czynnych **wykład 13**

Łukasz Berlicki

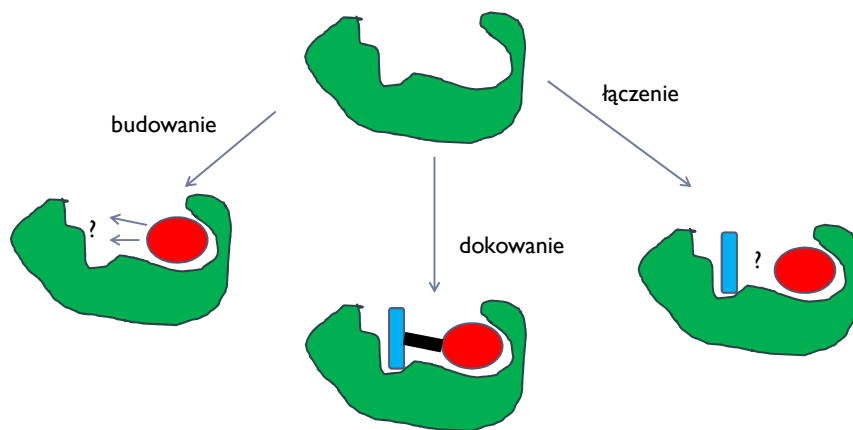
### Projektowanie wspomagane komputerowo

---

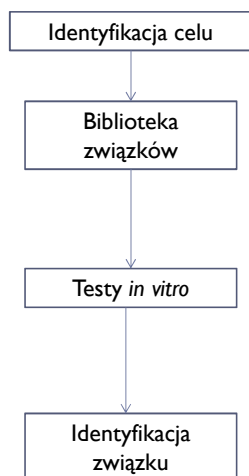
- ▶ Projektowanie leków oparte na strukturze często wykorzystuje metody komputerowe, aby:
  - ▶ Przeanalizować dużą liczbę struktur
  - ▶ Ocenić energię wiązania liganda
  - ▶ Zaprojektować nowe związki



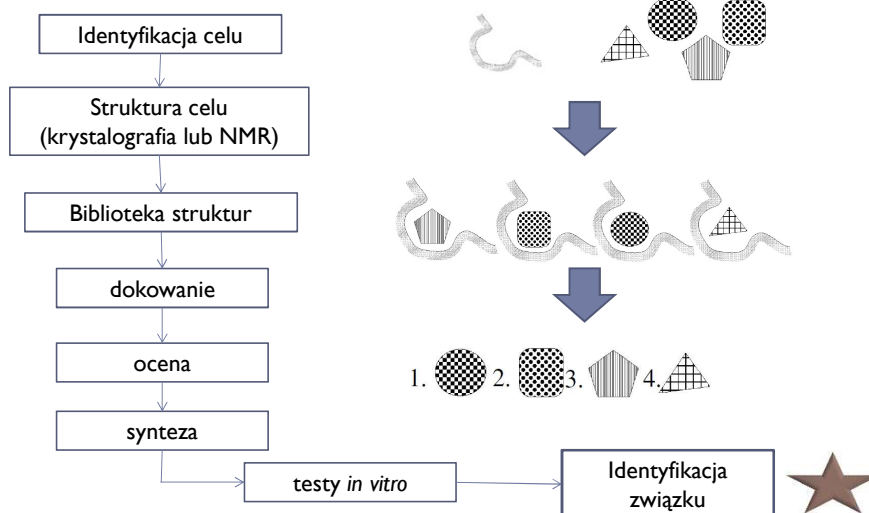
## Strategie poszukiwania ligandów



## podjęcie klasyczne



## Podejście wspomagane komputerowo



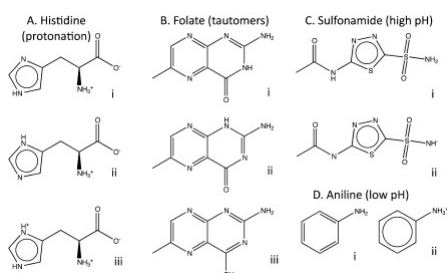
## Bazy danych związków

- ▶ Bazy publiczne, darmowe:
  - ▶ Baza NCI (*National Cancer Institute*), ok. 400 000 struktur
  - ▶ Baza ZINC (Uniwersytet Kalifornijski, San Francisco), ok. 35 000 000 struktur związków, które można kupić
- ▶ Bazy danych związków w firmach farmaceutycznych

The screenshot shows the ZINC 12 website interface. The header includes the UCSF logo and navigation links for 'About UCSF', 'Search UCSF', and 'UCSF Medical Center'. The main content area features a 'Molecule of the Month' section with a chemical structure and a 'Welcome to ZINC' message. The footer contains a search bar and navigation links for 'Structure/Draw', 'Physical Properties', 'Catalogs & Vendors', 'ZINC IDs', 'Targets', 'Rings', and 'Combination'.

