



Metody dokowania ligandów



Strategie projektowania leków

Ligand-based drug design

nieznana

Struktura
celu
molekularnego

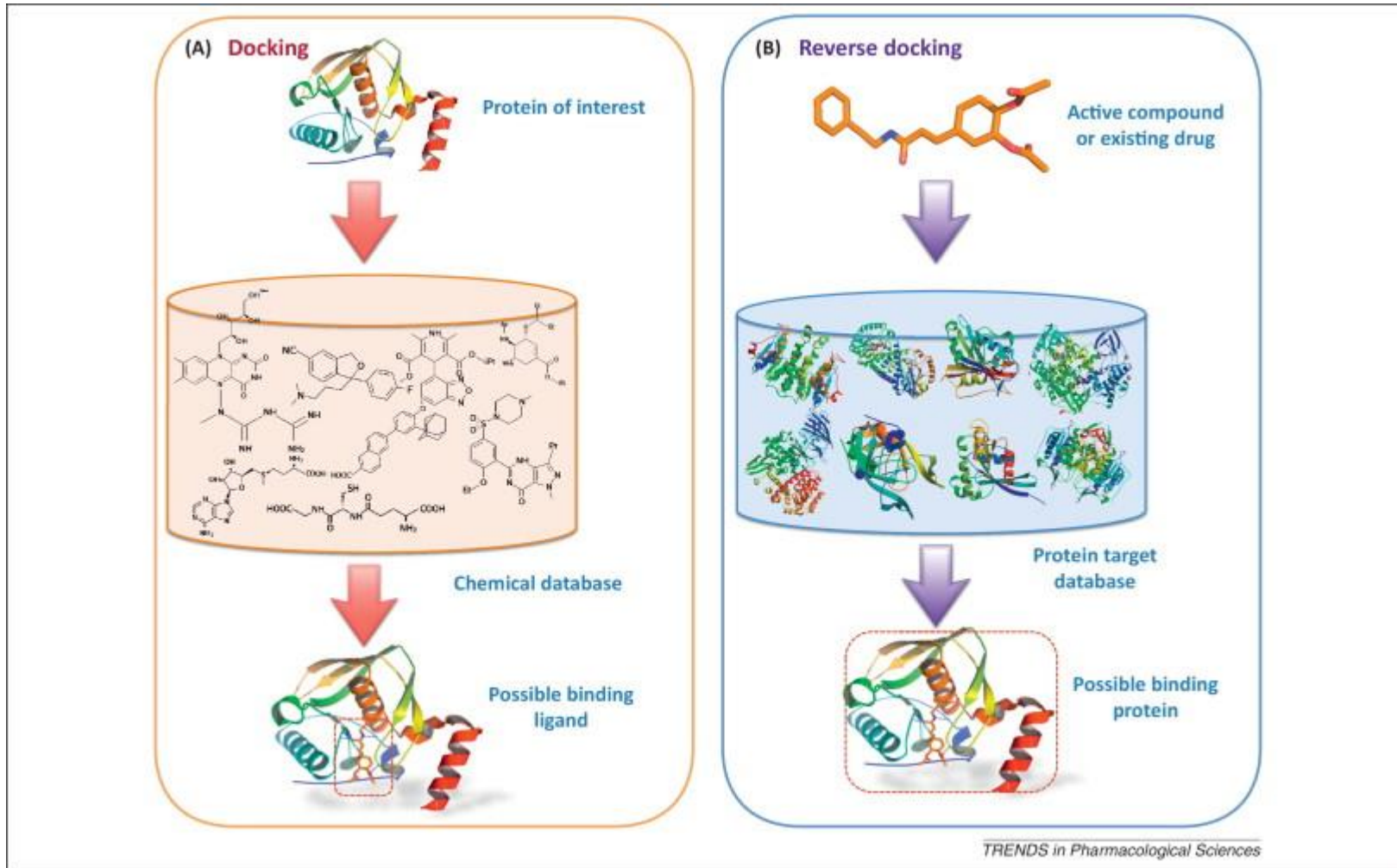
znana

Receptor-based drug design

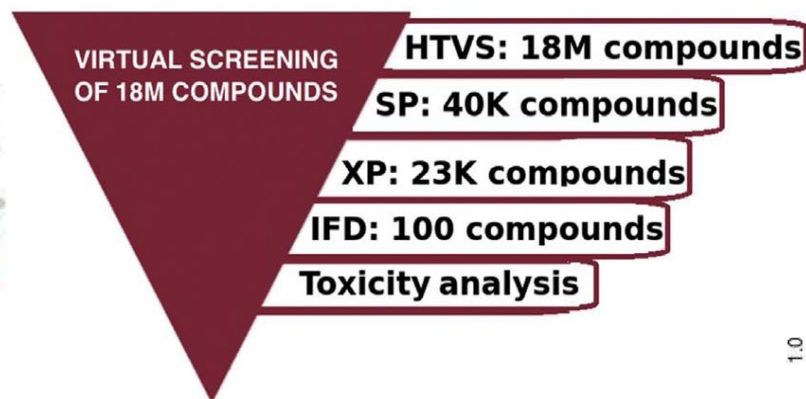
- Budowanie modelu miejsc aktywnych liganda (**farmakofofor**)
- Przeszukiwanie baz danych (**screening**)
- 1D i 3D **QSAR**
(pseudoreceptory i pola molekularne)

- Dopasowanie ligandów do miejsca aktywnego receptora (**dokowanie**)
- Budowa nowych ligandów *ab-initio*
- **Dynamika** kompleksu receptor-ligand

Dokowanie i odwrotne dokowanie

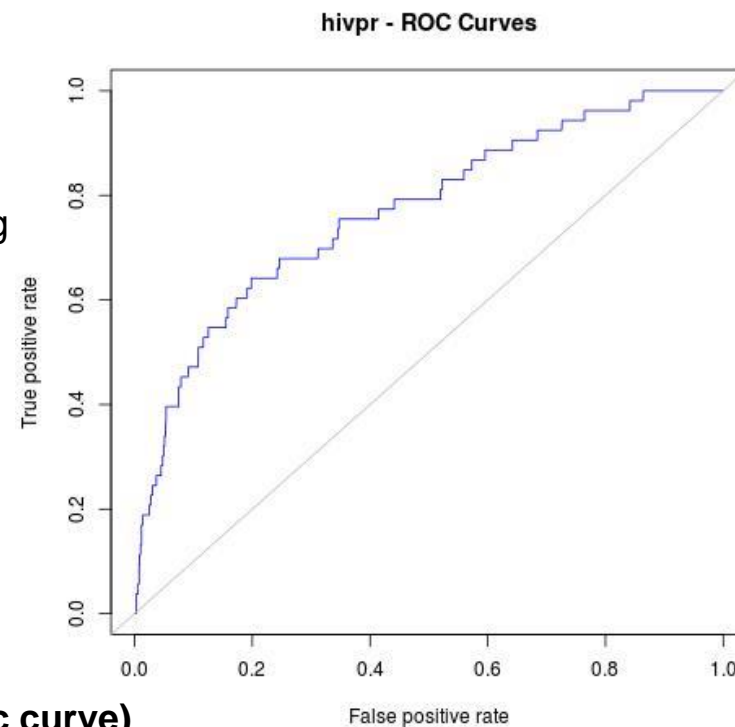


Przesiewanie baz danych ligandów



Pakiet Schrödinger:

HTVS – High-Throughput Virtual Screening
 SP – Standard Precision
 XP – Extra Precision
 IFD – induced Fit Docking

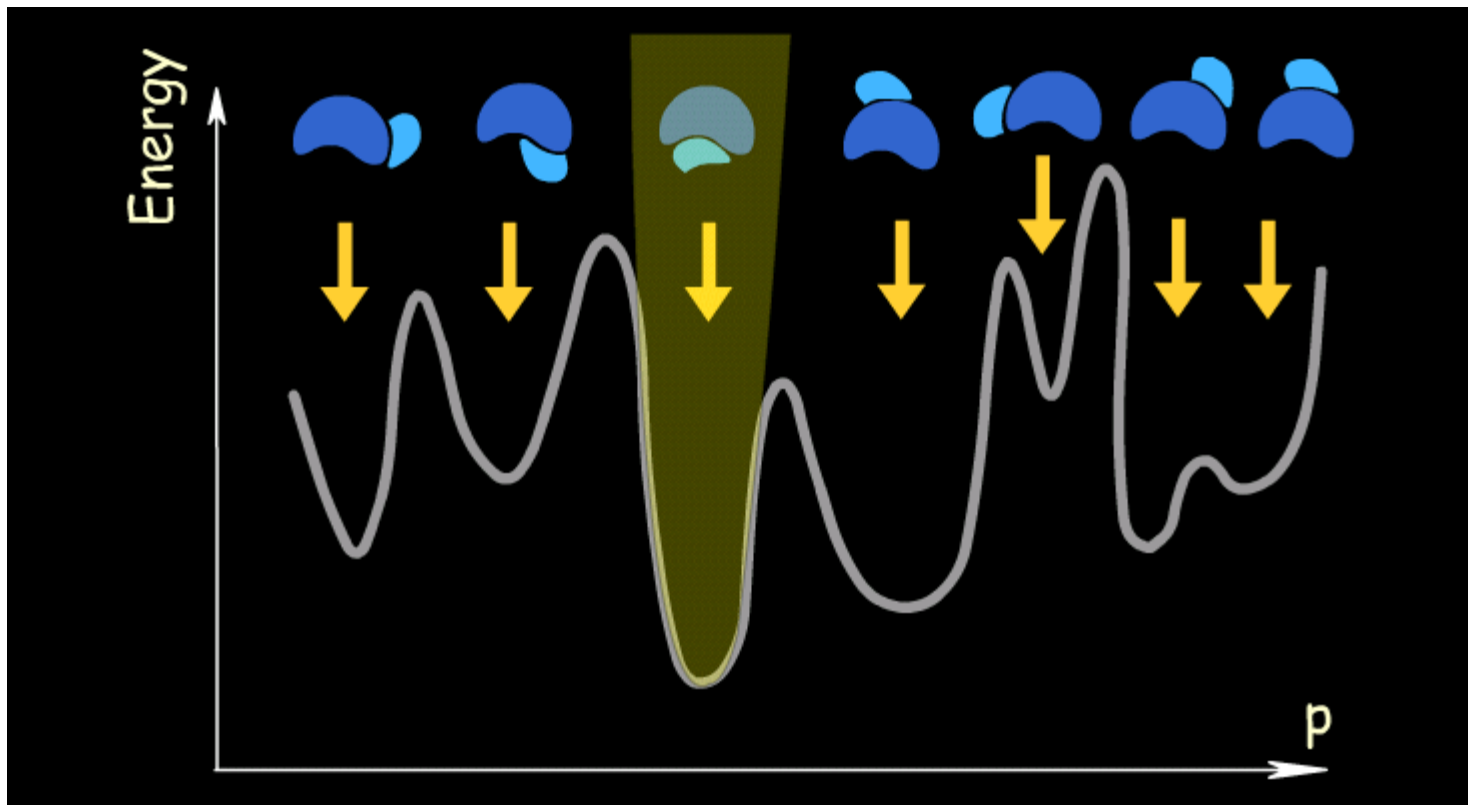


ROC curve (receiver operating characteristic curve)

– ocena jakości algorytmu do klasyfikacji (związek aktywny/nieaktywny)



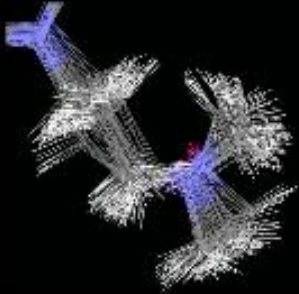
Dokowanie

znajdowanie struktury kompleksu o najniższej energii (najbardziej dopasowany)



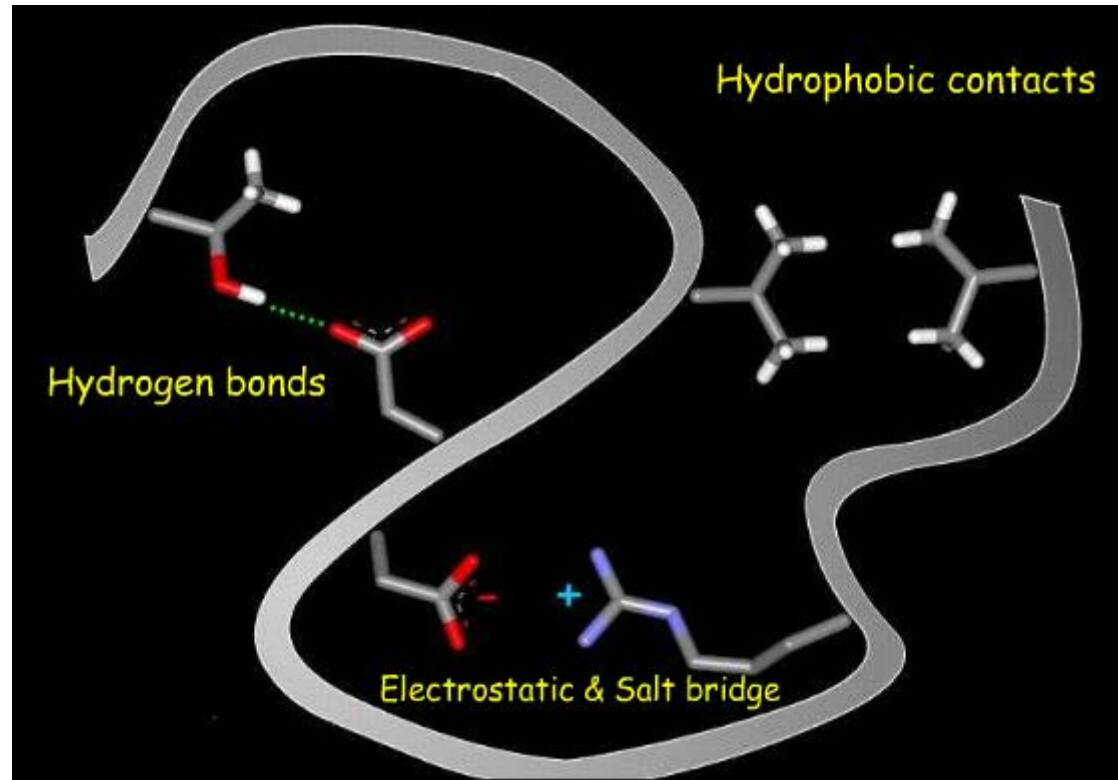
Energia lub tzw. „scoring function” (kształt i właściwości fizykochemiczne)

Czynniki wpływające na wiązanie ligandów

- Interaction energies 
- Solvation 
- Entropy 

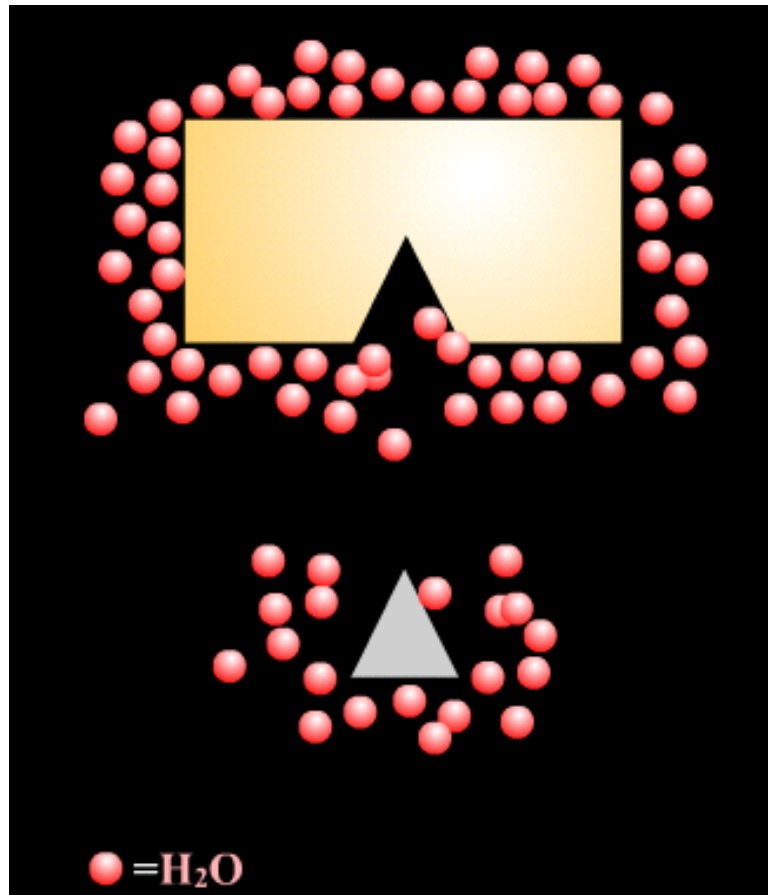
Korzystne oddziaływania

Oddziaływania
kierunkowe
i silne



Oddziaływania
van der Waalsa:
bezkierunkowe
i słabe lecz liczne

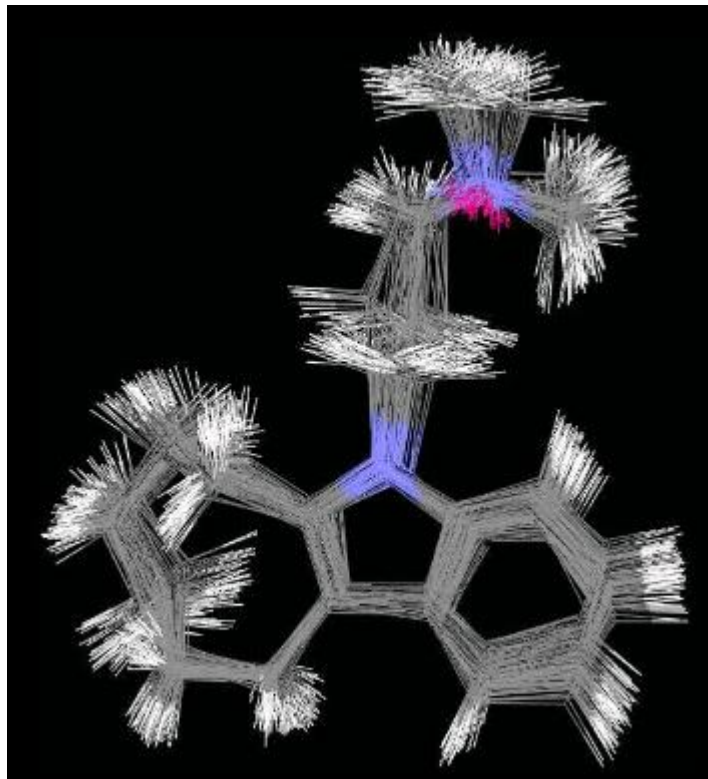
Energie częściowej desolvatacji liganda i miejsca aktywnego



Proces kierowany termodynamicznie
(efekty entropowe)

Części hydrofobowe oddziałują ze sobą aby zmniejszyć oddziaływania z otaczającą wodą

Konformacyjne efekty entropowe



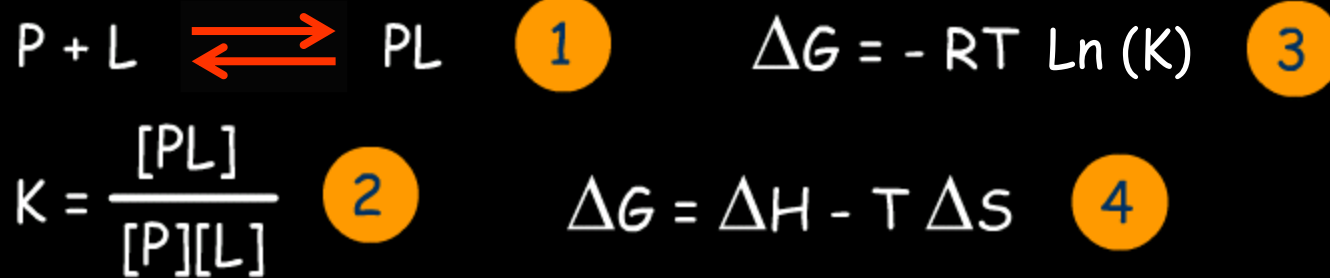
Konformacja bioaktywna (w miejscu aktywnym) może być inna niż na zewnątrz

Dotyczy także konformacji białka: głównie konformacji łańcuchów bocznych aminokwasów z miejsca aktywnego

Konieczne jest dokowanie wielu różnych konformacji liganda



Wiązanie liganda jako reakcja równowagowa



P = Protein L = Ligand

R = Gas constant (1.987 cal/mol*K)

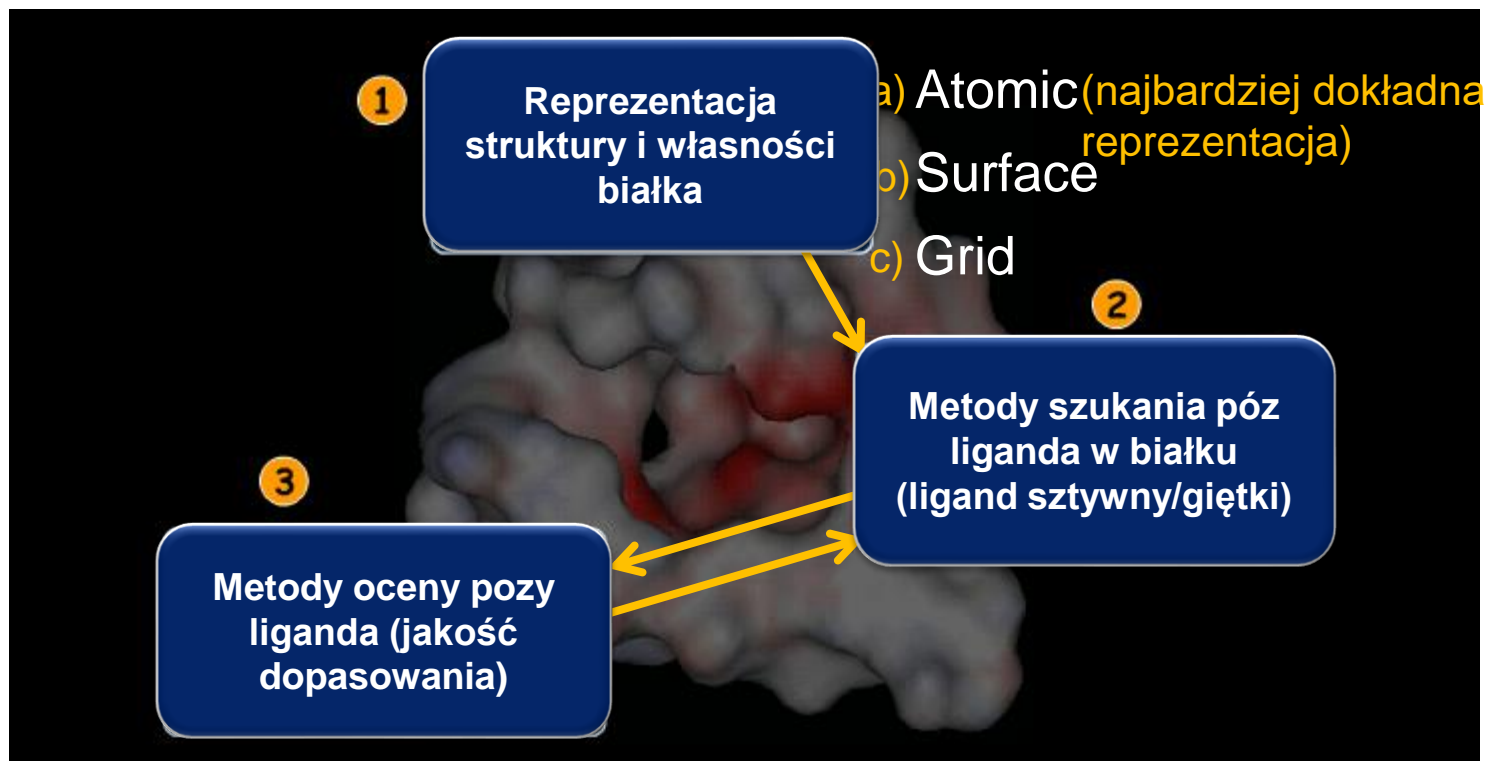
T = absolute temperature in degrees Kelvin (25° = 298 K)

Stała K jest mierzona doświadczalnie.

