

Evolucija proteinov

Predvideva se, da v naravi obstaja okoli 1000 – 2000 različnih vzorcev zvitja.

Domena je osnovna enota terciarne strukture proteina in predstavlja del polipeptidne verige s samostojnim stabilnim zvitjem.

Celoten repertoar proteinov se je razvil z mehanizmi:

1. Divergentne evolucije

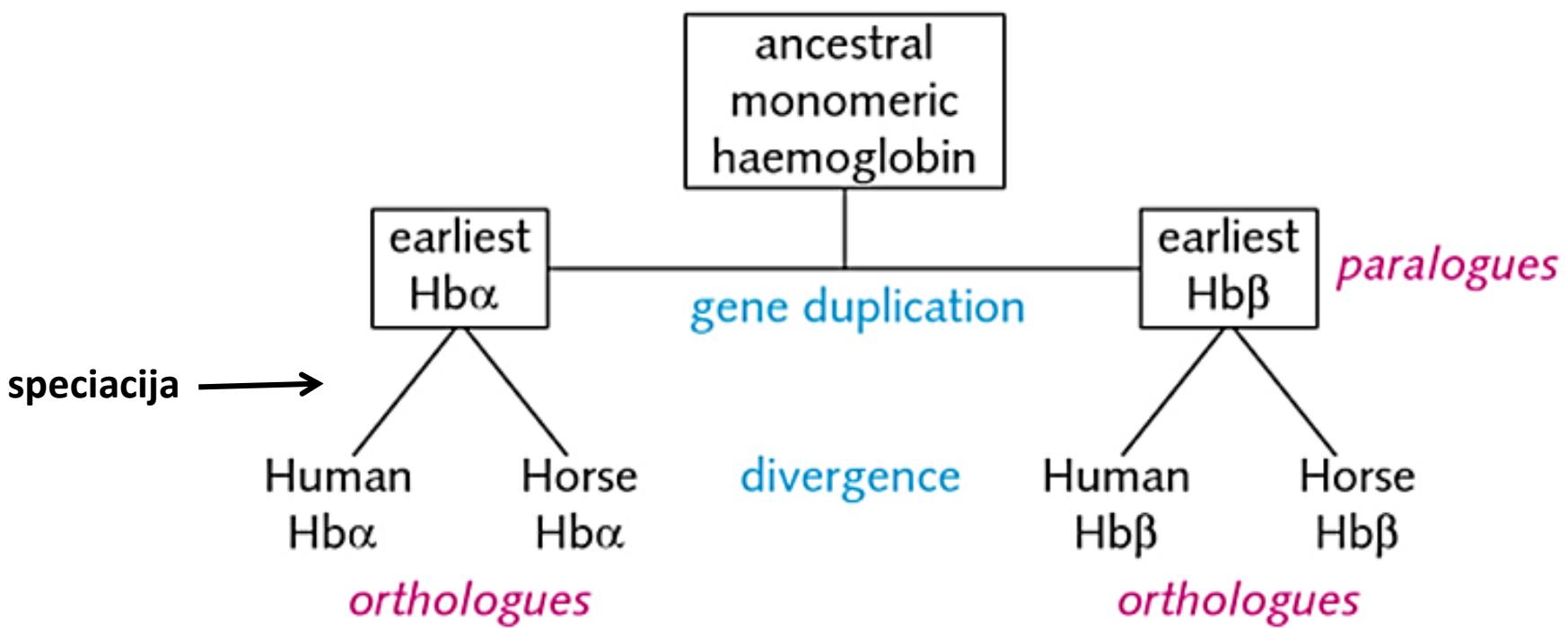
Podobnost – skupne lastnosti na nivoju primarne, sekundarne in terciarne strukture

Homologija – izvor iz skupnega prednika

2. Tvorbo kompleksnih struktur iz enostavnejših elementov

- a. Podaljševanje polipeptidne verige znotraj enote zvitja
- b. Tvorba oligomerov iz monomerov
- c. Tvorba večdomenskih proteinov

Evolucija proteinov

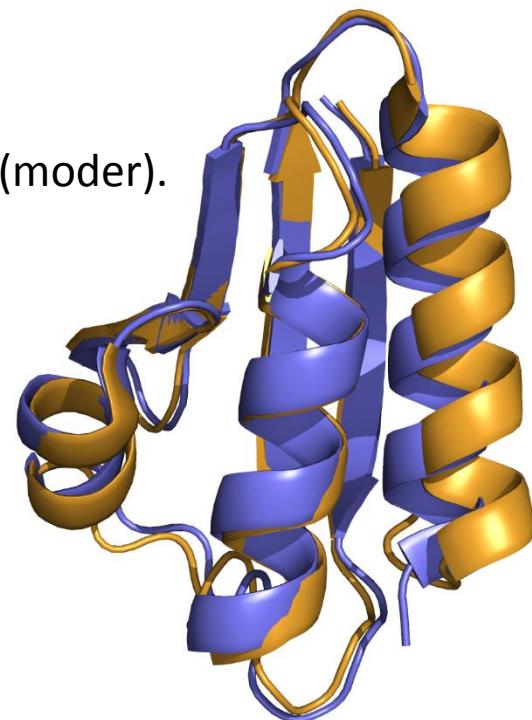


Zvitje proteinov

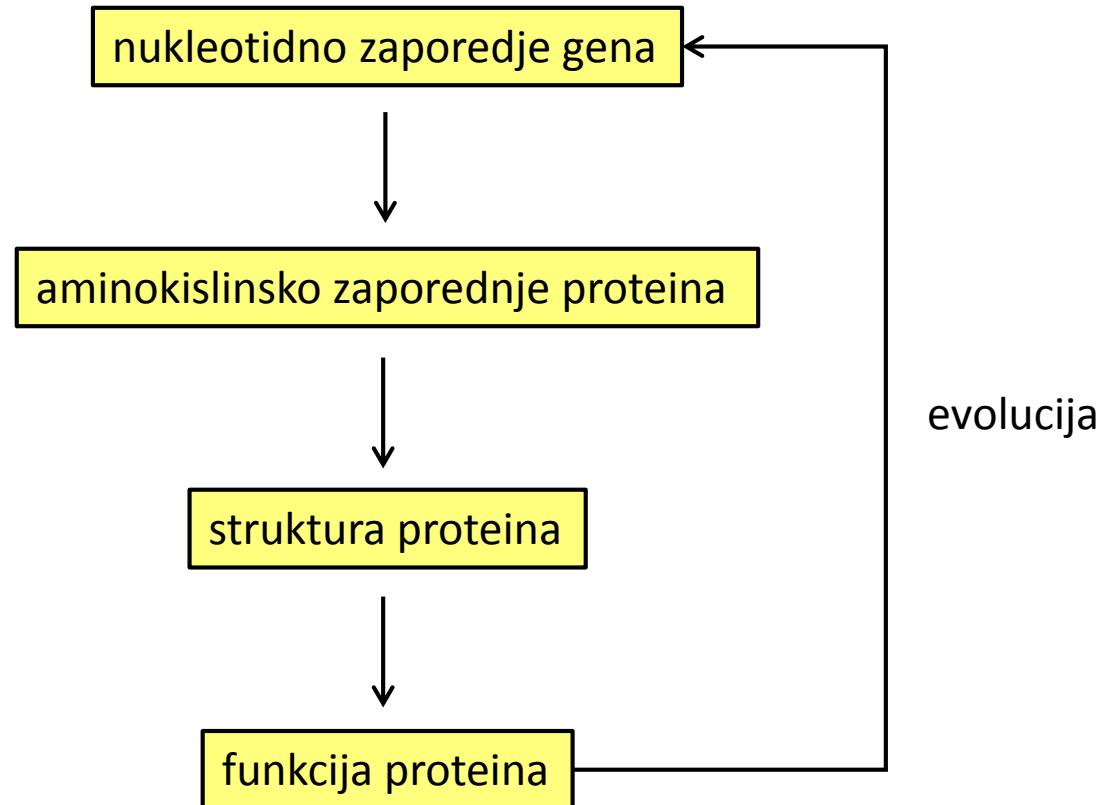
Primer homolognih proteinov:

Proteina HPr iz *Escherichie coli* (oranžen) in *Streptococcus faecalis* (moder).

Ohranjeno ujemanje terciarne strukture in elementov sekundarne strukture, le 38% ujemanje na nivoju primarne strukture (zaporedja).



Povezava med strukturo, funkcijo in evolucijo proteinov

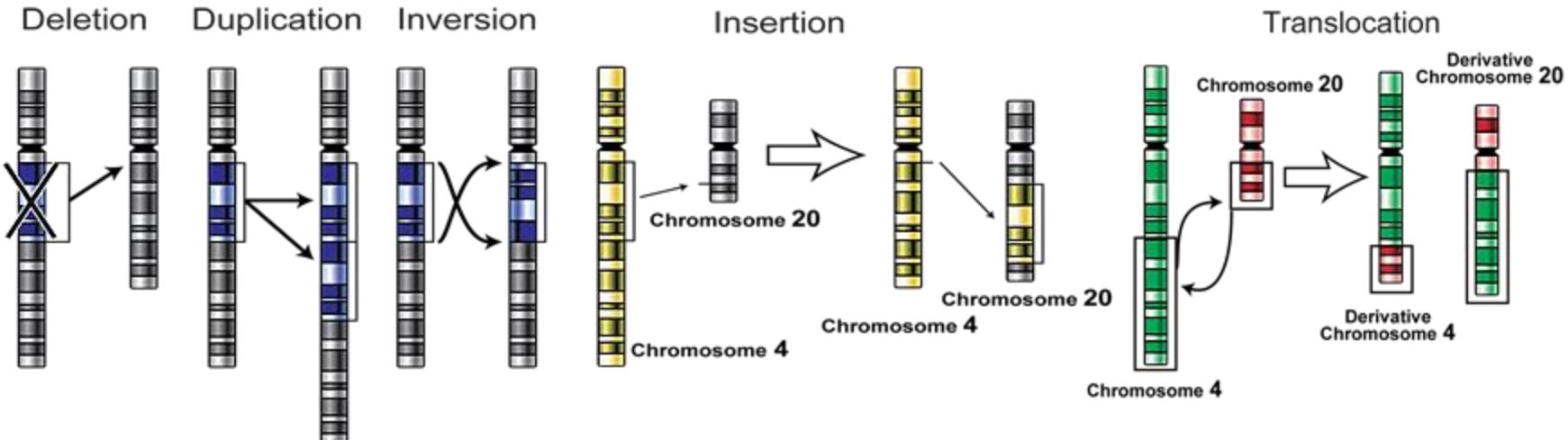


Mutacije

Mutacije so gonilna sila evolucije.

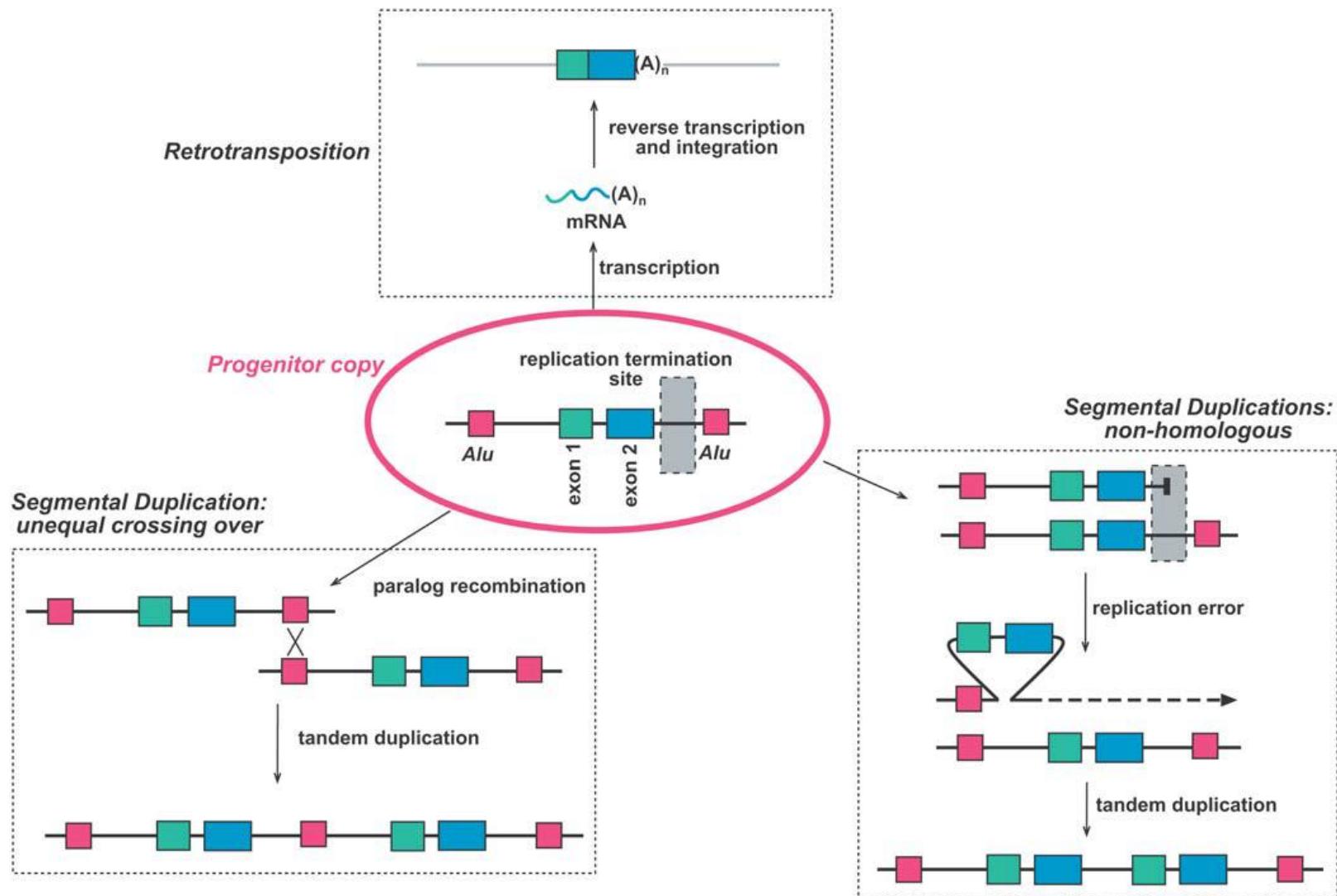
No mutation	Point mutations			
	Silent	Nonsense	Missense	
DNA level	TTC	TTT	ATC	TCC
mRNA level	AAG	AAA	UAG	AGG
protein level	Lys	Lys	STOP	Arg
			conservative	non-conservative

