

Struktūrinė bioinformatika vaistų paieškoje



Struktūrinė bioinformatika

➤ Kas yra SBI?

“Struktūrinė bioinformatika yra bioinformatikos šaka, kurios objektai yra biologinės molekulės – DNR, RNR, baltymai, ligandai ir kitokie kompleksai.”

<http://biology.sdsc.edu/strucb.html>

SBI vaistų paieškoje

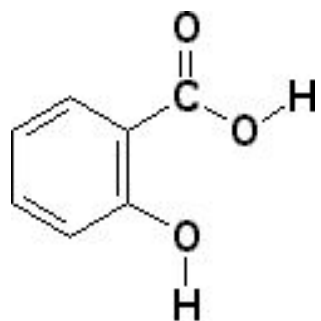
- SBI gali būti naudojama ištirti:
vaistų taikinius (drug targets), paprastai baltymus
Ligandų susirišimą



“racionalus” vaistų kūrimas

(piusai = sutaupyta laikas ir \$\$\$)

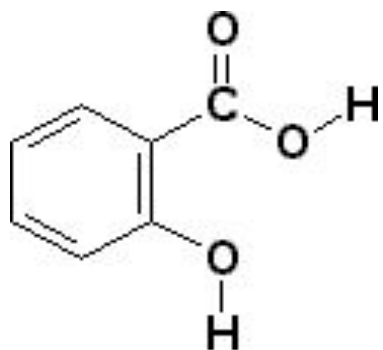
Tradicioniniai metodai



Natūralus
(naudojant augalinės
kilmės produktus)
ligos gydymas



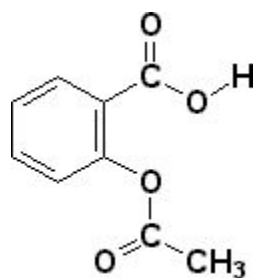
aktyvaus junginio
išskyrimas
(maža org. molekulė)



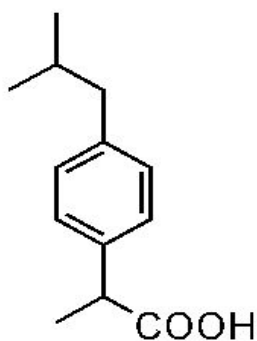
Junginio sintezė



Struktūros
manipuliacija norint
gauti geresnį vaistą
(didesnį efektyvumą,
mažiau šalutinių efektų)



Aspirin



Ibuprofen

Modernūs metodai

Kas daroma kitaip?

- Vaisto paieška pradedama nuo ligos, o ne nuo gydymo.
- Naudojant ligos modelį surandami realiausi genetiniai/biologiniai komponentai (pvz. Vaistų taikiniai)

Modernūs metodai

liga → genetinis biologinis taikiny



Vedančios (“lead”) molekulės paieška

- reikia sukurti platformą taikinio funkcijos patikrinimui
- Naudojant platformą surasti faktorius, įtakojačius taikinio funkciją



Didelio pralaidumo patikra, high throughput screen (HTS)

- skirtas surasti “hitus” (junginius su nM iki mažų μ M prijungimo giminingumu)

Modernūs metodai

Mažamolekuliniai geri junginiai



Manipuliacijos su struktūra norint pagerinti savybes
Pvz. padidinti K_i norint padidinti giminingumą



tolesnė “vedančios” (lead) molekulės optimizacija iki
tikro vaisto

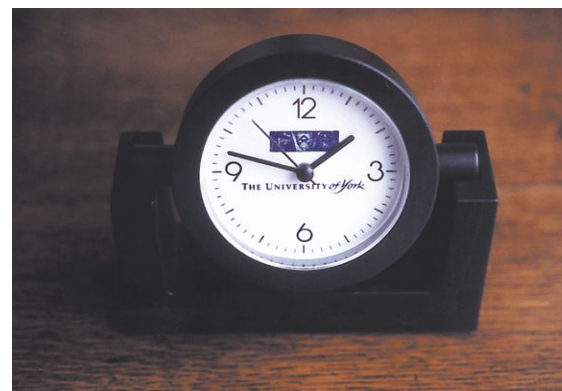
Reikalingų farmakologinių savybių įvedimas:
Absorbcija, metabolizmas, saugumas, vartojama forma



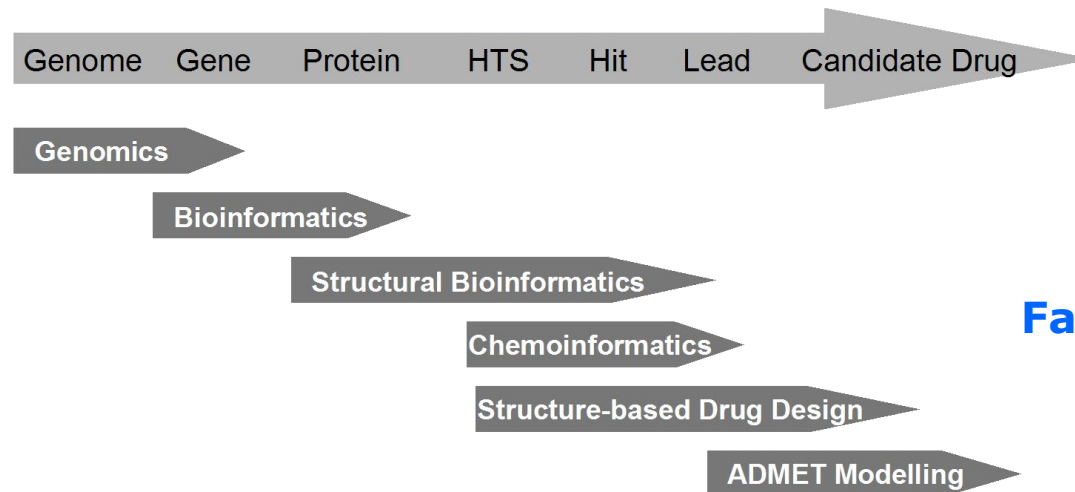
Klinikiniai tyrimai

Idomūs faktai...

- Virš 90% vaistų patenkančių iki klinikinių tyrimų stadijos būna atmetami
- Vidutinė naujo vaisto sukūrimo kaina yra apie **\$770 million**

