

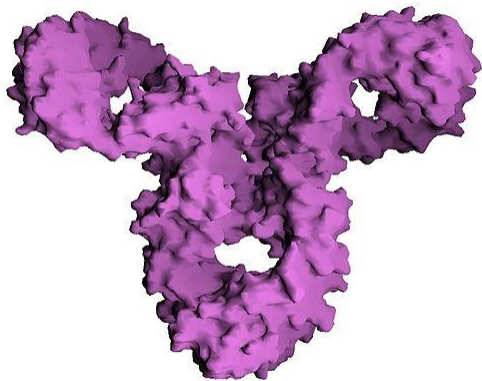
Analiza modelowania homologicznego, dokowanie molekularne i analiza oddziaływań

Semestr Letni 2018

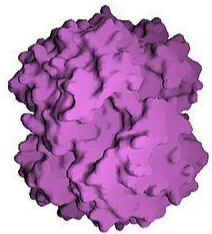
Magdalena Mozolewska

Białka

- *Wielkocząsteczkowe (masa cząsteczkowa od ok. 10 000 do kilku mln Daltonów) biopolimery, a właściwie biologiczne polikondensaty, zbudowane z reszt aminokwasów połączonych ze sobą wiązaniami peptydowymi -CONH-. Występują we wszystkich żywych organizmach oraz wirusach.*



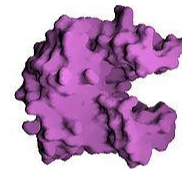
Przeciwciało (IgG)



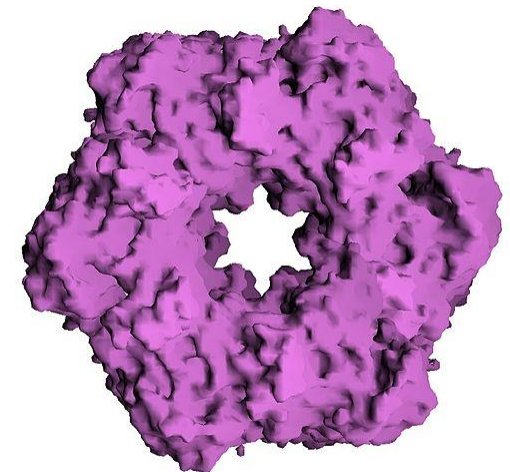
Hemoglobina



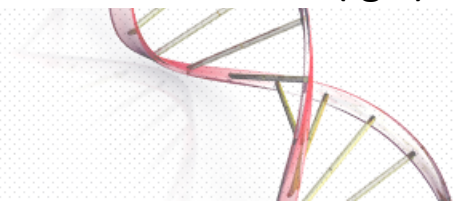
Insulina



Kinaza AK1



Ligaza glutaminy



Białka

- Stanowią aż 75% suchej masy tkanek miękkich naszego ciała
- Zróżnicowanie właściwości fizykochemicznych poszczególnych reszt aminokwasowych umożliwia tworzenie niezwykle różnorodnych struktur białek o skomplikowanej topologii i różnych funkcjach



STRUKTURALNE

Odpowiedzialne za mechaniczną stabilność narządów i tkanek. Do białek strukturalnych zalicza się także histony pełniące kluczową rolę w upakowaniu DNA w chromatynie

ZAPASOWE

Niektóre białka budujące mięśnie mogą być wykorzystywane jako materiał energetyczny; także wiele białek roślinnych pełni funkcję zapasową

ODPOIM

To białka układu (np. immunologiczne) organizmowi chronią przed chorobami

FUNKCJE BIAŁEK

8457

ENZYMATYCZNE

przyspieszające specyficzne reakcje chemiczne

TRANSPORTUJĄCE

To białka uczestniczące w transportowaniu różnych substancji odżywczych jak np. hemoglobina uczestnicząca w transporcie tlenu i CO₂.

REGULACYJNE

To niektóre hormony, a także receptory uczestniczące w percepcji różnych cząsteczek sygnałowych; białkami regulatorowymi są także czynniki transkrypcyjne, regulujące ekspresję genów

Foods High in Protein



Meat and fish



Cheese



Eggs



Beans



Bread

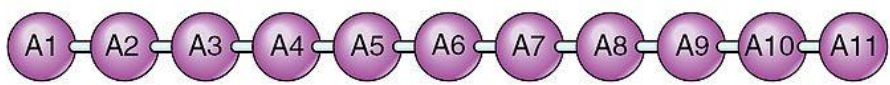


Hummus



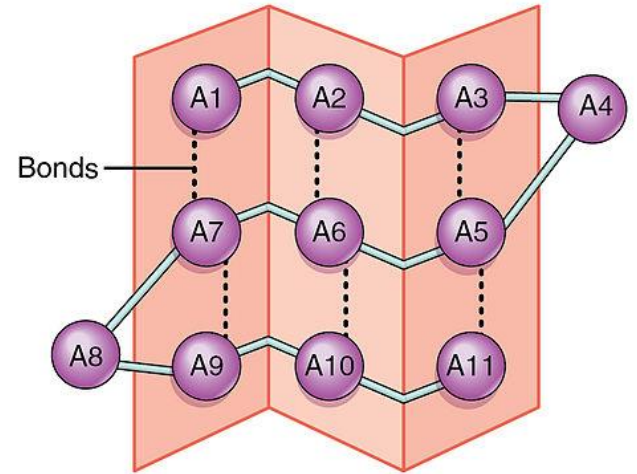
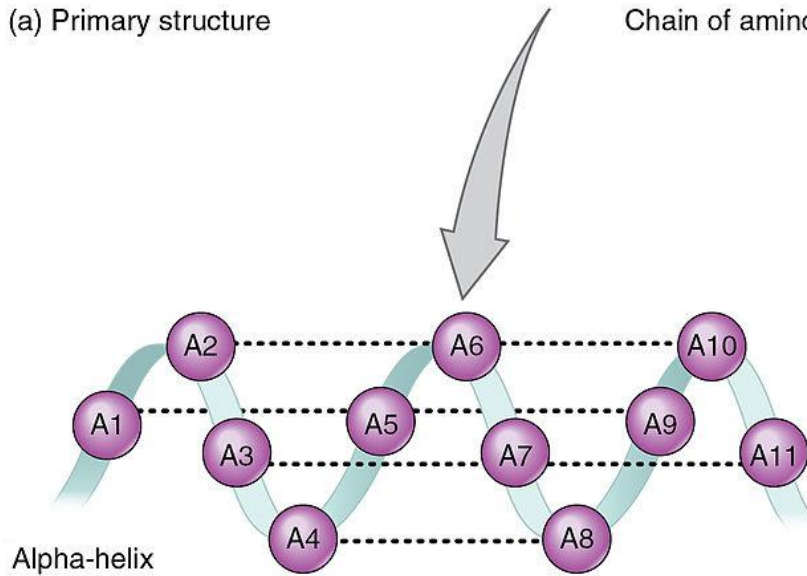
Nuts and seeds

©Nutrientsreview.com

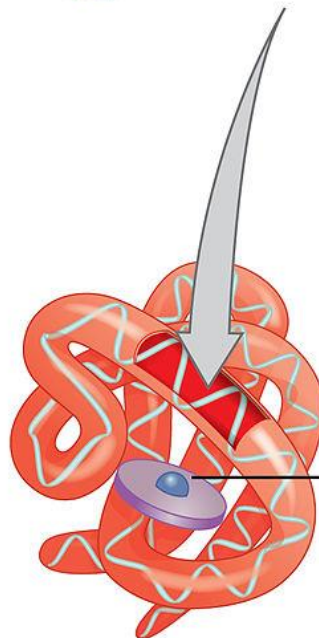


(a) Primary structure

Chain of amino acids

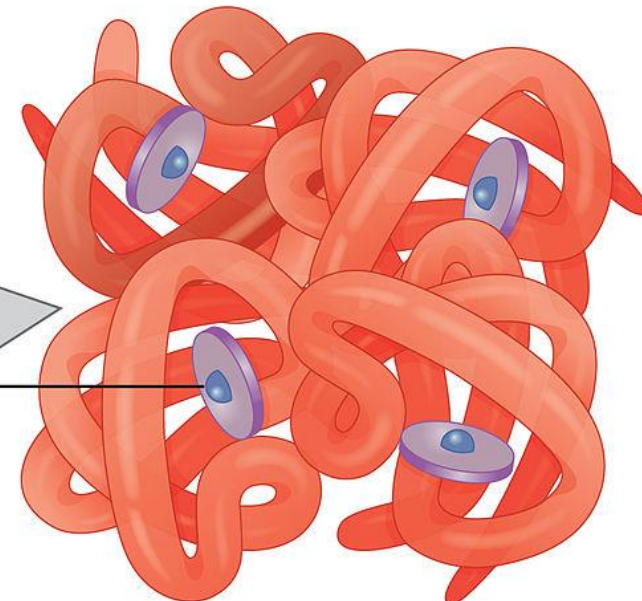


(b) Secondary structure (pleated sheet)



(c) Tertiary structure

Heme units



(d) Quaternary structure

Hemoglobin
(globular protein)

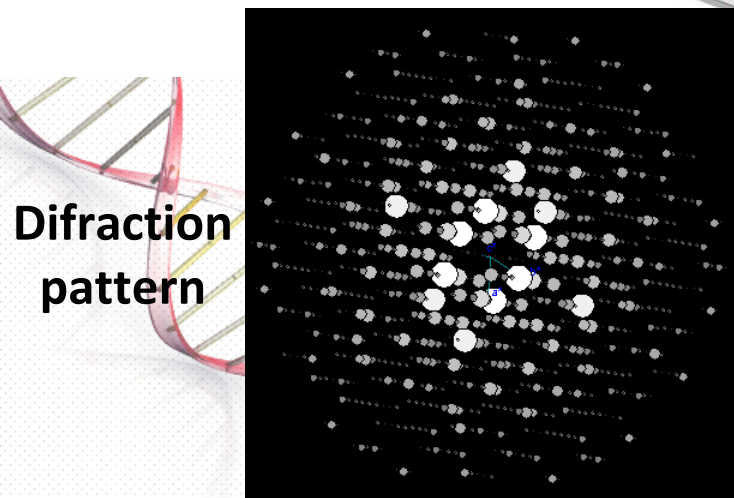
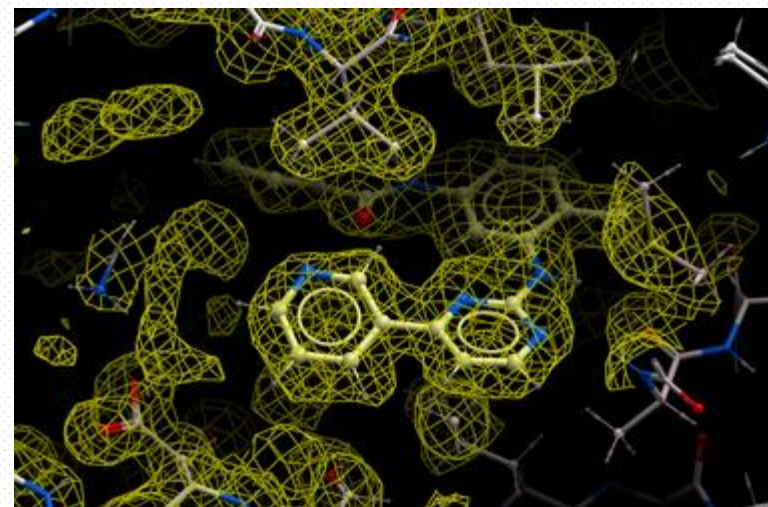
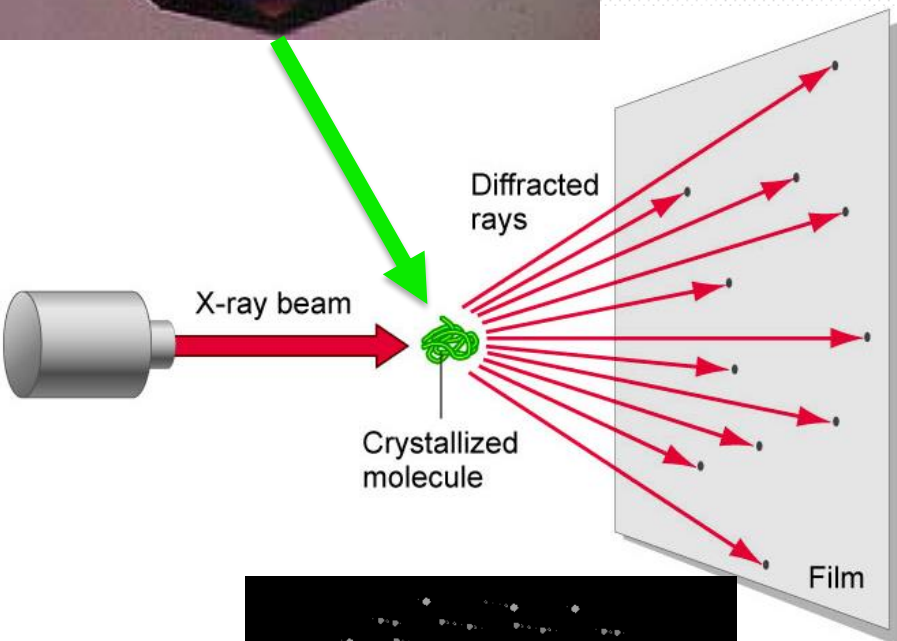
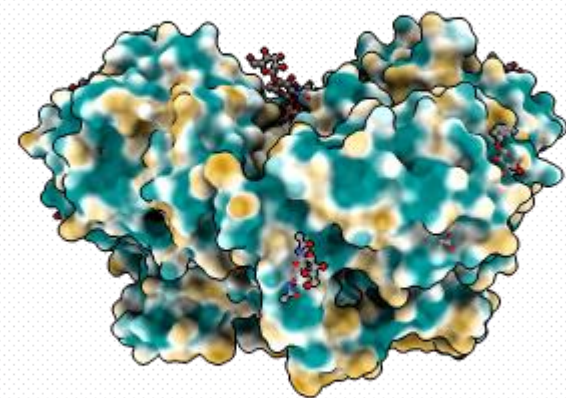
Protein structure

- Solving:
 - X-ray crystallography
 - Requires crystals (problematic and expensive)
 - Very good resolution
 - Protein size is not a big problem
 - Nuclear magnetic resonance spectroscopy
 - 3D and higher orders spectra
 - Only small proteins
 - Cryo-electron microscopy
 - Low resolution
 - Mostly large structures
 - FRET, SAXS, EPR, MS cross-linking, etc. – very low resolution methods
- Prediction:
 - Modeling and bioinformatical tools