K_F – stała tworzenia [M^{-1}]

K_D – stała dysocjacji [M]

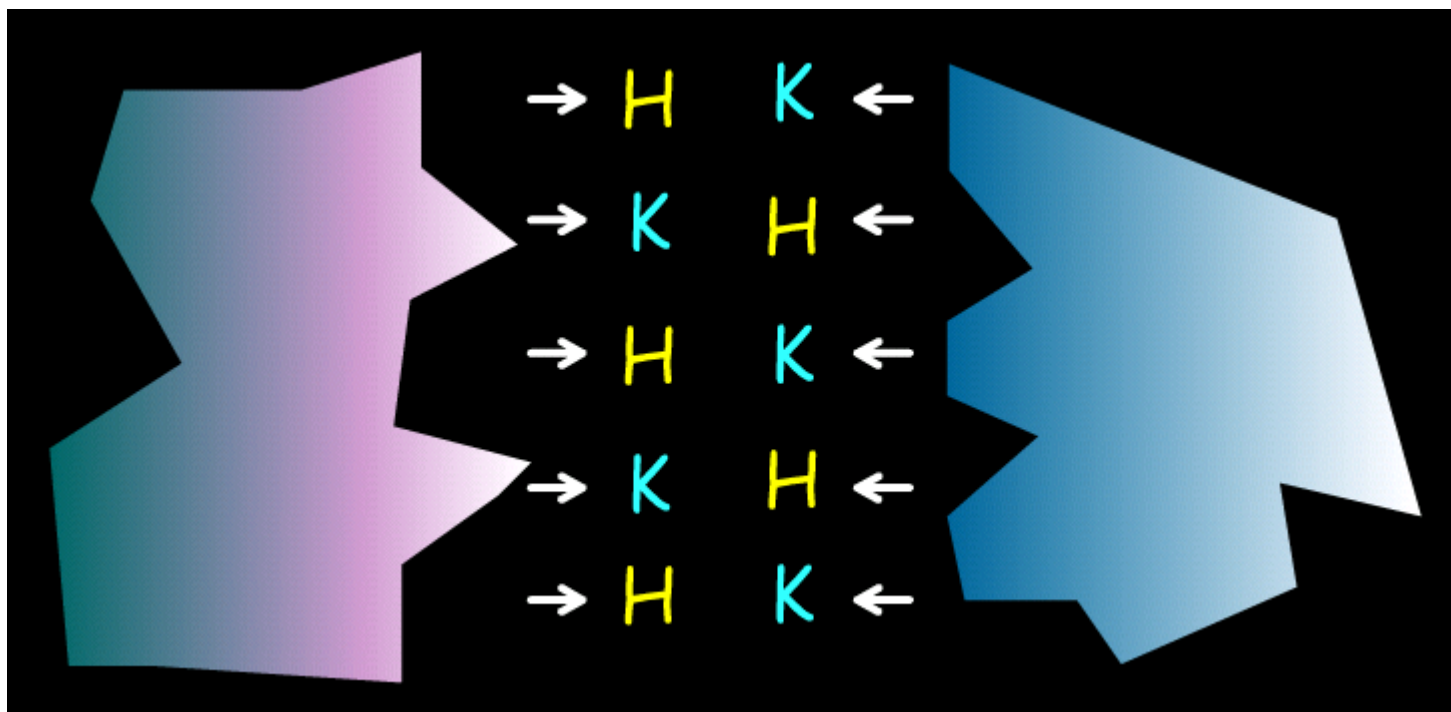
Trzy składniki programów do dokowania



Algorytmy szukania konformacji i wyznaczania energii są ściśle ze sobą związane

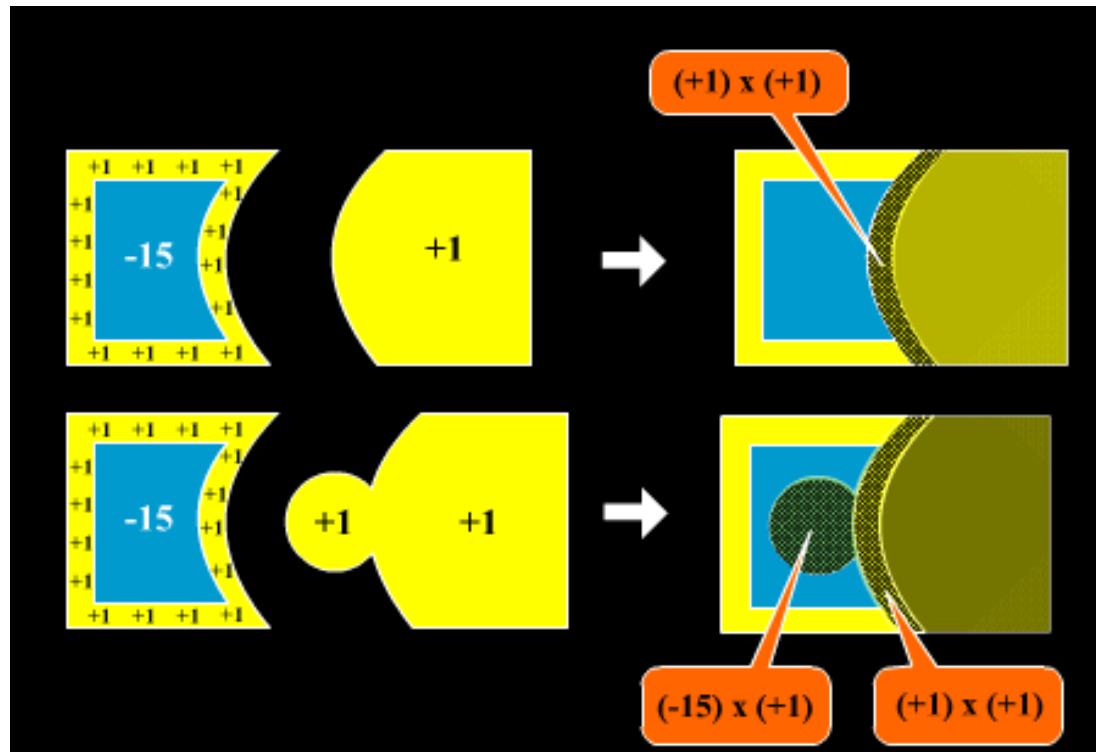
1. Reprezentacja struktury i własności białka

b) Dokowanie przez komplementarność powierzchni



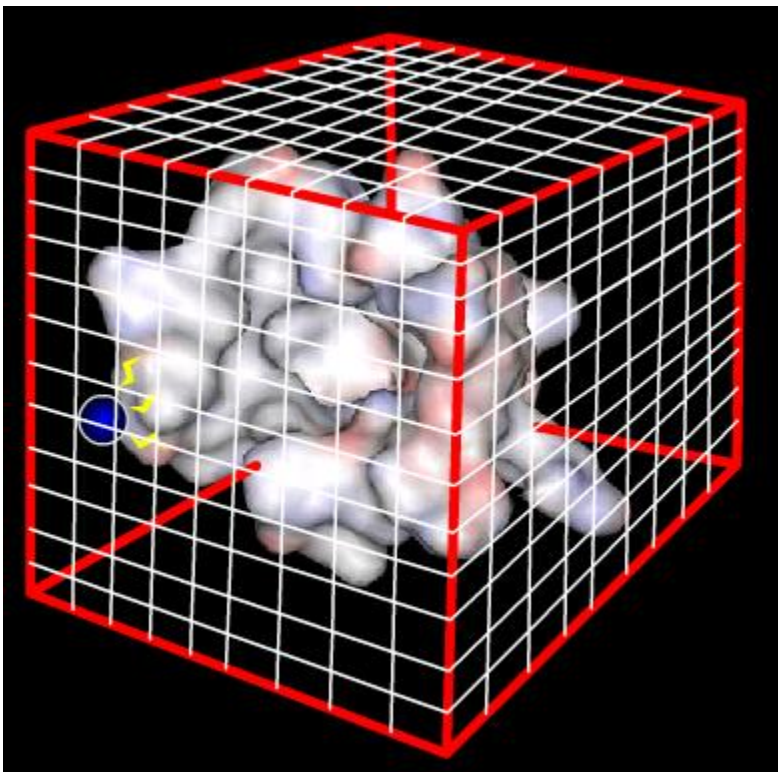
H – hole, K – knob; sposób używany w b. wielu programach dokujących

Prosta metoda szacowania komplementarności powierzchni



Metoda jest oparta na 3D grid

c) Reprezentacja białka przez grid (sieć)



W każdym punkcie sieci mierzymy wielkość oddziaływań niewiążących:

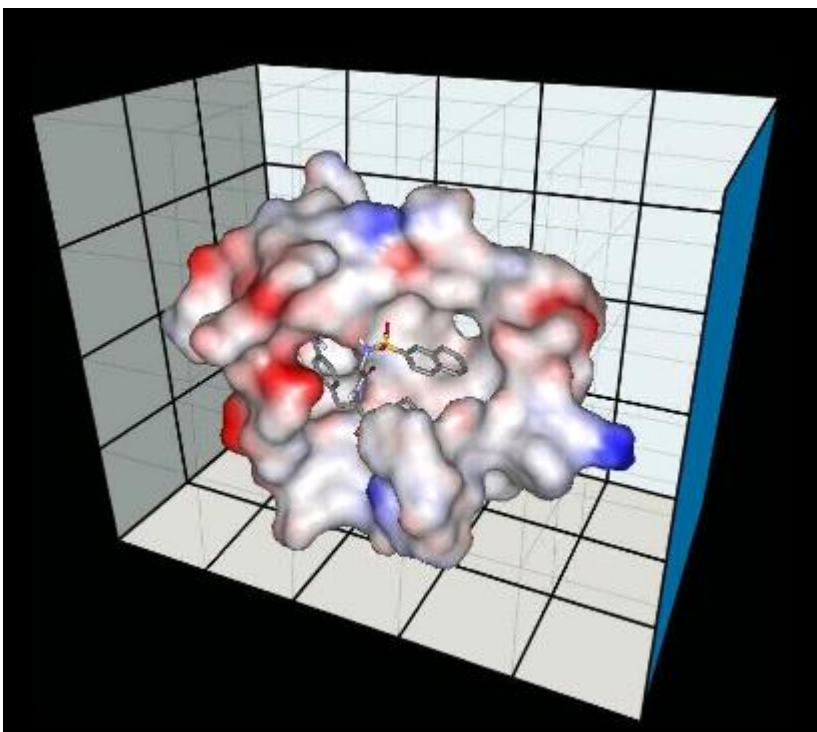
- Van der Waalsa
- Elektrostatyczne

Oczka sieci ok. 2 Å

Zmierzone wielkości są zapamiętywane w tabeli

Pierwsze użycie: Goodford 1985 r.

Użycie gridu do uproszczenia dokowania



Dla każdego atomu liganda liczone jest oddziaływanie od otaczających punktów sieci (odcięcie *cutoff* ok. 10 Å)

Dodatkowe przyspieszenie:

Dla każdego typu atomu liganda w każdym węźle sieci energia jest obliczana i przechowywana w tabeli

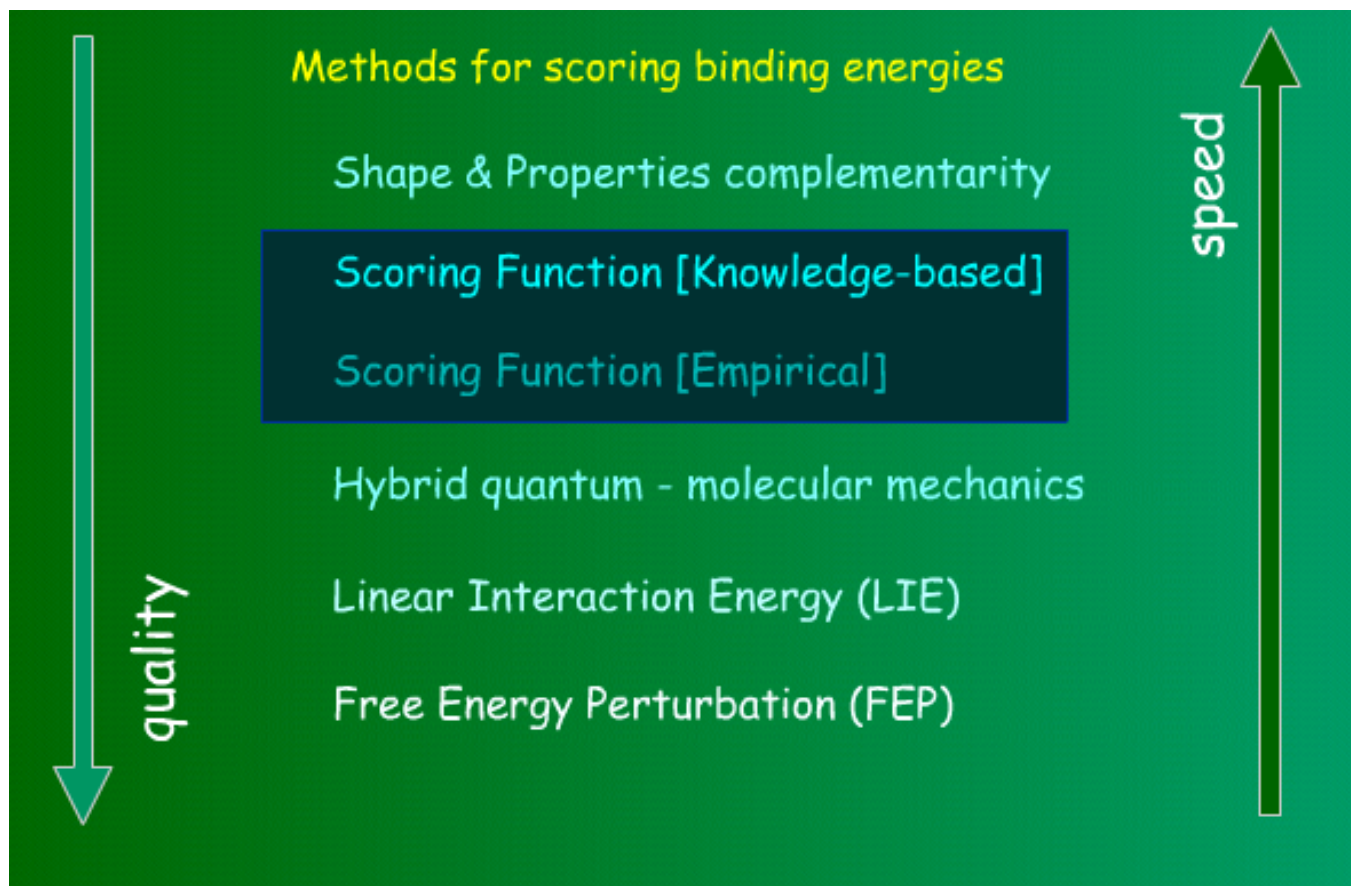
Potencjał dla danego atomu liganda jest obliczany przez interpolację z otaczających punktów gridu

Metoda często używana w symulacjach typu *virtual reality*

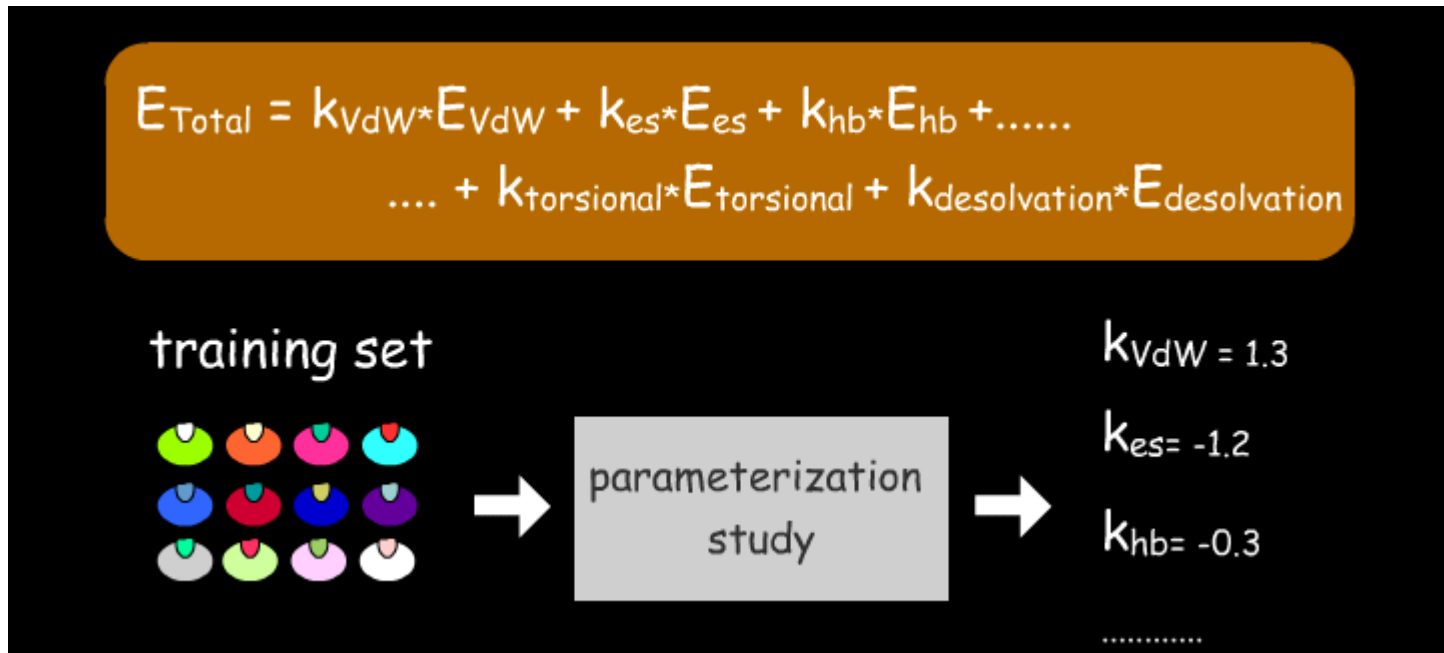


3. Funkcje oceniające wiązanie ligandów

dokładność vs. szybkość



Empiryczne funkcje oceniające



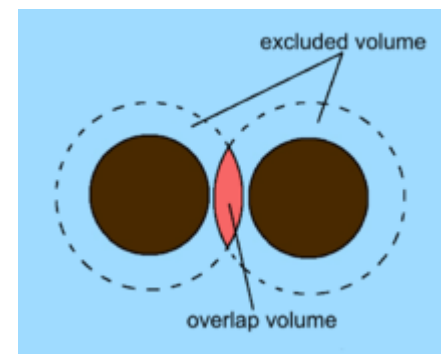
Współczynniki wagowe k uzyskuje się w procesie parametryzacji (kalibracji) wybranego pola siłowego (k mogą być dodatnie lub ujemne)

Pole siłowe skalibrowane dla kompleksów ze zbioru treningowego:
znane są struktury 3D i aktywności biologiczne ligandów w tych kompleksach

Przykład empirycznej funkcji oceniającej z programu AutoDock

Energy Term	Coefficient
Van der Waals	0.148
electrostatic	0.114
hydrogen bond	0.065
torsional rotation	0.311
desolvation	0.171

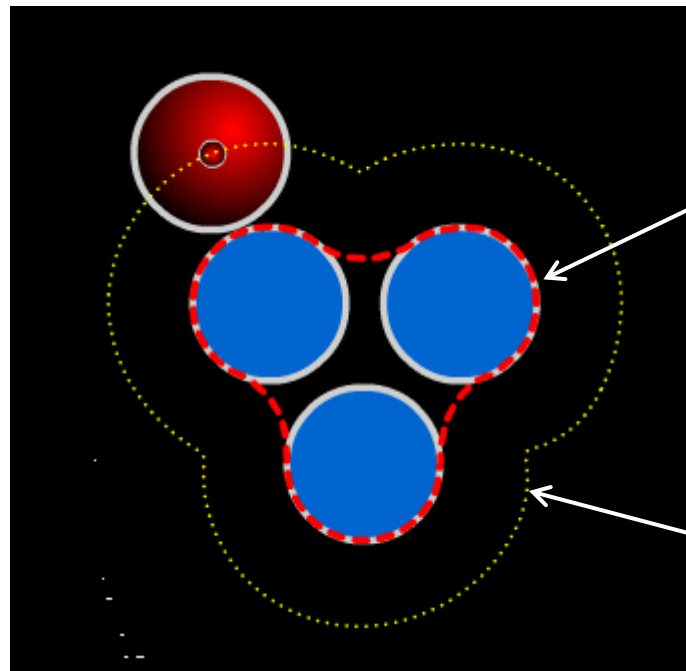
Coefficients of the AutoDock 3
Free Energy Model



(obliczane z *solvent excluded volumes*)

Obliczane są tylko niektóre składowe pola siłowego

Powierzchnia dostępna dla rozpuszczalnika



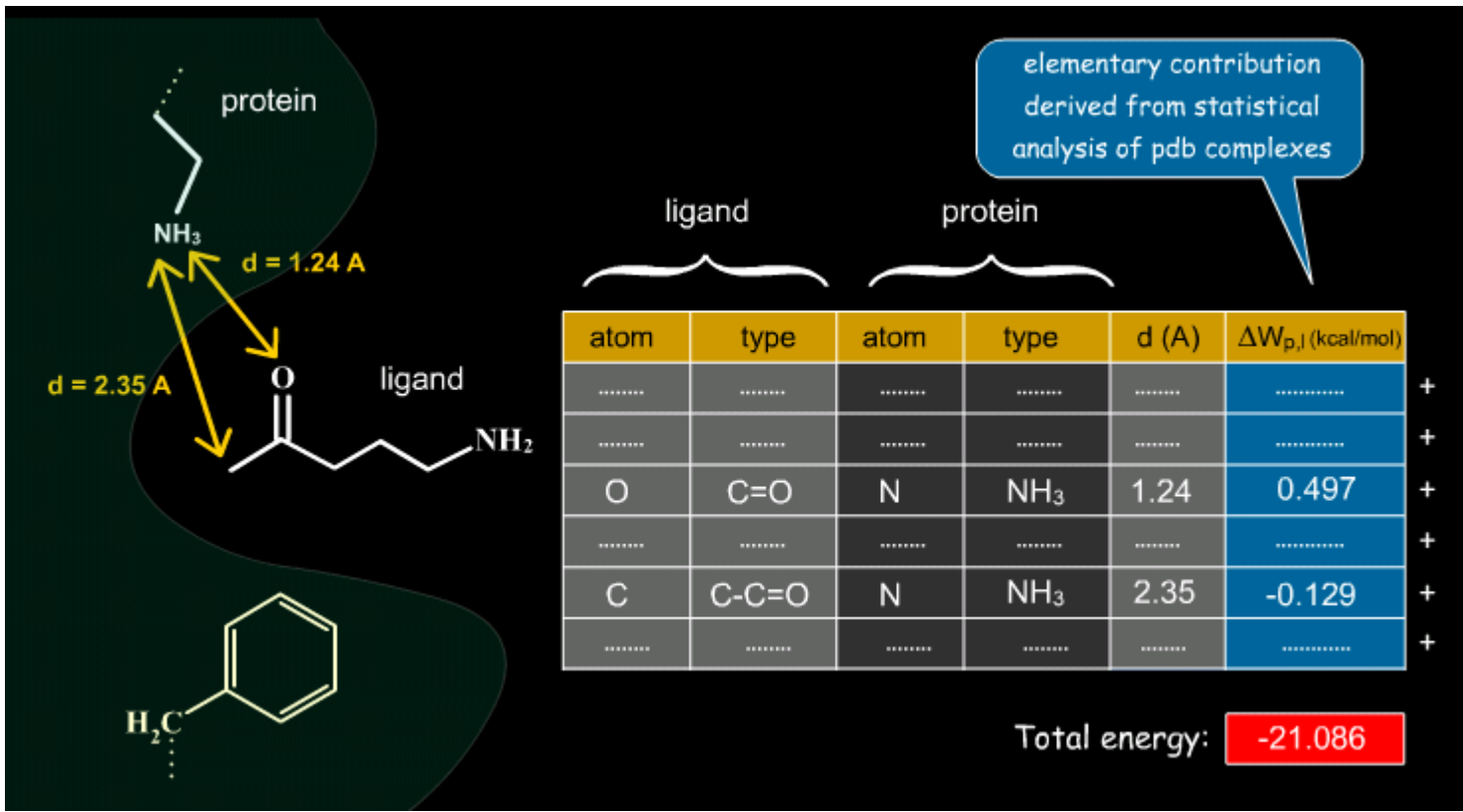
Connolly surface
(wygładzona powierzchnia
molekularna)