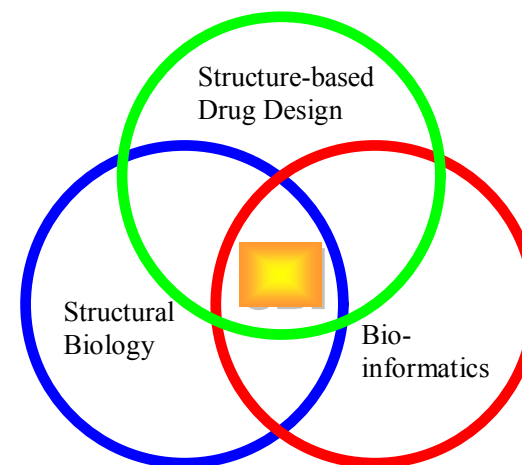


Struktūrinės biologijos įtaka vaistų kūrimei

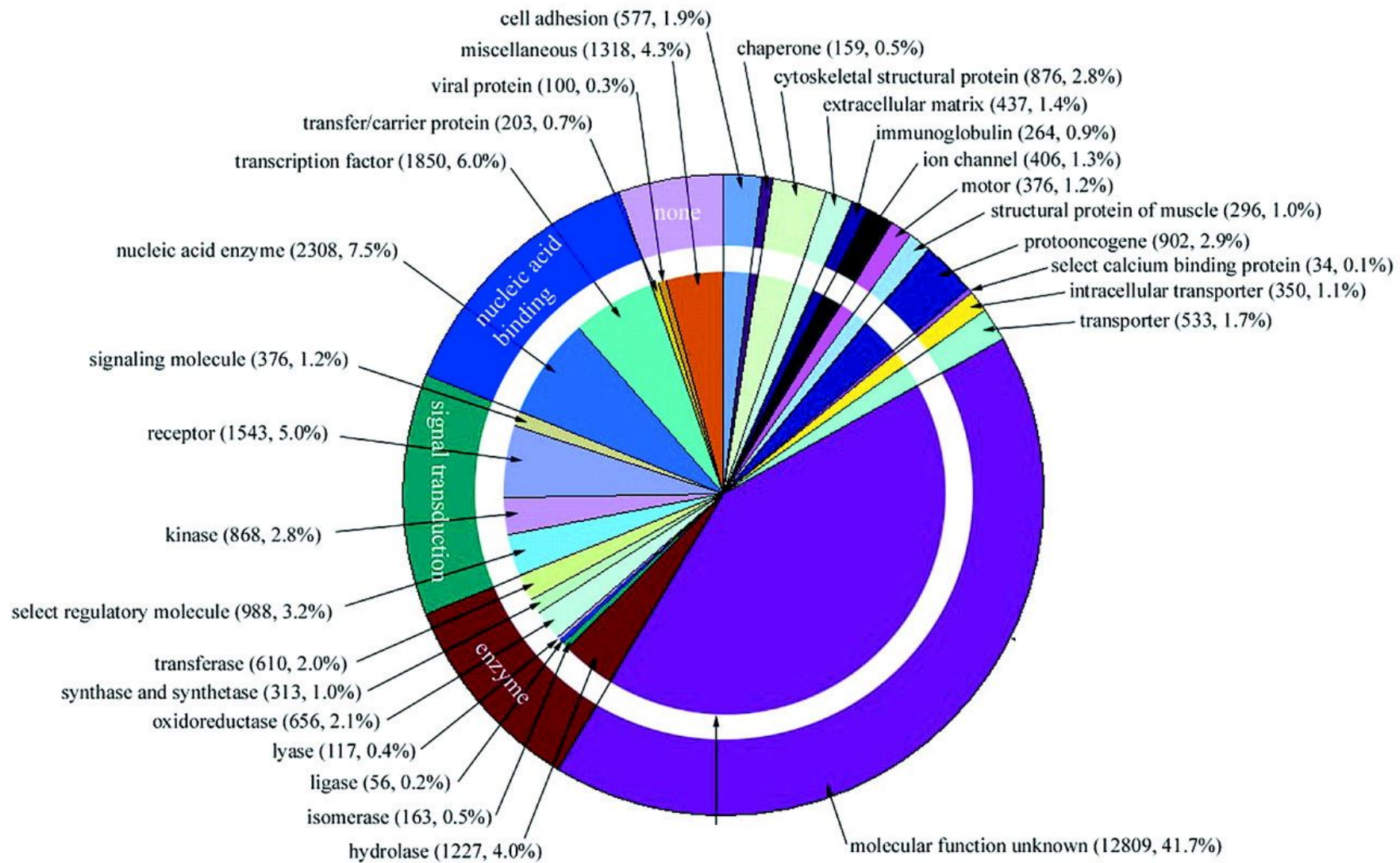


**Pagal
Fauman et al.**

- Pagreitina vaistų kūrimo procesą apjungiant šias sritis

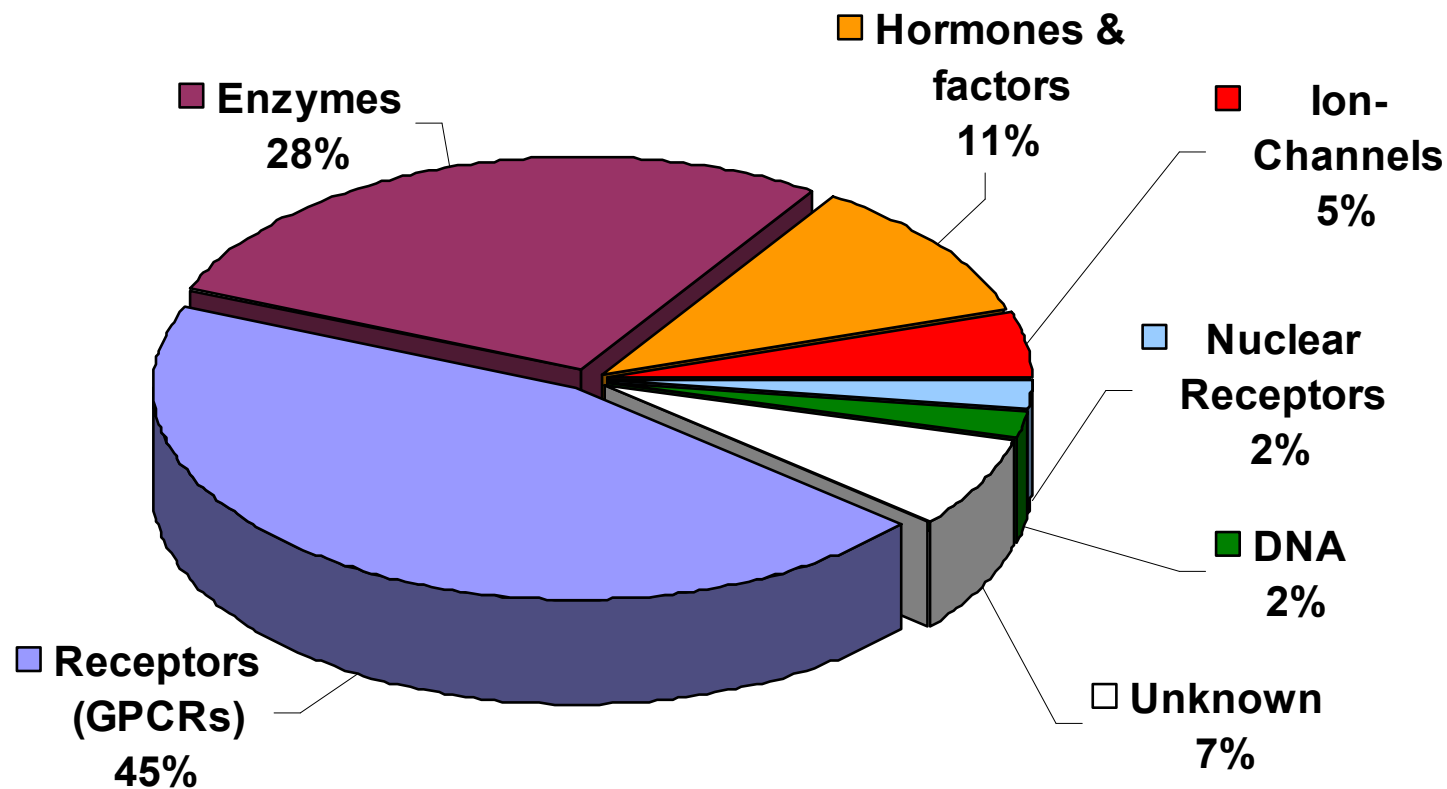


Žmogaus genome yra daugybė tinkamų taikinių



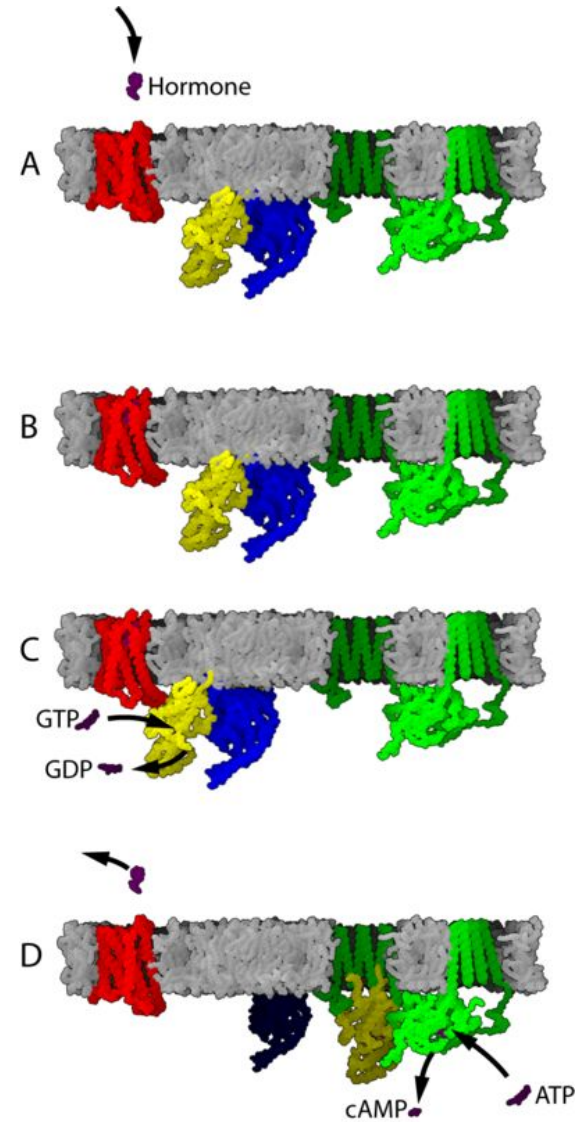
Dabar naudojamų vaistų taikiniai

Remiantis 483 vaistų iš Goodman and Gilman's "The Pharmacological basis of therapeutics"



GPRC

➤ G-protein



Žmogaus genomas

polisacharidai

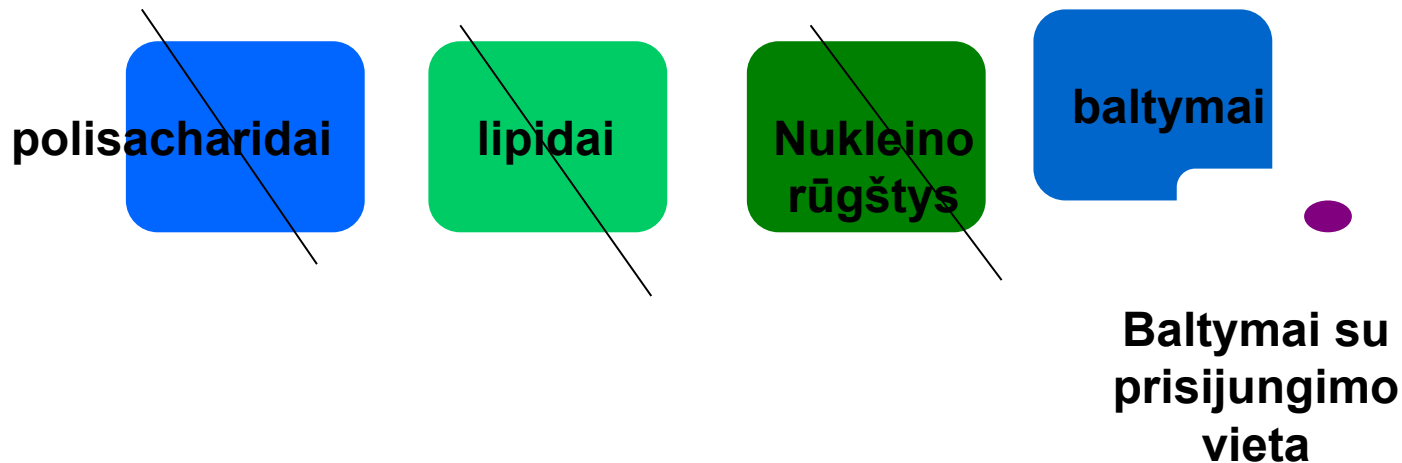
lipidai

Nukleino
rūgštys

baltymai

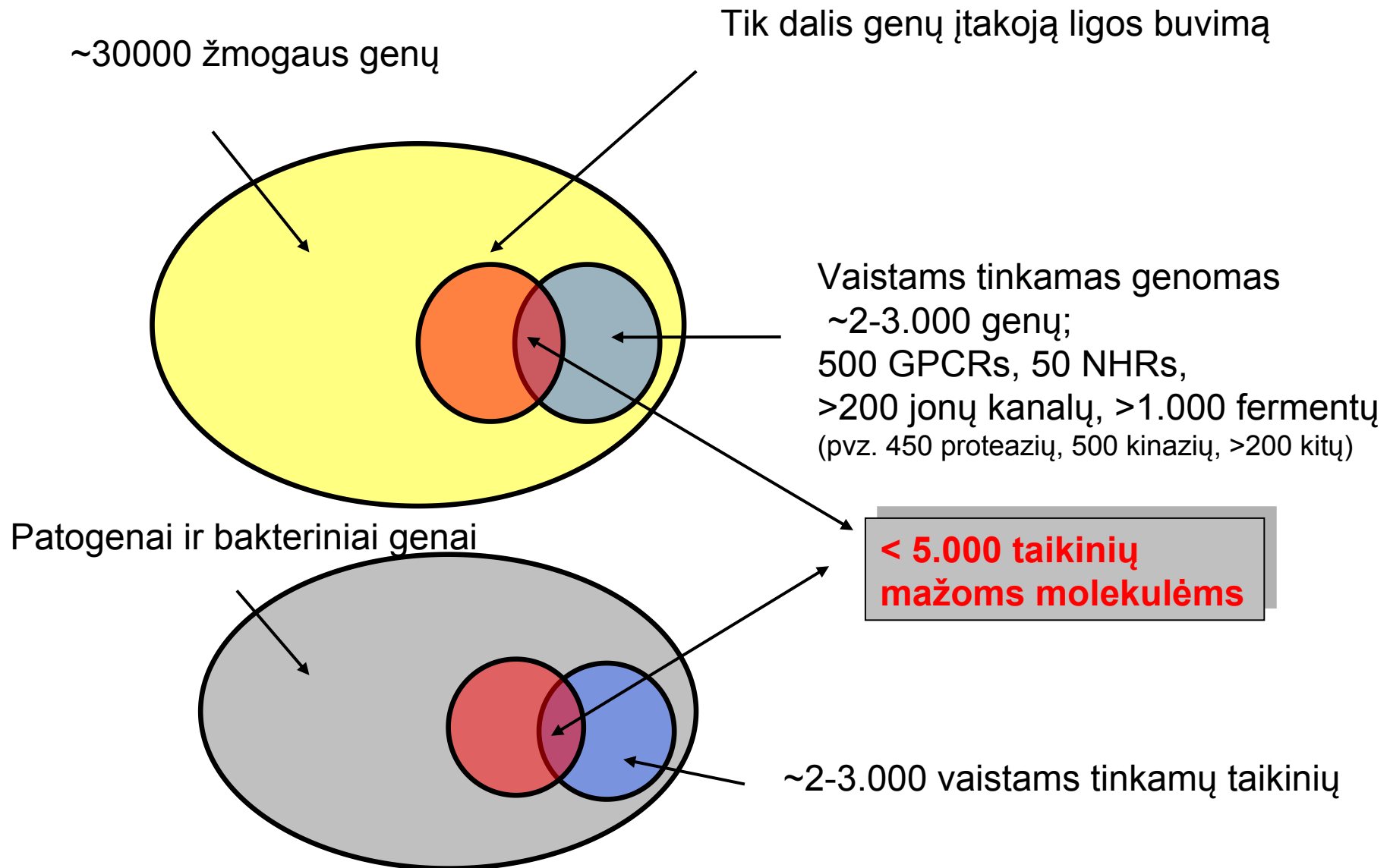
Problemos dėl toksiškumo,
specifiškumo pirmos trys
kategorijos atkrenta...

