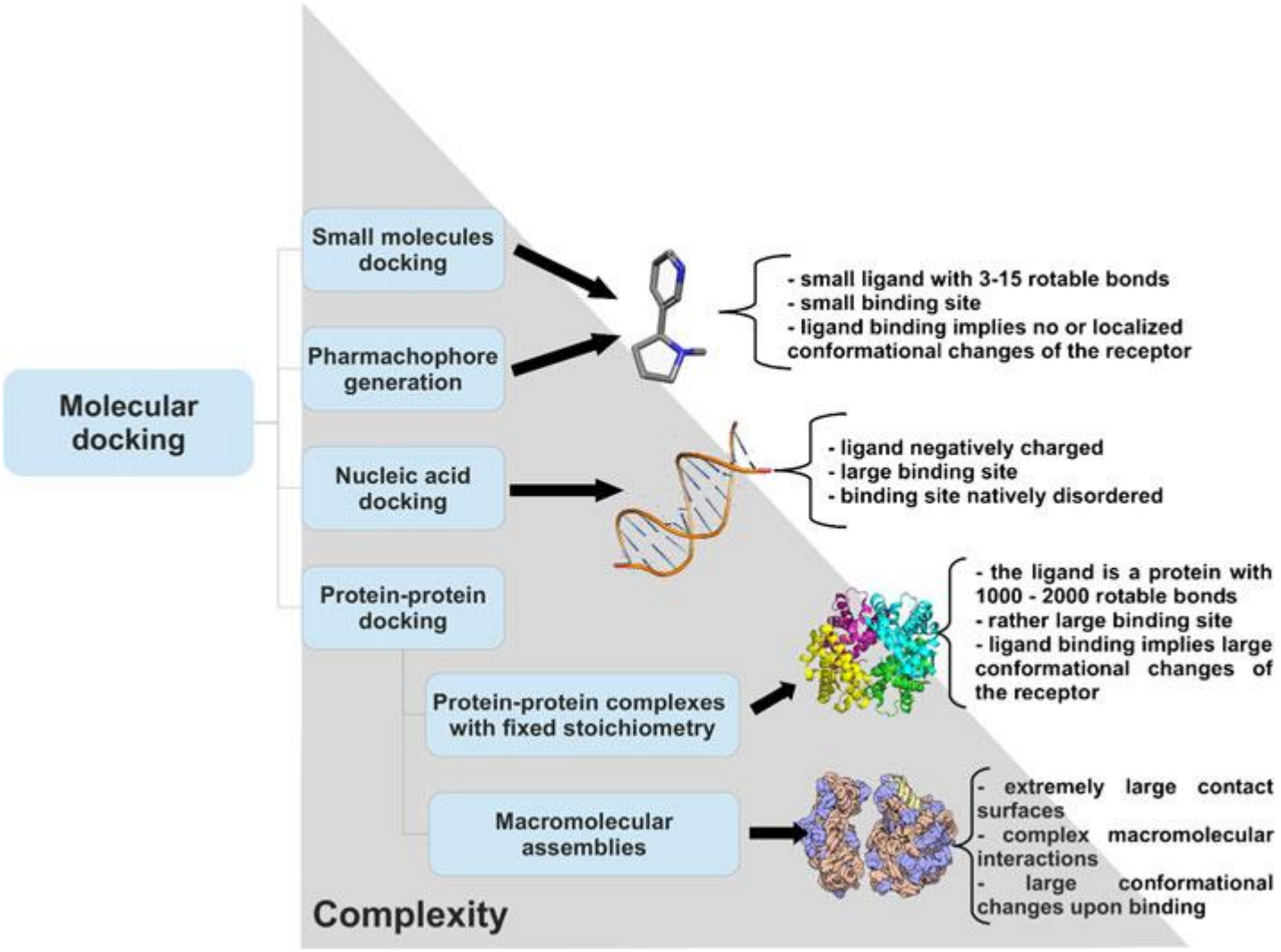
Molekulska umestitev

„Docking“ – računalniške metode za napovedovanje struktur kompleksov med molekulami z znanimi strukturami.



Mihasan. 2012. *J. Biosci*. 37(6)

Molekulska umestitev

„Docking“ – računalniške metode za napovedovanje struktur kompleksov med molekulami z znanimi strukturami.

1. Makromolekulska umestitev – protein-protein, protein-DNA
2. Umestitev malega liganda na/v makromolekulo

1. Umestitev togih teles („rigid-body docking“)
2. Umestitev gibljivih teles („flexible-body docking“)

1. *Ab initio* umestitev
2. Usmerjena umestitev (z uporabo eksperimentalnih podatkov)

Zgodovina: kot začetnike računalniške molekulske umestitve štejemo Levinthala in sod. (1975), ki so zgradili model interakcij med molekulami hemoglobina pri anemiji srpastih celic.

What is Docking?

*"Predicting the **best** ways two molecules will interact."*

- We need to *quantify* or *rank* solutions;
- We need a ***Scoring Function*** or force field.

*"Predicting the best **ways two molecules will interact**."*

- *(ways—plural)* The experimentally observed structure may be amongst one of *several predicted solutions*.
- We need a ***Search Method***.

algoritem za iskanje rešitev + funkcija za ovrednotenje rešitev

Molekulska umestitev – algoritmi za iskanje

TABLE 1

Search strategies in protein–protein docking

Search algorithms	Examples of docking programs
Exhaustive global search	
FFT-based search	FTDock, GRAMM, DOT, ZDOCK, MolFit, PIPER, F2DOCK, SDOCK, ASPDock, Cell-Dock
Spherical Fourier transform-based search	HEX, FRODOCK
Direct search in Cartesian space	SOFTDOCK, BIGGER, SKE-DOCK
Local shape feature matching	
Distance geometry algorithm	DOCK
Geometric hashing	PatchDock, SymmDock, LZerD
Genetic algorithm	GAPDOCK
Randomized search	
Monte Carlo search	RosettaDock, ICM, ATTRACT, HADDOCK
Particle swarm optimization	SwarmDock
Genetic algorithm	AutoDock
Post-docking approach	
Using advanced scoring functions	RPScore, ZRANK, PyDock, EMPIRE, DARS, DECK, SIPPER, PIE, MDockPP, etc.
Considering protein flexibility	MultiDock, SmoothDock, RDOCK, FireDock, FiberDock, EigenHex, etc.
Other ranking protocols	SDU, CyClus, CONSRANK, etc.

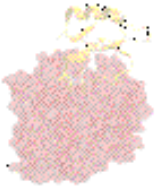
TABLE 2

Features of search algorithms in protein–protein docking

Search algorithms	Exhaustive search	Global search	Local search	Rigid docking	Flexible docking	Molecular representation	Computational cost
FFT-based correlation	√	√		√		Grid based	Low
SFT-based search	√	√		√		Harmonic surface	Low
Direct global search	√	√	√	√		Grid based	Medium–high
Local shape matching		√	√	√		Grid or surface	Medium
Randomized search		√	√	√	√	Atom based	High

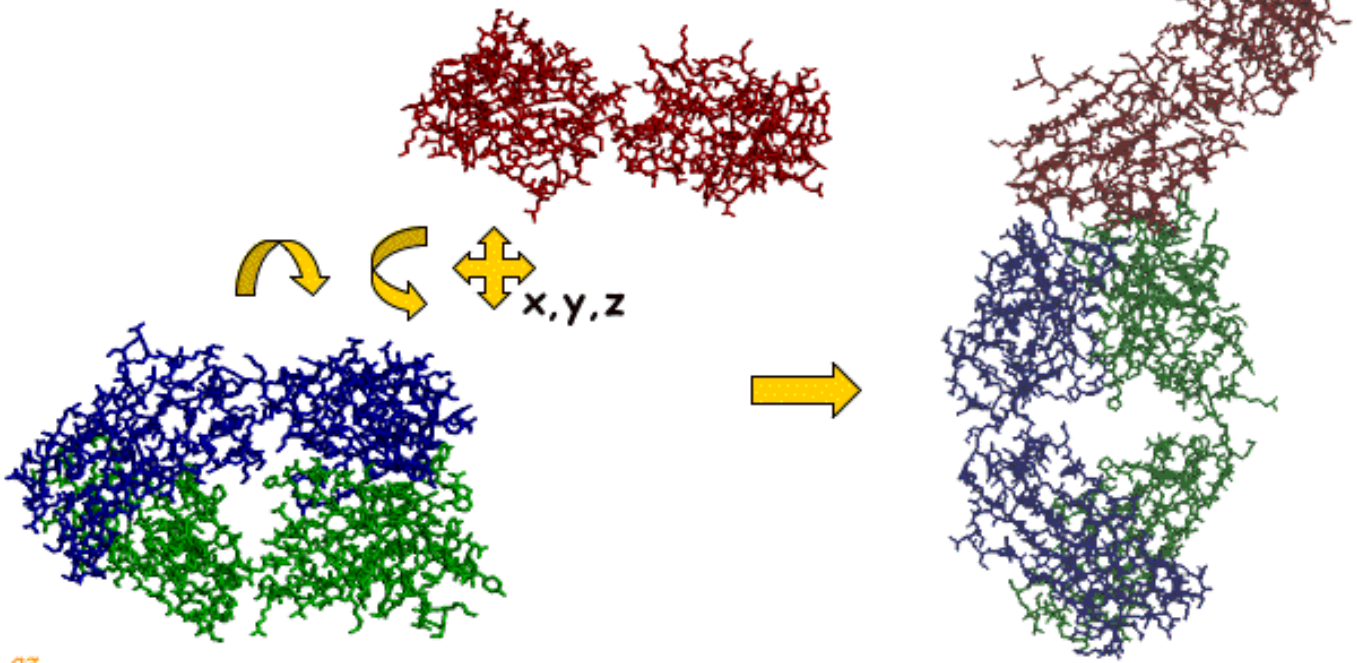
Huang. Drug Discov Today. (2014)

Molekulska umestitev – algoritmi za iskanje



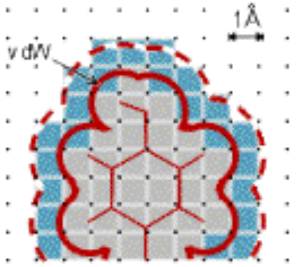
Explicit representation of the system

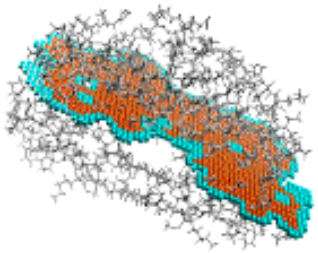
- x,y,z, coordinates of each atom for both molecules
- Search method will be in real space

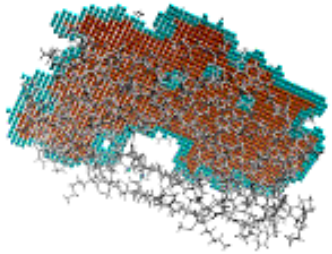


Grid-based representation of the system

- Discretise of the 3D structure of a protein onto a grid







(source: bigger/ Krippahl)

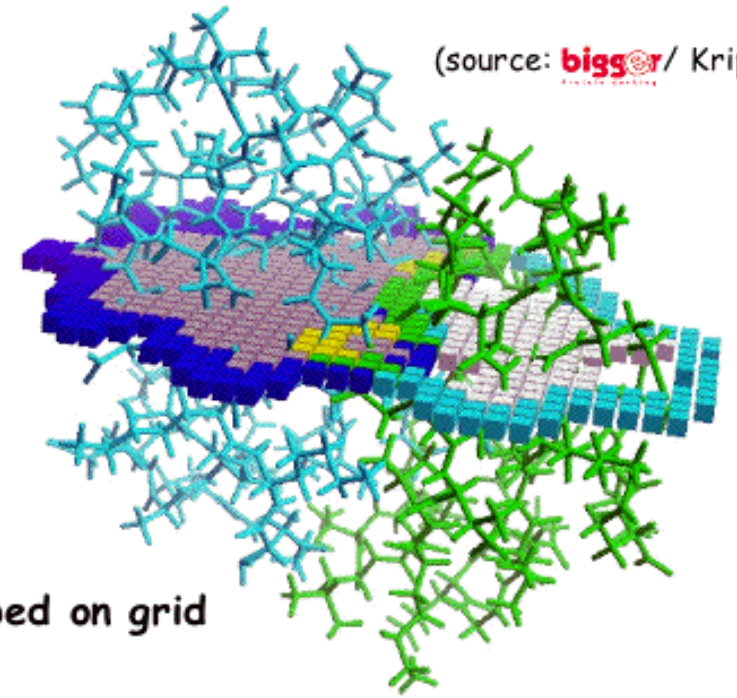
- "Shape representation" of the protein
- Resolution defined by grid spacing
- Docking will require to match the shapes ("geometric matching")
- Search in real or Fourier space

Grid-based representation of the system

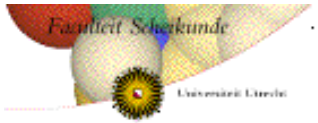
- Matching:
  - Surface overlap -> bonus
  - Core overlap -> penalty

Core
Surface
Core overlap

(source: biggar/ Krippahl)



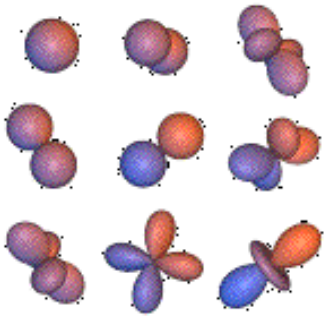
- Properties (e.g. charges, hydrophobicity) can be mapped on grid

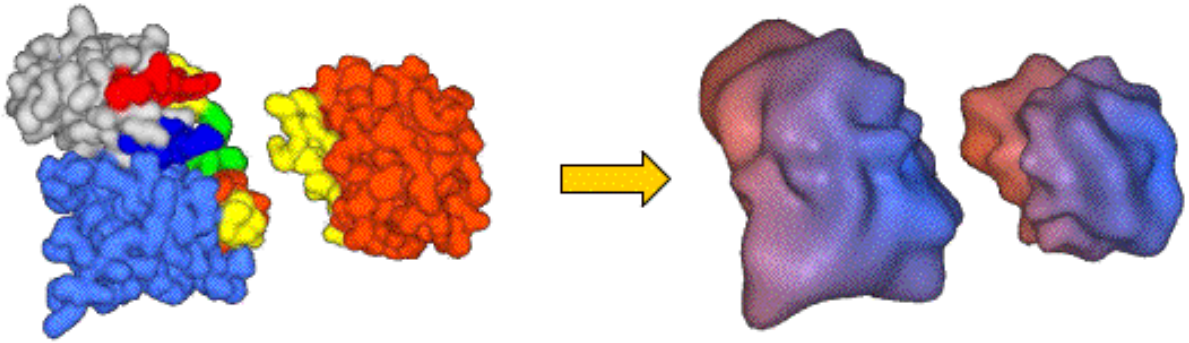


Surface representation of the system: spherical harmonics

- Surface of protein described by an expansion of spherical harmonics, e.g.

$$r(\theta,\phi) = \sum_{l=0}^{15} \sum_{m=-1}^{1} a_{lm} \psi_{lm}(\theta,\phi)$$

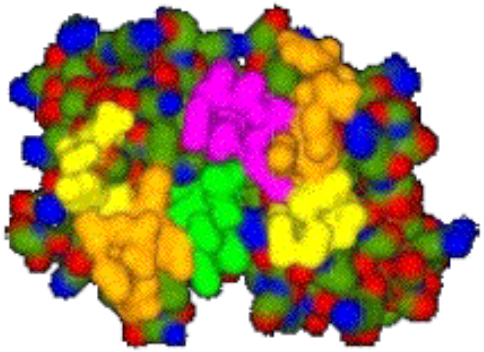


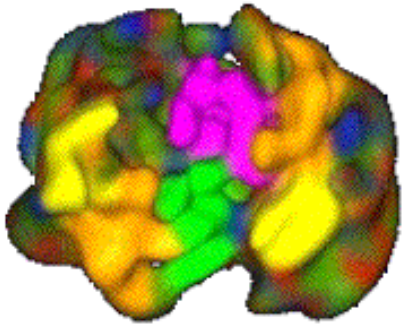


(source: HEX / Richie)

Surface representation of the system: spherical harmonics

- By varying the number of terms in the expansion the resolution can be tuned



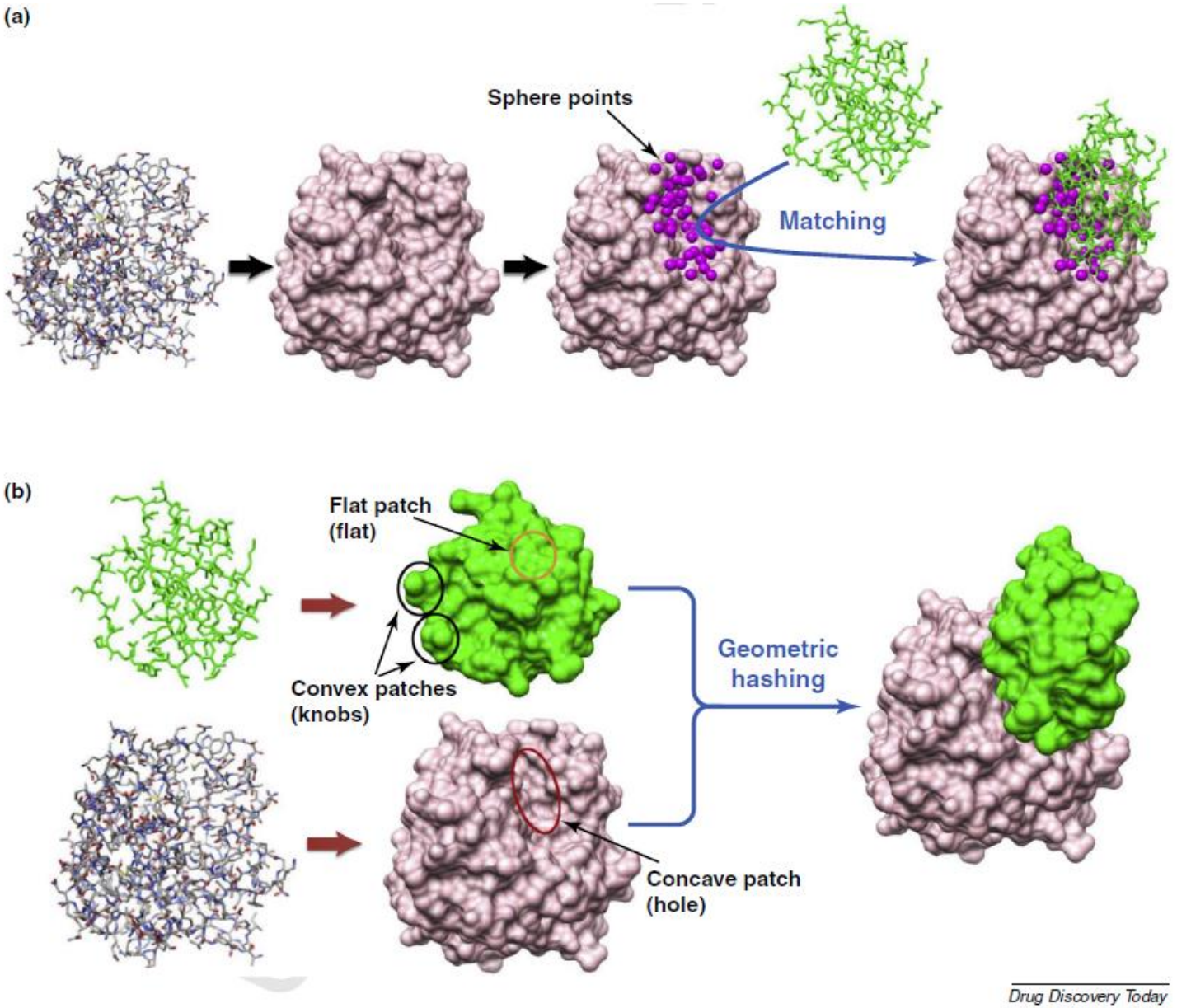


(source: HEX / Richie)