Types of Mutations



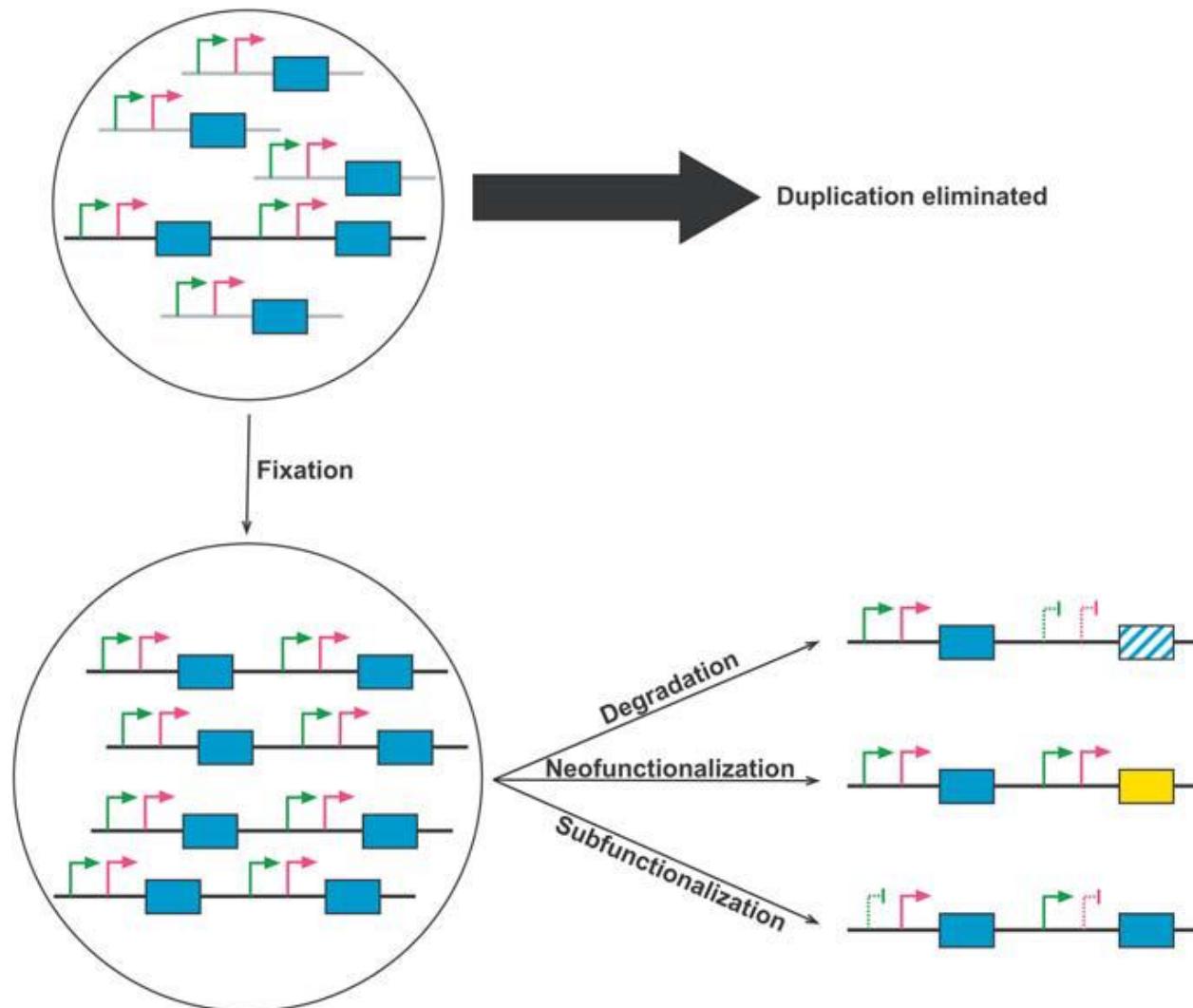
Duplikacije genov

Z duplikacijami genov nastajajo novi paralogi, ki lahko pridobijo nove funkcije.



Duplikacije genov

Z duplikacijami genov nastajajo novi paralogi, ki lahko pridobijo nove funkcije.



Usoda mutacij

Usoda mutacija je odvisna od njenega vpliva na **fitnes** genotipa.

Mutacije so lahko nevtralne, uničujoče ali prednostne za genotip.

Na usodo mutacij vplivata dva osnovna efekta:

Naravna selekcija favorizira genotipe z boljšim fitnesom. Neugodne mutacije se odstranijo z *negativno selekcijo*. Prednostne se v populaciji fiksirajo s *pozitivno selekcijo*. Signifikantne so tudi malenkostne razlike (1%).

Naključni genetski tok omogoča naključno spremembe pogostosti alelov v populaciji.

Usoda mutacij

Usoda mutacija je odvisna od njenega vpliva na **fitnes** genotipa.

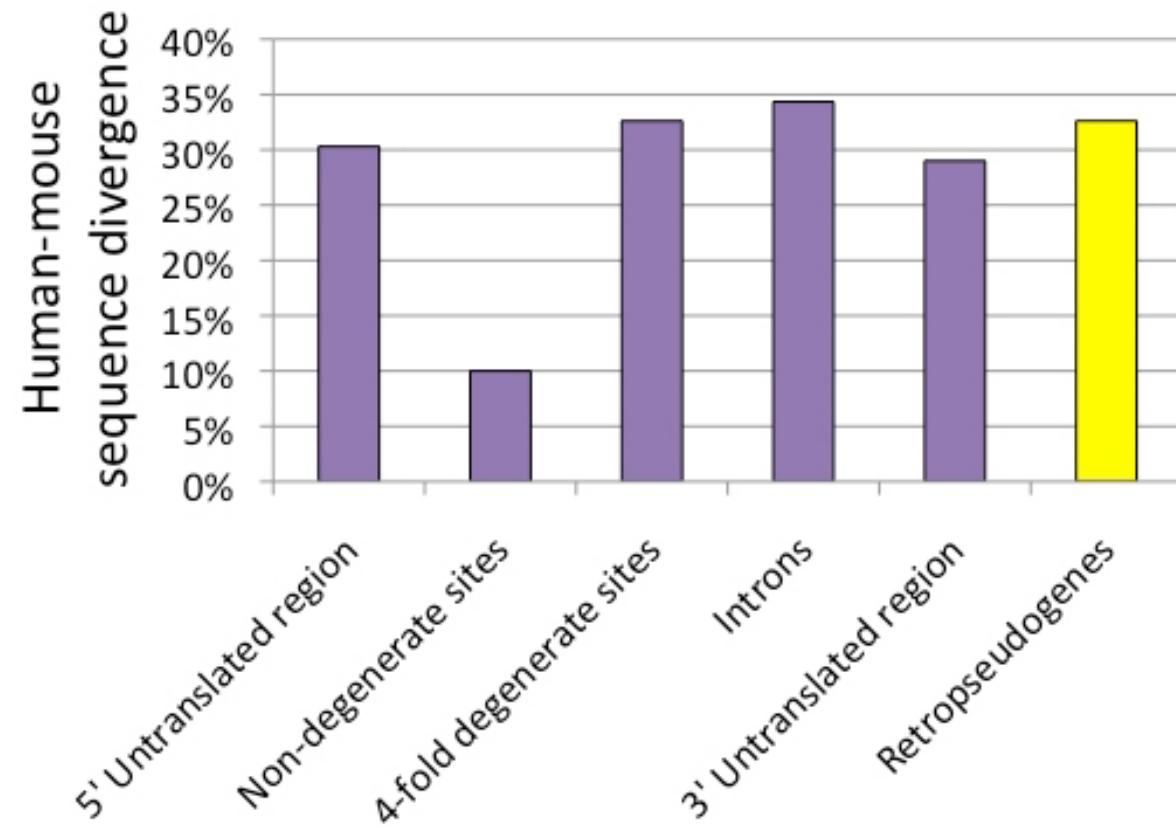
Mutacije so lahko nevtralne, uničajoče ali prednostne za genotip.

Neo-Darwinizem pravi, da ima ključno vlogo v evoluciji naravna selekcija, naključni faktorji pa so manj pomembni. **Selekcionizem** smatra naravno selekcijo za edino gonilno silo. Nov alel se lahko fiksira le, če izboljša fitnes organizma.

Kimurina nevtralna teorija molekulske evolucije pravi, da je večina mutacij nevtralnih ali skoraj nevtralnih, in da je gonilna sila evolucije naključni genetski tok, v manjši meri selekcija.

Usoda mutacij

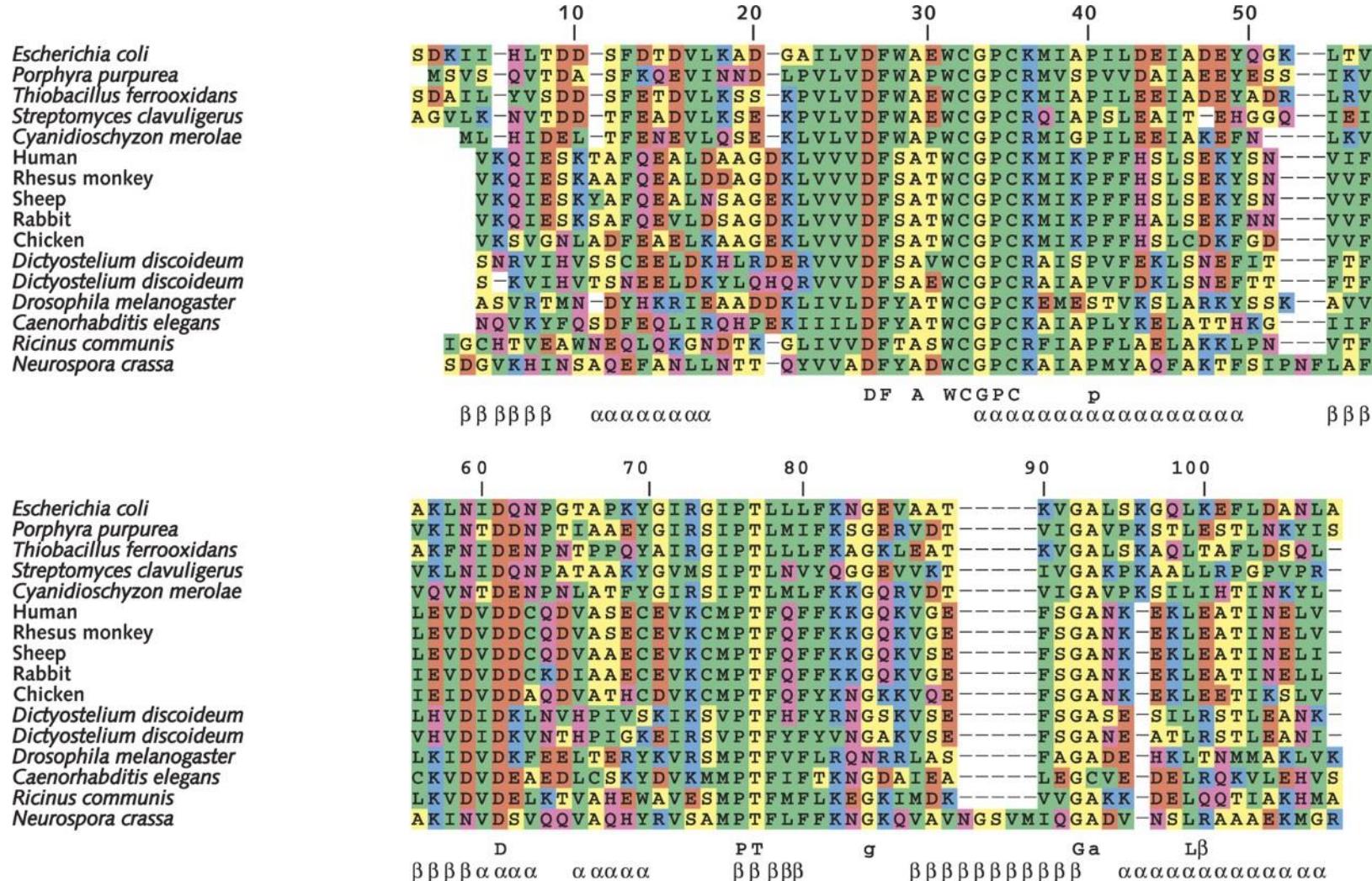
Neenaka pogostost divergence med posameznimi regijami med človeškim in mišjim genomom potrjuje veljavnost nevtralne teorije evolucije.



Poravnavo proteinikh zaporedij

Evolucijskim dogodkom sledimo s strukturno poravnavo homolognih proteinov.

Primer: poravnavo zaporedij tioredoksinov iz različnih organizmov



Poravnavo proteinskih zaporedij

Evolucijskim dogodkom sledimo s strukturno poravnavo homolognih proteinov.

Primer: poravnavo zaporedij tioredoksinov iz različnih organizmov

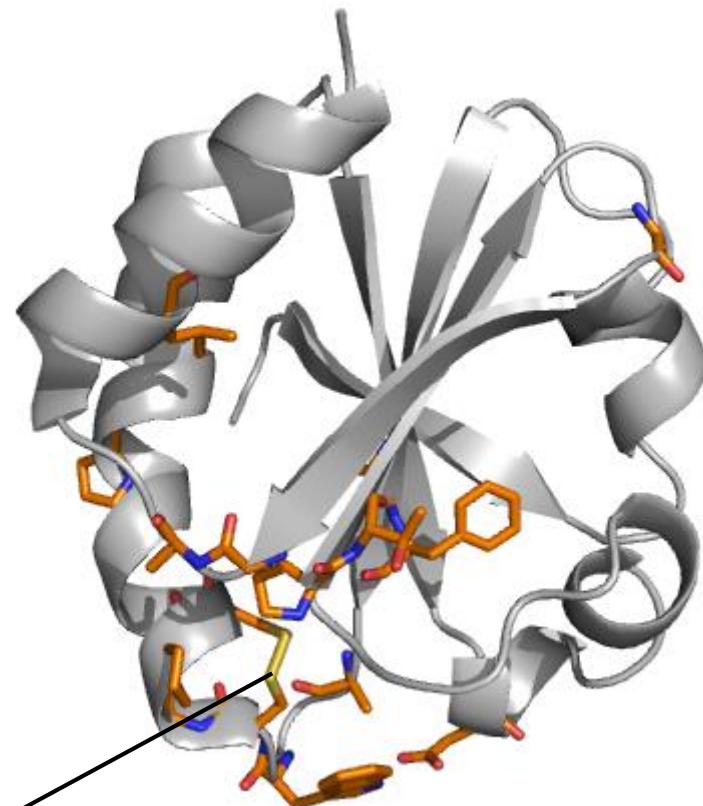
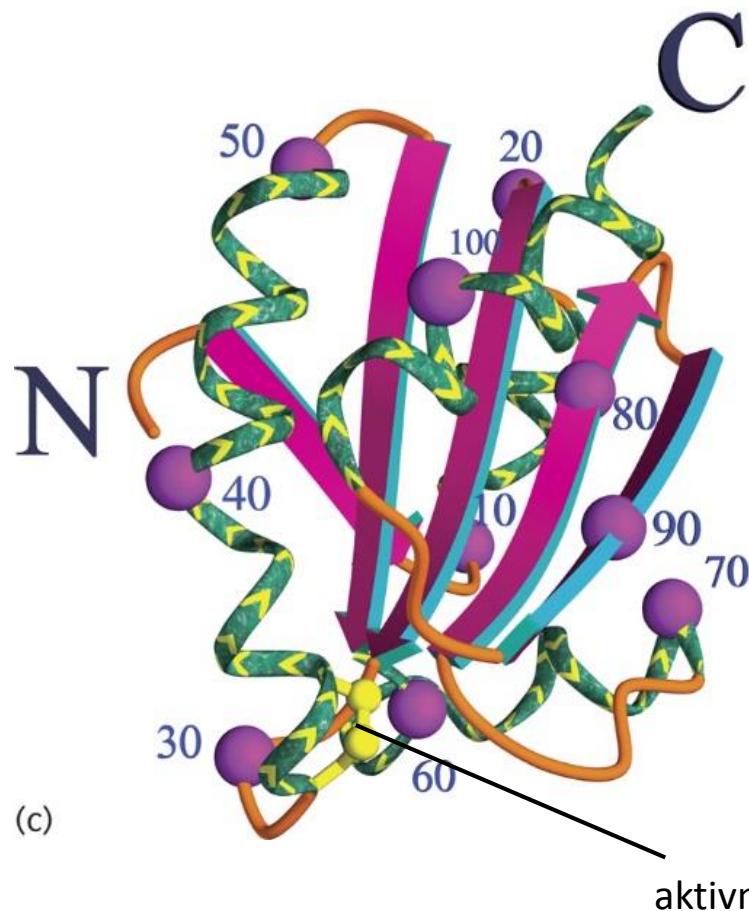
Thioredoxin



Poravnavo proteinских zaporedij

Evolucijskim dogodkom sledimo s strukturno poravnavo homolognih proteinov.