Žmogaus genomas



“vaistams tinkamas genomas” = aibė genų, kurie ekspresuoja baltymus, tinkamus vaistinių medžiagų susirišimui

Vaistams tinkamų taikinių mažoka?



Dabartinės kryptys

- Literatūra
- Nežinomų taikinių paieška – sąrašų prioretizavimas
- Vieno taikinio atsisakymas
 - Taikinių įvertinimas ir palyginimas
 - Susijusių, nesugretintų duomenų šaltinių apjungimas
 - Geresnis supratimas apie taikinio “kaiminystę”
 - Ligos mechanizmas
 - Biomarkeriai
 - Toksikologija

Kontekstinės informacijos šaltiniai

➤ Struktūruota

20%

- Internal Chemical Dbs
- Internal Biological Dbs
- External, Commercial Dbs
 - GVK Bio, Ingenuity IPA...
- External Public Dbs
 - EMBL, PDB, SNPdb, etc

Išvystyta technologija

➤ nestrukūruota

80%

- Internal Docs:
 - Tox Reports, Clinical Trial Reports.
- External Docs:
 - Patents; USPTO, WIPO, EP, etc
 - Literature; Medline, Embase
 - Press Releases:
 - competitor, supplier, collaborator, academic (etc)
 - Government Agencies
 - Conference Proceedings
 - News Feeds

Besivystanti technologija

Tekstinių duomenų analizė

➤ Šaltiniai

- Literatūra
- patentai
- Nuosavi duomenys

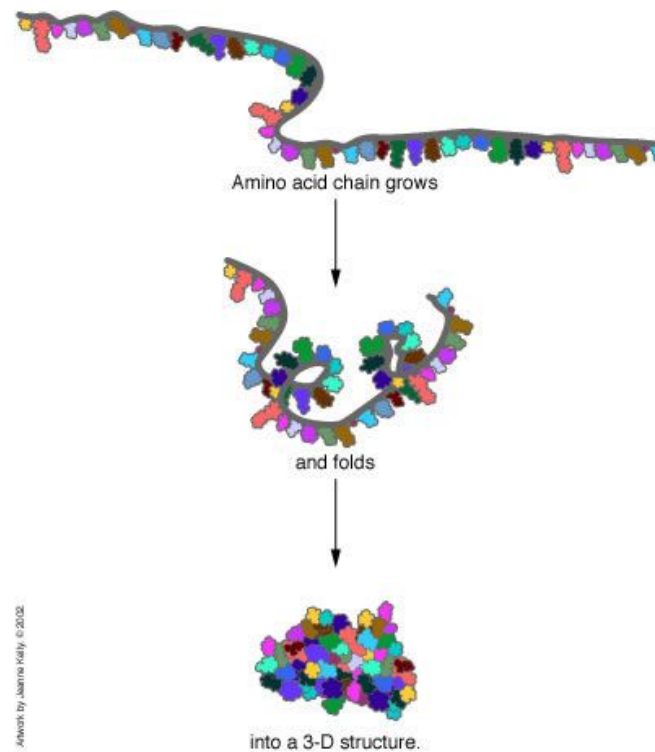
➤ Informacija

- Baltymų-baltymų sąveikos
- Ekspresija audiniuose
- Farmakologiniai skirtumai
 - Polimorfizmai, alternatyvus “splice’ingas”
 - Rūšys
- Toksikologija
- kita

Baltymų struktūrų modeliavimas

Baltymų struktūros

- Baltymų susipakavimas: susipakuoja į minimalios energijos būseną

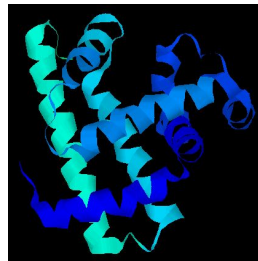


Baltymų struktūros

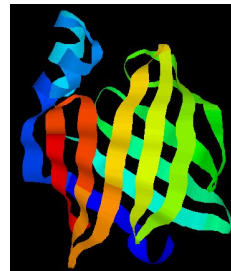
- Pirminė struktūra – amino r. seka

MTYKLILNGKTKGETTTEAVDAATAEKVFQYANDNGVDGEWTYTE

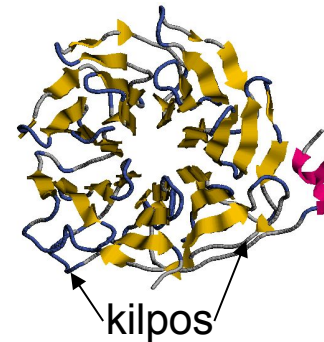
- Antrinės struktūros



spiralės
helices



lakštai
strands

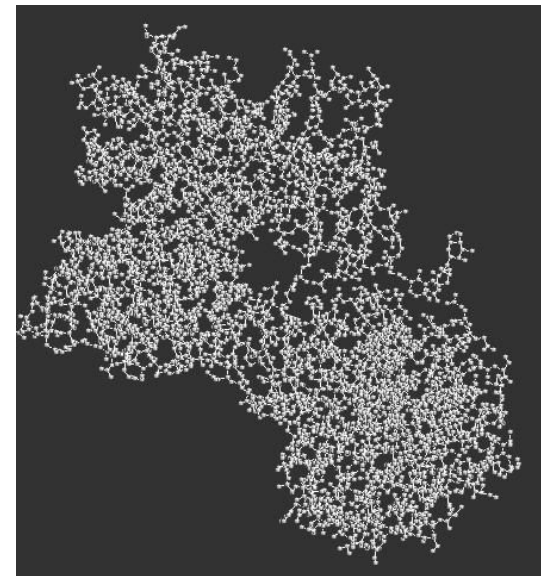
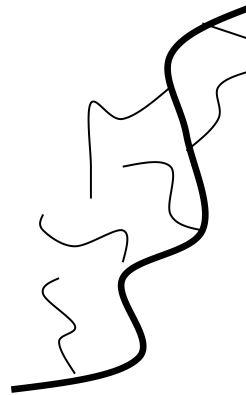


kilpos
loops

- Tretinės struktūros

Baltymų struktūros

- Pagrindinė grandinė (“Stuburas”) vs visų atomų modelis

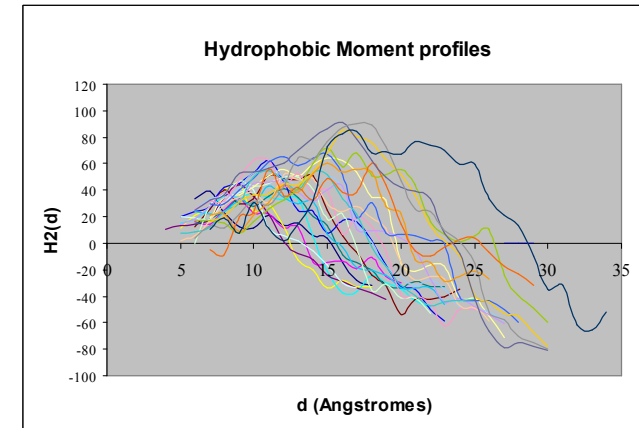


Pagrindinė grandinė + šoninė grandinė =
modelis iš visų atomų

Pagrindinė grandinė == struktūrinis elementas(fold)

Baltymų struktūros

- Baltymų struktūros
 - Paprastai kompaktiškos

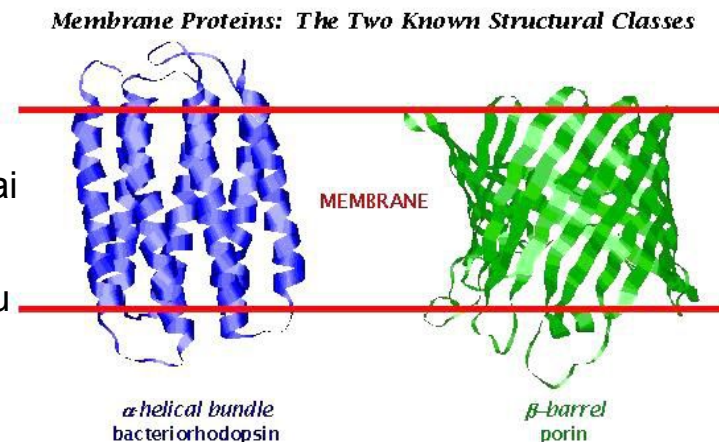


- **Tirpios struktūros**
 - Individualūs domenai paprastai globuliniai
 - pasižymi panašiomis charakteristikomis, pvz. hidrofobiškumo profiliu

- **Membraniniai baltymai**

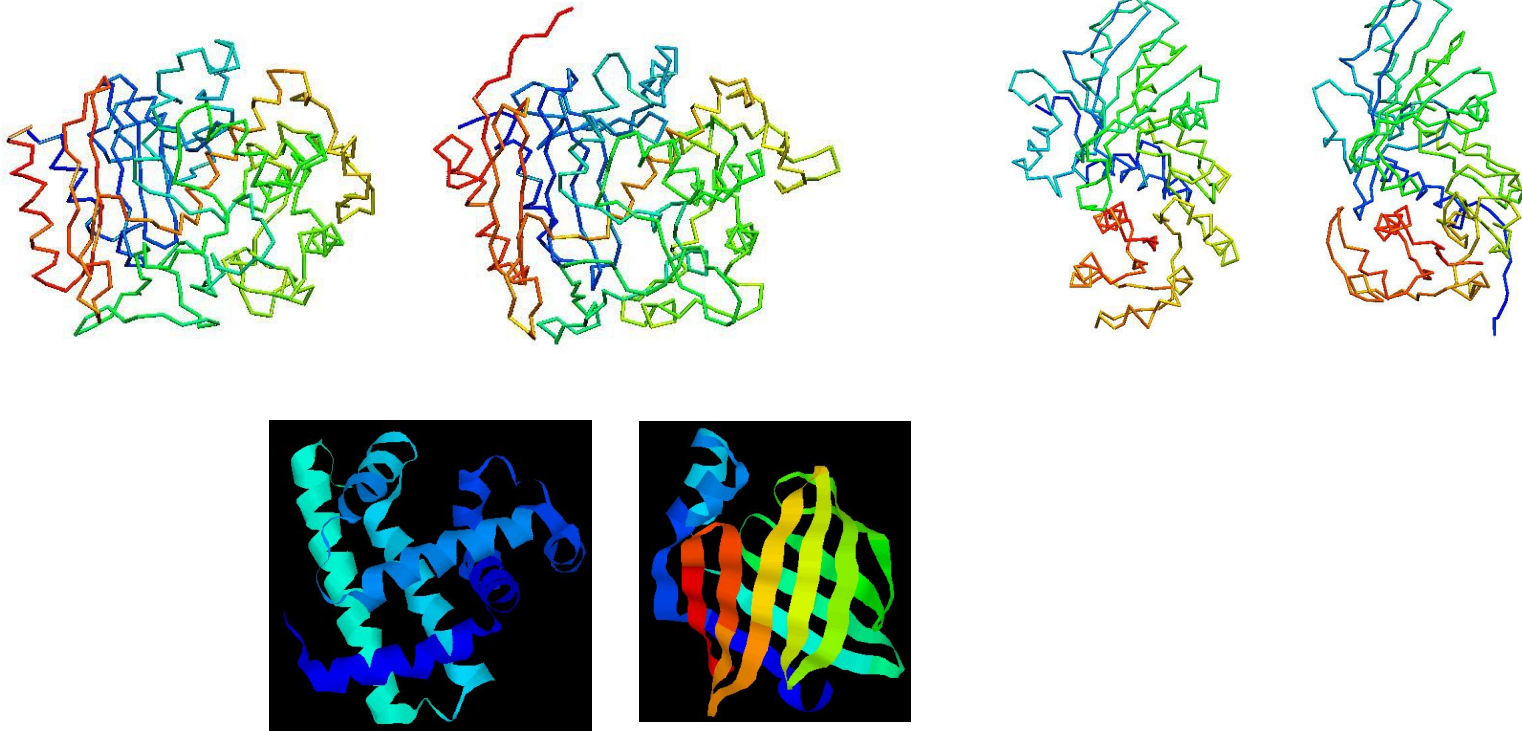
Transmembraninio segmento šoninės grandinės paprastai nepolinės