Molekulska umestitev – algoritmi za iskanje

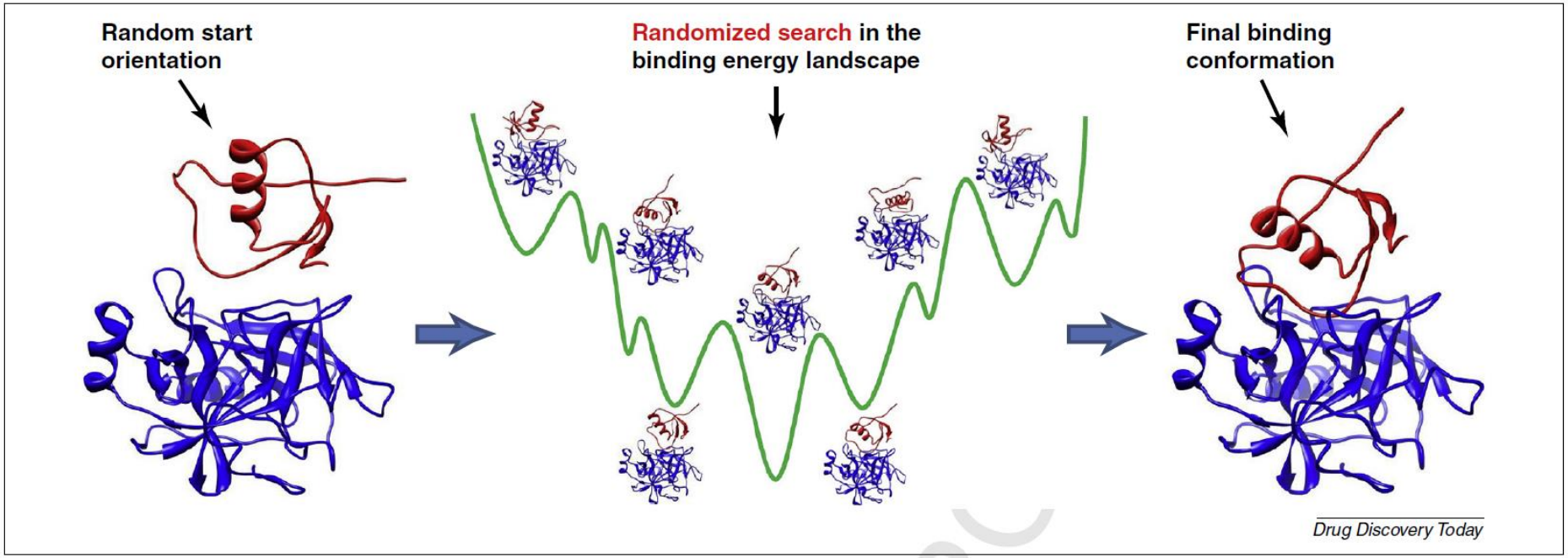
sphere matching

geometric hashing



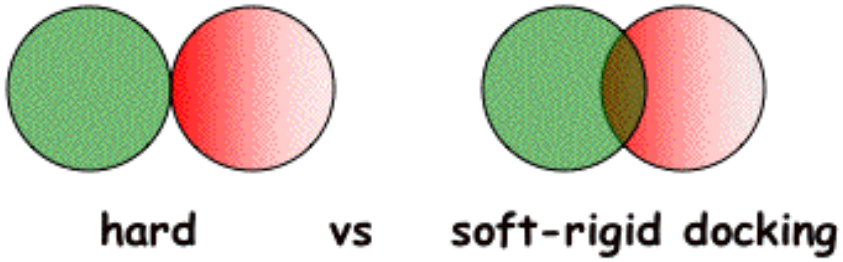
Molekulska umestitev – algoritmi za iskanje

naključno iskanje



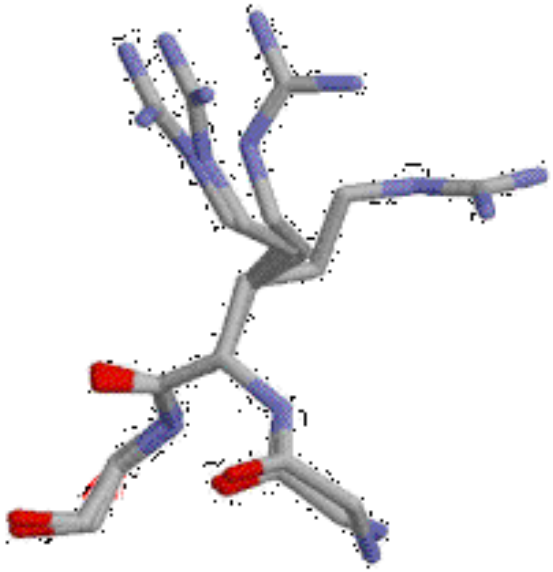
Molekulska umestitev – fleksibilnost

1. Implicitna fleksibilnost – umestitev togih teles s prekrivanjem

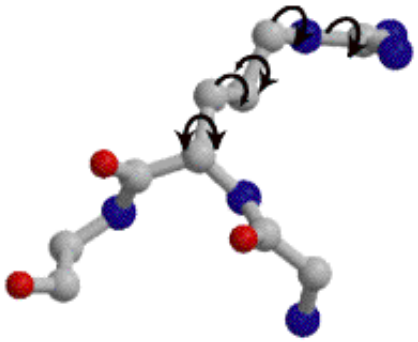


Rešitve je potrebno optimirati in odstraniti trke

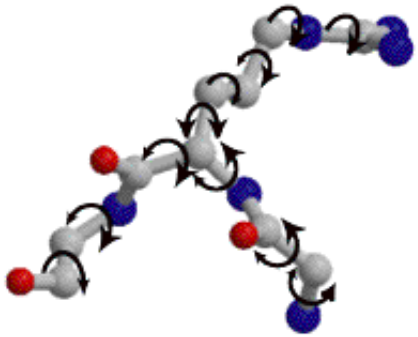
2. Zbirka energijsko ugodnih konformacij



3. Eksplicitna fleksibilnost

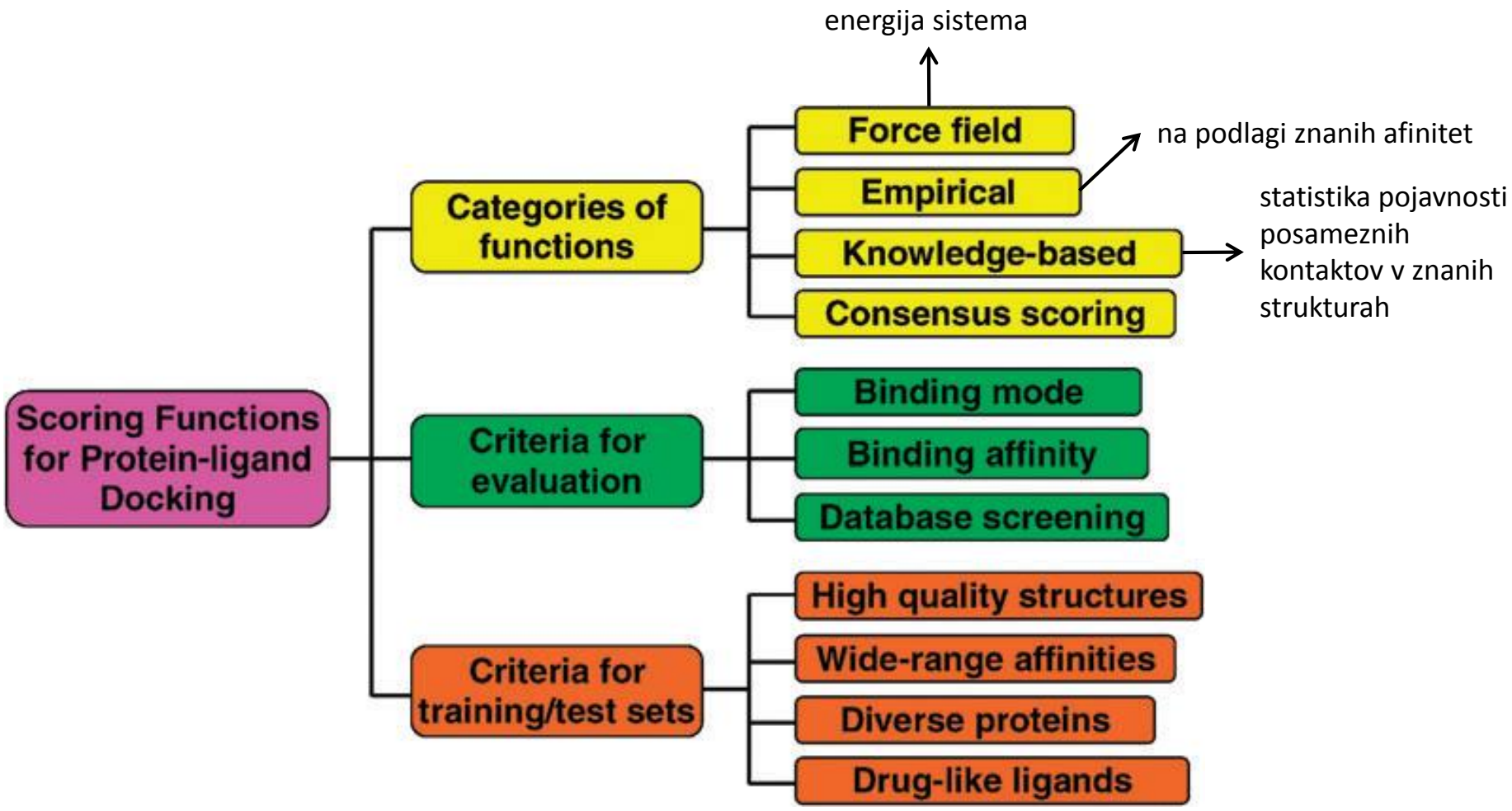


samo stranska veriga



stranska in glavna veriga

Molekulska umestitev – funkcije za vrednotenje



Molekulska umestitev – funkcije za vrednotenje

Examples of scoring function formulae

Scoring function formulae

$$V = W_{vdw}\sum_{i,j}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond}\sum_{i,j}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec}\sum_{i,j}\frac{q_i q_j}{\epsilon(r_{ij})r_{ij}} + W_{sol}\sum_{i,j}\left(S_i V_j + S_j V_i\right)e^{\left(-r_{ij}^2/2\sigma^2\right)}$$

Extended force-field-based scoring function from AutoDock.

For two atoms i, j, the pair-wise atomic energy is evaluated by the sum of van der Waals, hydrogen bond, coulomb energy and desolvation. W are weighted factors for calibrate the empirical free energy.

$$\Delta G = \Delta G_0 + \Delta G_{rot} \times N_{rot} + \Delta G_{hb}\sum_{neutral\ H-bond} f(\Delta R, \Delta a) + \Delta G_{io}\sum_{ion\ init.} f(\Delta R, \Delta a) + \Delta G_{aro}\sum_{aro\ int.} f(\Delta R, \Delta a) + \Delta G_{lipo}{}_{lipo}$$

Empirical scoring function from FlexX.

ΔG is the estimated free energy of binding; ΔG_0 is the regression constant; ΔG_{rot} , ΔG_{hb} , ΔG_{io} , ΔG_{aro} and ΔG_{lipo} are regression coefficients for each corresponding free energy term; $f(\Delta R,\Delta\alpha)$ is scaling function penalizing deviations from the ideal geometry; N_{rot} is the number of free rotate bonds that are immobilized in the complex.

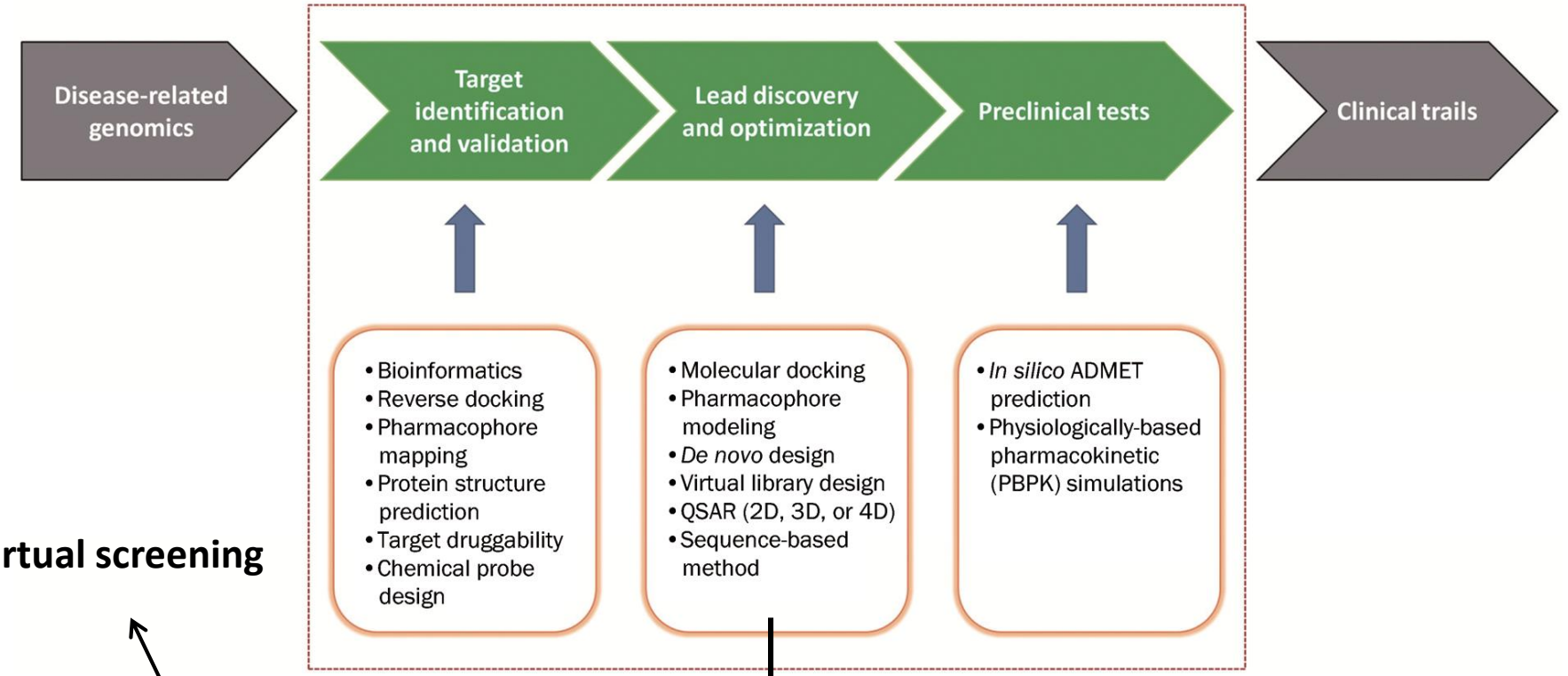
$$PM\_score = \sum_{\substack{kl \\ r<r_{cut-off}^{ij}}} A_{ij}(r) \qquad A_{ij}(r) = -k_B T \ln\left[f^{j}_{Vol\_corr}(r)\frac{\rho_{seg}^{ij}(r)}{\rho_{bulk}^{ij}}\right]$$

Knowledge-based scoring functions PMF.

k_B is the Boltzmann constant; T is the absolute temperature; r is the atom pair distance. $f^{j}_{Vol\_corr}(r)$ is the ligand volume correction factor; $\frac{\rho_{seg}^{ij}(r)}{\rho_{bulk}^{ij}}$ designates the radial distribution function of a protein atom of type i and a ligand atom of type j.

Umestitev majhnih molekul

Se najpogosteje uporablja pri iskanju novih učinkovin.

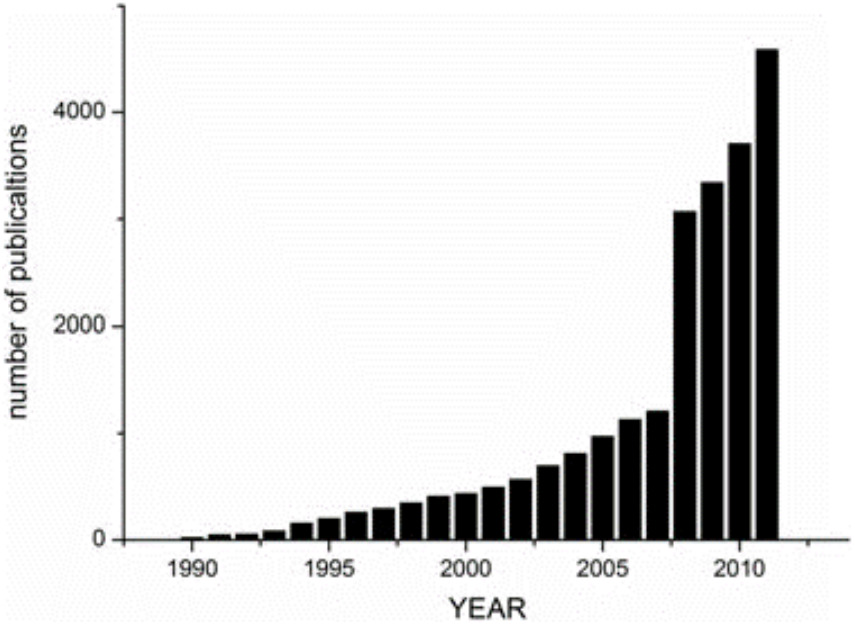


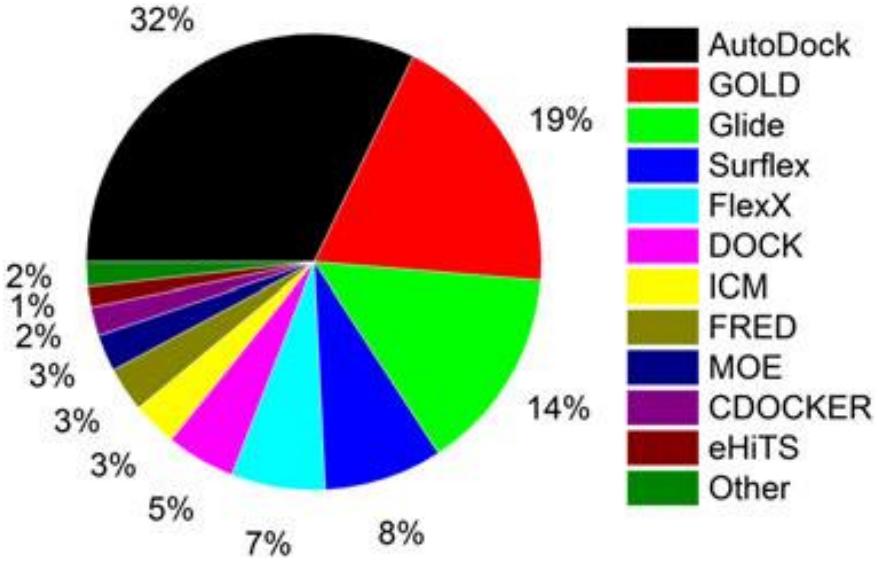
Virtual screening

high-throughput screening → **hit** to **lead** (H2L) → lead optimization

10^{-6} M 10^{-9} M

Umestitev majhnih molekul





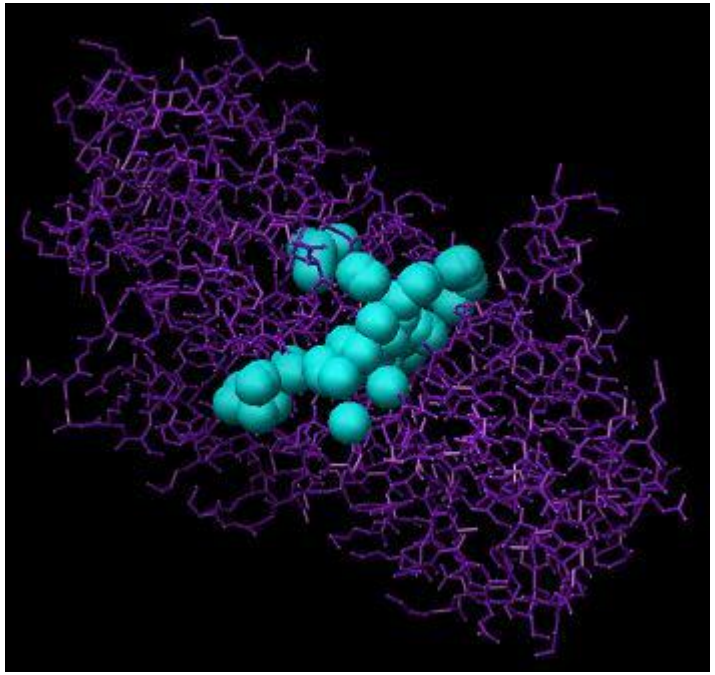
Število publikacij, ki vsebujejo rezultate umestitev.

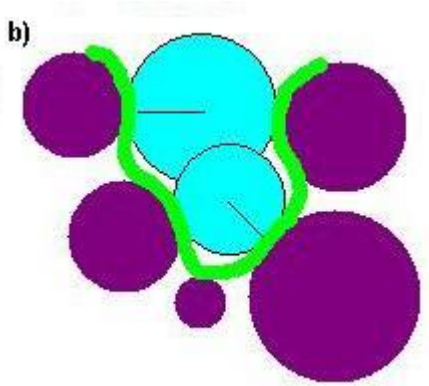
Dve temeljni obliki uporabe

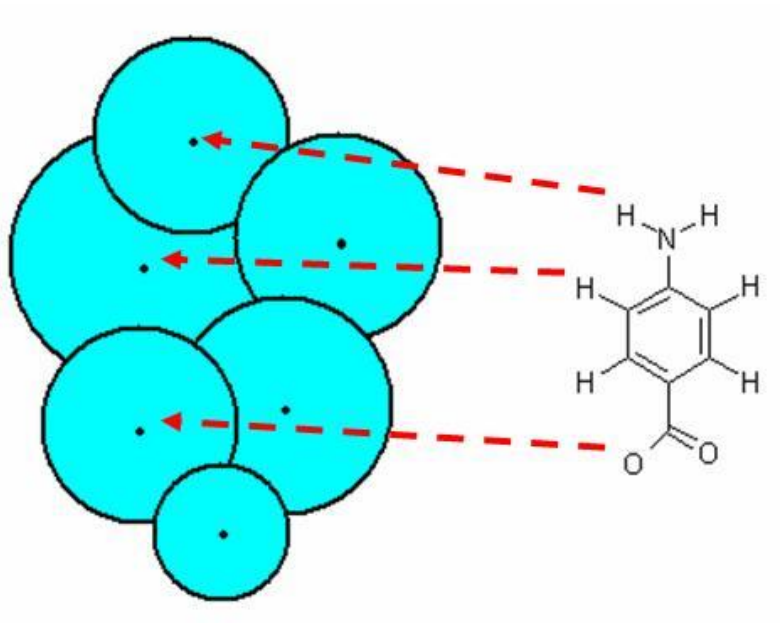
- Umestitev knjižnic spojin – hitreje, manj natančno
- Umestitev posameznih spojin – počasneje, bolj natančno

UCSF DOCK – rigidna umestitev

ustvarimo negativno sliko površne vezavnega mesta na receptorju s sferami definirane velikosti.