Primer: poravnavo zaporedij tioredoksinov iz različnih organizmov



Poravnavo proteinskih zaporedij

Evolucijskim dogodkom sledimo s strukturno poravnavo homolognih proteinov. V primeru dostopnosti tridimenzionalnih struktur, je najbolje narediti poravnavo le-teh, zlasti v primeru bolj divergentnih proteinov. Strukturna poravnava nato služi kot vodilo za poravnavo zaporedij.

Strukturna poravnava pomeni prileganje vseh izbranih atomov enega proteina (ponavadi C α) atomom drugega proteina z operacijama rotacije in translacije. Kriterij za optimalno poravnavo je minimalna vrednost korena kvadratov razdalj med vsemi atomi na ekvivalentnih pozicijah.

$$r.m.s.d. = \sqrt{\frac{\sum_i d_i^2}{N}}$$

kjer so d_i razdalje med koordinatami ekvivalentnih atomov v obeh modelih po končani superpoziciji.

Strežniki za prostorske poravnave: DALI, MUSTANG

Poravnavo proteinskih zaporedij

Z višanjem evolucijske oddaljenosti med dvema homologoma prihaja do vedno večje divergencije - podobnost na nivoju primarne in terciarne strukture pada.

Primer: papain (iz papaje, *Carica papaya*) in aktinidin (iz kivija *Actinidia chinensis*)

```

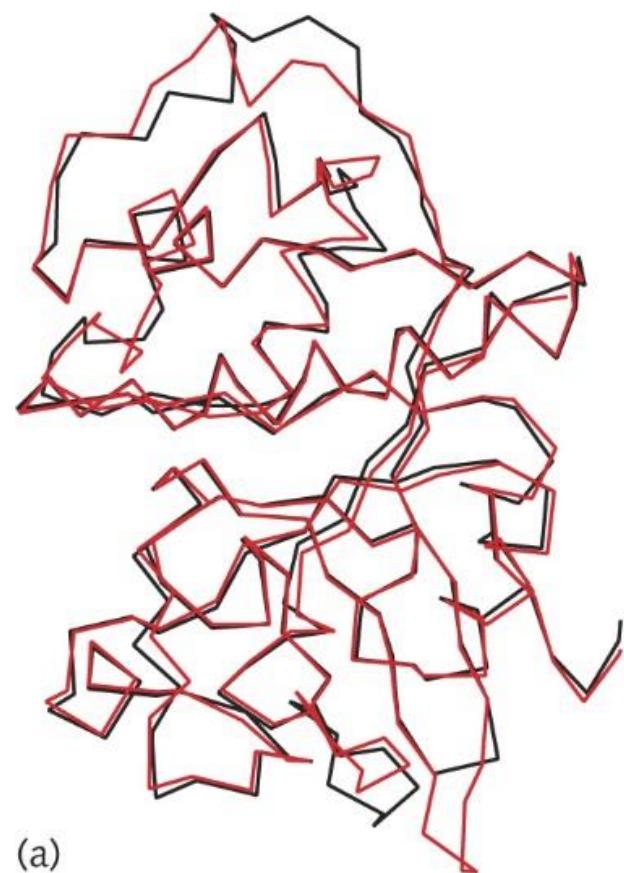
9pap IPEYVDWRQKGAVTPVKNQGSCGSCWAFSAVVTIEGIKIRTGNLNQYSEQEL
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2act LPSYVDWRSAGAVVDIKSQGECGGCWAFSAIATVEGINKITSGSLISLSEQEL

9pap LDCCDRR--SYGCNGGYPWSALQLVAQY-GIHYRNTYPYEGVQRYCRSR-EKGP
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2act IDCGRQTQNTRGCDGGYITDGFQFIINDGGINTEENYPYTAQDGDCD--VAL--

9pap ---YAAKTDGVRQVQPYNQGALLYSIANQPVSVVLQAAGKDFQLYRGGIFVGP
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2act QDQKYVTIDTYENVPYNNEWALQTAVTYQPVSVALDAAGDAFKQYASGIFTGP

9pap CGNKVDHAVA AVGYGP-----NYILIKNSWGTGWGNGYIRIKRGTGNSYG
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2act CGTAVDHAIVIVGYG-TEGGVDYWIVKNSWDTTWGEEGYMRILRNV-GGAG

```



Poravnavo protinskih zaporedij

Z višanjem evolucijske oddaljenosti med dvema homologoma prihaja do vedno večje divergence - podobnost na nivoju primarne in terciarne strukture pada.

Primer: papain (iz papaje, *Carica papaya*) in človeški katepsin L

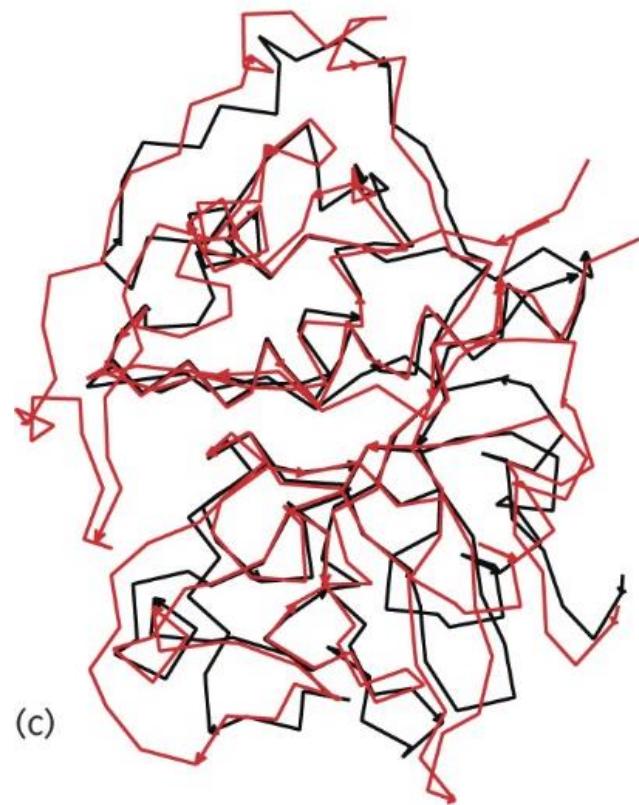
9pap	IPEYDWRQKGAVTPVKNQGSCGSCWAFSAVVTIEGIKIRTGNLNQYSEQEL
1cj1	-----VDWREKGYVTPVKNQGQCGSSWAFSATGALEGQMFRKTGRLISLSEQNL
9pap	LDCDR--RSYGCNGGYPWSALQLVAQY-GIHYRNTYPYEGVQRYCRSREKGP-
1cj1	VDCSGPEGNEGCGNGLMDYAFQYVQDNGGLDSEESYPYEATEEESCKY--N-PK
9pap	-YAAK--TDGVRQVQPYNQGALLYSIA-NQPVSVVLQAAGKDFQLYRGGIFVG
1cj1	YS-VANDA-GFVDIPK-QEKALMKAVATVGPISVAIDAGHESFLFYKEGIYFE
9pap	P-CGN--KVDHAVAAVGYGPNYILIKNSWGTGWGENGYIRIKRG TGNSYGVCG
1cj1	PDC-SSEDMDHGVLVVGYG-----
9pap	LYTSSFYPVKN---
1cj1	-----FES



Poravnavo proteinskih zaporedij

Z višanjem evolucijske oddaljenosti med dvema homologoma prihaja do vedno večje divergencije - podobnost na nivoju primarne in terciarne strukture pada.

Primer: papain (iz papaje, *Carica papaya*) in človeški katepsin B



Poravnavo protienskih zaporedij

Z višanjem evolucijske oddaljenosti med dvema homologoma prihaja do vedno večje divergencije - podobnost na nivoju primarne in terciarne strukture pada.

Primer: papain (iz papaje, *Carica papaya*) in stafopain (*S. aureus*)

9pap	IPE----YVDWRQKGAVTPVKNQGSCGSCWAFSAVVTIEGIKIRTGNLNQ-YS
1cv8	---NEQYVNKL--E-NFKIRETQGNNGWCAGYTMSALLNATYN----T-NKYH
9pap	EQELLDLCDRRSYG-----CNGGYPW-S---ALQLVAQYGIHYRNTYPYEG
1cv8	AEAVMRF LH----PNLQGQQFQFTGLT-PREMIYFGQ-T-----
9pap	VQR YCRSREKGPYAAKTDGVRQVQP-Y-NQ--GALL-YSIA-NQ-PVSVVLQAA
1cv8	-----QG-RSPQL-LNRMTTYNE--VDNL-TKNNKGIAILG---
9pap	GKDFQLYRGGIFVGPGCGNKV-----DHAVA AVGYGP----NYILIKN
1cv8	-----SRVESRNGMHAGHAMAVVGNAKLNNGQEVI IIWN
9pap	SWG TGW GENG YIRIKRGTGNSYG-VCGLY-----TSSFYPVKN
1cv8	PWD N----G-FMTQDA-K----NN----VIPVSNGDHYQWYSSIYGY---



Podobnost zaporedij je krepko v **coni somraka** (10-30% identičnosti). Brez strukturnih in funkcijskih podatkov bi težko sklepali na homologijo.

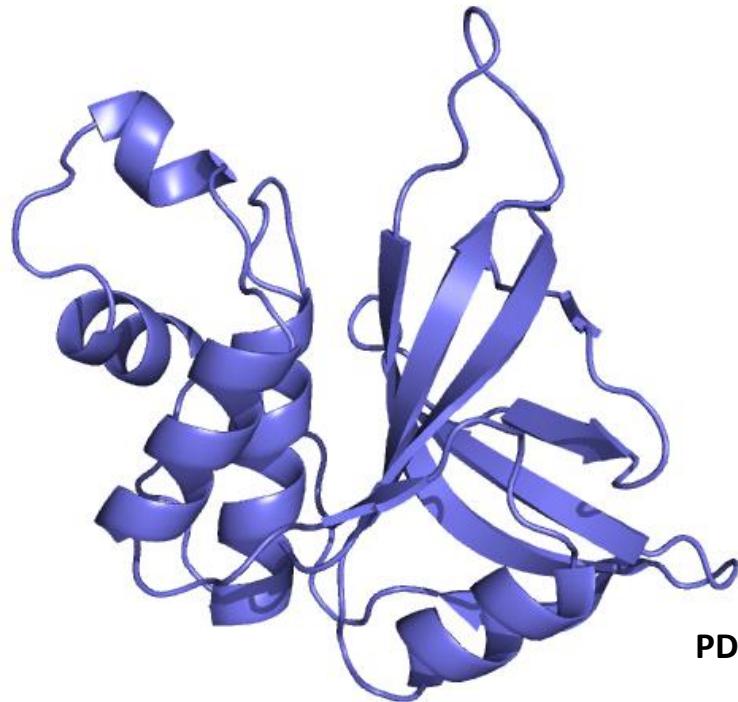
Poravnavo protinskih zaporedij

Z višanjem evolucijske oddaljenosti med dvema homologoma prihaja do vedno večje divergencije - podobnost na nivoju primarne in terciarne strukture pada.

Primer: papain (iz papaje, *Carica papaya*) in stafopain (*S. aureus*)



