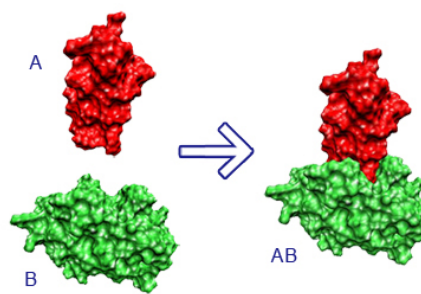
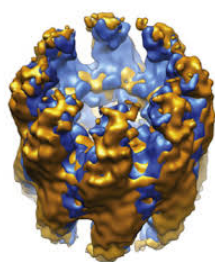


Analiza 3D struktur makromolekul in modeliranje



Miha Pavšič
20. 5. 2014

Analiza 3D struktur makromolekul in modeliranje

Pregled obravnavanih tem:

6. Primerjava 3D struktur makromolekul

- b. Zvitje in klasifikacija struktur ← nadaljevanje od prejšnjic
- c. Primerjanje in iskanje podobnih struktur

7. Metode za določevanje struktur makromolekul (nizka ločljivost)

- a. Ozkokočno sipanje rentgenske svetlobe (SAXS)
- b. Krio-EM

8. Modeliranje makromolekul

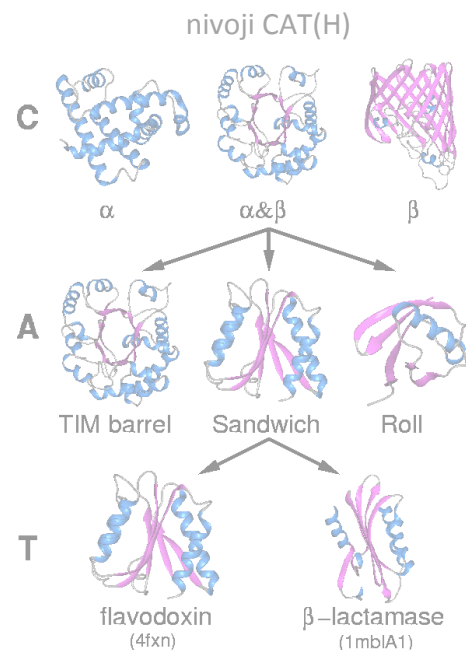
- a. Modeliranje posameznih makromolekul
 - Modeliranje na osnovi homologije (*homology modeling*)
 - Prepoznavanje zvitja (*protein threading*)
 - Modeliranje *ab initio* (*ab initio modeling*)
- b. Modeliranje zank (*loop modeling*)
- c. FoldIt
- d. Baze modelov
- e. CASP (*Critical Assessment of protein Structure Prediction*)
- f. Umestitev (*docking*) in CAPRI (*Critical Assessment of PRediction of Interactions*)
- g. Simulacija molekulske dinamike (*molecular dynamics simulations*)

6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Polipeptidne verige s podobnim ak-zaporedjem zavzamejo podobno 3D strukturo – značilno zvitje (*fold*), ki ga opišemo z zaporedjem in relativno orientacijo elementov sekundarne strukture.

Dve osnovni klasifikaciji:

- **SCOP** - Structural Classification of Proteins (nekoliko zastarela in neosvežena; nova verzija **SCOP2**)
- **CATH** – opis nivojev strukture:
 - CLASS - razred (alfa, beta, mešano alfa/beta, ali pa z malo sekundarne strukture)
 - ARCHITECTURE - arhitektura (celotna oblika glede na orientacijo elementov sekundarne strukture)
 - TOPOLOGY - topologija (družina zvitja)
 - HOMOMOLOGY - homologna superdružina (združuje proteinske domene)



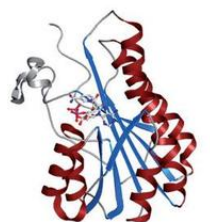
6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Klasifikacija SCOP (Structural Classification Of Proteins): <http://scop2.mrc-lmb.cam.ac.uk>

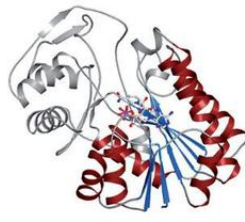
- strukture iz PDB → ročna klasifikacija
- **osnovna enota klasifikacije je domena**
- Pri domenah razlikujemo **zvitja (fold)**: **pri istem zvitju imamo enake glavne elemente sekundarne strukture, so enako urejeni relativno en glede na drugega, enake so tudi povezave med njimi (topologija).**

Primer zvitja: Rossmanovo zvitje (Rossman fold)

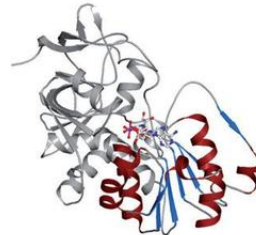
Opis: 3 plasti a/b/c, paralelna β -površina iz 5 trakov (vrstni red 32145)



3 alpha, 20 beta-hydroxysteroid dehydrogenase (*Streptomyces exfoliatus*)



UDP-galactose 4-epimerase (*Escherichia coli*)



alcohol dehydrogenase (*Equus caballus*)



mannitol 2-dehydrogenase (*Pseudomonas fluorescens*)

6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Klasifikacija SCOP (Structural Classification Of Proteins)

Nivoji:

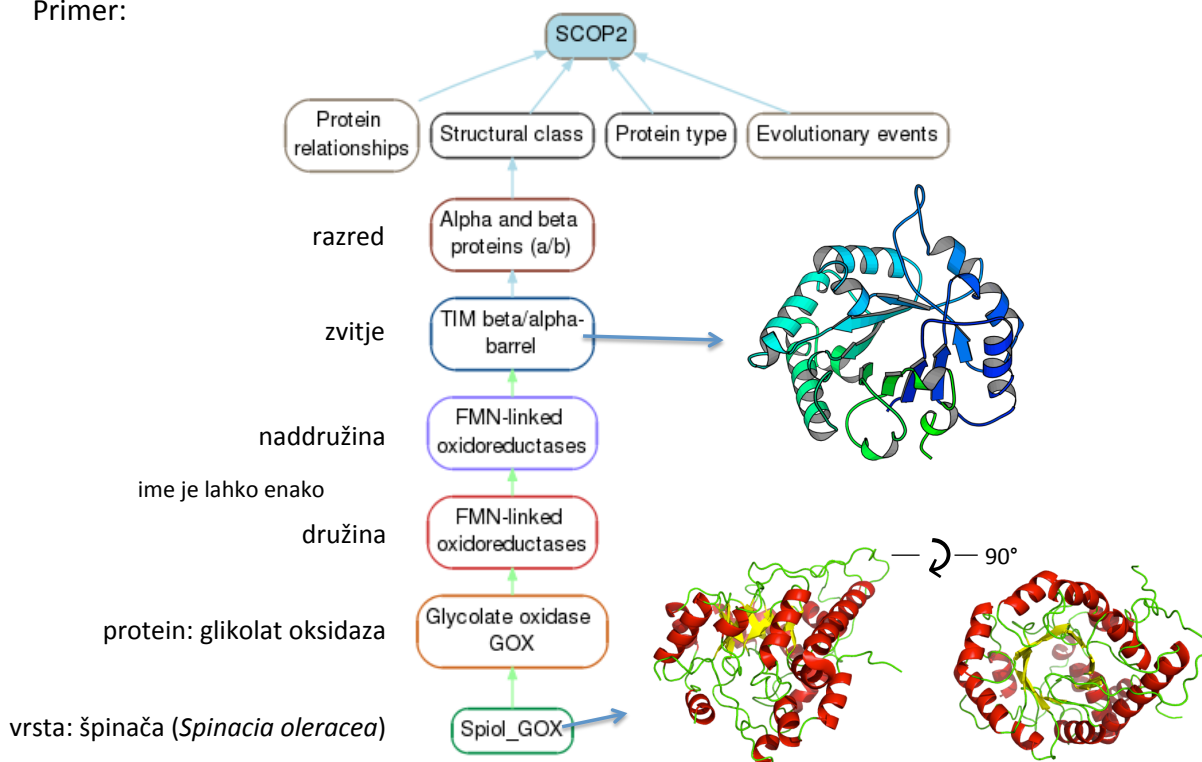
nivo	
RAZRED (class)	splošne lastnosti (npr. samo α , samo β , α/β in $\alpha+\beta$, membranski proteini, majhni proteini, ...), ni evolucijske povezanosti
ZVITJE (fold)	podobna topologija, morda gre za evolucijsko povezanost
NADDRUŽINA (superfamily)	jasna strukturna homolgija *
DRUŽINA (family)	jasna homologija na osnovi zaporedja *
PROTEIN	funkcionalna podobnost
VRSTA (species)	posamezna zaporedja