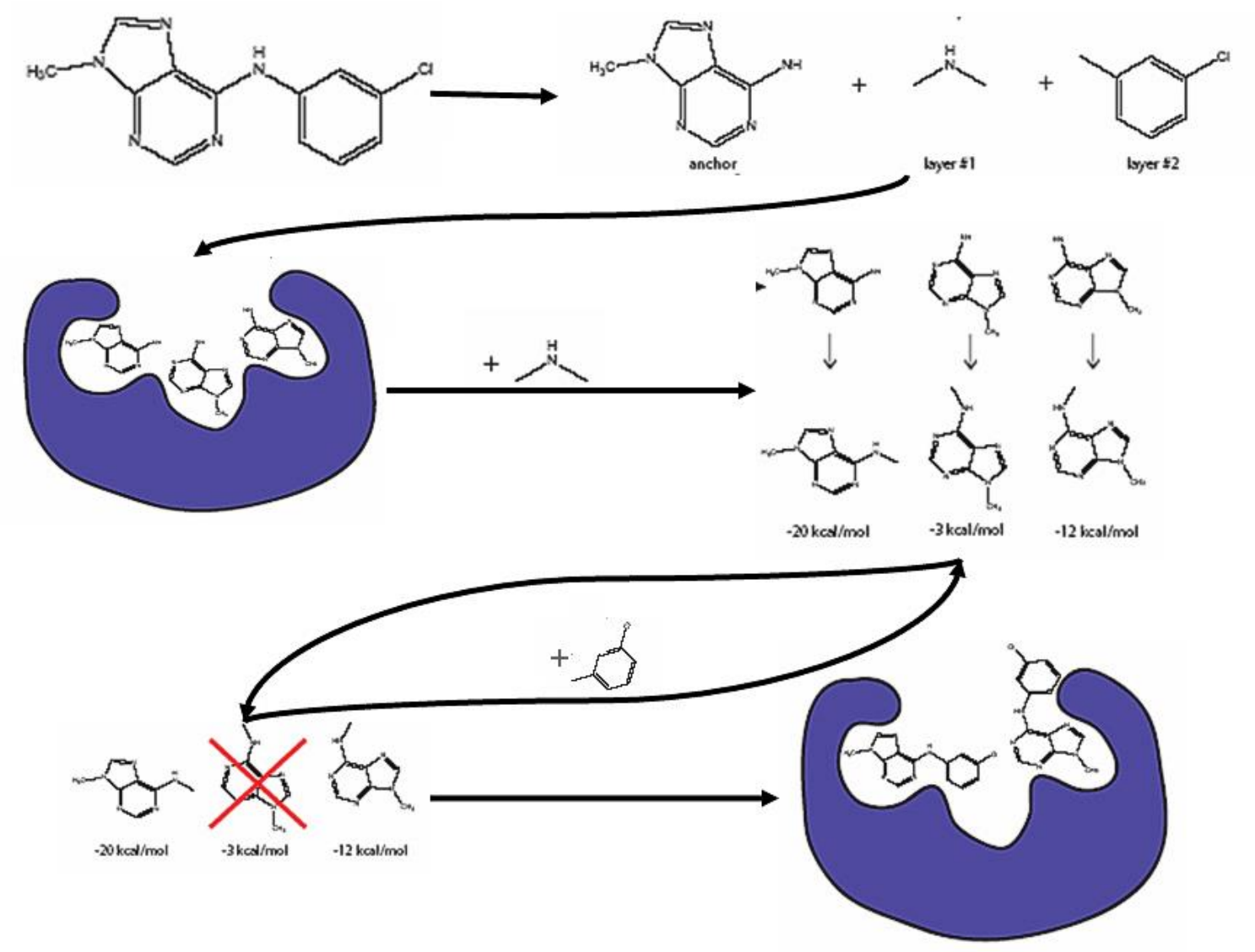


Rigidno molekulo liganda poskušamo pozicionirati v sfere, tako da težki atomi sovpadajo s centri sfer. Program ustvari nekaj sto različnih orientacij in za vsako izračuna energijo.

$$E = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right)$$

UCSF DOCK – fleksibilna umestitev

Postopek „anchor-and-grow“



ZINC

Baza vseh komercialno dostopnih spojin na planetu.

zinc.docking.org

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

Shoichet Laboratory docking.org

ZINC 12

Not Authenticated – sign in

Active cart: Temporary Cart (0 items)

About Search Subsets Help Social g+1 76 Quick Search Bar... Go

Welcome to ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 35 million purchasable compounds in ready-to-dock, 3D formats. ZINC is provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF). To cite ZINC, please reference: Irwin, Sterling, Mysinger, Bolstad and Coleman, *J. Chem. Inf. Model. 2012* DOI: 10.1021/ci3001277. The original publication is Irwin and Shoichet, *J. Chem. Inf. Model. 2005;45(1):177-82* PDF, DOI. We thank NIGMS for financial support (GM71896).

ZINC ID, Drug Name, SMILES, Catalog, Vendor Cod Go

Structure/Draw Physical Properties Catalogs & Vendors ZINC IDS Targets Rings Combination

What's NEW? Feedback Like us

@chem4biology Blog RSS

Video Walkthroughs

ZINC Database

Like 493

Quick Links

Download Search

Target focused Thanks

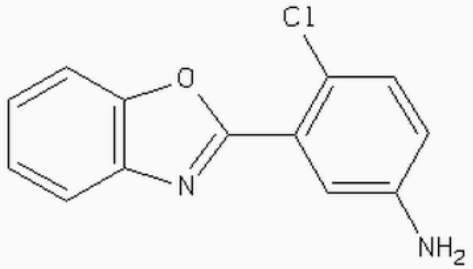
Natural Products Special Subsets

Search By Target PBCs

Your Carts

Create an account or login to have multiple carts.

Molecule of the Week 122974



ZINC

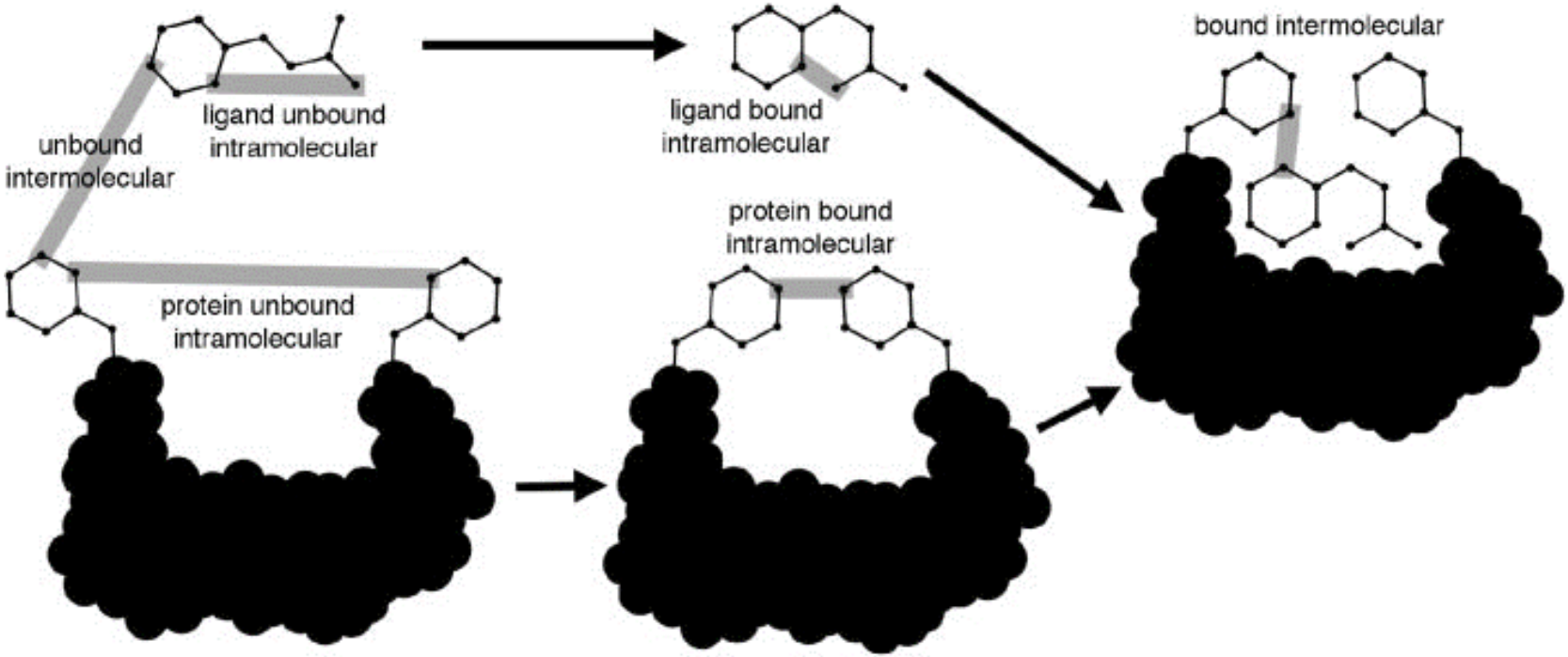
Standard: For delivery in 0-10 weeks, including in-stock, make-on-demand and agent compounds
Clean: Stricter filtering rules have been applied — e.g. aldehyes and thiols have been removed
In Stock: For immediate delivery — typically less than 2 weeks
Boutique: Boutique compounds — you should be willing to "pay any price". These are NOT included in standard, clean and in stock subsets, where the target price is $100 or less per sample

More about ZINC subsets is here. ZINC may be used free of charge for research by individuals and institutions. **Whereas you are free to share the results of a ZINC search or a screen of molecules from ZINC, you may not redistribute major portions of ZINC without the express written permission of John Irwin.**

	Lead-Like	Fragment-Like	Drug-Like	All	Shards
Standard Size Updated	**Lead-Like** 6,687,370 2013-03-11	**Fragment-Like** 1,389,525 2013-10-25	**Drug-Like** 15,798,630 2013-02-08	**All Purchasable** 22,724,825 2013-12-18	**Shards** 85,247 2013-10-20
Clean Size Updated	**Clean Leads** 5,735,035 2013-11-05	**Clean Fragments** 148,310 2013-11-05	**Clean Drug-Like** 13,195,609 2013-11-05	**All Clean** 16,403,865 2013-12-18	**Clean Shards** 60,021 2013-11-05
In Stock Size Updated	**Leads Now** 2,419,472 2013-11-01	**Frags Now** 527,585 2013-10-25	**Drugs Now** 7,397,957 2013-11-11	**All Now** 9,046,036 2013-04-04	**Shards Now** 63,861 2013-10-20
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	**Boutique Shards** 80,698 2013-11-08
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	Type I binding sites
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 <= 10		p.mwt < 160

AutoDock

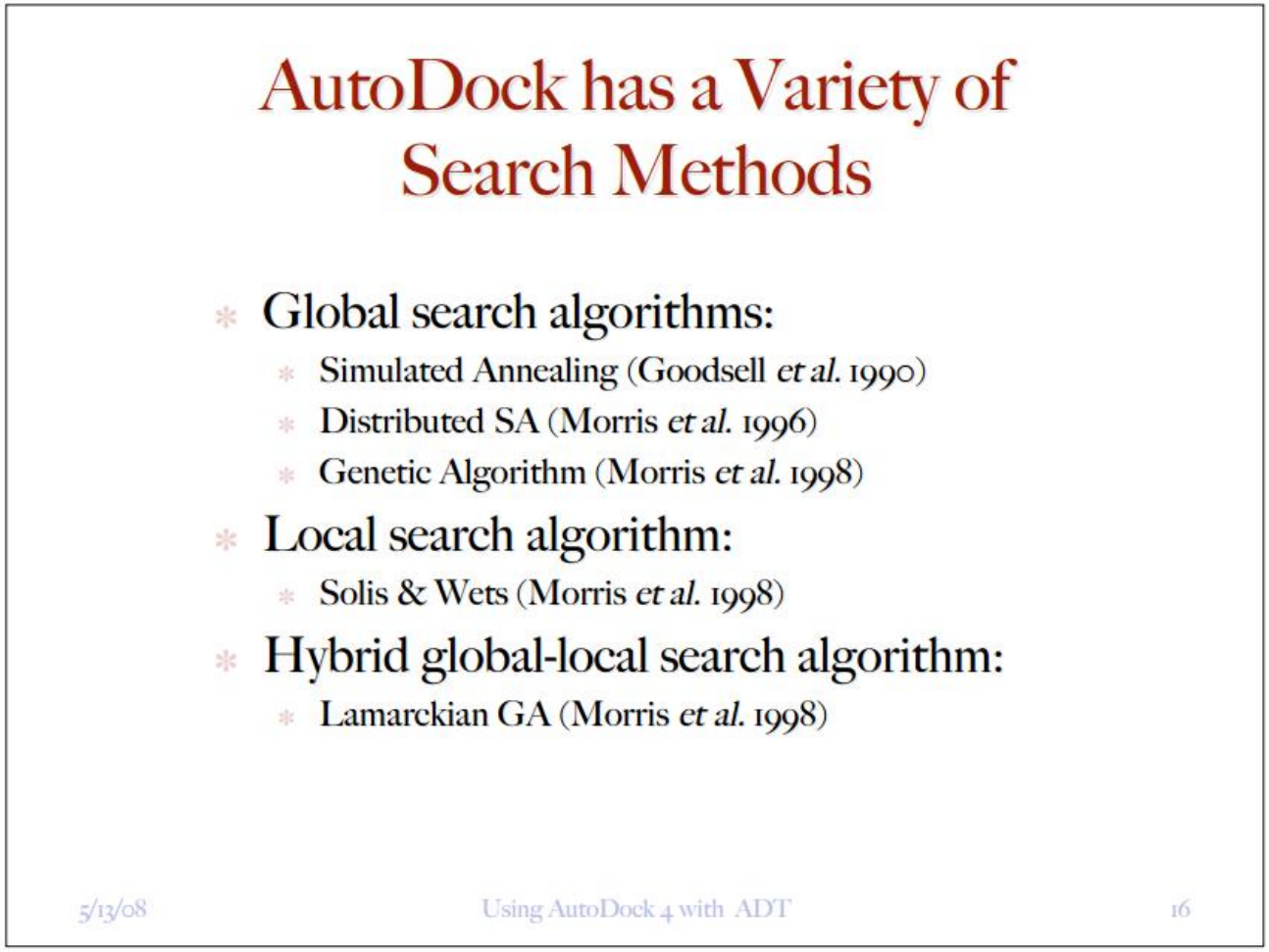
Uporablja semiempirično funkcijo za vrednotenje in več metod naključnega (stohastičnega) iskanja. Površino receptorja definira z mrežo (AutoGrid), kar pospeši kalkulacije.



$$\Delta G = (V^{\mathrm{L-L}}_{\mathrm{bound}} - V^{\mathrm{L-L}}_{\mathrm{unbound}}) + (V^{\mathrm{P-P}}_{\mathrm{bound}} - V^{\mathrm{P-P}}_{\mathrm{unbound}}) + (V^{\mathrm{P-L}}_{\mathrm{bound}} - V^{\mathrm{P-L}}_{\mathrm{unbound}} + \Delta S_{\mathrm{conf}})$$

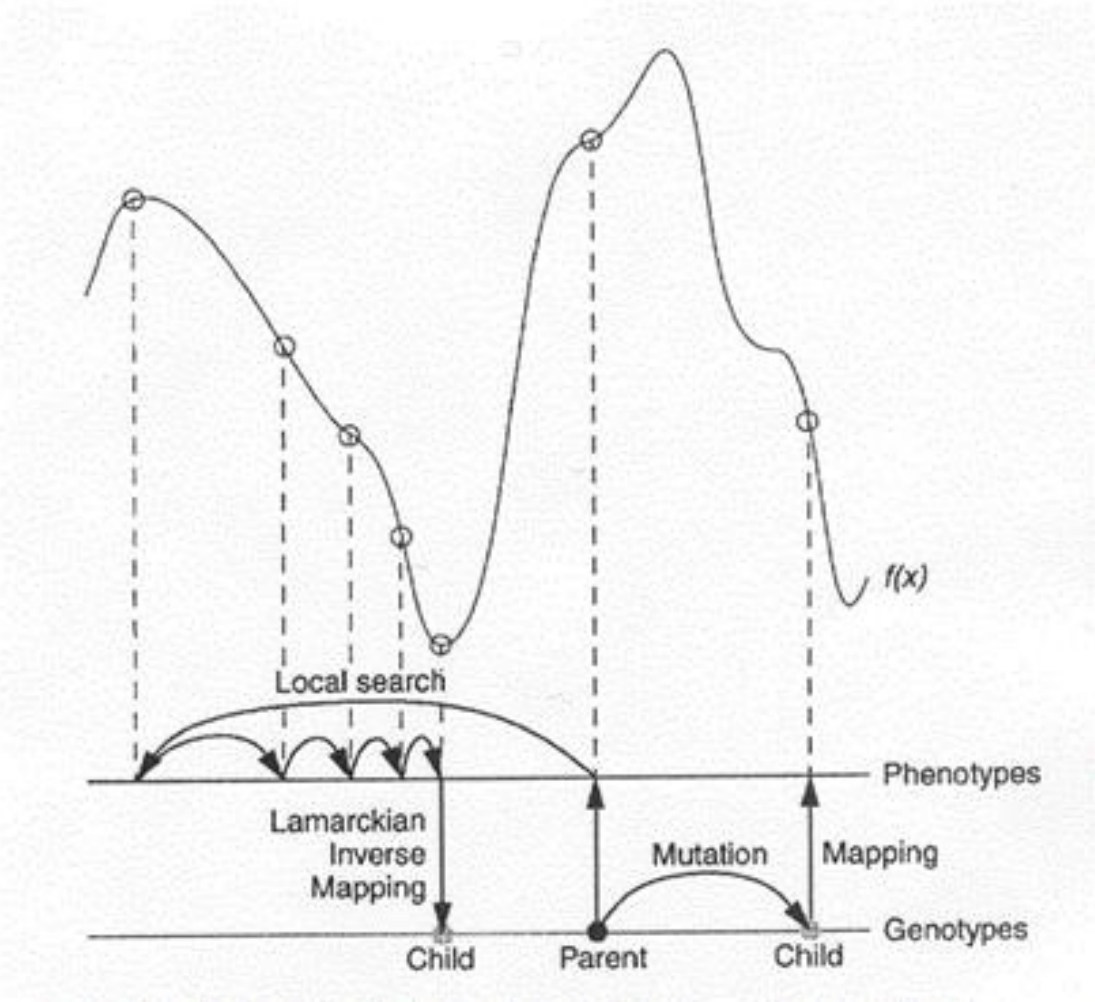
AutoDock

Uporablja semiempirično funkcijo za vrednotenje in več metod naključnega (stohastičnega) iskanja. Površino receptorja definira z mrežo (AutoGrid), kar pospeši kalkulacije.



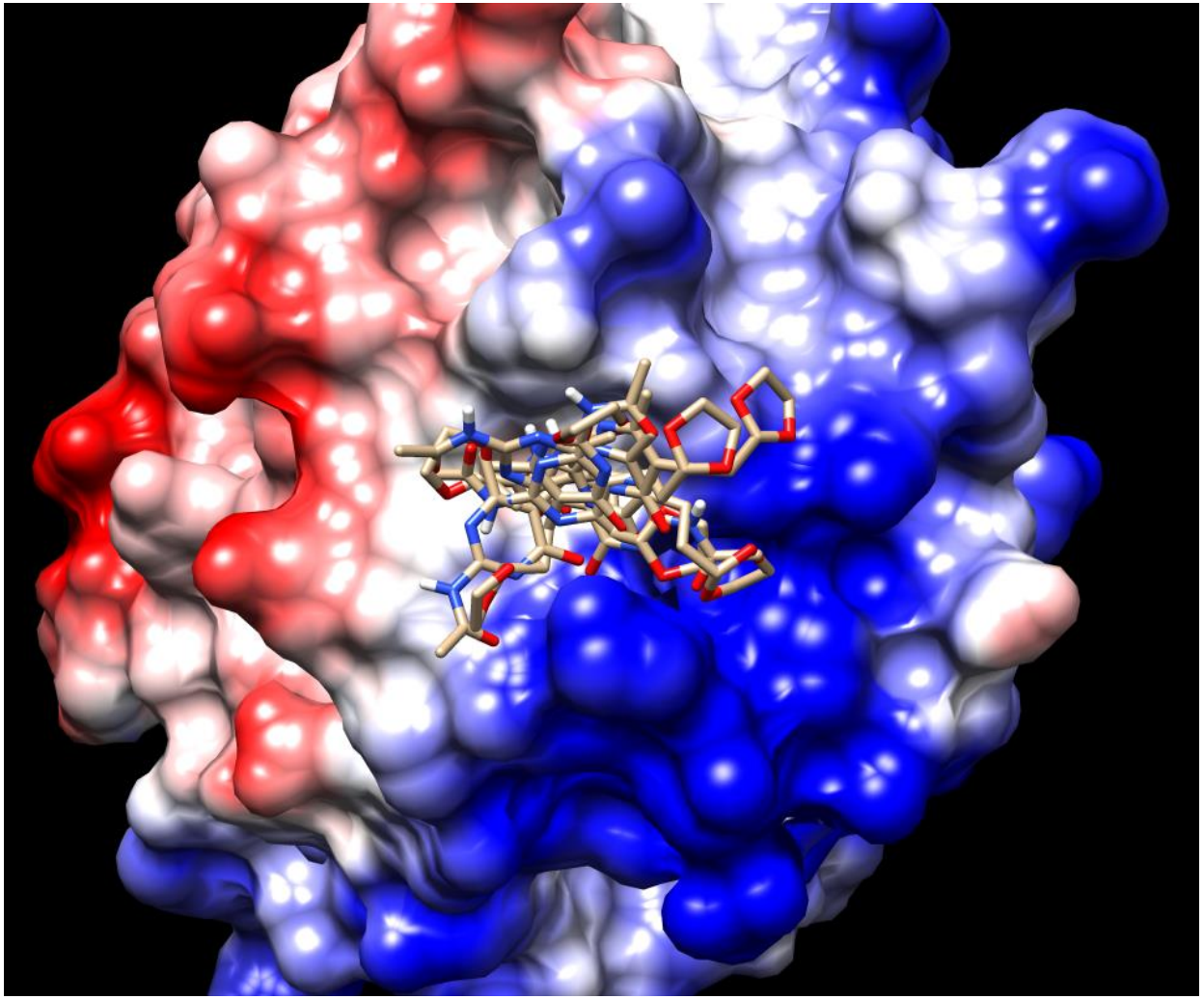
AutoDock

Uporablja semiempirično funkcijo za vrednotenje in več metod naključnega (stohastičnega) iskanja. Površino receptorja definira z mrežo (AutoGrid), kar pospeši kalkulacije.



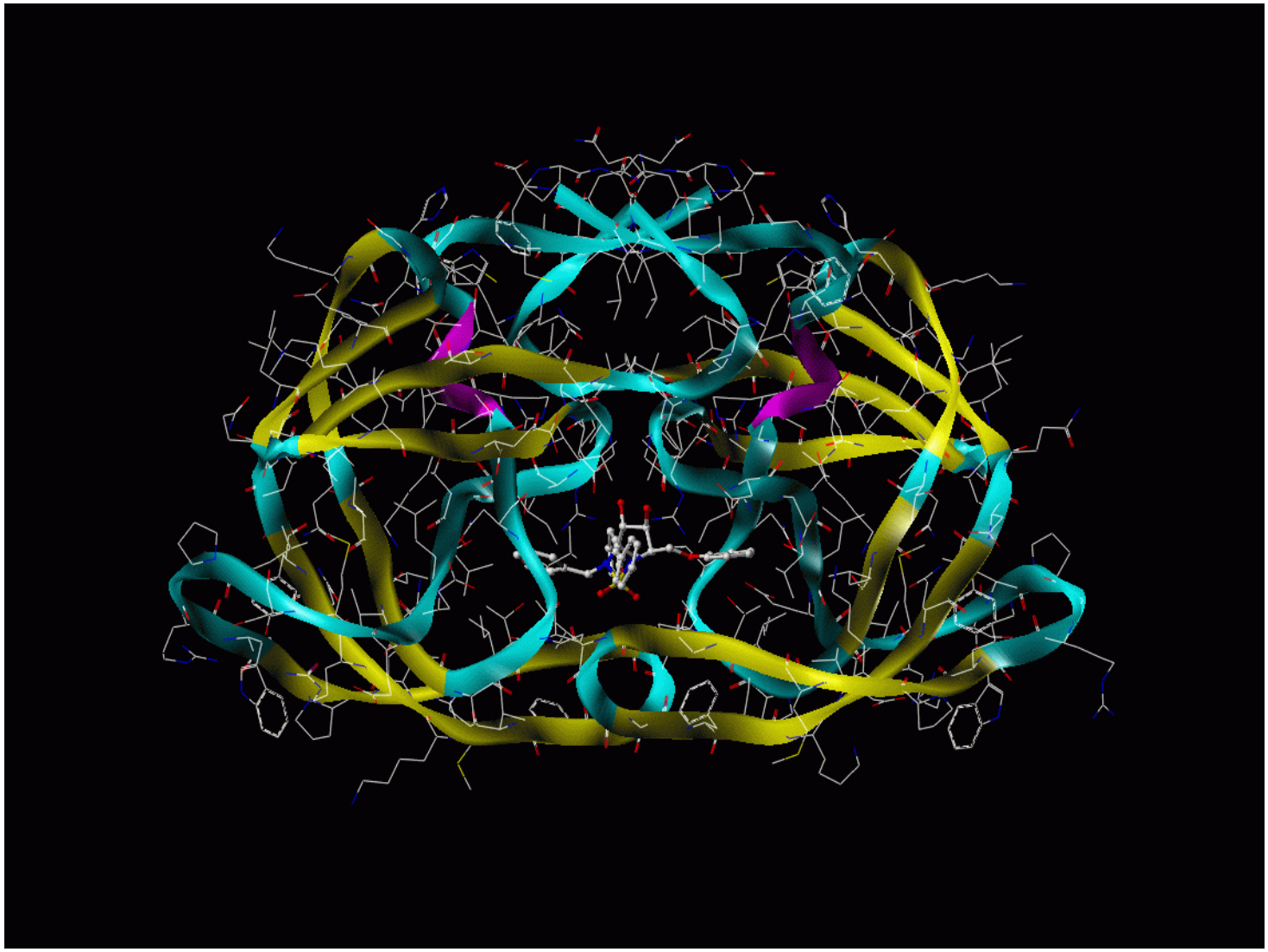
AutoDock

AutoDock izračuna več rešitev in jih združi v skupine.