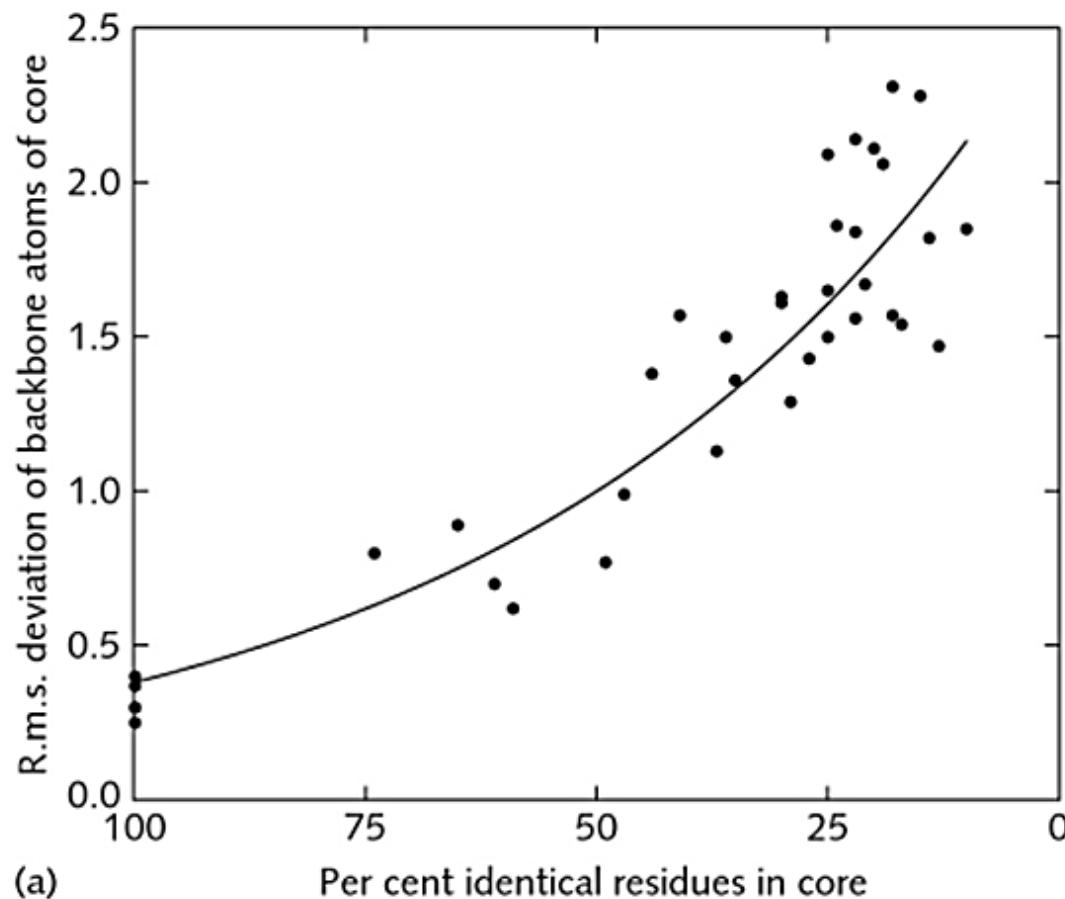
PDB 9PAP



PDB 1CV8

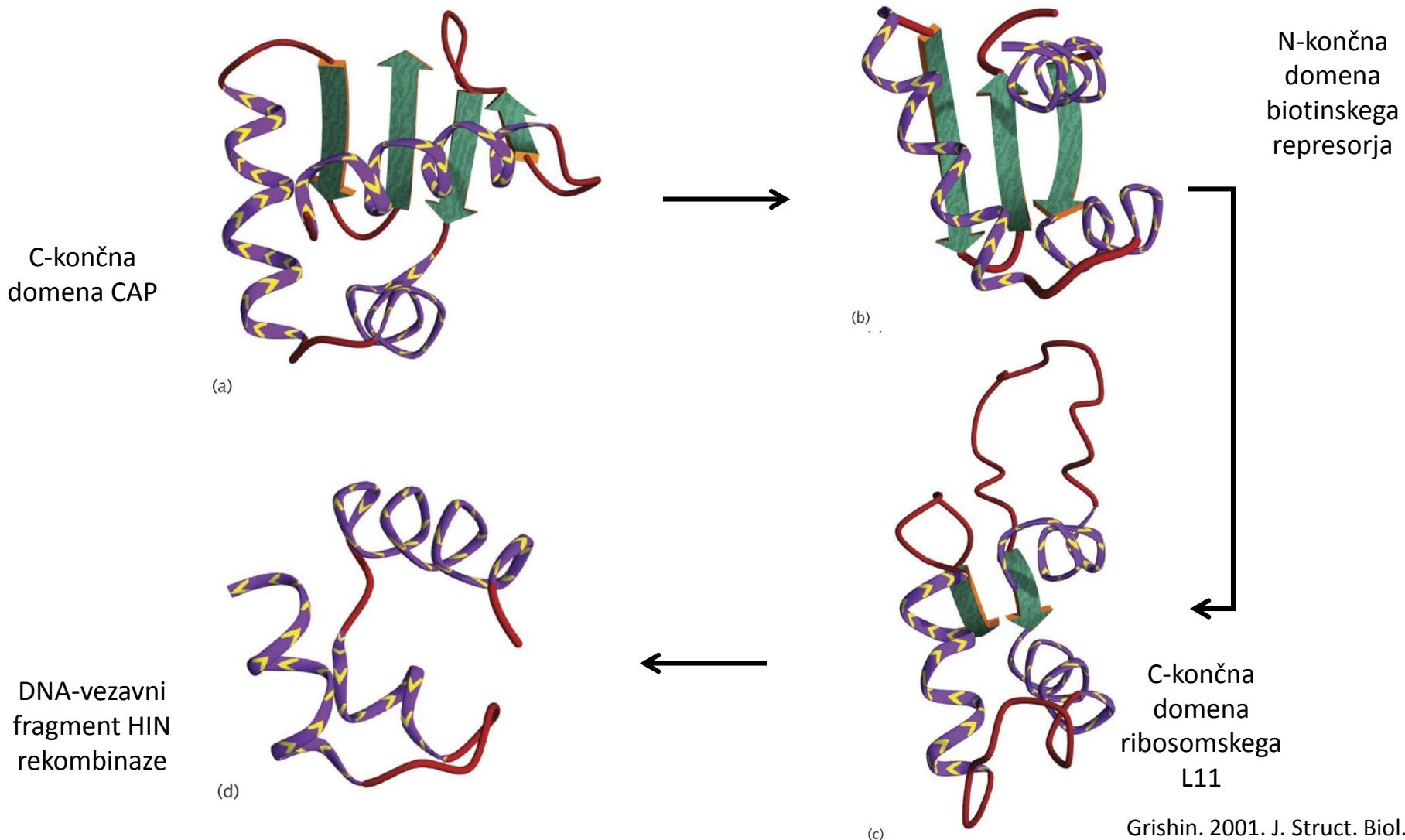
Divergenca struktur proteinov

Za globularne proteine lahko v splošnem rečemo, da so sestavljeni iz bolje ohranjenega jedra in variabilnih perifernih elementov (zank). Struktura jedra se na nivoju primarne in terciarne strukture manj spreminja kot zanke, zato jo lažje sledimo čez večje evolucijske razdalje.



Divergenca struktur proteinov

Obstajajo tudi posamezni proteini, kjer ne glede na odsotnost podobnosti na nivoju primarne in terciarne strukture ne moremo z gotovostjo trditi, da niso homologi.



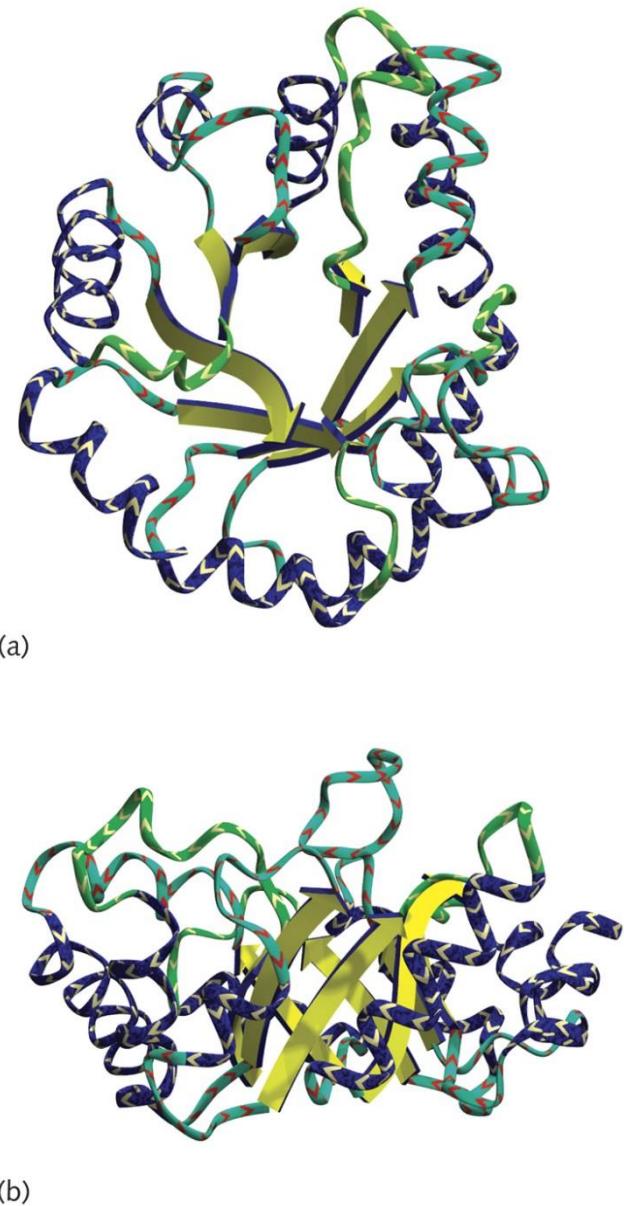
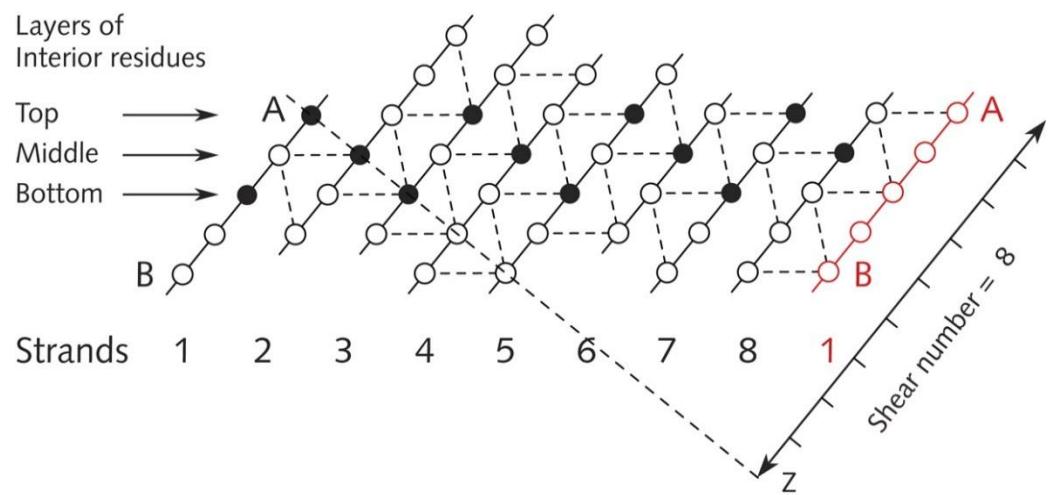
Omejitve evolucije proteinov

Faktorji, ki lahko omejujejo evolucijo proteinov:

- stabilnost – pakiranje proteina; zamenjave posameznih ostankov se dobro tolerirajo/kompenzirajo
- elementi sekundarnih struktur se krajšajo/podaljšujejo na koncih strukture
- vzorci kontaktov med elementi sekundarnih struktur (vijačnica/ β -ploskev) se ohranjajo
- ohranitev funkcije (vloga duplikacij)
- nekatera zvitja so omejena s samo geometrijo zvitja (β -sodčki)

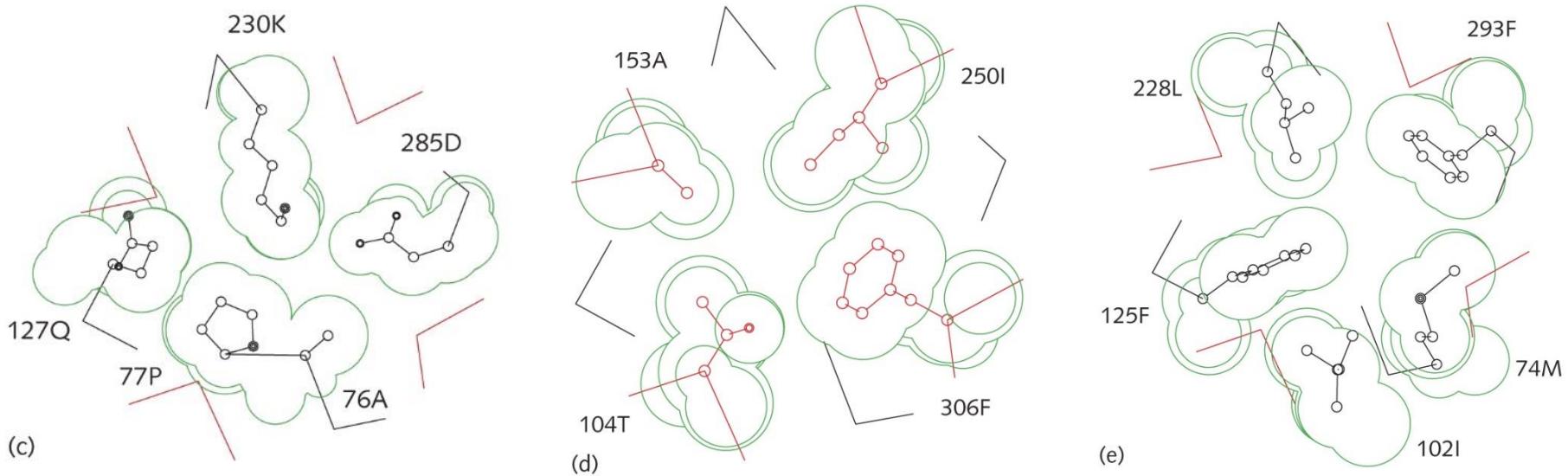
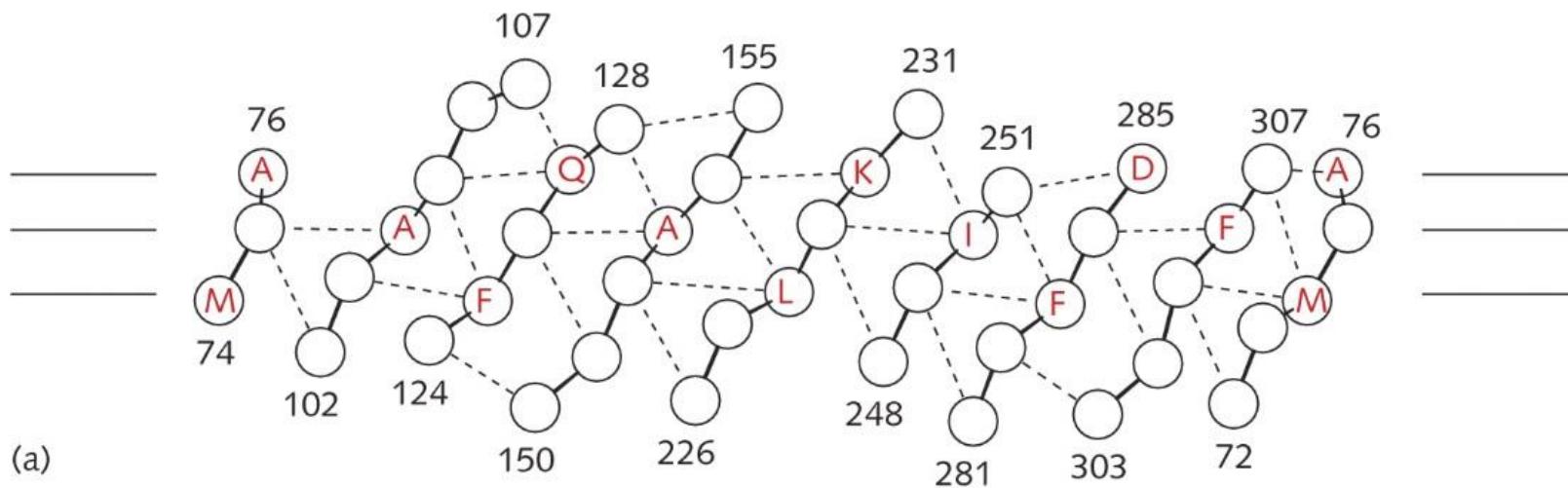
TIM sodčki

Sodčki vsebujejo ohranjen vzorec vodikovih vezi in ostankov obrnjenih v notranjost sodčka.



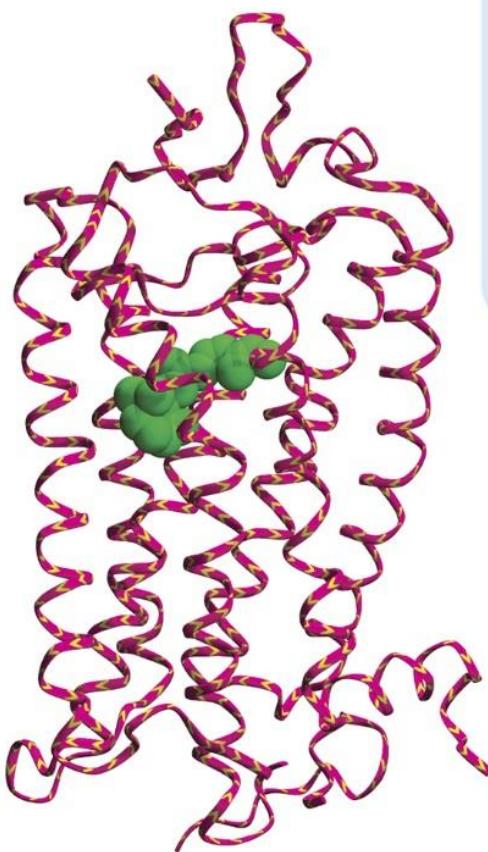
TIM sodčki

Primer: glikoksilat oksidaza



Vidni pigmenti

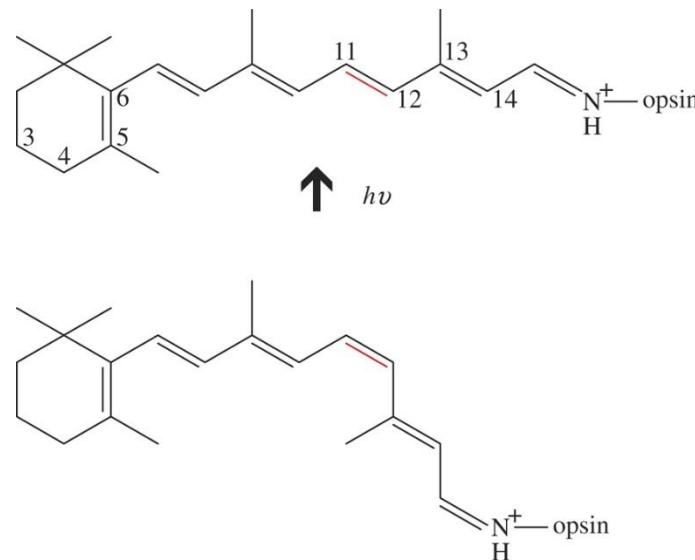
Vidni pigmenti so se razvili na način, da omogočajo zaznavanje svetlobe v različnih pogojih, glede na življenjsko okolje organizmov. Zgrajeni so iz proteina opsina in vezanega *cis*-11-retinala.