* Strukturna podobnost evolucijsko oddaljenih struktur je večja kot podobnost njihovih aminokislinskih zaporedij.

6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Klasifikacija SCOP (Structural Classification Of Proteins)

Primer:



6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Klasifikacija CATH (Class-Architecture-Topology-Homology)

Semi-avtomatska klasifikacija

Primer:

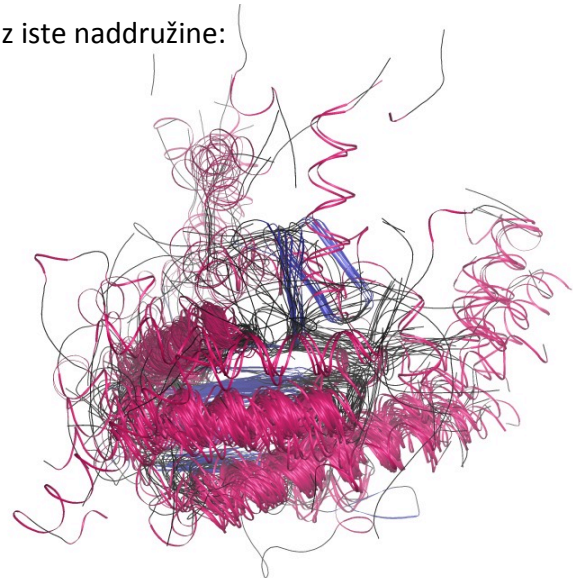
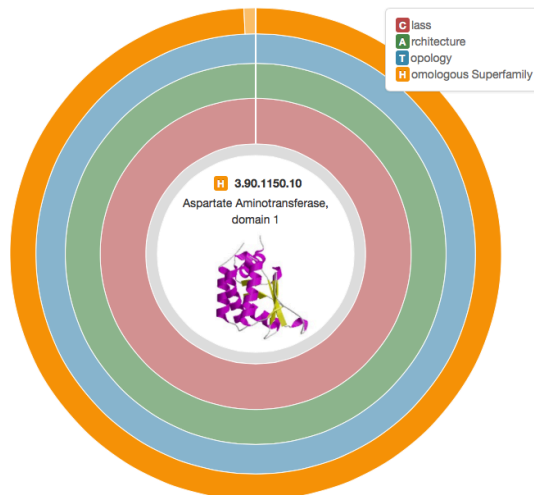
 [Top of CATH hierarchy](#)

C [3 Alpha Beta](#)

A [3.90 Alpha-Beta Complex](#)

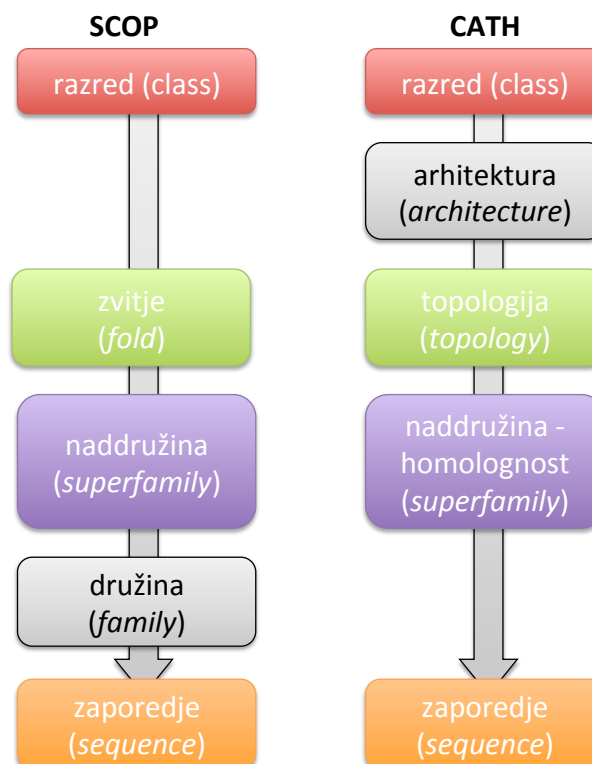
T [3.90.1150 Aspartate Aminotransferase, domain 1](#)

superpozicija struktur iz iste naddružine:



6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

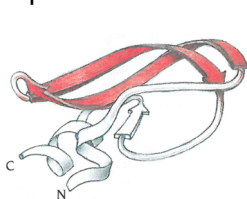
SCOP in CATH – primerjava nivojev



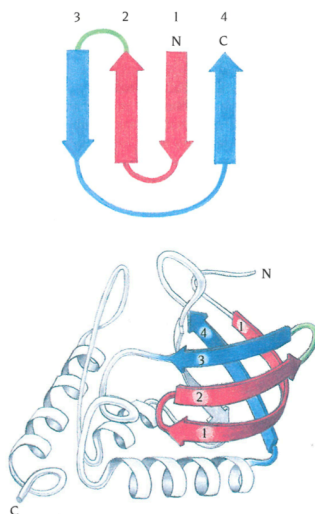
6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Motivi

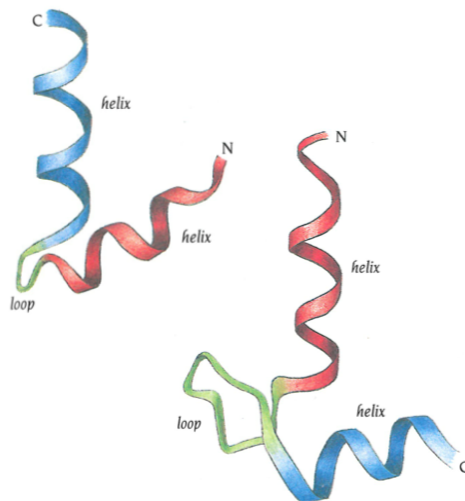
- supersekundarna (nadsekundarna) struktura, ki se lahko pojavlja v različnih zvitjih
- primeri:



β -lasnica
(β -hairpin)



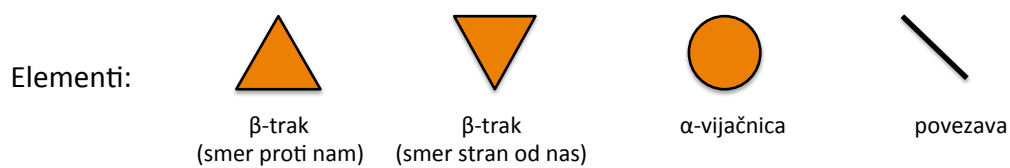
grški ključ
(Greek key)



vijačnica-zavoj-vijačnica
(*helix-turn-helix, HTH*)