Solvent accessible
surface area

Promień kulki dla wody = 1.4 Å

Funkcje oceniające oparte na wiedzy

(*knowledge based scoring functions*) - wykorzystanie

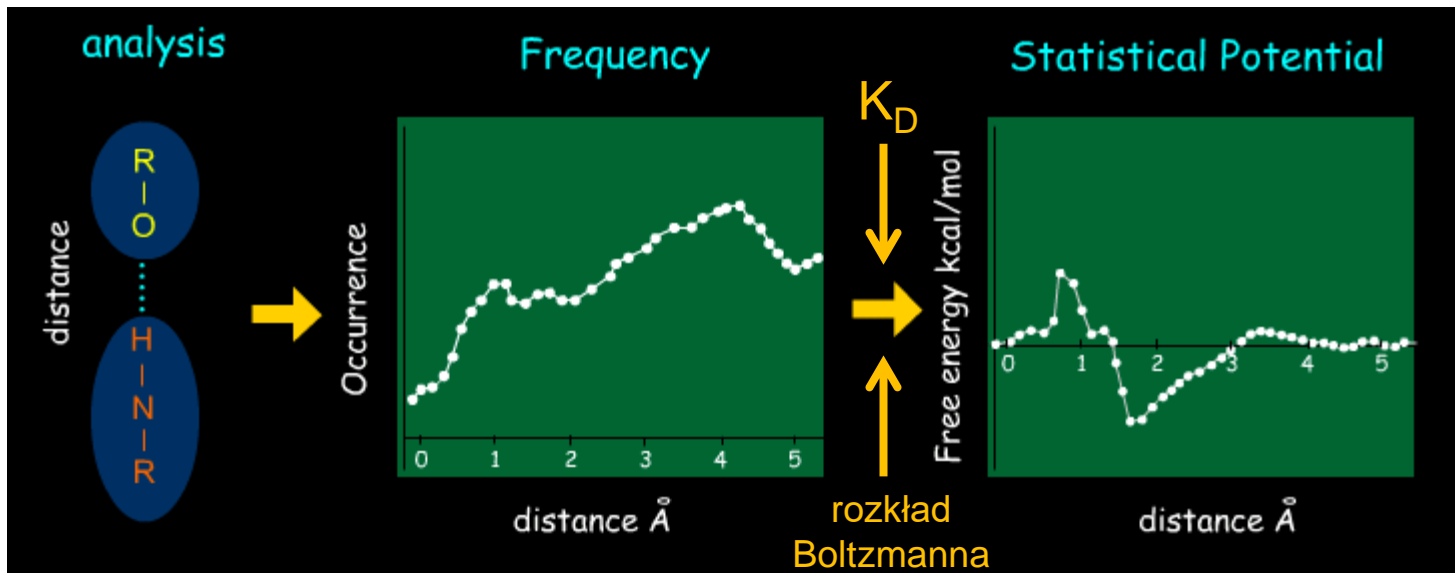


Bazują na odległościach w strukturach 3D kompleksów ligand-cel molekularny

Funkcje oceniające oparte na wiedzy

otrzymywanie

Dla każdej pary oddziałujących atomów (typów atomów) obliczany jest wykres PMF (Potential of Mean Force)



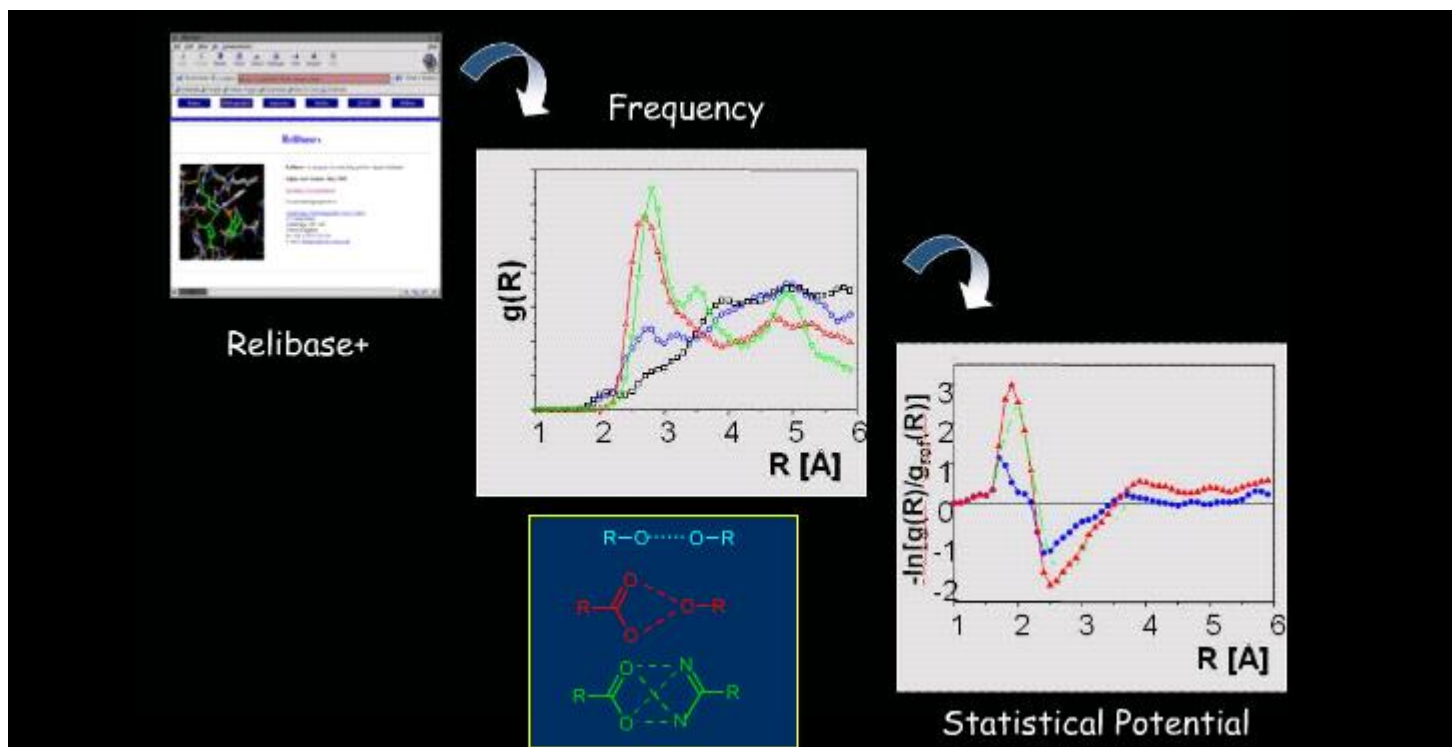
$$g_{ij} \sim \exp(-E_{ij}/kT)$$

$$E_{ij} \sim \ln(g_{ij}) * kT$$

$$\Delta E_{ij} \sim \ln(g_{ij} / g_{ref})$$

Dla 126 typów atomów w białku i 34 typów atomów w ligandach obliczanych jest 4284 wykresów PMF

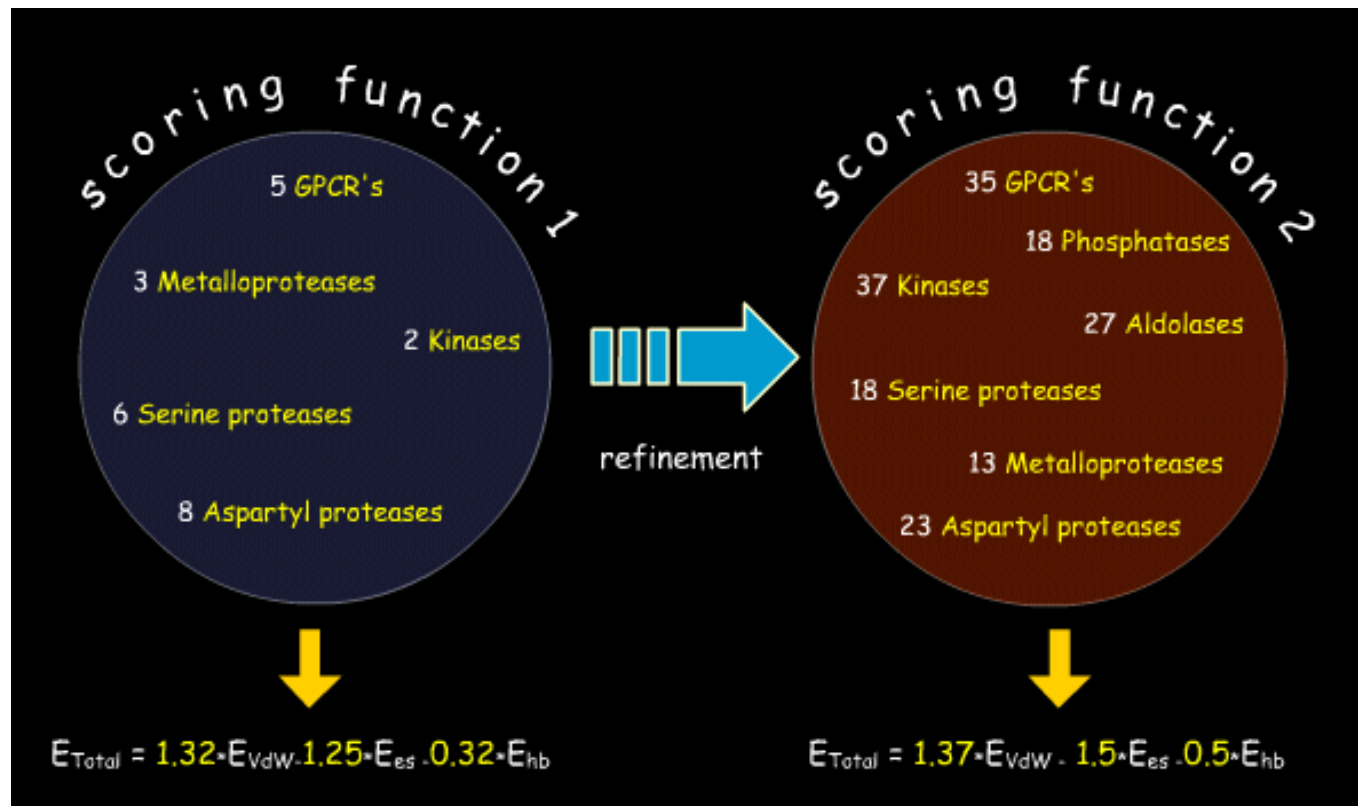
Przykład użycia funkcji oceniającej opartej na wiedzy w programie DrugScore



1400 kompleksów ligand-białko z bazy PDB (G. Klebe, Marburg University)

$G_{ref}(R)$ – wykres dla wszystkich ligandów (kolor czarny)

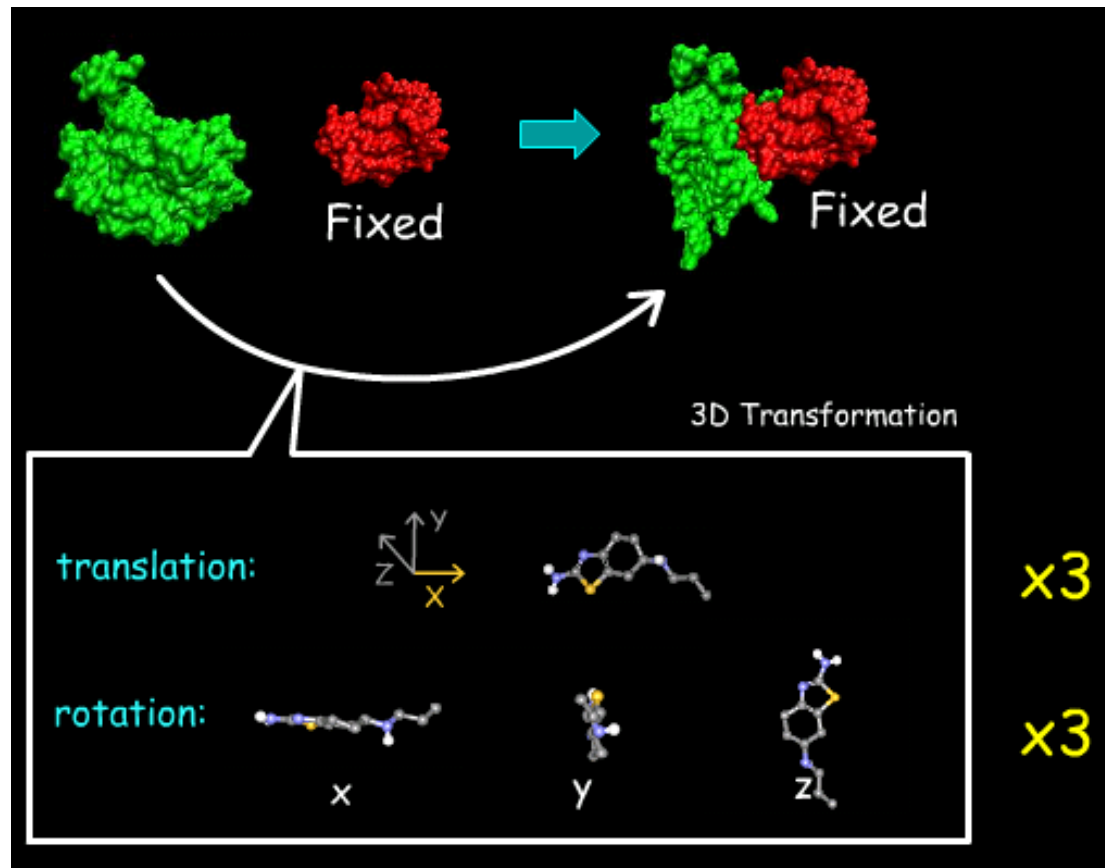
Funkcje oceniające są wrażliwe na wybór kompleksów ligand-białko do zbioru treningowego



Większy zbiór kompleksów zapewnia większą dokładność funkcji oceniającej dla średniego białka

2. Metody szukania póz liganda w białku

Zmiana położenia sztywnego liganda w miejscu aktywnym

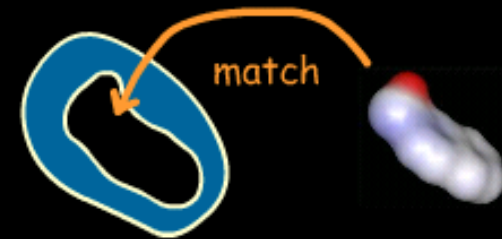


Ligand i białko traktowane jako sztywne obiekty

Podział algorytmów dokujących

- Feature-Based Matching Methods

Algorytm szuka komplementarnych własności

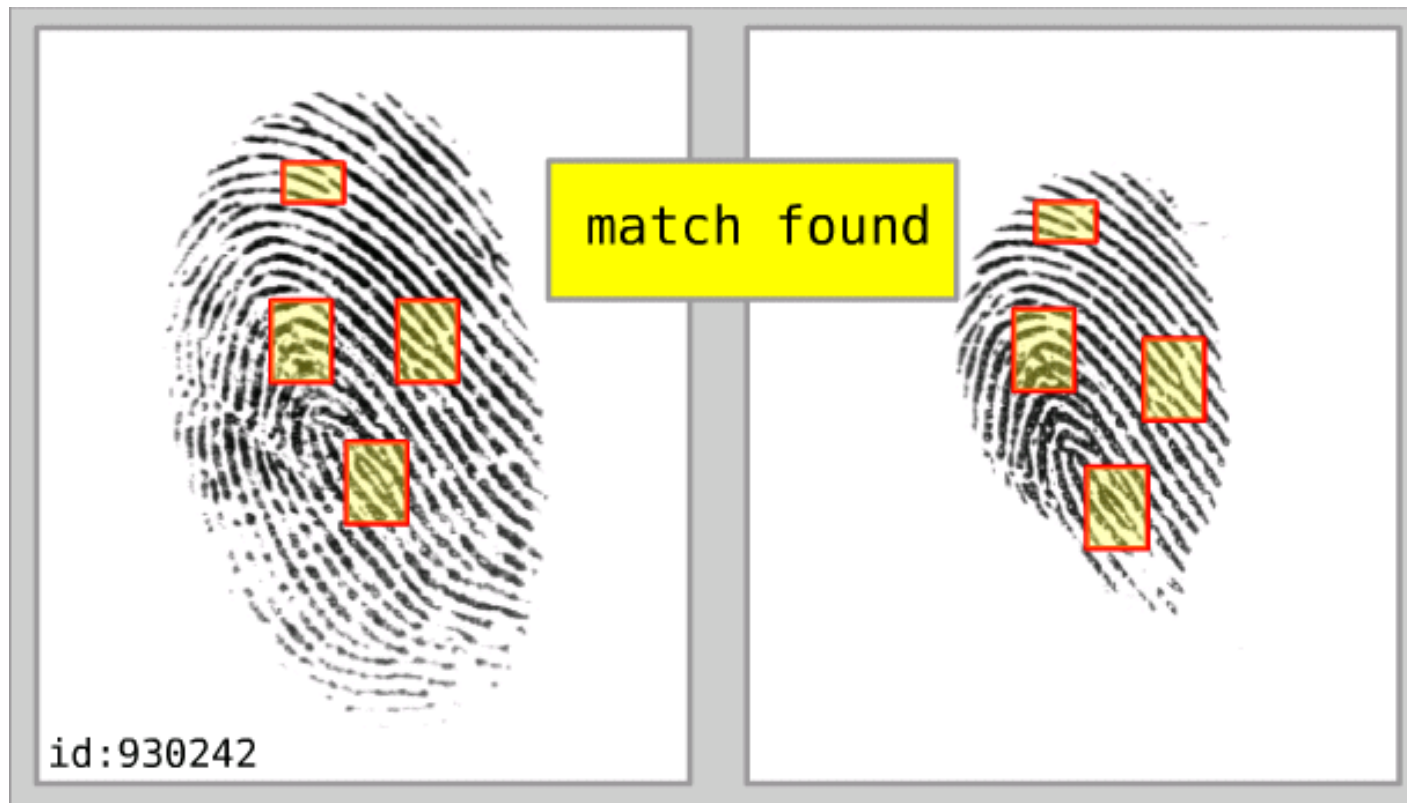


- Stepwise Search Methods

Algorytm eksploruje całą dostępną przestrzeń

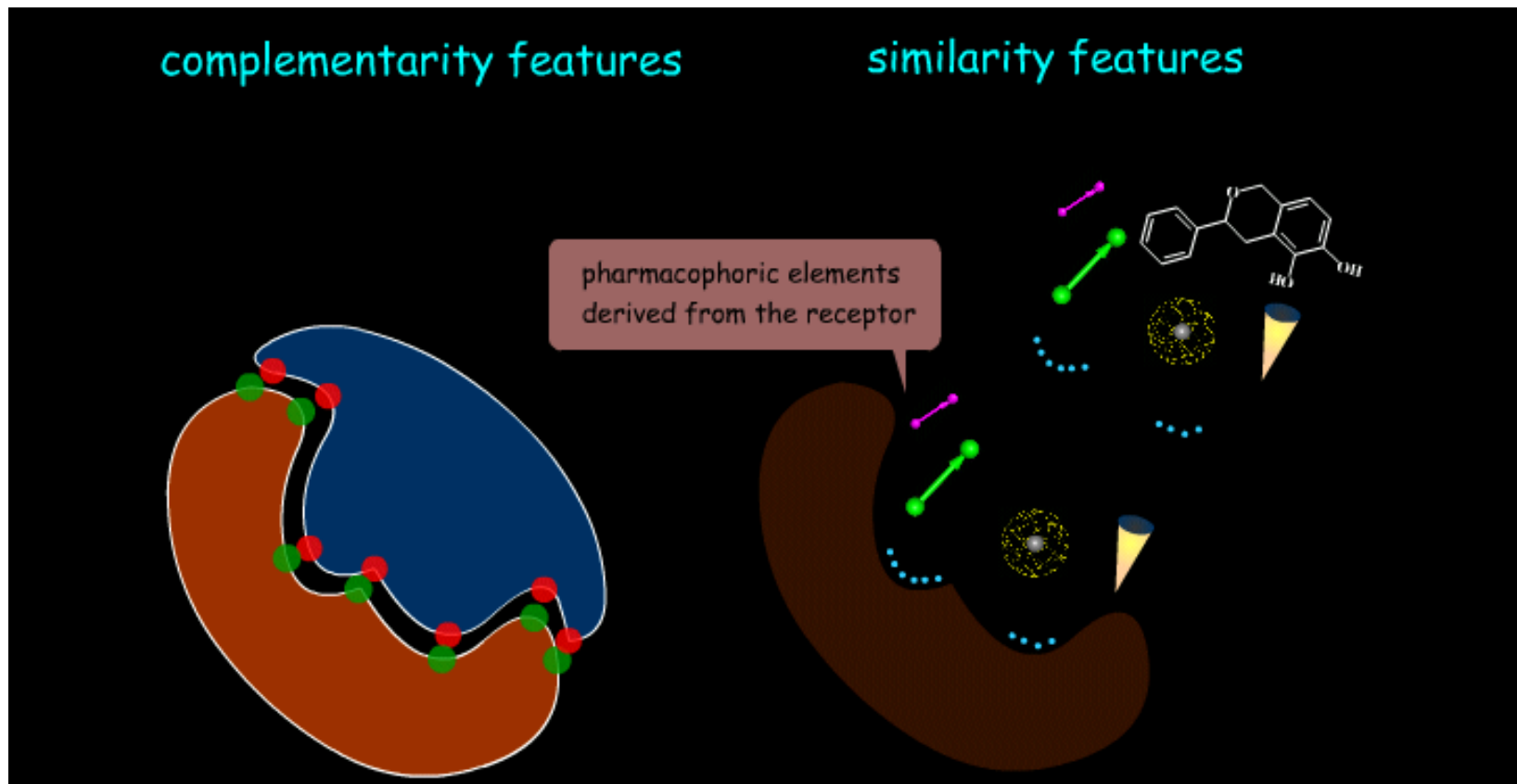


Feature-based matching - idea

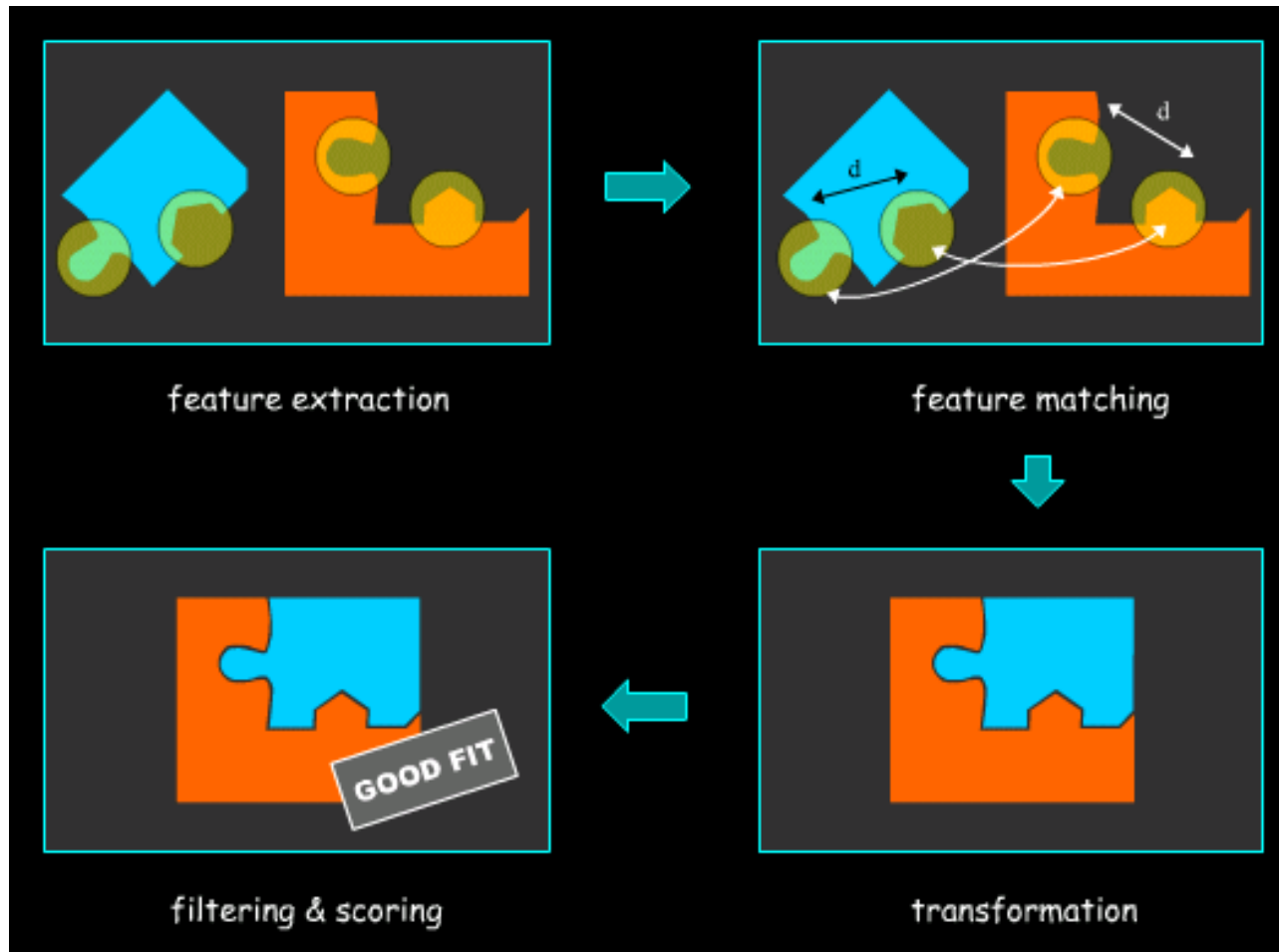


Pattern recognition (algorytmy z analizy obrazów, robotyki, ...)