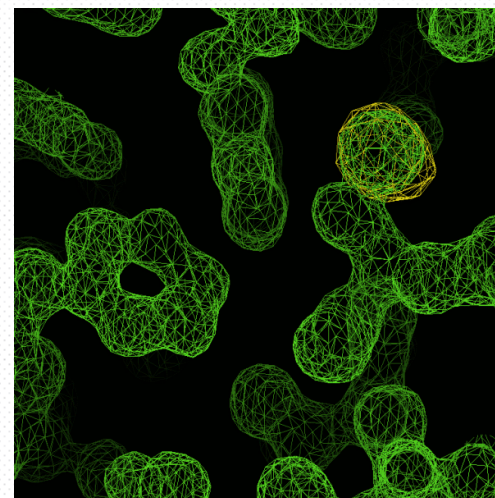




X-ray

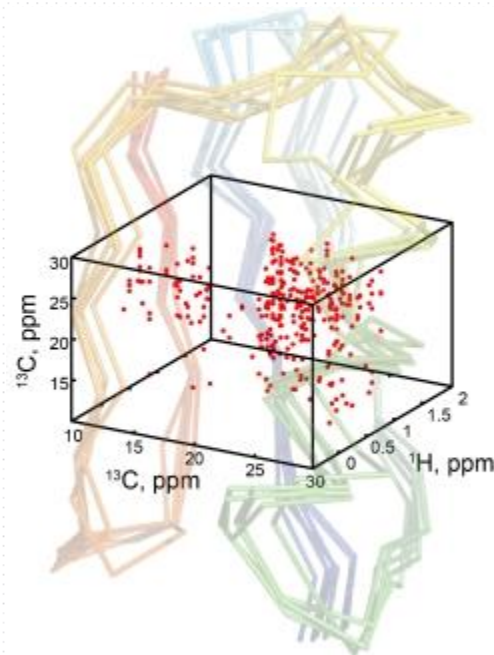
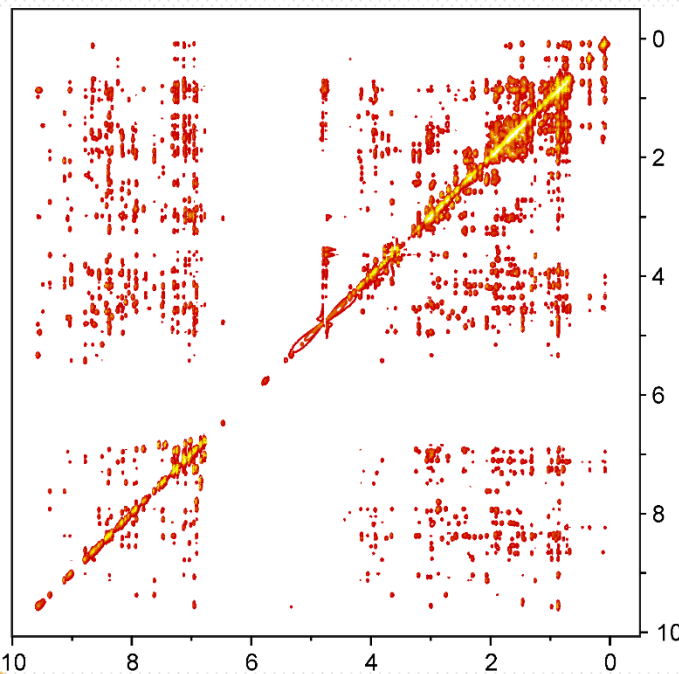


(Inverse) Fourier Transformation

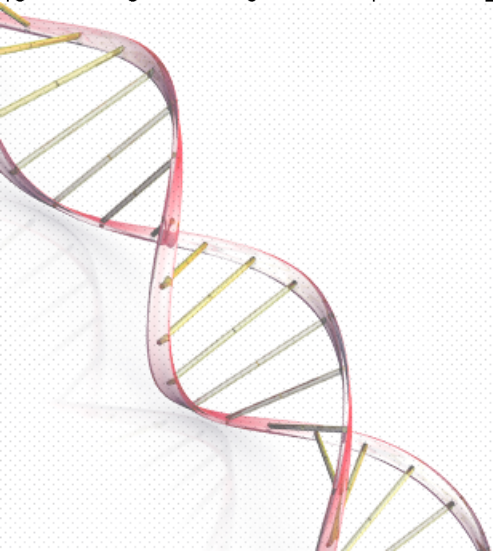
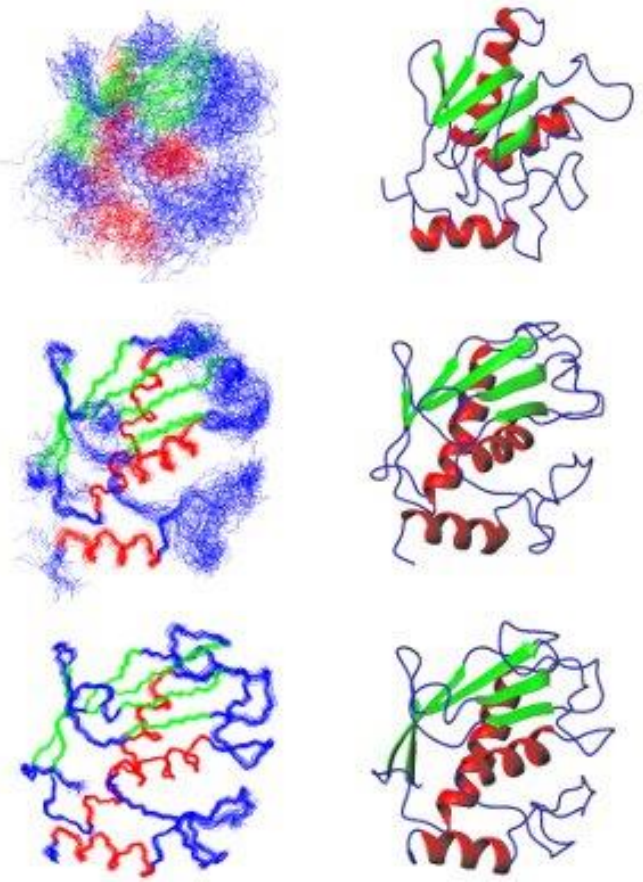


Electron density

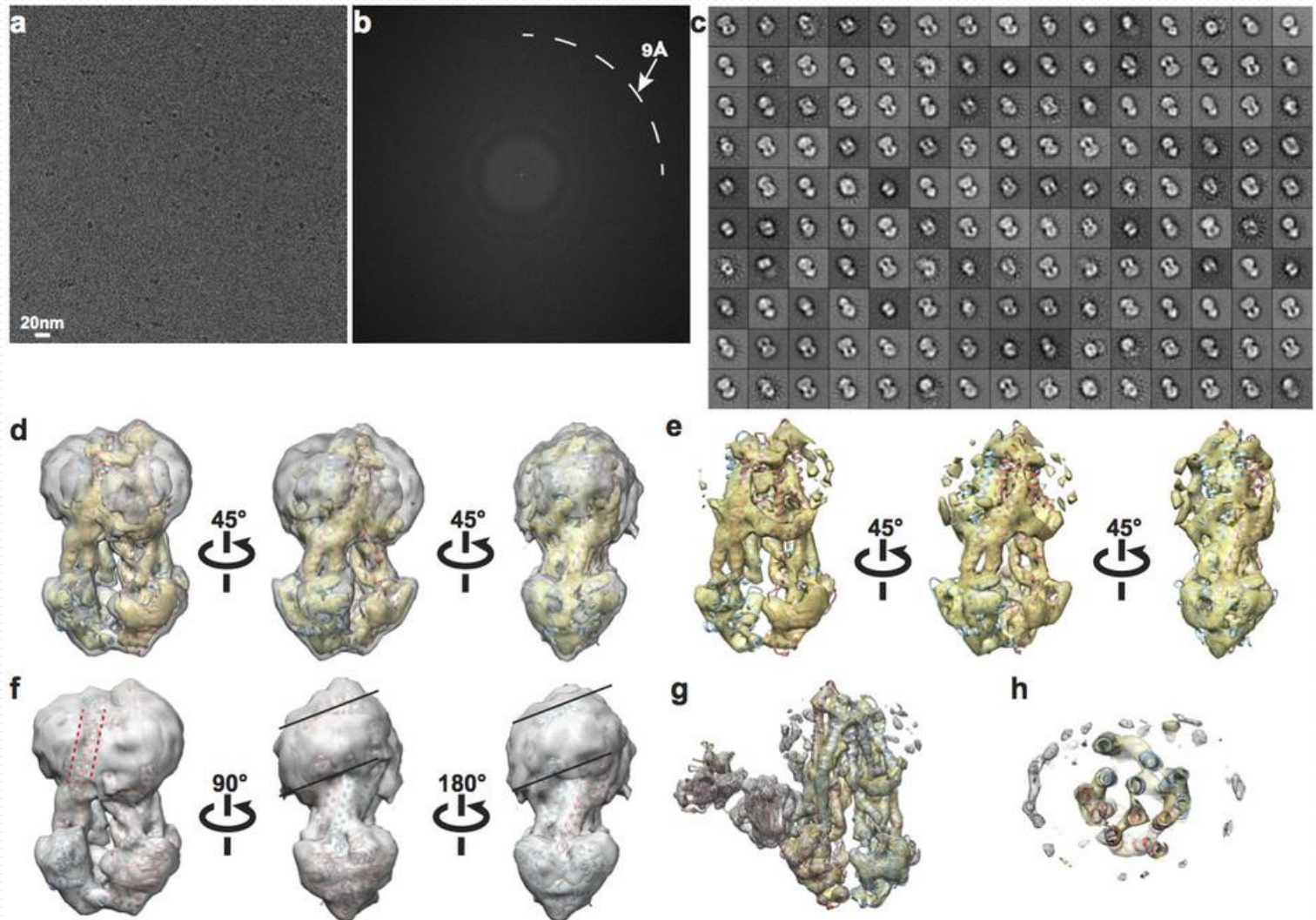
NMR – jądrowy rezonans magnetyczny

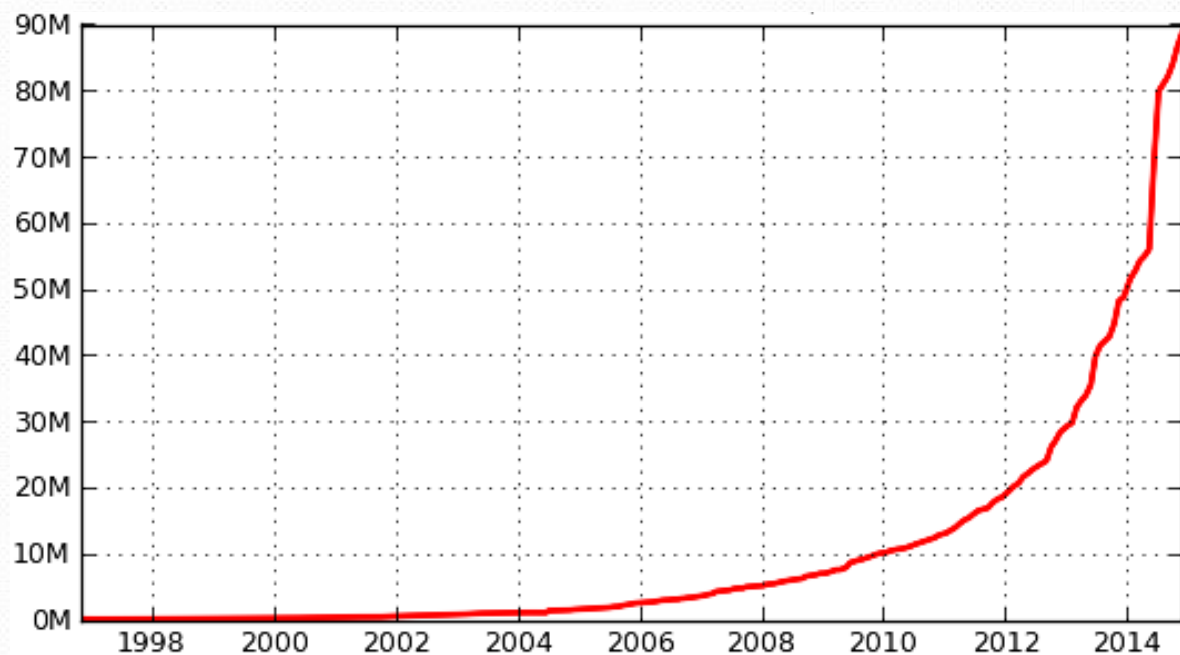


MMP-1 NMR Structure as a Function of the Total Number of Restraints used in the Simulation



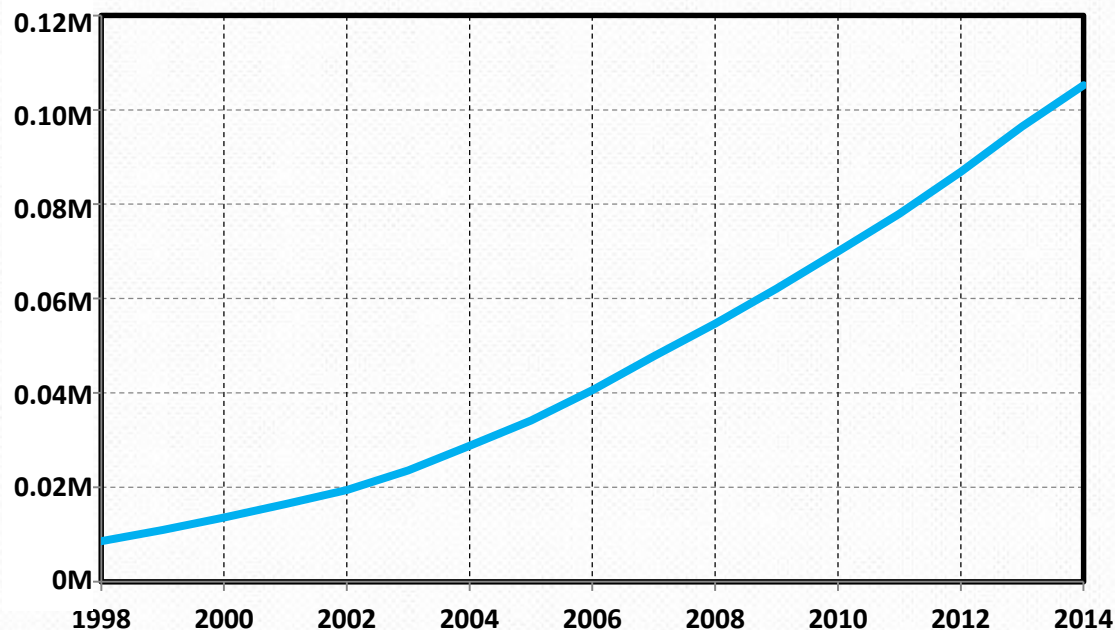
Kriomikroskopia elektronowa





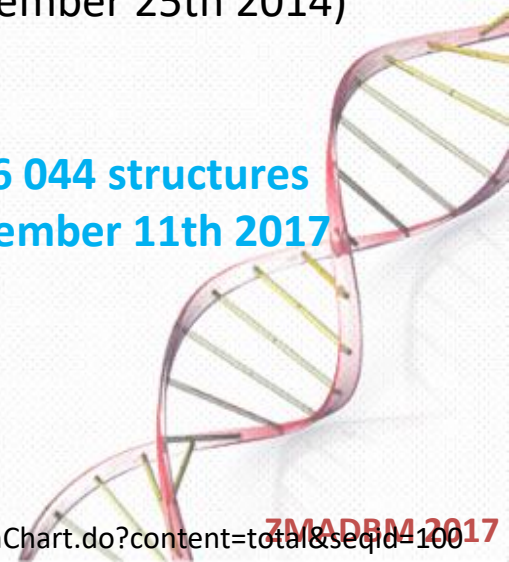
Number of **protein sequences** in
UniProtKB database
(November 26th 2014)