6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Shematski prikaz topologije – diagram TOPS



antiparalelna β -površina

strukturo orientiramo tako, da so β -trakovi pravokotno na ravnino/papir/ekran; če se povezava začne za objektom, to pomeni od spodaj, če je pa pred njim, pa od zgoraj



6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

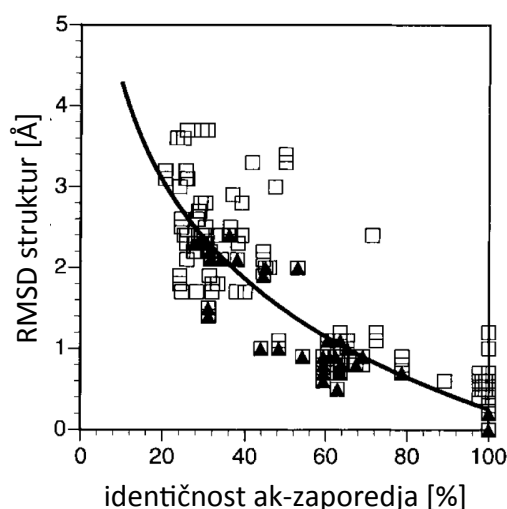
Povezava podobno zaporedje – podobna struktura

HOMOLOGIJA: 20-30 % identičnost (na splošno)

Pozor – odvisno od dolžine zaporedij!

- poravnave kratkih zaporedij imajo pogosto >30% identičnost
- zaporedja, krajša od 10 ak-ostankov – 100% identičnost ni signifikantna

Primer za homologna zaporedja modrih baker-vezavnih proteinov:



6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Povezava podobno zaporedje – podobna struktura

Primer: tiroglobulinske domene (TY) različnih človeških proteinov (1/2)

Poravnava zaporedij:

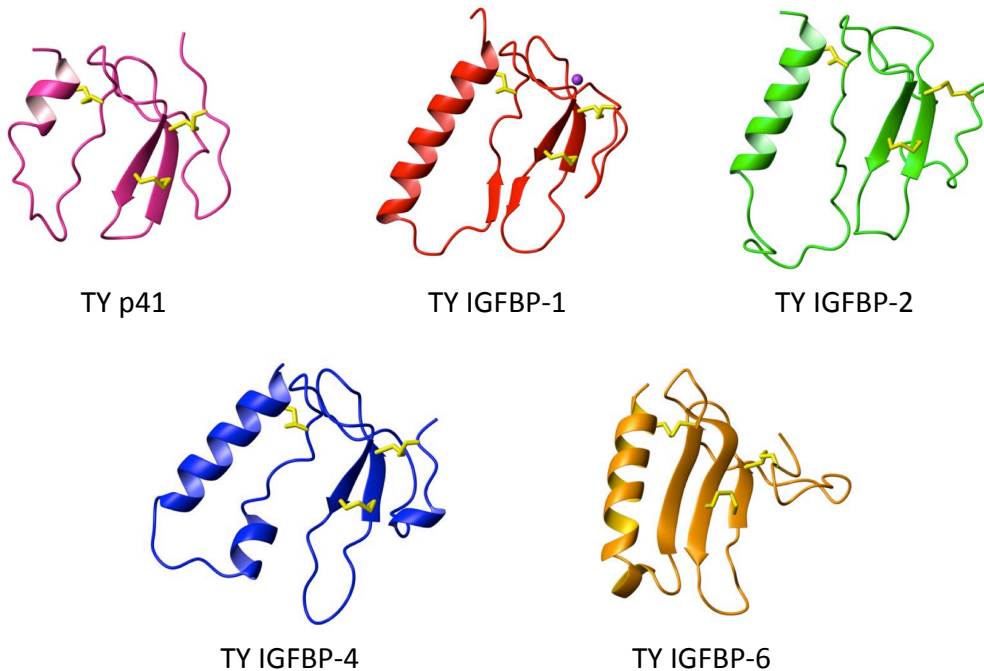
tiroglobulin [34-92]	CELQRETAFLK-----QADYVPQCAEDGSFQTVQQND---GRSCWCVG-ANGSEVLGSRQP-G-RPVAC
tiroglobulin [96-160]	CQLQKQKILLSGYIN-----STDTSYLPQCQDSGDYAPVQCDVQ---QVQCWCVD-AEGMEVYGRQL-G-RPKRC
nidogen-1 [849-919]	CQHEREHILGAAGATD-----PQRPPIPLFVPECDAHGHYAPTOCHGS---TGYSWCVD-RDGREVEGTRTRPGMTPP-C
nidogen-2 [940-1005]	CEQQQRHAQAQYAYPG-----AR-----FHIPQCDEQGNFLPLQCHGS---TGFQCWVD-PDGEHEVPGTTPPGSTPPHC
nidogen-2 [1019-1084]	CERWRENLLHYGGT-----PR---DDQYVPQCDDLGHFIPLQCHGK---SDFCWCVD-KDGREVQGTRSQPGTTPA-C
SMOC-1 [95-158]	CRLEQAQALEQAKK-----PQEA VFPVECGEDGSFTQVQCHTY---TGYSWCVTPD-GKPISGSSVQNK--TPVC
SMOC-1 [227-292]	CDQERQSALEEAQQN-----PREGIVIPCEAPGGLYKPVQCHQS---TGYSWCVLVDTRPLPGTSTRYV--MPSC
SMOC-2 [190-153]	CVAERKYTQEQARK-----EFQVQVFIPECNDDGTYSQVQCHSY---TGYSWCVTPN-GRPISGTAVAHK--TPRC
SMOC-2 [216-281]	CDQEHSQSALEEAQK-----KNDNVVIPECAHGGLYKPVQCHPS---TGYSWCVLVDTRPLPGTSTRYE--QPKC
testikan-1 [313-376]	CQNEMNRIQKLSKGG-----SLGAFIPRCNEEGYKATQCHGS---TGQCWCVD-KYGNELAGSRKQ-G--AVSC
testikan-2 [313-376]	CLAELERIQIQEAAK-----KKPGIFIPSCDEEDGYRKMCDQS---SGDCWCVD-QLGLELTGRTRH-G--SPDC
testikan-3 [317-380]	CQTELSNIQKRQGVK-----KLLGQYIPLCDEDGYKPTQCHGS---VGQCWCVD-RYGNVEMGSRIN-G--VADC
Trop-1 [66-135]	CLVMKAEMNGSKLGR-RAKP-EGALQNNGLYDPCDESLGFLKAKQNGT---SMCWCVN-TAGVRRTDK-DTEIT---C
Trop-2 [73-145]	CLLLKARMSAPKNARTLVRPSEHALVDNDGLYDPCDEPGEFRKARQCNQT---SVCWCVN-SVGVRRTDKGDLSLR---C
IGFBP-1 [176-251]	CRIELYRVVESLA--KAQET-SG--EESKIFYLPNCNKNFGYHSRQETSMDGEAGLCWCVYPWNGKRIPGSPPIRG--DPNC
IGFBP-2 [230-309]	CQQLDQVLERISTMRLPDE-RGPLEHLYSLHIPNCDKHGLYNLKKCKMSLNGQRGECWCVNPNNTGKLIQGAPTIRG--DPEC
IGFBP-4 [174-249]	CQSELHRALERLA----ASQ-SRTHEDLYIIPINCDRNGNFHPKQCHPALDGGQRGKWCVDRKTGVKLPGGLEPKG--ELDC
IGFBP-3 [213-285]	CRREMEDTLNHLK-----FLNVLSPRGVHIPNCDKKGFKYKPKCRPSKGRKRGFCWCVD-KYQQLPGYTTKGE-DVHC
IGFBP-5 [191-263]	CRRHMEASLQELK-----ASPRMVPRAYVLPNCDRKGFKYKPKCRPSRGRKRGICWCVD-KYGMKLPMEYVDG--DFQC
IGFBP-6 [163-234]	CRRHLDSVLQQLQ-----TEVYRGAQTLVYVPCDHRGFYKPKCRSSQGGRRGFCWCVD-RMGKSLPGSPDNG--SSSC
p41 [213-271]	CQEEVSHIPAVH-----PGSFPRKCDENGNLPLQCYGS---IGYWCVVF-PNGTEVPNTRSR-G--HHNC
	* * * * *
	CQ+E+ER+LEQ+K+K-LA-P-+G+LE+LDG+YIPNCDDEDG+YKPVQCHGSLGG+TGYSWCVDPK+GRE+PGTRTRPGTTPP+C