Dopasowanie oparte na komplementarności i podobieństwie



Etapy *feature-based matching*



Podobne do układania puzzli



Programy stosujące *feature-based matching*

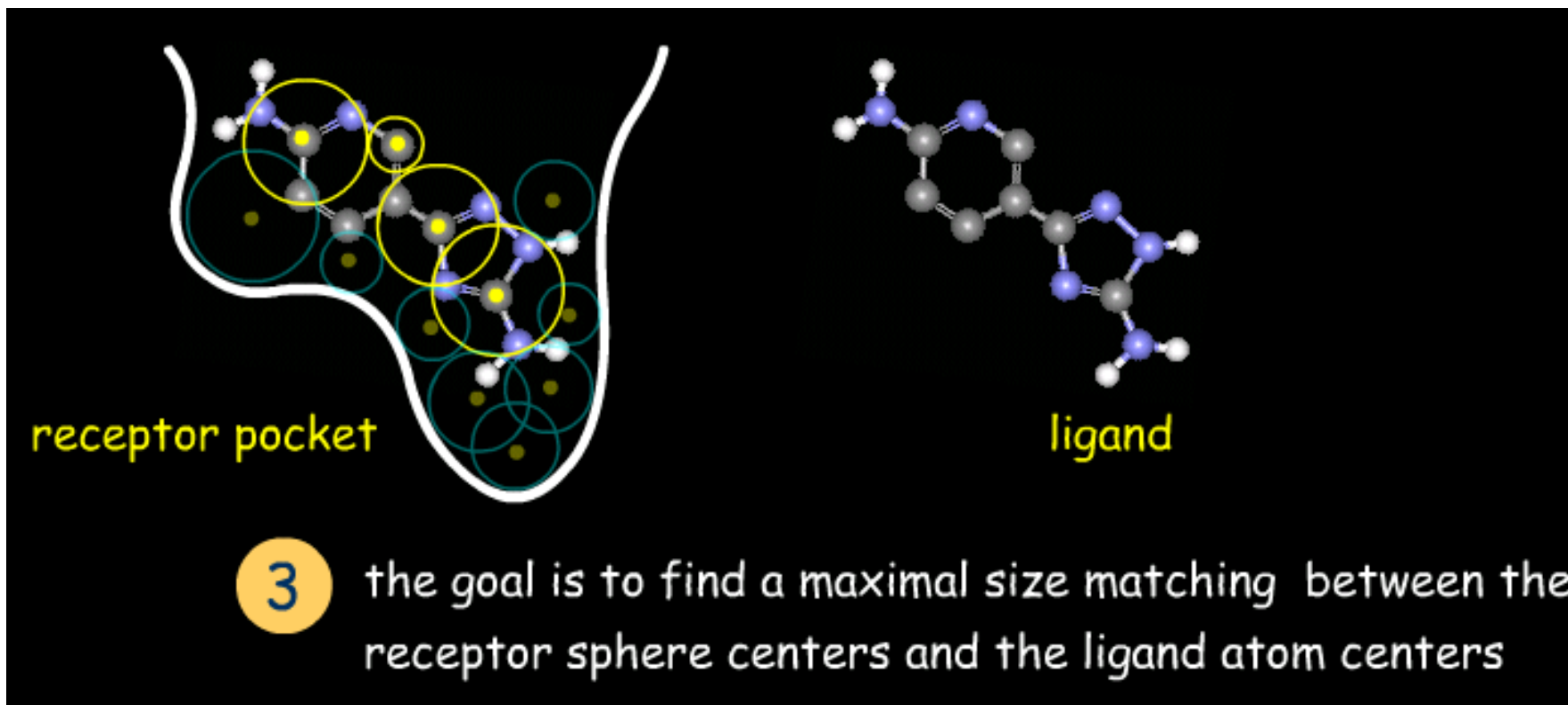
DOCK	distance matching	docking
FLEX	interaction site matching	docking
FRED	shape matching	docking
CAVEAT	vector matching	de novo design
ADAM	H-bond matching	docking
NEW LEAD	vector matching	de novo design
CLIX	steric effect matching	docking
SANDOCK	shape and chemical matching	docking
LIGIN	shape matching	docking
LUDI	interaction site matching	de novo design

Pierwszy program DOCK z 1982 r.



Program DOCK

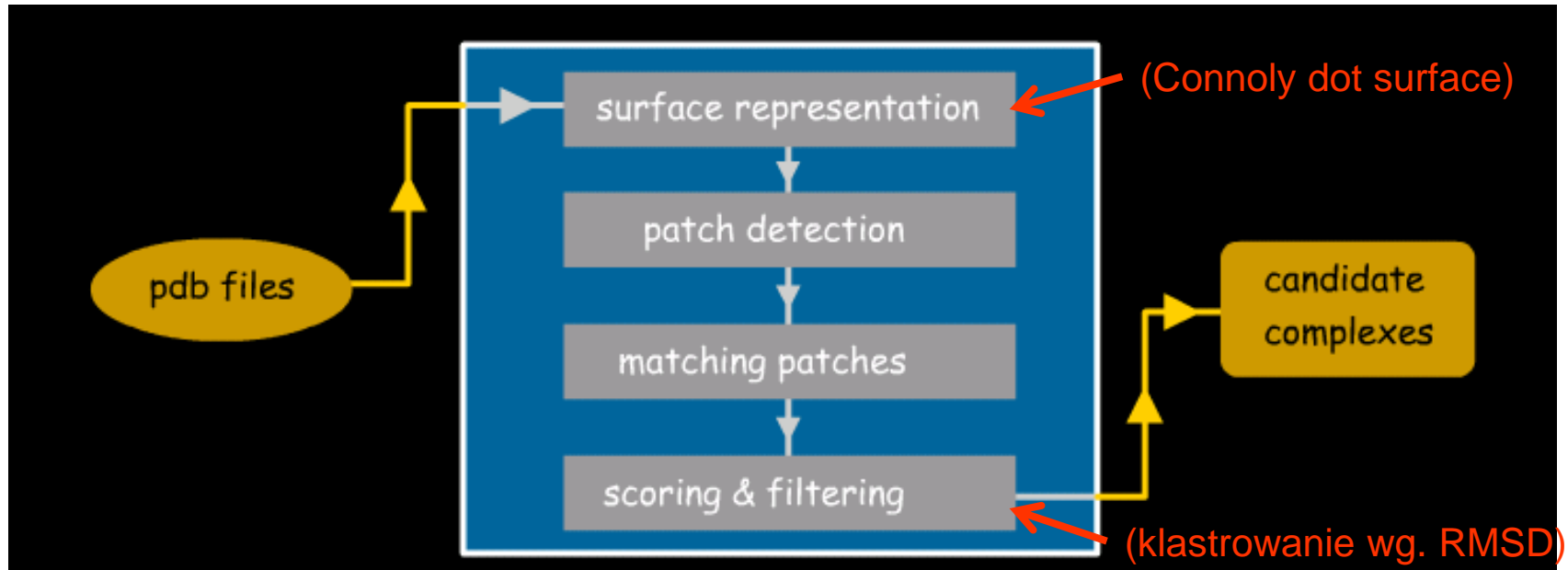
(algorytm *clique-search*)



- kieszeń receptora jest odwzorowana sferami;
- min. 4 atomy liganda muszą się nałożyć na środki sfer.

Program PatchDock

dokowanie ligand-białko i białko-białko



PatchDock – dokowanie „łatek”



Metody optymalizacyjne znajdowania najlepszego dopasowania liganda

Stochastic methods for solving
complex optimization problems

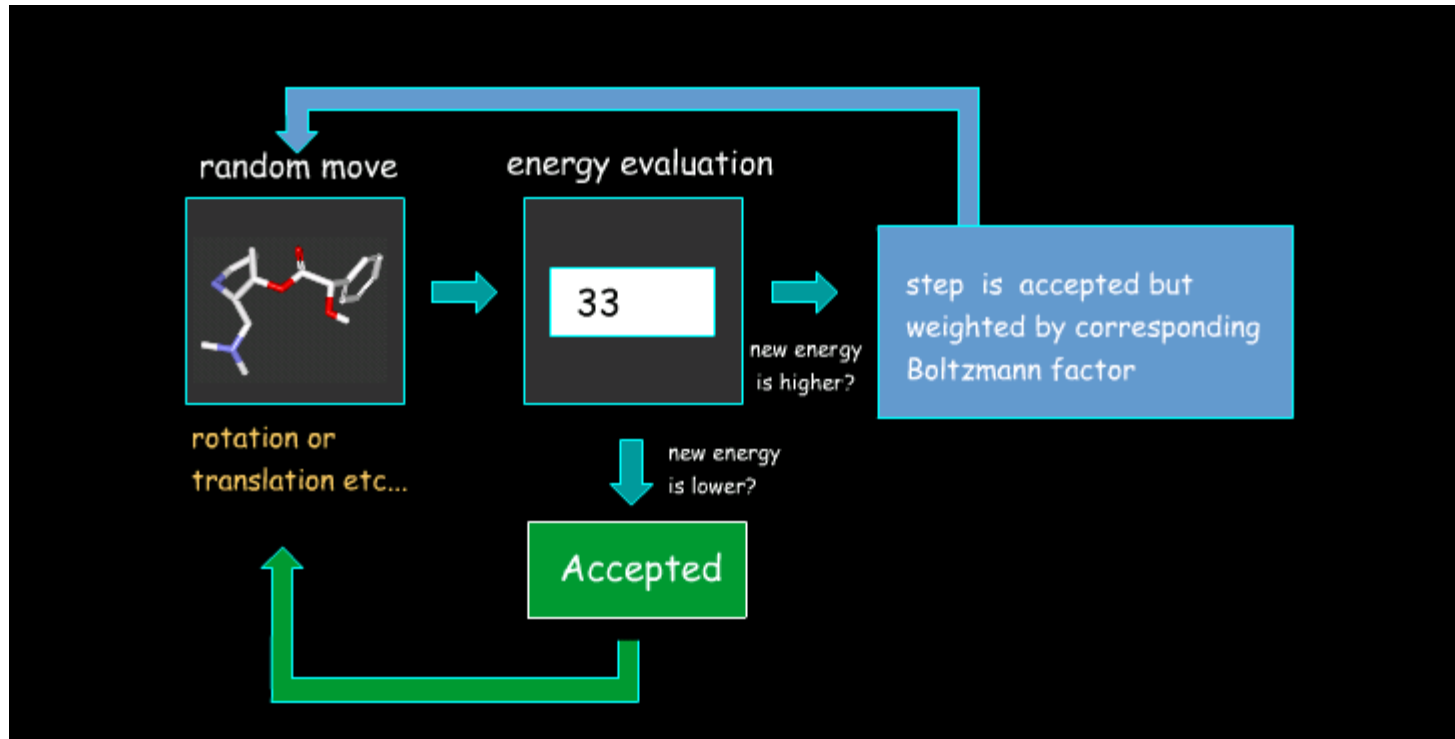
Monte Carlo (MC)

Genetic algorithms (GA)

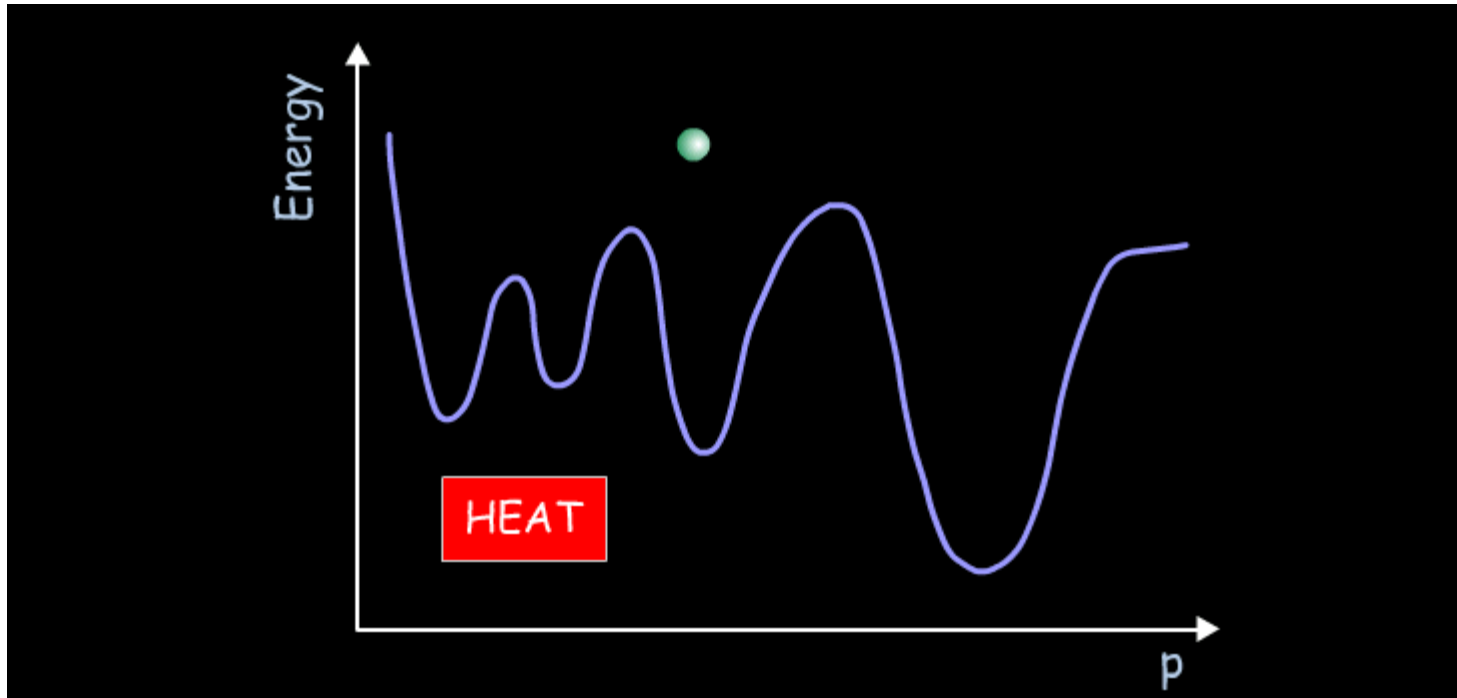
Simulated Annealing (SA)

Tabu Search (TS)

Schemat metody Monte Carlo



Schemat metody Simulated Annealing

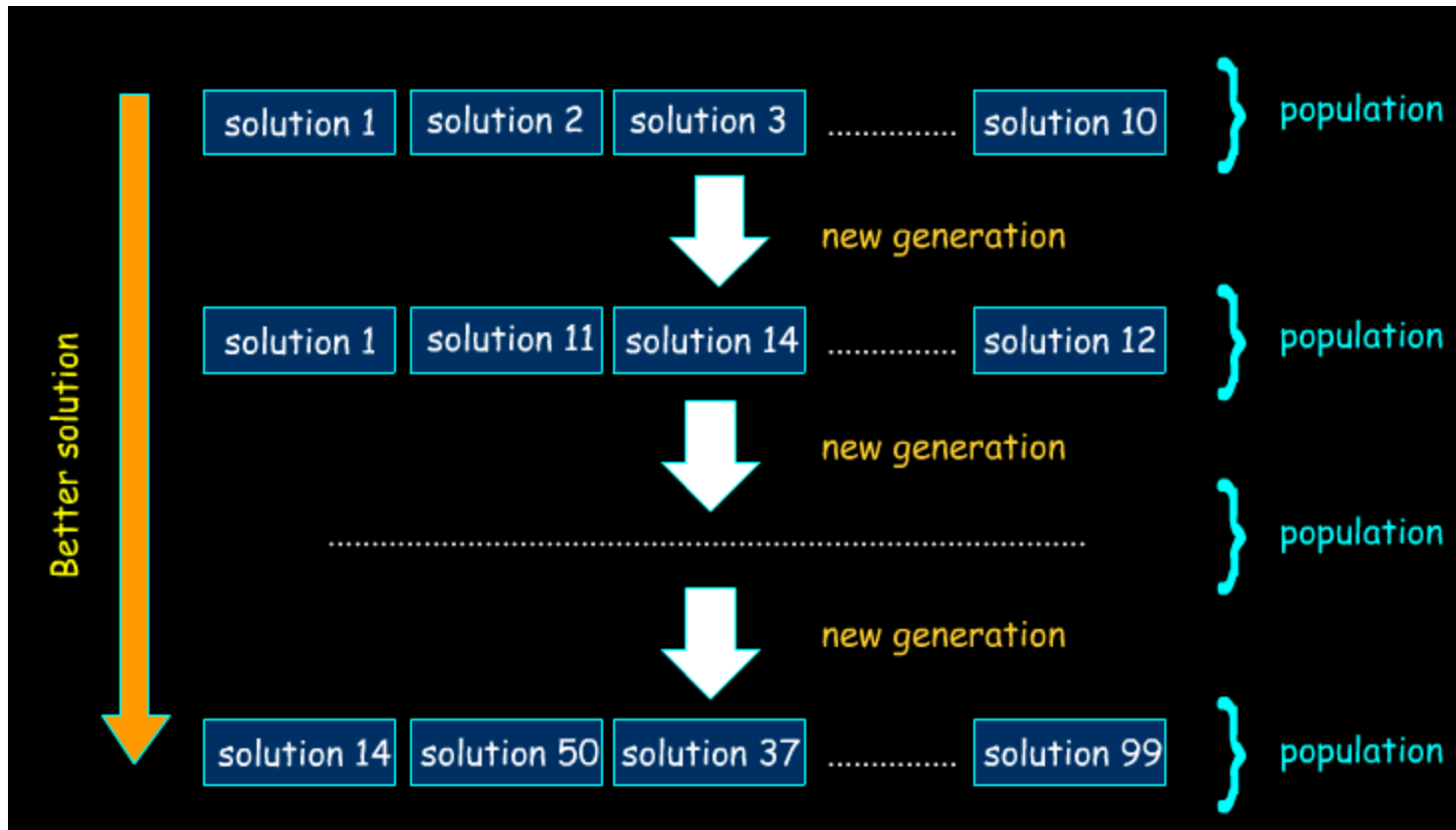


Cykl ogrzewania i powolnego oziębiania układu jest powtarzany wiele razy aż do znalezienia wielu prawie optymalnych rozwiązań.

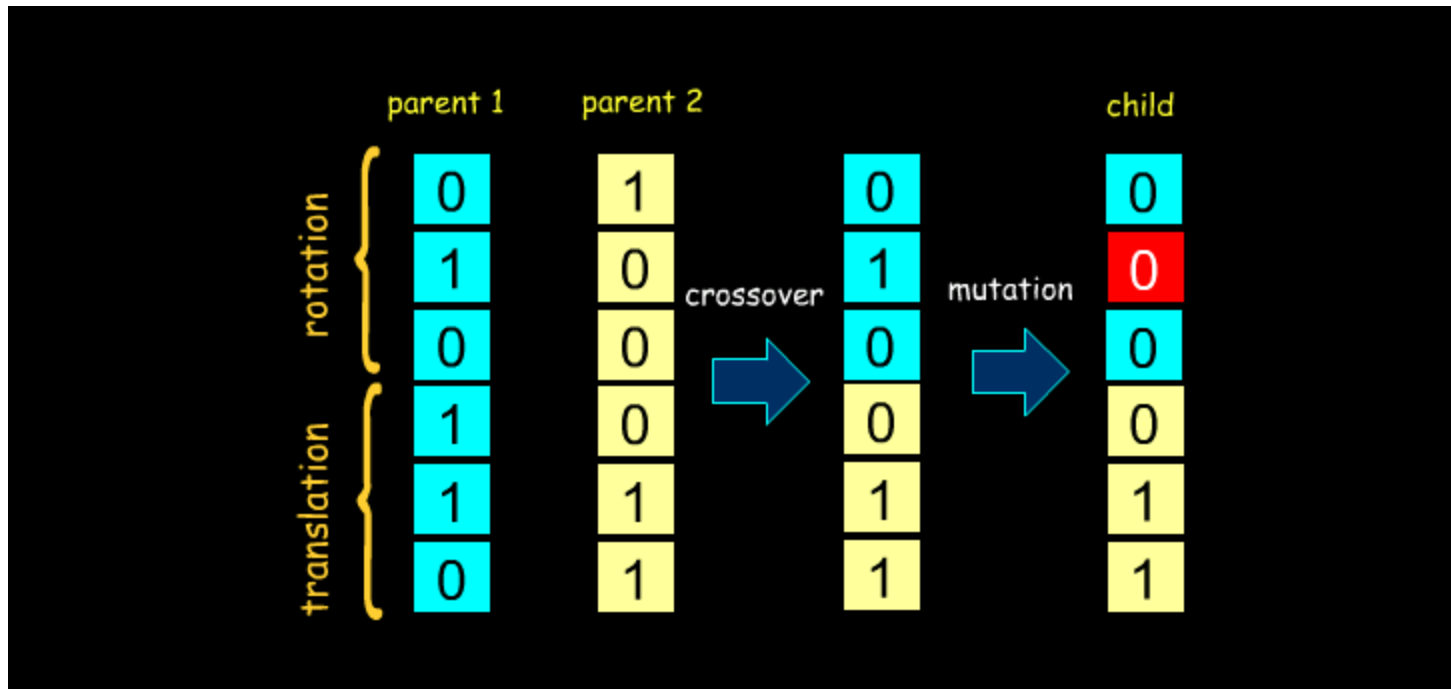
Ogrzewanie powoduje przekraczanie barier energetycznych a oziębianie pozwala na znalezienie układów z najniższą energią.

Zaimplementowana po raz pierwszy w AutoDocku.