556 196 reviewed
98 705 220 unreviewed
December 11th 2017



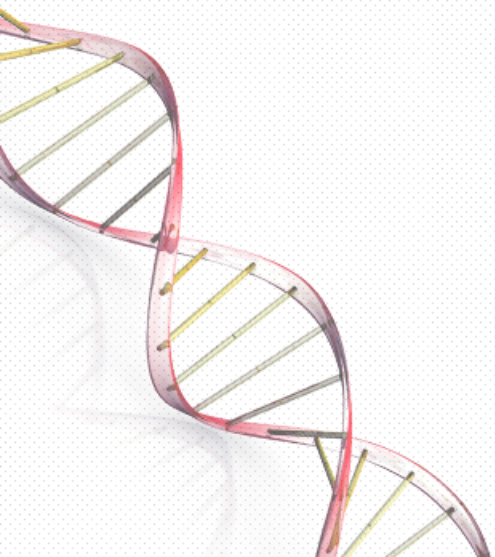
Number of **protein structures** in
RCSB database
(November 25th 2014)

126 044 structures
December 11th 2017



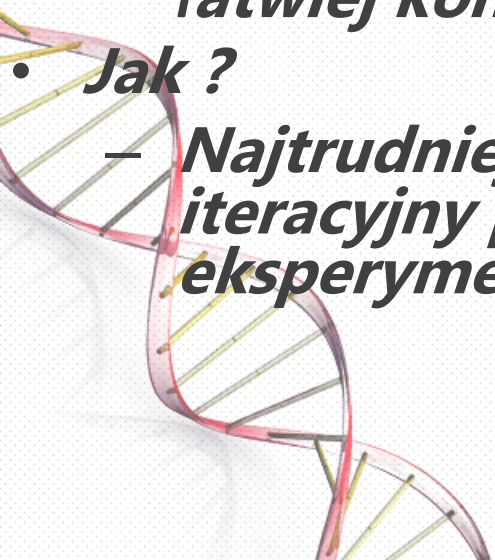
Modelowanie molekularne – co to jest ?

- Opis skomplikowanych układów na poziomie atomowym
- Zrozumienie lub przewidywanie własności makroskopowych układu poprzez analizę zachowania układu w skali mikroskopowej



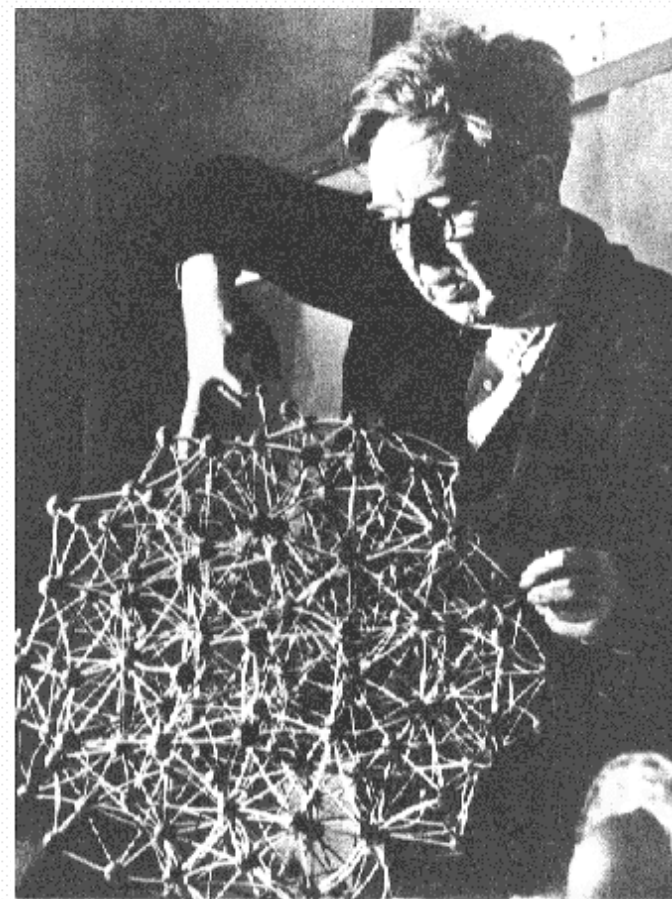
Chemia obliczeniowa

- Co?
 - *Probuje odpowiedzieć na te same pytania co chemia eksperymentalna w laboratorium, ale wykorzystuje podstawowe prawa fizyczne, intuicję chemiczną i coraz szybsze komputery.*
- Dlaczego ?
 - *Mniej uciążliwa, często tańsza; niektóre zjawiska trudno obserwować w laboratorium, w symulacji łatwiej kontrolować warunki.*
- Jak ?
 - *Najtrudniejsze jest właściwe postawienie pytań, iteracyjny proces pomiędzy nowymi teoriami a eksperymentem.*



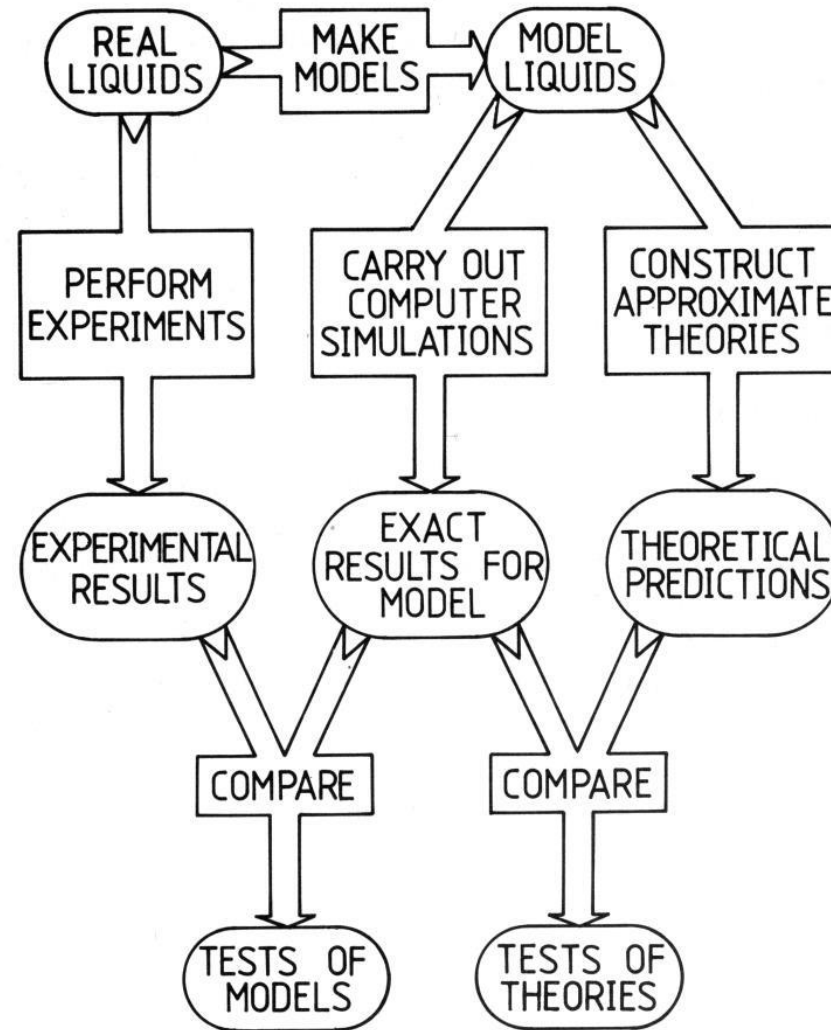
Chemia obliczeniowa i chemia teoretyczna

- Prawa natury są wyrażone równaniami których nie potrafimy rozwiązać analitycznie z wyjątkiem najprostszych układów
- Modelowanie pozwala na sprawdzanie prawidłowości stawianych hipotez dla rzeczywistych układów
- Interakcja teoria-obliczenia-eksperyment



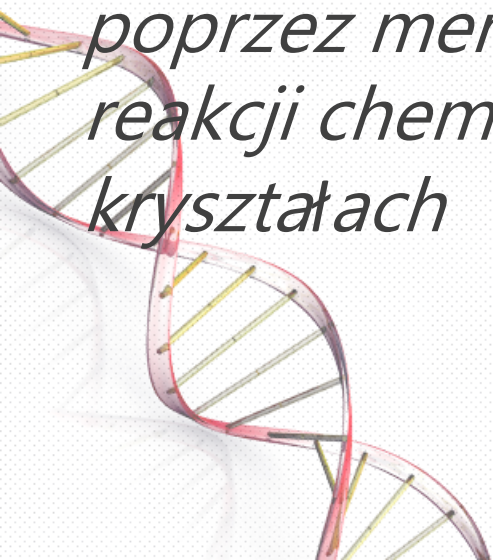
Model struktury wody
zbudowany przez J.D.Bernal

Schemat interakcji teoria-obliczenia-eksperyment



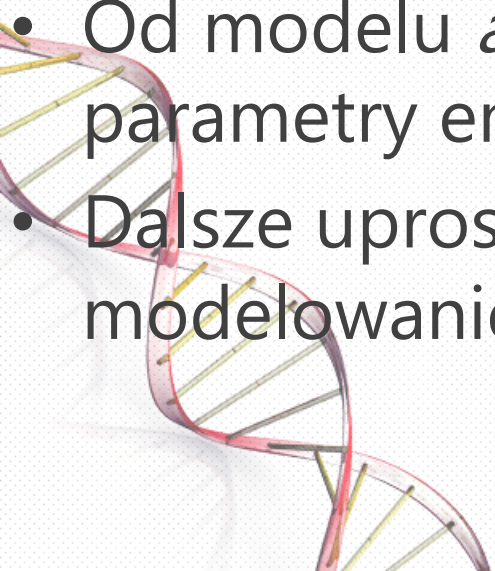
Modelowanie molekularne pozwala na wyznaczenie:

- Statycznych – równowagowych własności układu: *np. stała wiązania inhibitora do enzymu, średnia energia potencjalna układu, radialna funkcja rozkładu w cieczy*
- Dynamicznych – nierównowagowych własności układu: *np. lepkość cieczy, szybkość dyfuzji poprzez membrany, przejścia fazowe, kinetyka reakcji chemicznych, dynamika defektów w kryształach*



Wybór modelu i metod obliczeniowych

- Jakie stawiamy pytania oraz jakiej dokładności i wiarygodności żądamy ?
- Konieczność uwzględnienia przybliżeń w zależności od stopnia skomplikowania układu i czasu w jakim zachodzą zjawiska
- Od modelu *ab initio* do modelu opartego o parametry empiryczne
- Dalsze uproszczenia- analiza porównawcza, QSAR, modelowanie oparte na podobieństwie



Prediction of protein structure

- Homology (template-based) modeling
 - Based on the databases
 - Often servers
- *De novo*
 - Based on physics of interactions
 - Molecular dynamics (MD) or Monte Carlo (MC) simulations
 - Protein folding
- Mixed
 - Both approaches with possible additions (e.g. genetic information)



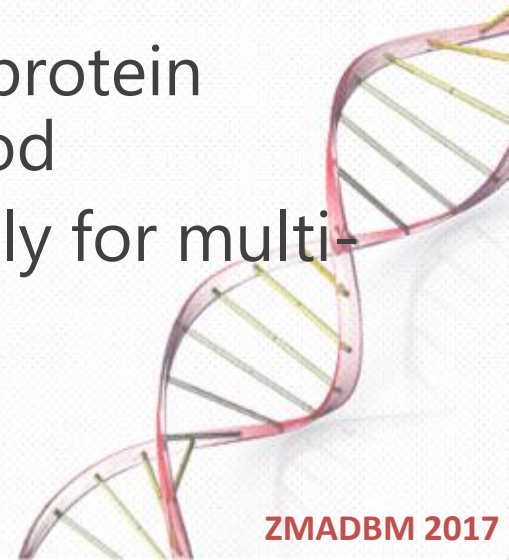
***De novo* methods**

- Based on physics of interactions
- Using molecular dynamics (MD) simulations allows to study protein folding and conformational changes
- Include dynamics and thermodynamics of proteins
- Do not rely on databases



TBM and Mixed methods

- Template-based modeling:
 - Best performance for majority of the proteins
 - Bad performance for non-standard proteins
- Mixed methods:
 - Combines template-based and *de novo* methods
 - Difficult to assess which part of the protein should be predicted by which method
 - Excellent peak performance especially for multi-domain proteins



CASP

C
A
S
P
12



- The Critical Assessment of protein Structure Prediction.
- Completely blind prediction (only protein sequence is provided).
- Over 100 groups are involved in the predictions.
- Started at 1994 (CASP1) and last edition in 2016 (CASP12) – protein structure problem still not solved.
- Released proteins have CASP codes: T0XXX.