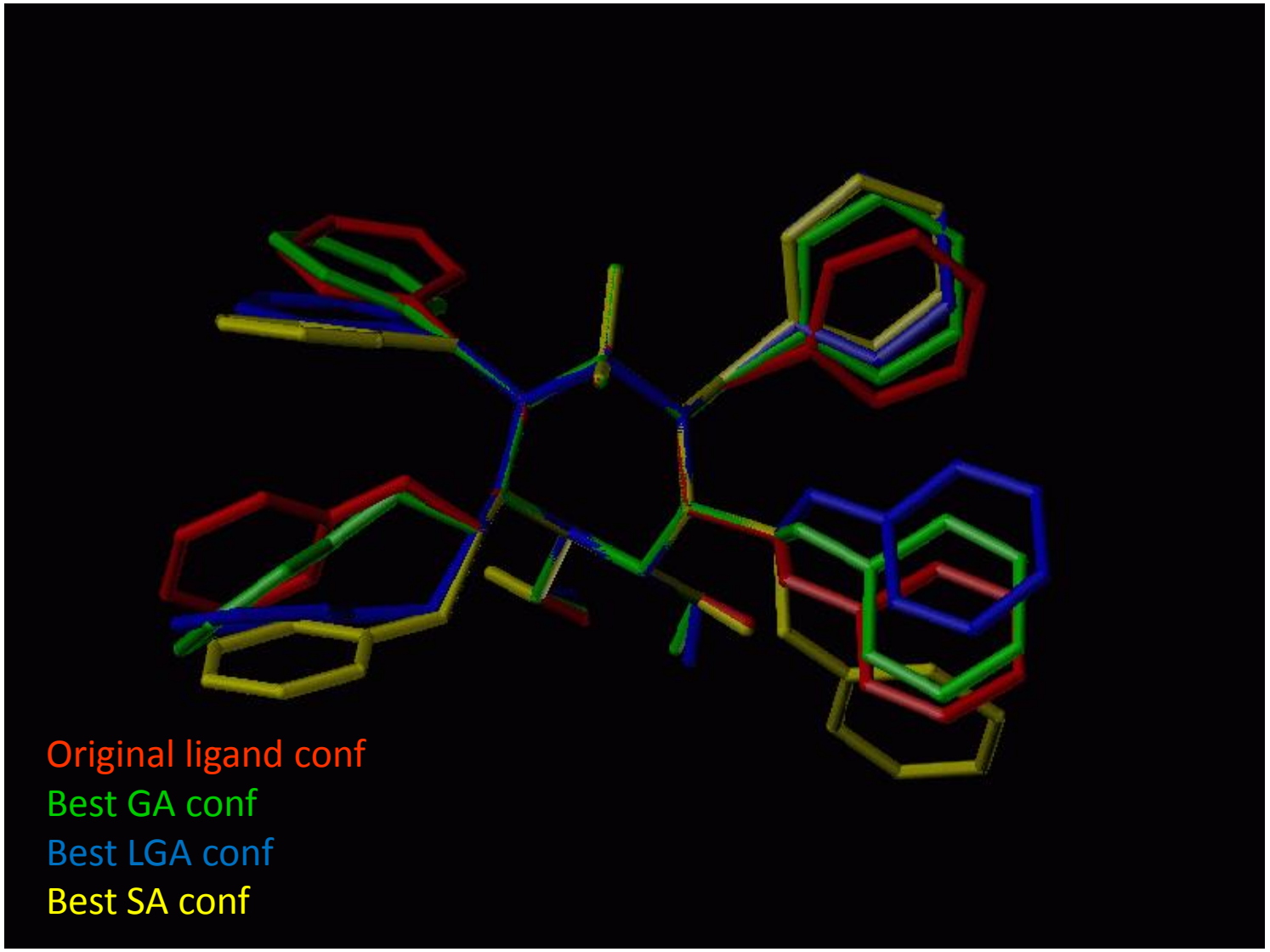
AutoDock

Primer: inhibitor HIV-1 proteaze AHA006.



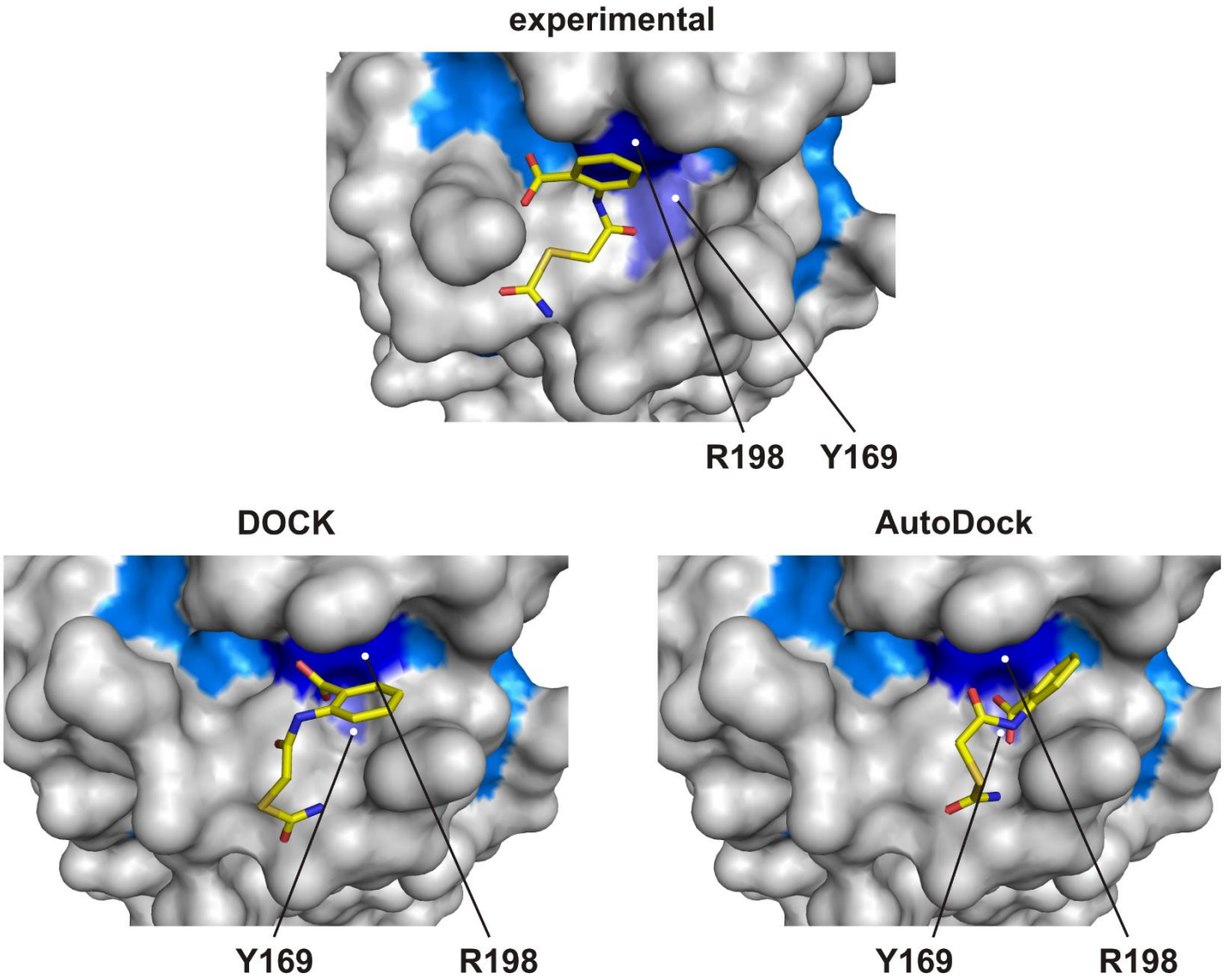
AutoDock

Primer: inhibitor HIV-1 proteaze AHA006.



DOCK in AutoDock

Primer: inhibitor NSC13345 vezan na katepsin K.



DOCK in AutoDock

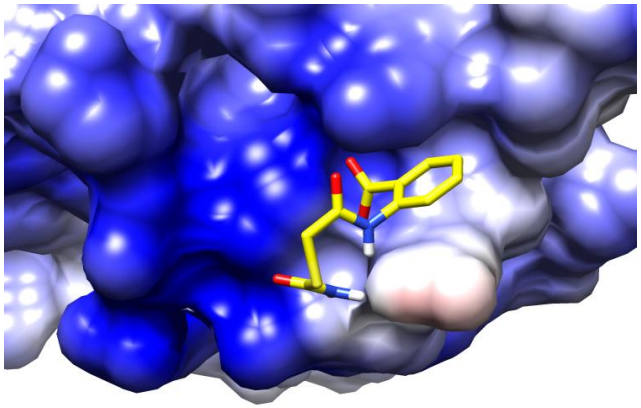
Primer: inhibitor NSC13345 vezan na katepsin K.

Rezultati AutoDock:

256 ponovitev

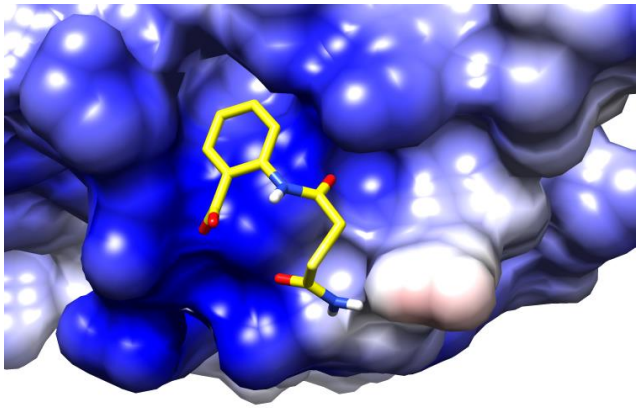
Clus-ter Rank	Lowest Binding Energy	Run	Mean Binding Energy	Num in Clus
1	-5.10	235	-4.70	103
2	-4.70	9	-4.00	79
3	-4.58	207	-4.48	42
4	-4.18	5	-3.65	11
5	-3.95	24	-3.66	19
6	-3.44	112	-3.44	1
7	-3.34	13	-3.34	1

Cluster 1



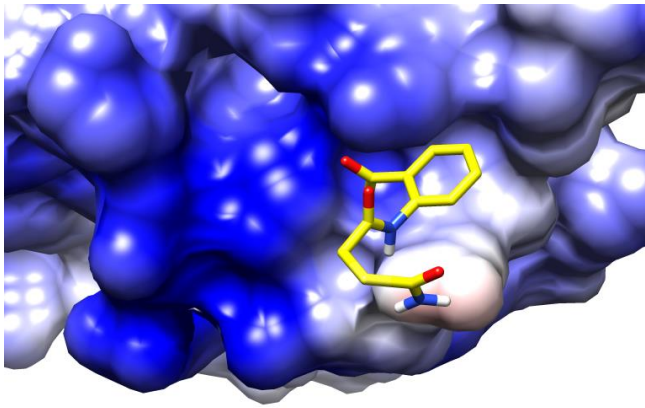
K_i = 182 μM

Cluster 2



K_i = 360 μM

Cluster 3



K_i = 438 μM

DOCK in AutoDock

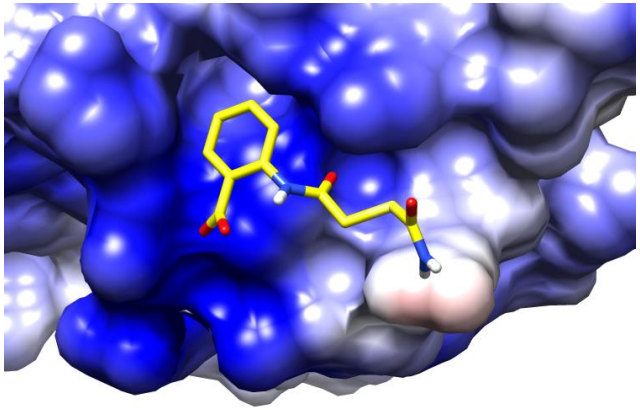
Primer: inhibitor NSC13345 vezan na katepsin K.

Rezultati AutoDock:

256 ponovitev

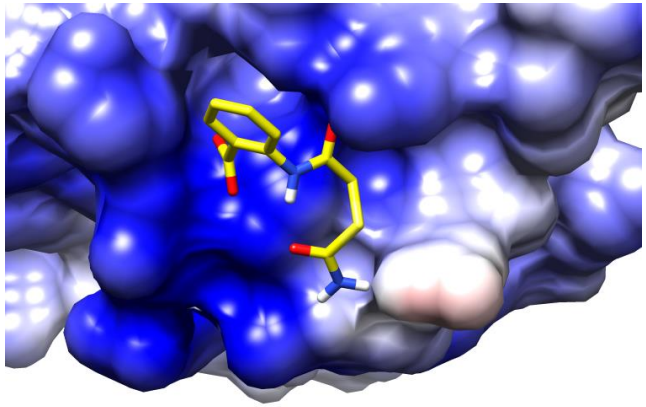
Clus-ter Rank	Lowest Binding Energy	Run	Mean Binding Energy	Num in Clus
1	-5.10	235	-4.70	103
2	-4.70	9	-4.00	79
3	-4.58	207	-4.48	42
4	-4.18	5	-3.65	11
5	-3.95	24	-3.66	19
6	-3.44	112	-3.44	1
7	-3.34	13	-3.34	1

Cluster 4



K_i = 866 μM

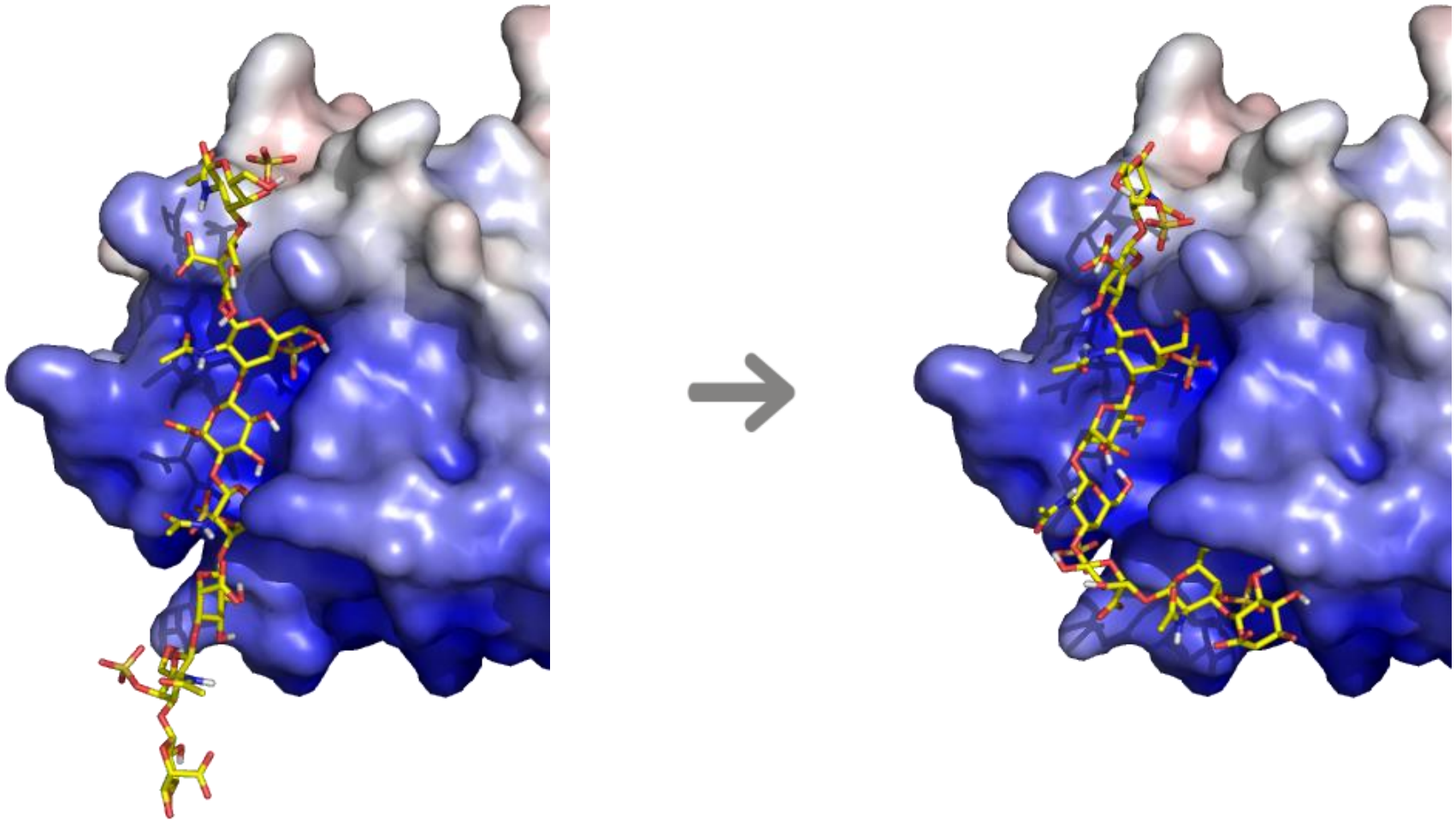
Cluster 5



K_i = 1270 μM

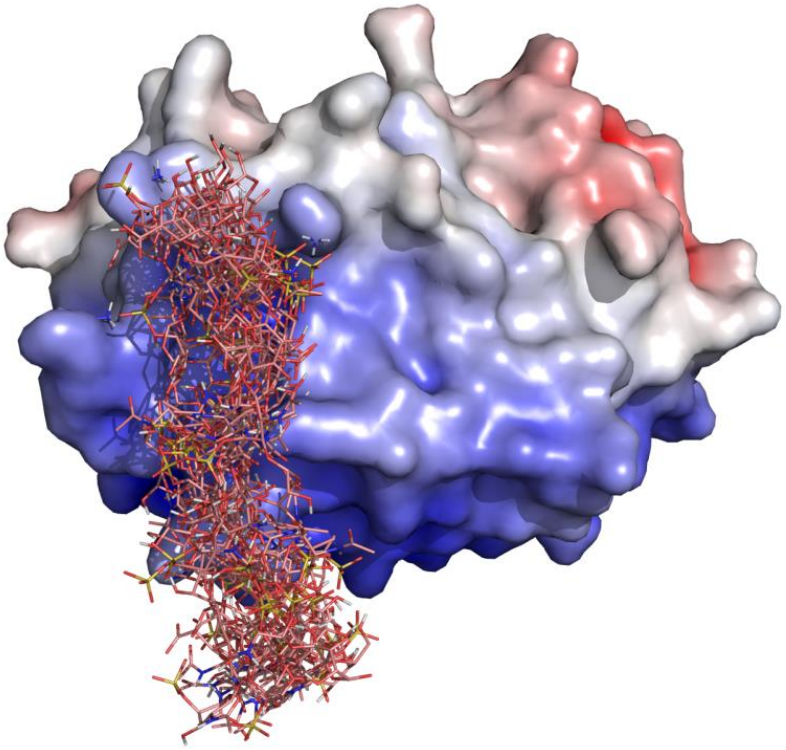
AutoDock

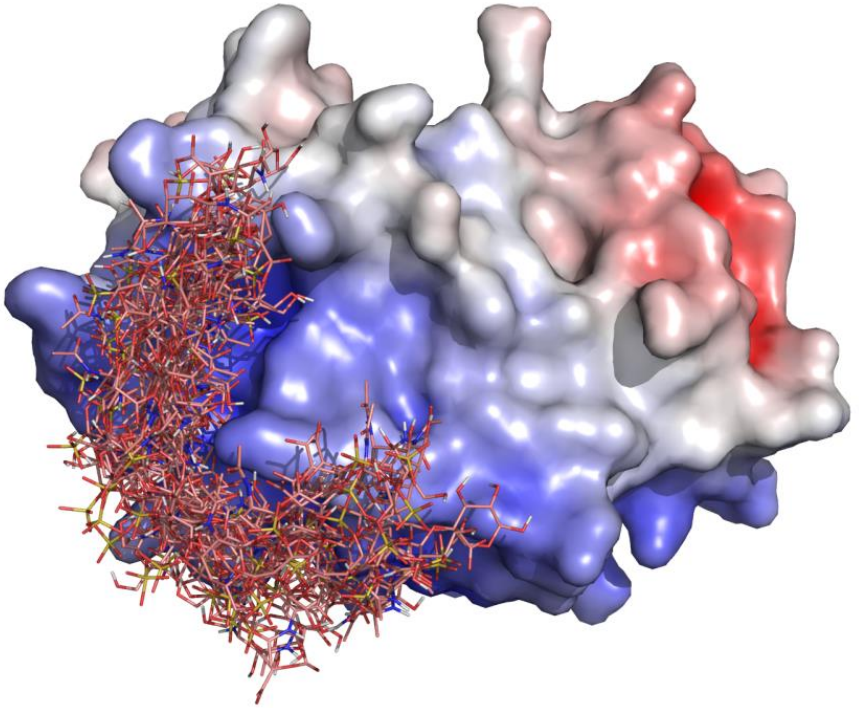
Primer: konformacijska sprememba dermatansulfata ob vezavi na katepsin K.