6.b. Primerjava 3D struktur makromolekul: ZVITJE IN KLASIFIKACIJA STRUKTUR

Povezava **podobno zaporedje – podobna struktura**

Primer: tiroglobulinske domene (TY) različnih človeških proteinov (2/2)

Strukture:



6.c. Primerjava 3D struktur makromolekul: PRIMERJANJE IN ISKANJE PODOBNIH STRUKTUR

Glavne metode za primerjanje struktur (poravnava struktur, *structural alignment*):

- **superpozicija rigidnih objektov** (*rigid body superposition*)

translacija
večih struktur

- ponavadi transliramo COM (*center-of-mass*)
- pomaknemo na isto izhodišče

identifikacija
ekvivalentnih
položajev

- s pomočjo poravnave ak-zaporedij
- ročno (npr. aktivno mesto)

rotacija
struktur
(relativno)

- optimalna poravnava ekvivalentnih položajev

- **metoda na osnovi elementov sekundarne strukture**

teorija: delecije/insercije se zgodijo predvsem v zankah na površini
implementirana npr. v GRATH, SSM; precej hitra metoda

identifikacija
el. sek. str.

diagram razdalj
in kotov med
njimi

primerjava

- **metoda na osnovi razdalj med ak-ostanki in kontakti med njimi**

implementirana v npr. DALI, CE, SSAP

strukture

fragmenti

lokalna geometrija,
razdalje, medsebojni
kontakti

združitev in
optimizacija

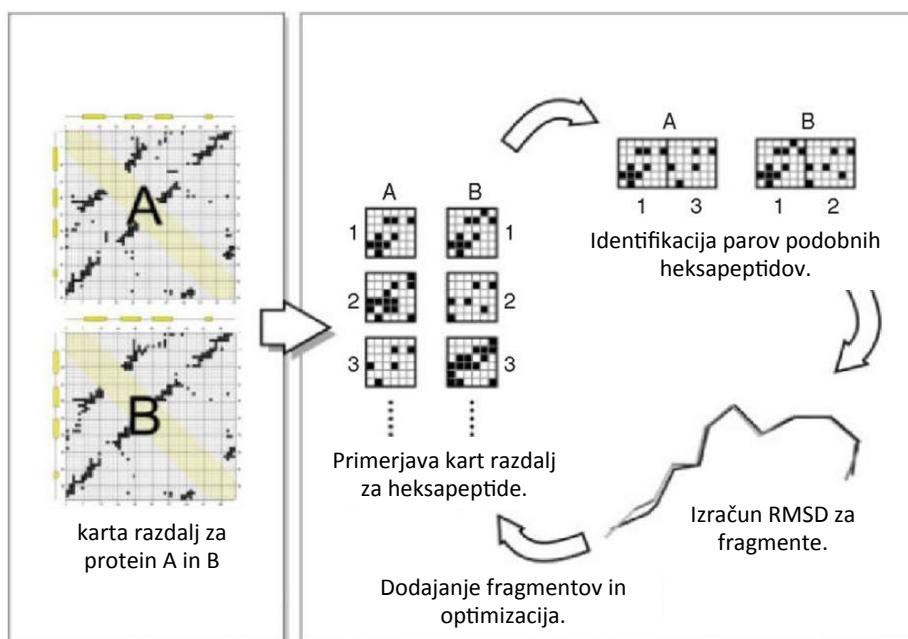
6.c. Primerjava 3D struktur makromolekul: PRIMERJANJE IN ISKANJE PODOBNIH STRUKTUR

DALI (Distance mAtrix aLignment)

http://ekhidna.biocenter.helsinki.fi/dali_server/

Metoda na osnovi razdalj med ak-ostanki in kontakti med njimi

Izračun razdalj med ak-ostanki → karta razdalj → primerjamo razdalje znotraj heksapeptidov.



<http://what-when-how.com/proteomics/structure-comparison-and-protein-structure-classifications-proteomics/>

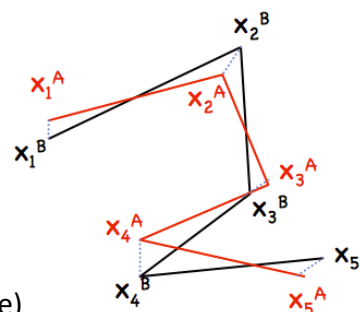
6.c. Primerjava 3D struktur makromolekul: PRIMERJANJE IN ISKANJE PODOBNIH STRUKTUR

Vrednotenje podobnosti struktur

Vrednost RMSD (RMSD score)

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2}$$

- enota dolžine (Å)
- močno odvisna od dolžine
- občutljiva za kratke nepodobne fragmente (npr. zanke)



Vrednost Z (Z-score)

- statistični model
- koliko st. deviacij je vrednost poravnave od povprečja

Vrednost GDT (Global Distance Test Total Score, GDT TS)

- glavna ocena kvalitete modelov pri CASP ("tekmovanje" v modeliranju struktur)
- definirana je kot največje št. atomov C_{α} modela, ki so oddaljeni za manj kot X od njihovih položajev v eksperimentalni strukturi
- izračuna se za več vrednosti X
- konvergenca: št. atomov se z naraščanjem X ne spremeni

6.c. Primerjava 3D struktur makromolekul: PRIMERJANJE IN ISKANJE PODOBNIH STRUKTUR

PDBeFold (SSM) <http://www.ebi.ac.uk/msd-srv/ssm/>



- ne upošteva ak-zaporedja oz. poravnave ak-zaporedij
- princip: ujemanje sekundarnih struktur (SSM – Secondary Structure Matching)
- vrednotenje: vrednost Q (poravnava C_{α}), P (RMSD + dolžina poravnave + vrzeli + ujemanje v elementih sek. strukture + povezave med el. sek. strukture), Z (statistična metoda), RMSD in % identičnosti ak-zaporedja
- uporaba:
 - primerjava dveh ali večih struktur
 - primerjava strukture s strukturami v PDB ali bazo SCOP

##	Scoring			RMSD	Nalign	N _{seg}	%seq	Query	Target (PDB entry)	Title		
	Q	P	Z									
1	1.00	28.8	16.0	0.00	96	0	100	100	100	100	96	DETERMINATION AND RESTRAINED LEAST-SQUARES REFINEMENT OF THE CRYSTAL STRUCTURES OF RIBONUCLEASE SA AND ITS COMPLEX WITH 3'-GUANYLIC ACID AT 1.8 ANGSTROMS RESOLUTION
2	1.00	12.2	10.5	0.10	96	0	100	86	100	100	96	COMPLEX OF RIBONUCLEASE FROM STREPTOMYCES AUREOFACIENS WITH 2'-GMP AT 1.7 ANGSTROMS RESOLUTION
3	1.00	12.6	10.7	0.13	96	0	99	86	100	100	96	HYDROLASE, GUANYLORIBONUCLEASE

7. Metode za določevanje struktur makromolekul (nizka ločljivost)

Metode nizkih ločljivosti so posebej pomembne pri raziskovanju **velikih kompleksov**, ki jih z rentgensko kristalografijo in NMR ne moremo proučevati → **omejitve za umestitev**.