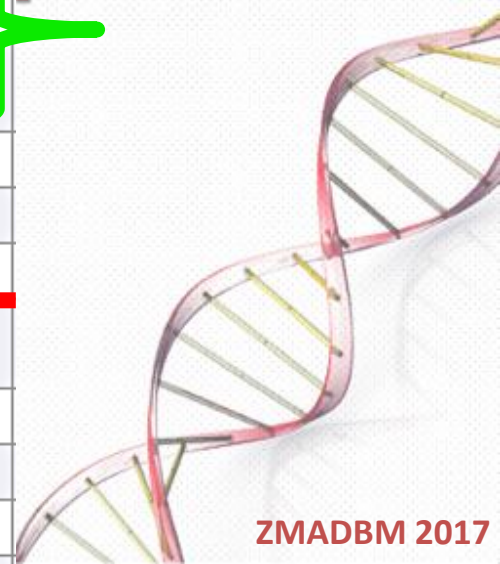


- ☒ Analysis on the models designated as "1"
 - ☐ Analysis on the models with the best scores
- ☐ All groups on 'all groups' targets
 - ☒ Server groups on 'all groups' + 'server only' targets
- The ranking of the groups is based on the analysis of zscores for GDT_TS.
 - ☒ TBM
 - ☒ TBM/FM
 - ☒ FM

Show

#	GR code	GR name	Domains Count	SUM Zscore (>-2.0)	Rank SUM Zscore (>-2.0)	AVG Zscore (>-2.0)
1	479	Zhang-Server	96	98.6995	1	1.0281
2	183	QUARK	96	93.4959	2	0.9739
3	005	BAKER-ROSETTASERVER	96	92.3649	3	0.9621
4	220	GOAL	94	79.6005	4	0.8894
5	092	RaptorX	96	67.2517	5	0.7005
6	048	ToyPred_email	95	63.8482	6	0.6931
7	236	MULTICOM-CONSTRUCT	96	36.3879	7	0.3790
8	287	MULTICOM-CLUSTER	96	34.2757	8	0.3570
9	345	MULTICOM-NOVEL	96	33.2208	9	0.3461
10	405	IntFOLD4	96	26.3897	10	0.2749

**Best
servers
for
protein
structure
prediction**



CAMEO

CAMEO | Continuous Automated Model EvaluatiOn

YLDVGFDTTRVAVIQFVLSGK
SDFSNDVFPEFADRSQDVRAP
SVVVKRGGAVPIGIGMADITIS

★ Try our BETA!

Home

3D - Protein Structure

QE - Model Quality Estimation

CP - Contact Prediction

More ▾

Login +

CAMEO continuously evaluate the accuracy and reliability of predictions

3D - Protein Structure

309 weeks, 5484 targets, 31 predictors.

QE - Model Quality Estimation

193 weeks, 27652 structural models, 20 predictors.

CP - Contact Prediction

64 weeks, 68 targets, 4 predictors.

★ Predictions in all categories are evaluated against reference structures released by the PDB on a weekly basis.

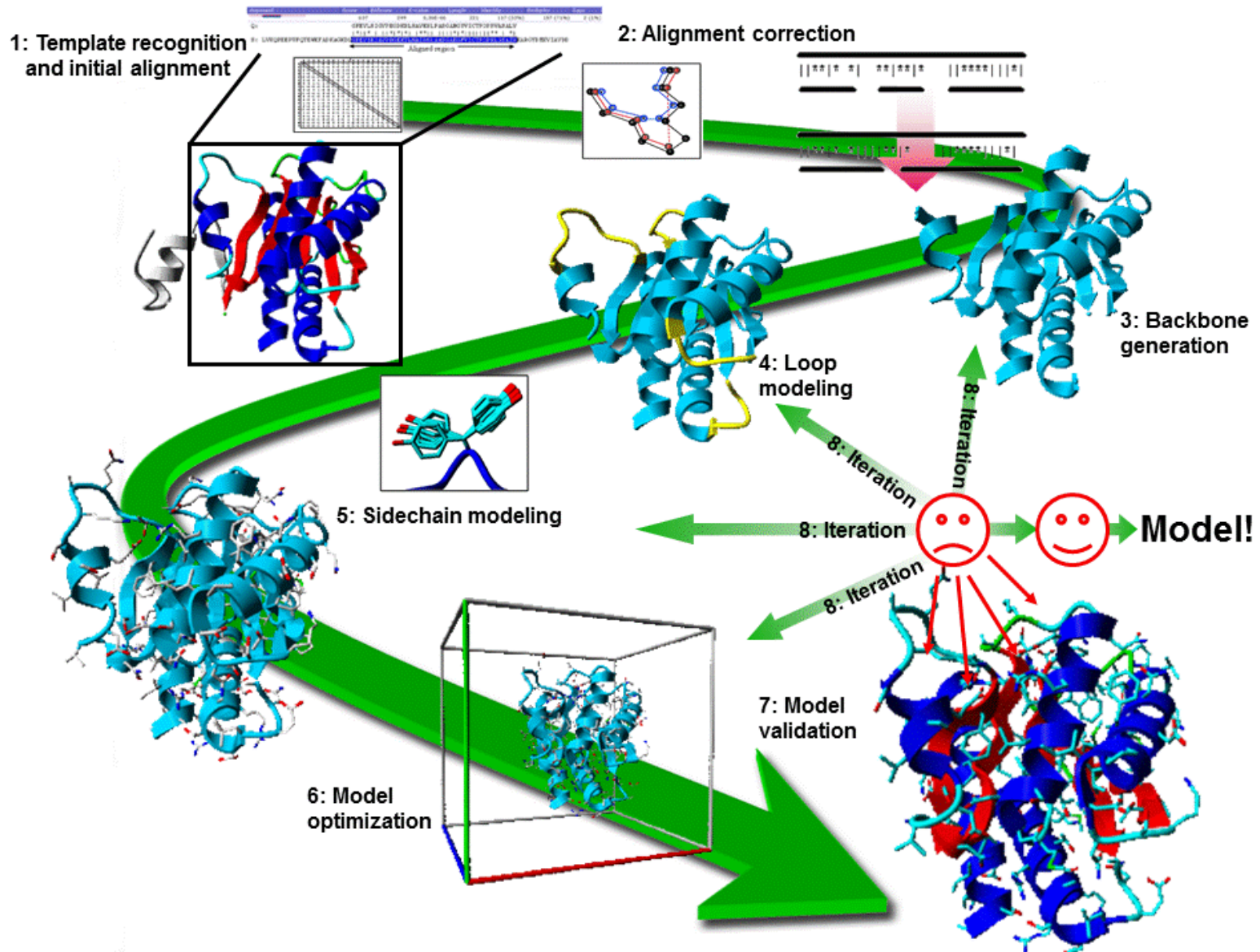


CAMEO is a community project

→ CAMEO continuously applies quality assessment criteria established by the protein structure prediction community. Since the accuracy requirements for different scientific applications vary, there is no "one fits all" score. CAMEO therefore offers a variety of scores - assessing different aspects of a prediction (coverage, local accuracy, completeness, etc.) to reflect these requirements.

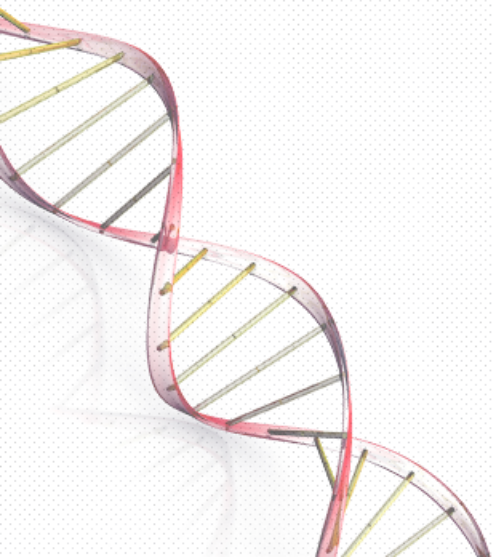
→ CAMEO is a community project - please feel free to suggest additional/alternative ways how CAMEO can support users and developers of structure prediction.

TBM methods



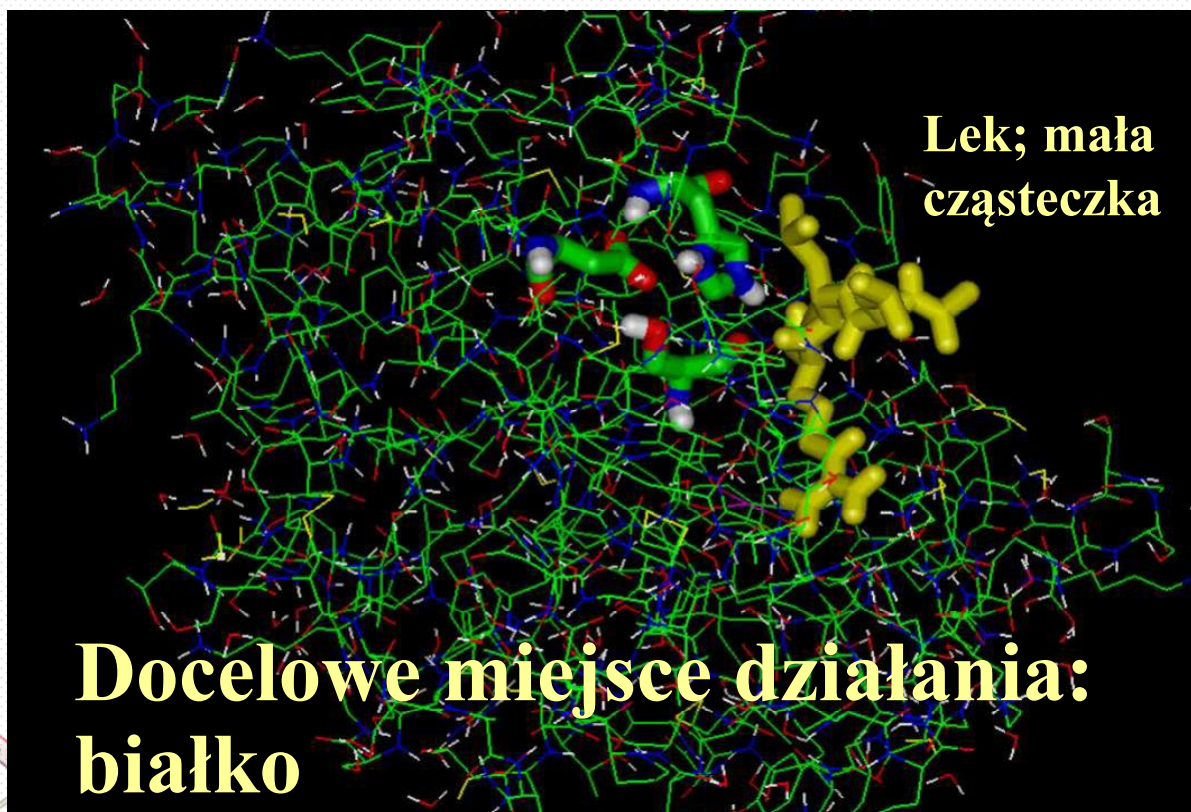
Dokowanie molekularne i analiza oddziaływań

- Semestr Letni 2018



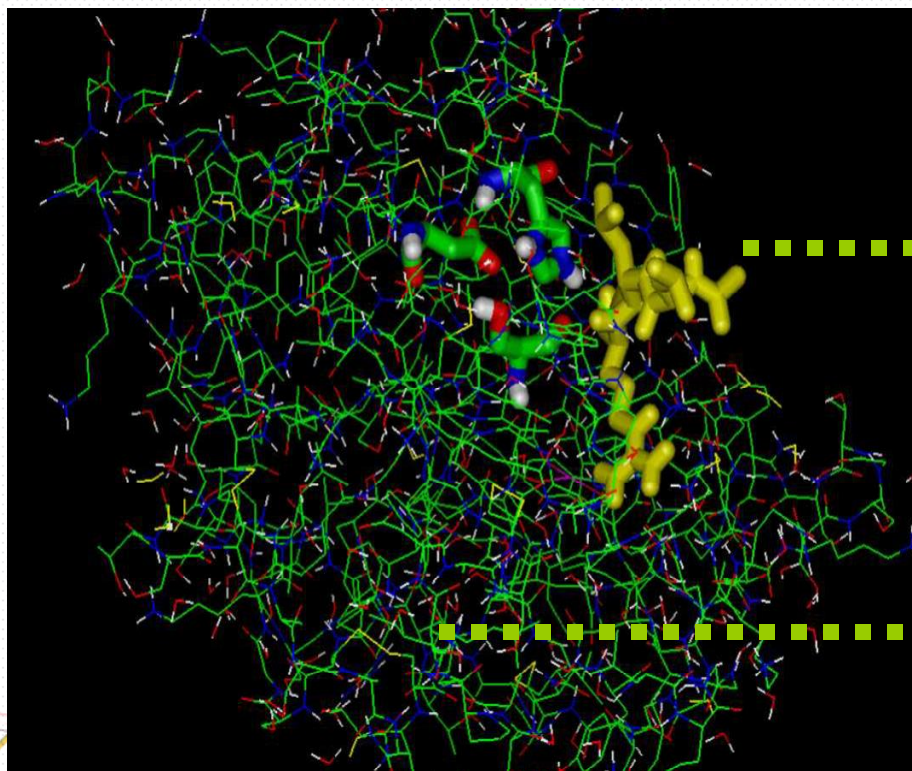
Magdalena Mozolewska

Jak działają leki ?



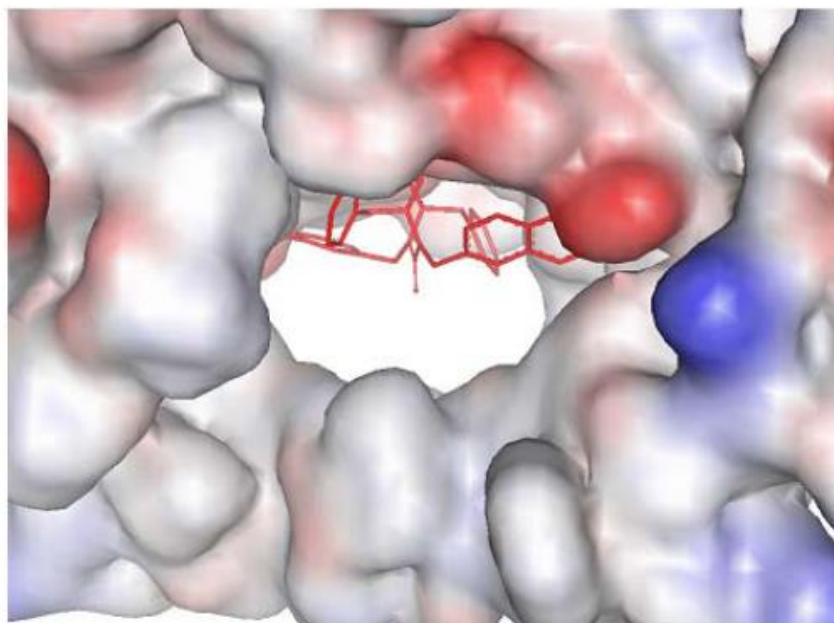
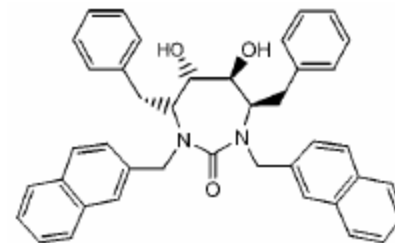
Białko i ligand : zamek i klucz ?

➤ Jak zaprojektować lek ?