## Baza ZINC

- ▶ <http://zinc.docking.org/>
- ▶ Suma katalogów (295) firm dostarczających związki chemiczne
- ▶ Podzielona na bazy częściowe
- ▶ Opisuje różne stany uprotonowania
- ▶ Format przygotowany do dokowania



## Baza ZINC

	Lead-Like	Fragment-Like	Drug-Like	All
<b>Standard</b> Size Updated	<a href="#">Lead-Like</a> 6,053,287 2014-09-29	<a href="#">Fragment-Like</a> 2,098,733 2014-11-24	<a href="#">Drug-Like</a> 17,900,742 2014-11-24	<a href="#">All Purchasable</a> 22,724,825 2014-11-28
<b>Clean</b> Size Updated	<a href="#">Clean Leads</a> 4,591,276 2014-09-25	<a href="#">Clean Fragments</a> 1,611,889 2014-09-24	<a href="#">Clean Drug-Like</a> 13,195,609 2013-11-05	<a href="#">All Clean</a> 16,402,865 2013-12-18
<b>In Stock</b> Size Updated	<a href="#">Leads Now</a> 3,687,621 2014-06-25	<a href="#">Frgs Now</a> 1,768,827 2014-09-17	<a href="#">Drugs Now</a> 10,639,555 2014-11-24	<a href="#">All Now</a> 12,782,590 2014-05-01
<b>Boutique</b> Size Updated	<a href="#">Boutique Leads</a> 5,114,169 2012-12-24	<a href="#">Boutique Frags</a> 2,755,555 2013-11-08	<a href="#">Boutique Drugs</a> 10,292,210 2012-11-27	<a href="#">All Boutique</a> 12,217,845 2012-11-27
Comments/Citation	<a href="#">Teague, Davis, Leeson, Oprea, Angew Chem Int Ed Engl. 1999 Dec 16;38(24):3743-3748.</a>	<a href="#">Carr RA, Congreve M, Murray CW, Rees DC. Drug Discov Today. 2002 Jul; 15:1014-1987</a>	<a href="#">Lipinski, J Pharmacol Toxicol Methods. 2000 Jul-Aug;44(1):235-49.</a>	Purchasable chemical space
Filtering Criteria	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	



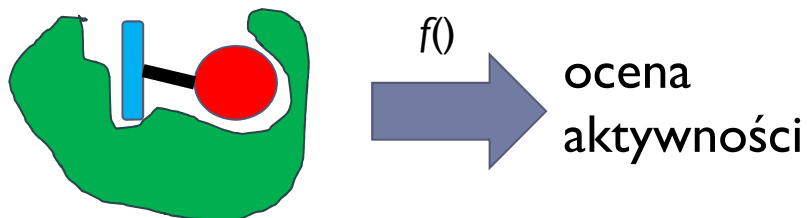
## Programy do dokowania

- ▶ Założenia dla dokowania:
  - ▶ Rozpuszczalnik pomijany
  - ▶ Białko w ustalonej konformacji (sztywne)
  - ▶ Ligand może przyjmować różne konformacje
- ▶ Typy algorytmów:
  - ▶ Przeszukiwanie systematyczne
  - ▶ Stochastyczne (Monte Carlo i genetyczne)
  - ▶ Dynamika molekularna



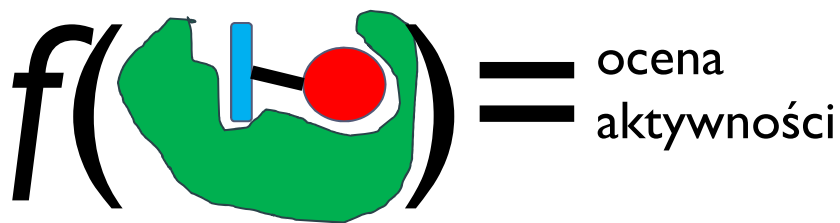
## Funkcje oceniające

- ▶ **Funkcja oceniająca** (*scoring function*) – algorytm oceny stałej wiązania liganda do białka dla danego kompleksu ligand-białko