ločljivost tipično < 3 Å

Rentgenska kristalografija (X-ray)

Jedrsko magnetna resonanca (NMR)

struktura podenot

struktura podenot

oblika podenot

oblika podenot

stiki med poenotami

stiki med poenotami

razdalje med podenotami

razdalje med podenotami

stehiometrija podenot

stehiometrija podenot

simetrija kompleksa

simetrija kompleksa

oblika kompleksa

oblika kompleksa

struktura kompleksa

struktura kompleksa

ločljivost > 3 Å

Ozkokotno sipanje rtg. svetlobe (SAXS)

Krioelektronska mikroskopija (Cryo-EM)

oblika podenot

oblika podenot

stiki med poenotami

stiki med poenotami

razdalje med podenotami

razdalje med podenotami

stehiometrija podenot

stehiometrija podenot

simetrija kompleksa

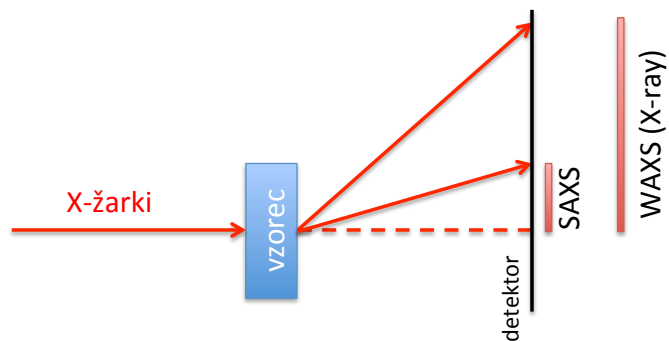
simetrija kompleksa

oblika kompleksa

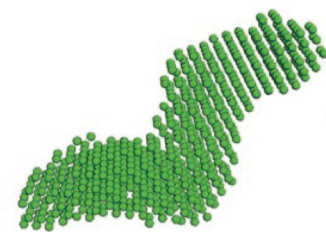
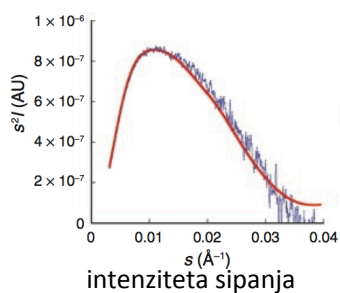
oblika kompleksa

7.a. Metode za določevanje struktur makromolekul (nizka ločljivost): SAXS

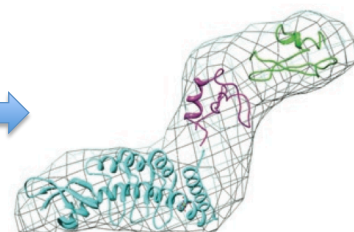
Ozkokotno sipanje rentgenske svetlobe (*Small Angle X-ray Scattering, SAXS*)



- makromolekula je v raztopini
- dobimo podatke o obliki kompleksa
- gradnja modela z umeščanjem rigidnih teles (struktur podenot)



dummy atom model



rekonstrukcija

DOI 10.1038/nsmb.1836

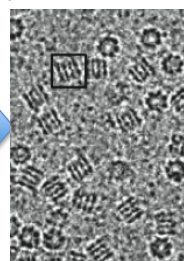
7.b. Metode za določevanje struktur makromolekul (nizka ločljivost): Krio-EM

Krio-elektronska mikroskopija (*Cryo-EM*)

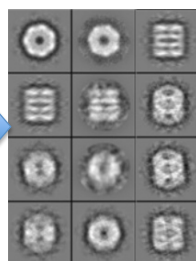
- kriogeni pogoji (lqN₂, lqHe), transmisijski elektronski mikroskop
- običajna EM: za doseg visoke ločljivosti bi potrebovali močan tok elektronov → poškodbe vzorca
- metode:
 - krio-elektronska kristalografija (sipanje elektronov na urejenih delcih, skoraj atomska ločljivost)
 - krio-elektronska tomografija (posamezen delec slikamo v različnih orientacijah → rekonstrukcija)



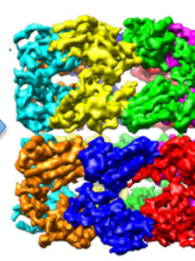
EM



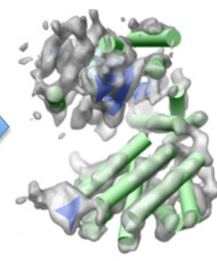
surovi podatki



procesiranje

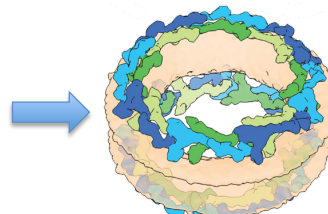
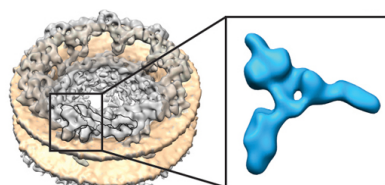


rekonstrukcija



analiza in gradnja modela

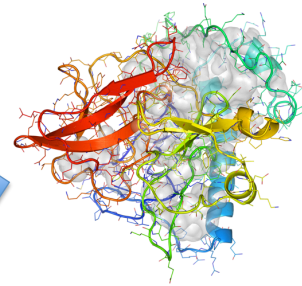
jedrna pora (EMBL, Beck):



8.a. Modeliranje makromolekul: MODELIRANJE POSAMEZNIH MAKROMOLEKUL

Velikokrat nimamo na voljo modela strukture (atomske ločljivosti) na osnovi eksperimentalnih podatkov (X-ray, NMR) → modeliranje.

- fizikalno-kemijske omejitve (dolžine in koti med vezmi, ...)
- napoved sekundarne strukture
- strukture homolognih proteinov
- druge omejitve



...AHGTYCSPLVRRREQNTSGALYW...