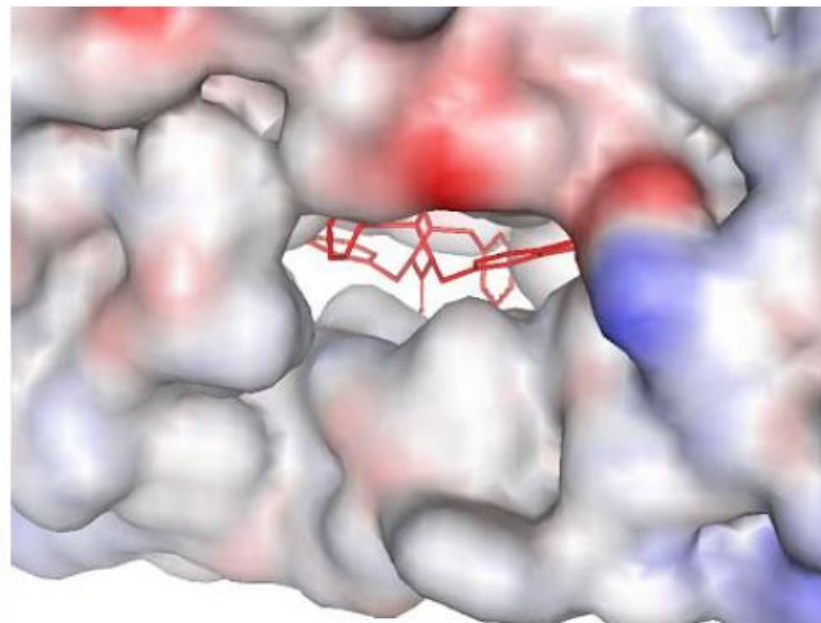
Białko i ligand : zamek i klucz ?

Dokowanie z wykorzystaniem strategii „**Induced fit**”



HIV protease inhibitor

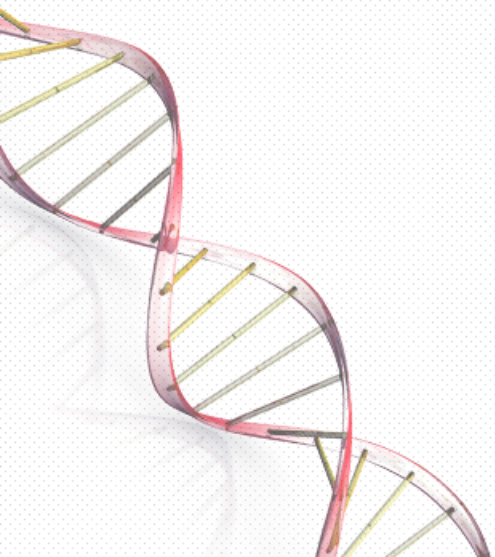
AutoDock – rigid environment



VAMP – flexible environment

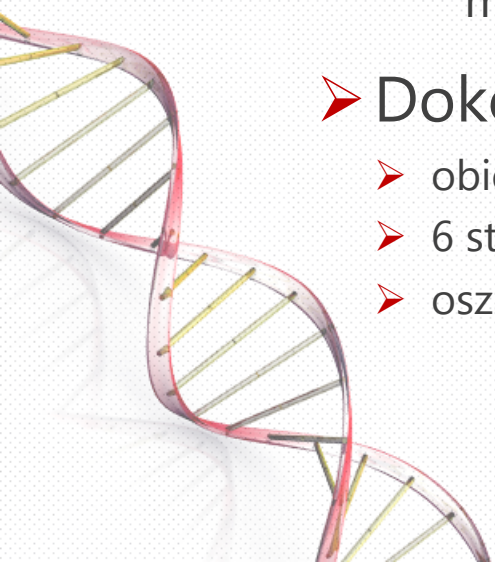
Znaczenie biologiczne wiązania „ligandów”

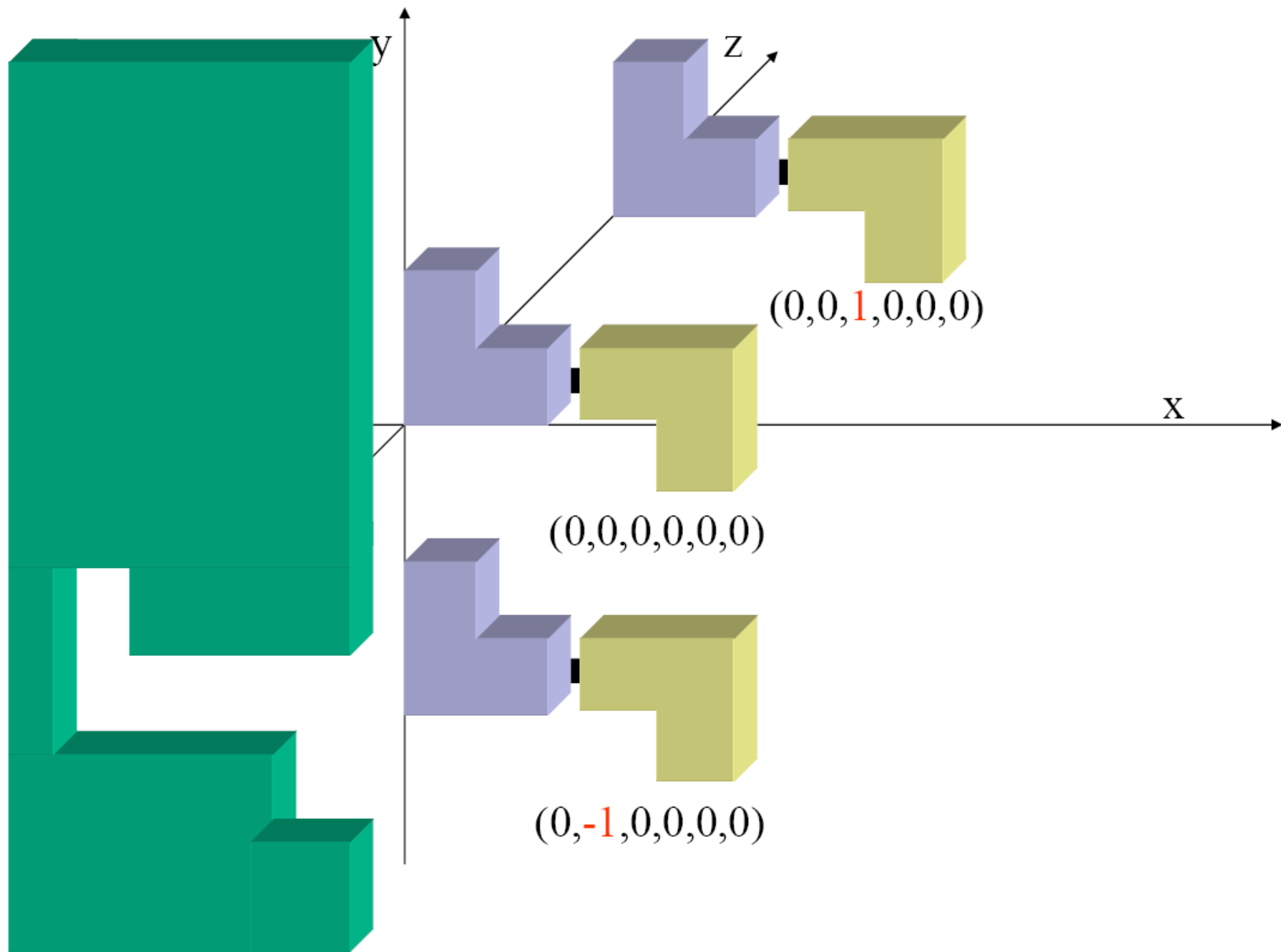
- Oddziaływanie z substratami, reakcje enzymatyczne
- Łączenie inhibitorów, regulacja metaboliczna, transdukcja sygnału
- Agregacja (samorzutne tworzenie większych aglomeratów – kompleksów – białek)

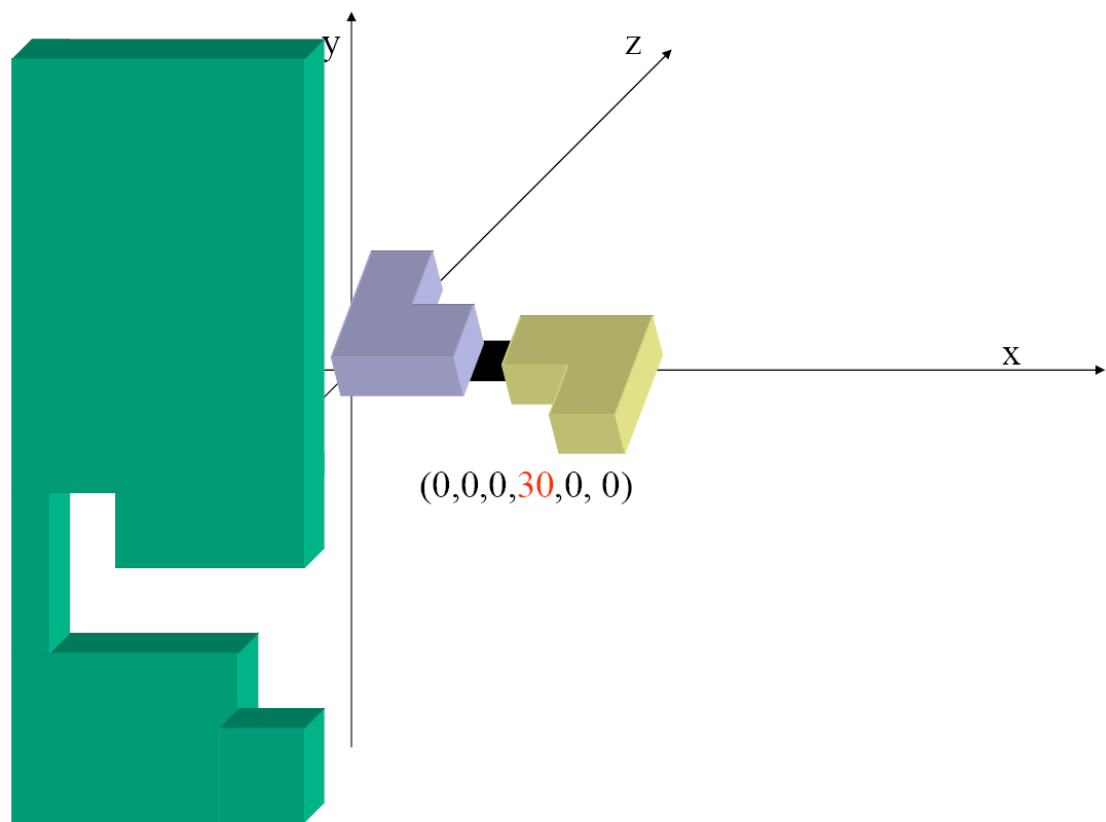
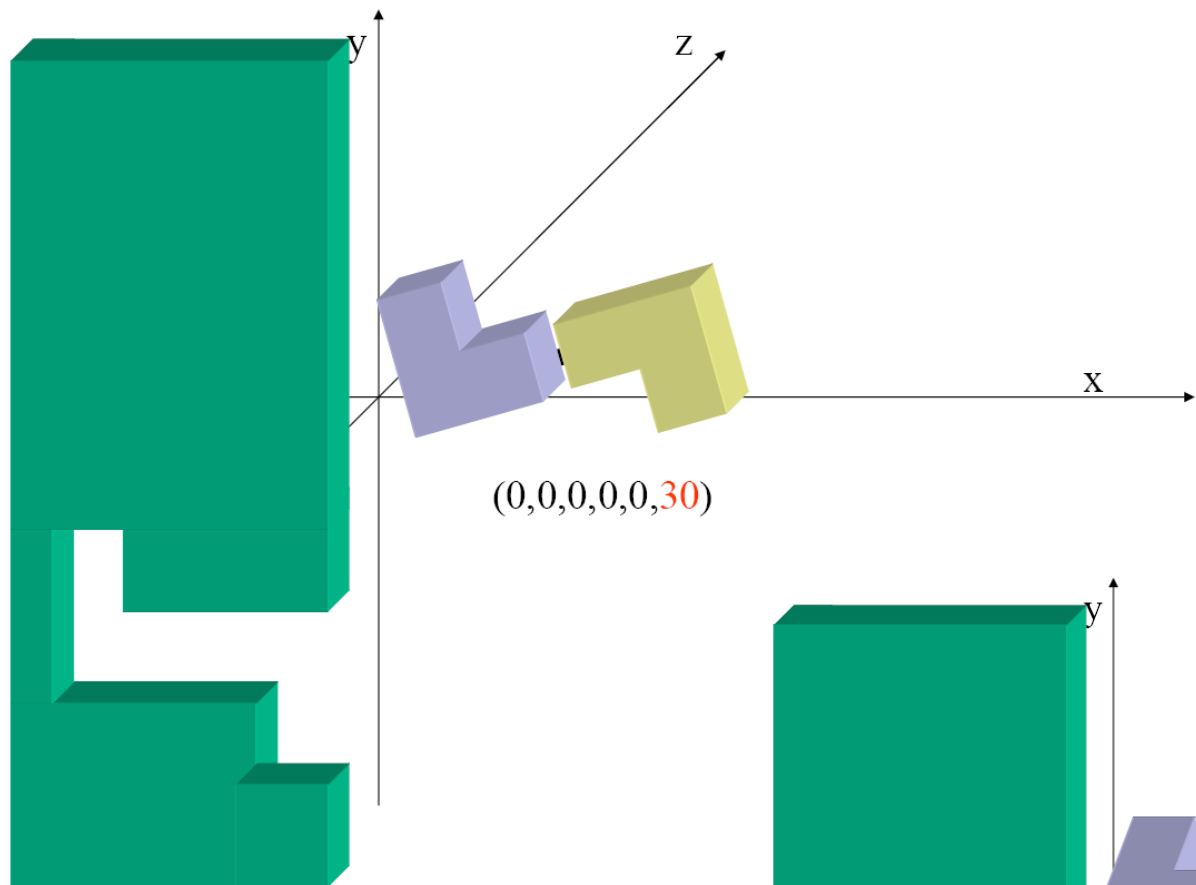


Dokowanie

- Dokowanie: szukanie wzajemnej orientacji dwóch cząsteczek o znanej budowie.
- Dokowanie białko-ligand
 - zwykle ruchomy ligand, nieruchomy receptor
 - duża liczba stopni swobody
 - uproszczona reprezentacja liganda (sztywne fragmenty połączone obracalnymi wiązaniami)
 - poszukiwanie konfiguracji kompleksów przy użyciu mechaniki molekularnej (Monte Carlo lub dynamika molekularna)
- Dokowanie białko-białko
 - obie cząsteczki zwykle nie zmieniają konformacji
 - 6 stopni swobody, 3 dla rotacji, 3 dla translacji
 - oszacowanie energii oddziaływania białek

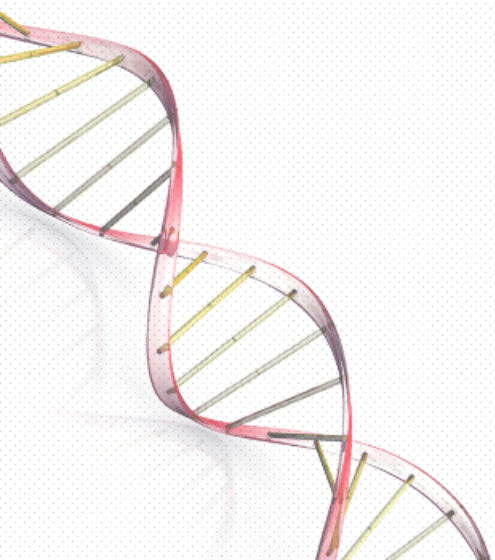






Dokowanie

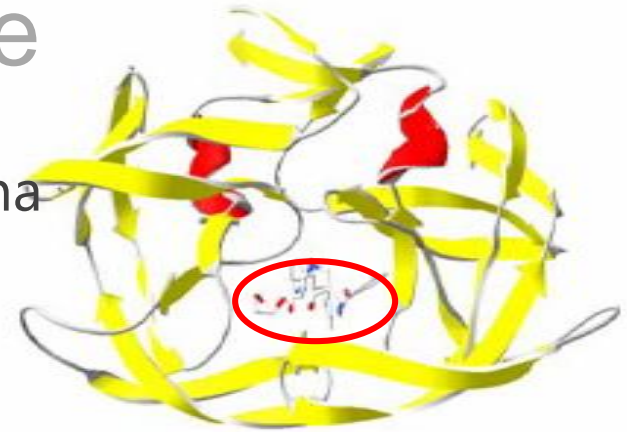
- Dopasowanie geometryczne:
 - zapobieganie nakładaniu atomów,
 - maksymalne dopasowanie konformacji liganda lub powierzchni pomiędzy ligandem i miejscem wiążącym.
- Dopasowanie chemiczne:
 - maksymalna liczba sprzyjających oddziaływań fizycznych (chemicznych), minimalna liczba oddziaływań niesprzyjających.