Polinės grupės dalyvauja sudarant vandenilinius ryšius su kitomis grandinėmis



Baltymų struktūros

- Duomenų bazėje Protein Data Bank (PDB) 2010 lapkričio 23d. buvo 69510 baltymų struktūros
- ~1200-1300 unikalūs foldai



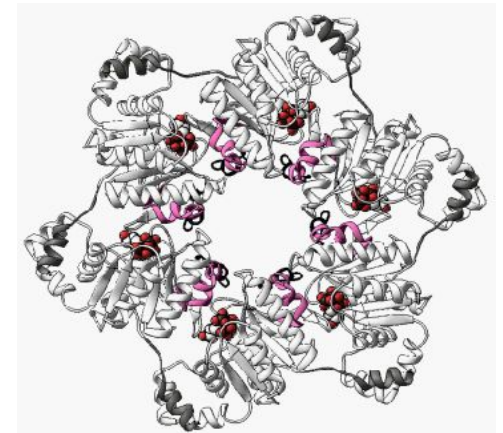
Baltymų struktūros

- Baltymo struktūroje yra užkoduota informacija apie jo funkciją
- Nors ir iš sekos galima bandyti nuspėti funkciją, tačiau tik iš tretinės struktūros galima numanyti veikimo mechanizmą

Baltymo struktūros nustatymas

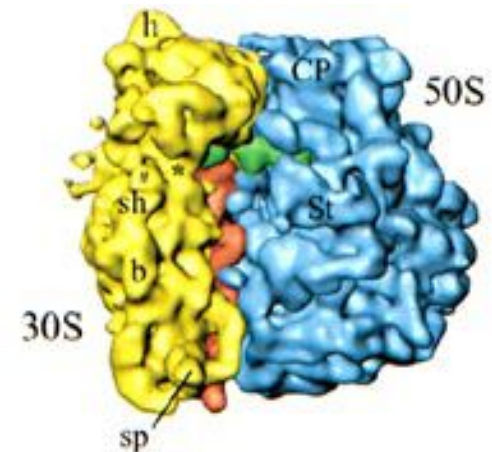
➤ Didelio tikslumo metodai

- Iš Rentgeno kristalografinių duomenų (~1Å)
- Branduolių magnetinio rezonanso (NMR) (~1-2.5Å)



➤ Nedidelio tikslumo metodai

- Cryo-EM (elektroninė mikroskopija) ~10-15Å



Baltymo struktūros nustatymas

- Rentgeno kristalografinis
 - tiksliausias
 - in vitro
 - Reikia baltymo kristalų
 - > ~100K per structure

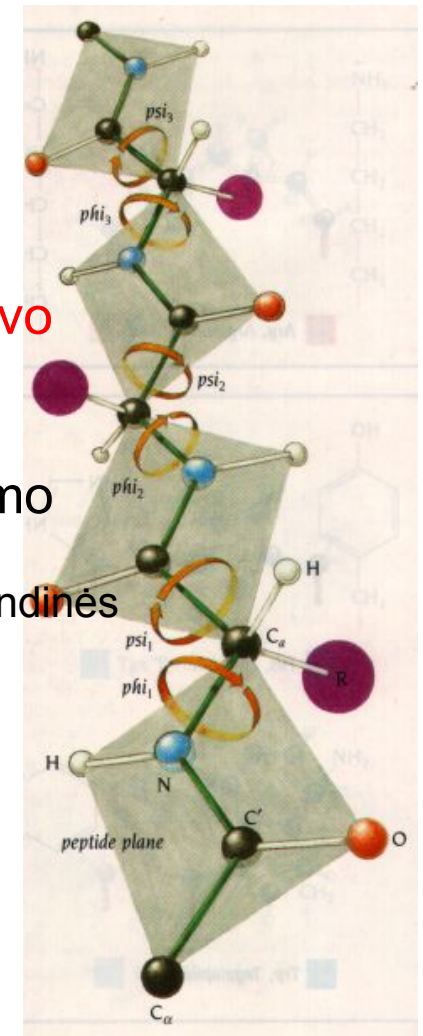
- - Gana tikslus
 - in vivo
 - Nereikia kristalų
 - Tinka tik nedideliems baltymams

- - Atvaizdavimo technologijos
 - Mažos raiškos

Baltymo struktūros nustatymas

- Teoriškai baltymo struktūra gali būti apskaičiuota
- Baltymas susisuka į trimatę struktūrą kad sumažintų savo potencinę energiją
- Problema gali būti suformuluota kaip energijos minimumo paieška
 - Paieškos erdvė gali būti apribota iki psi/phi kampų ir šoninės grandinės rotamerų
 - Paieškos erdvė yra milžiniška net mažiems baltymams
 - Lokalių minimumų skaičius auga eksponentiškai didėjant šon. grandinių skaičiui

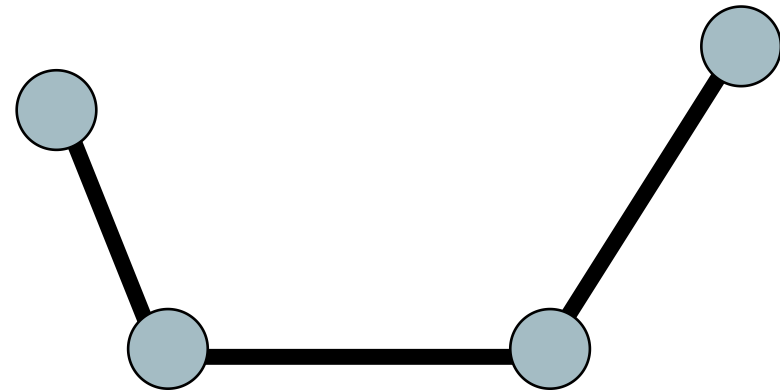
Todėl tai yra sudėtingas uždavinys



Skaičiavimo metodai baltymo struktūros numatyme

➤ Energijos funkcija baltymo aprašymui

- Jungties energija
- Jungčių kampo energija
- Dvisienių kampų energija
- van der Waals'o energija
- Elektrostatinė energija



➤ Struktūros apskaičiavimas naudojant energiją

➤ Nepraktiškas

- Reikalauja daug resursų
- Mažas tikslumas

Duoda informaciją apie susisukimo kelią ir tai kokia yra “galutinė” struktūra

Skaičiavimo metodai baltymo struktūros numatyme



- **Protein threading** – struktūros numatymas atliekamas “gerus” sekos-struktūros atitikmenis
- **Homologinis modeliavimas** – identifikuojami homologiniai baltymai panaudojant sekų palyginimus. Struktūra numatoma padedant atitinkamas a.r. į reikalingą vietą

Protein Threading

- Pagrindinis teiginys

Unikalių sąsukų (structural fold) skaičius gamtoje yra nedidelis, gali siekti tik kelis tūkstančius

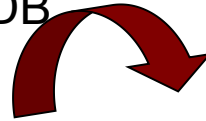
- Statistika iš “Protein Data Bank” (~30,000 struktūrų)

90% naujai įvestų struktūrų jau turi analogus su panašia sąsuka

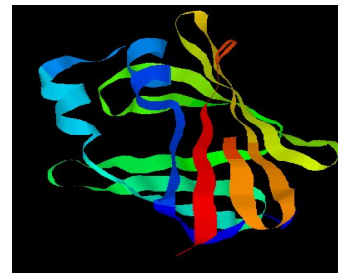
- Šansas, kad baltymas yra natyvioje būsenoje yra gan didelis
 - Baltymai su panašia sąsuka gali būti homologai arba analogai

Protein Threading

- **Tikslas:** surasti “tinkamą” sekos-struktūros palyginį tarp tiriamos sekos ir natyvių sąsukų iš PDB



MTYKLILN NGVDGEWTYTE



- **Energijos funkcija** – labiau paremta žiniomis ar statistika nei fizikiniu modeliu
 - Turėtų padėti atskirti teisingą struktūrinę sąsuką nuo neteisingos
 - Turėtų atskirti teisingą sekos-struktūros palyginį nuo neteisingo

Protein threading



- Panašios sekos turi panašias struktūras, bet būna išimčių
- Panašios struktūros gali turėti labai skirtingas sekas
Struktūriniai homologai ir analogai

Protein Threading – pagrindiniai komponentai



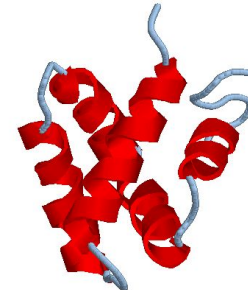
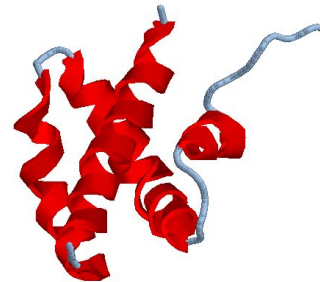
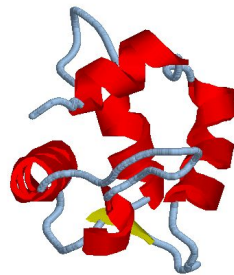
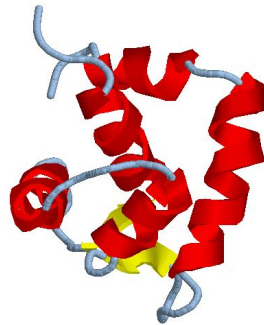
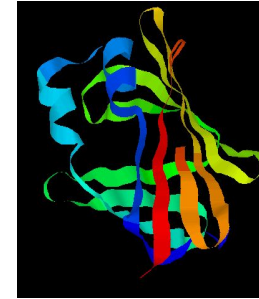
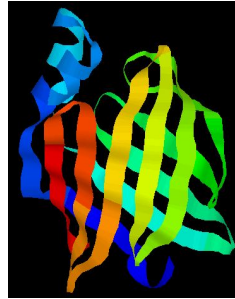
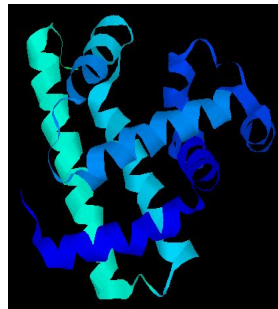
➤ Energijos funkcija



