

Slide 1

STRUKTURNE PODATKOVNE ZBIRKE

STRUKTURA PROTEINOV

NAČINI DOLOČEVANJA STRUKTURE PROTEINOV

STRUKTURNE PODATKOVNE ZBIRKE

PRIMERJAVE PROTEINSKIH STRUKTUR

Slide 2

ZAKAJ?

1. Mehanizem delovanja (katalitski aminokislinski ostanki, vezavne površine....).
2. Inženiring za izboljšanje lastnosti (bolj stabilni encimi, bolj stabilni proteini, neaktivni proteini....)
3. Farmacevtska industrija (heterologno pridobivanje pomembnih proteinov, načrtovanje učinkovin- "*drug design*")
4. Vzorci molekularne evolucije
5. Raziskave procesa zvitja proteinov

```

    graph TD
      A[napoved, eksperiment] --> B[struktura]
      B --> C[eksperimentiranje]
      B --> D[modeliranje]
      C --> E[razlaga procesov]
      D --> E
  
```

Slide 3

Določevanje strukture

- EKSPERIMENTALNO

Slikanje proteinskih kristalov z X-žarki

NMR

- In silico*

Homologno modeliranje

Ab initio predikcija

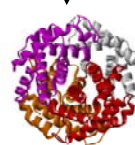
npr. Rosetta

MVHLTPEEKSAVTALWGKLVNDEVGGEALGR
 LLVVYPWTQRFFESFGDLSTPDVAMGNPVKA
 HGKVLGAFSDGLAHLNLTGTFATLSELHC
 DKLHVDPENFRLLGNVLCVLAHFGKEFTF
 PVQAAYQKVVAGVANALAHKYH

↓

↓

STRUKTURNA GENOMIKA



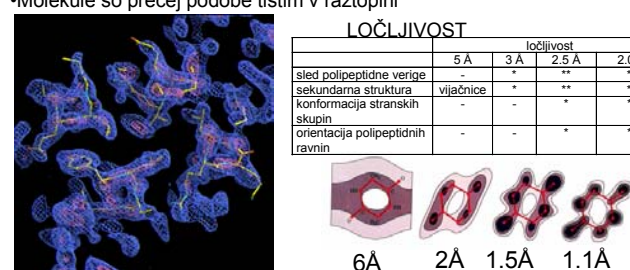
Slide 4

Slikanje proteinskih kristalov z X-žarki

- Največ 3D struktur z dobro ločljivostjo
- Potrebuješ proteine v kristalni obliki (čistost, koncentracije- kristalizacija ozko grlo!)
- Difrakcija rentgenske svetlobe na elektronih, iz intenzit uklonov lahko izračunamo elektronsko gostoto v katero prilegamo model
- Možna so popačenja zaradi pakiranja v kristale.
- Molekule so precej podobne tistim v raztopini

LOČLJIVOST

	ločljivost			
	5 Å	3 Å	2.5 Å	2.0 Å
sled polipeptidne verige	-	*	**	**
sekundarna struktura	-	vijačnice	*	**
konformacija stranskih skupin	-	-	*	**
orientacija polipeptidnih ravnin	-	-	*	**



6Å 2Å 1.5Å 1.1Å

Slide 5

Jedrska magnetna resonanca, NMR

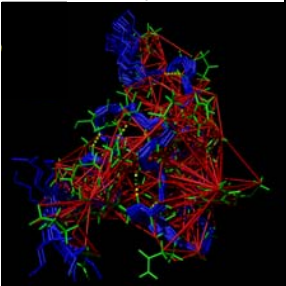
Spremljamo jedra ^1H , ^{15}N , ^{13}C , ^{31}P ...

Določanje struktur v raztopini (visoka zopnost 0.2-1mM, čistost)

Prednost: metabolno markiranje (^{15}N , ^{13}C)

Tudi informacija o dinamiki molekule v raztopini ali njenih delov

1. določitev jeder v NMR spektrih
2. identifikacija razdalj med protoni, krajših od $\sim 5 \text{ \AA}$ (NOE)
3. Izračun struktur, ki ustrezajo eksperimentalne podatkom



Slide 6

RAZKORAK MED ŠTEVILOM ZAPOREDIJ IN ŠTEVILOM 3D STRUKTUR

>30 000 000	GeneBank
>140 000	SwissProt
24 444	Protein Data Bank

Problemi pri pripravi vzorca (količina, izražanje), lastnosti (netopni, membranski), kristalizaciji

Npr. MsbA iz *E. coli* ABC transporter

- >20 homologov iz 12 bakterijskih vrst izražajo
- >20 detergentov z vsemi, skupaj 96 000 testov
- 35 kristalnih oblik
- 1 kristal, ki dobro sipa
-
- 1 SCIENCE članek (Chang G and Roth CB (2001) Science 293, 1793-1800)