BOX
7.6

Events in the molecular mechanism of vision

- Absorption of light by retinal.
- Isomerization of retinal from *11-cis* to *all-trans* (Fig. 7.40). The *all-trans* isomer has a different shape, which doesn't fit the binding site well. The protein changes conformation and releases the retinal.
- The new conformation of the rhodopsin activates the G protein transducin.
- Transducin activates a phosphodiesterase .
- The phosphodiesterase hydrolyses cyclic GMP.
- Cyclic-GMP gated cation channels close.
- Buildup of potential difference across cell membrane.
- Potential difference transmitted to adjacent neuron (in vertebrates, via a bipolar cell), triggering a nerve impulse.
- Feedback mechanisms affecting each of these steps restore the system to its resting state.



Vidni pigmenti

Vidni pigmenti so se razvili na način, da omogočajo zaznavanje svetlobe v različnih pogojih, glede na življenjsko okolje organizmov. Zgrajeni so iz proteina opsina in vezanega *cis*-11-retinala.

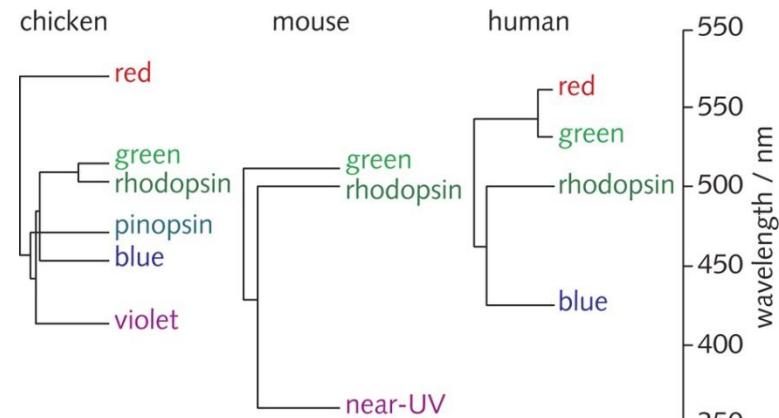
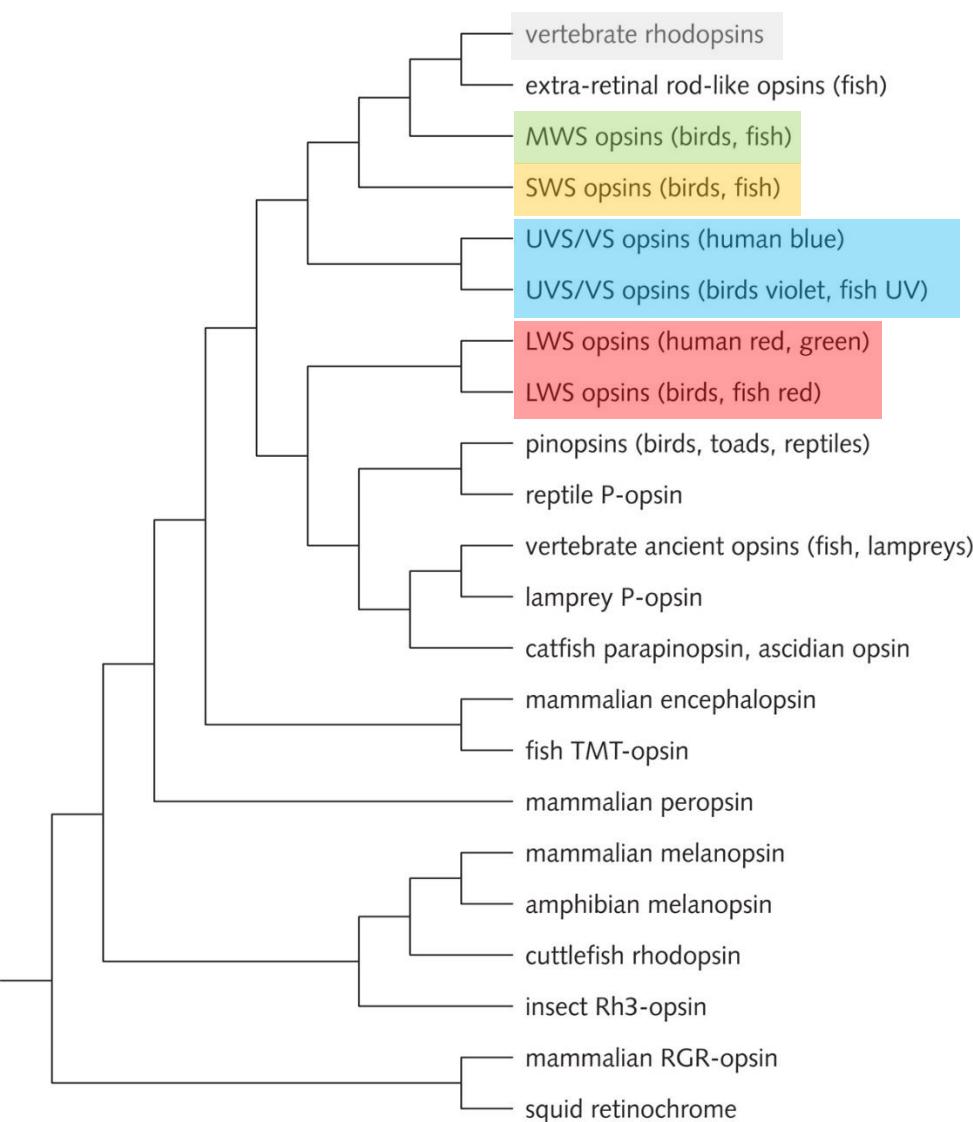
Geni za opsine pri človeku:

Chromosomal locations and numbers of introns of the nine human opsin genes

	Opsin	Chromosomal location	Number of introns
paličnice	Rhodopsin	3q22.1	4
čepnice	Blue opsin	7q32.1	4
	Red opsin	Xq28	5
	Green opsin	Xq28	5
	Encephalopsin	1q43	3
	Melanopsin	10q23.2	9
	Peropsin	4q25	6
	RGR	10q23.1	6
	Neuropsin	6p12.3	6

Vidni pigmenti

Filogenetska razmerja med opsin:

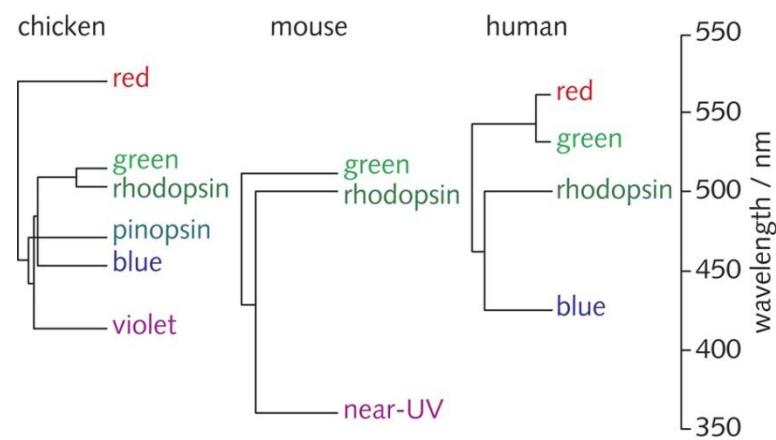
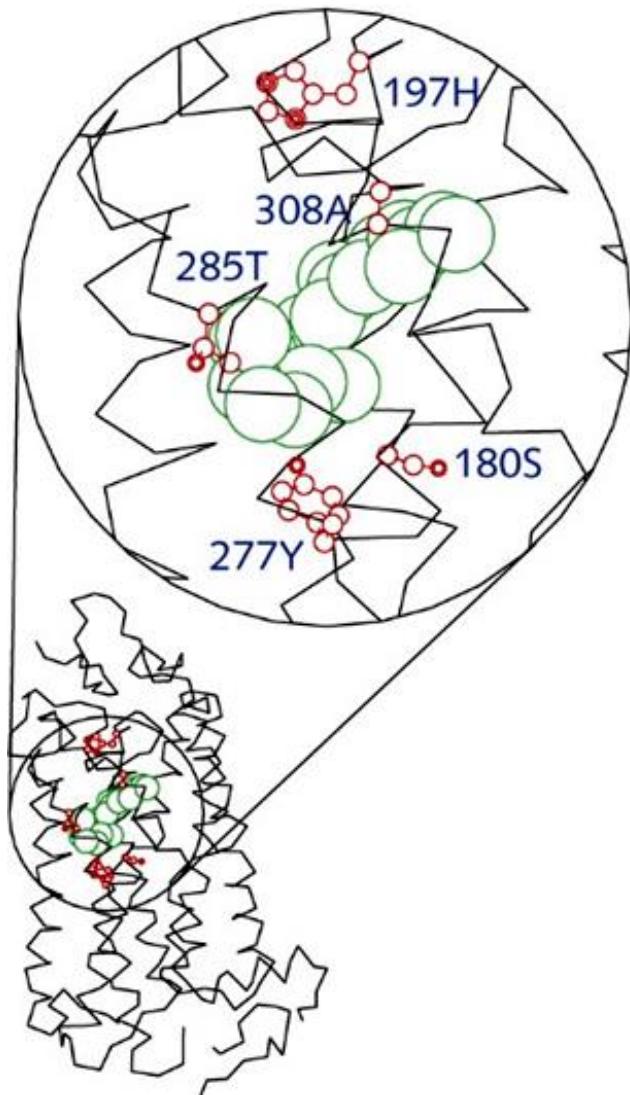


Pri človeku se gena za rdeči in zelen opsin nahajata v tandemu na X-kromosomu. Identičnost je 96%.

Identičnost z modrim opsinom (na kromosomu 7) je 40%.

Vidni pigmenti

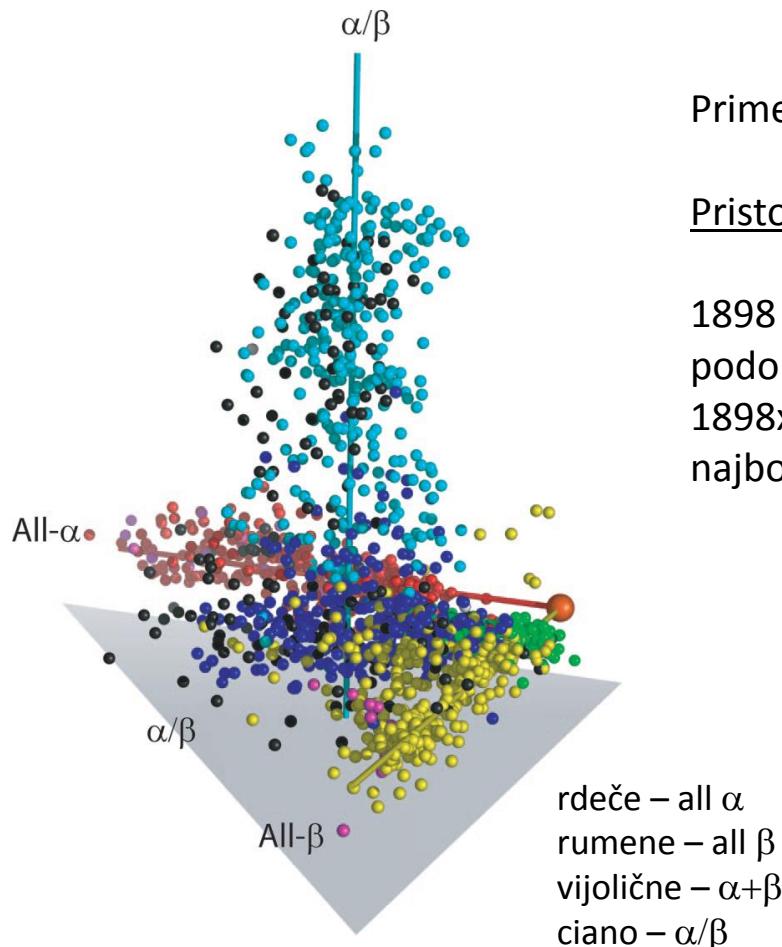
Kaj določa valovno dolžino absorbirane svetlobe?



	pozicija					λ_{\max}
	180	197	277	285	308	
človek	A	H	Y	T	A	552
človek	A	H	F	A	A	530
am. opica	S	H	Y	T	A	561
am. opica	A	H	Y	A	A	539
veverica	S	Y	Y	T	A	532
delfin	A	H	Y	T	S	524
konj	A	H	F	T	A	545
morski praš.	S	Y	Y	A	A	516

Prostor proteinov

Koncept **prostora proteinov**¹ je širok pojem, ki omogoča predstavitev lastnosti proteinov v večdimenzijsnem prostoru. V okviru evolucije lahko govorimo o prostoru zaporedij proteinov ali prostoru struktur proteinov. S primerjavo vseh znanih proteinskih struktur lahko zgradimo **vesolje proteinov**.