➤ Prognozavimo patikimumo įvertinimas

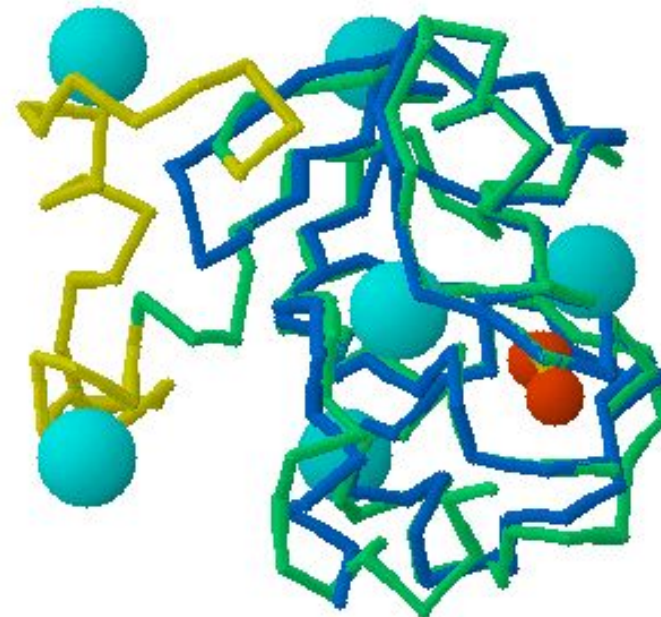
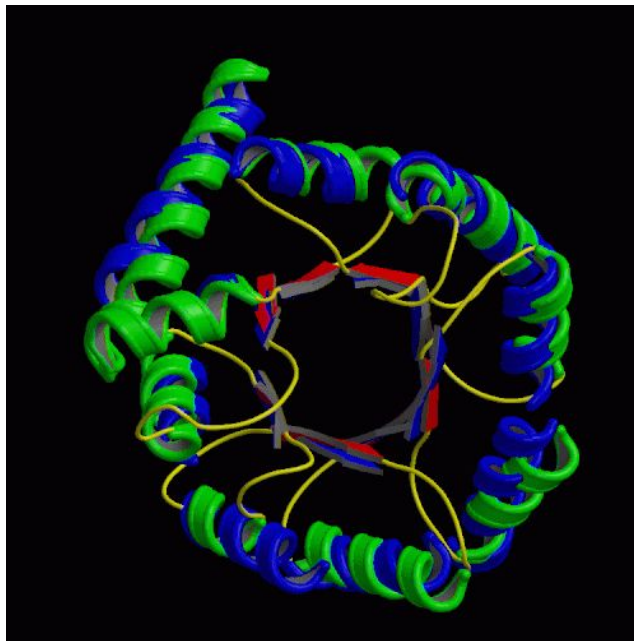
Protein Threading – struktūrų duomenų bazė

- Susikurti šablonų duomenų bazę



Protein Threading

- Dažnai pakanka naudoti unikalias struktūrines PDB sąsukas kaip šablonus

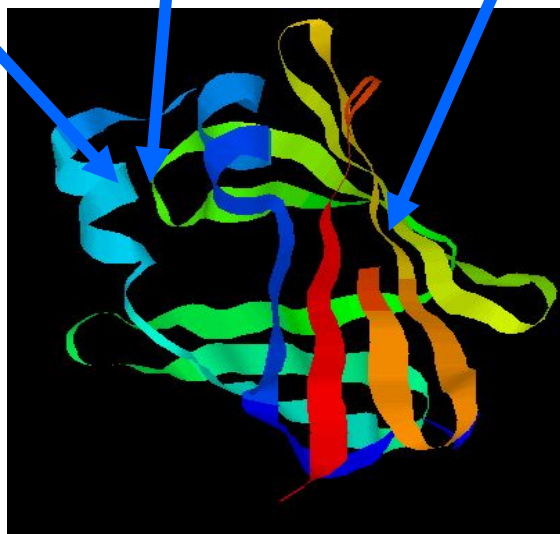


Protein Threading — energijos funkcija

MTYKLILNGKTKGETTTEAVDAATAEKVFQYANDNGVDGEWTYTE

Kaip geriausia būtų dvi
a.r. padėti šalia: E_p

Sulyginimo tarpo
bauda: E_g



Kaip a.r. “įsipaišo” į
aplinką: E_s

Suminė energija: $E_p + E_s + E_g$

**Surasti tokį struktūros-sekos
palyginį, kad energija būtų minimali**

Protein Threading



- Antrinė struktūra: alfa-spiralė, beta-lakštas, kilpa
- Prieinamumas solventui: 0, 10, 20, ..., 100% prieinamumo
- Kiekviena šių parametrų kombinacija apibūdina struktūrinę aplinką

E.g., (alpha-helix, 30%), (loop, 80%), ...

➤ E_s : įverčių matrica, pvz. 30 aplinkų ir 20 a.r.

- Pvz., $E_s((\text{kilpa}, 30\%), A)$

➤ $E_s(S, X) = -\log (FE(S, X)/FO(S, X))$

- $FE()$: spėjamas dažnumas
- $FO()$: stebimas dažnumas

“Singletoninis”
energijos terminas

Helix

	Buried	Inter	Exposed
ALA	-0.578	-0.119	-0.160
ARG	0.997	-0.507	-0.488
ASN	0.819	0.090	-0.007
ASP	1.050	0.172	-0.426
CYS	-0.360	0.333	1.831
GLN	1.047	-0.294	-0.939
GLU	0.670	-0.313	-0.721
GLY	0.414	0.932	0.969
HIS	0.479	-0.223	0.136
ILE	-0.551	0.087	1.248
LEU	-0.744	-0.218	0.940
LYS	1.863	-0.045	-0.865
MET	-0.641	-0.183	0.779
PHE	-0.491	0.057	1.364
PRO	1.090	0.705	0.236
SER	0.350	0.260	-0.020
THR	0.291	0.215	0.304
TRP	-0.379	-0.363	1.178
TYR	-0.111	-0.292	0.942
VAL	-0.374	0.236	1.144

Sheet

	Buried	Inter	Exposed
ALA	0.010	0.583	0.921
ARG	1.267	-0.345	-0.580
ASN	0.844	0.221	0.046
ASP	1.145	0.322	0.061
CYS	-0.671	0.003	1.216
GLN	1.452	0.139	-0.555
GLU	0.999	0.031	-0.494
GLY	0.177	0.565	0.989
HIS	0.306	-0.343	-0.014
ILE	-0.875	-0.182	0.500
LEU	-0.411	0.179	0.900
LYS	2.109	-0.017	-0.901
MET	-0.269	0.197	0.658
PHE	-0.649	-0.200	0.776
PRO	1.249	0.695	0.145
SER	0.303	0.058	-0.075
THR	0.156	-0.382	-0.584
TRP	-0.270	-0.477	0.682
TYR	-0.267	-0.691	0.292
VAL	-0.912	-0.334	0.089

Loop

	Buried	Inter	Exposed
ALA	0.023	0.218	0.368
ARG	0.930	-0.005	-0.032
ASN	0.030	-0.322	-0.487
ASP	0.308	-0.224	-0.541
CYS	-0.690	-0.225	1.216
GLN	1.326	0.486	-0.244
GLU	0.845	0.248	-0.144
GLY	-0.562	-0.299	-0.601
HIS	0.019	-0.285	0.051
ILE	-0.166	0.384	1.336
LEU	-0.205	0.169	1.217
LYS	1.925	0.474	-0.498
MET	-0.228	0.113	0.714
PHE	-0.375	-0.001	1.251
PRO	-0.412	-0.491	-0.641
SER	-0.173	-0.210	-0.228
THR	-0.012	-0.103	-0.125
TRP	-0.220	-0.099	1.267
TYR	-0.015	-0.176	0.946
VAL	-0.030	0.309	0.998

Protein Threading

- E_p : įverčių matrica iš 20 a.r. x 20 a.r.
- $E_p(X, Y, C) = -\log (FE(X, Y, C)/FO(X, Y, C))$
 - $FE()$: spėjamas dažnumas
 - $FO()$: stebimas dažnumas
 - X, Y : amino rūgštys
 - C : sąlyga – pvz., atstumas, sant. kampas, ...

Dviejų a.r. tarpusavio sąveikos energijos terminas

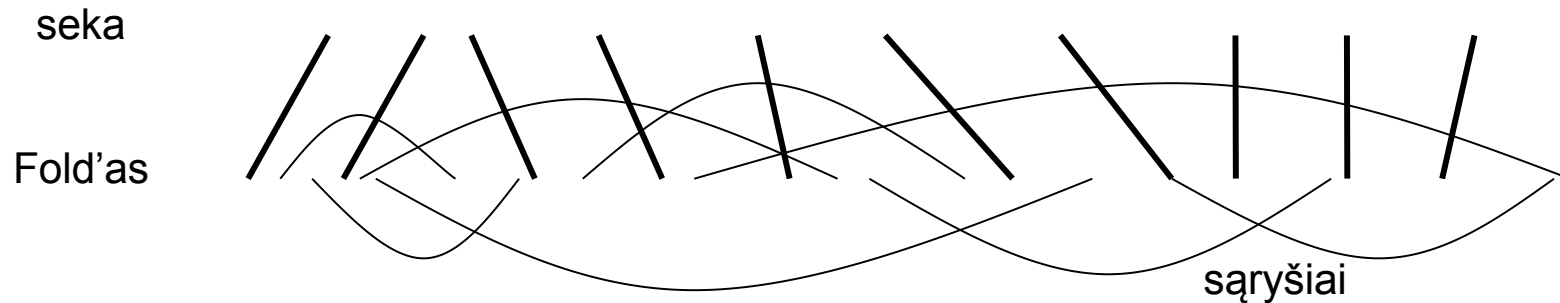
- E_g : sulyginimo tarpo bauda

$$-\log (\text{FE}(\mathbf{S}, \mathbf{X})/\text{FO}(\mathbf{S}, \mathbf{X}))$$

ALA	-140																				
ARG	268	-18																			
ASN	105	-85	-435																		
ASP	217	-616	-417	17																	
CYS	330	67	106	278	-1923																
GLN	27	-60	-200	67	191	-115															
GLU	122	-564	-136	140	122	10	68														
GLY	11	-80	-103	-267	88	-72	-31	-288													
HIS	58	-263	61	-454	190	272	-368	74	-448												
ILE	-114	110	351	318	154	243	294	179	294	-326											
LEU	-182	263	358	370	238	25	255	237	200	-160	-278										
LYS	123	310	-201	-564	246	-184	-667	95	54	194	178	122									
MET	-74	304	314	211	50	32	141	13	-7	-12	-106	301	-494								
PHE	-65	62	201	284	34	72	235	114	158	-96	-195	-17	-272	-206							
PRO	174	-33	-212	-28	105	-81	-102	-73	-65	369	218	-46	35	-21	-210						
SER	169	-80	-223	-299	7	-163	-212	-186	-133	206	272	-58	193	114	-162	-177					
THR	58	60	-231	-203	372	-151	-211	-73	-239	109	225	-16	158	283	-98	-215	-210				
TRP	51	-150	-18	104	52	-12	157	-69	-212	-18	81	29	-5	31	-432	129	95	-20			
TYR	53	-132	53	268	62	-90	269	58	34	-163	-93	-312	-173	-5	-81	104	163	-95	-6		
VAL	-105	171	298	431	196	180	235	202	204	-232	-218	269	-50	-42	46	267	73	101	107	-324	
	ALA	ARG	ASN	ASP	CYS	GLN	GLU	GLY	HIS	ILE	LEU	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	VAL	

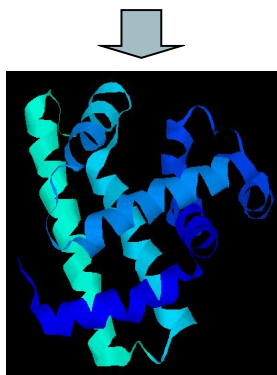
Protein Threading -- algoritmas

- **Threading algorithm** – reikia surasti minimalią energiją
 - Atsižvelgiant į visus energijos narius

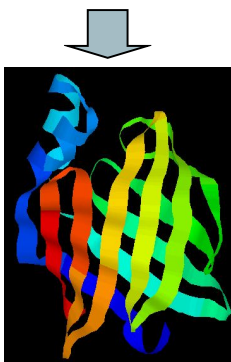


Prognozės patikimumo įvertis

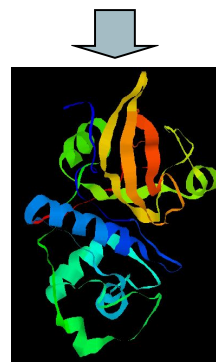
MTYKLILNGKTKGETTTEAVDAATAEKVFQYANDNGVDGEWTYPE



Score = -1500



Score = -720



Score = -1120



Score = -900

Kuri iš struktūrinių sąsukų yra teisinga?