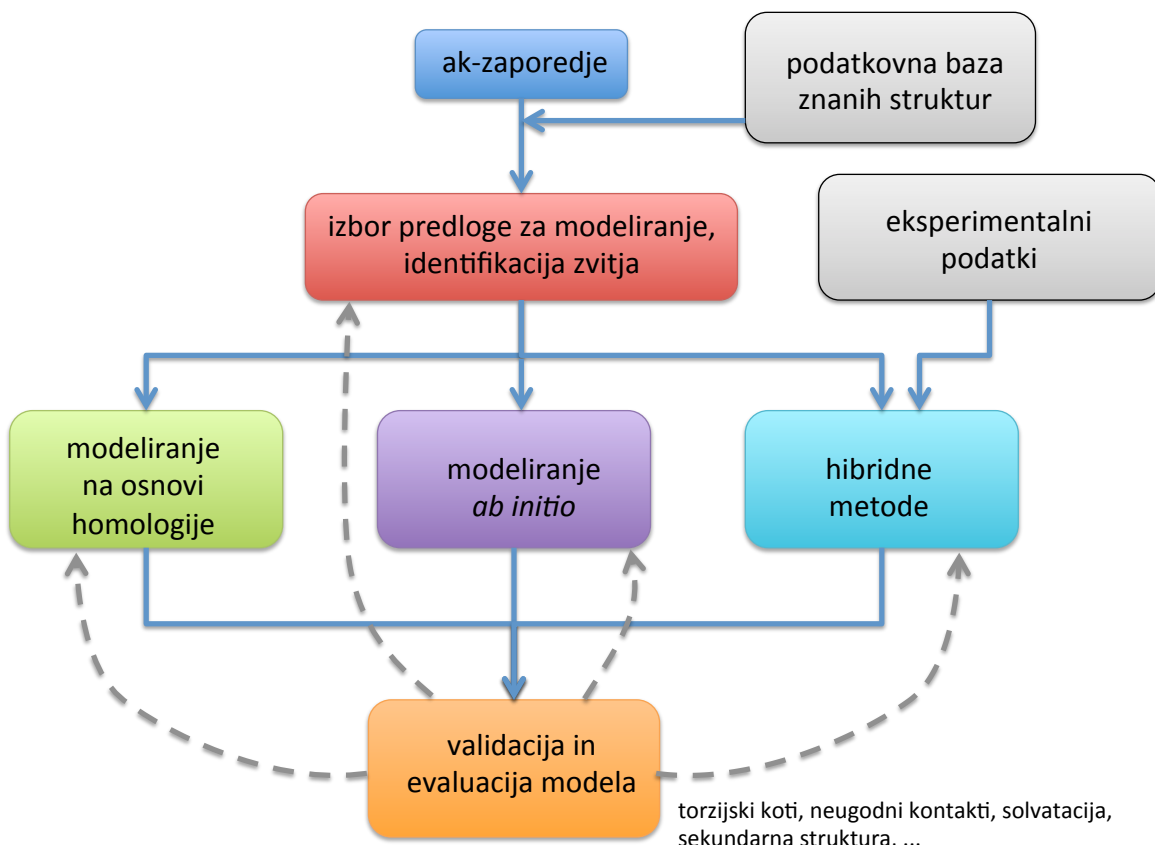
ak-zaporedje

model

Osnovni pristopi za gradnjo modelov makromolekul:

- **modeliranje na osnovi podobnosti**
 - modeliranje na osnovi homologije (homology modeling)
 - identifikacija zvitja (protein threading)
- **modeliranje *ab initio* oz. *de novo* ("od začetka")**
 - metode, osnovane na enegijskih potencialih
 - metode, osnovane na iskanju fragmentov
 - evolucijske metode za napoved stikov v prostoru

8.a. Modeliranje makromolekul: MODELIRANJE POSAMEZNIH MAKROMOLEKUL



8.a. Modeliranje makromolekul: MODELIRANJE POSAMEZNIH MAKROMOLEKUL

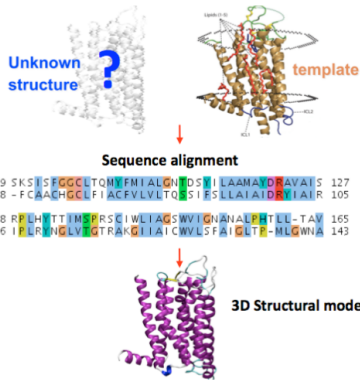
Modeliranje na osnovi homologije (*homology modeling, comparative modeling*)

Dejstva:

- 3D-struktura je določena z ak-zaporedjem (načeloma).
- Evolucijsko gledano so strukturne spremembe manjše kot spremembe z njo povezanega ak-zaporedja.

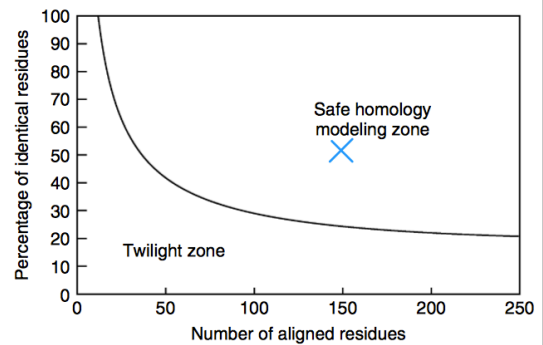


podobno zaporedje → podobna struktura



Koraki:

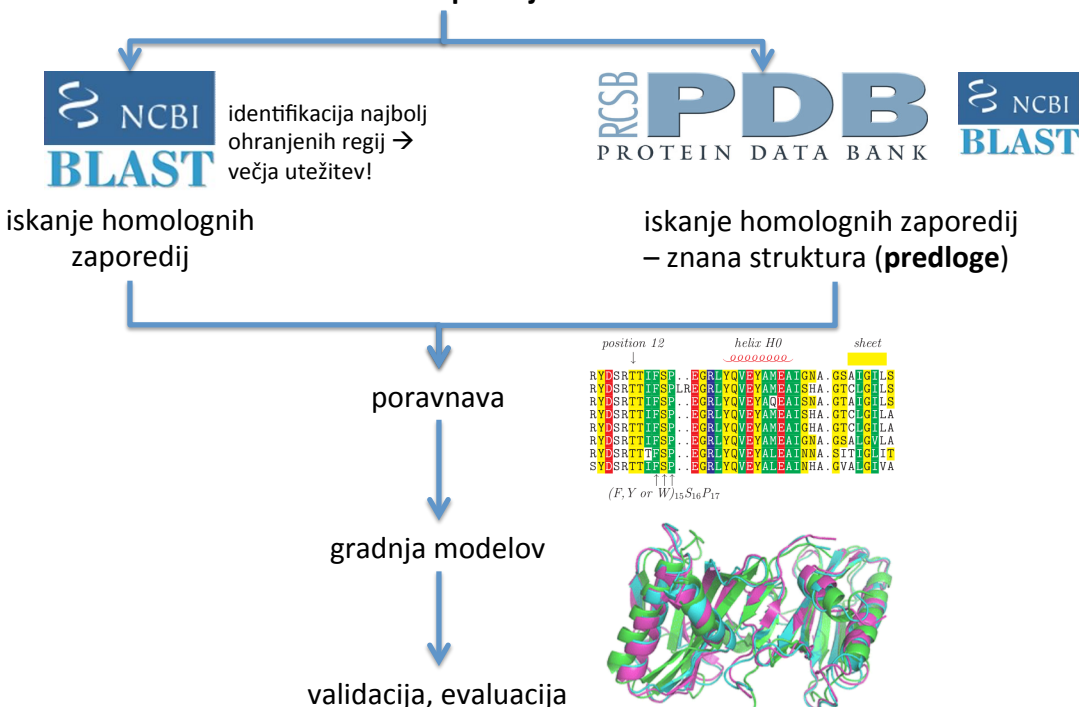
- Identifikacija predloge in priprava poravnave (prileganja).
- Gradnja ogrodja (backbone – atomi C_{α}).
- Modeliranje zank.
- Modeliranj stranskih skupin ak-ostankov.
- Optimizacija modela.
- Validacija modela.



8.a. Modeliranje makromolekul: MODELIRANJE POSAMEZNIH MAKROMOLEKUL

Modeliranje na osnovi homologije (*homology modeling, comparative modeling*)

tarčno zaporedje



8.a. Modeliranje makromolekul: MODELIRANJE POSAMEZNIH MAKROMOLEKUL

Modeliranje na osnovi homologije (*homology modeling, comparative modeling*)

- implementirano je v mnogih programih za modeliranje
- programi za modeliranje (tudi taki, ki so v osnovi *ab initio*) pogosto sami poiščejo, ali obstajajo primerne predloge za vsaj kakšno regijo/domeno
- primeri programov: **MODELLER**, SWISS-MODEL, PHYRE2, ...