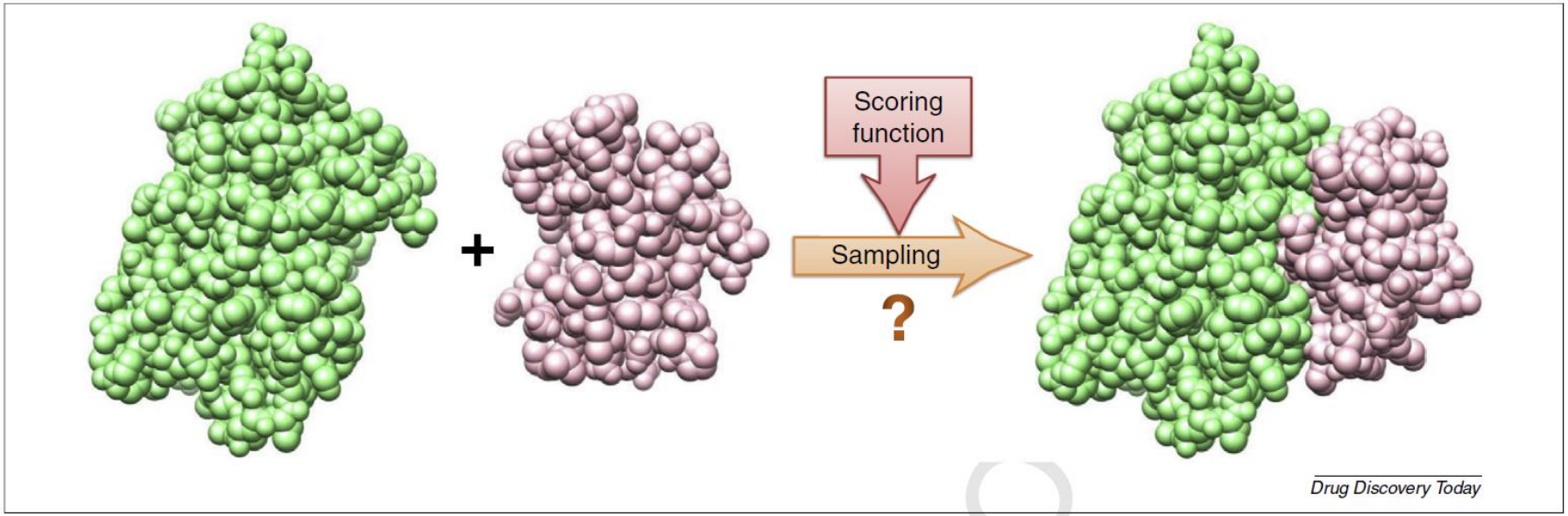
AutoDock

Primer: konformacijska sprememba dermatansulfata ob vezavi na katepsin K.





Makromolekulska umestitev



algoritem za iskanje rešitev + funkcija za ovrednotenje rešitev

Makromolekulska umestitev

TABLE 5

List of available protein–protein docking servers

Server	Algorithm	Website
ClusPro	FFT based	http://cluspro.bu.edu/
GRAMM-X	FFT based	http://vakser.bioinformatics.ku.edu/resources/gramm/grammx/
ZDOCK	FFT based	http://zdock.umassmed.edu/
3D-Garden	FFT based	http://www.sbg.bio.ic.ac.uk/~3dgarden/
HEX Server	SFT based	http://hexserver.loria.fr/
PatchDock	Geometric hashing	http://bioinfo3d.cs.tau.ac.il/PatchDock/
HADDOCK	Randomized search	http://haddock.science.uu.nl/services/HADDOCK/
RosettaDock	Randomized search	http://rosettadock.graylab.jhu.edu/
SwarmDock	Randomized search	http://bmm.cancerresearchuk.org/SwarmDock/
DOCK-PIE	Post-docking	http://clsb.ices.utexas.edu/web/dock.html
FiberDock	Post-docking	http://bioinfo3d.cs.tau.ac.il/FiberDock/
FireDock	Post-docking	http://bioinfo3d.cs.tau.ac.il/FireDock/
pyDockWeb	Post-docking	http://life.bsc.es/servlet/pydock/home/

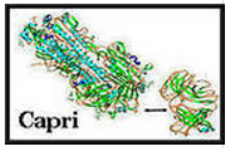
Huang. Drug Discov Today. (2014)

CAPRI

www.ebi.ac.uk/msd-srv/capri/capri.html

EMBL-EBI | Services | Research | Training | About us

CAPRI: Critical Assessment of PRediction of Interactions



Databases > PDBe > Services > Capri-Home

> contact PDBe

- PDB idcodes for past targets
- Call For Targets
- Capri Rules 2007
- Original Capri Rules 2001
- Management
- Formats
- ROUND 29
- ROUND 28
- ROUND 27
- ROUND 26
- ROUND 25
- ROUND 24
- ROUND 23
- ROUND 22
- ROUND 21
- ROUND 20
- ROUND 19
- ROUND 18
- ROUND 17

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

Hosted by the Protein Data Bank in Europe (PDBe) Group

5th CAPRI Evaluation meeting, Utrecht Report

The 5th CAPRI meeting was organiosed in Utrecht (April 17-19, 2013). The summary report on the Round Table discussion on CAPRI's future directions is now available.

ROUND 29 ANNOUNCEMENT

New CAPRI ROUND 29

- **18th Nov 2013** - Registration opens for Round 29
- **29th Nov 2013** - Round 29 opens - Prediction of targets 65-67 (protein-peptide complexes)
- **15th Dec 2013** - Deadline for submitting models for target 65 (a protein-peptide complex)
- **20th Dec 2013** - Deadline for submitting models for targets 66-67 (protein-peptide complexes)

ROUND 27 RESULT ANNOUNCEMENT

ROUND 28 results for target 59 are now availabe

To download the target coordinates you must agree to the following conditions at the time of download.

Agreement (pdf)

To Register please email
Shashana Wodak - shoshana AT sickkids.ca, Sameer Velankar - sameer AT ebi.ac.uk

CAPRI

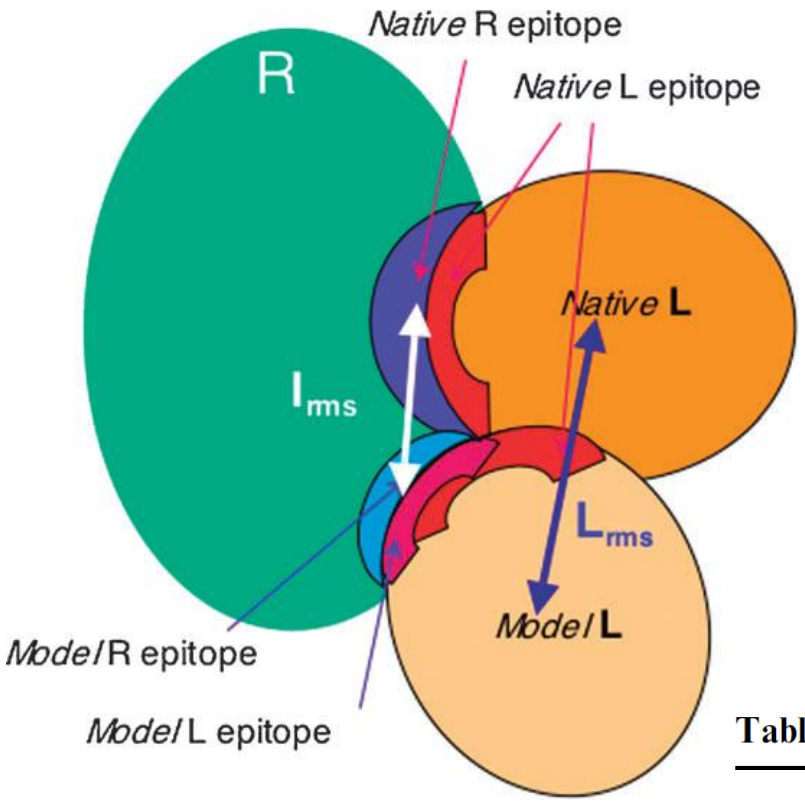


Table 2 The CAPRI star system

Model quality[a]	f_{nc}	I_{rms}/Å	L_{rms}/Å	Number of targets[b]
Three-star (high)	>0.5	<1.0	<1.0	16
Two-star (medium)	>0.3	<2.0	<5.0	12
One-star (acceptable)	>0.1	<4.0	<10.0	7
Incorrect	<0.1	or >4.0	or >10.0	6
Total[c]				41

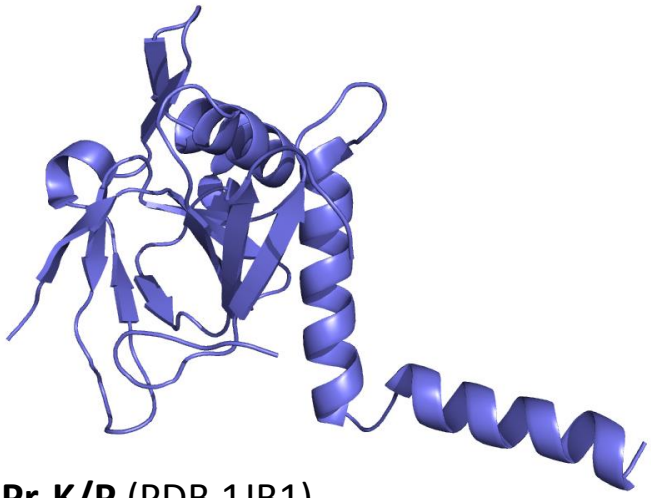
f_{nc} ... delež pravilno določenih kontaktov protein-protein

Moreira et al. 2010. *J Comput Chem*.

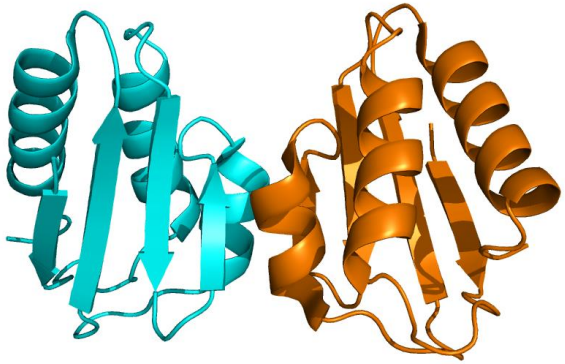
CAPRI

Primer: trial 1 – kompleks HPr kinaze/fosforilaze in fosfoprenašalca HPr

Izhodni strukturi:

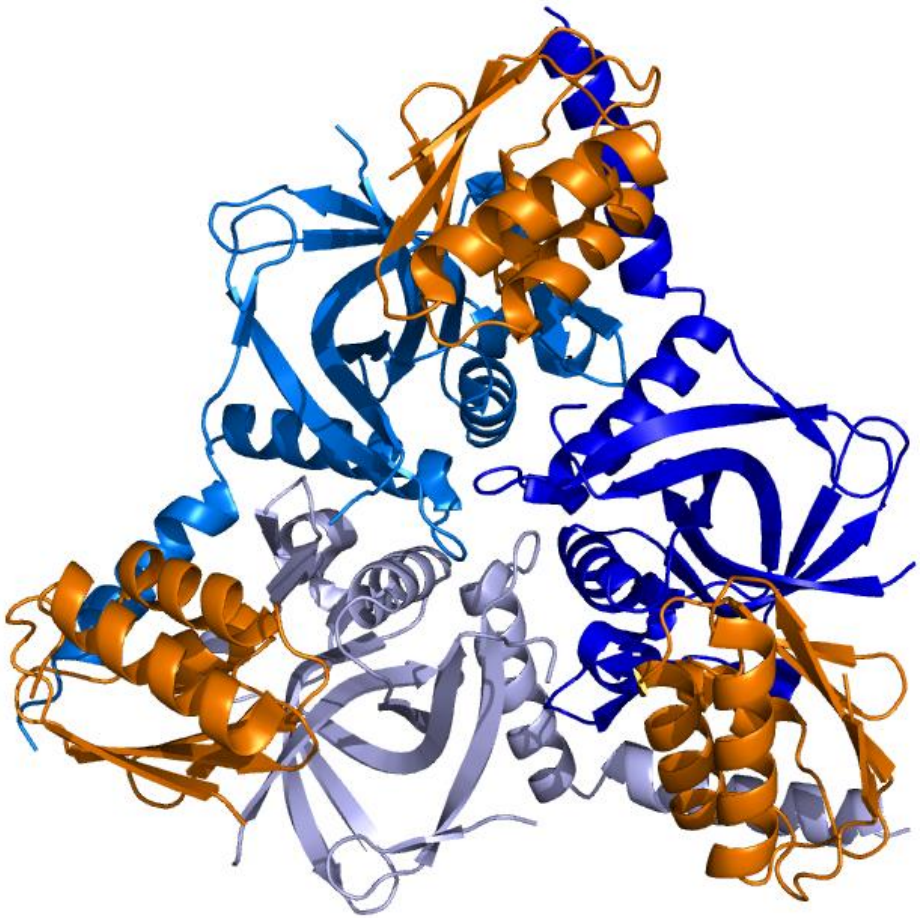


HPr-K/P (PDB 1JB1)



HPr (PDB 1SPH)

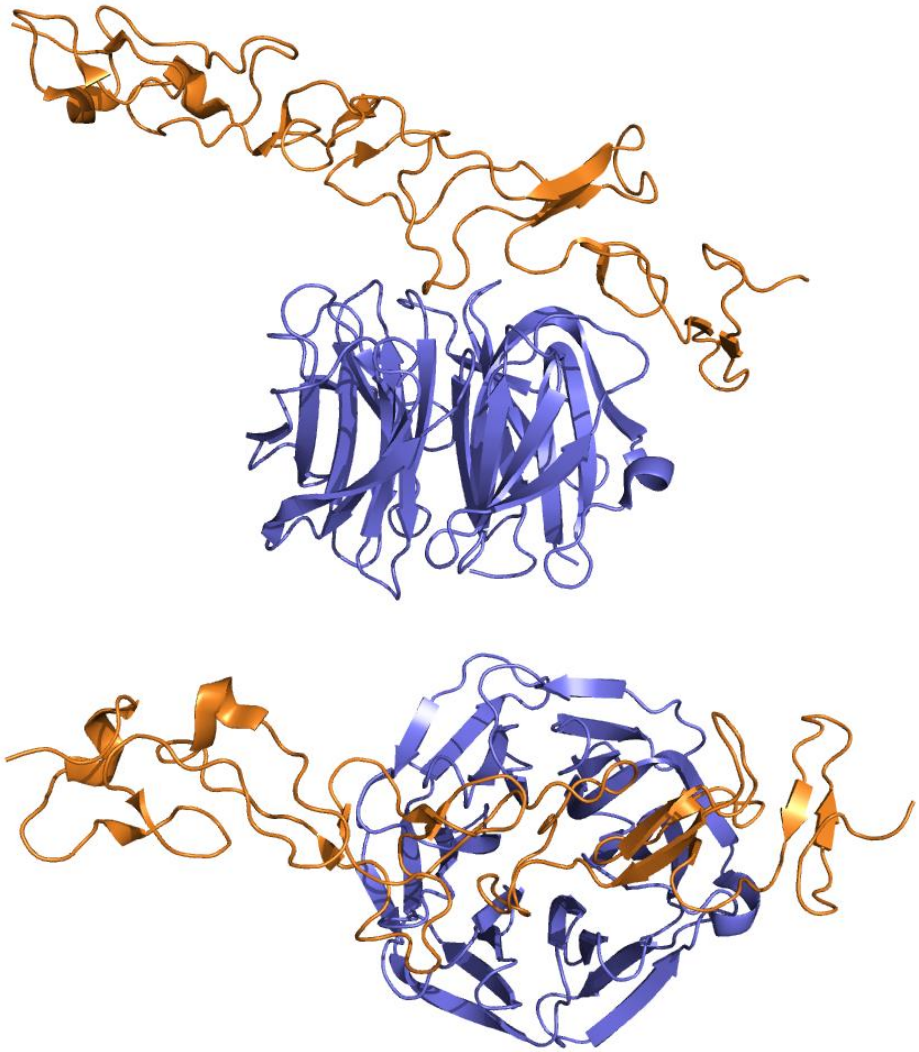
Rezultat:



(PDB 1KKL)

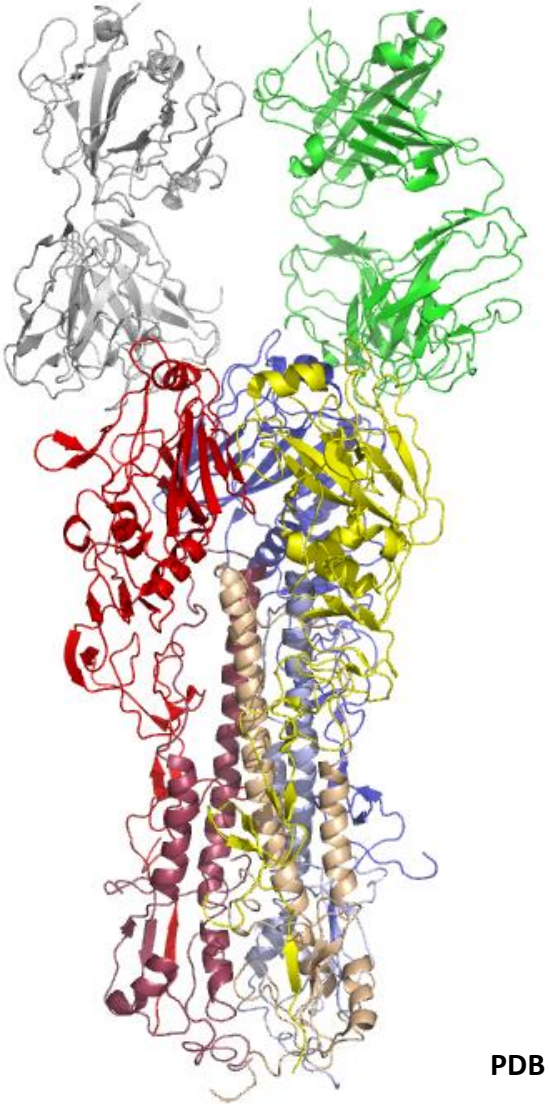
CAPRI

Trial 8:
Nidogen G3 domena (moder)/laminin (oranžen)



PDB ID 1NPE

Trial 3:
Hemaglutinin + Fab



PDB ID 1KEN

Makromolekulska umestitev

TABLE 4

List of top-performing groups in the previous CAPRI experiments

Algorithms	Group	Summary[f]
Rounds 1–2 (seven targets, 2001–2003)		
ICM-DISCO	Abagyan	3/2**/1***
SmoothDock	Camacho	3/2***
MolFit	Eisenstein	3/1***
3D-DOCK/MULTIDOCK	Sternberg	3/1***
DOT	Ten-Eyck	3/1**
Rounds 3–5 (nine targets, 2003–2005)		
ICM-DISCO	Abagyan	8/4**/2***
PatchDock/FlexDock	Wolfson	8/3**
ZDOCK	Weng	7/3**/3***
Modified 3D-DOCK, MultiDock	Bates	7/3**
RosettaDock	Baker[a]	6/2**/4***
Rounds 6–12 (seven targets, 2005–2007)		
ZDOCK, ZRANK	Weng	5/2(**)
HADDOCK	Bonvin	4/2(**)
MolFit	Eisenstein	3/1(***)
MolFit, 3D-Dock, RosettaDock, SMD refinement	Smith	3/2(**)
ClusPro	Vajda	3/2(**)
Rounds 13–19 (13 targets, 2007–2009)		
Cluspro, PIPER, SDU	Vajda	6/4***/2**
ATTRACT	Zacharias	6/4***/1**
MDockPP	Zou[b]	6/3***/2**
MolFit	Eisenstein[c]	6/3***/1**
PatchDock, FlexDock, FiberDock	Wolfson[c]	6/3***/1**
ZDOCK, ZRANK	Weng[d]	6/2***/2**
meta-PPISP, ZDOCK, ZRANK, CHARMM, HADDOCK	Zhou[d]	6/2***/2**
Rounds 20–27 (ten targets, 2010–2012)		
HADDOCK	Bonvin	9/1***/3**
SwarmDock1Markov-chain model	Bates	8/2**
Template-based docking	Vakser	7/1***
ClusPro 2.0/PIPPER	Vajda	6/2***/3**
pyDock	Fernandez-Recio[e]	6/1***/3**
ClusPro 2.0 1 SDU	Shen[e]	6/1***/3**

Huang. Drug Discov Today. (2014)

Makromolekulska umestitev

Nekatere komplekse je lažje napovedati kot druge.