Ar su didžiausiu įverčiu ?

Prognozės patikimumo įvertis

Seka: AAAA

Šablonas #1: AATTAATACATTAATATAATAAAAATTACTGA

Geresnis
šablonas?

Šablonas #2: CGGTAGTACGTAGTGTTTAGTAGCTATGAA

Kuri iš šių dviejų sekų duos geresnius
rezultatus po atsitiktinio išmaišymo?

Prognozės patikimumo įvertis

- Skirtingi šablonai gali turėti skirtingas vidutines bazines reikšmes todėl jų negalima tiesiogiai lyginti
- Rezultatai gali būti palyginami tik lyginant kiek įverčiai “išsišoka” iš vidutinių reikšmių

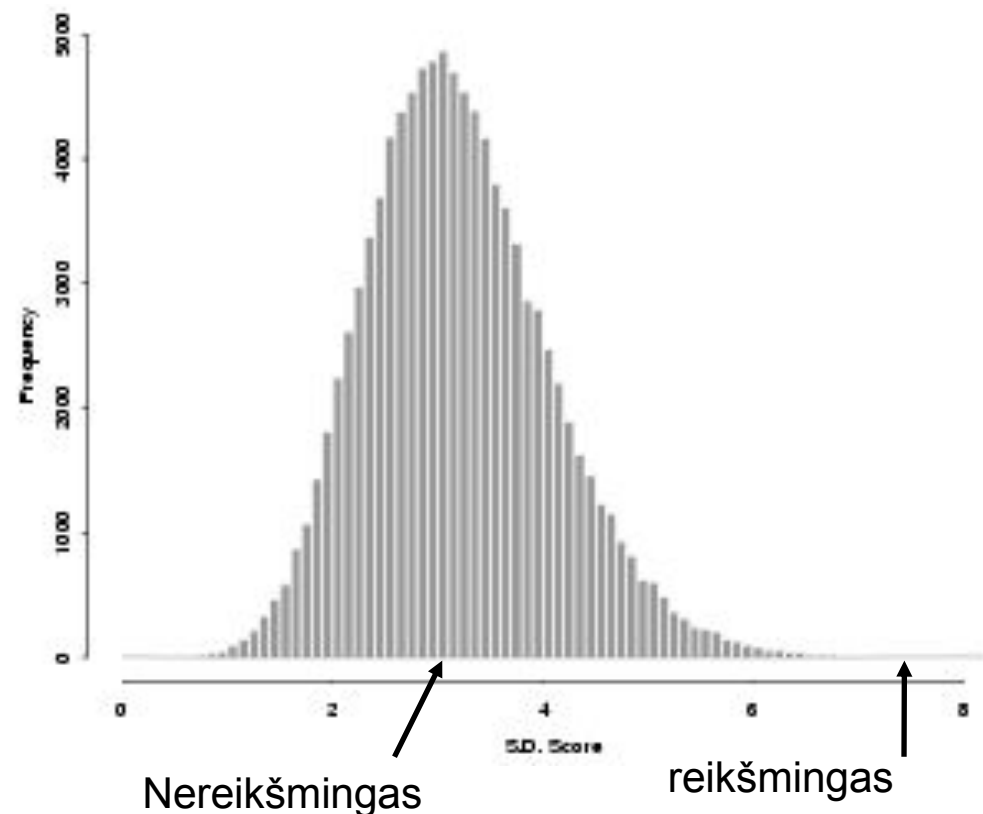
Prognozės patikimumo įvertis

"Pratepjam" 100,000 atsitiktinių sekų per šabloninę struktūrą ir gaunam atsitiktinių sekų įverčių pasiskirstymą.

Naudojant gautą pasiskirstymą nustatome kiek reikšmingas yra mūsų gautas įvertis.

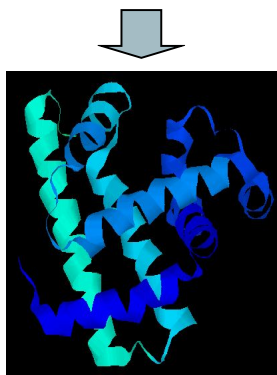
$$Z\text{-score} = \frac{\text{įvertis} - \text{vidurkis}}{\text{standartinis nuokrypis}}$$

E-value



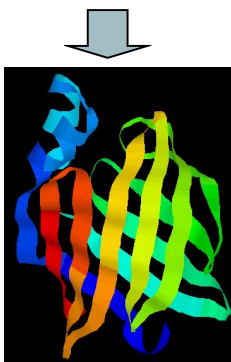
Prognozės patikimumo įvertis

MTYKLILNGKTKGETTTEAVDAATAEKVFQYANDNGVDGEWTYPE



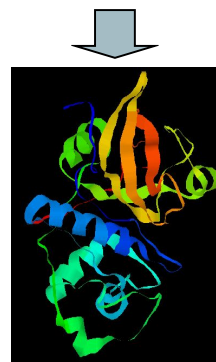
Score = -1500

E-value = e^{-1}



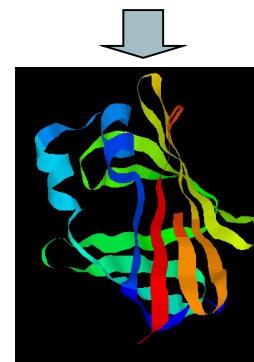
Score = -720

E-value = e^{-2}



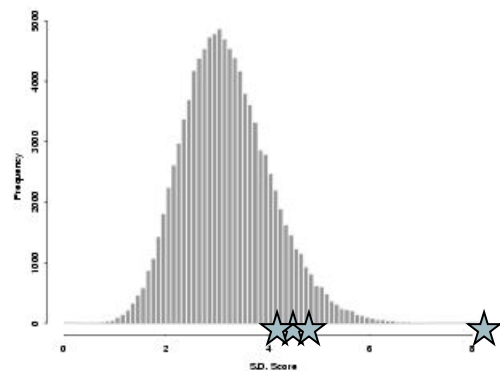
Score = -1120

E-value = $0.5 e^{-1}$



Score = -900

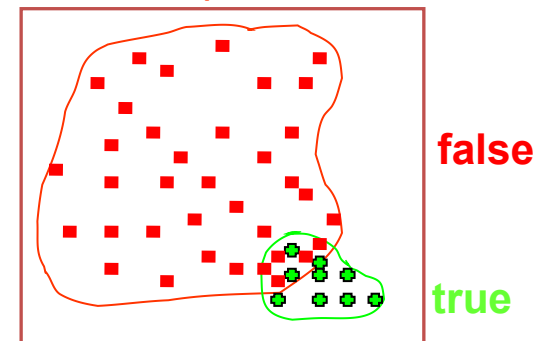
E-value = e^{-21}



Jeigu nėra patikimų ir reikšmių, atmesti
skaičiavimą

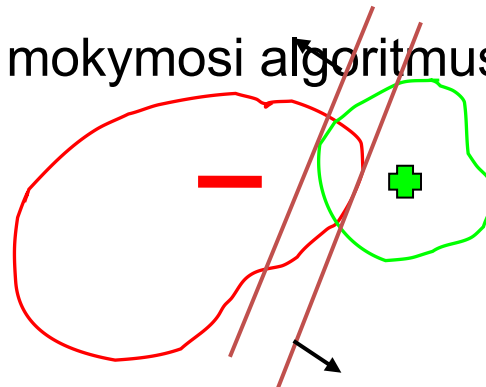
- Žar e įverčių skaičiavimas - labai resursams imlus procesas $O(NP)$. Bandoma išsiversti naudojant mokymosi algoritmus.
- Remdamesi žinomomis struktūromis ir jų sekomis įvertinam duotos struktūros "pratepimo" įverčio komponentus ir kuriam prognostinį modelį patikimiems įverčiams išskirti:
(E_s , E_p , antrinės struktūros įvertis, hidrofobiškumo įvertis ir t.t. dropobic moment score,) vs **true/false foldo atpažinime.**

-2000, -500, -35, -90,, true
 -1000, -201, -11, -500,, false
 -5020, -900, -20, -75,, true
 -1050, -185, -18, -320,, false



.....

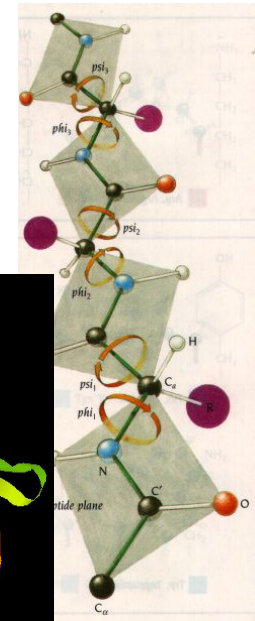
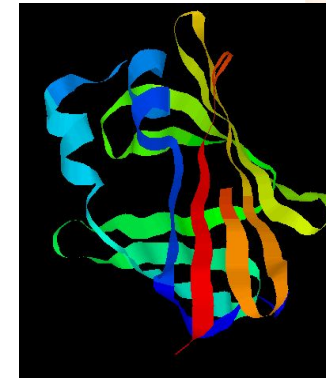
- Naudojant mokymosi algoritmus sukuriamas klasifikatorius.
Pvz SVM



Baltymų struktūros prognozė

- Leisti paiešką ant šablonų duomenų bazės
- Palikti rezultatus su gerom e-reikšmėm, pvz., $< e^{-10}$