Slide 7


STRUKTURNE PODATKOVNE ZBIRKE

- PDB (Protein Data Bank)
Podatkovna zbirka 3D zgradb proteinov in nukleinskih kislin. Eksperimentalno določene strukture. Primarni vir strukturnih podatkov. Kazalci iz drugih podatkovnih zbirk ali strežnikov (SwissProt, SRS). 29326 struktur (25. 1. 2005)
<http://www.rcsb.org/pdb/>
- PDBsum
Vsa potrebna dodatna informacija na voljo: zaporedje aminokislin, podatki o sekundarni strukturi, opis in prikaz 3D zgradbe. Povezave na ostale podatkovne zbirke. Iskanje s ključnimi besedami in PDB kodami
<http://www.biochem.ucl.ac.uk/bsm/pdbsum/>
- FSSP (Fold Classification based on Structure-structure Alignment of Proteins)
Podatkovna zbirka zvijit proteinov, ki temelji na primerjavi 3D struktur proteinov (s programom DALI).
<http://www.ebi.ac.uk/dali/fssp/fssp.html>

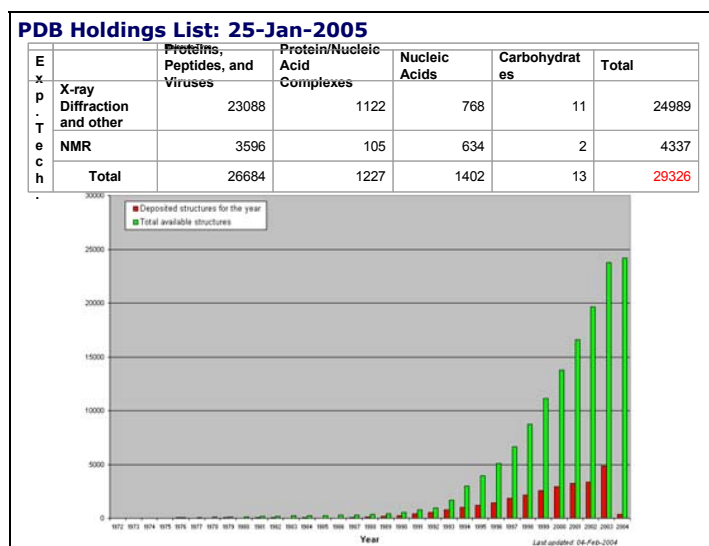
Slide 8

Brookhaven Protein Data Bank proteini 1973

Research Collaboratory for Structural Bioinformatics
Rutgers, the State University of New Jersey, National Institute of Standards and Technology (NIST), San Diego Supercomputer Center (SDSC)



Slide 9



Slide 10

FORMAT ZAPISA

PDB

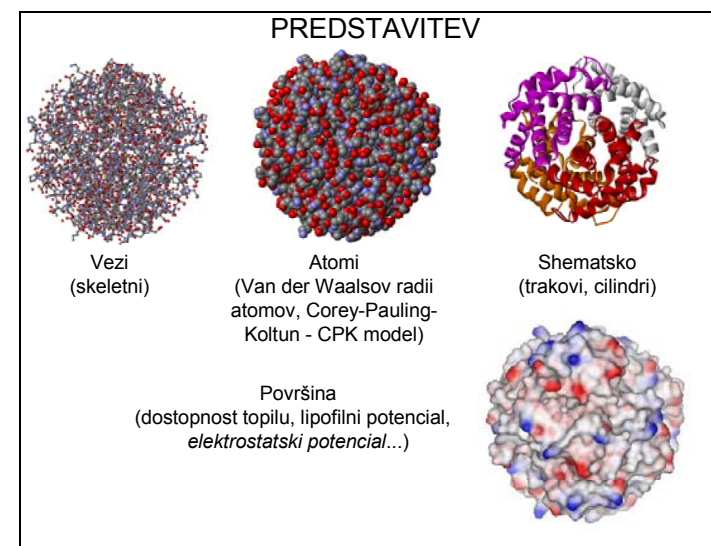
Koordinate atomov brez definiranih kemijskih vezi, npr.

```

...
ATOM 17 N HIS 4 10.809 -0.520 7.814 1.00 23.20 2 1CA2 149
ATOM 18 CA HIS 4 9.940 0.115 8.847 1.00 22.64 2 1CA2 150
ATOM 19 C HIS 4 9.218 -0.939 9.627 1.00 20.94 2 1CA2 151
ATOM 20 O HIS 4 9.362 -2.180 9.350 1.00 21.32 2 1CA2 152
ATOM 21 CB HIS 4 9.034 1.208 8.184 1.00 24.47 2 1CA2 153
ATOM 22 CG HIS 4 9.777 2.520 8.351 1.00 26.46 2 1CA2 154
ATOM 23 ND1 HIS 4 9.878 3.168 9.582 1.00 27.37 2 1CA2 155
ATOM 24 CD2 HIS 4 10.440 3.297 7.469 1.00 27.10 2 1CA2 156
ATOM 25 CE1 HIS 4 10.565 4.296 9.420 1.00 27.45 2 1CA2 157
ATOM 26 NE2 HIS 4 10.944 4.385 8.141 1.00 27.36 2 1CA2 158
ATOM 27 N TRP 5 8.439 -0.552 10.640 1.00 18.42 1CA2 159
ATOM 28 CA TRP 5 7.704 -1.477 11.481 1.00 15.91 1CA2 160
ATOM 29 C TRP 5 6.729 -2.255 10.632 1.00 15.53 1CA2 161
ATOM 30 O TRP 5 6.252 -1.753 9.576 1.00 15.52 1CA2 162
ATOM 31 CB TRP 5 7.001 -0.749 12.644 1.00 14.85 1CA2 163
ATOM 32 CG TRP 5 5.817 0.042 12.226 1.00 14.11 1CA2 164
ATOM 33 CD1 TRP 5 5.782 1.347 11.843 1.00 13.95 1CA2 165
ATOM 34 CD2 TRP 5 4.443 -0.411 12.177 1.00 13.56 1CA2 166
ATOM 35 NE1 TRP 5 4.527 1.717 11.523 1.00 13.30 1CA2 167
ATOM 36 CE2 TRP 5 3.672 0.676 11.750 1.00 13.50 1CA2 168
...