## Funkcje oceniające

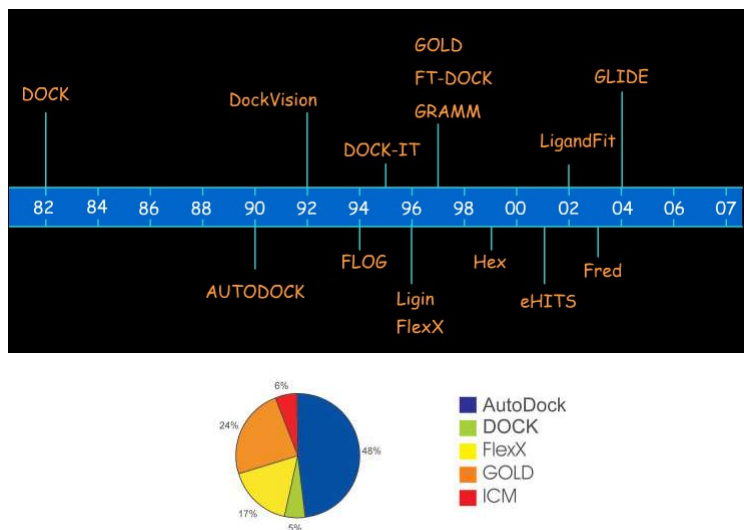
- ▶ **Funkcja oceniająca** (*scoring function*) – algorytm oceny stałej wiązania liganda do białka dla danego kompleksu ligand-białko



## Funkcje oceniające

- ▶ **Typy funkcji oceniających:**
  - ▶ **Bazujące na energii oddziaływania** (*force field-based*) - obliczonej na podstawie równań opisujących mechanikę molekularną (z zastosowaniem pola siłowego).
  - ▶ **Empiryczne** (*empirical*) – równia opisujące poszczególne oddziaływania opracowane na podstawie znanych kompleksów ligand-białko
  - ▶ **Bazujące na wiedzy** (*knowledge-based*) – równania korelujące znane aktywności inhibitorowe z strukturami kompleksów na podstawie odległości poszczególnych par atomów.
  - ▶ **Konsensusowe** (*consensus scoring*) – połączenie kilku funkcji oceniających.

## Programy do dokowania

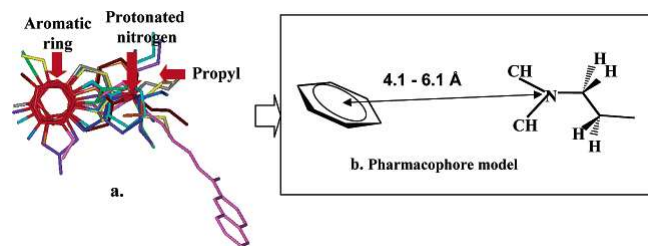


## Virtual screening vs HTS

- ▶ Poszukiwanie inhibitorów fosfatazy tyrozyn białkowych (*protein tyrosine phosphatase*, PTP 1B)
  - ▶ wielkoskalowe testy przesiewowe (HTS)
    - ▶ 400 000 związków,
    - ▶ 300 związków o aktywności  $< 300 \mu\text{M}$ ,
    - ▶ 85 przetestowanych związków  $\text{IC}_{50} < 100 \mu\text{M}$ ,
    - ▶ Współczynnik sukcesu (*hit rate*) = **0.021 %**
  - ▶ virtual screening (program DOCK v. 3.5)
    - ▶ 235 000 związków,
    - ▶ 365 struktur z wysokimi ocenami,
    - ▶ 127 przetestowanych związków  $\text{IC}_{50} < 100 \mu\text{M}$ ,
    - ▶ Współczynnik sukcesu (*hit rate*) = **34.8 %**

## Ligandy receptora dopaminowego D<sub>3</sub>

- ▶ Virtual screening
  - ▶ 250 251 związków (baza NCI)
  - ▶ Dopasowanie do farmakoforu

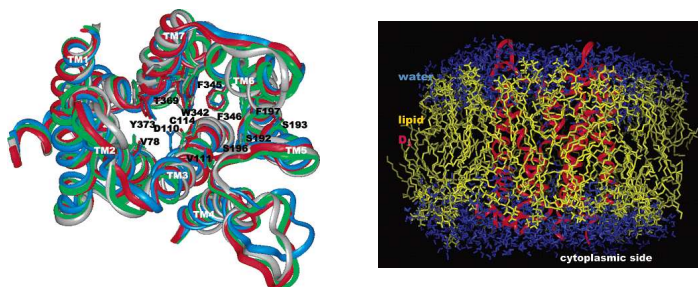


- ▶ 6 727 związków pasujących do farmakoforu



## Ligandy receptora dopaminowego D<sub>3</sub>

- ▶ Dokowanie do 4 konformacji receptora



- ▶ Wybrano 20 związków, które reprezentowały nowe chemotypy i miały wysoką ocenę dla dokowań dla co najmniej 2 konformacji receptora