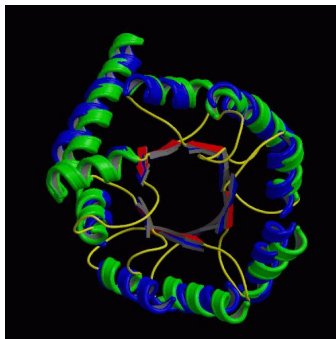
FMFTAIGEEVVQRSRKIL- - - DDLVELVK
AVLTRYGQRRLIQLYDLLAQIQQKAFDVLS



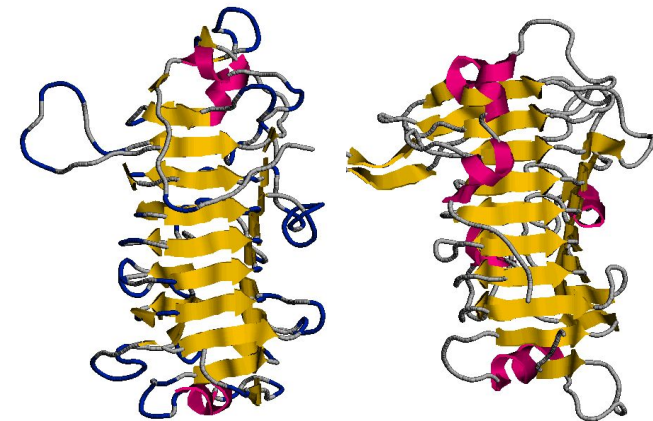
- “Padėti skeleto atomus į reikiamas vietas pagal a.r. padėti
a.r. kurių nepavyko
padėti, atmesti

Baltymų struktūros prognozė

- “Protein threading” metodas tinkamas tik skeleto struktūros prognozei (šoninių grandinių padėtys turi būti skaičiuojamos kitais metodais)



Mėlyna: tikra struktūra
Žalia: prognozuota



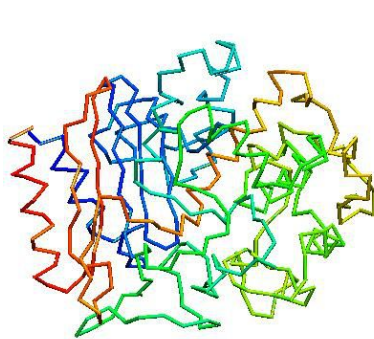
prognozuota

tikra

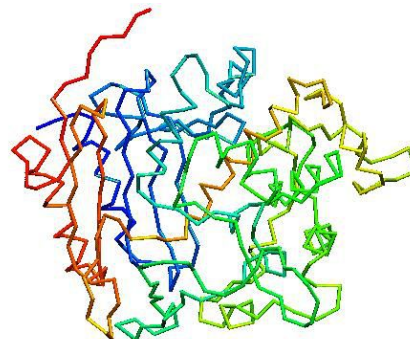
- Kuo mažesnė e-reikšmė, tuo tikslesnė prognozė

Baltymų struktūros prognozė

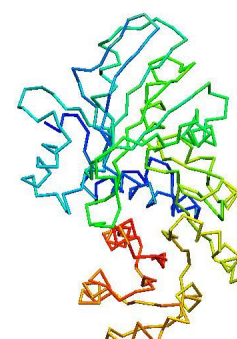
➤ Pavyzdžiai



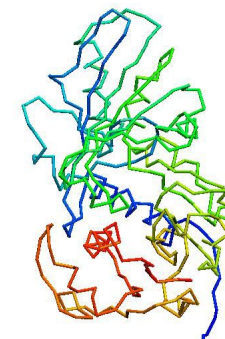
Tikra



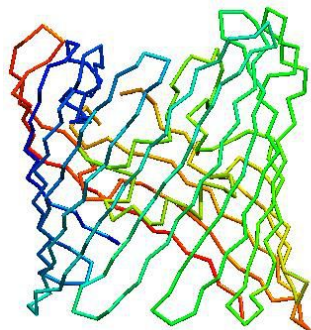
prognozuota



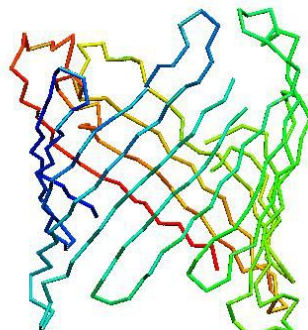
Tikra



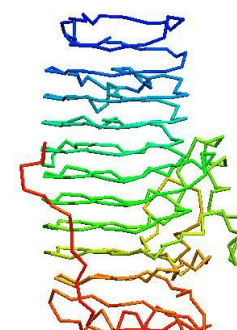
prognozuota



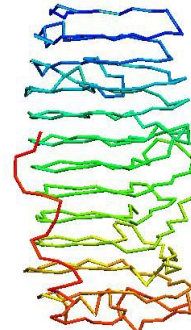
Tikra



prognozuota



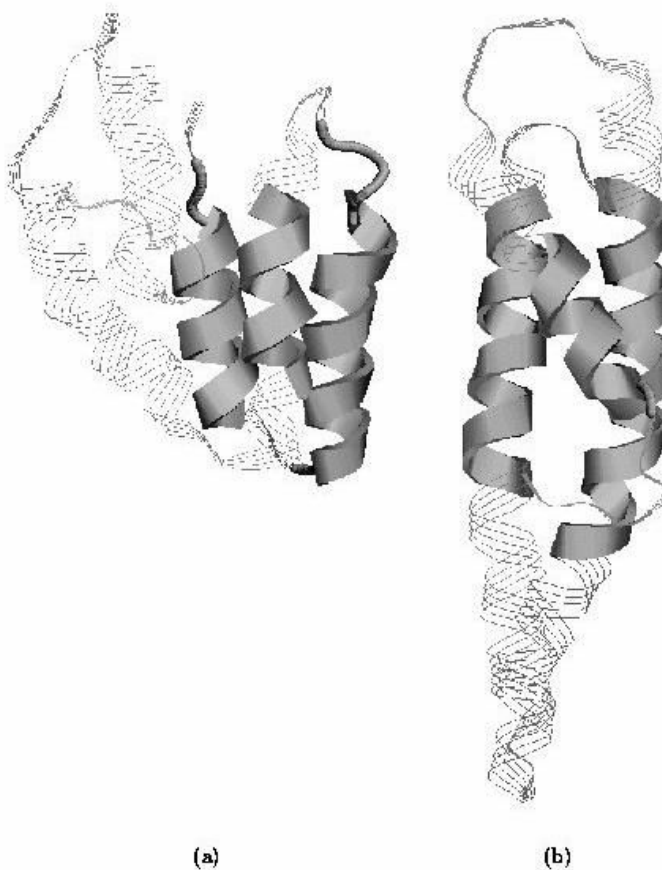
Tikra



prognozuota

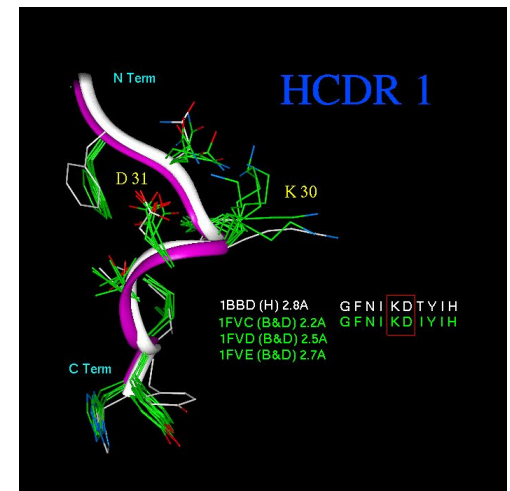
Baltymų struktūros prognozė

- Bet gali gautis ir taip:



Baltymų struktūros prognozė

- Struktūros su visais atomais gali būti prognozuojama:
 - Prognozuojant pagrindinę grandinę (skeletą)
 - Prognozuojant šoninių radikalų išsidėstymą:
 - Nuo skeleto priklausantys rotamerai
 - Ab initio* šoninių grandinių prognozė



- Pasiekimai – vis dar šoninių radikalų prognozės yra iššūkis

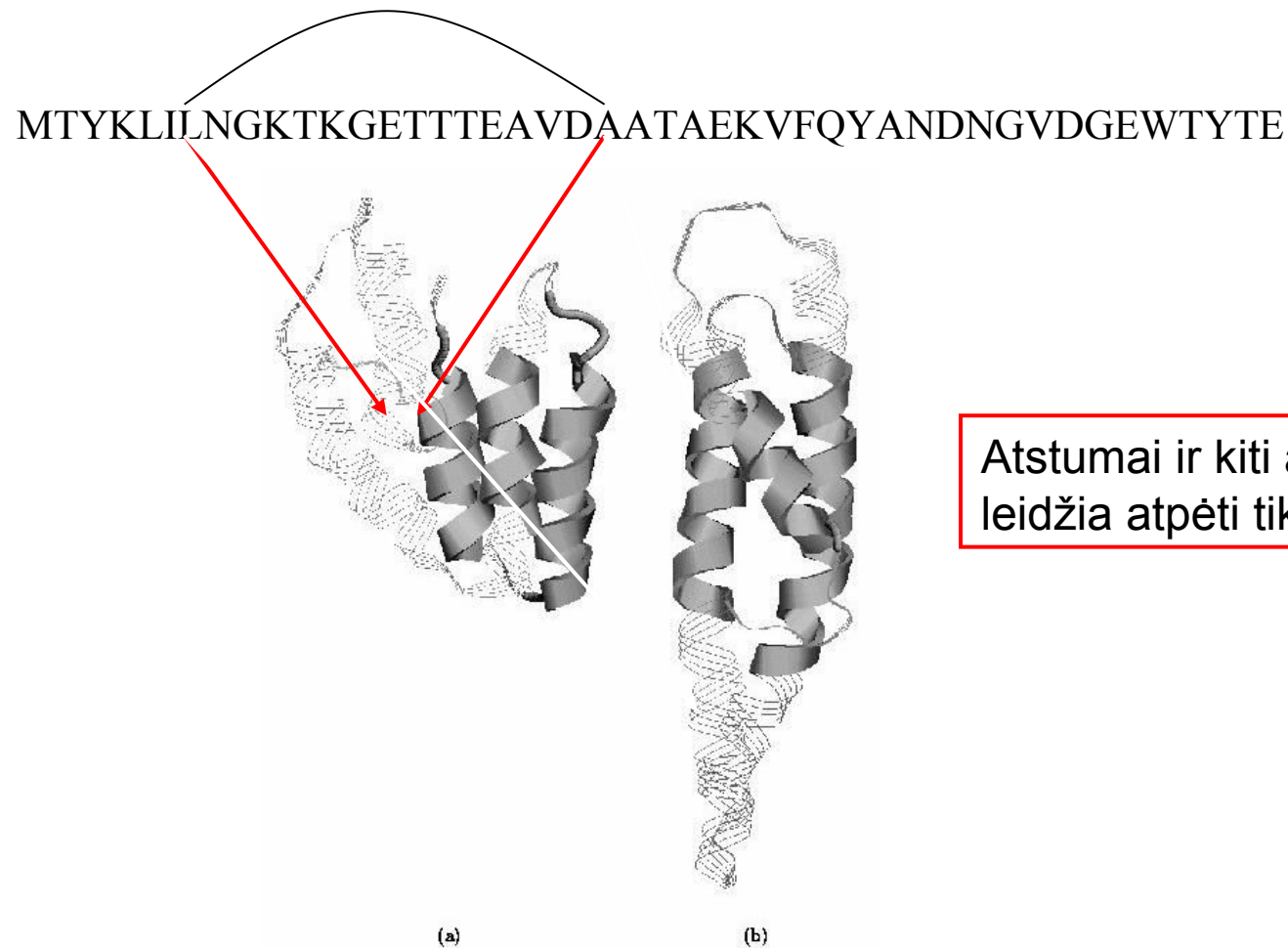
Struktūros prognozė naudojant papildomą informaciją