Zasada działania Algorytmów Genetycznych



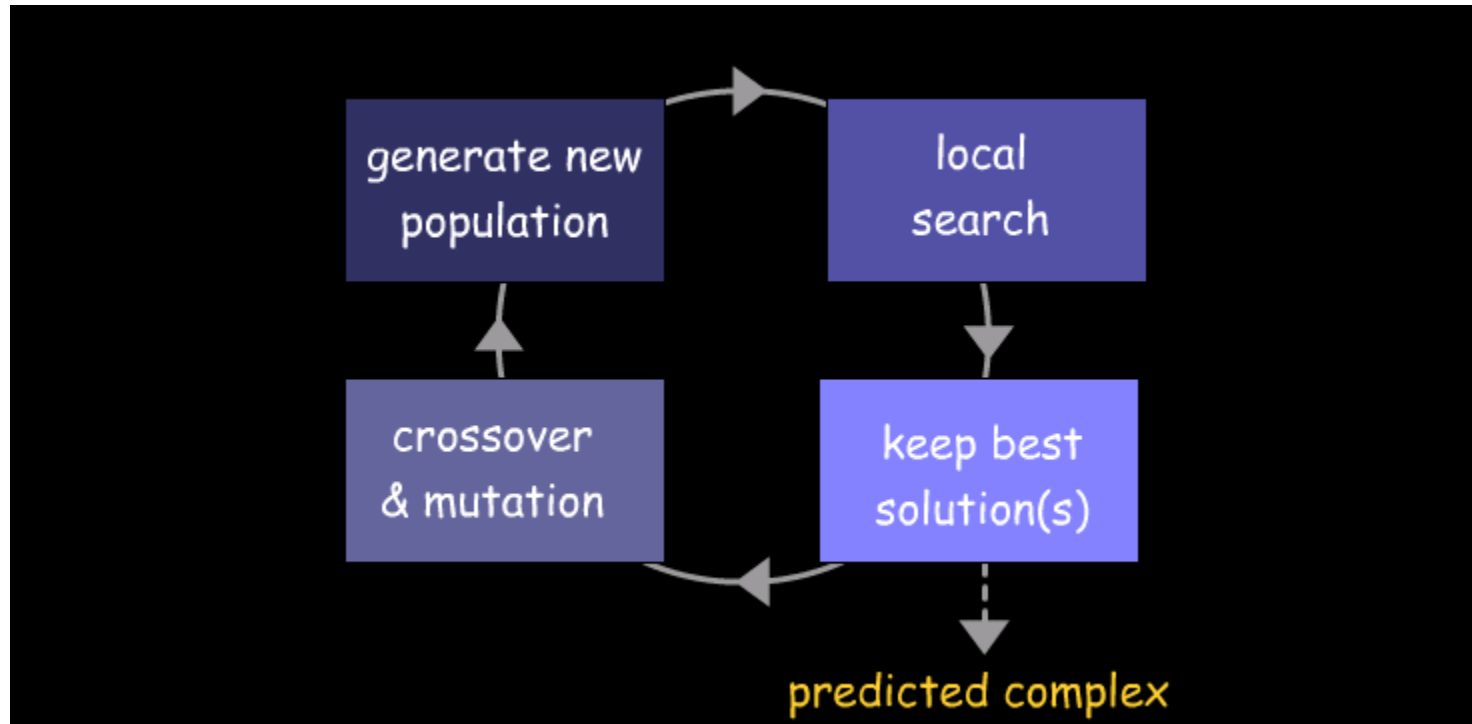
Krzyżowanie i mutacje w GA przy dokowaniu



Dokowanie sztywne (tylko rotacje i translacje)

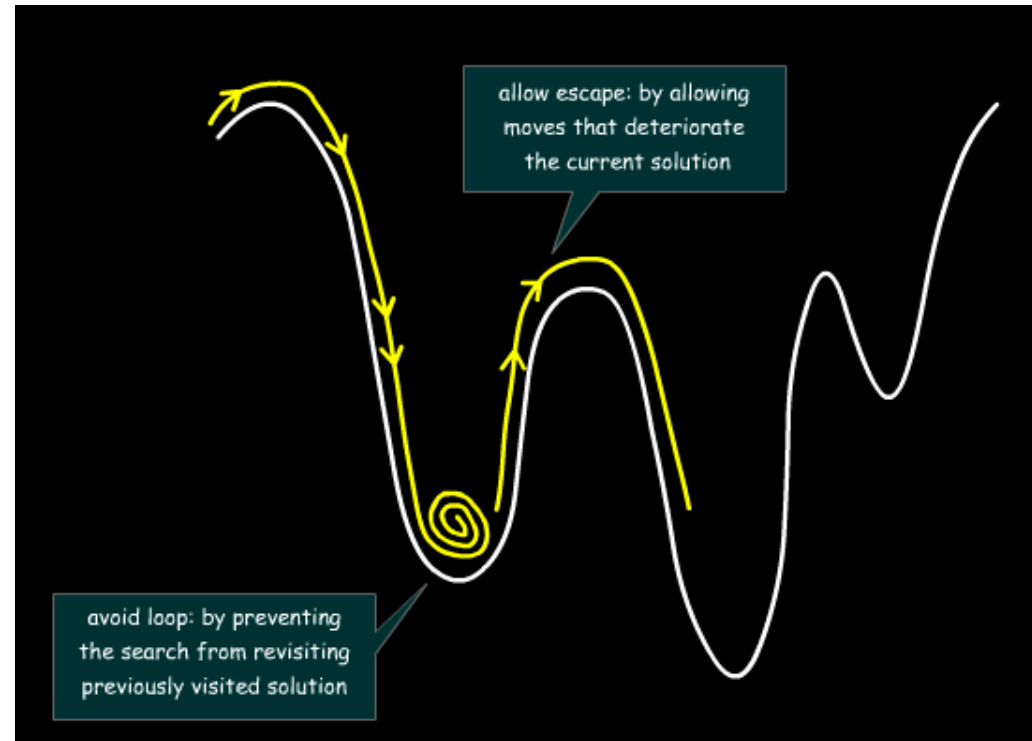
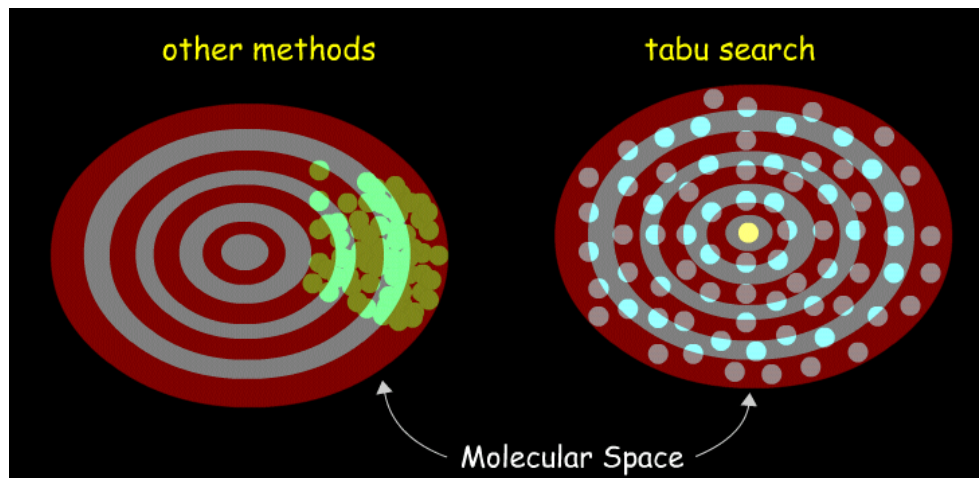
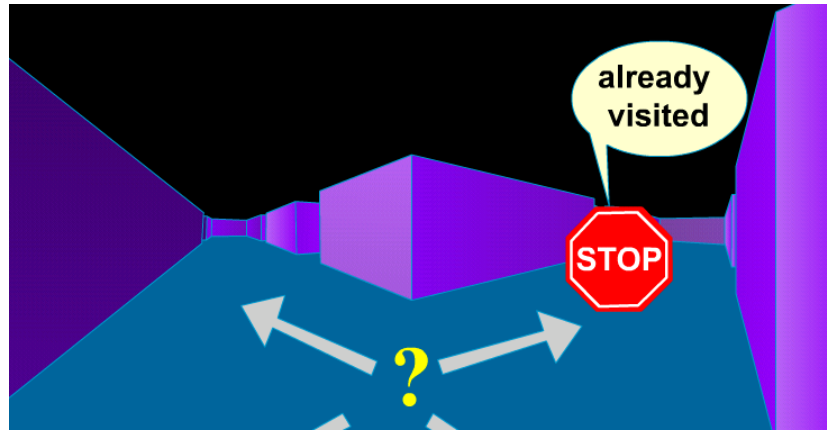
Algorytmy Genetyczne Lamarcka (LGA)

z lokalną optymalizacją



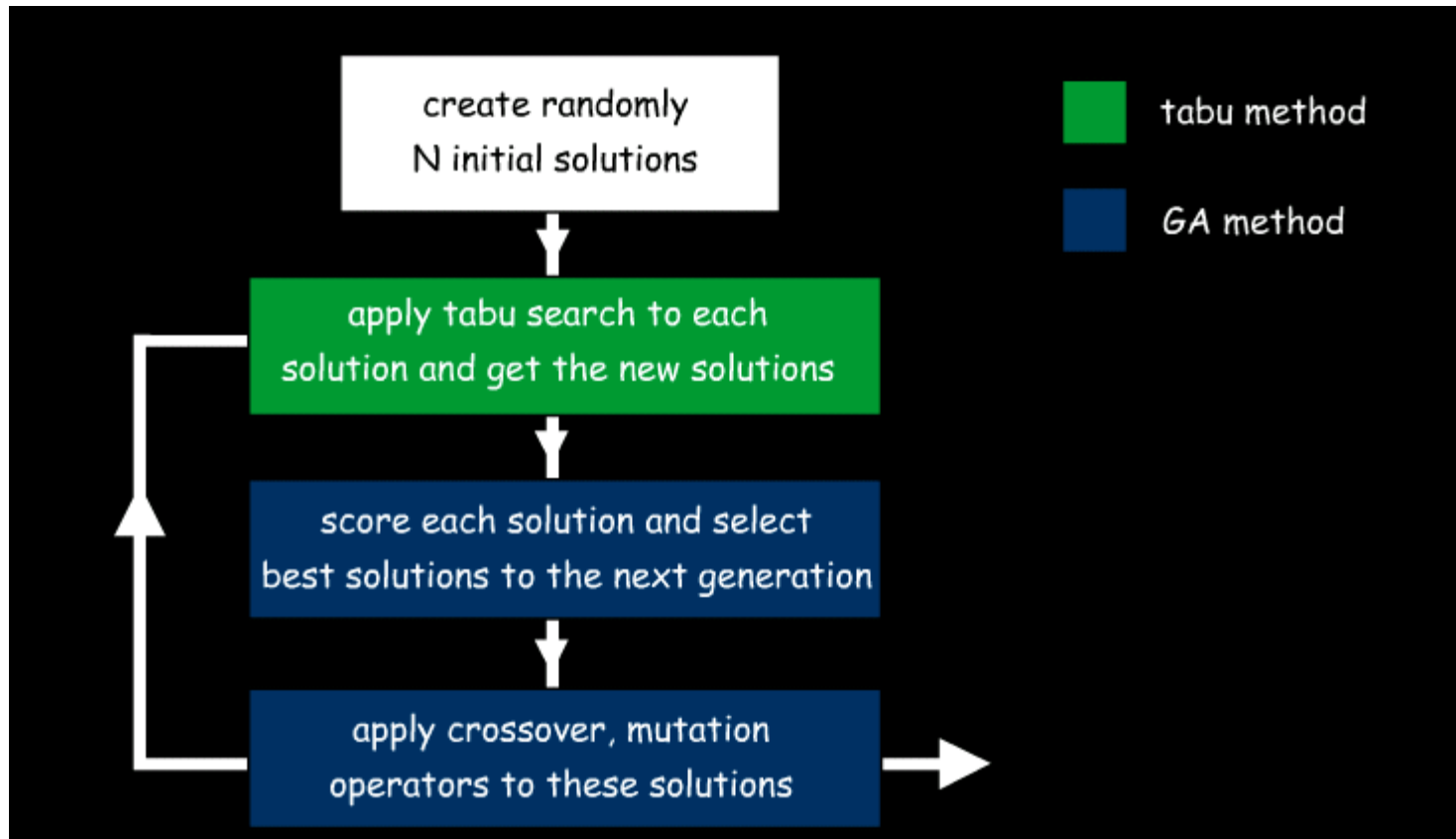
Procedura stosowana obecnie w AutoDocku

Tabu search (TS)

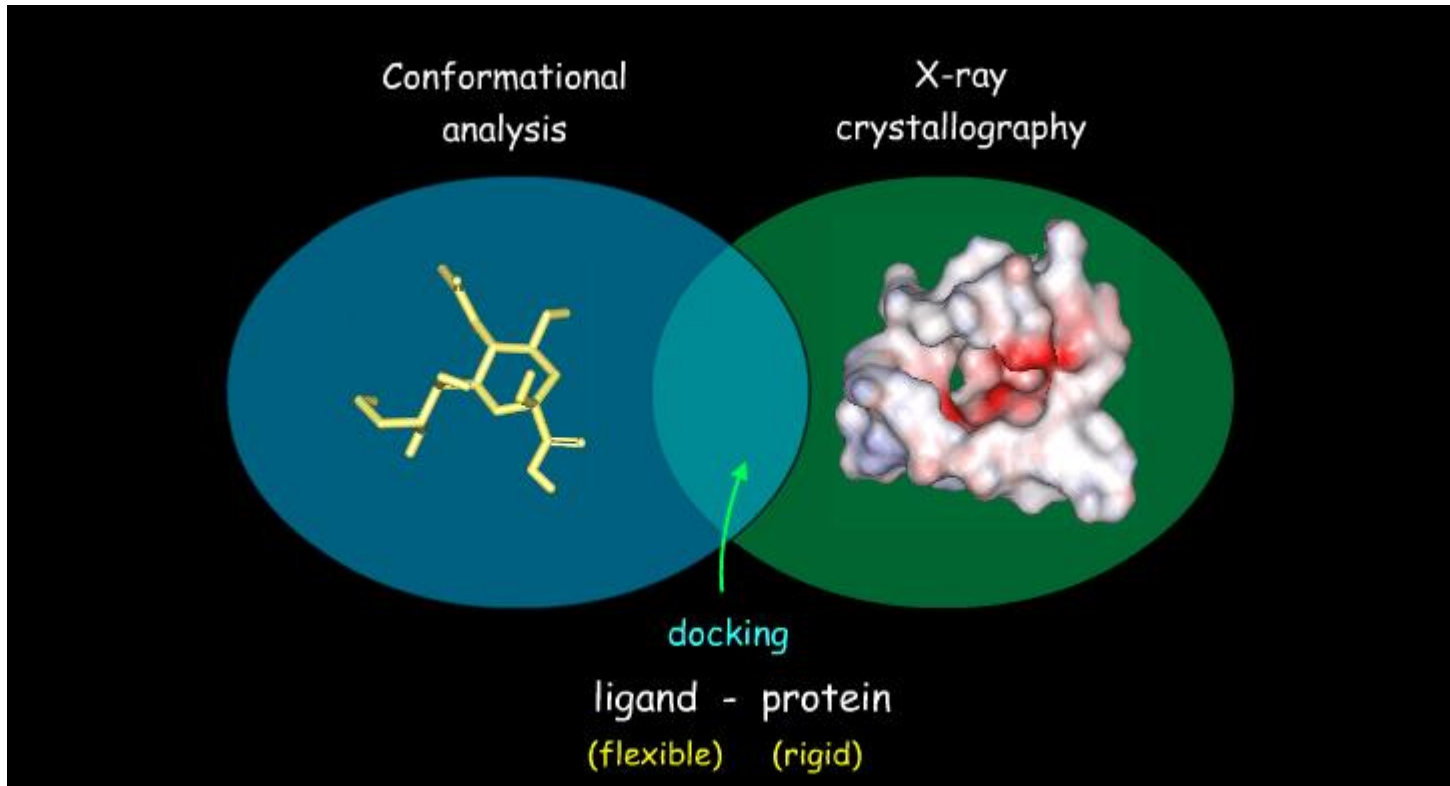


Pozwala na wyjście z lokalnych minimów kosztem przejściowego zaburzenia układu (stopniowe zasypywanie lokalnych minimów energii)

Metoda hybrydowa dokowania (TS+GA)



Dokowanie giętkich ligandów



Rozwój metod giętkiego dokowania zapoczątkował nową dziedzinę projektowania leków: structure-based drug design

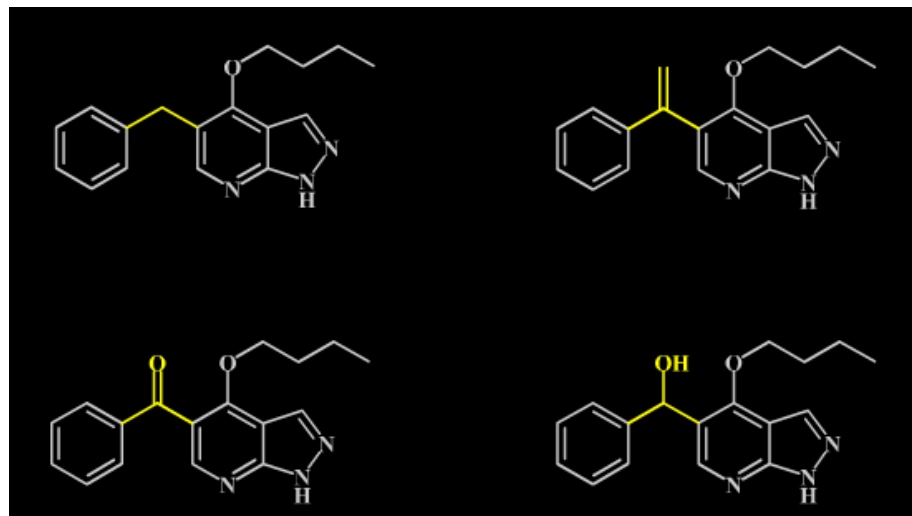
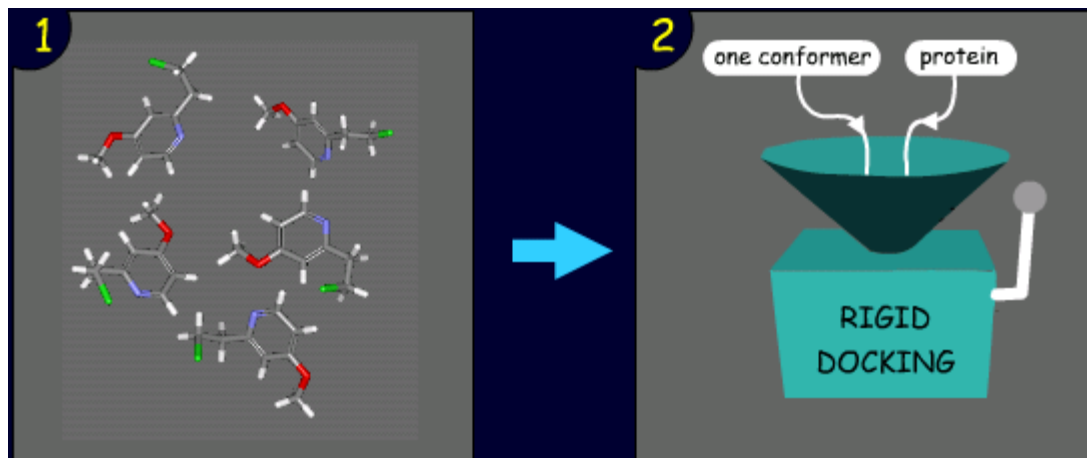


Metody do giętkiego dokowania ligandów

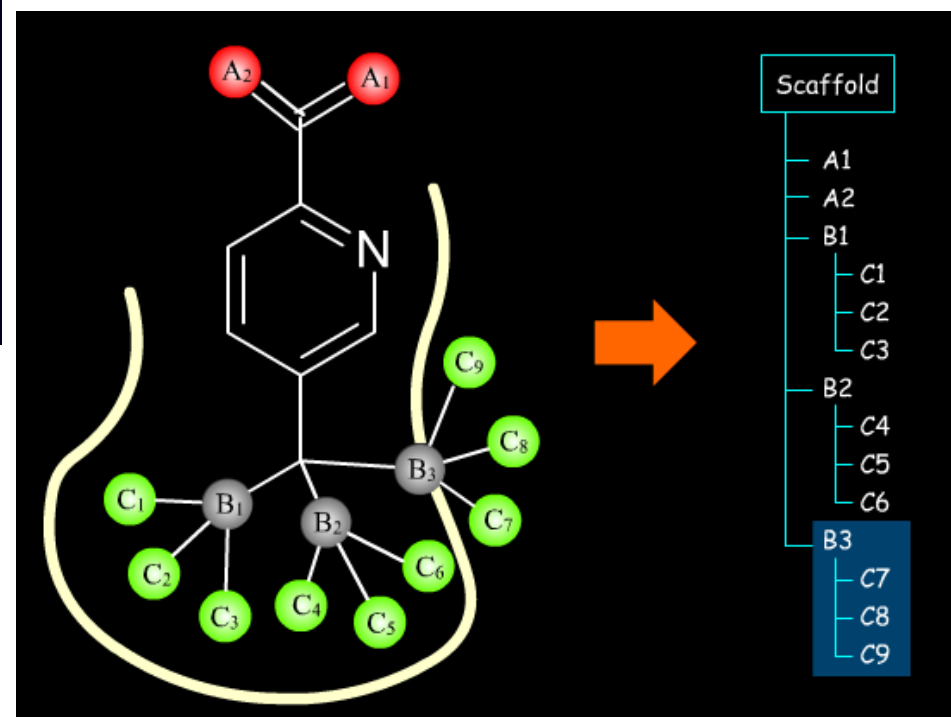
- the ensemble docking method
- the fragmentation docking method
- the stochastic search approach

Dokowanie zespołowe/całościowe

(ensemble docking)

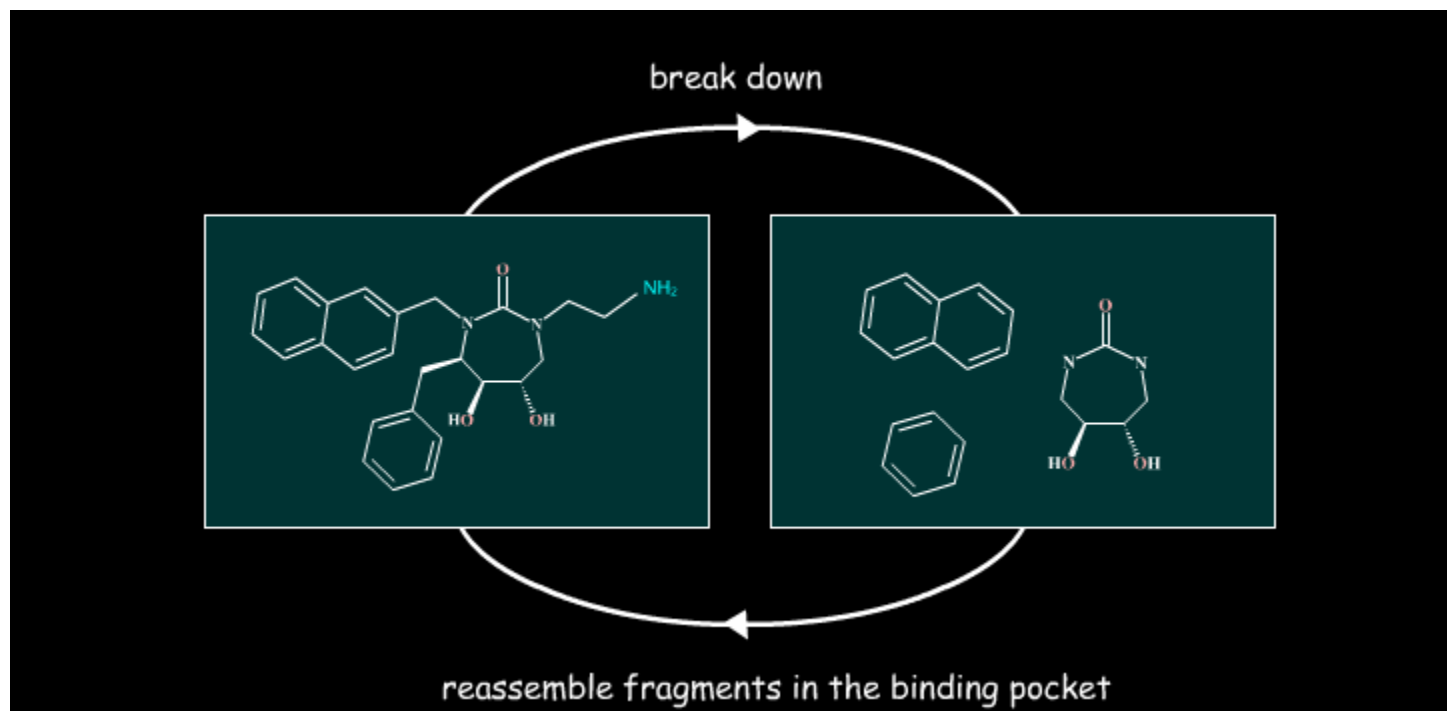


Łatwa implementacja do dokowania podobnych związków z dużych baz danych



Efektywne unikanie błędnych dokowań

Dokowanie fragmentowe (fragmentation docking)