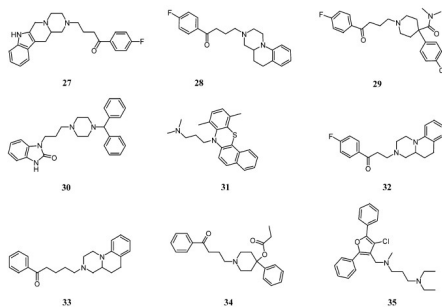




## Ligandy receptora dopaminowego D<sub>3</sub>

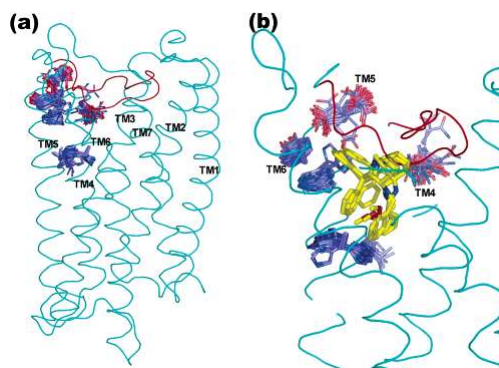
- ▶ 8 związków wykazywało aktywność < 0.5 μM
- ▶ Najbardziej aktywny związek miał  $K_i = 11 \text{ nM}$

NCI no.	rank at each D <sub>3</sub> conformer					K <sub>i</sub> ± SD (nM) <sup>b</sup>	
	I (%)	II (%)	III (%)	IV (%)	rank (%)		
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	>10000



## Antagoniści receptora neurokininy 1

- ▶ Receptor neurokininy I jest GPCR
- ▶ Odpowiada za wiązanie neurotransmiterów i przekazywanie bólu.
- ▶ Struktura krystaliczna nie jest znana
- ▶ Dla potrzeb *virtual screening* opracowano model na podstawie homologii



## Antagoniści receptora neurokininy 1

- ▶ Zastosowano 8 strukturalnych baz danych – **826 952** związki
- ▶ Filtry: masa molowa i ilość wiązań rotacyjnych – **419 747** związków
- ▶ Filtry: cechy hydrofobowe i donor/akceptor wiązania wodorowego – **131 967** związków
- ▶ Dopasowanie do farmakoforu – **11 109** związków

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

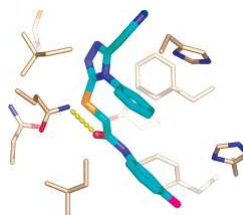
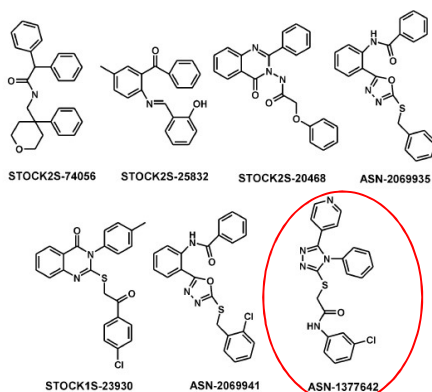
  

filter step	ChemStar		IBS		LEADQUEST		Σ	
	no. of compds <sup>a</sup>	%	no. of compds <sup>a</sup>	%	no. of compds <sup>a</sup>	%	no. of compds <sup>a</sup>	%
1. rotatable bonds/MW	57927	100.00	158942	100.00	52002	100.00	826952	100.00
2. requested no. of hydrophobic, donor, and acceptor properties	28712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



## Antagoniści receptora neurokininy 1

- ▶ Dokowanie do modelu receptora (program FlexX) i ocena (Drug Score)
- ▶ **1 000** związków o najwyższej ocenie sprawdzonych ręcznie
- ▶ **7** związków wybranych do testów

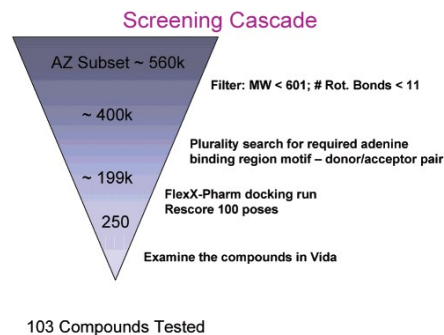


IC<sub>50</sub> = 250 nM



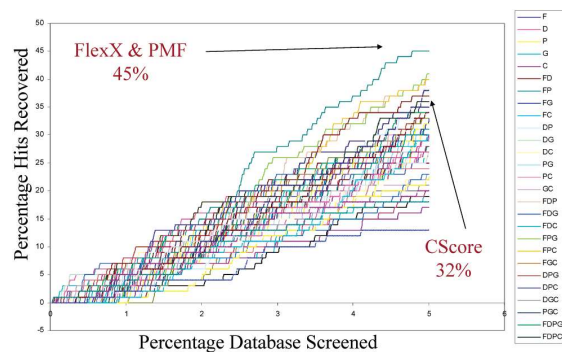
## Inhibitory kinazy Chk1

- ▶ **560 000** związków z bazy danych AstraZeneca
- ▶ Zastosowanie filtrów:
  - ▶ masa molowa < 601
  - ▶ liczba wiązań rotacyjnych < 11
- ▶ **400 000** związków,
- ▶ Obecność pary donor/akceptor wiązania wodorowego (motyw wiążący)
- ▶ **199 000** związków
- ▶ Dokowanie do miejsca wiązania ATP