Primer seta 1898 različnih proteinskih struktur

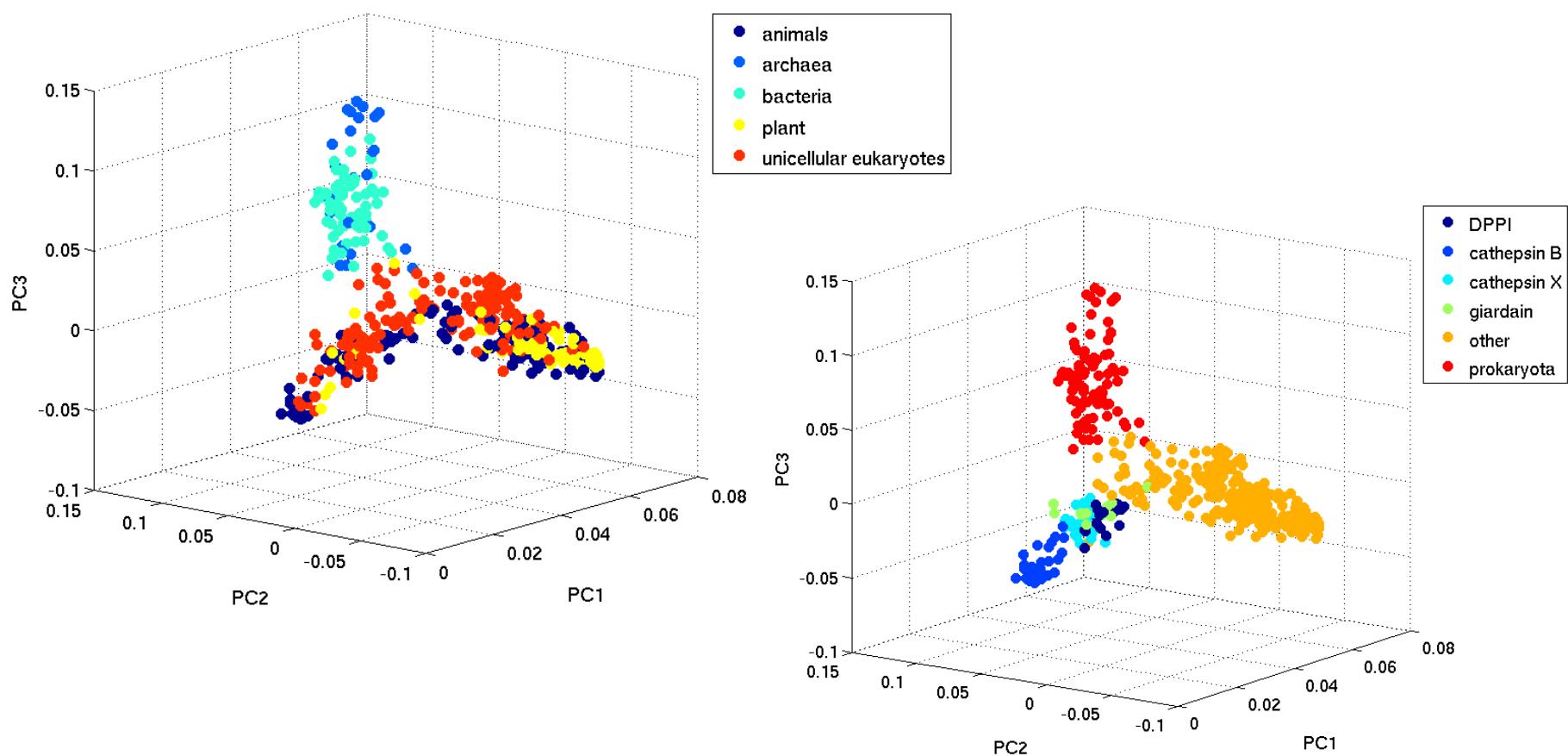
Pristop:

1898 struktur z $< 25\%$ identičnostjo iz PDB → izračunamo podobnost med strukturami z DaliLite → matrika 1898×1898 → analiza glavnih komponent → mapiramo tri najbolj raznolike komponente na 3D graf.

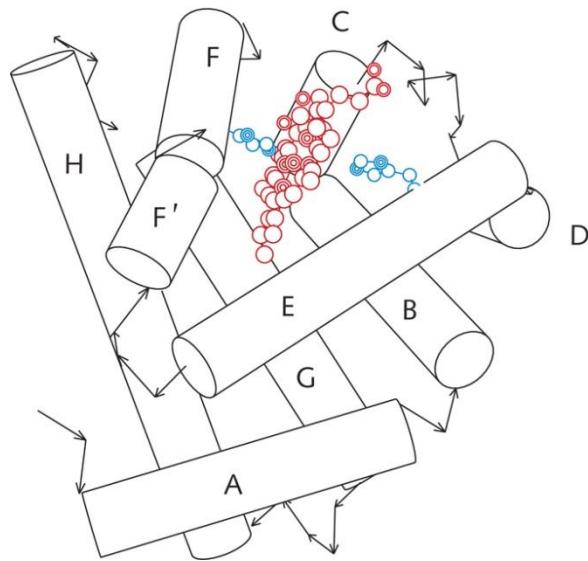
Prostor proteinov

Koncept prostora proteinov lahko uporabimo tudi za analizo zaporedij znotraj družine proteinov – analiza poravnave zaporedij z metodo poglavitnih osi. Metoda je npr. vključena v program Jalview. Praktična je za analizo velikih setov podatkov, kjer filogenetska drevesa ne dajo pregledne informacije.

Primer: analiza 400 zaporedij papainu podobnih proteaz:



Evolucija globinov



Sesalčji mioglobin in hemoglobin

Human Haemoglobin α-chain
Horse Haemoglobin α-chain
Human Haemoglobin β-chain
Horse Haemoglobin β-chain
Sperm-whale myoglobin

	10	20	30	40	50	60
Human Haemoglobin α-chain	VLSPADKTNVKAAWGKVGA	HAGEYGAEEALERMF	SFPTTKTYFPHF	DLS		HGS
Horse Haemoglobin α-chain	VLSAADKTNVKAAWSKVGG	HAGEYGAEEALERMF	LGFPTTKTYFPHF	DLS		HGS
Human Haemoglobin β-chain	VHLTPEEKSAVTALWGKV	--NVDEVGGGEALGRL	LVVYPWTQRF	ESFGDLS	TPDAVMGN	
Horse Haemoglobin β-chain	VQLSGEEKAAVLALWDKV	--NEEEVGGEALGRL	LVVYPWTQRF	FDSFGDLS	NPGAVMGN	
Sperm-whale myoglobin	VLSEGEWQLVLHVWAKEA	-DVAGHQDILIRLFKSH	PETLEKFDRKHLKTE	AEMKAS		
	L k V a W KV	e G ea L R	P T	F F d LS		g

Human Haemoglobin α-chain
Horse Haemoglobin α-chain
Human Haemoglobin β-chain
Horse Haemoglobin β-chain
Sperm-whale myoglobin

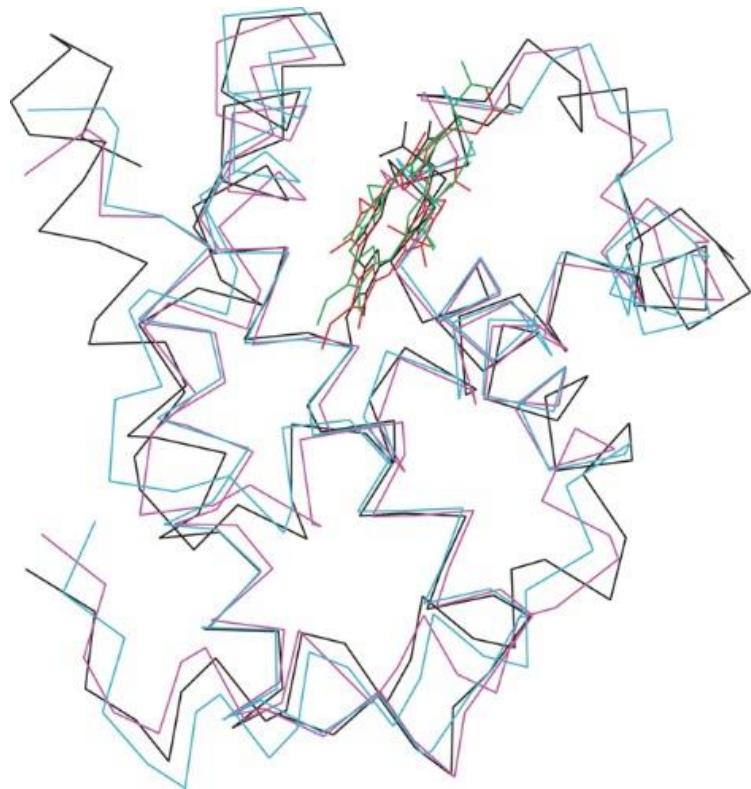
	70	80	90	100	110	120
Human Haemoglobin α-chain	AQVKGHGKKVADALTNAVAHV	-----D-DMPNALSALS	SDLHAKLRVDPVNFKLLSHCLLV			
Horse Haemoglobin α-chain	AQVKAHGKKVKGDALTLAVGHL	-----D-DLPGALSNLSDL	HAKLRVDPVNFKLLSHCLLS			
Human Haemoglobin β-chain	PKVKAHGKKVKGDALTLAVGHL	-----D-NLKGTFTATLSEL	HCDKLHVDPENFRLLGNVLVC			
Horse Haemoglobin β-chain	PKVKAHGKKVKGDALTLAVGHL	-----D-NLKGTFAALSEL	HCDKLHVDPENFRLLGNVLVV			
Sperm-whale myoglobin	EDLKKHGVTVLTALEGAILKK	-----KGHEAELKPLAQSHAT	KHKIPIKYLEFISEAIIH			
	v K HGkkV	h	d	Ls l H	Kh v dp n f	11 1

Human Haemoglobin α-chain
Horse Haemoglobin α-chain
Human Haemoglobin β-chain
Horse Haemoglobin β-chain
Sperm-whale myoglobin

	130	140	150	160			
Human Haemoglobin α-chain	TIAAHLP-A-EFTPAPVHASLDKFLASVSTVLTSKYR						
Horse Haemoglobin α-chain	TIAVHLP-N-DFTPAPVHASLDKFLSSVSTVLTSKYR						
Human Haemoglobin β-chain	VLAHHFG-K-EFTPAPVQAAVQKVVAGVANALAHKYH						
Horse Haemoglobin β-chain	VLAHRFG-K-DFTPAPVQASVQKVVAGVANALAHKYH						
Sperm-whale myoglobin	VLHSRHP-G-DFGADAOQGAMNKALELFRKDIAAKYKELGYQG						
	La h	Ftp	a	K	v	l	KY

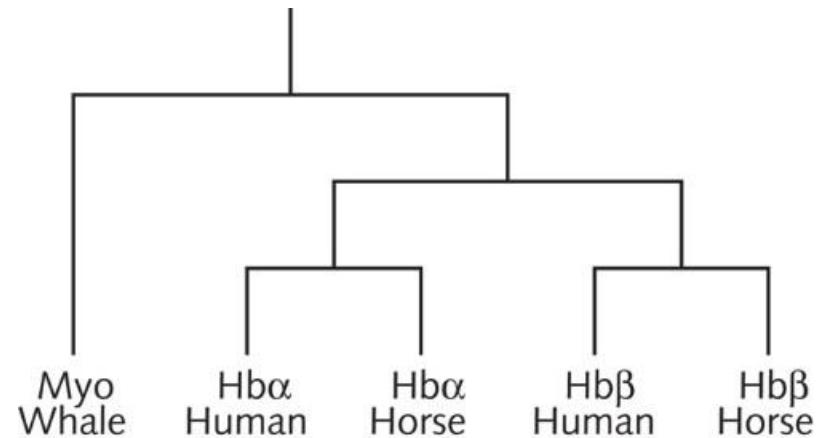
Evolucija globinov

Strukturna poravnava α in β verig
hemoglobina ter mioglobina

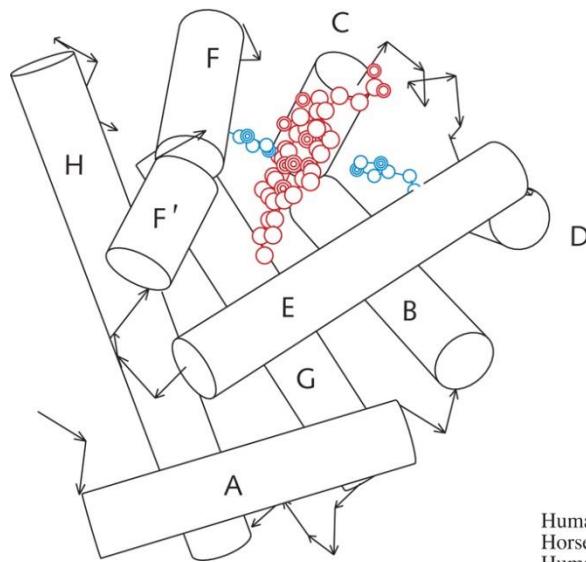


črn – mioglobin
ciano – α -hemoglobin
magenta – β -hemoglobin

Filogenetska razmerja med
posameznimi proteini

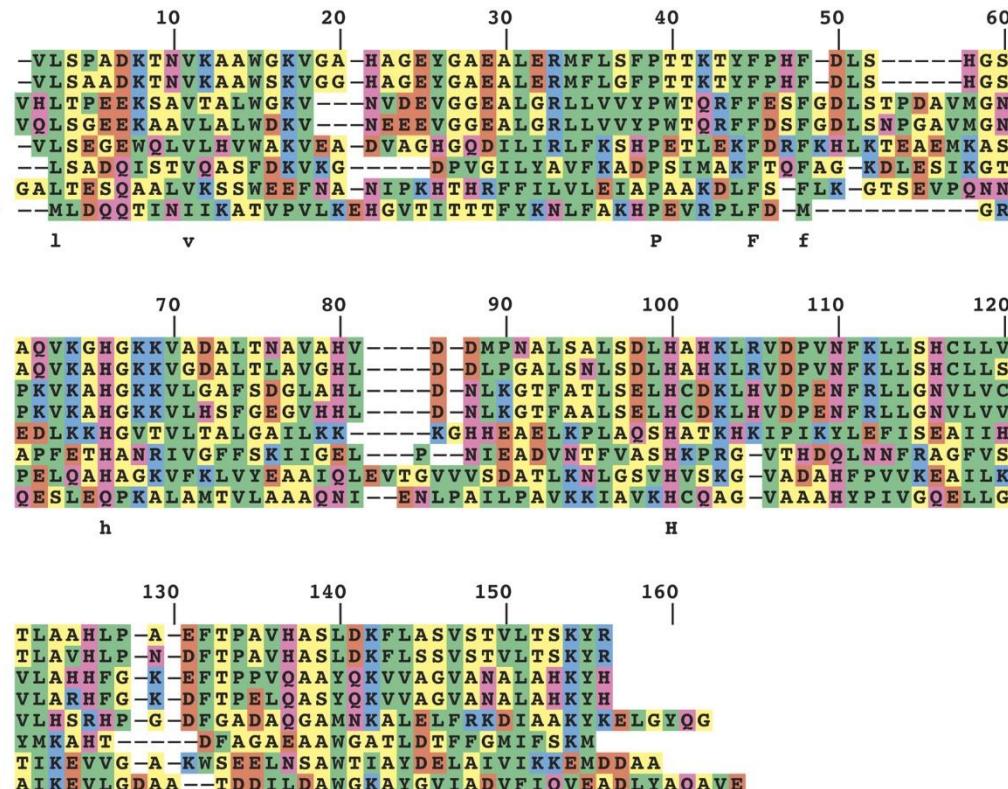


Evolucija globinov



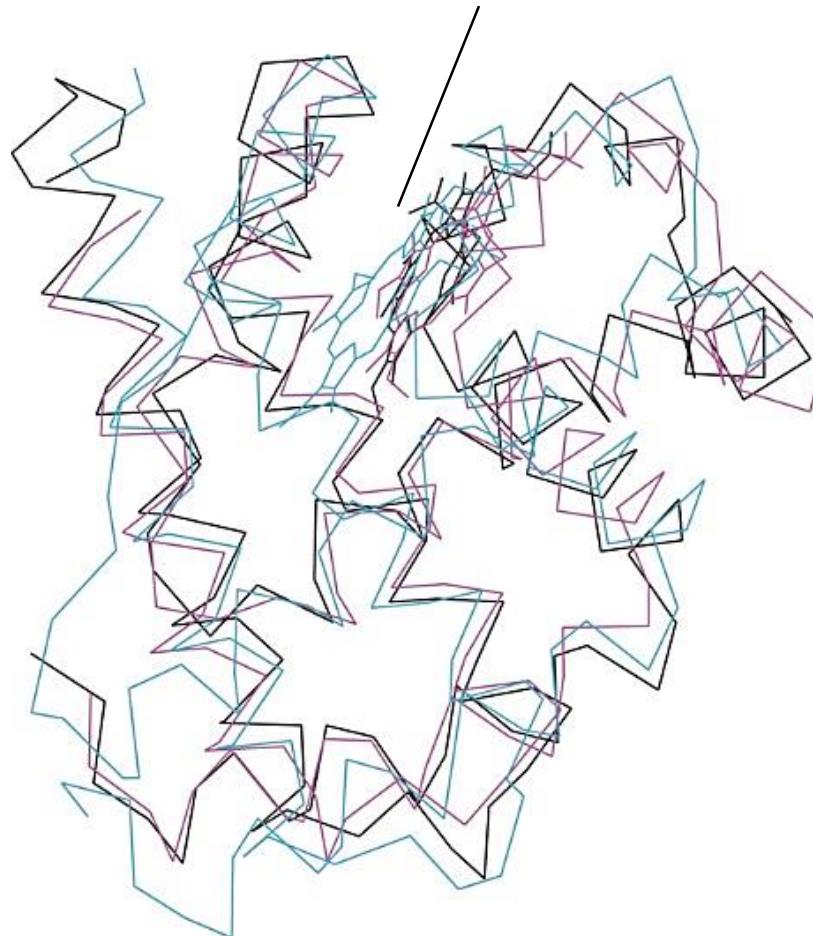
Sesalčni mioglobin in hemoglobin, insektni globin (komar *Chironomus*), rastlinski globin (volčji bob – lupin) in bakterijski globin

Human Haemoglobin α-chain
Horse Haemoglobin α-chain
Human Haemoglobin β-chain
Horse Haemoglobin β-chain
Sperm-whale myoglobin
Chironomus erythrocyruorin
Lupin leghaemoglobin
Bacterial globin (*Vitrosicilla* sp.)