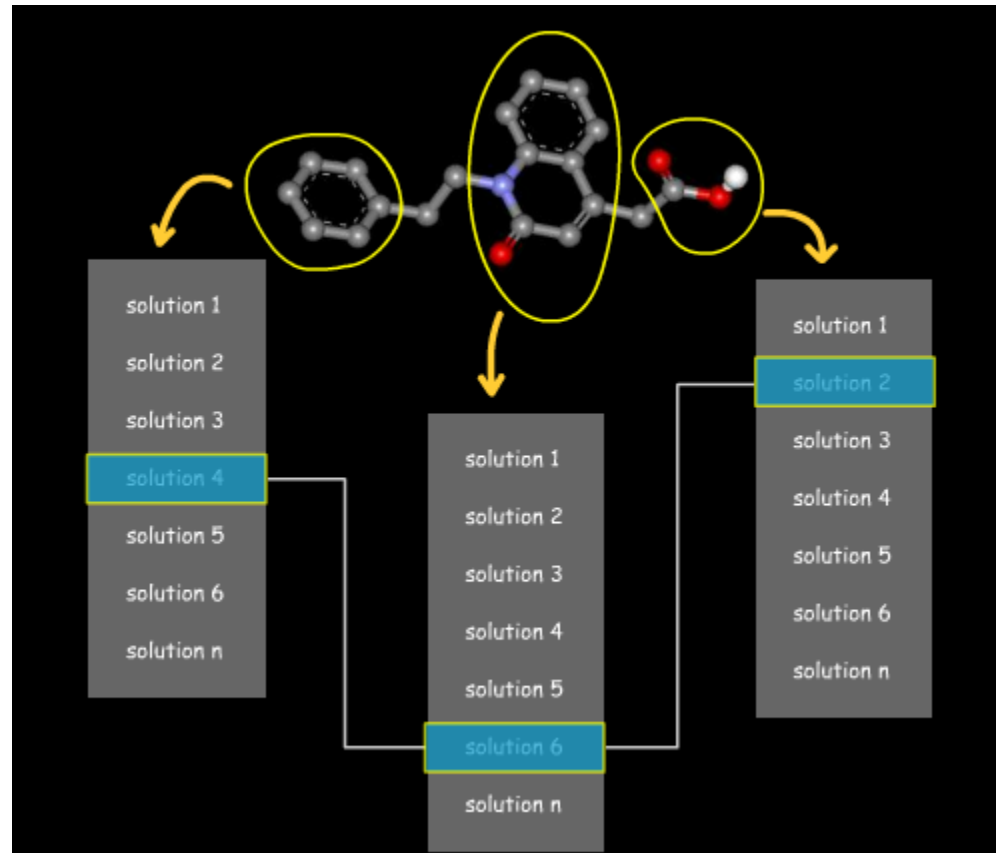


Metody dokowania fragmentowego:

- Place-and-join (fragmenty dokowane niezależne)
- Incremental approach (fragmenty dokowane zależne od siebie)

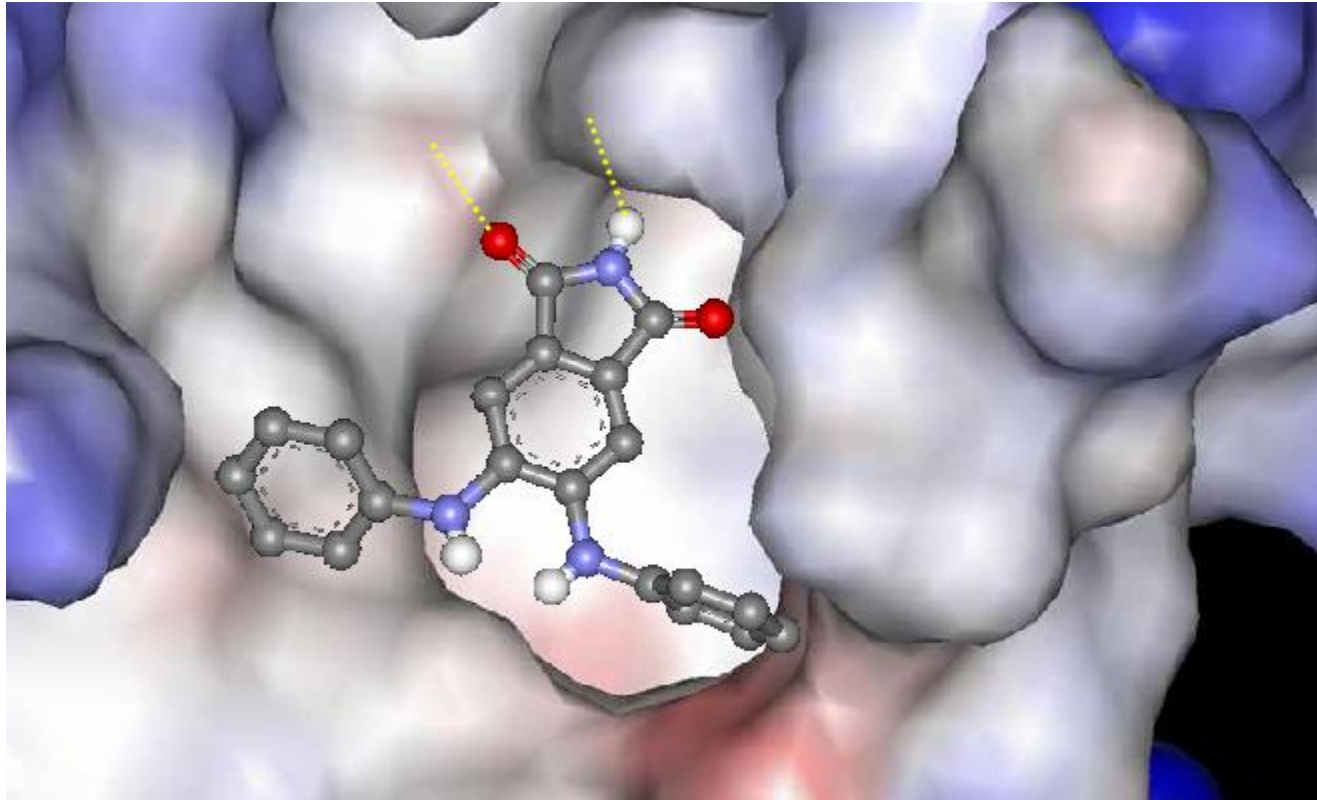
Metoda ułóż i połącz

(place-and-join)



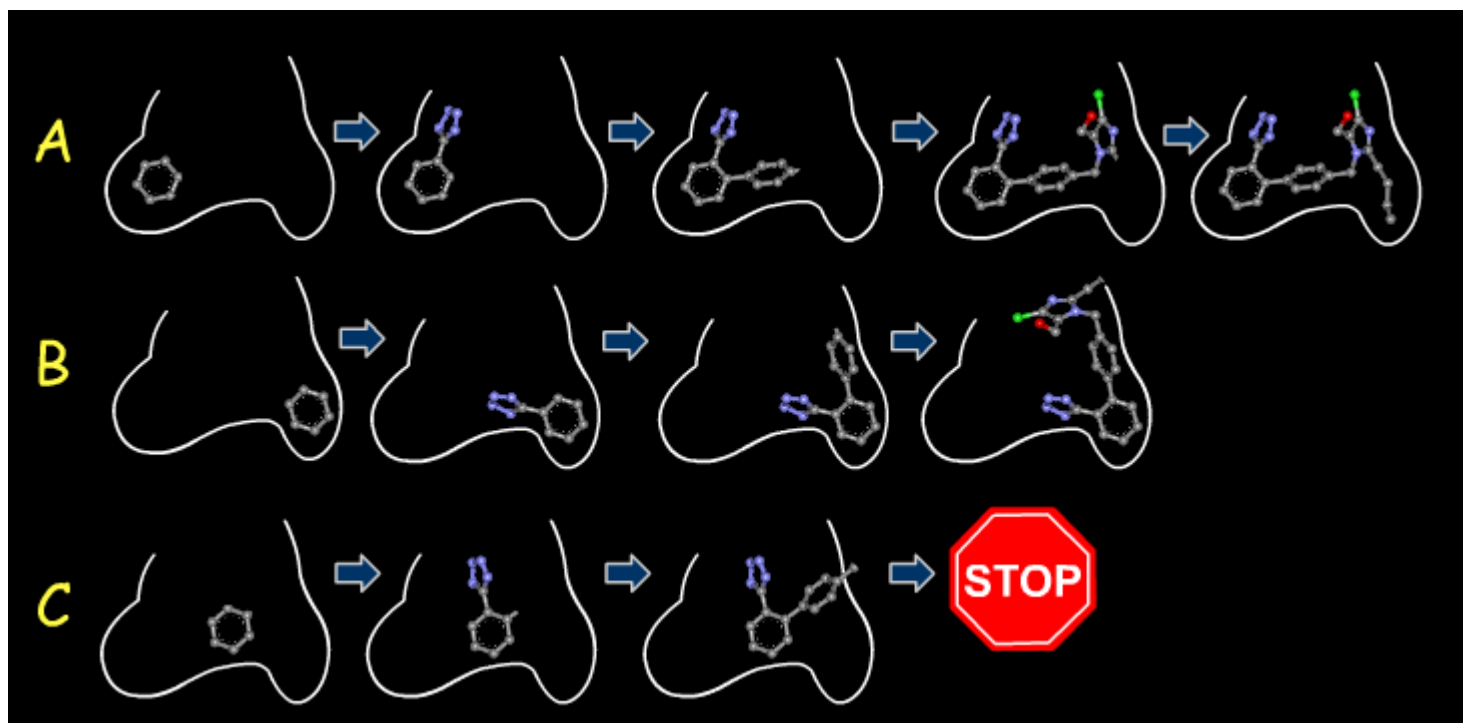
Wybór nieoptymalnych dokowań cząstkowych aby połączyć fragmenty (wskazówki do modyfikacji struktury liganda)

Metoda stopniowego dokowania fragmentów (incremental approach)



Najpierw dokuje się rdzeń liganda a następnie po kolei pozostałe fragmenty minimalizując energię powstającego liganda

Wybór położenia pierwszego fragmentu ma wpływ na wynik dokowania



Wybór najlepszego dopasowania na podstawie funkcji
oceniającej poszczególne pozy



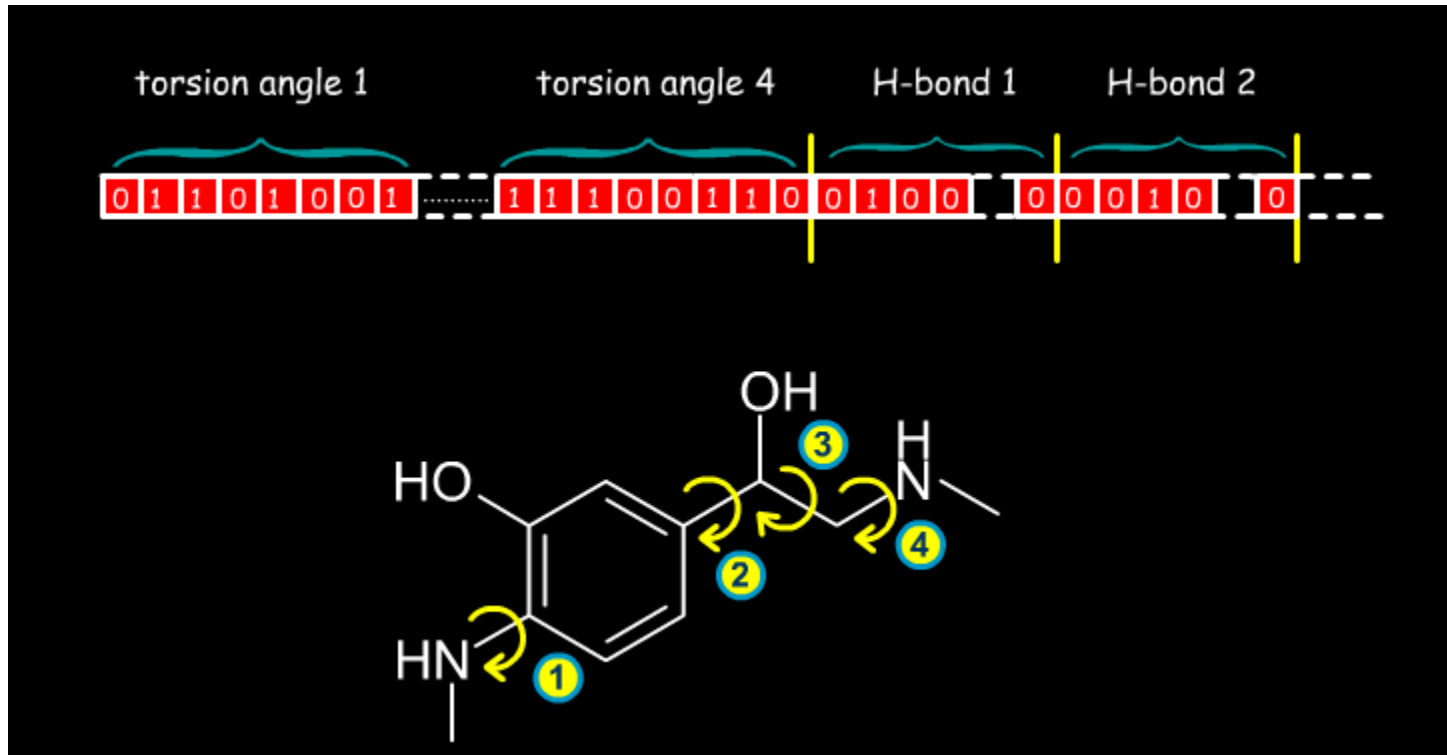
Przykłady zastosowań **dokowania giętkiego** w programach

- **GOLD** – algorytmy genetyczne (GA)
- **MOE-Dock** – Monte Carlo (MC), Tabu search
- **PRO_LEADS** – algorytmy genetyczne
- **AutoDock** – algorytmy genetyczne, Tabu search

Kąty torsyjne giętkiego liganda wprowadza się do chromosomów w GA lub do parametrów do optymalizacji w MC

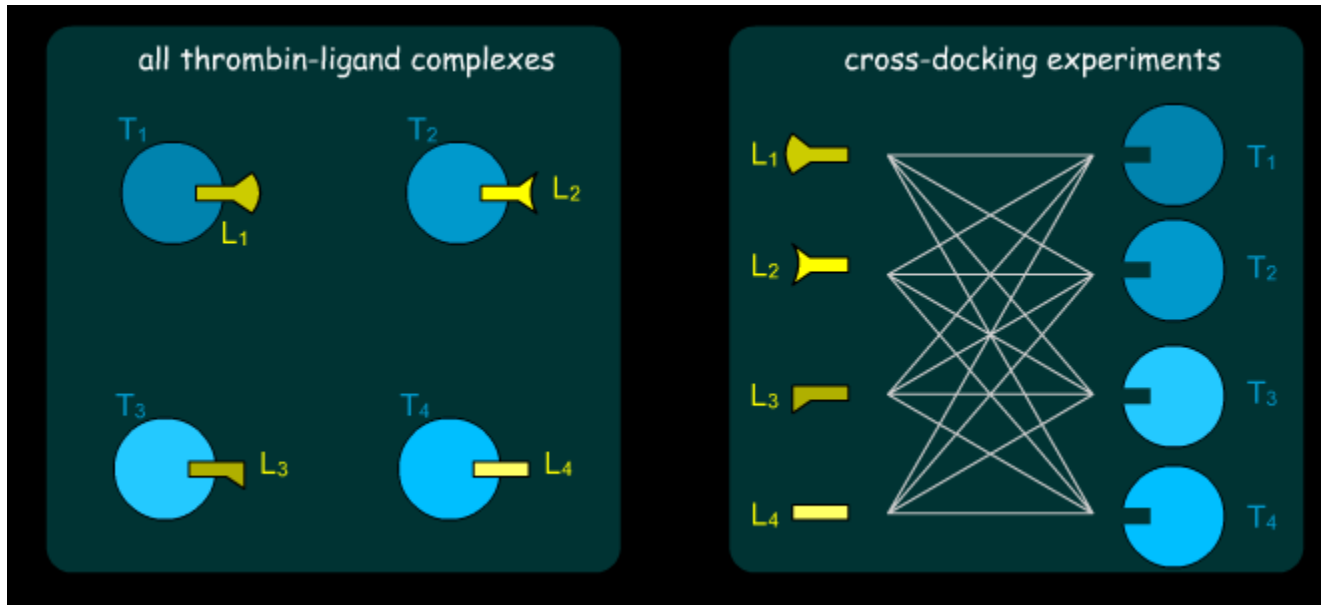
GOLD – skład chromosomów w GA

część chromosomu dotycząca giętkiego dokowania



Dodatkowe śledzenie każdego potencjalnego wiązania wodorowego pomiędzy ligandem i białkiem w celu maksymalizacji ich liczby

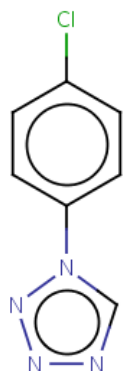
Ważność giętkiego dokowania



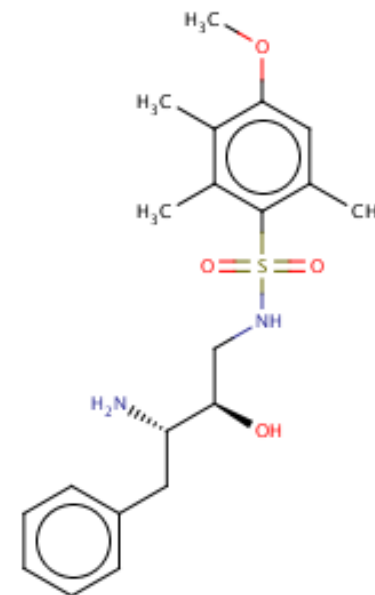
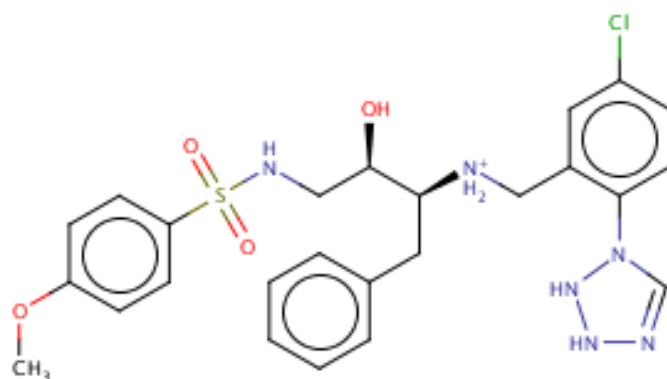
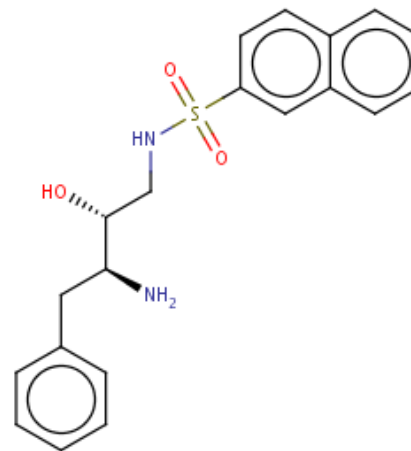
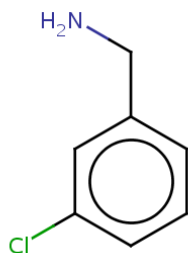
W 50% przypadków szttywne kros-dokowanie prowadziło do błędnych wyników

Inhibitory trombiny

(proteaza serynowa – tnie fibrynogen – krzepliwość krwi)