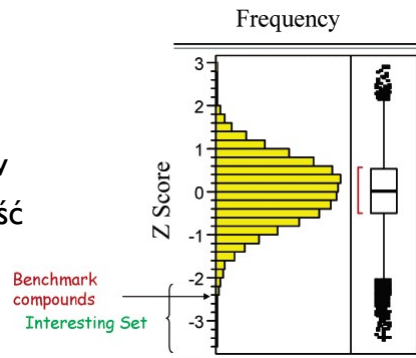
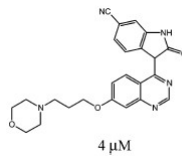
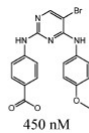
## Inhibitory kinazy Chk1

- ▶ Ocena dokowania na podstawie **funkcji konsensusowej**
- ▶ Opracowanie funkcji konsensusowej na podstawie oceny inhibitorów kinazy CDK2 (100 inhibitorów wśród 8 000 związków)
- ▶ Sprawdzenie różnych kombinacji funkcji oceniających pozwoliło wskazać tą, która wskazywała najwięcej inhibitorów.
- ▶ Kombinacja funkcji FlexX i PMF wskazywała **45%** inhibitorów w **5%** najlepszych wyników.



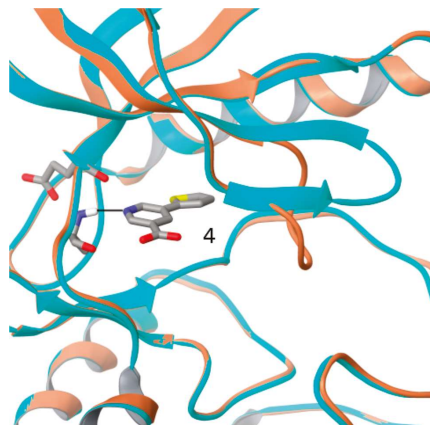
## Inhibitory kinazy Chk1

- ▶ Na podstawie oceny funkcją konsensusową wybrano **250** związków
- ▶ Przetestowani **103** związków
- ▶ **36** związków miało aktywność w zakresie **0.11 do 68  $\mu\text{M}$** .



## Inhibitory kinazy FGFR

- ▶ Kinaza FGFR (*Fibroblast Growth Factor Receptor 1 Kinase*) pełni rolę w rozwoju różnych nowotworów.
- ▶ Miejsce aktywne może przyjmować dwie konformacje
- ▶ Zastosowano bazę ZINC z 2.2 mln związków.
- ▶ Dokowanie przeprowadzono do obu konformacji przy użyciu programu Glide.

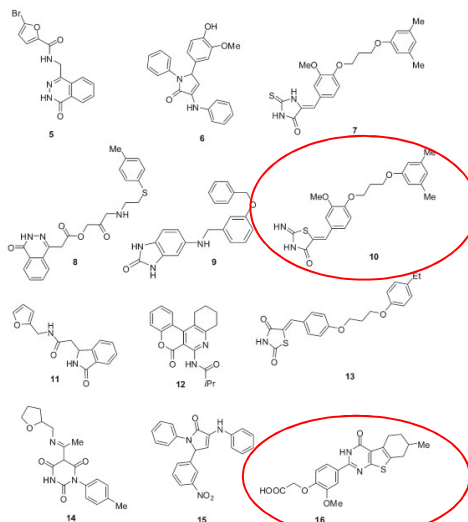


## Inhibitory kinazy FGFR

- ▶ Wszystkie związki (**2.2 mln**) z bazy ZINC zostały zadokowane w trybie standardowej precyzji do obu konformacji.
- ▶ **40 000** najlepiej ocenionych związków zadokowano ponownie w trybie wysokiej precyzji do obu konformacji.
- ▶ **1 000** najlepiej ocenionych związków zadokowano do innych kinaz (EGFR, InsR, VEGFR2, Src i MEK).
- ▶ **100** najlepiej ocenionych związków, które nie były dobrze ocenione dla innych kinaz.
- ▶ Przetestowano **24** związki.
- ▶ **2** związki wykazywały aktywność (IC<sub>50</sub> **23** i **50** μM).