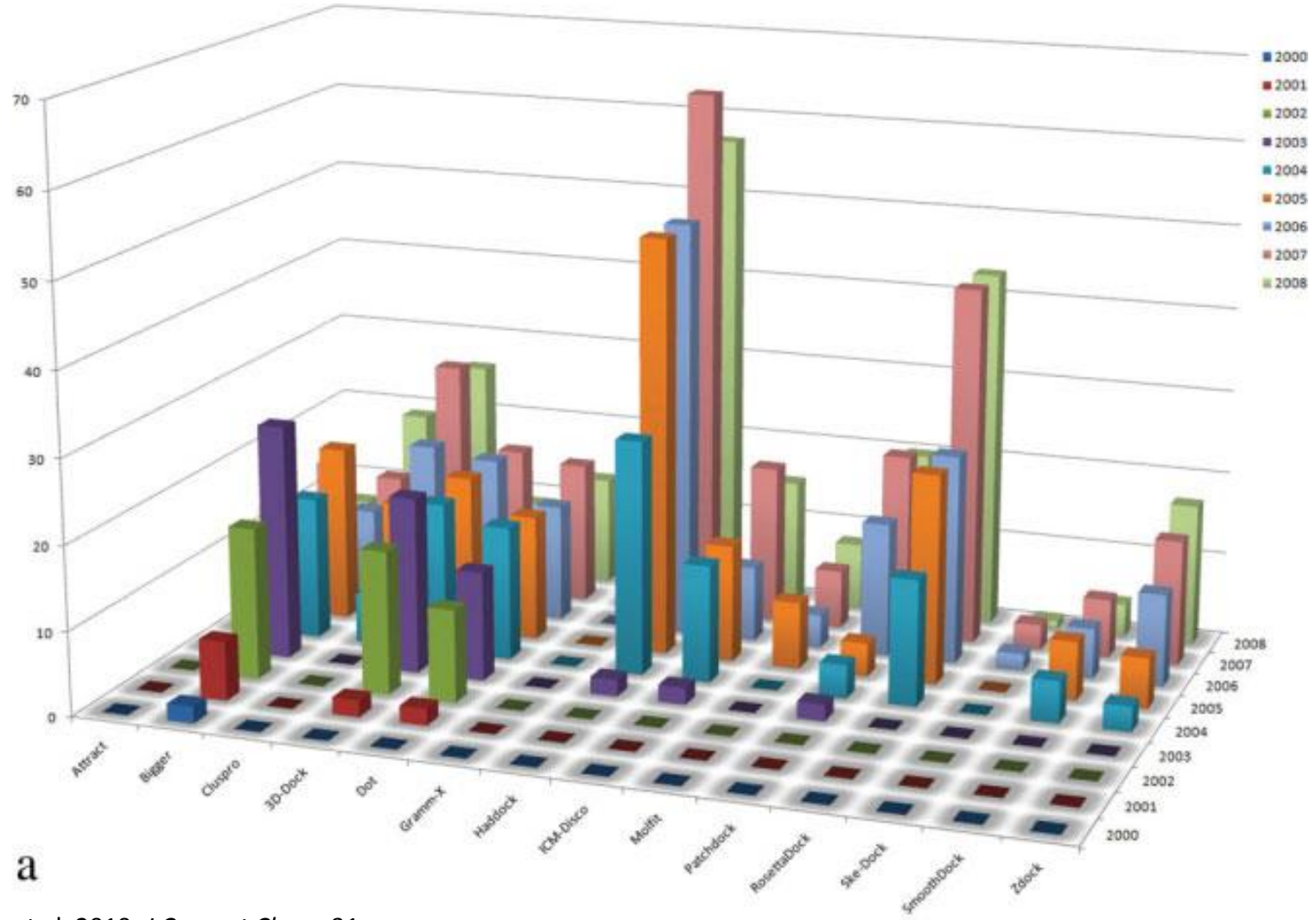
Table 3. Classification of Proteic Complexes on the Basis of Docking Difficulty.

Type	ΔASA (Å^2)	ΔG_{des} (kcal/mol)	Cα RMSD	Expected difficulty of docking
I	1400–2000	<−4.0		Easy, unless key side-chains are in the wrong conformations
II	2000–3000		<2.0	Moderated difficulty
III	1400–2000	>−4.0		Very difficult almost unpredictable
IV	<1400			Very difficult
V	>2500		>2.0	Rigid-body methods always seem to fail

Moreira et al. 2010. *J Comput Chem*. 31

Makromolekulska umestitev

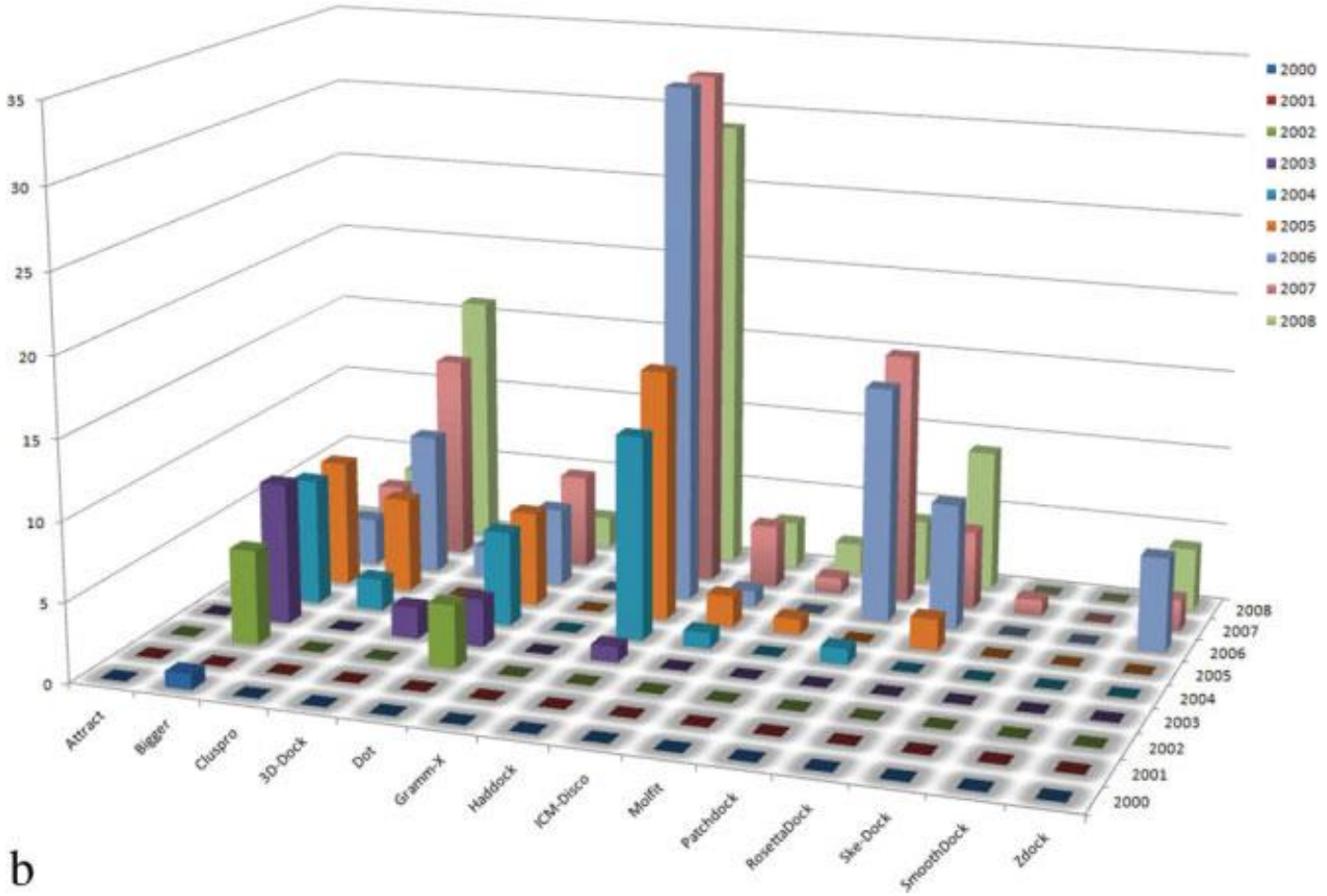
Število vseh člankov, ki so objavili rezultate umestitev protein-protein:



Moreira et al. 2010. *J Comput Chem*. 31

Makromolekulska umestitev

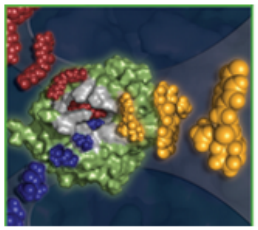
Število člankov, ki so objavili rezultate umestitev protein-protein z uporabo eksperimentalnih omejitev:



HADDOCK – 800+ citatov

Moreira et al. 2010. *J Comput Chem*. 31

Rosetta



The Baker Laboratory

NEWS | RESEARCH | MEMBERS | PAST MEMBERS | PUBLICATIONS | EMPLOY

COMMUNITY COMPUTING ALLOWS EVERYONE TO GET INVOLVED FROM HOME

foldit BETA
Solve Puzzles for Science

Foldit is a new computer game which enables you to contribute to cutting edge scientific research. Join this free online game and help us to design new proteins to cure diseases, create new materials, and develop new ways of capturing and storing energy.

Rosetta@home
Protein Folding, Design, and Docking

Rosetta@home needs your help to determine the 3-dimensional shapes of proteins in research that may ultimately lead to finding cures for some major human diseases. By running the Rosetta program on your computer while you don't need it you will help us speed up and extend our research in ways we couldn't possibly attempt without your help. You will also be helping our efforts at designing new proteins to fight diseases such as HIV, Malaria, Cancer, and Alzheimer's (See our Disease Related Research for more information). Please join us in our efforts! ***Rosetta@home is not for profit.***

ROBETTA BETA
Full-chain Protein Structure Prediction Server

Robetta: Full-chain Protein Structure Prediction

USER LOGIN

Username: *

Password: *

LOG IN

Create new account
Request new password

RosettaDock

https://www.rosettacommons.org

Home | Software | Manual | Forum | Support | Publications | Positions/REU

Search

RosettaCommons

News

Foldit in the news

Cooper *et al.* 2010 Predicting protein structures with a multiplayer online game, *Nature* **466** , 756
see also video

New positions available!

Rosetta Commons is seeking for talented and highly motivated applicants. If you are interested in computational biology and

Free Servers

ROSIE: Rosetta Server that Includes Everyone

Robetta -- Protein Structure Prediction Server

RosettaDesign -- Protein Sequence Design Server

RosettaBackrub - Flexible Backbone Modeling and Design Server

Rosetta – The premier software suite for macromolecular modeling

Rosetta is the premier software suite for modeling macromolecular structures. As a flexible, multi-purpose application, it includes tools for structure prediction, design, and remodeling of proteins and nucleic acids. Since 1998, Rosetta web servers have run billions of structure prediction and protein design simulations.

Researchers use Rosetta to better understand treatments of infectious diseases, cancers, and autoimmune disorders. Further applications involve the development of vaccines, new materials, targeted protein binders, and enzyme design.

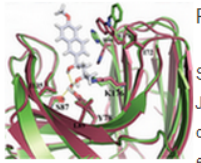
» Learn about the latest software releases

» Read the documentation

Rosetta began as a structure prediction tool, and has consistently been a strong performer in Critical Assessment of Structure Prediction (CASP) competitions. It has grown to offer a wide variety of effective sampling algorithms to explore backbone, side-chain and sequence space. Rosetta boasts broadly tested scoring (energy) functions and contains an unparalleled breadth of applications from folding to docking to design.

Rosetta is freely available to academic and government laboratories, with over 10,000 free licenses already in use. An active support forum allows users to easily collaborate within the broad research community of Rosetta users.

» Obtain a Rosetta license



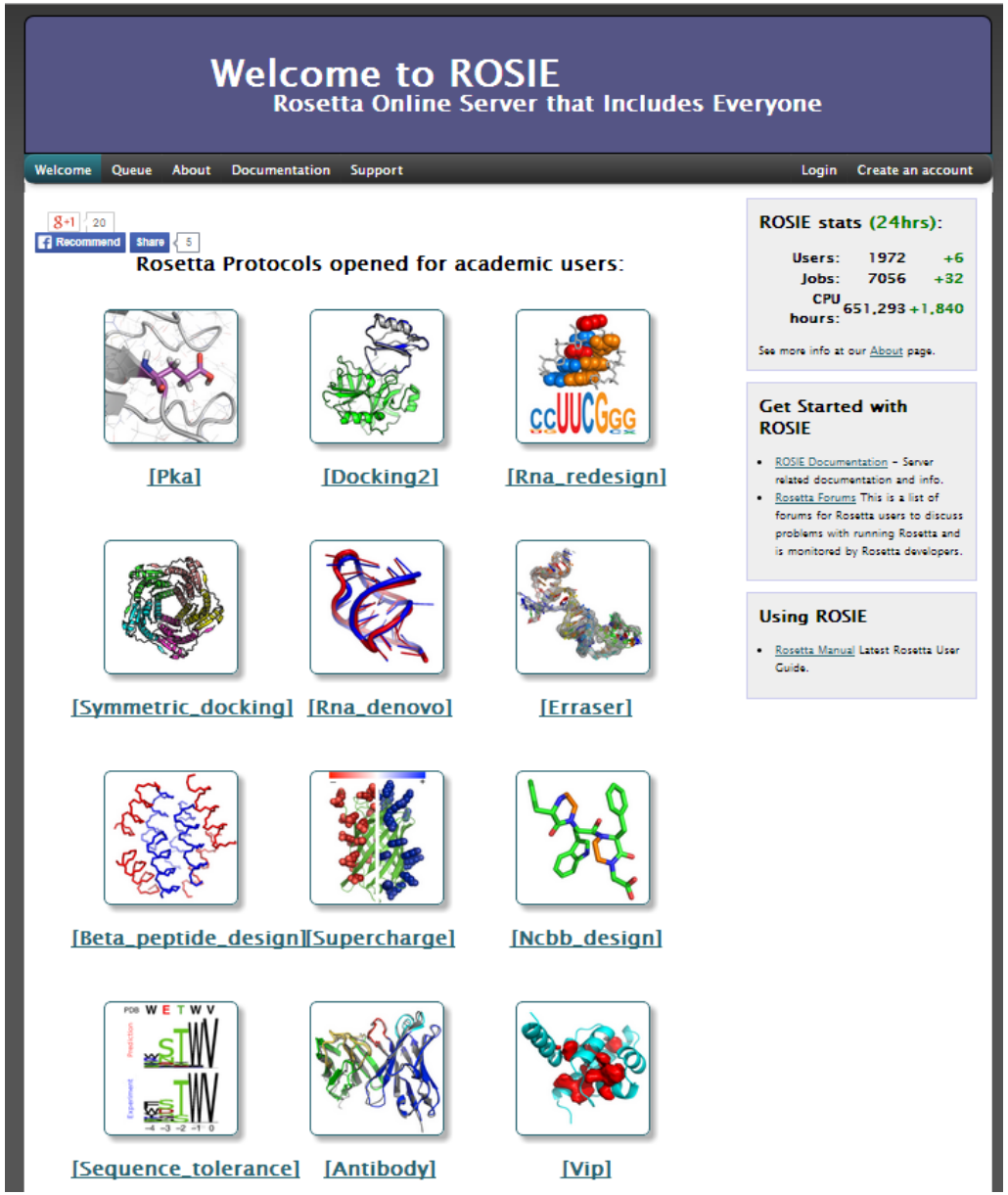
Related research includes:

Structures of designed enzymes. Jiang L, *et al* (2008). De novo computational design of retro-aldol enzymes.Science 319, 1387-91.

1.6 Å C{alpha}-RMSD blind structure prediction for CASP6 target T0281, hypothetical protein from Thermus thermophilus Hb8.(Bradley P, Misura KM, Baker D, (2005) Science. 309:1868-71.)

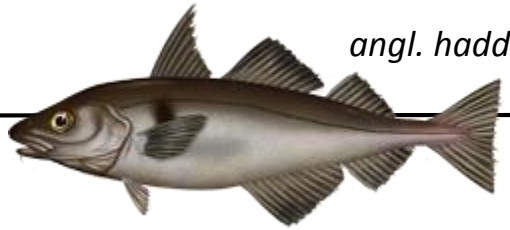
Ribbon diagrams of Top7 with residues 46 to 76 highlighted in red -- A novel protein structure created with RosettaDesign.(Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D, Science 302, 1364-8.)

RosettaDock

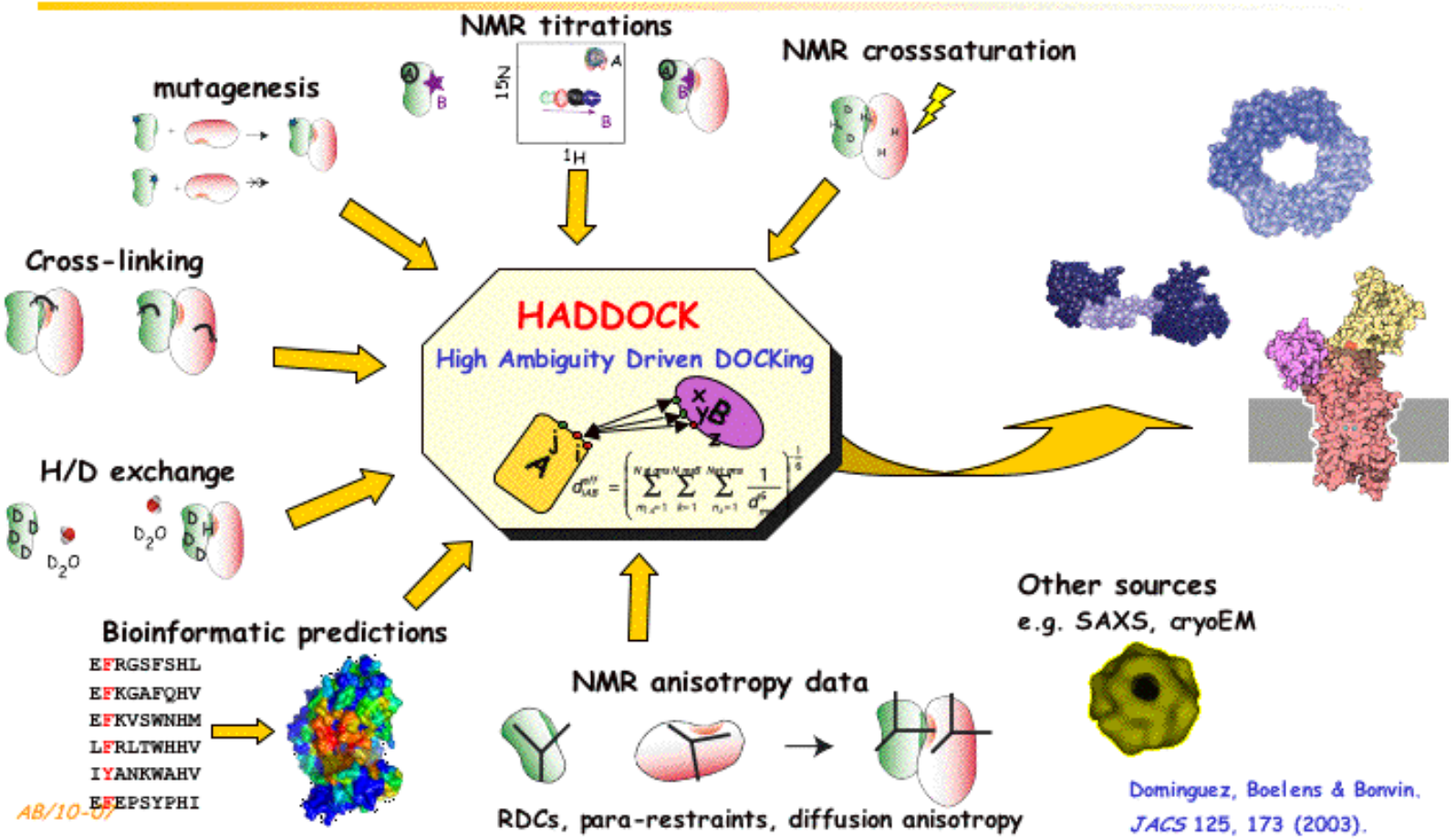


HADDOCK

angl. haddock = trska



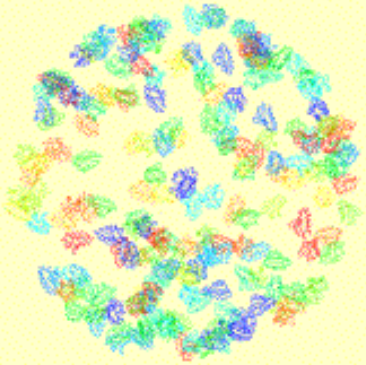
Studying biomolecular interactions



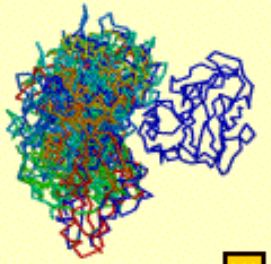
HADDOCK

HADDOCK docking protocol

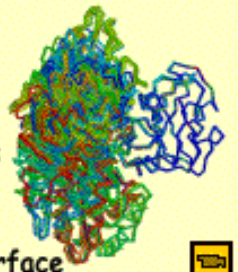
Position proteins 150Å away from each other and apply random rotations



Rigid body energy minimization:
- first only rotations
- then rotations + translations



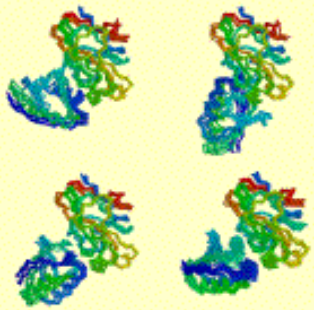
Semi-flexible simulated annealing in torsion angle space.
1. Rigid body dynamics
2. SA with flexible side-chains at the interface
3. SA with flexible backbone and side-chains at the interface



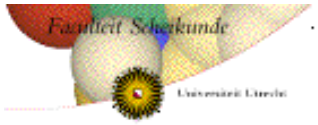
Final refinement in explicit solvent

Clustering and analysis

Scoring of clusters according to the intermolecular energy (Evdw+Eelec+EAIR)



AB



Energetics & Scoring (HADDOCK2.0)

- OPLS non-bonded parameters (Jorgensen, *JACS* 110, 1657 (1988))
- 8.5Å non-bonded cutoff, switching function, ε=10
- Ranking of rigid body docking solutions based on HADDOCK score defined as:

Rigid: $Score = 0.01\ E_{air} + 0.01\ E_{vdW} + 1.0\ E_{elec} + 1.0\ E_{desolv} - 0.01\ BSA$

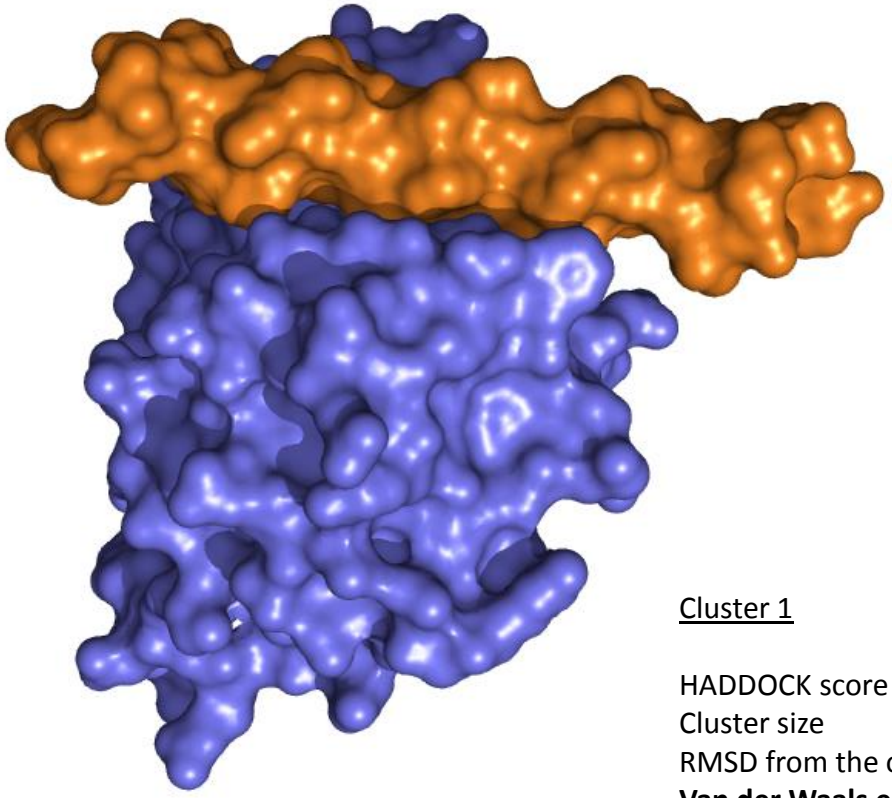
Flexible: $Score = 0.1\ E_{air} + 1.0\ E_{vdW} + 1.0\ E_{elec} + 1.0\ E_{desolv} - 0.01\ BSA$

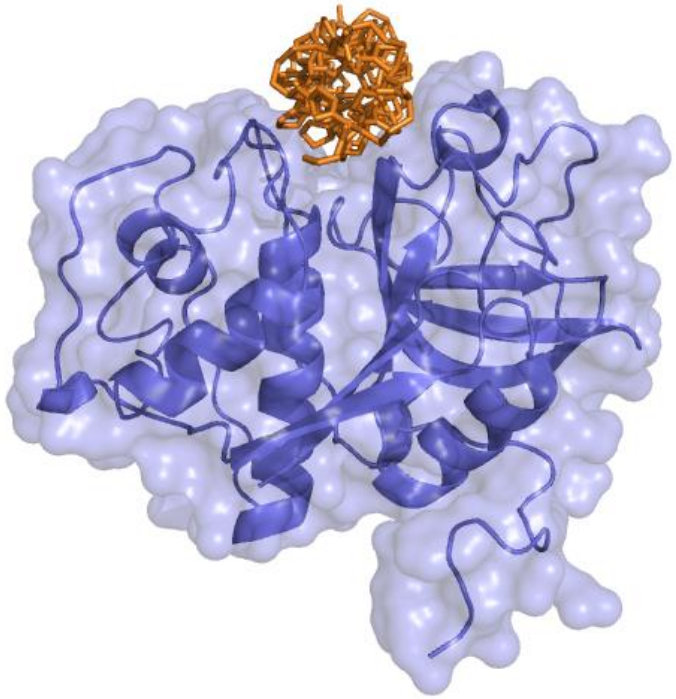
Water: $Score = 0.1\ E_{air} + 1.0\ E_{vdW} + 0.2\ E_{elec} + 1.0\ E_{desolv}$

- E_{air}: ambiguous interaction restraint energy
- E_{desolv}: desolvation energy using Atomic Solvation Parameters (Fernandez-Recio et al *JMB* 335, 843 (2004))
- BSA: buried surface area

HADDOCK

Primer: umeščanje trojne vijačnice kolagena tipa 1 v aktivno mesto katepsina K. Omejitve pri katepsinu so ostanki aktivnega mesta, pri kolagenu trije zaporedni ostanki na sredini trojne vijačnice.