Modeller <https://salilab.org/modeller/>

lokalna instalacija ali spletni vmesnik

Modeller

Program for Comparative Protein Structure Modelling by Satisfaction of Spatial Restraints

```

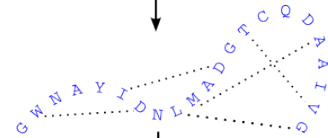
A I L V G S M P R R D G M E R K D L L K A N V K I F K C Q G A
V E V C P V D C P Y E G P N F L V I H P D E C I D C A L C E P
S A C K P E C P N I L Q G S - - L Y A I D A D S S I G C S S
S - - L A C G A C K P E C P V N I L Q G S - - L Y A I D A D S
    
```



1. Align sequence with structures

Template structure(s) SWQTYVDTNLVGTGAVTQA - - AI
Target sequence - GWNAYIDNLMADGTCQDAIVG

2. Extract spatial restraints



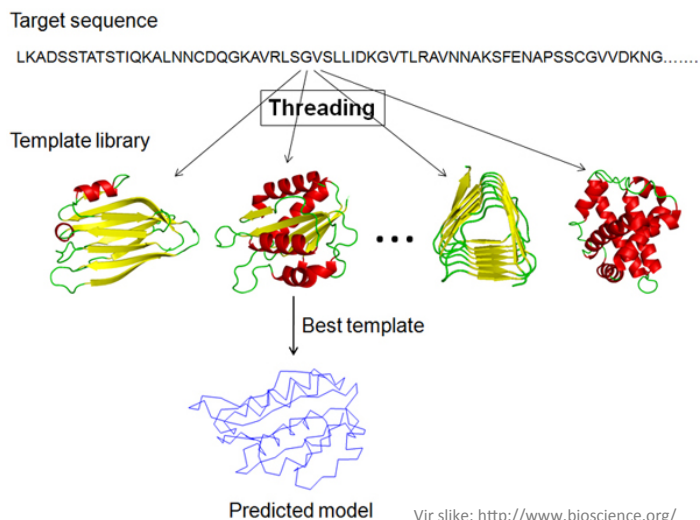
3. Satisfy spatial restraints



8.a. Modeliranje makromolekul: MODELIRANJE POSAMEZNIH MAKROMOLEKUL

Identifikacija zvitja (*protein threading*)

- Temelji na statistični analizi struktur v PDB in tarčnim zaporedjem.
- Na splošno: uporabimo ga lahko takrat, ko ne najdemo dobre predloge za modeliranje na osnovi homologije.
- Osnovna ideja:
 - v naravi je omejeno število zvitij
 - izbor izmed možnih zvitij temelji na poznavanju preferenc posameznih ak-ostankov, da se nahajajo v določenih strukturnih okoljih

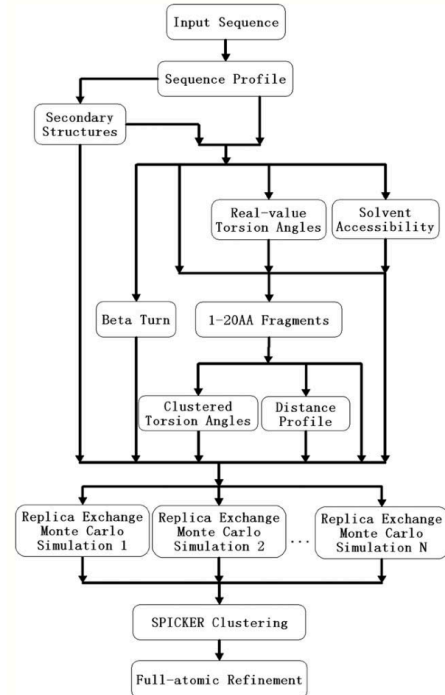


- Ni enako modeliranju na osnovi homologije – metodo lahko uporabimo na zaporedjih, za katere ne obstaja homologno zaporedje z znano 3D-strukturo.
- Orodja: HHpred, RAPTOR, PHYRE2, MUSTER, ...

8.a. Modeliranje makromolekul: MODELIRANJE POSAMEZNIH MAKROMOLEKUL

Modeliranje *ab initio* / *de novo* (*ab initio* / *de novo* modeling)

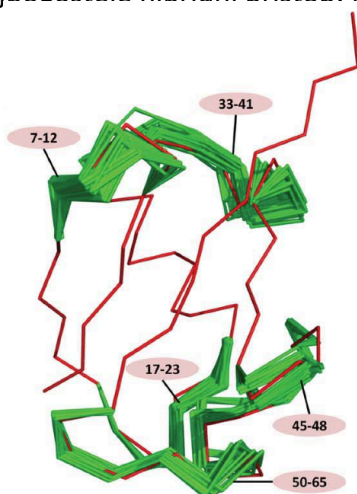
- fizikalno-kemijska osnova (+ evolucijska kovarianca intramolekulskih stikov)
- več principov, nor. globalna optimizacija primerne energijske funkcije (Ie-ta opisuje intramolekulske interakcije)
- metoda je računsko zelo zahtevna
- pogosto hibridni pristop (kombinacija metod, npr. modeliranje na osnovi homologije + *ab initio*)
- uporabnost:
 - manjši proteini
 - zanke in krajši fragmenti homolognih modelov, za katere nimamo ustrezne predloge
- orodja: ROSETTA, I-TASSER (QUARK), ...



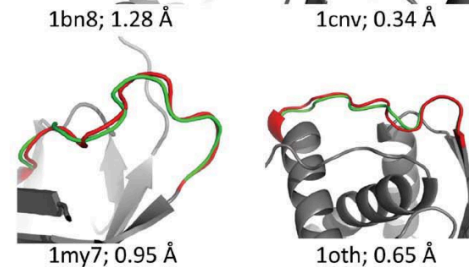
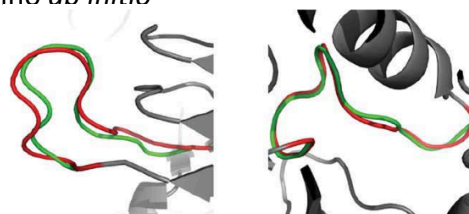
8.b. Modeliranje makromolekul: MODELIRANJE ZANK

Modeliranje zank (*loop modeling*)

- gre predvsem za zanke na površini, za katere pogosto nimamo homologne predloge
- najpogosteje se uporablja pristop na osnovi predlog:
 - v bazi fragmentov (na osnovi znanih struktur) poiščemo podobne fragmente
 - optimizacija (dihedralni koti, rotameri) → optimalno pakiranje in energija
- lahko tudi *ab initio*
- najpogosteje hibridni pristop: fragmenti + delno *ab initio*



izračun celega ensembela modelov (različne konformacije)



primeri dobrih modelov + X-ray struktur

8.c. Modeliranje makromolekul: FoldIt

FoldIt <http://fold.it>

- računalniška "igra", pri kateri igralci prispevajo k razvoju znanosti
- ročno in s pomočjo orodij zvijemo polipeptidno verigo z danim zaporedjem
- poseben algoritem za točkovanje (*clashing + packing + hiding + bonding + backbone + sidechain + disulfides ...*)