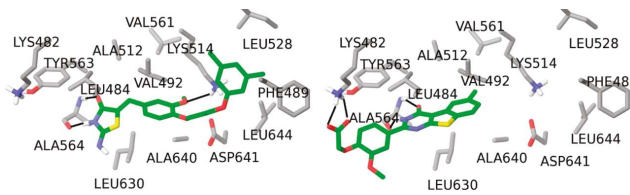
## Inhibitory kinazy FGFR

compd	XP rank	XP score	SP rank	IC <sub>50</sub> (μM) <sup>2</sup>
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



## Inhibitory kinazy FGFR

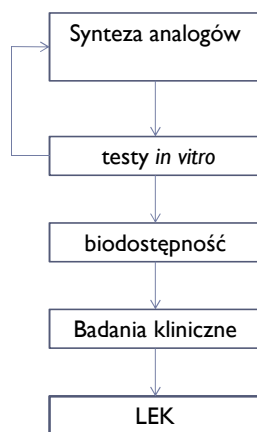
- Pomimo zastosowania procedury zwiększającej selektywność otrzymane związki nie wykazywały znaczącej selektywności.



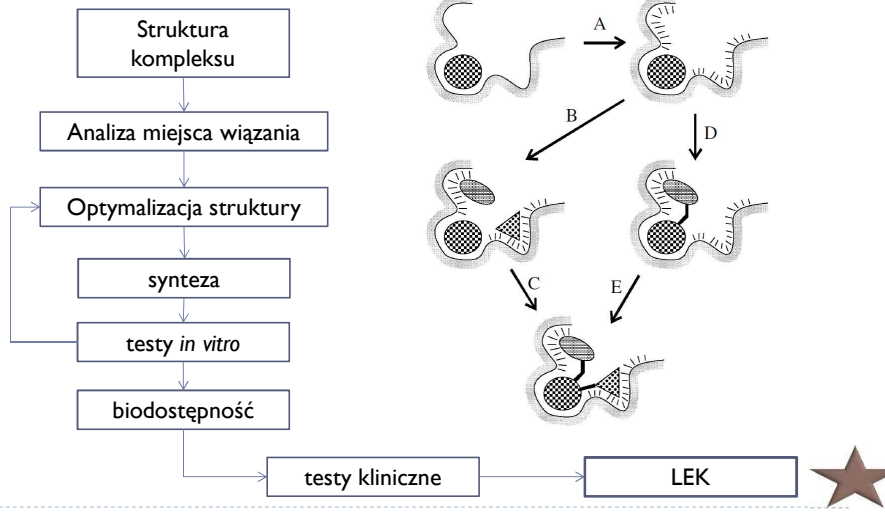
compd	IC <sub>50</sub> (μM)			
	FGFR1	EGFR	Src	InsR
10	23	56	10	47
40	1.9	2.4	1.9	na



## Optymalizacja – podejście klasyczne

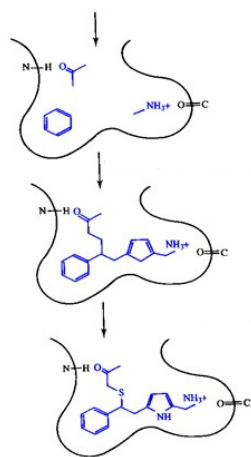


## Optymalizacja wspomaganą komputerowo

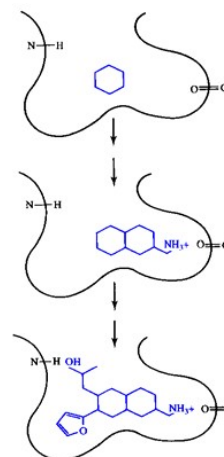


## Strategie optymalizacji

### ▶ Łączenie fragmentów



### ▶ Budowanie



## Strategie optymalizacji

- ▶ **Przylączenie atomów**
  - ▶ **Zalety:**
    - ▶ Duża różnorodność strukturalna związków
    - ▶ Efektywne wykorzystanie wszystkich możliwych oddziaływań
    - ▶ Brak konieczności użycia bazy danych fragmentów
  - ▶ **Wady:**
    - ▶ Zaprojektowane związki mogą być trudne do syntezy lub niestabilne
- ▶ **Przylączenie fragmentów**
  - ▶ **Zalety:**
    - ▶ Związki składają się z fragmentów, które są syntezowalne
  - ▶ **Wady:**
    - ▶ Różnorodność strukturalna związków zależy od wykorzystanej bazy