Evolucija globinov

spremenjen položaj hema

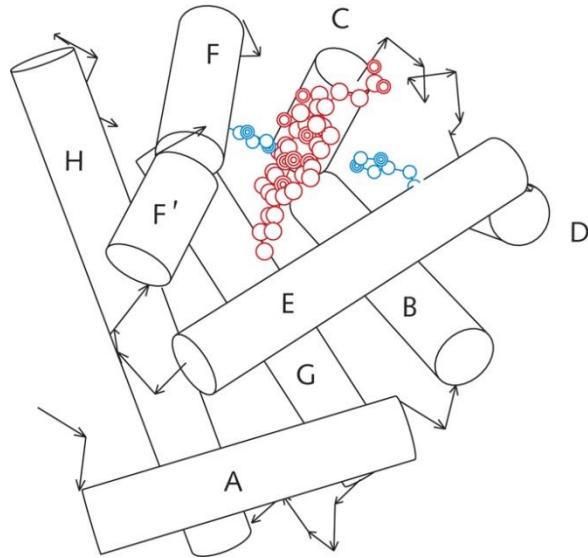


črn – mioglobin

ciano – rastlinski globin

magenta – insektni globin

Evolucija globinov



Sperm whale myoglobin
P. caudatum
C. eugametos
M. tuberculosis

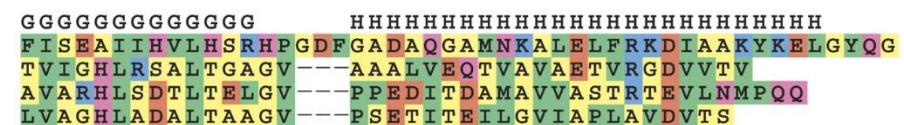
Sperm whale myoglobin
P. caudatum
C. eugametos
M. tuberculosis

Sperm whale myoglobin
P. caudatum
C. eugametos
M. tuberculosis

Skrájšani globini



H



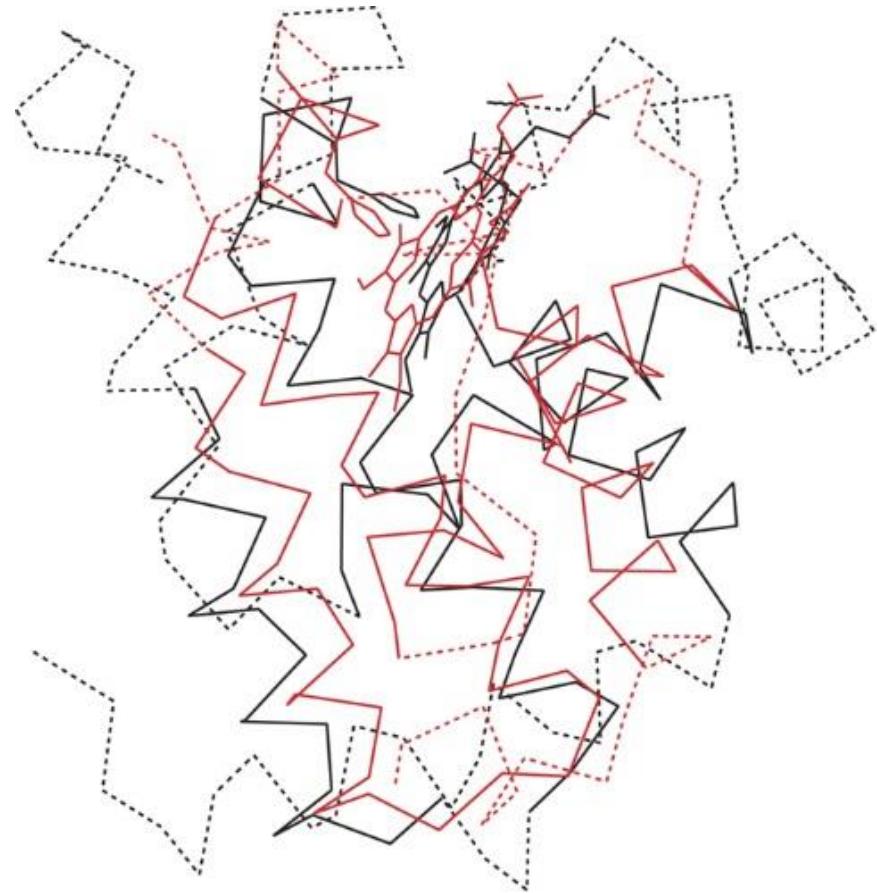
I

črn – mioglobin

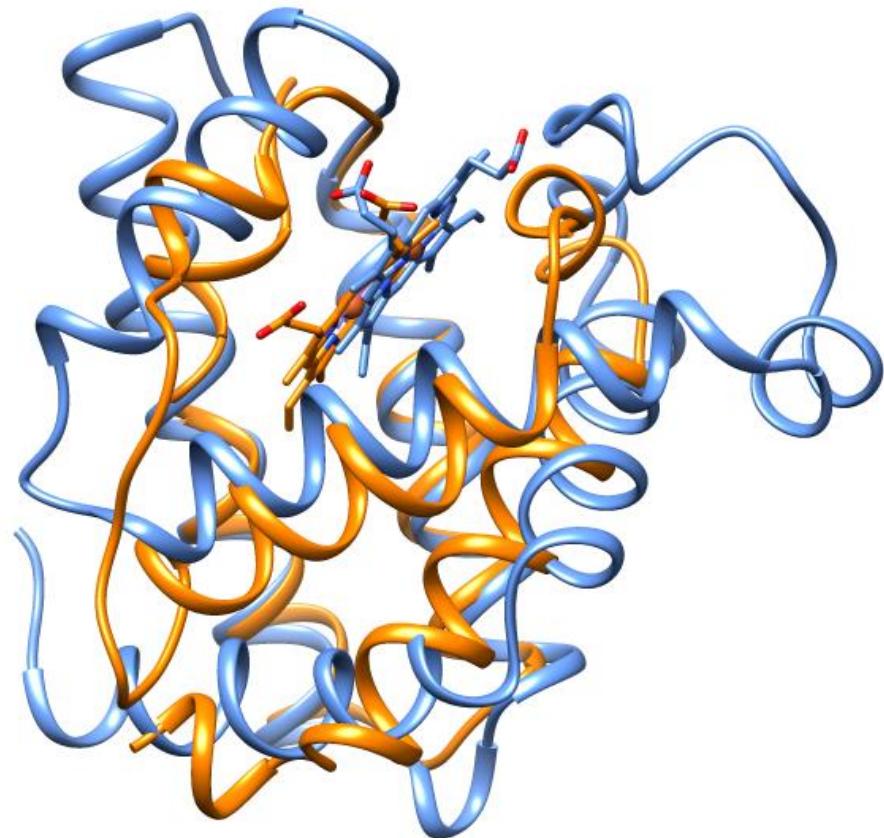
ciano – rastlinski globin

magenta – insektní globin

Evolucija globinov



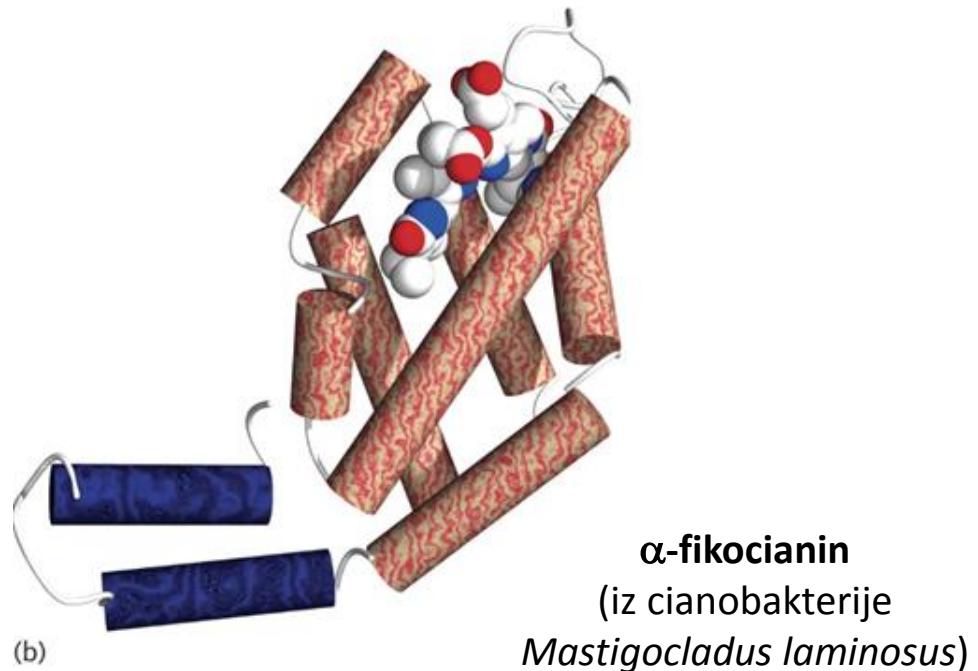
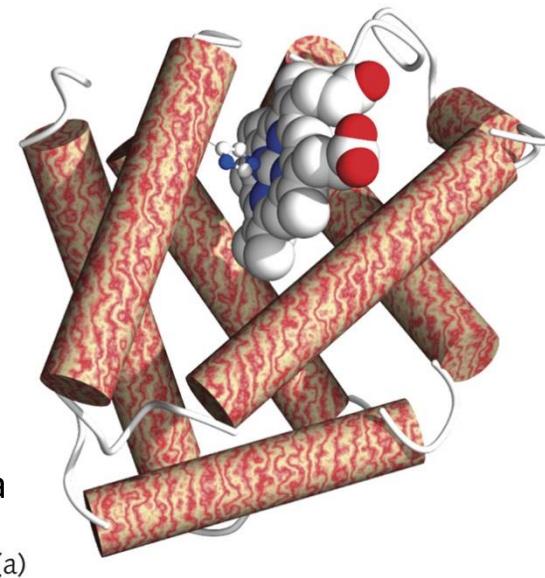
črn – mioglobin
rdeč – t-globin



moder – mioglobin (PDB 1MBO)
oranžen – t-globin (PDB 1DLW)

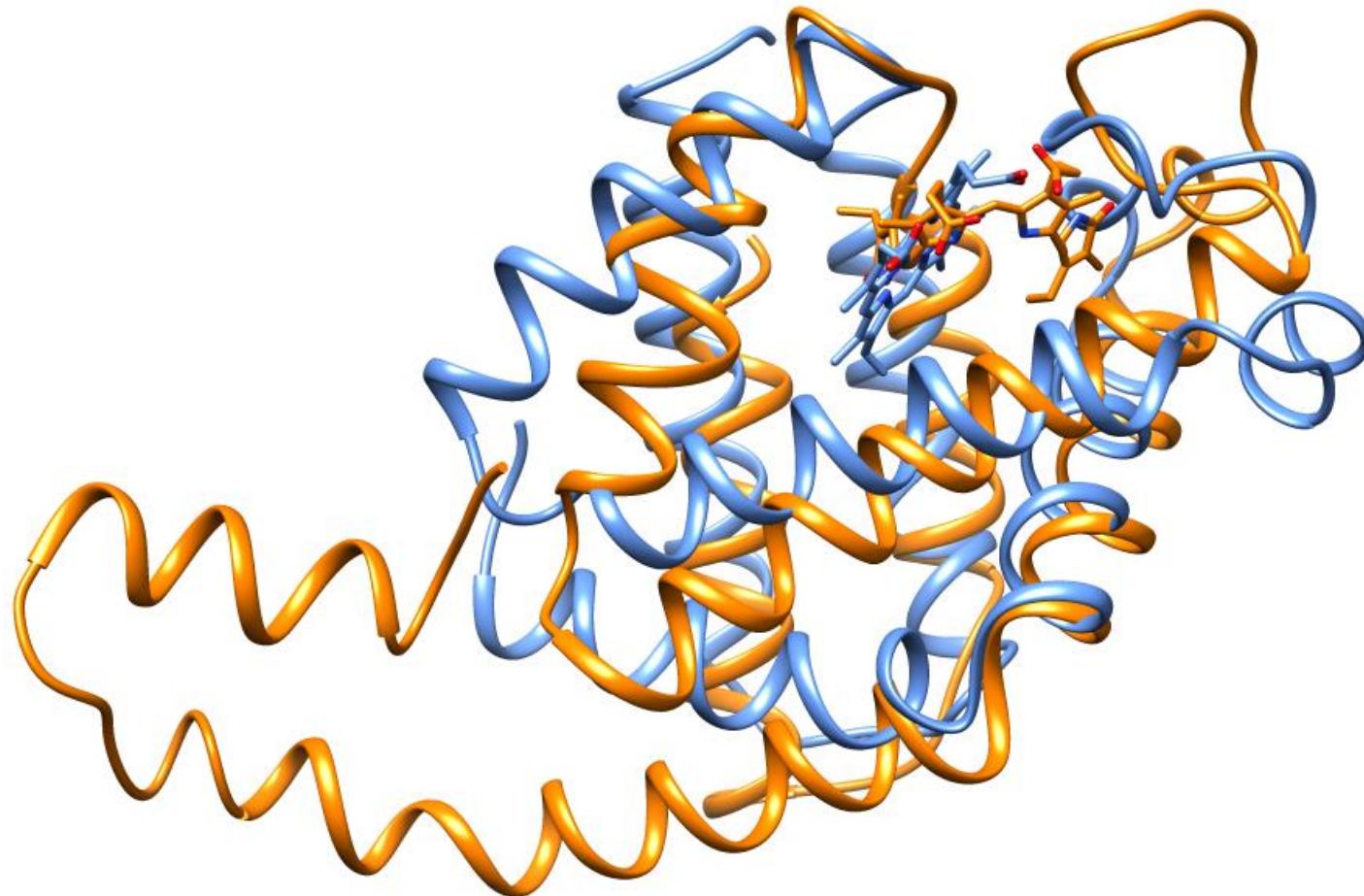
Evolucija globinov

Evolucijsko razmerje med fikocianini in globini? Podobna struktura, različna funkcija.