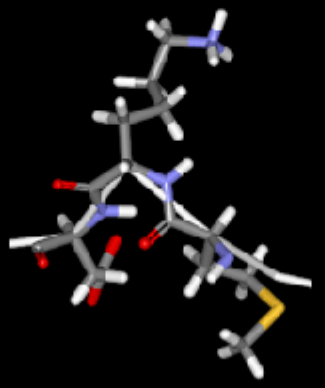

poходna tetrazolowa



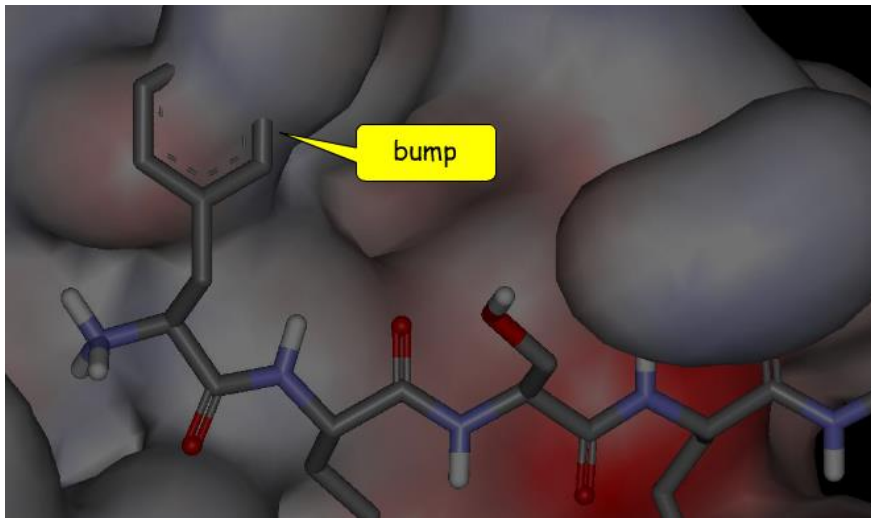
→ porównanie struktur trombiny + CXCR4



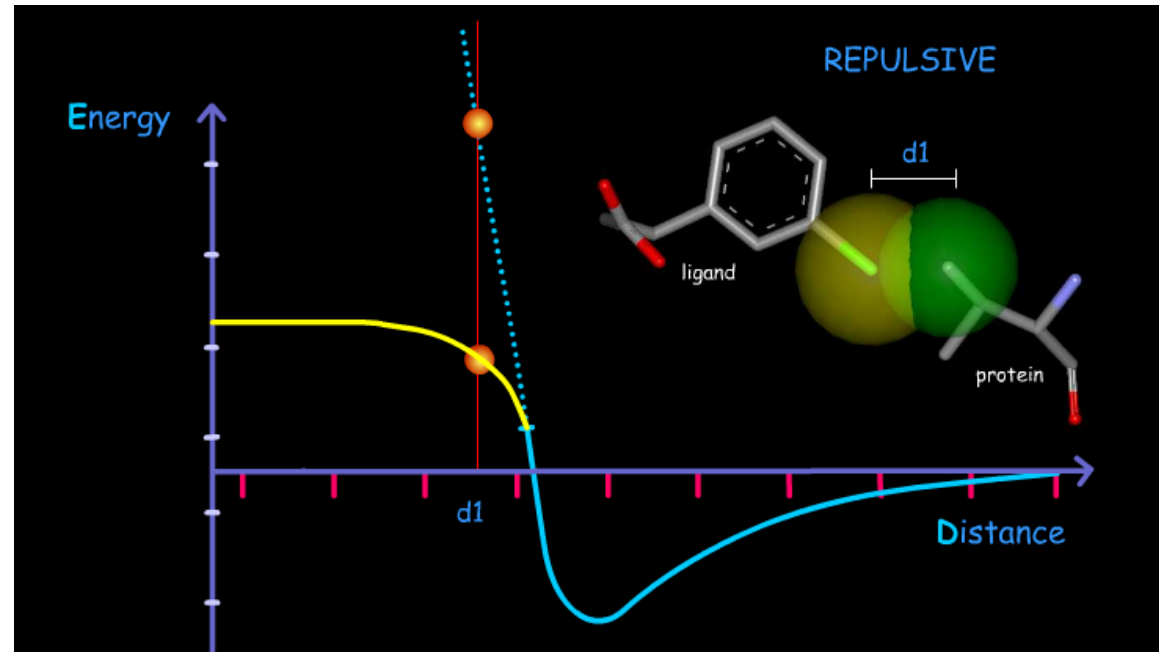
Uwzględnianie giętkości białka

pseudo flexibility	side chain flexibility	full flexibility
		
soft docking	partial flexibility	backbone & domain

Miękkie dokowanie ligandów

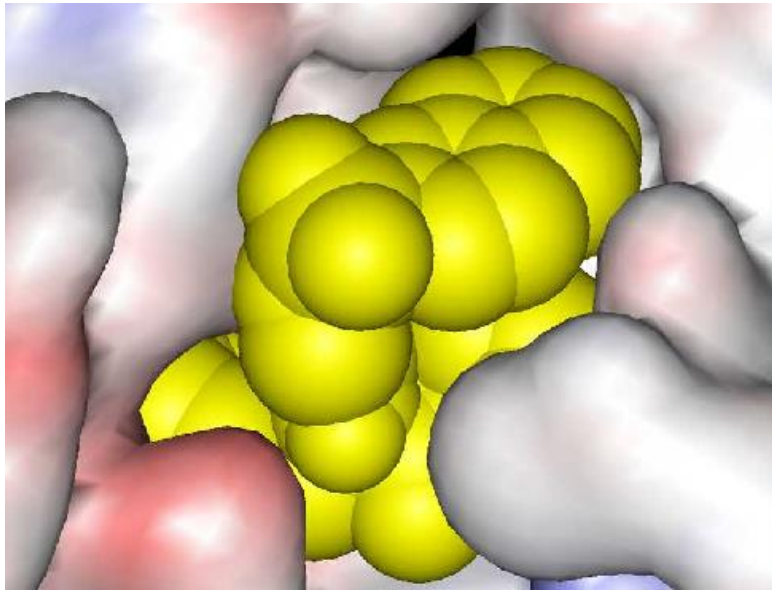


Pozwala na niewielkie nakładanie się struktur liganda i receptora

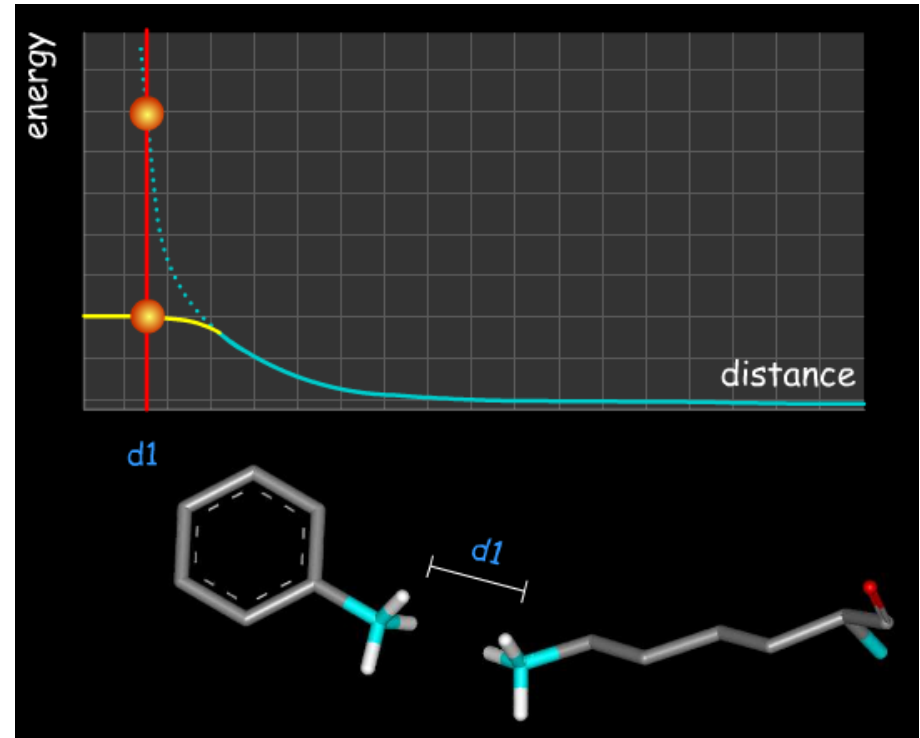


Modyfikacja potencjału Lennarda-Jonesa (LJ) aby siły van der Waalsa były mniej odpychające (plastyczność receptora bez zmiany jego konformacji)

Miękkie dokowanie ligandów



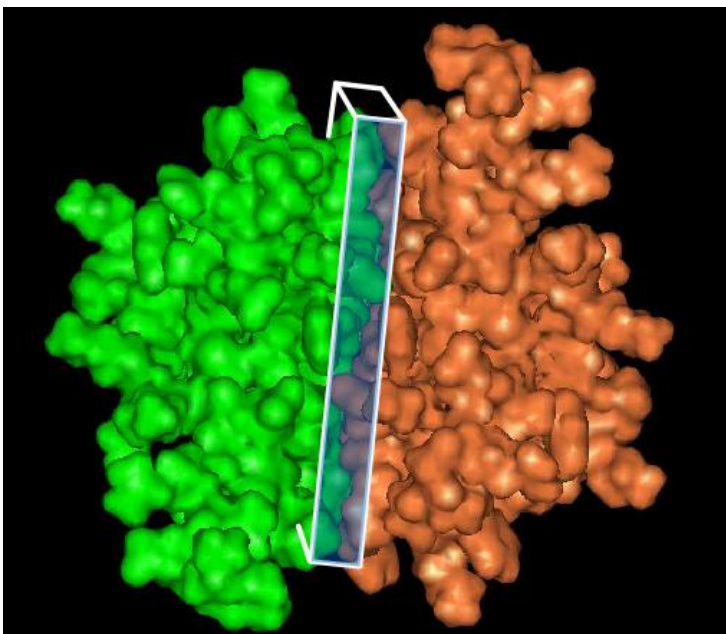
Zmniejszanie promieni van der Waalsa (rozmiarów atomów) dla liganda i receptora



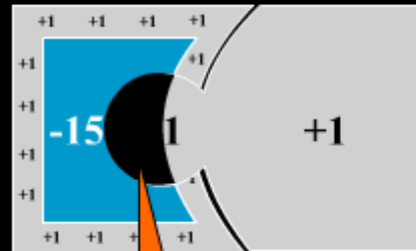
Modyfikacja potencjału elektrostatycznego

„Miękkie” funkcje oceniające

(soft scoring)

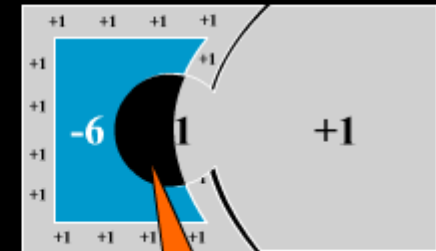


mapped-grid method



the penalization for
this overlap is
 $(-15) \times (+1) = -15$

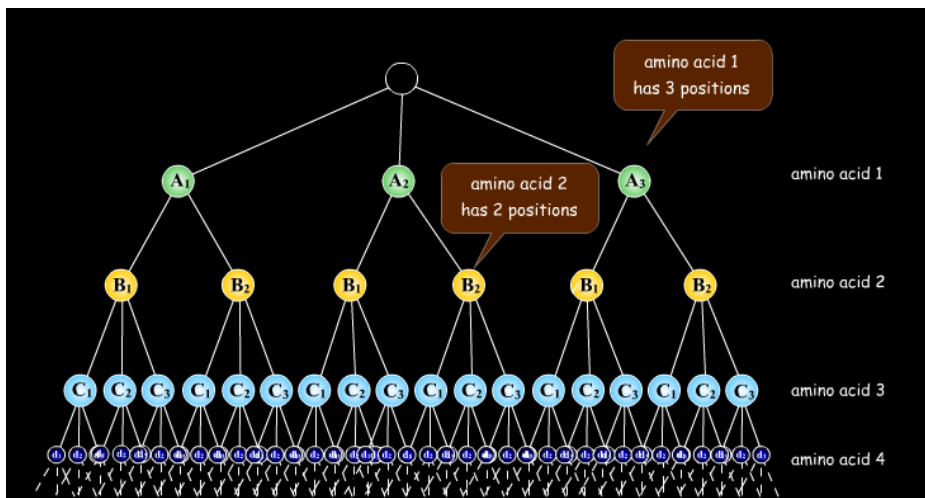
mapped-grid with soft-docking



the penalization for
this overlap is only
 $(-6) \times (+1) = -6$

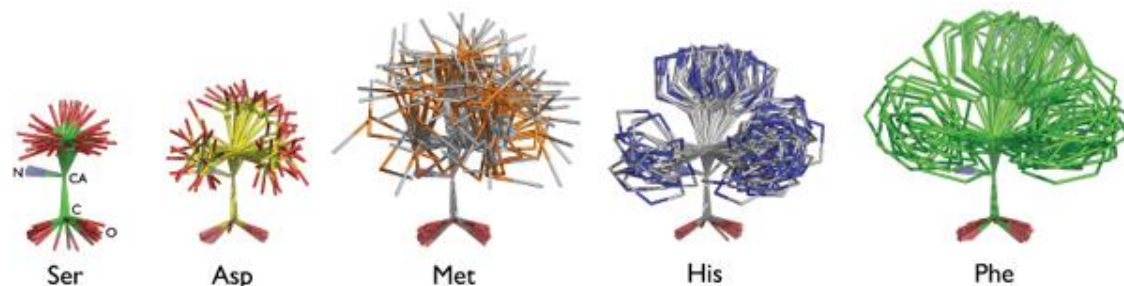
Soft scoring jest stosowana jako wstępna funkcja oceniająca do szybkiego odsiania niepasujących ligandów

Uwzględnianie ruchów łańcuchów bocznych w białku

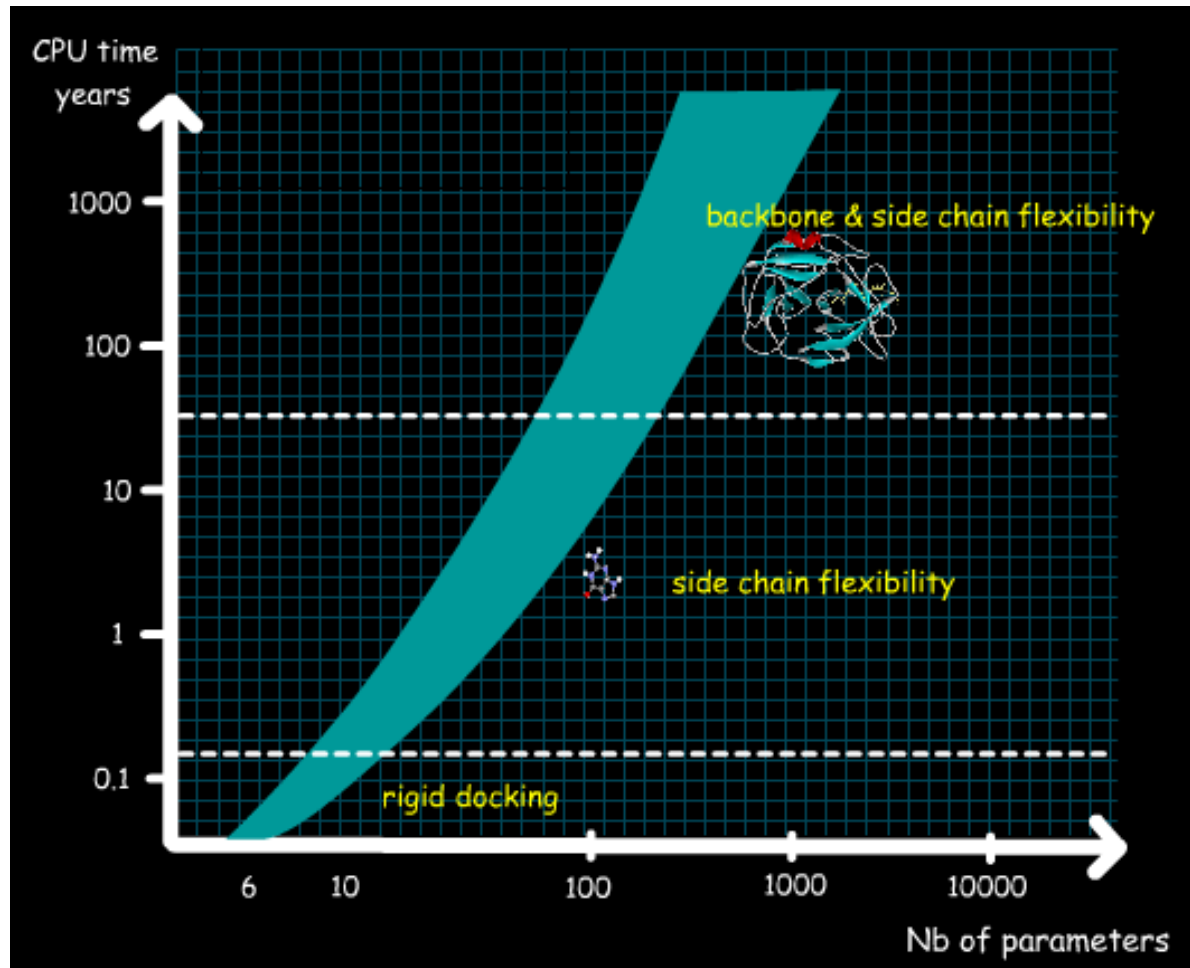


eksplozja konformacyjna

Rozwiązanie:
biblioteka konformerów aminokwasów
(1987 Ponder & Richards)

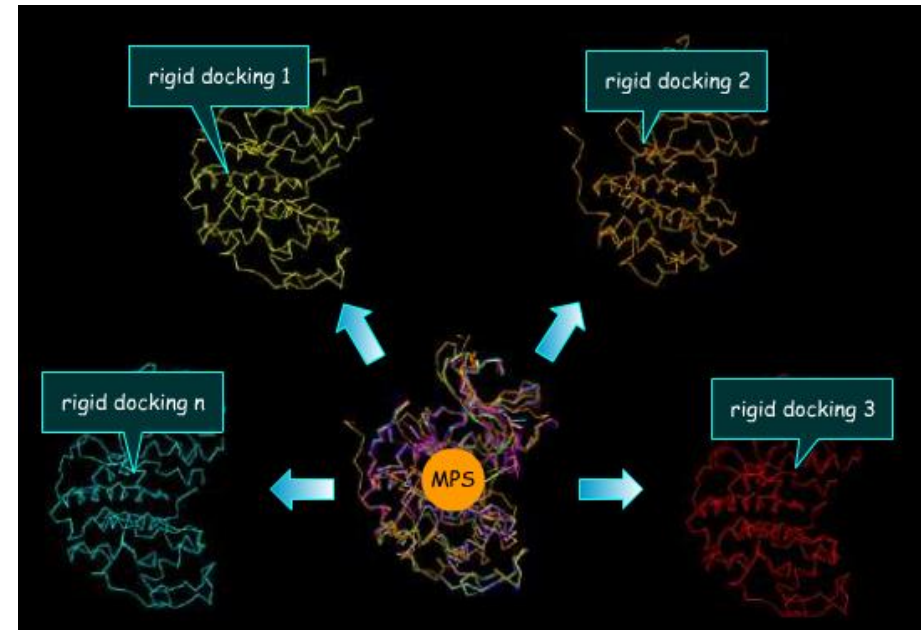
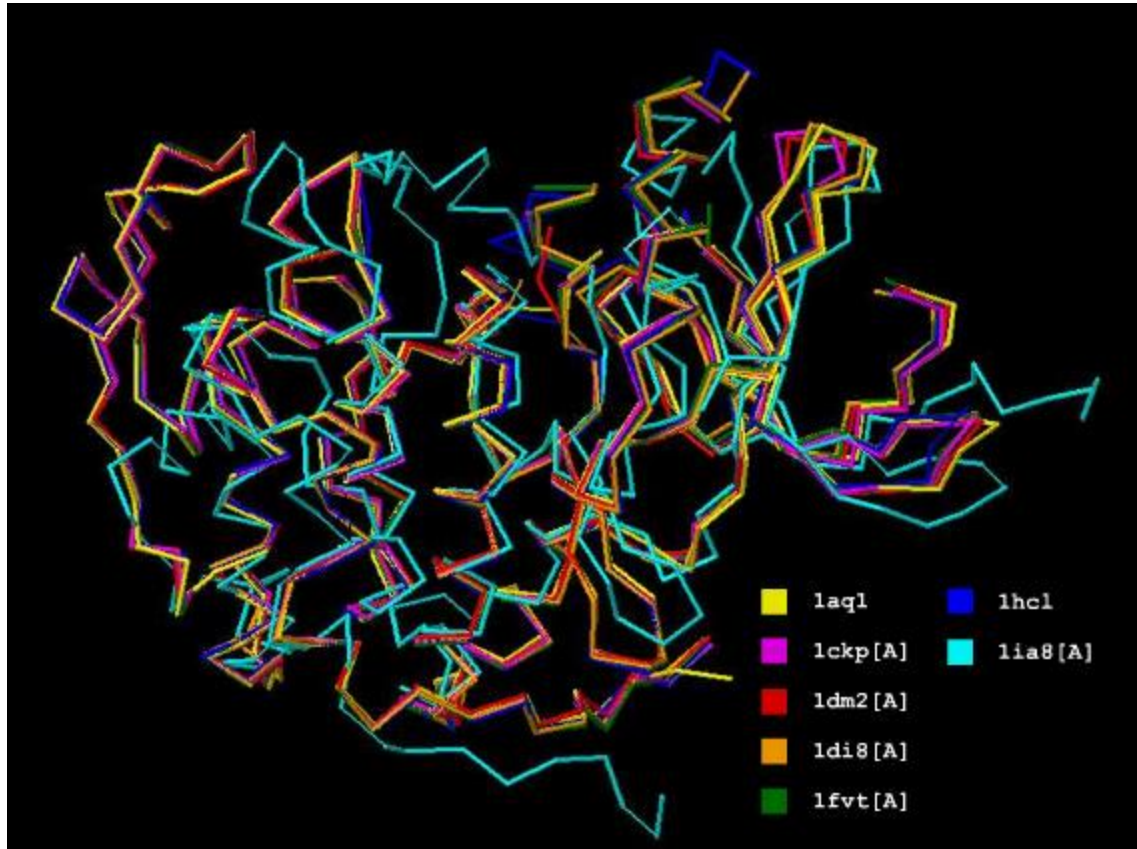


Złożoność obliczeniowa przy giętkim dokowaniu



Oszacowanie na jeden CPU

Metoda *Multiple protein structure (MPS)*

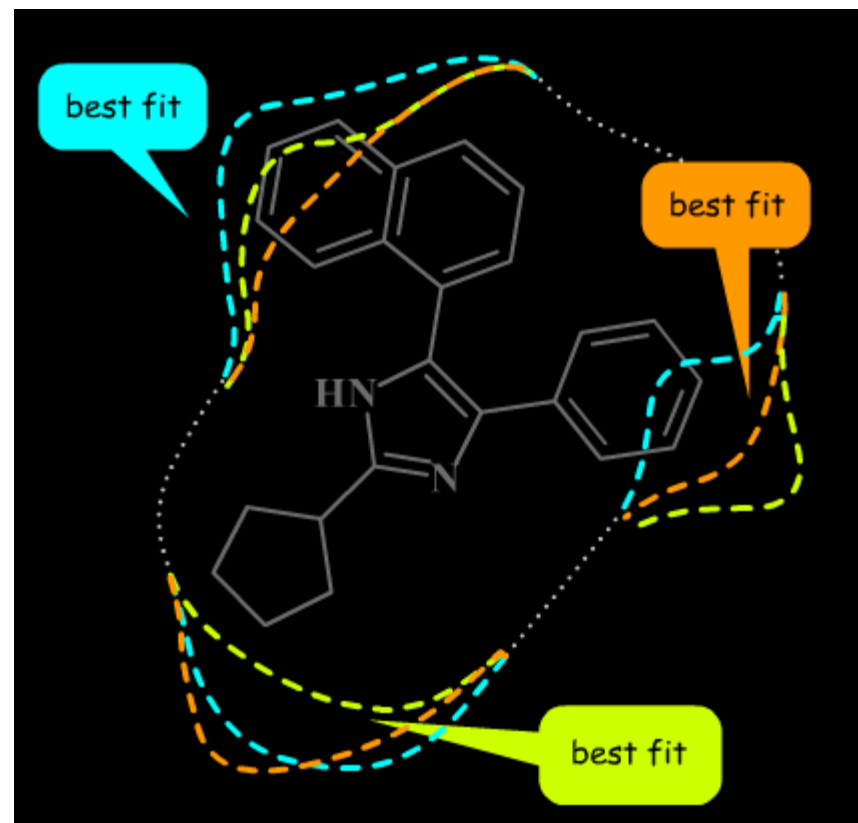
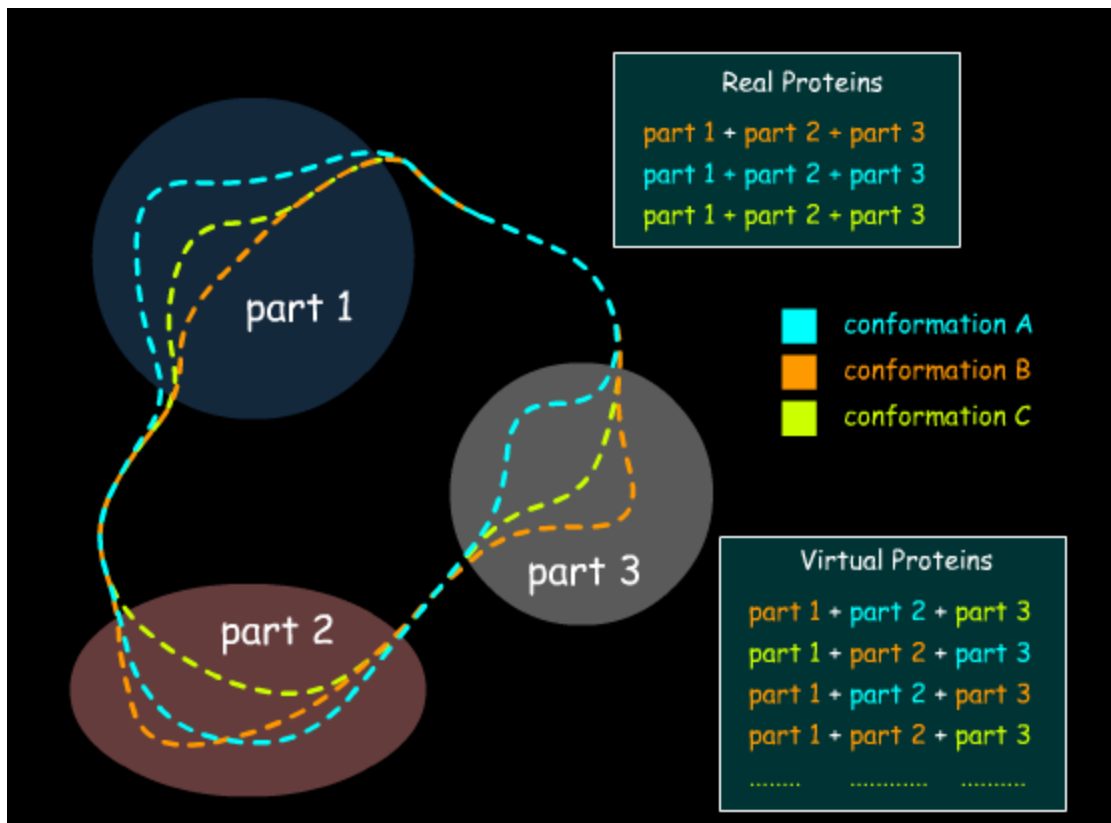