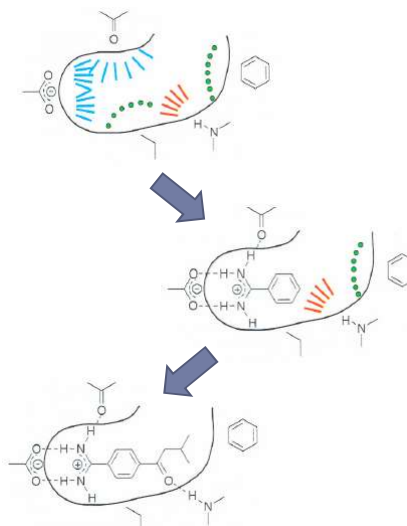
## Programy do projektowania

▶ <b>GENSTAR</b>	atomy	rozbudowuje cząsteczki na bazie struktury enzymu
▶ <b>GROUPBUILD</b>	fragmenty	wzrost sekwencyjny
▶ <b>GROW</b>	aminokwasy	wzrost sekwencyjny
▶ <b>GROWMOL</b>	fragmenty	wzrost sekwencyjny
▶ <b>HOOK</b>	fragmenty	poszukiwanie stochastyczne
▶ <b>LEGEND</b>	atomy	łączenie fragmentów
▶ <b>LUDI</b> kombinatoryczne	fragmenty	poszukiwanie stochastyczne
▶ <b>MCSS</b>	fragmenty	poszukiwanie
▶ <b>MCSS</b>	fragmenty	próbkowanie stochastyczne
▶ <b>PRO-LIGAND</b>	fragmenty	wzrost sekwencyjny
▶ <b>PRO-SELECT</b>	fragmenty	podejście szkielet-linker
▶ <b>SKELGEN</b>	małe fragmenty	poszukiwanie Monte-Carlo
▶ <b>SPROUT</b>	fragmenty	wzrost sekwencyjny
		poszukiwanie kombinatoryczne

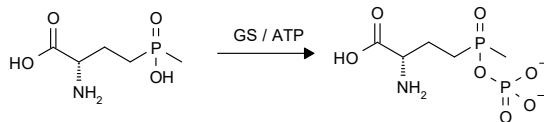
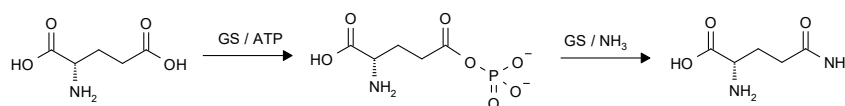


## Program LUDI

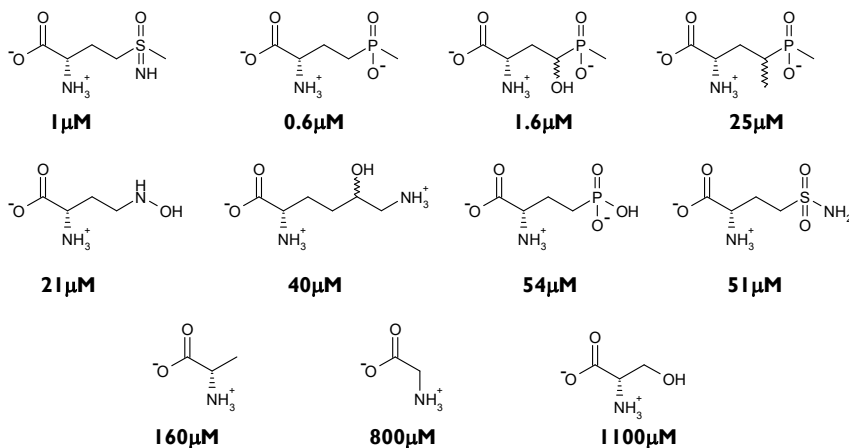
- ▶ Znajdowanie potencjalnych miejsc oddziaływania:
  - ▶ Lipofilowe (zielone)
  - ▶ Donor wiązania wodorowego (czerwone)
  - ▶ Akceptor wiązania wodorowego (niebieskie)
- ▶ Nakładanie fragmentów na miejsca oddziaływania
- ▶ Przyłączanie kolejnych fragmentów (nakładających się z cząsteczką bazową i miejscami oddziaływania)