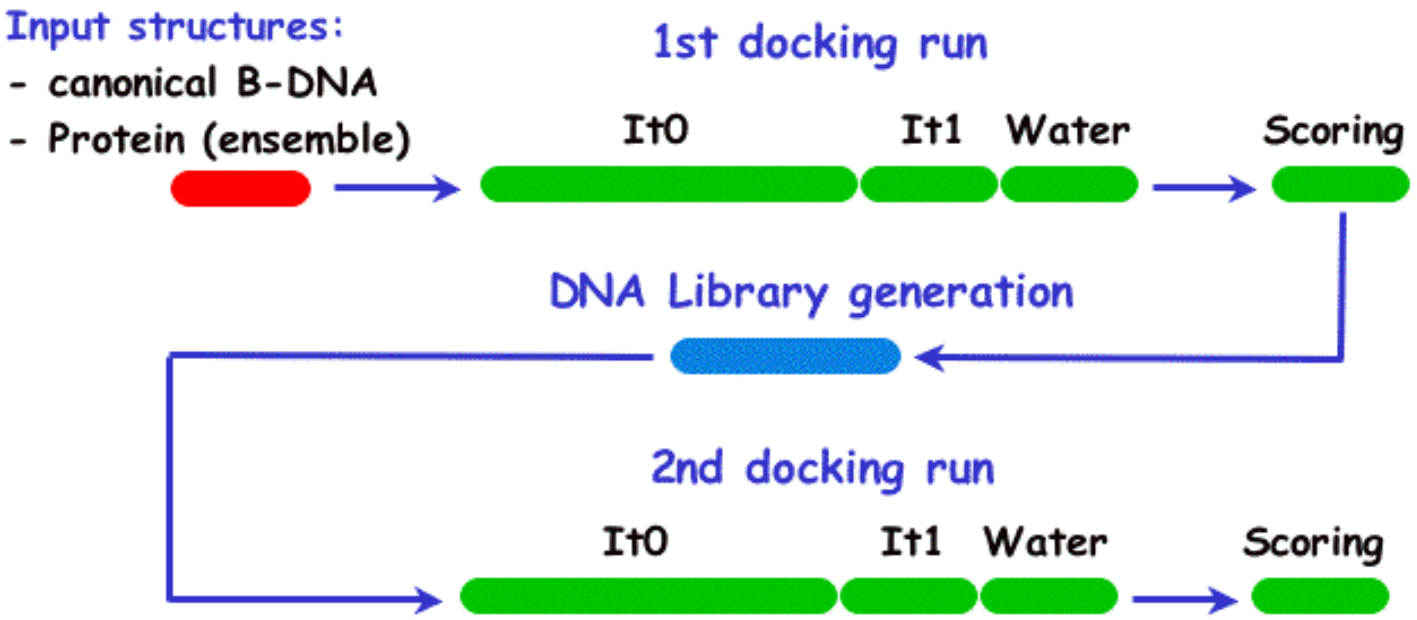




Cluster 1

HADDOCK score	-83.7 +/- 3.2
Cluster size	182
RMSD from the overall lowest-energy structure	2.9 +/- 0.7
Van der Waals energy	**-62.5 +/- 3.5**
Electrostatic energy	**-17.6 +/- 6.2**
Desolvation energy	**-23.0 +/- 3.3**
Restraints violation energy	53.4 +/- 30.69
Buried Surface Area	1431.1 +/- 63.8
Z-Score	-1.0

Protein-DNA HADDOCKing protocol

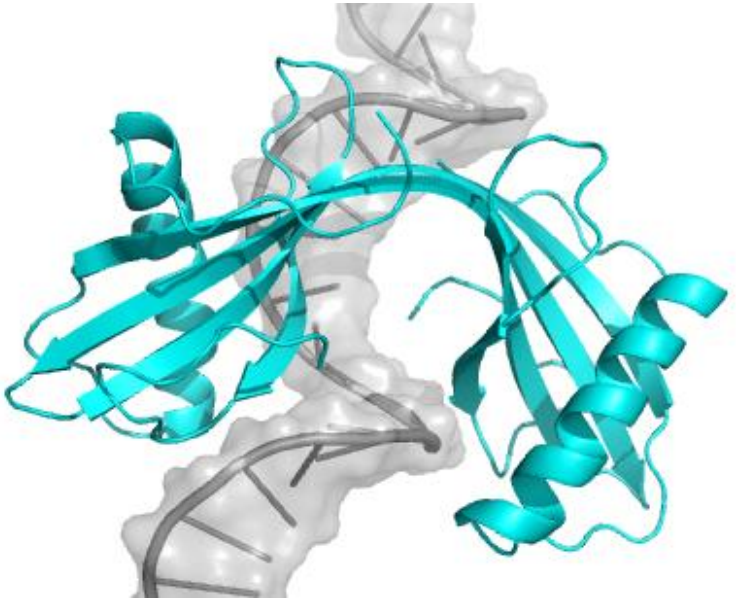


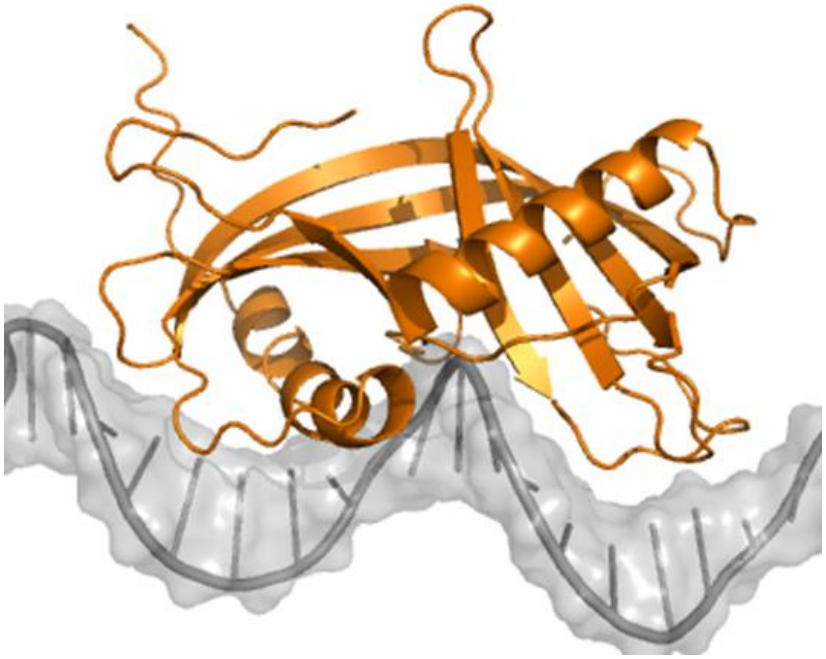
It0: rigid body docking It1: semi-flexible refinement Water: final refinement explicit solvent

Van Dijk *et al. Nucl. Acid. Res.* 2006

HADDOCK

Primer: vezava dimera stefina B na ssDNA (30G). „Eksperimentalne“ omejitve - napoved DNA-vezavnih motivov v stefinu B s temu namenjimi strežniki

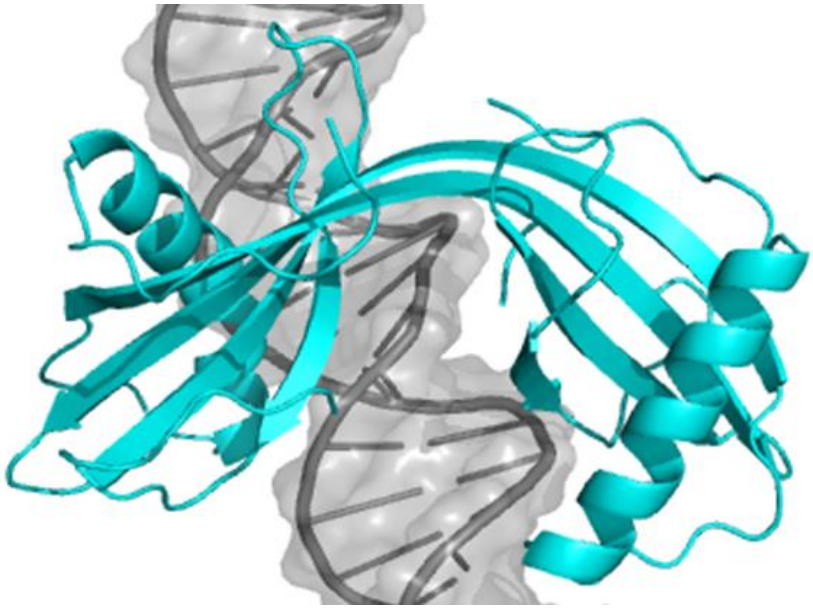


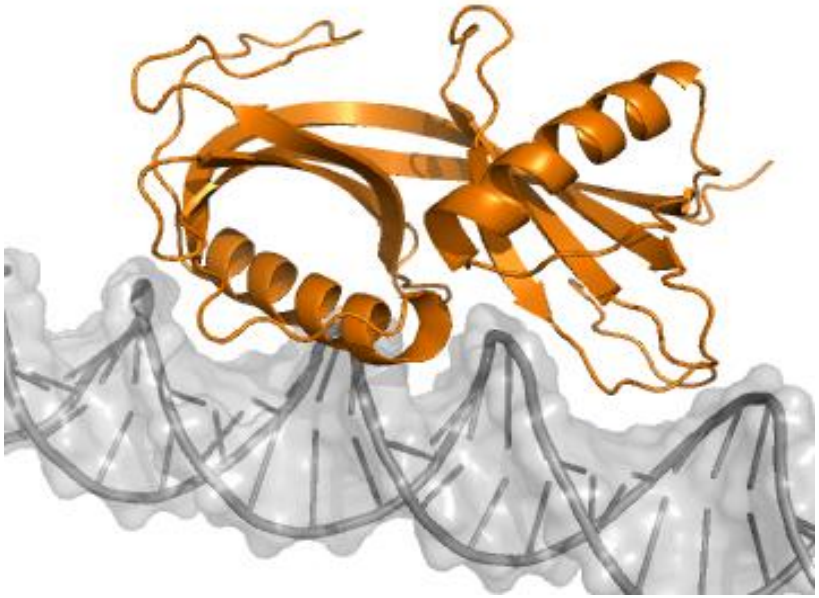


	Kompleks 1 (levo)	Kompleks 2 (desno)
Van der Waalsova energija (kcal/mol)	-87.6 +/- 9.6	-87.0 +/- 4.9
Elektrostatska energija (kcal/mol)	-460.8 +/- 48.5	-531.2 +/- 25.6
Desolvatacijska energija (kcal/mol)	50.6 +/- 10.1	56.3 +/- 6.1

HADDOCK

Primer: vezava dimera stefina B na dsDNA (30GC). „Eksperimentalne“ omejitve - napoved DNA-vezavnih motivov v stefinu B s temu namenjimi strežniki

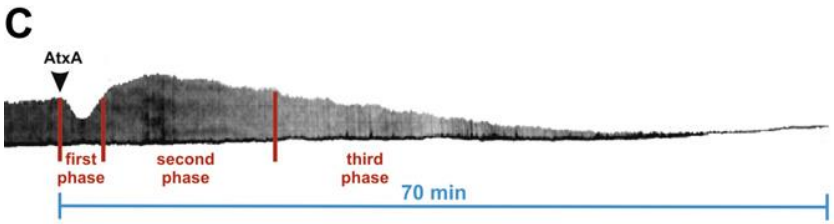


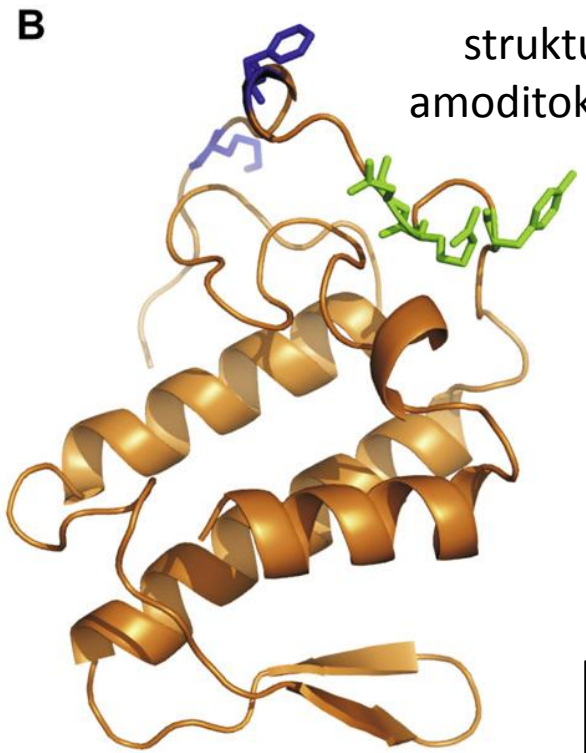


	Kompleks 3	Kompleks 4
Van der Waalsova energija (kcal/mol)	-44.4 +/- 5.6	-56.1 +/- 2.8
Elektrostatska energija (kcal/mol)	-434.6 +/- 39.7	-470.2 +/- 47.7
Desolvatacijska energija (kcal/mol)	48.7 +/- 3.7	69.6 +/- 3.7

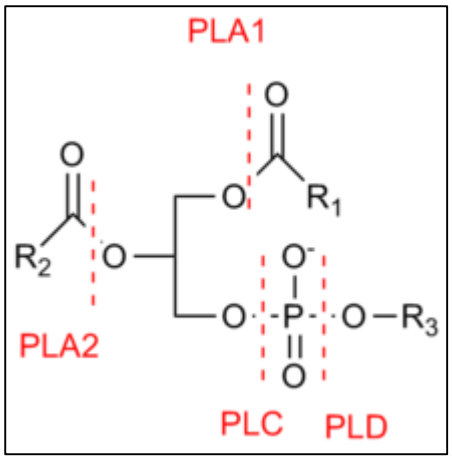
Kompleks amoditoksin/kalmodulin

Amoditoksini so presinaptično nevrotoksične sekretorne fosfolipaze A_2 iz strupa modrasa.





struktura amoditoksinov

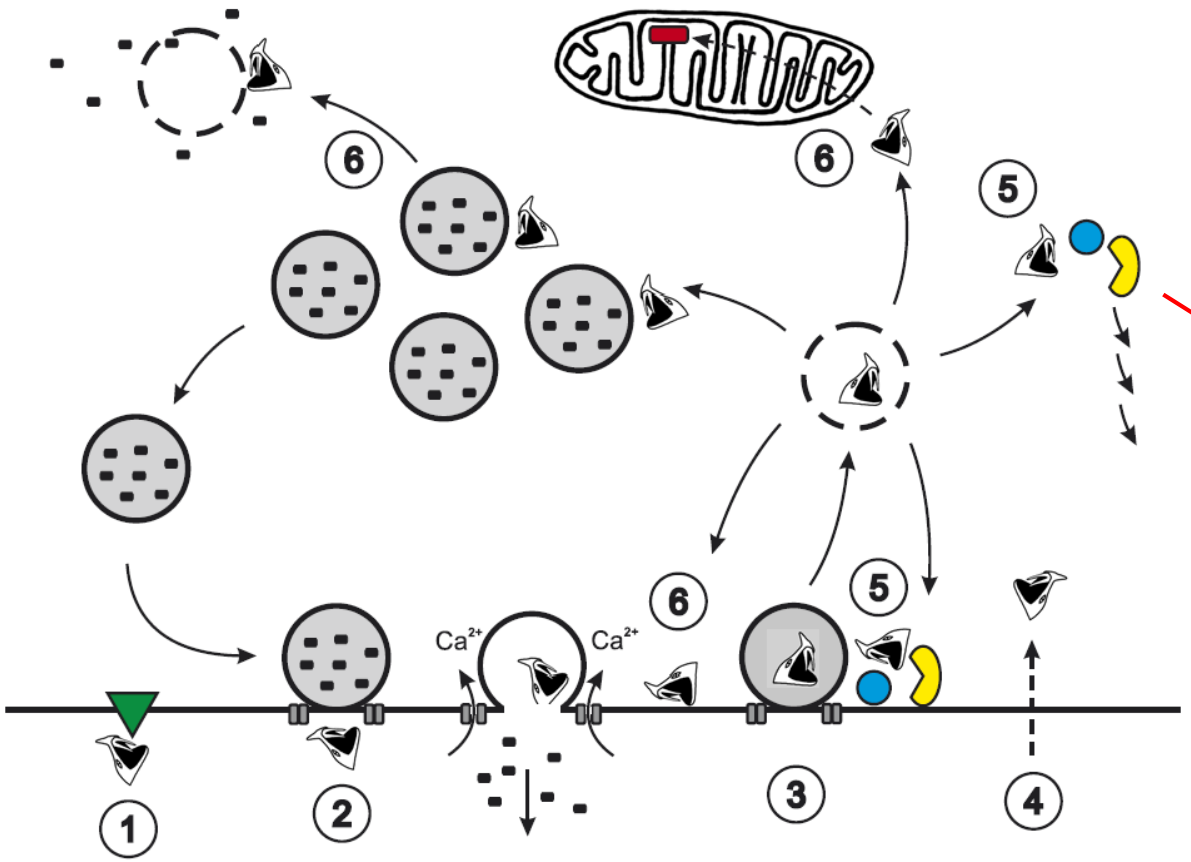


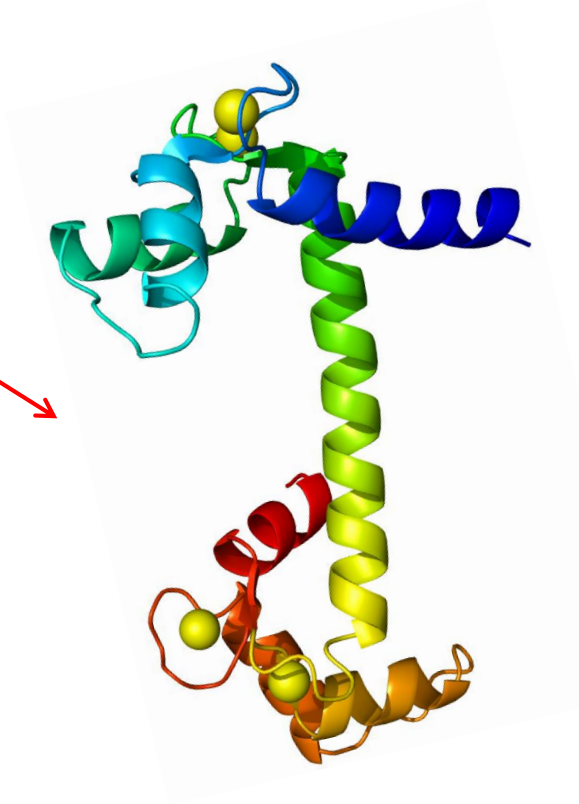
specifičnost fosfolipaz

Križaj (2011) Toxicon. 58

Kompleks amoditoksin/kalmodulin

Amoditoksini so presinaptično nevrotoksične sekretorne fosfolipaze A_2 iz strupa modrasa.