Dokowanie

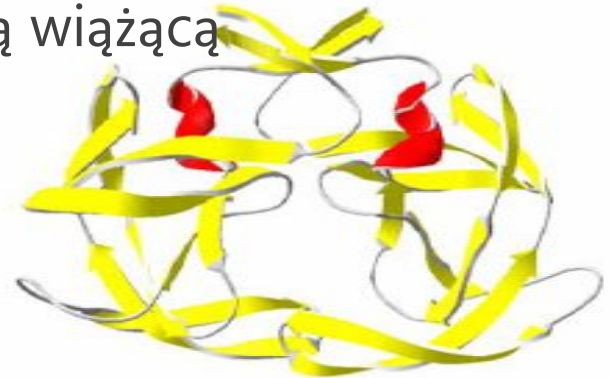
Struktura liganda w kieszeni wiążącej jest znana

- Jak inne ligandy oddziałują z białkiem ?
- Jak silne jest oddziaływanie białko-ligand?
- ⇒ Programy dokujące (głównie ruchome ligandy)



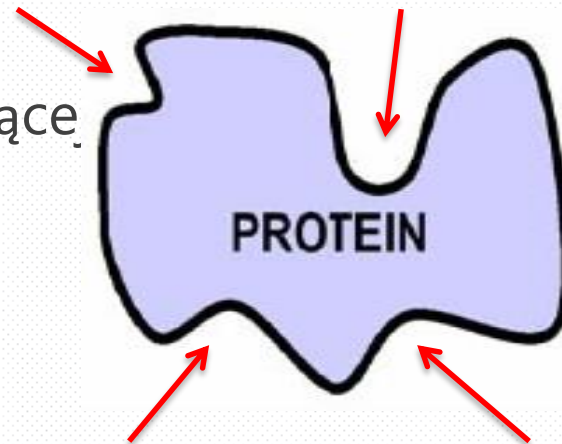
Znana jest tylko struktura białka z pustą kieszenią wiążącą

- W jaki sposób umieszczenie liganda
- wpływa na kieszeń wiążącą?
- ⇒ Programy dokujące, ruchome ligandy, optymalizacja struktury kompleksu



Nie jest znane położenie kieszeni wiążącej

- Poszukiwanie kieszeni wiążącej ligand
- Dokowanie wykonywane bez znajomości kieszeni wiążącej



Dokowanie białko-ligand

➤ Dokowanie lokalne:

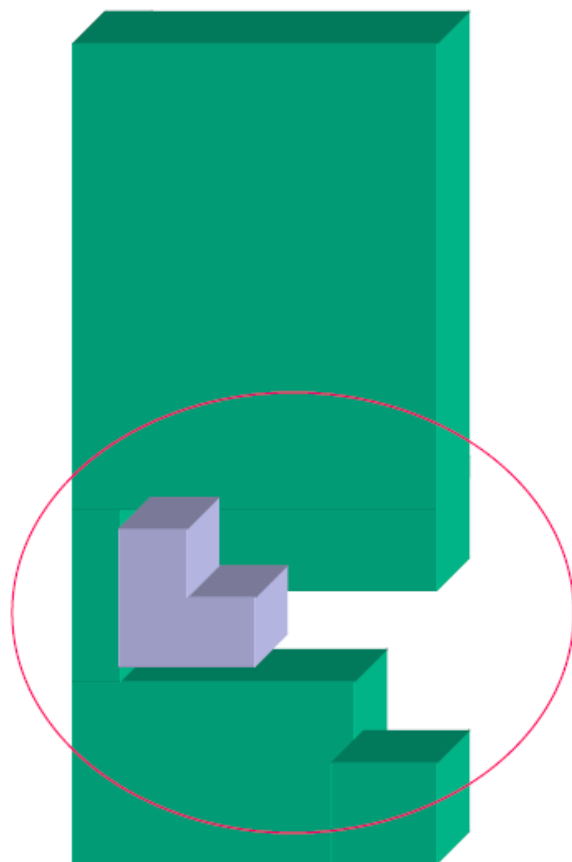
- Poszukiwanie przypuszczalnego miejsca przyłączenia liganda (jeśli nie jest ono wcześniej znane).
- Poszukiwanie jest skuteczne tylko jeśli znana jest konformacja białka w jakiej ono oddziałuje z ligandem
- Optymalne zorientowania liganda w miejscu wiążącym.

➤ Dokowanie globalne:

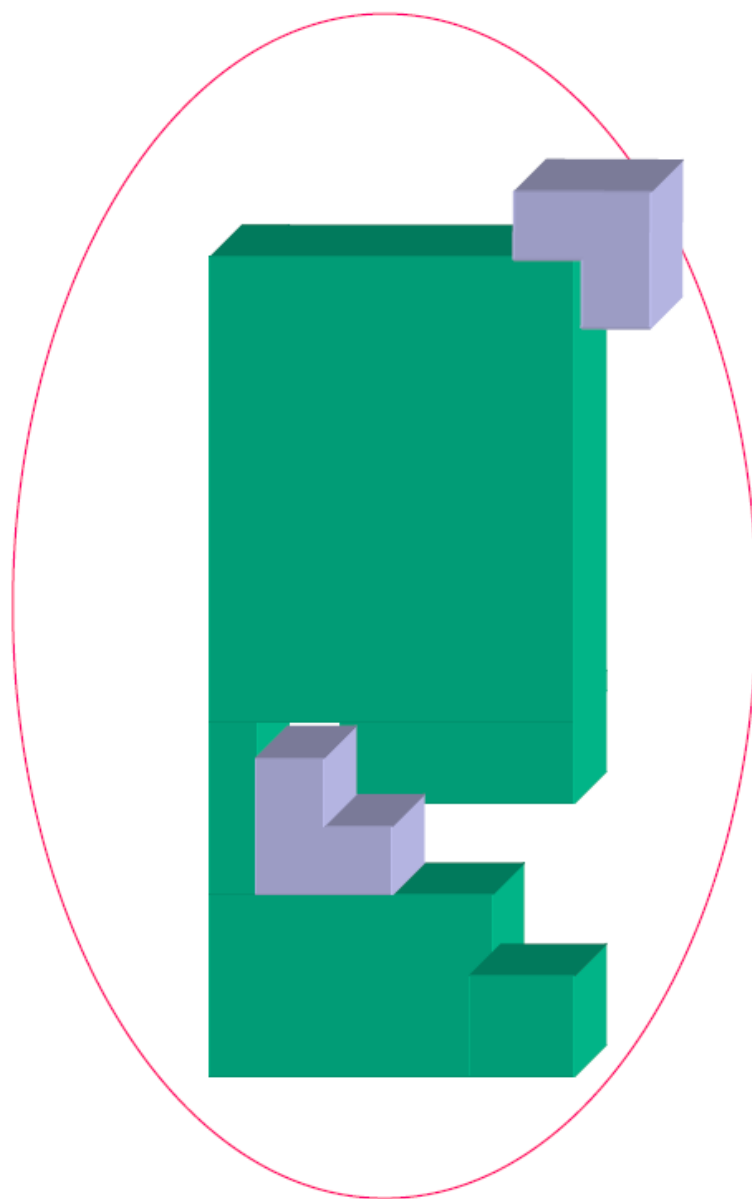
- Wykorzystanie zagłębień na powierzchni lub wolnych przestrzeni wewnątrz cząsteczki białka.
- Poszukiwanie orientacji liganda z uwzględnieniem oddziaływań w przypuszczalnym miejscu wiążącym.



Local docking

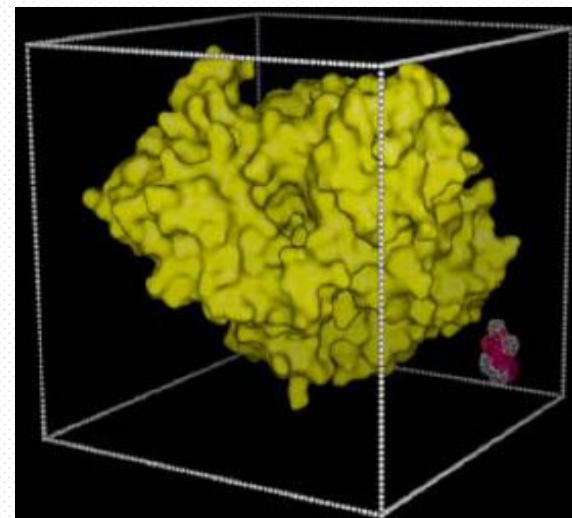
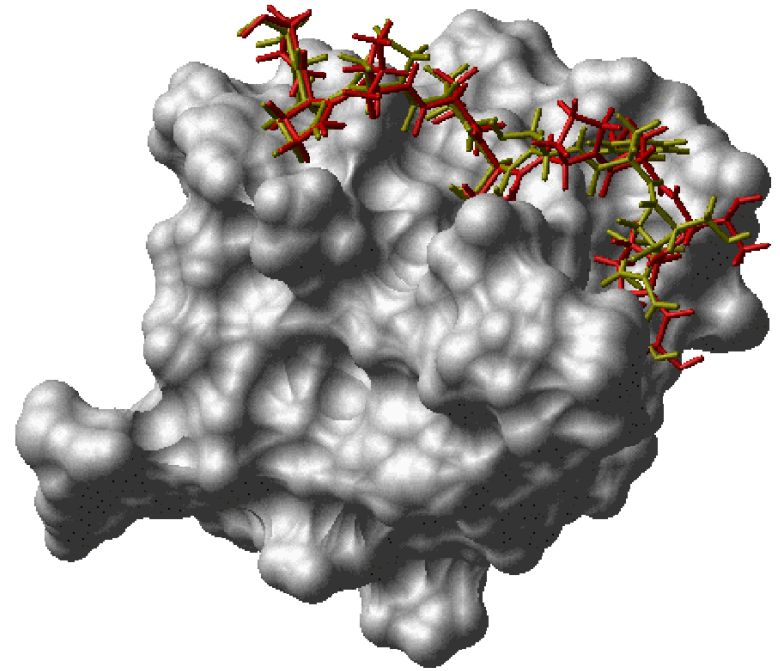


Global docking

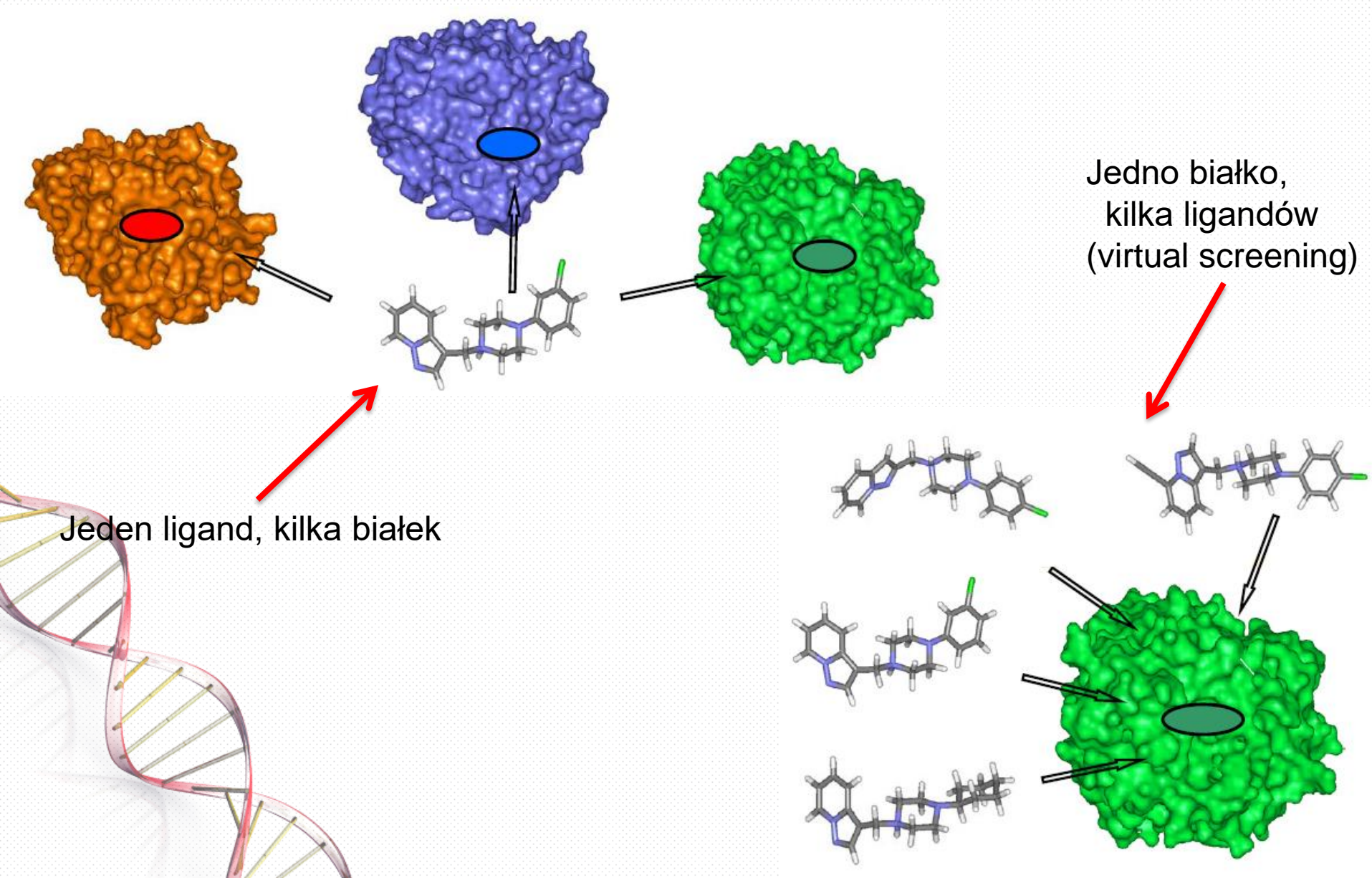


Miejsce przyłączenia liganda

- Gdzie znajduje się miejsce wiążące ligand?
- Jak znaleźć konformację liganda oddziałującą z receptorem?
- Jak zaprojektować optymalnego liganda?
- Co można zrobić gdy nie jest znana struktura receptora?
- Systematyczne poszukiwania są bardzo czasochłonne