8.c. Modeliranje makromolekul: FoldIt

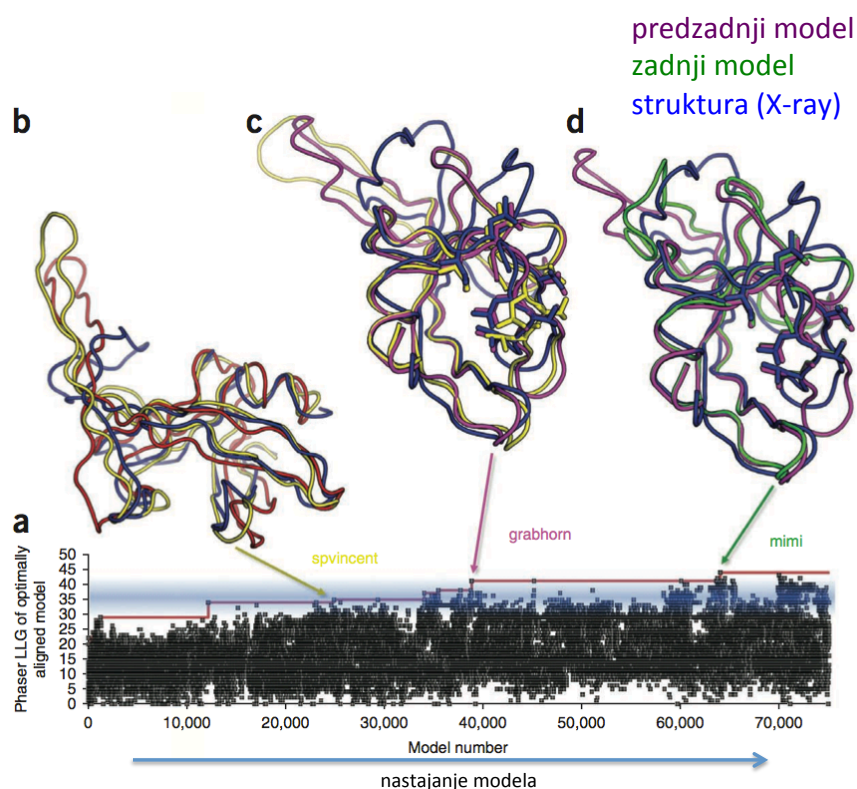
FoldIt

Crystal structure of a monomeric retroviral protease solved by protein folding game players

Firas Khatib¹, Frank DiMaio¹, Foldit Contenders Group, Foldit Void Crushers Group, Seth Cooper², Maciej Kazmierczyk³, Mirosław Gilski^{3,4}, Szymon Krzywda⁵, Helena Zabranska⁵, Iva Pichova⁵, James Thompson¹, Zoran Popović², Mariusz Jaskolski^{3,4} & David Baker^{1,6}

Following the failure of a wide range of attempts to solve the crystal structure of M-PMV retroviral protease by molecular replacement, we challenged players of the protein folding game FoldIt to produce accurate models of the protein. Remarkably, Foldit players were able to generate models of sufficient quality for successful molecular replacement and subsequent structure determination. The refined structure provides new insights for the design of antiretroviral drugs.

- Strukture monomerne retrovirusne proteaze ni bilo moč rešiti z molekularno zamenjavo (MR; ena izmed tehnik faziranja) na osnovi znanih homolognih struktur.
- Igralci FoldIt so pripravili dober model → uspešna molekularna zamenjava in rešitev kristalne strukture.



8.d. Modeliranje makromolekul: BAZE MODELOV

Baze modelov

- obstajajo že izdelane podatkovne baze modelov struktur za zaporedja, za katera ni znane eksperimentalno določene strukture
- previdnost pri uporabi – modeli so zgenerirani popolnoma avtomatsko!!!
- največji portal – Protein Model Portal (<http://www.proteinmodelportal.org>)

PSI | The Protein Model Portal

Home Interactive Modeling Quality Estimation Protein Modeling 101 More ▾

Welcome to the Protein Model Portal (PMP)

PMP gives access to various models computed by comparative modeling methods provided by different partner sites, and provides access to various interactive services for model building, and quality assessment.

Please enter your query.

Search Examples: (UniProt AC) (UniProt ID) (RefSeq) (PDBID) (Sequence) (Free Text)

Modeling Highlights (Show all)

FRAGFOLD with contacts

0.58 0.68 0.66

PDB

De novo structure prediction of globular proteins aided by sequence variation-derived contacts.

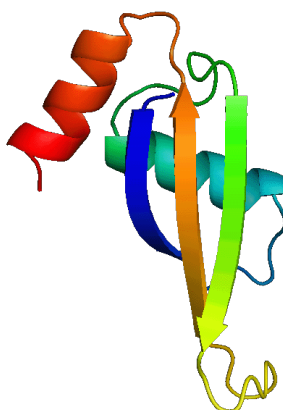
Koscioliek T., Jones DT. PLoS One. (2014) 9(3):e92197

8.e. Modeliranje makromolekul: CASP

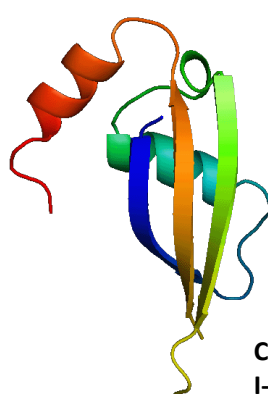
Kritična ocena metod za napoved strukture proteinov – CASP (Critical Assessment of protein Structure Prediction)

- “tekmovanje” algoritmov/programov za napoved strukture
- vsaki dve leti od leta 1994 naprej
- objektivni test; niti tekmovalci niti organizatorji ne poznajo eksperimentalno določenih struktur tarčnih proteinov!
- tarčne strukture: strukture *on hold* v PDB ter strukture, ki bodo kmalu rešene
- zmagovalec CASP10 (2012): I-TASSER (Zhang lab)

<http://predictioncenter.org>



napovedan model *ab initio*



CASP9
I-TASSER

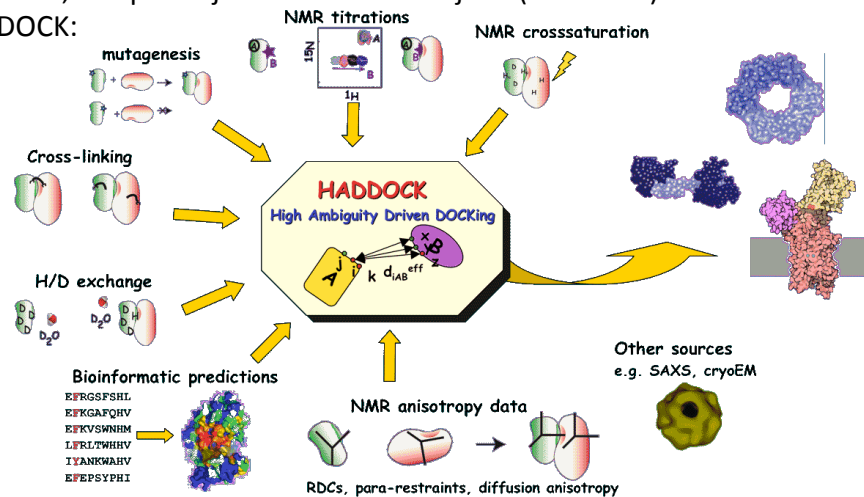
X-ray struktura

8.f. Modeliranje makromolekul: UMESTITEV in CAPRI

Umestitev (docking)

- napoved kompleksov:
 - vezava majhnih molekul na protein (farmacija!)
 - težje: medproteinski kompleksi (ponavadi večje konformacijske spremembe)
- identifikacija potencialnih interakcijskih regij → različne relativne orientacije komponent kompleksa → ovrednotenje različnih orientacij
- fleksibilnost stranskih skupin ak-ostankov; fleksibilnost peptidnega ogrodja navadno ne upoštevamo
- najboljše rezultate dobimo, če uporabljamo dodatne omejitve (restraints)
- primer: program HADDOCK:

- podobno kot pri modeliranju posameznih proteinov – tekmovanje CAPRI (*Critical Assessment of Prediction of Interactions*)
- CLUSPRO > HADDOCK > GRAMM-X > ...



8.g. Modeliranje makromolekul: MOLEKULSKA DINAMIKA

Simulacija molekulske dinamike (molecular dynamics simulations)

- računalniška simulacija časovnega razvoja sistema – študij dinamičnih procesov: difuzije, kemijskih reakcij, konformacijskih sprememb, minimizacija struktur, piljenje interakcijskih površin, ...
- upoštevanje fizikalno-kemijskih principov + energijska funkcija (*force field*)
- pri velikih sistemih zelo veliko št. atomov → grobi modeli (coarse grained models); npr. ena kroglica predstavlja več atomov
- primer: proučevanje homo-dimerizacije transmembranskih heliksov