## Inhibitory syntetazy glutaminy

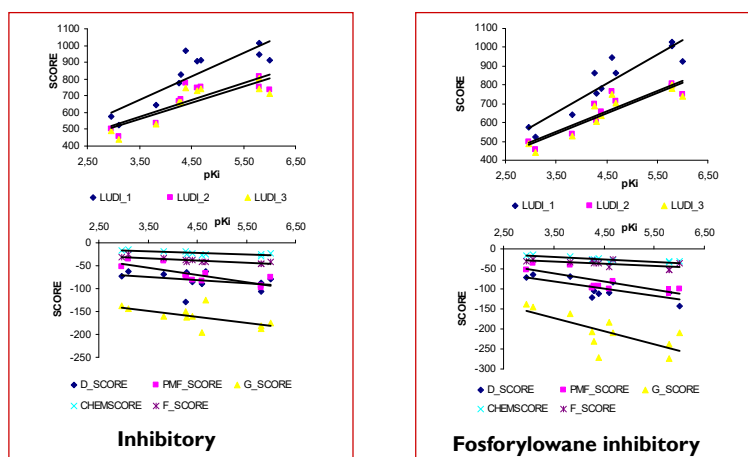


## Inhibitory syntetazy glutaminy



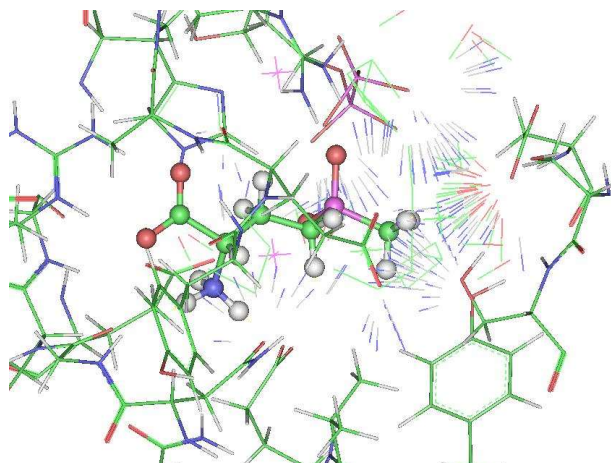
## Inhibitory syntetazy glutaminy

### ► Ocena kompleksów inhibitor-enzym



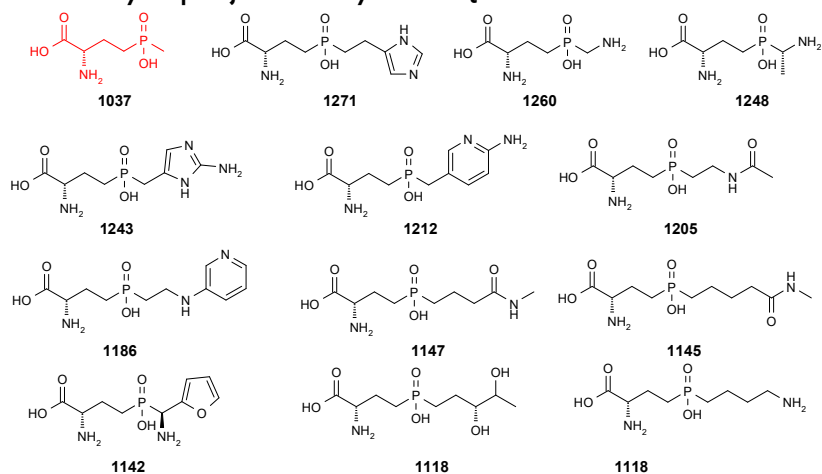
## Inhibitory syntetazy glutaminy

### ► Miejsca oddziaływania (program LUDI)



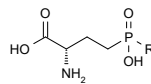
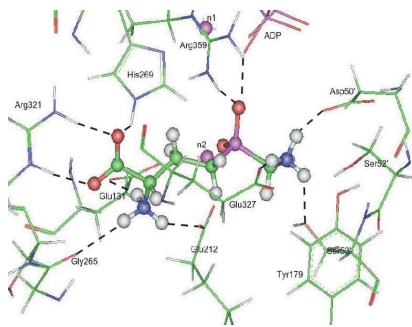
## Inhibitory syntetazy glutaminy

### ► Struktury zaprojektowanych związków



## Inhibitory syntetazy glutaminy

### ► Struktura kompleksu enzym-inhibitor



	$K_i$ [ $\mu\text{M}$ ]		
	<i>E. coli</i>	<i>M. tuberculosis</i>	kukurydza
-CH <sub>3</sub>	0.6	ND	1.1
-CH <sub>2</sub> NH <sub>2</sub>	0.59	0.8	1.8
-CH <sub>2</sub> OH	2.1	5.7	8.5
-CH(CH <sub>3</sub> )NH <sub>2</sub>	33	3.4	21.6
-CH <sub>2</sub> CH <sub>2</sub> NH <sub>2</sub>	55	ND	320
-CH <sub>2</sub> CH <sub>2</sub> COOH	970	ND	ND

## Podsumowanie

- Wspomagane komputerowo poszukiwanie nowych cząsteczek aktywnych jest znacząco wydajniejsze niż klasyczne techniki HTS.
- Projektowanie nowych cząsteczek prowadzi do struktur trudnych do otrzymania innymi metodami.