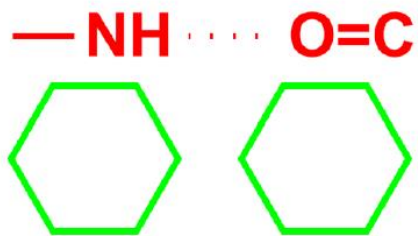
Selektywność liganda



Składowe (uproszczonej) funkcji energii wiązania liganda z receptorem



oddziaływania hydrofobowe



wiązania wodorowe

oddziaływania $\pi - \pi$



oddziaływania kation $-\pi$



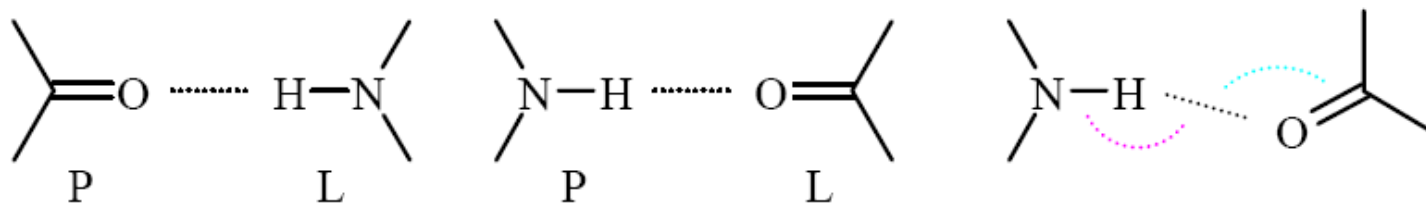
mostki solne



wiązania w kompleksie jon metalu – ligand

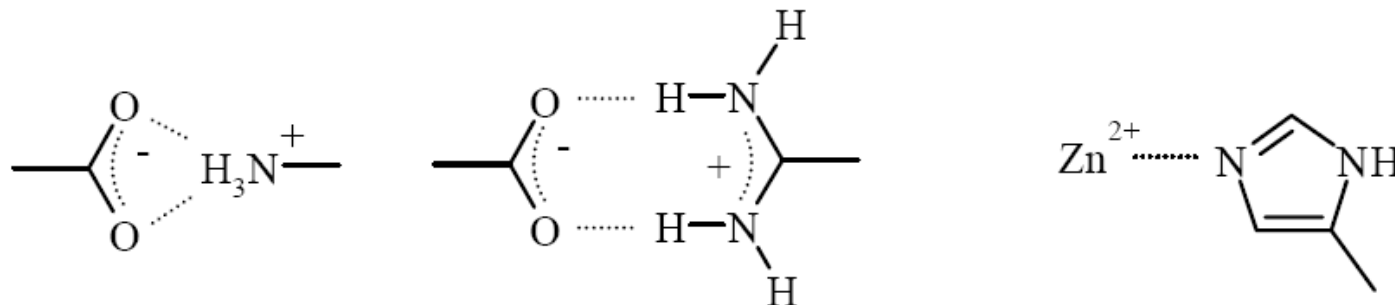
Składowe (uproszczonej) funkcji energii wiązania liganda z receptorem

wiązania wodorowe



(N-H...O: $>150^\circ$, C=O...H: $100-180^\circ$)

mostki solne




Metody poszukiwania optymalnej konfiguracji składników kompleksu

- Dynamika molekularna (MD)
- Metoda *Monte Carlo* (MC)
- Algorytmy genetyczne (GA)
- Metody oparte na rekonstrukcji liganda z fragmentów (FB)
- Lokalna komplementarność
- *Distance Geometry*
- Wykorzystanie metody *Tabu Search*
- Poszukiwanie systematyczne
- Łączenie wielu algorytmów
- Wirtualne „przesiewanie” (*ang. Virtual screening*)



Dynamika molekularna

- Wykorzystanie pełnych pól siłowych Amber lub Charmm
- Metoda była wykorzystywana [1-6] do dokowania ruchomego liganda w ruchomym modelu białka
- Trudno wyprowadzić ligand z minimum lokalnego
- Może być łączona z algorytmami umożliwiającymi wyjście z minimum lokalnego

- 
1. Di Nola, A., Roccatano, D. and Berendsen, H.J.C. Proteins, 19 (1994) 174–182.
 2. Nakajima, N., Higo, J., Kidera, A. and Nakamura, H. Chem. Phys. Lett., 278 (1997) 297–301
 3. Mangoni, R., Roccatano, D. and Di Nola, A. Proteins, 35 (1999) 153–162.
 4. Pak, Y. and Wang, S. J. Phys. Chem. B, 104 (2000) 354–359.
 5. Carlson, H.A. and McCammon, J.A. Mol. Pharmacol., 57 (2000) 213–218.
 6. Wei, B. Q., Weaver L. H., Ferrari A. M., Matthews B. W. and Shoichet B. K., J. Mol. Biol., 337(2004), 1161-1182.

Monte Carlo

- Może być użyta prostsza [*](niż w symulacjach MD) postać funkcji energii.
- Łatwiejsze przechodzenie przez bariery energii.
- Losowe przejścia pomiędzy konfiguracjami, akceptowanie lub odrzucenie konfiguracji na podstawie rozkładu Boltzmann.



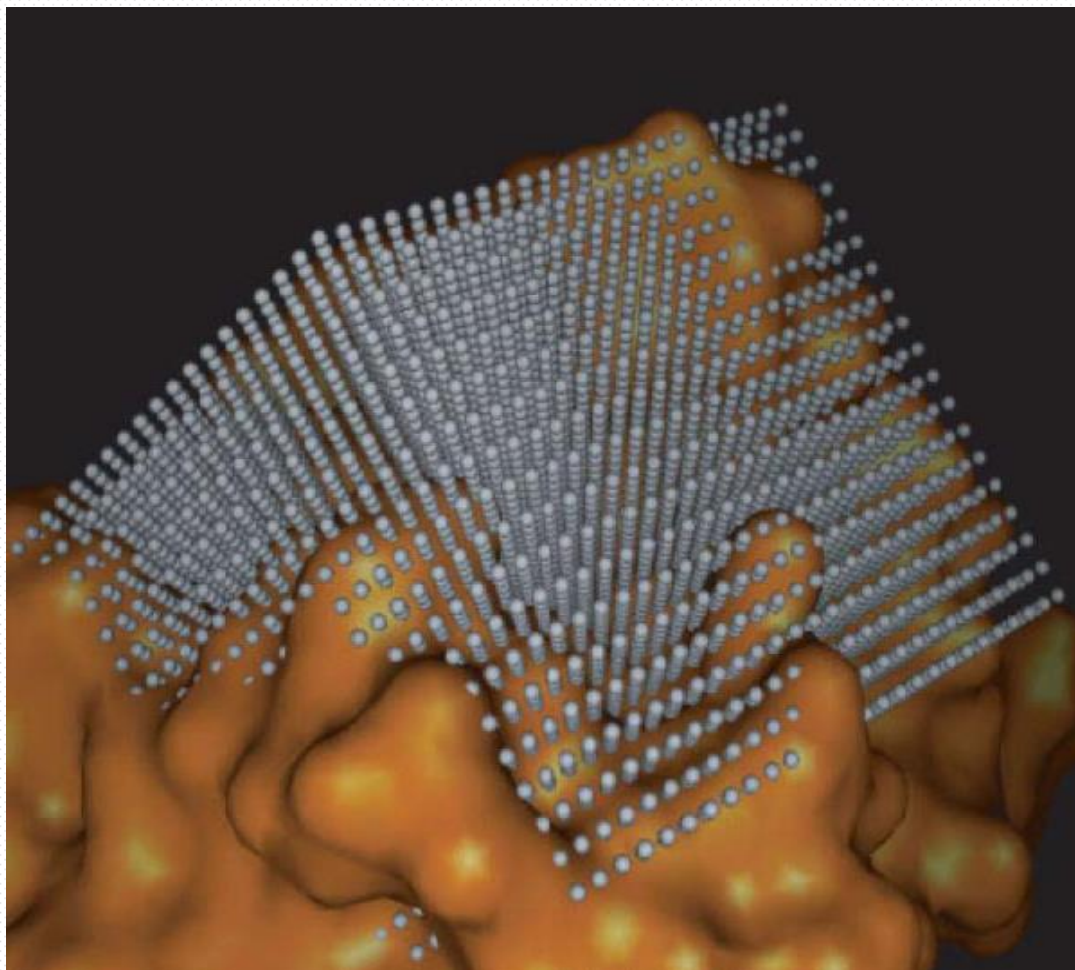
*David R. Westhead, David E. Clark, Christopher W. Murray: A comparison of heuristic search algorithms for molecular docking. Journal of Computer-Aided Molecular Design 11(3): 209-228 (1997)

AutoDock 2.4 (MC)

- Ruchomy ligand dokowany do nieruchomego receptora[*].
- Posługuje się polem sił Amber.
- Odtworzył 6 testowanych kompleksów.
- Najniżej energetyczne struktury nie zawsze odpowiadają konformacji znalezionej w badaniach krystalograficznych.

*Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998), Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function *J. Computational Chemistry*, **19**: 1639-1662.

AutoDock 2.4 (MC)



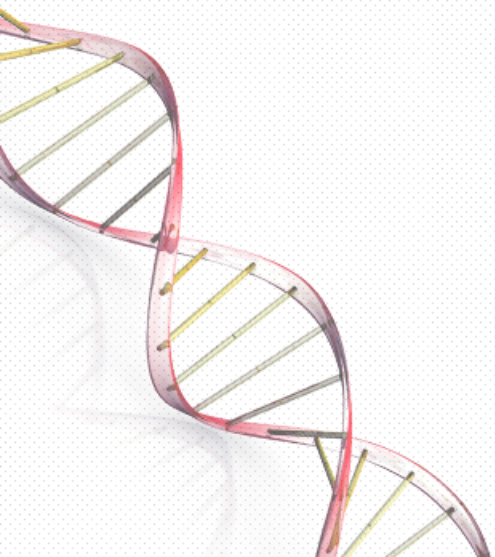
*Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998), Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function *J. Computational Chemistry*, **19**: 1639-1662.

Prodock (MC)

- Ruchomy ligand dokowany do ruchomego receptora.
- Po każdym losowym wyborze lokalnej zmiany konformacji wykonywana jest gradientowa optymalizacja struktury.
- Posługuje się polami sił ECEPP/3 lub Amber.
- Program MC Caflisch, podobny do programu Prodock, posługuje się polem sił CHARMM.

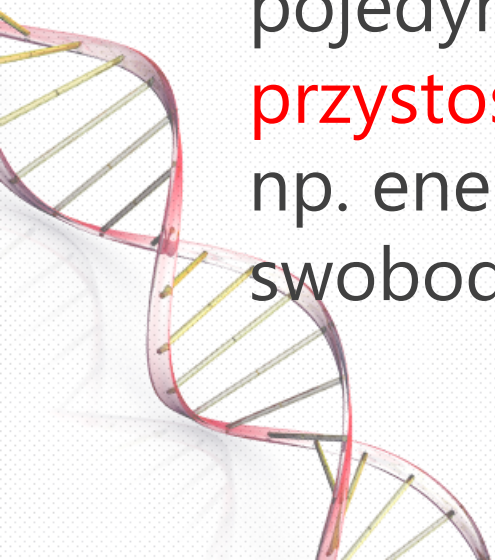
Inne programy oparte na MC

- ICM
- MCDOCK
- DockVision
- QXP
- Affinity
- Glide



Algorytmy Genetyczne

- Metody te posługują się pojęciami zaczerpniętymi z genetyki
- Wymagają początkowego **zbioru** konformacji (**populacji** początkowej)
 - metody MC i MD wymagają tylko jednej konfiguracji początkowej kompleksu.
- Stopnie swobody są zakodowane w „genotypie” pojedynczego przedstawiciela, **funkcja przystosowania** opiera się na **funkcji oceniającej** np. energii, entalpii swobodnej lub energii swobodnej.



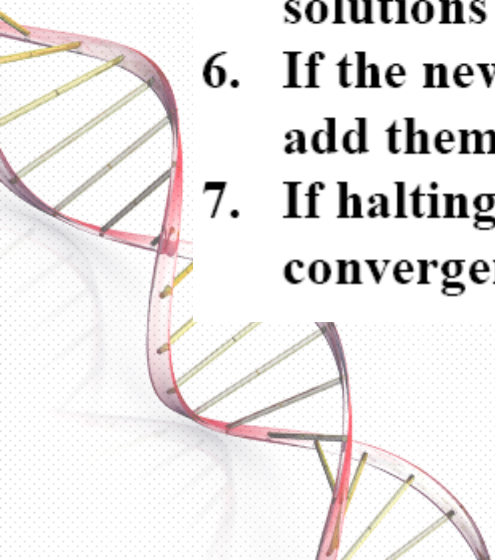
Algorytmy genetyczne

- Stosowanie **operatorów genetycznych**, specyficznych dla poszukiwanej postaci rozwiązań.
- Zmiany właściwości **wielu** przedstawicieli populacji w celu **równoległego** przeszukiwania przestrzeni rozwiązań.
- **Ukierunkowanie** procesu przeszukiwania albo na podstawie jakości aktualnych rozwiązań, albo z wykorzystaniem dodatkowych informacji o układzie.
- Celowe (losowe) zaburzanie nieoptymalnych rozwiązań.