- Kai kokia struktūrinė informacija gali būti panaudota prieš spendžiant problemą
 - Disulfidinės jungtys
 - Aktyvios vietos
 - Nustatyta amino rūgščių santykinė padėtis (paslėpta residues identified buried/exposed)
 - (dalinė) antrinė struktūra
 - Daliniai NMR duomenys
 - Grupių tarpusavio atstumai iš susiuvimo ir masių spektrometrijos duomenų
 - Bendra išorinė forma iš cryo-EM
 -

- Šie duomenys gali suteikti labai naudingų apribojimų struktūros prognozavime

Struktūros prognozė naudojant papildomą informaciją

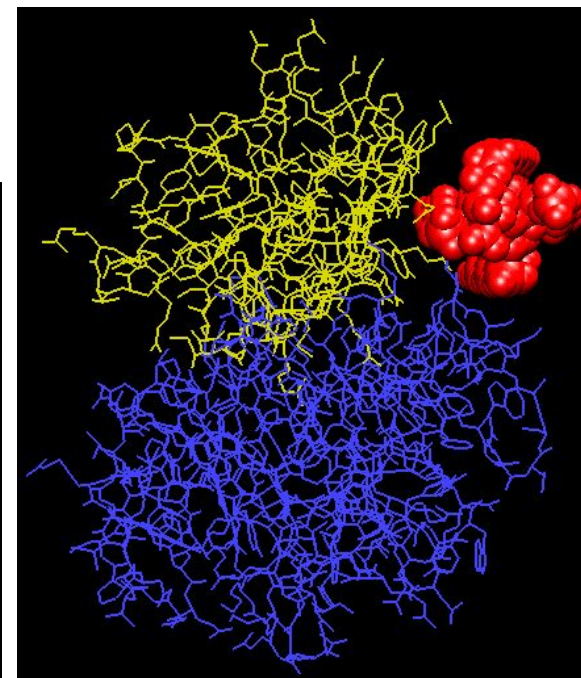
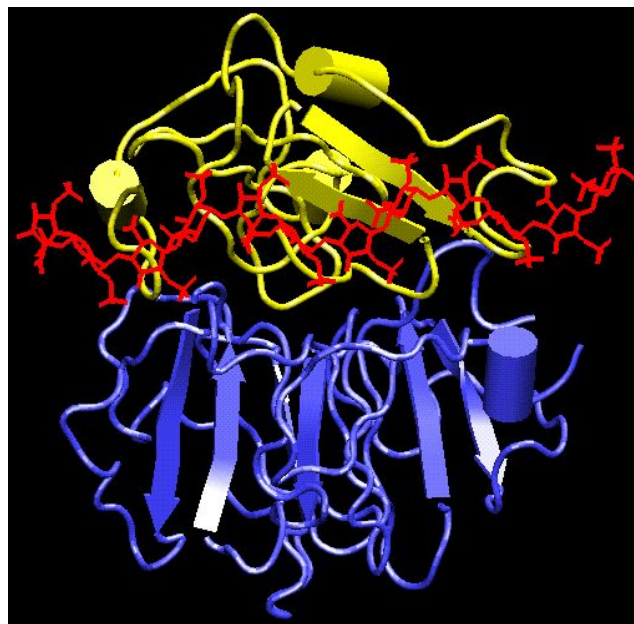
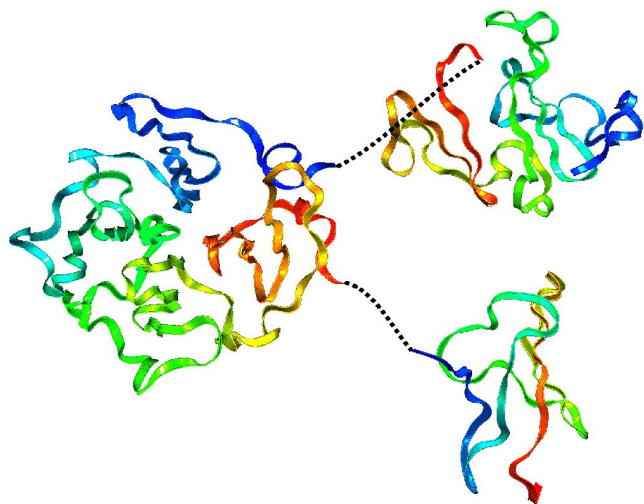
➤ Pagrindinė idėja



Atstumai ir kiti apribojimų tipai
leidžia atpėti tiksliau

Struktūros prognozė naudojant papildomą informaciją

- Viktronektinas (victronectin): trijų domenų baltymas
 - Įvairūs struktūriniai duomenys gali būti gauti iš eksperimentų, pvz.: disulfidinės jungtys, **aktyvios vietos**, heparino prisijungimo vietos, **kirpimo vietos**,
 - Taigi duomenys apribojami prognozuojant



Esančios struktūros prognozavimo programos

➤ PROSPECT

- https://csbl.bmb.uga.edu/protein_pipeline

➤ FUGU

- <http://www-cryst.bioc.cam.ac.uk/~fugue/prfsearch.html>

➤ THREADER

- <http://bioinf.cs.ucl.ac.uk/threader/>

CASP <http://predictioncenter.org/>

www.predictioncenter.org/index.cgi

images-1253.jpg Naudojimas naršyklėje Lankomiausi Labai šokoladinis pyra...

Protein Structure Prediction Center

Menu

- [Home](#)
- [FORCASP Forum](#)
- [PC Login](#)
- [PC Registration](#)
- CASP Experiments**
 - [CASP ROLL](#)
 - CASP11 (2014)**
 - [Home](#)
 - [My CASP11 profile](#)
 - [Targets](#)
 - [CASP11 in numbers](#)
 - [CASP10 \(2012\)](#)
 - [CASP9 \(2010\)](#)
 - [CASP8 \(2008\)](#)
 - [CASP7 \(2006\)](#)
 - [CASP6 \(2004\)](#)
 - [CASP5 \(2002\)](#)
 - [CASP4 \(2000\)](#)
 - [CASP3 \(1998\)](#)
 - [CASP2 \(1996\)](#)
 - [CASP1 \(1994\)](#)
- Initiatives**
- Data Archive**
- Local Services**

Welcome to the Protein Structure Prediction Center!

Our goal is to help advance the methods of identifying protein structure from sequence. The Center has been organized to provide the means of objective testing of these methods via the process of blind prediction. The Critical Assessment of protein Structure Prediction (CASP) experiments aim at establishing the current state of the art in protein structure prediction, identifying what progress has been made, and highlighting where future effort may be most productively focused.

There have been ten previous CASP experiments. The eleventh experiment will start in May 2014. Description of these experiments and the full data (targets, predictions, interactive tables with numerical evaluation results, dynamic graphs and prediction visualization tools) can be accessed following the links:

[CASP1 \(1994\)](#) | [CASP2 \(1996\)](#) | [CASP3 \(1998\)](#) | [CASP4 \(2000\)](#) | [CASP5 \(2002\)](#) | [CASP6 \(2004\)](#) | [CASP7 \(2006\)](#) | [CASP8 \(2008\)](#) | [CASP9 \(2010\)](#) | [CASP10 \(2012\)](#) | [CASP11 \(2014\)](#)

Raw data for the experiments held so far are archived and stored in our [data archive](#).

In November 2011 we have opened a new rolling CASP experiment for all-year-round testing of ab initio modeling methods:

[CASP ROLL](#)

Details of the experiments have been published in a scientific journal *Proteins: Structure, Function and Bioinformatics*. [CASP proceedings](#) include papers describing the structure and conduct of the experiments, the numerical evaluation measures, reports from the assessment teams highlighting state of the art in different prediction categories, methods from some of the most successful prediction teams, and progress in various aspects of the modeling.

Prediction methods are assessed on the basis of the analysis of a large number of blind predictions of protein structure. Summary of numerical evaluation of the methods tested in the latest CASP experiment can be found [on this web page](#). The main numerical measures used in evaluations are described in the papers [\[1\]](#), [\[2\]](#). The latter paper also contains explanations of data handling procedures and guidelines for navigating the data presented on this website.

Some of the best performing methods are implemented as [fully automated servers](#) and therefore can be used by public for protein structure modeling.

Message Board

Ground transportation address

[To conference attendees: Please make sure you arrive at Paraiso Maya, not Lindo. The ground transportation address is: Iberostar Paraiso Maya, km 309 on the coast highway. CASP organizers ...](#)

Program, restaurants, posters

[Program: As you may have noticed, the final program of the meeting has been posted at our webpage. Restaurants at the hotel: Breakfast and lunch do not require a reservation. Also for dinner, t ...](#)

Final preparations for the meeting in Mexico

[Transfers from /to Cancun airport: There are taxis, shuttles, rental cars \(sometimes very cheap + hotel parking is free\). You may use the transportation ...](#)

centr

Viską paryškinti Skirti raidžių registrą 1 iš 11 atitikmenų

„Firefox“ automatiškai siunčia tam tikrus duomenis į „Mozilla“ programos gerinimo tikslais.

Pažirinkti, kas bus siunčiama

← →

zhanglab.ccmb.med.umich.edu/I-TASSER/

itasser

images-1253.jpg Naudojimasis naršykle Lankomiausi Labai šokoladinis pyra...

Online Services

I-TASSER

QUARK

LOMETS

COACH

COFACTOR

MUSTER

SEGMENT

FG-MD

ModRefiner

REMO

SPRING

COTH

BSpred

SVMSEQ

ANGLOR

BSP-SLIM

SAXSTER


ThreaDom

EvoDesign

gPCR-I-TASSER

TM-score

TM-align



I-TASSER ONLINE

Protein Structure & Function Predictions

(The server completed predictions for 203989 proteins submitted by 51057 users from 116 countries)
(The template library was updated on 2014/12/03)

I-TASSER server is an on-line platform for protein structure and function predictions. 3D models are built based on multiple-threading alignments by LOMETS and iterative template fragment assembly simulations; function insights are derived by matching the 3D models with BioLiP protein function database. I-TASSER (as 'Zhang-Server') was ranked as the No 1 server for protein structure prediction in recent CASP7, CASP8, CASP9, and CASP10 experiments. It was also ranked as the best for function prediction in CASP9. The server is in active development with the goal to provide the most accurate structural and function predictions using state-of-the-art algorithms. The server is only for non-commercial use. Please report problems and questions at I-TASSER message board and some members will study and answer the questions asap. ([>> More about the server ...](#))

Download I-TASSER Standalone Package (Version 4.2) NEW

[\[Queue\]](#) [\[Forum\]](#) [\[Download\]](#) [\[Example\]](#) [\[Search\]](#) [\[Registration\]](#) [\[About\]](#) [\[Statistics\]](#) [\[Remove\]](#) [\[Potential\]](#) [\[Decoys\]](#) [\[News\]](#)

Copy and paste your sequence here (<1,500 residues, in FASTA format):

Or upload the sequence from your local computer:

Parinkti... Neparinktas joks failas.

centr

Viską paryškinti Skirti raidžių registrą 1 iš 11 atitikmenų

„Firefox“ automatiškai siunčia tam tikrus duomenis į „Mozilla“ programos gerinimo tikslais.

Pagirinkti, kas bus siunčiama