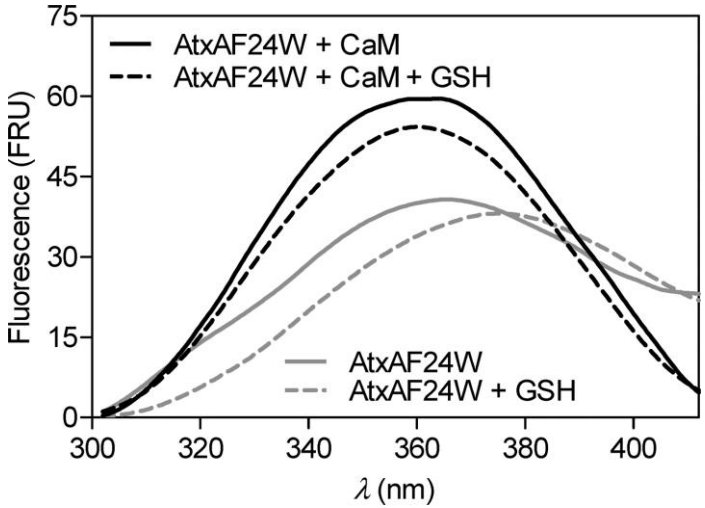




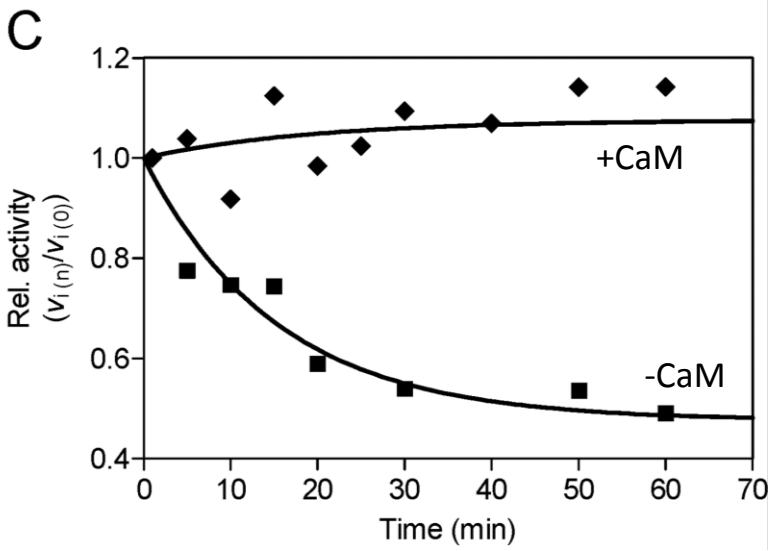
Križaj & Pungerčar (2007) Toxicon. 50

Kompleks amoditoksin/kalmodulin

Kalmodulin v citosolu tarčne celice deluje kot aktivator in stabilizator amoditoksinov.



Fluorescenčni spektri Atx v prisotnosti kalmodulina.

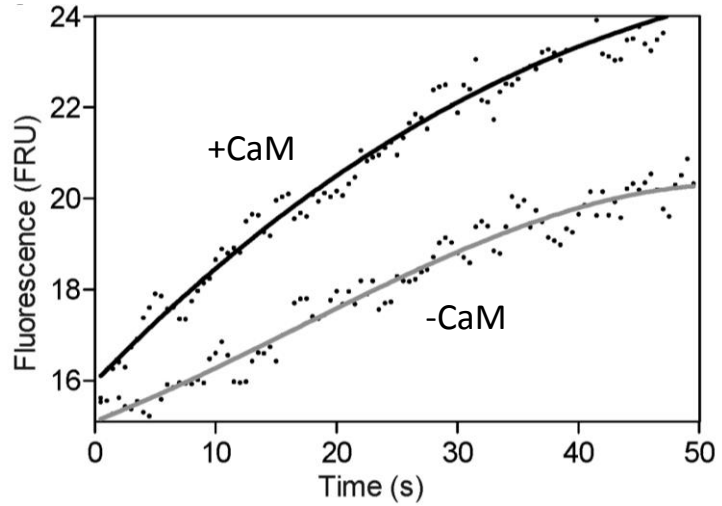


Kalmodulin poviša stabilnost Atx v citosolu.
(Merjena je preostala hitrost hidrolize sintetičnega substrata v citosolu podobnih pogojih)

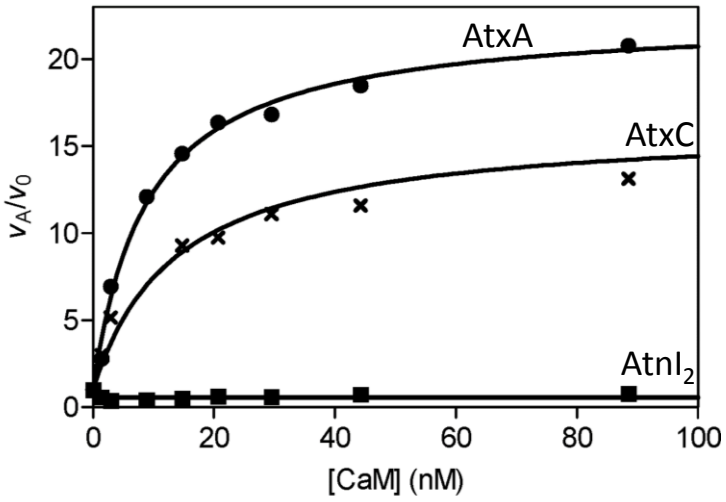
Kovačič et al. (2009) *Biochemistry*. 48(47)

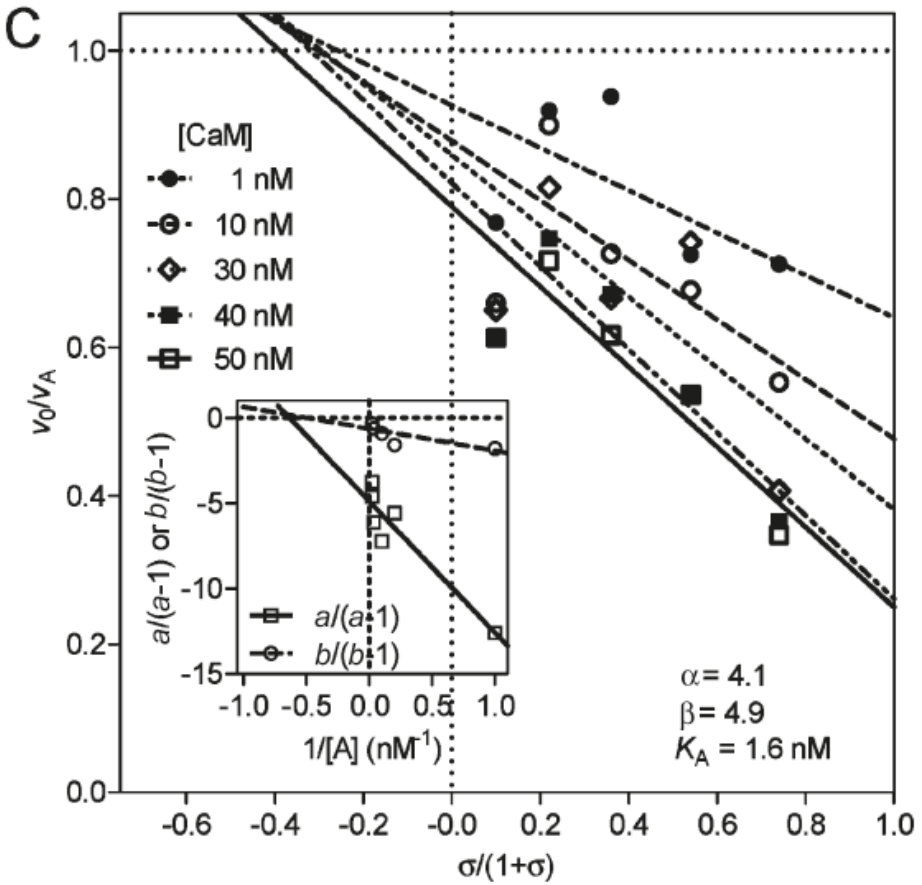
Kompleks amoditoksin/kalmodulin

Kalmodulin v citosolu tarčne celice deluje kot aktivator in stabilizator amoditoksinov.



„fitanje“ krivulj s polinomsko funkcijo za določanje v_0

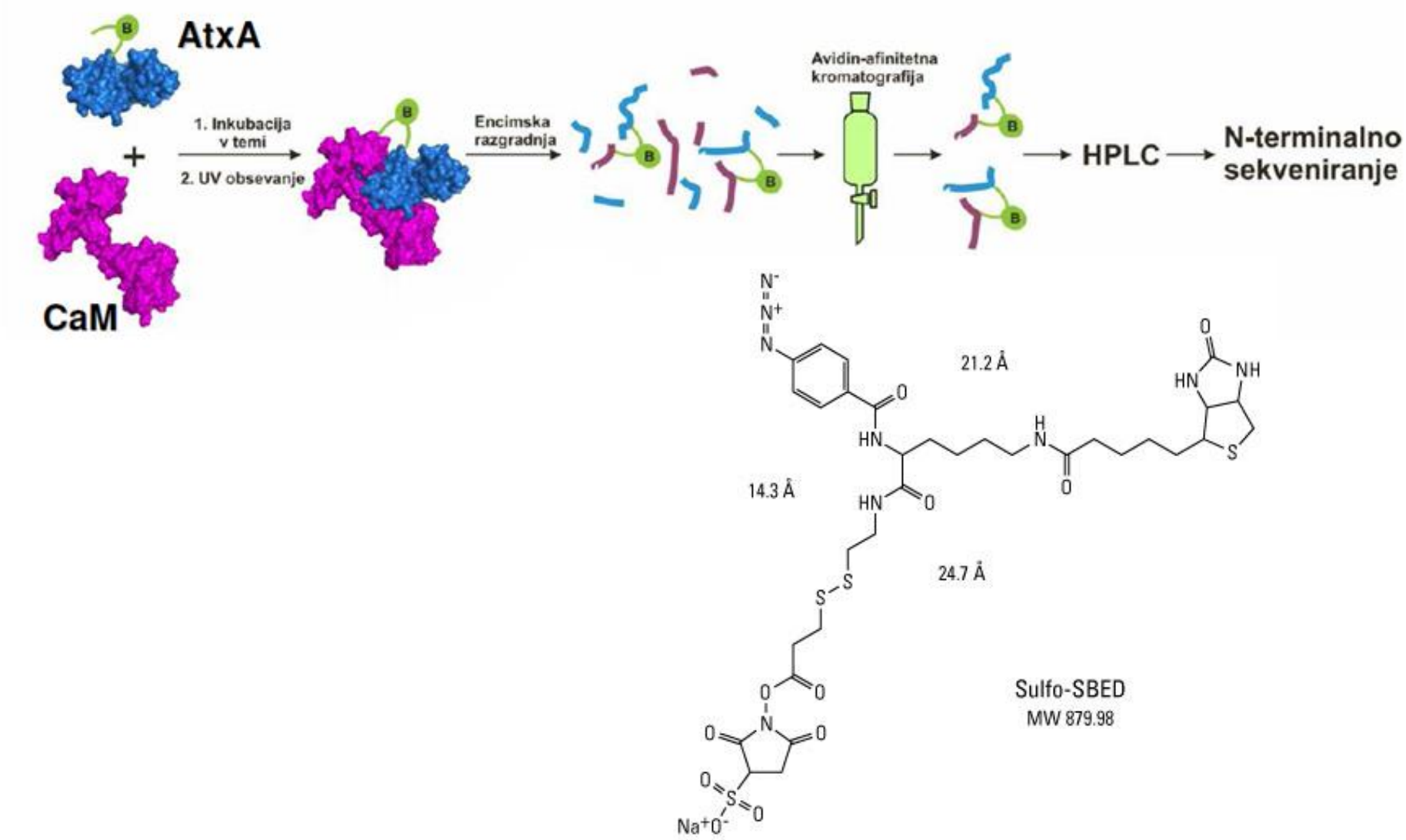




Kovačič et al. (2009) *Biochemistry*. 48(47)

Kompleks amoditoksin/kalmodulin

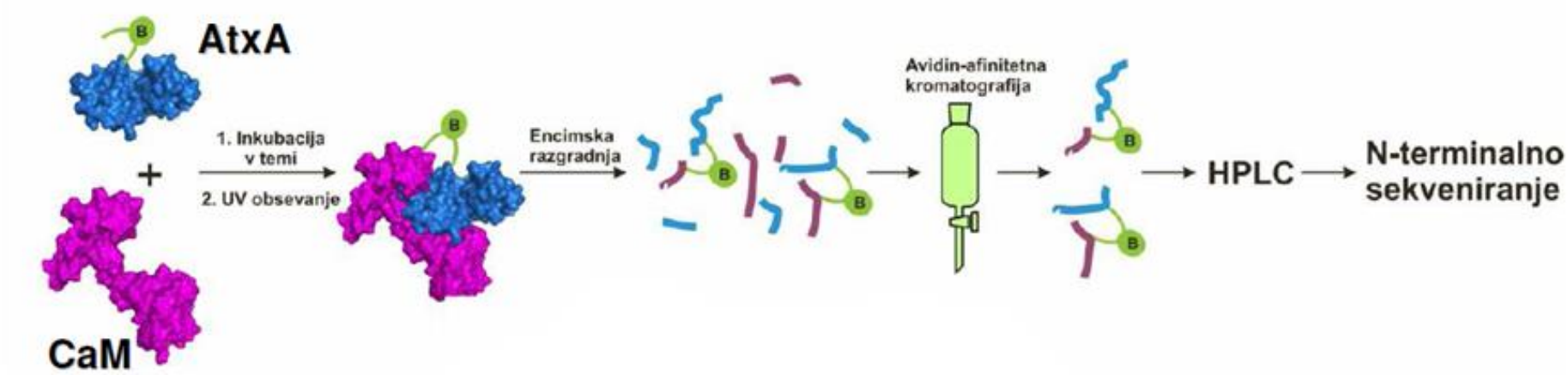
Mapiranje stične površine kompleksa med fosfolipazo amoditoksin A in kalmodulinom:



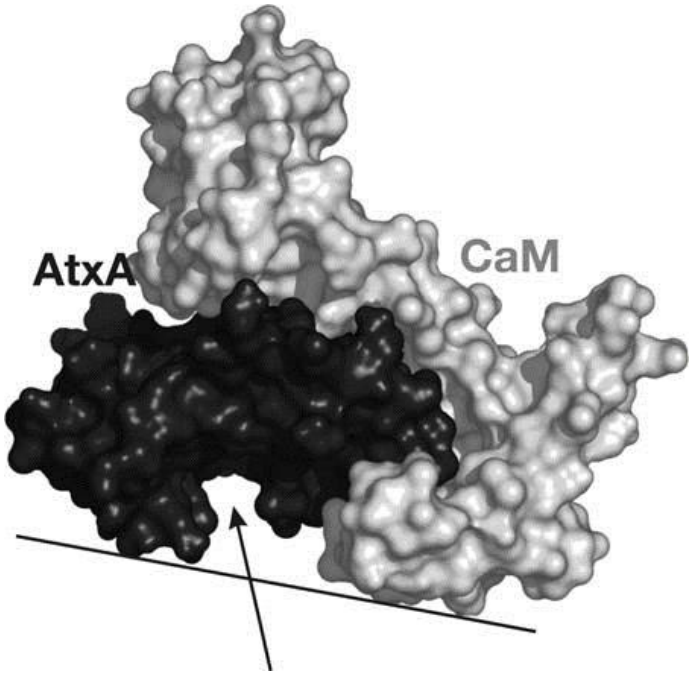
Kovačič et al. 2010. PEDS. 23(6)

Kompleks amoditoksin/kalmodulin

Mapiranje stične površine kompleksa med fosfolipazo amoditoksin A in kalmodulinom:



umestitev s programom Hex – ročna selekcija modelov, ki se skladajo z eksperimentalnimi podatki

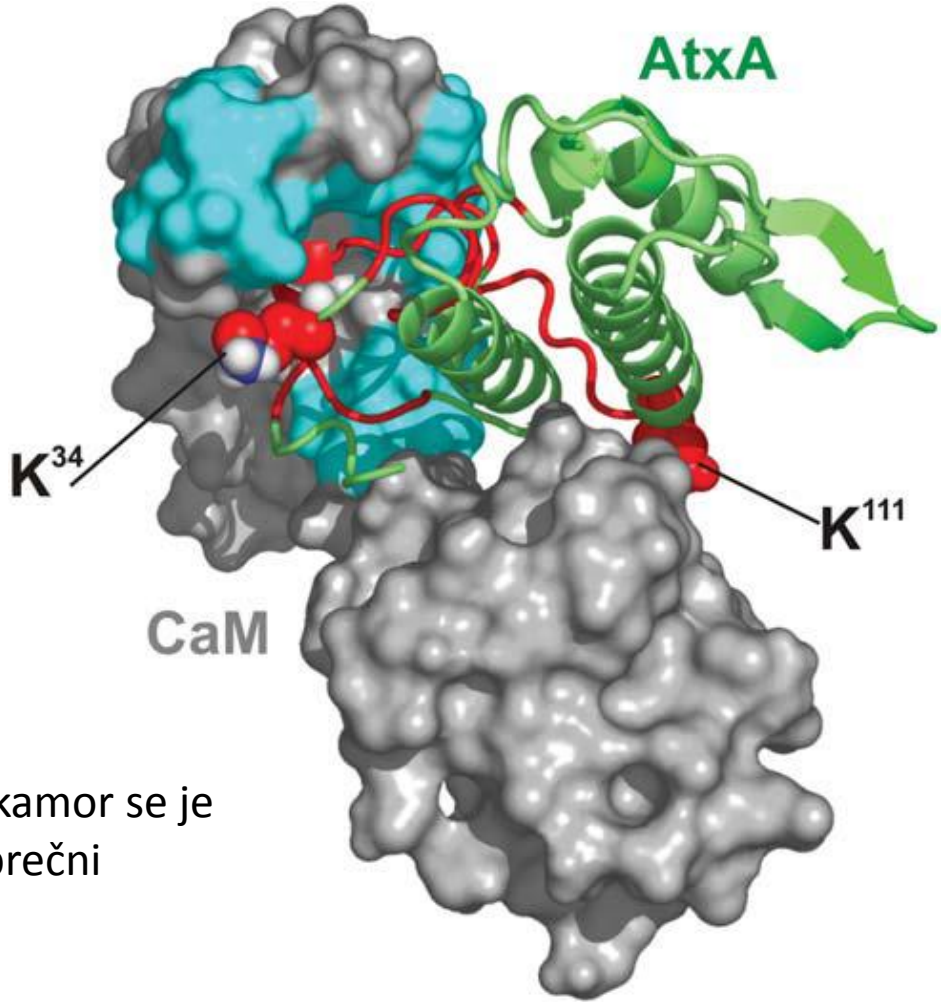


aktivno mesto je dostopno substratu

Kovačič et al. 2010. PEDS. 23(6)

Kompleks amoditoksin/kalmodulin

Mapiranje stične površine kompleksa med fosfolipazo amoditoksin A in kalmodulinom:



označena sta lizina, kamor se je najverjetneje vezal prečni povezovalec.

Kovačič et al. 2010. PEDS. 23(6)

Kompleks amoditoksin/kalmodulin

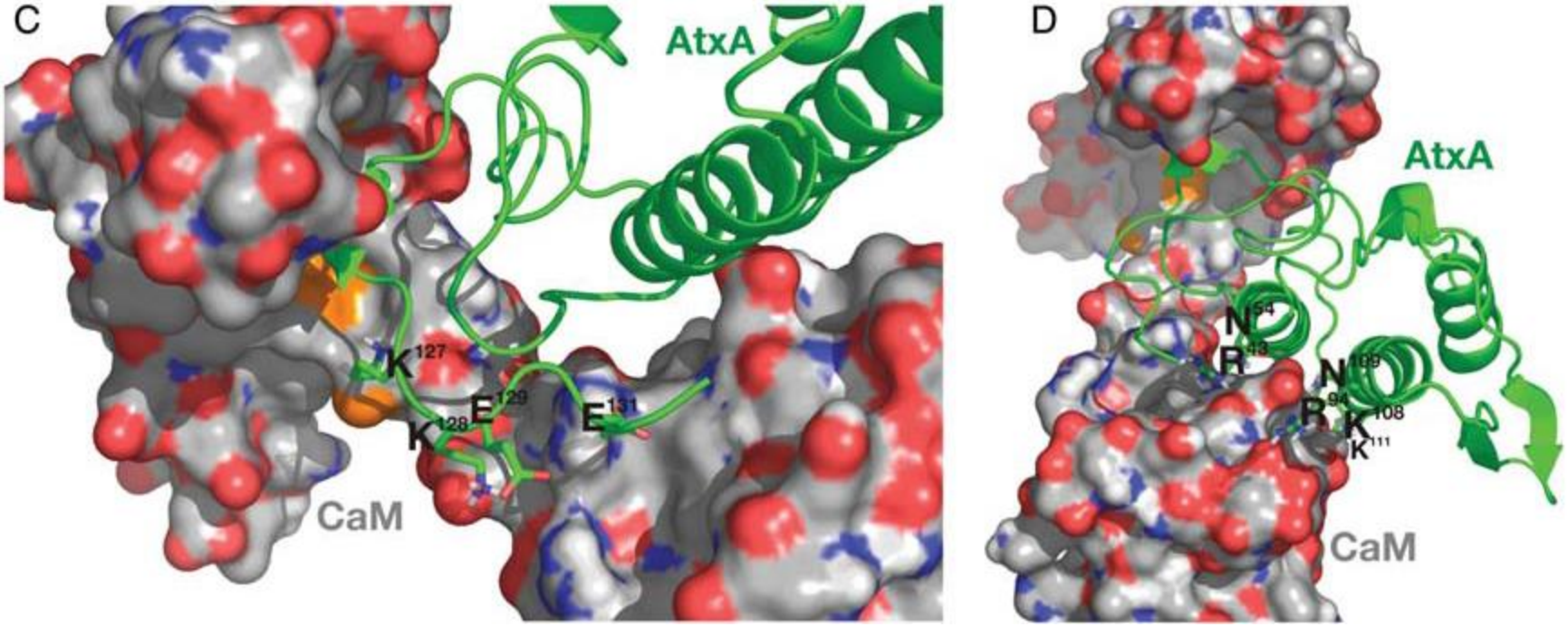


Table I. Interaction of snake venom and mammalian $sPLA_2$s with CaM and its effect on enzyme activity

Protein	Energy of the non-bonded interaction (kJ/mol)		$sPLA_2$–CaM complex formation	
	Protein alone	$sPLA_2$–CaM model	Prediction	Experiment
CaM	−1934			
AtxA	−2865	−5329	Forms	Forms[a]
AtxC	−2868	−5177	Forms	Forms[a]
$AtnI_2$	−2567	+5198	Does not form	Does not form[a]
notexin	−3356	+189 290	Does not form	Not shown
OS_2	−2752	+2526	Does not form	Does not form[a]
crotoxin	−2794	−110	Weak interaction	Weak interaction[a]
β-Butx	−3950	+19 886	Does not form	Does not form[a]
taipoxin	−2494	+5740	Does not form	Does not form[a]
Agtx	−2492	+7623	Does not form	Weak interaction[a]
pGIB $sPLA_2$	−2842	+4209	Does not form	Does not form[a]
huGIIA $sPLA_2$	−2730	+6016	Does not form	Does not form[a]
huGV $sPLA_2$	−2485	−4997	Forms	Forms[b]
mGX $sPLA_2$	−2678	−4564	Forms	Forms[b]

Kovačič et al. 2010. PEDS. 23(6)