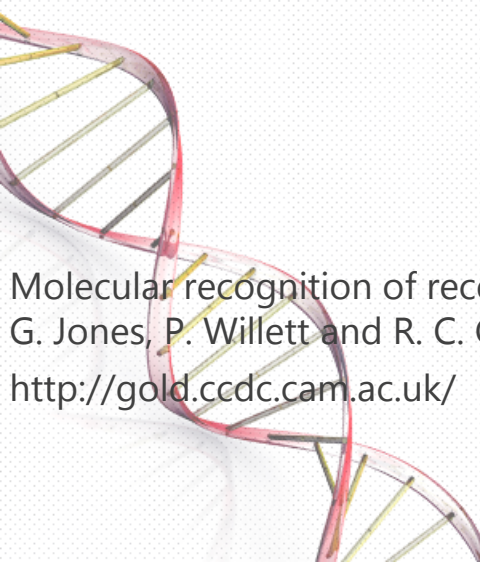
Genetic algorithms

1. For each operation choose a weight
2. Create initial population of solutions. Evaluate each one with the scoring function
3. Using a Rolette wheel and the weights of 1 choose an operation
4. Choose parents from the current generation
5. According to the parents and the operation create new solutions and evaluate them
6. If the new solutions are not represented in the population add them and remove the least fitted solutions
7. If halting condition is fulfilled (# of generation, convergence of scoring function) – stop. Otherwise go to 3



Program GOLD (GA)

- Ruchomy ligand, częściowo ruchomy receptor.
- Skuteczność 72% poprawnie zadokowanych ligandów.
- Na 100 testowanych w oryginalnej publikacji kompleksów 66 miało RMSD 2.0Å lub mniej.
- 72 miało RMSD 3.0Å lub mniej.

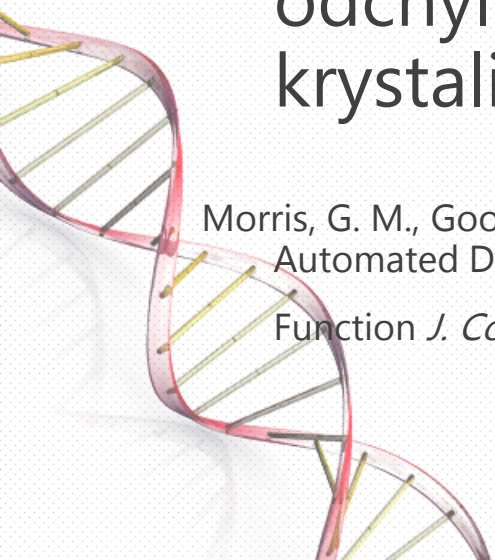


Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation
G. Jones, P. Willett and R. C. Glen, *J. Mol. Biol.*, **245**, 43-53, 1995
<http://gold.ccdc.cam.ac.uk/>

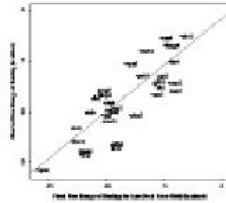
Autodock 3.0 (GA)

- Algorytm genetyczny jest używany do optymalizacji globalnej.
- Dla każdej nowoutworzonej populacji ułamek konfiguracji poddawany jest przeszukiwaniu lokalnemu.
- Testowany na 7 kompleksach. Wszystkie niskoenergetyczne konfiguracje miały wartość odchylenia RMSD 1.14Å od struktury krystalicznej.

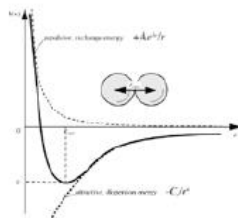
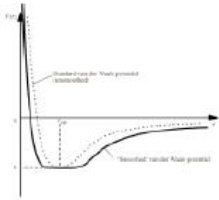
Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998), Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function *J. Computational Chemistry*, **19**: 1639-1662.



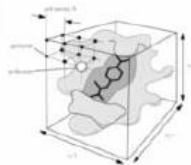
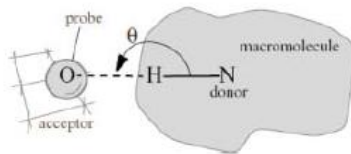
Autodock 3.0 (GA)



$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$



$$\epsilon(r) = A + \frac{B}{1 + ke^{-\lambda Br}}$$



• ΔG_{vdW}

Składnik 12-6 Lennarda-Jonesa

• ΔG_{elec}

Składnik elektrostatyczny ze zmienną przenikalnością dielektryczną

• ΔG_{hbond}

Składnik 10-12 z uwzględnieniem orientacji wiązań chemicznych

• ΔG_{desolv}

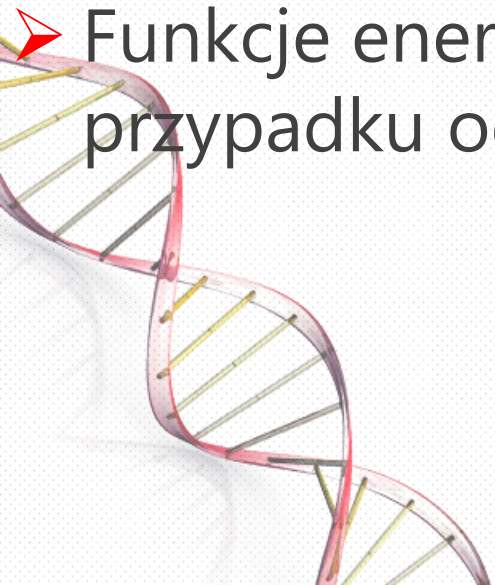
Składnik uwzględniający solwatację

• ΔG_{tors}



Składnik uwzględniający oddziaływania torsyjne

Dokowanie białko-białko

- Podejścia do dokowania są podobne jak dla dokowania białko-ligand.
- Metody opierają się na przeszukiwaniu konformacji (orientacji) i ocenianiu znalezionych konfiguracji za pomocą funkcji energii.
- Liczba stopni swobody jest większa.
- Funkcje energii są podobne (lub te same) jak w przypadku oddziaływań białko-ligand.



Dokowanie białko-białko



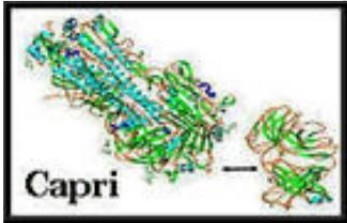
All Databases

Go

Give us feedback

DatabasesToolsEBI GroupsTrainingIndustryAbout UsHelpSite IndexRSSPrint

contact PDBe



- PDB idcodes for past targets
- Call For Targets
- Capri Rules 2007
- Original Capri Rules 2001
- Management
- Formats

Home > Databases > PDBe > Services > Capri-Home

CAPRI: Critical Assessment of PRediction of Interactions

CAPRI communitywide experiment on the comparative evaluation of protein-protein docking for structure prediction

Hosted By EMBL/EBI-PDBe Group

Fourth CAPRI Evaluation Meeting, Barcelona Dec. 9-11, 2009

The 4th CAPRI meeting will be organised in Barcelona (9-11 Dec 09) . The program and registration details can be found on the following web page:
<http://mmb.pcb.ub.es/capri2009/>



<http://www.ebi.ac.uk/msd-srv/capri/>

Dokowanie białko-białko

HADDOCK

v2.0

Main

Introduction

Project setup

Download

Installation

Publications

Software links

Manual

PDB files

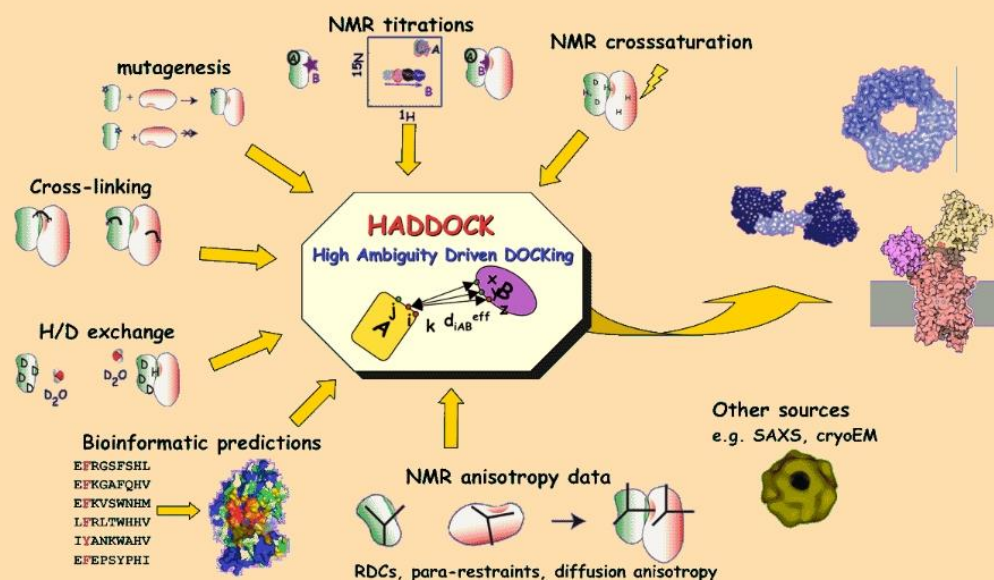
AIR restraints

RDC restraints

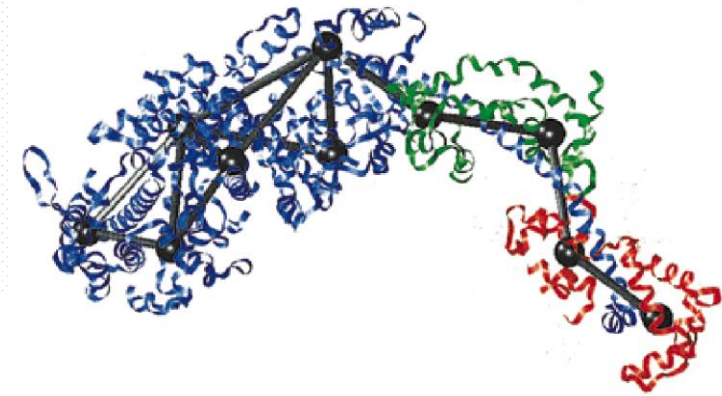
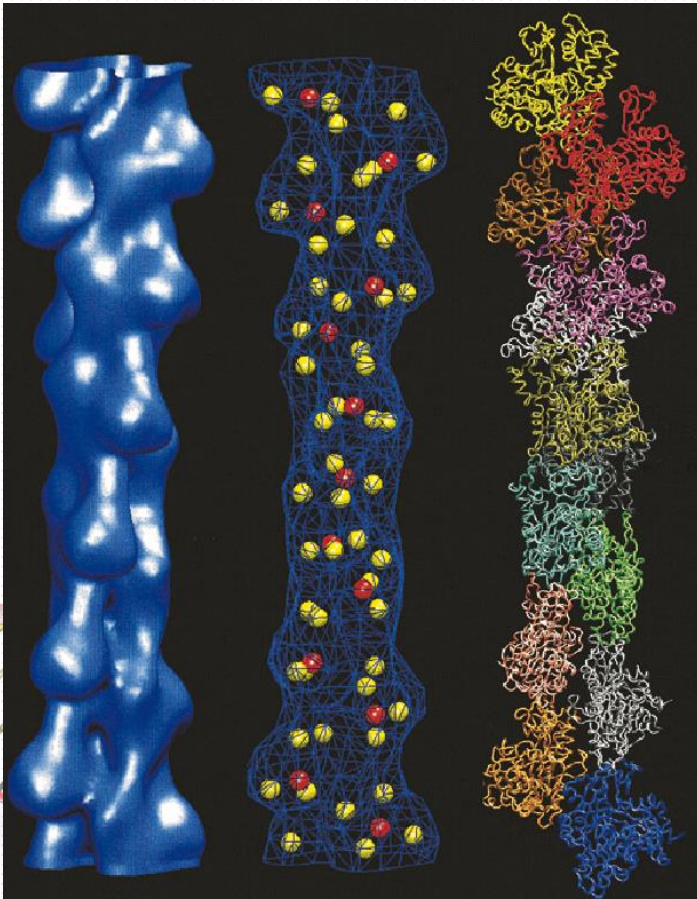
HADDOCK

High Ambiguity Driven biomolecular DOCKing


based on biochemical and/or biophysical information.



Dokowanie białko-białko



Wriggers, W., Milligan, R.A., Schulten, K., and McCammon, J.A., Self-organizing Neural Networks Bridge the Biomolecular Resolution Gap, *J. Mol. Biol.* (1998) 284, 1247-1254



Docking approach	Examples
Matching of descriptors	DOCK, QSDOCK, SLIDE
Incremental construction	FlexX, Hammerhead
Monte Carlo Simulated Annealing	AutoDock, MCDOCK
Monte Carlo Minimization	ICM, QXP
Molecular Dynamics	MDD
Genetic Algorithms	GOLD, AutoDock3

Pytania i odpowiedzi na zadania

m.mozolewska@ipipan.waw.pl

Zdjęcia, schematy i rysunki zostały zaczerpnięte z internetu.



Zadania (maksymalnie 10 punktów)

1. Przewidywanie struktury:
 - a) Dla podanej sekwencji przewidź białka homologiczne (bio3d)
 - i. MLPVITRFARPALMAIRPVNAMGVLRASSITKRLYHPKVIEHYTHPRNVGSLDKK
LPNVGTGLVGAPACGDVMRLQIKVNDSTGVIEDVKFKTFGCGSAIASSSYMTEL
VQGMTLDDAAKIKNTEIAKELSLPPVKLHCSMLAEDAIAKAIKDYKSKRNTPTML
S
 - b) Wrysować histogram RMSD (bio3d)
 - c) Wrysować dendogram RMSD (bio3d)
 - d) Wrysować heat mapę podobieństwa za pomocą pakietu AlignStat
2. Analiza kompleksu:
 - a) Zidentyfikuje za pomocą pakietu bio3d miejsce wiążące dla białka 4q21
3. Dla białka o kodzie PDB 5UOO:
 - a) Zidentyfikuj reszty aminokwasowe biorące udział w oddziaływaniu z cząsteczką CD161
 - b) Policz odległości pomiędzy zidentyfikowanymi resztami a resztami liganda
 - c) Na podstawie punktu b wykreśl heat mapę
4. Znajdź inne pakiety R do analizy struktury kompleksu białek obejmujące punkt 3 oraz dodatkowe funkcje jak: znalezienie reszt oddziałujących pomiędzy białkami, identyfikacja siły wiązania pomiędzy AA pomiędzy białkami.