Białko CDK2 (Cyclin-Dependent Kinase 2)



Metoda łączonych białek

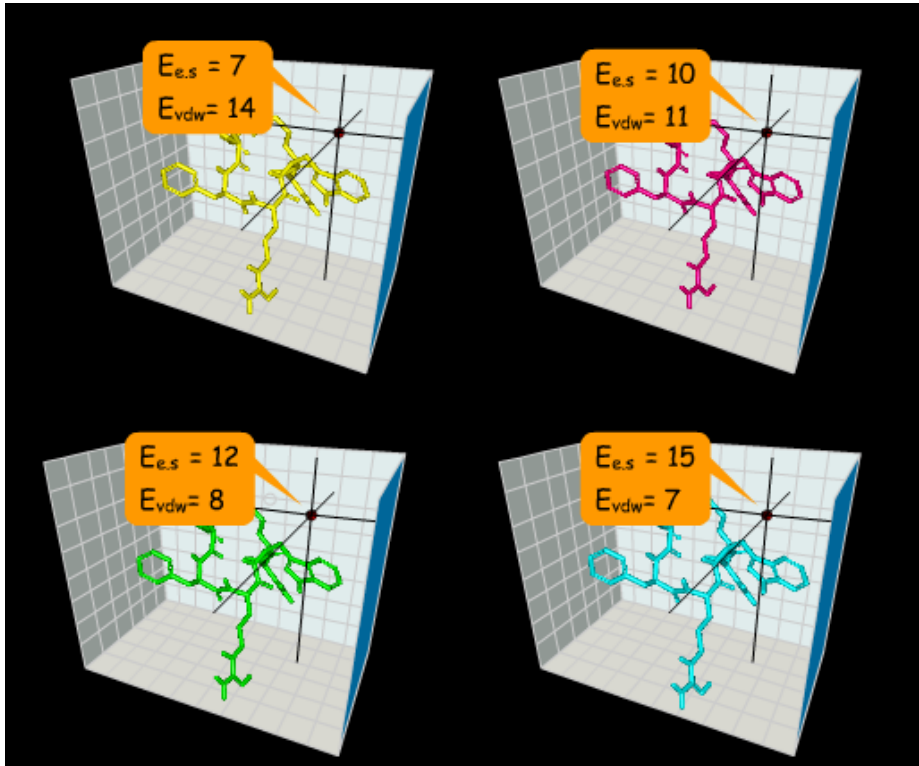
united protein approach



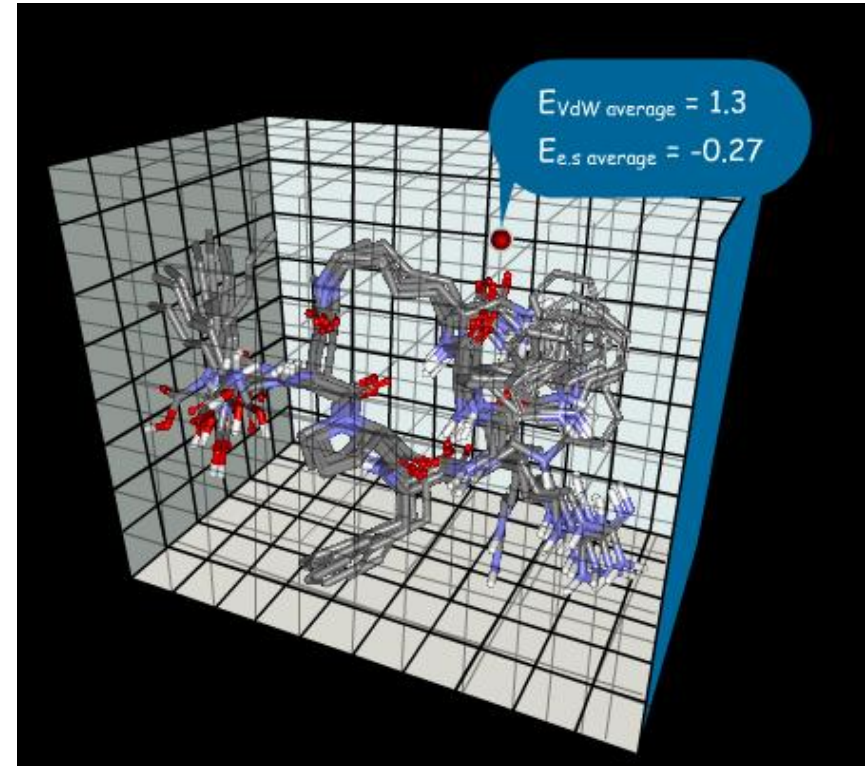
Program FlexE. Nałożone struktury białka pozwalają na zbudowaniu wielu wirtualnych konformacji

Sposób dokowania ligandów

Metoda uśrednionego gridu

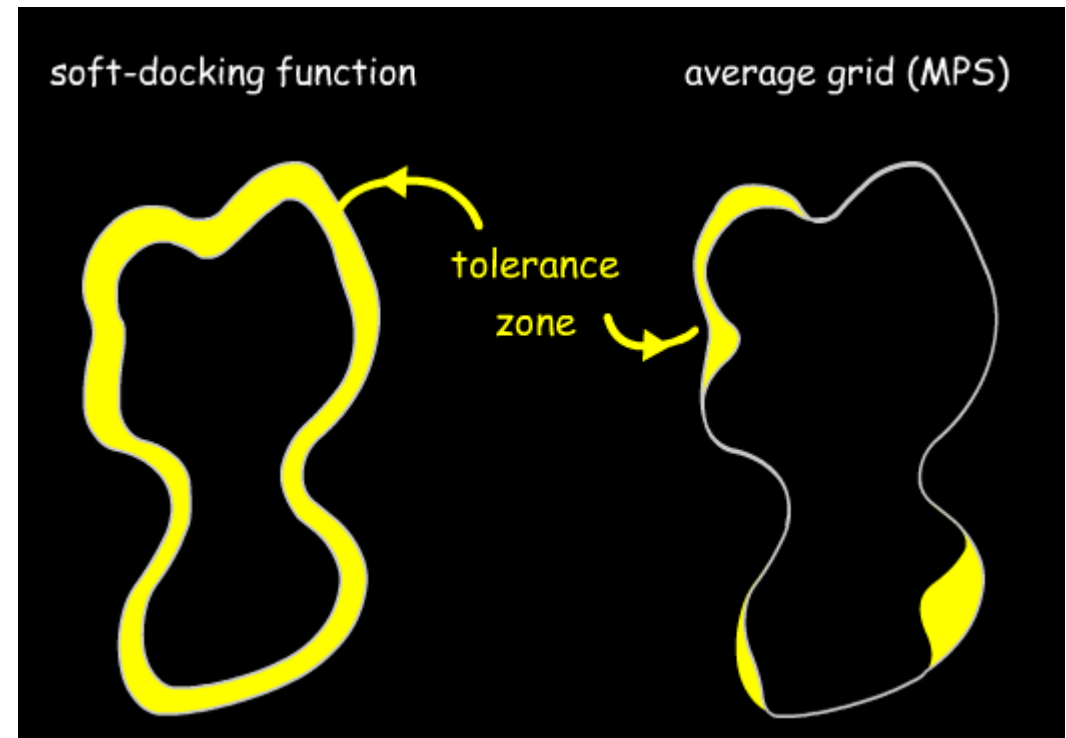
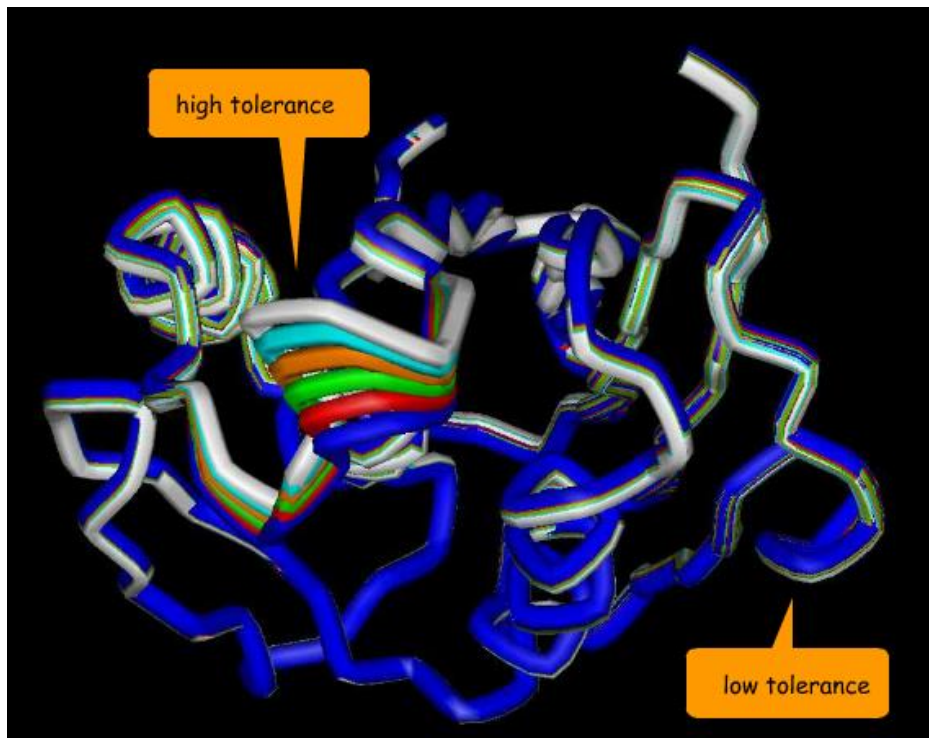


Energie oddziaływania liczone dla każdej struktury białka oddzielnie



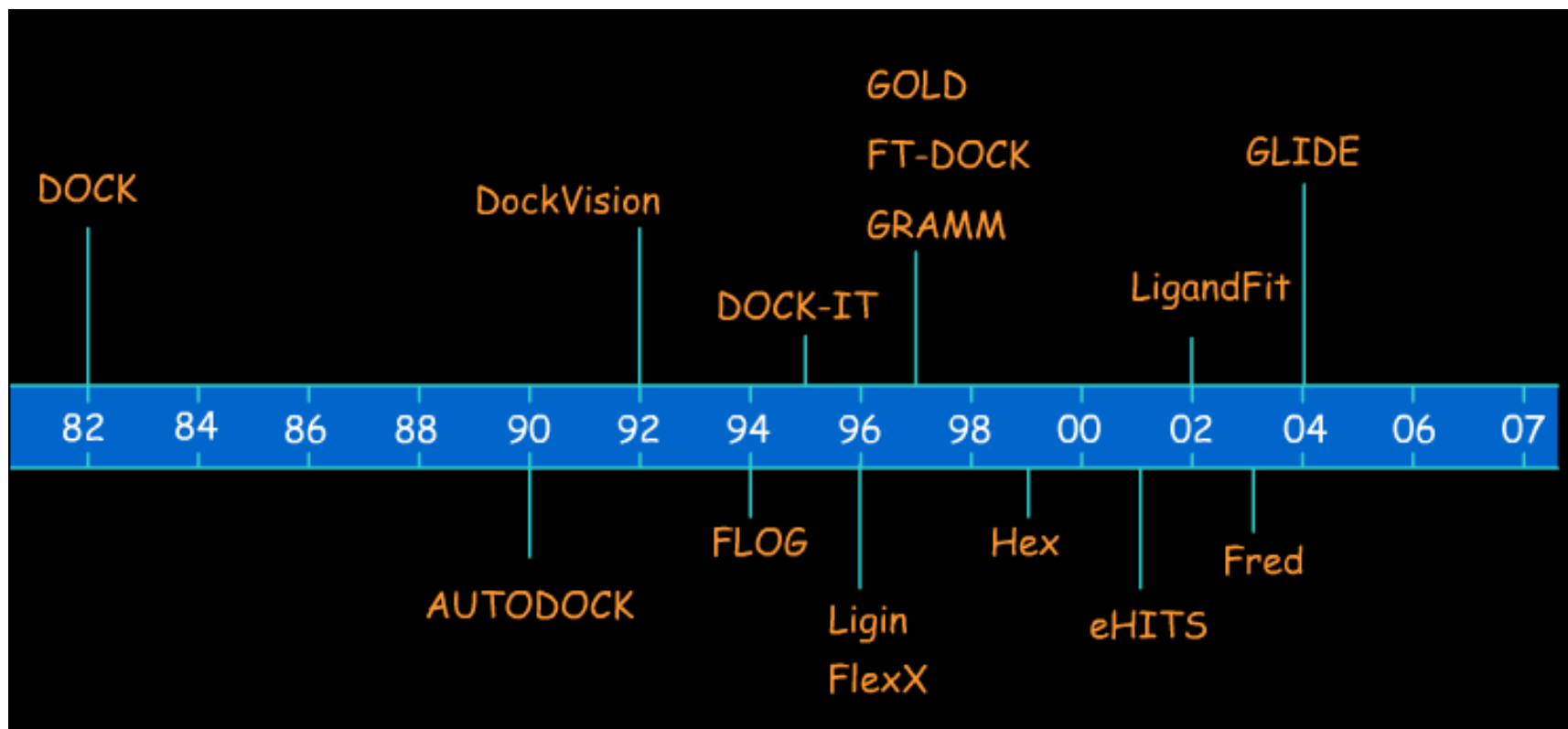
Uśredniony grid

Tolerancja dokowania ligandów

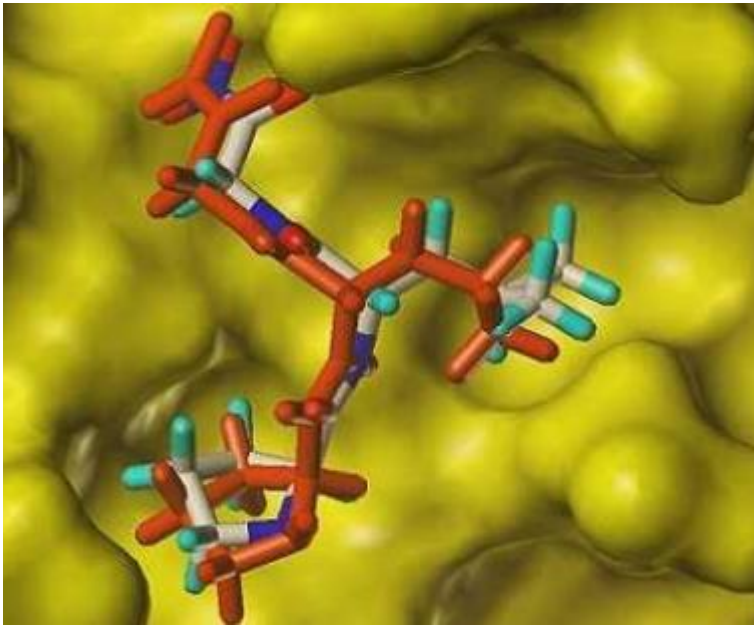




Programy do dokowania



Program GOLD

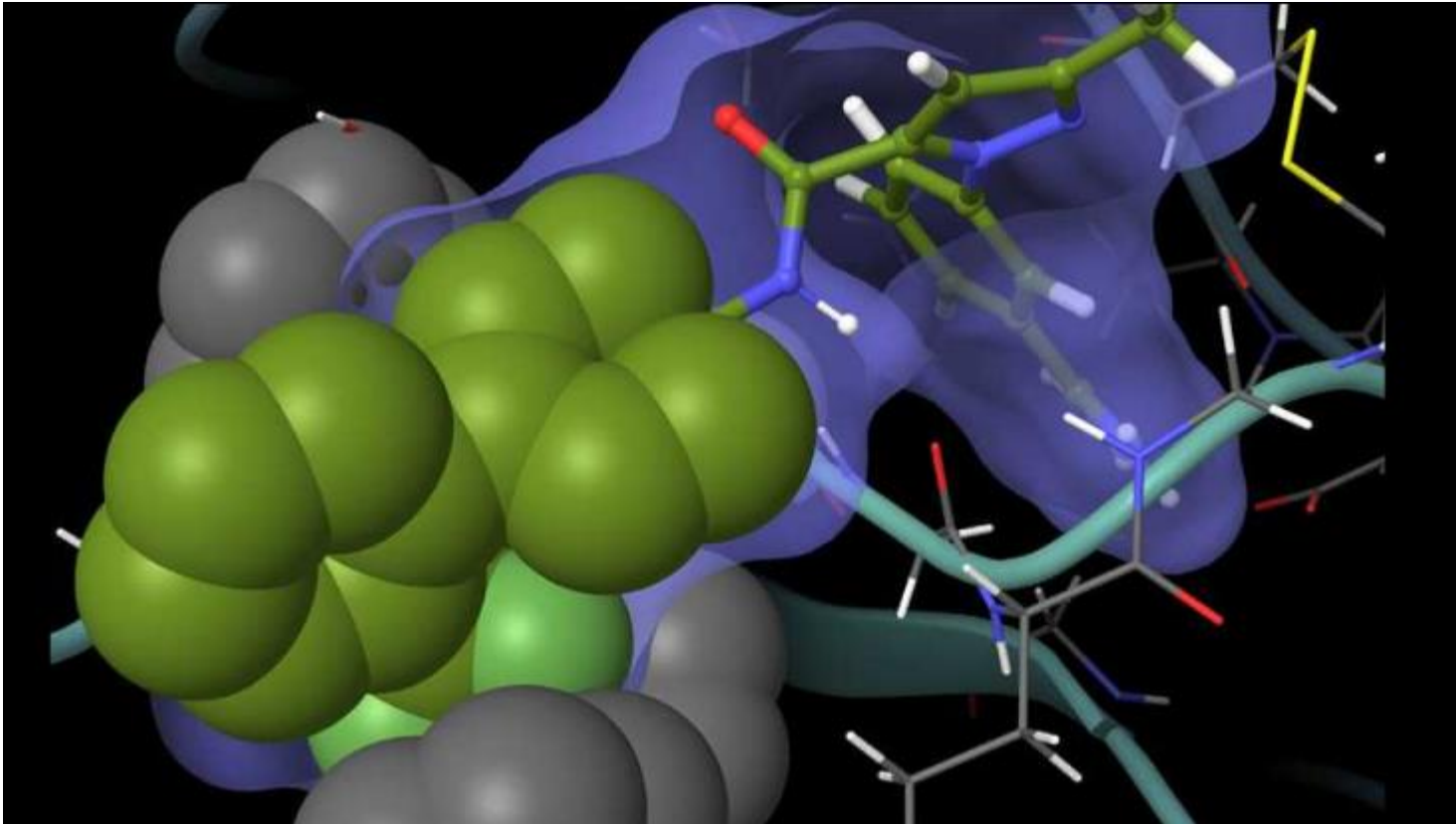


Dokowanie inhibitora do metaloproteazy
(porównanie ze strukturą krystaliczną)

Własności:

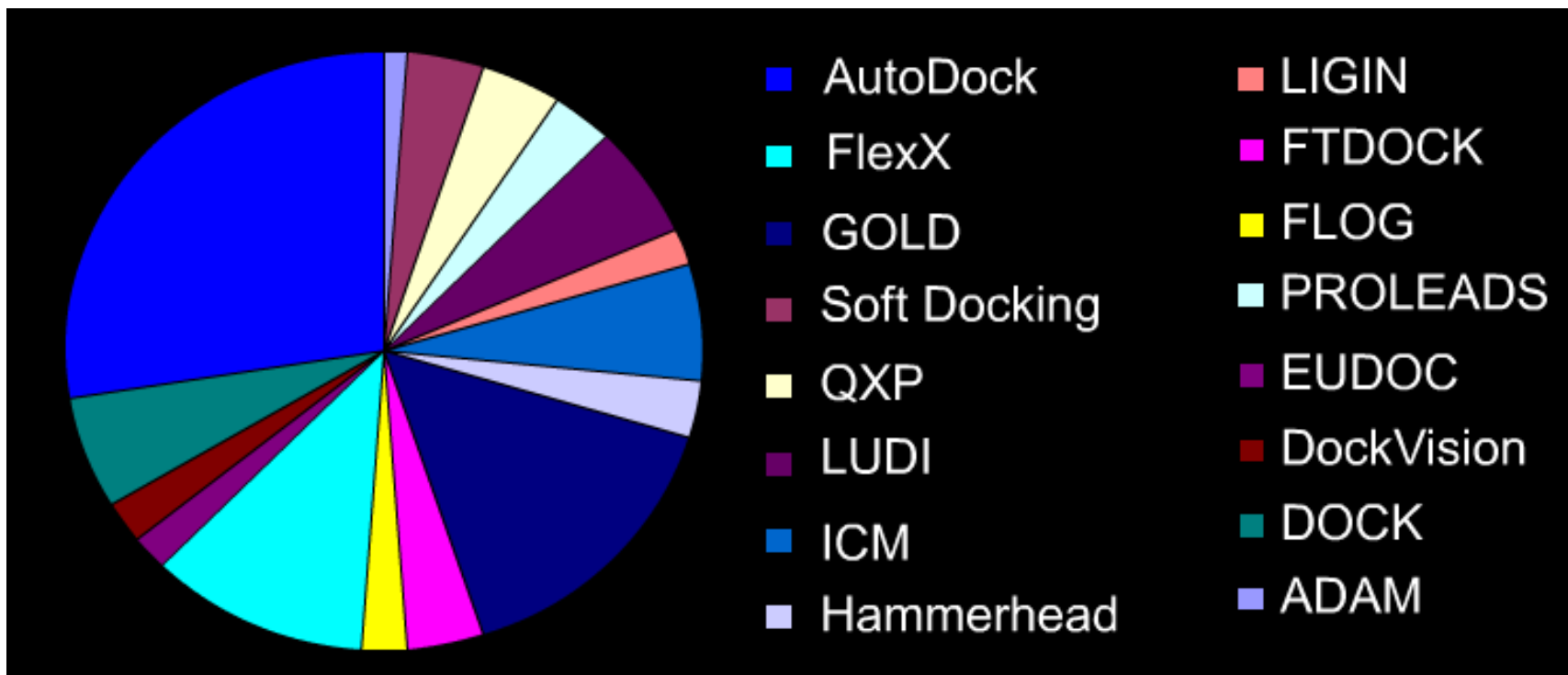
- Algorytmy genetyczne do dokowania ligandu
- Giętki ligand i częściowo giętkie białko
- Funkcje oceniające oparte na strukturach realnych ligandów
- Możliwość wyboru funkcji oceniającej

Program GLIDE (Schrödinger Inc.)



Stosowana szybka metoda systematycznego przeszukiwania
(powrót do tej metody za sprawą komputerów wieloprocessorowych
i obliczeń masywnie równoległych)

Procentowe użycie poszczególnych programów do dokowania





Przykłady leków uzyskanych przy użyciu dokowania

TARGET	DRUG	COMPANY
Carbonic anhydrase	dorzolamide	Merck
Influenza neuraminidase	zanamivir (relenza)	GSK
HIV-1 protease	amprenavir (agenerase)	GSK
Renin	aliskiren (tekturna)	Novartis
Thymidilate synthase	tomudex	AstraZeneca
HIV-1 protease	nelfinavir	Agouron
ABL tyrosine kinase	imitinab mesylate (gleevec)	Novartis
HIV-1 Protease	saquinavir	Roche
HIV-1 Protease	indinavir	Merck