```

Slide 11



Slide 12

PROGRAMI ZA PREDSTAVITEV STRUKTUR

Rasmol, Chime
 Verzija prirejena za splet
<http://www.umass.edu/microbio/rasmol/>

Swiss-PDBViewer- Deep view, homologno modeliranje, rezultati SwissModel, merjenje razdalj, kotov, mutacije, izračun površin, potencialov...
<http://www.expasy.org/spdbv/>

VMD
<http://www.ks.uiuc.edu/Research/vmd/>

Cn3D; Komercialni

PROGRAMI KOT PRIPOMOČKI PRI MODELIRANJU

Izbiranje posameznih atomov oz. skupin
 Merjenje razdalj, torzijskih in dihedralnih kotov, površin
 Sidranje dveh molekul (docking)
 Animacije

Slide 13

HOMOLOGNO MODELIRANJE

Ker je struktura ponavadi bolj ohranjena od aminokislinskega zaporedja, lahko napovemo strukturo neznanemu homologu, če obstaja zanj struktura homologa v podatkovnih zbirkah

1. Iskanje homologov v podatkovnih zbirkah
2. Iskanje struktur homologov v podatkovnih zbirkah
3. Napoved strukture (katalitskih aminokislin, a.k. na površini...) in napoved funkcije
4. Usmerjeno eksperimentiranje

Modeliranje

Slide 14

AVTOMATSKO MODELIRANJE

SWISSMODEL <http://www.expasy.ch/swissmod/SWISS-MODEL.html>
Modeller <http://sallab.org/modeller/modeller.html>
WHATIF <http://www.cmbi.kun.nl/gv/servers/WIWWWI/>

Modeliranje na osnovi podobnih proteinov (homology modeling)

1. Poiščemo podobna zaporedja z določeno 3D strukturo (definiramo maksimalno E vrednost v BLASTu)
2. Identifikacija predelov z >25% identičnosti na dolžini >20 AK in proteinov na osnovi katerih modeliramo variabilne predele
3. Priprava modela na osnovi različnih fragmentov proteinov, ki imajo najbolj podobno zaporedje.

ZBIRKA MODELOV- SWISS MODEL REPOSITORY
<http://swissmodel.expasy.org/repository/>

Models:534422 Modeli zaporedij pripravljani s programom UniProt:3.4 SWISS-MODEL. Za zaporedja, ki imajo PDB:20.11.2004 podobne proteine (>25% identičnosti) z Last update:7.01.2005 znano 3D strukturo.

Slide 15

PRIMERJAVE 3D STRUKTUR

Iskanje oddaljenih homologij, sklepanje na funkcijo....

Strežnik DALI
<http://www2.ebi.ac.uk/dali/>
 Internetni strežnik za primerjave 3D struktur proteinov.

Slide 16

```

COMPND COLICIN HA (C-TERMINAL DOMAIN) (PORE-FORMING DOMAIN)
AUTHOR H.M.PARKER,J.P.M.POSTMA,F.PATTUS,A.D.TUCKER,D.TSERNOGLOU
NR, STRID1 STRID2 Z RMSD LALI LSEQ2 ZIDE REVERS PERMUT NFRAG TOPO PROTEIN
! 4: 1co1-a 1hb1 6.2 3.1 119 157 9 0 0 10 S HEMOGLOBIN (SEA CUCUMBER)
    
```

138	H E == A H	101
139	H S == R H	102
140	H W == T H	103
141	H Y == H H	104
142	H L == .	105
143	T S == D H	106
144	T G == L H	106
145	I == N T	107
146	A == K T	108
147	H S == .	109
148	H S == .	110
149	H V == V	110
150	H A == G	111
151	H L == A H	111
152	H G == D H	112
153	H I == H H	113
154	H F == Y H	114
155	H S == N H	115
156	H A == L H	116
157	H T == F H	117
158	H L == A H	118
159	H G == K H	119

Holm, L., & Sander, C. (1993). Structural alignment of globins, phycocyanins and colicin A. FEBS Lett., 315, 301-306.