Evolucija globinov

Evolucijsko razmerje med fikocianini in globini? Podobna struktura, različna funkcija.



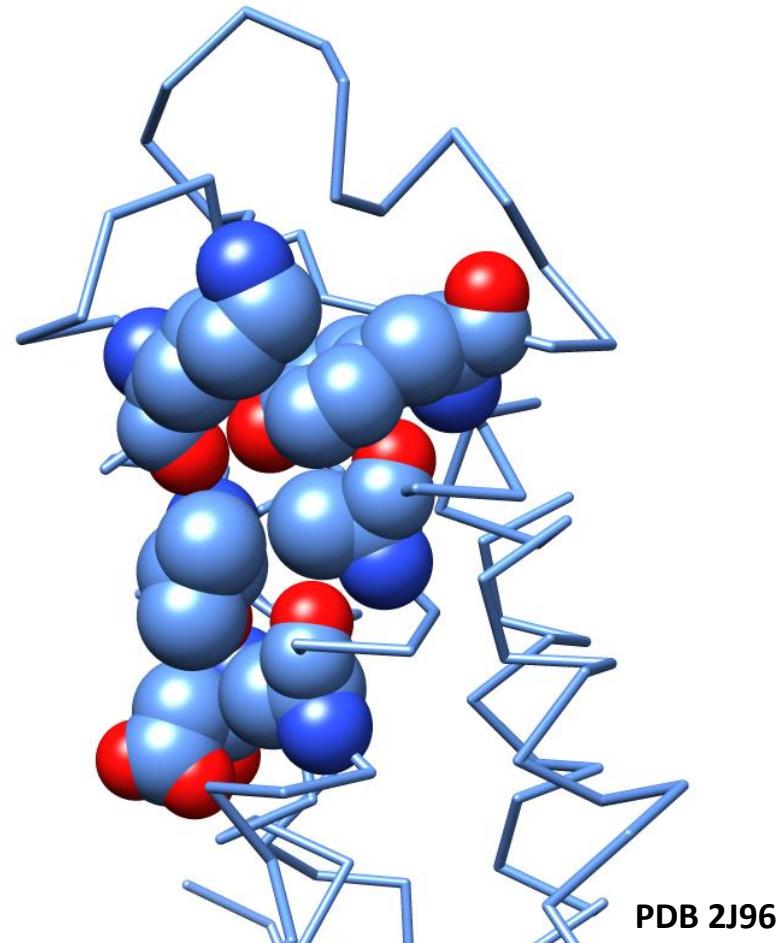
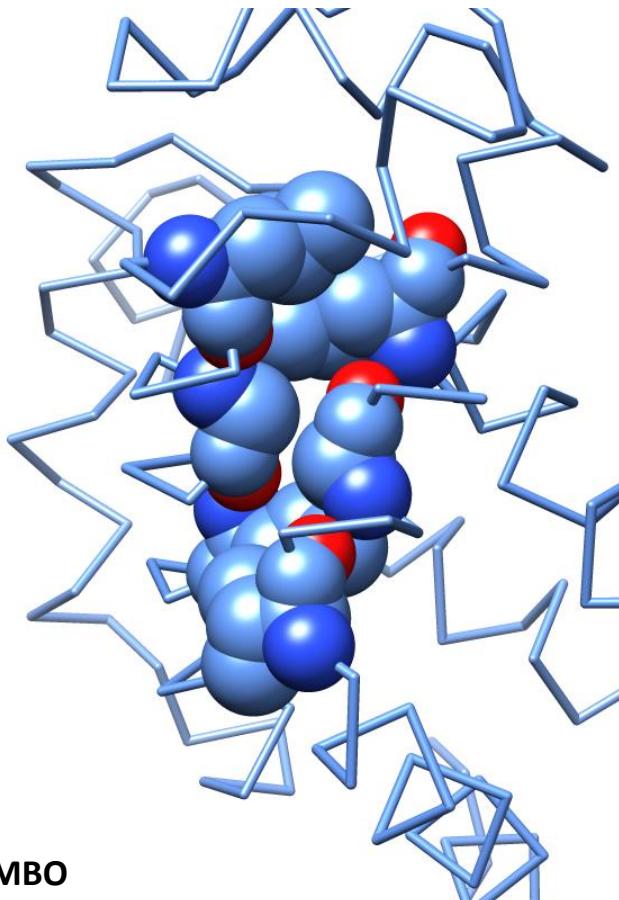
Moder – mioglobin

oranžen – fikocianin

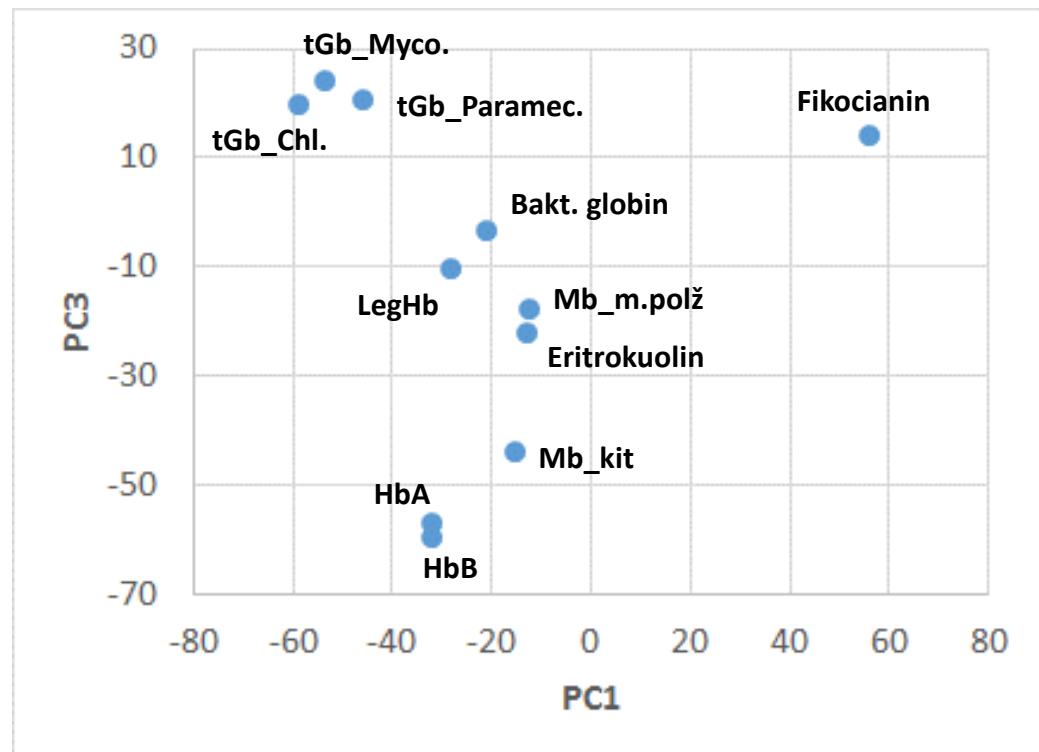
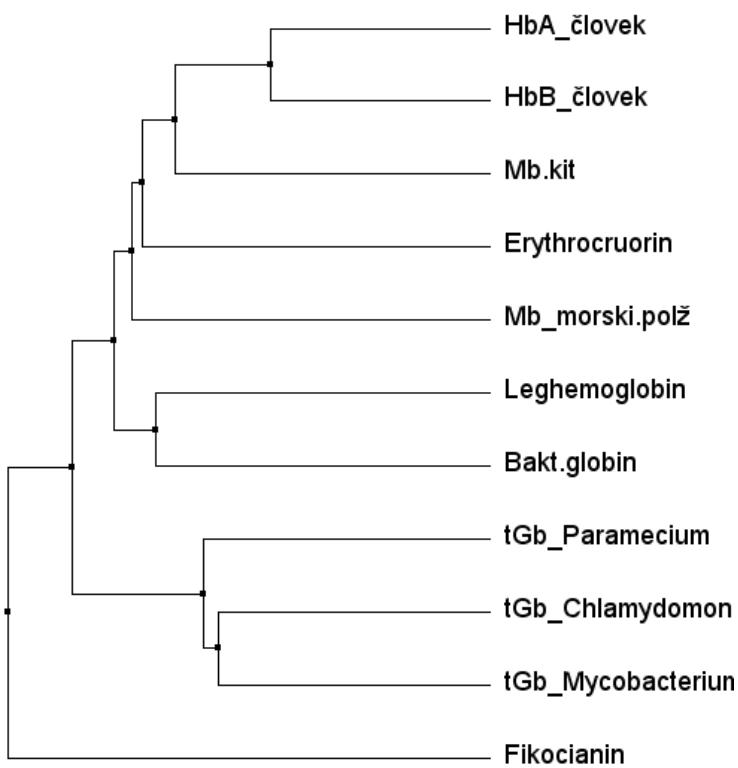
PDB 1MBA, 2J96

Evolucija globinov

Podobna vzorca pakiranja vijačnic B in E v mioglobinu (*levo*) in fikocianinu (*desno*):



Evolucija globinov



Večdomenski proteini

Večina proteinov je sestavljenih iz več domen.

Prokarionti:

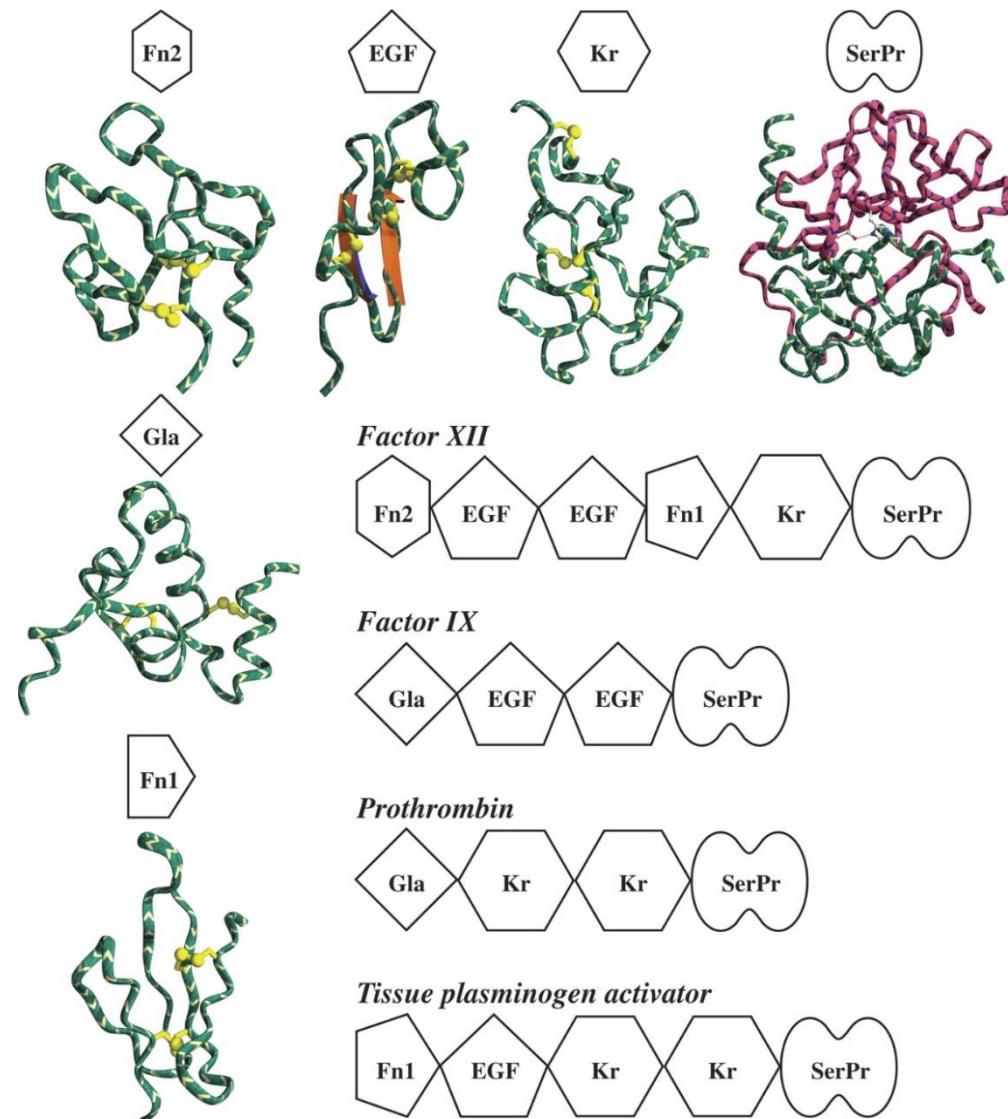
pribl. 2/3 vseh proteinov je večdomenskih sestavljeni iz okoli 400 različnih vrst domen

Evkarionti:

pribl. 3/4 vseh proteinov je večdomenskih sestavljeni iz 600-700 različnih vrst domen

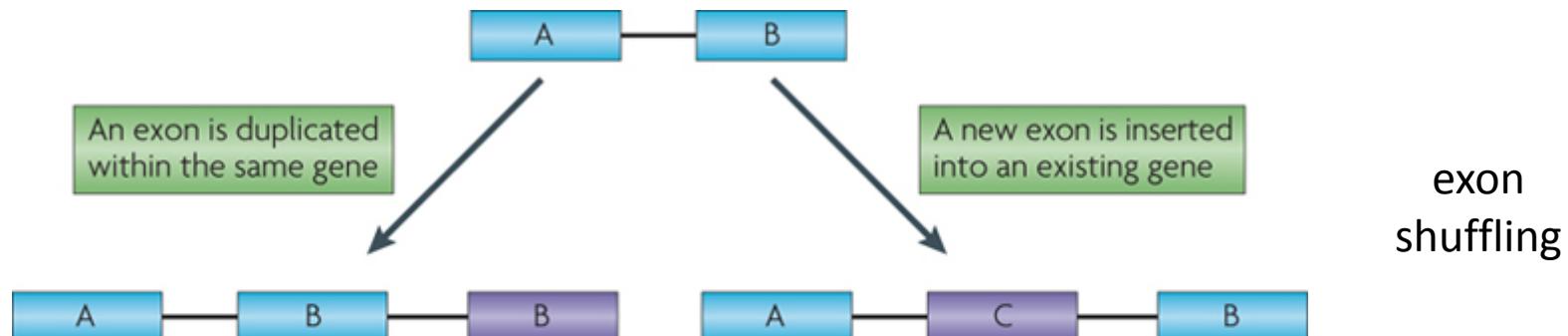
Primer: proteini kaskade strjevanja krvi so nastali s kombiniranjem serinskih proteaz z drugimi moduli.

Moduli so domene, ki jih najdemo kot samostojne enote znotraj funkcionalno in strukturno nesorodnih proteinov.

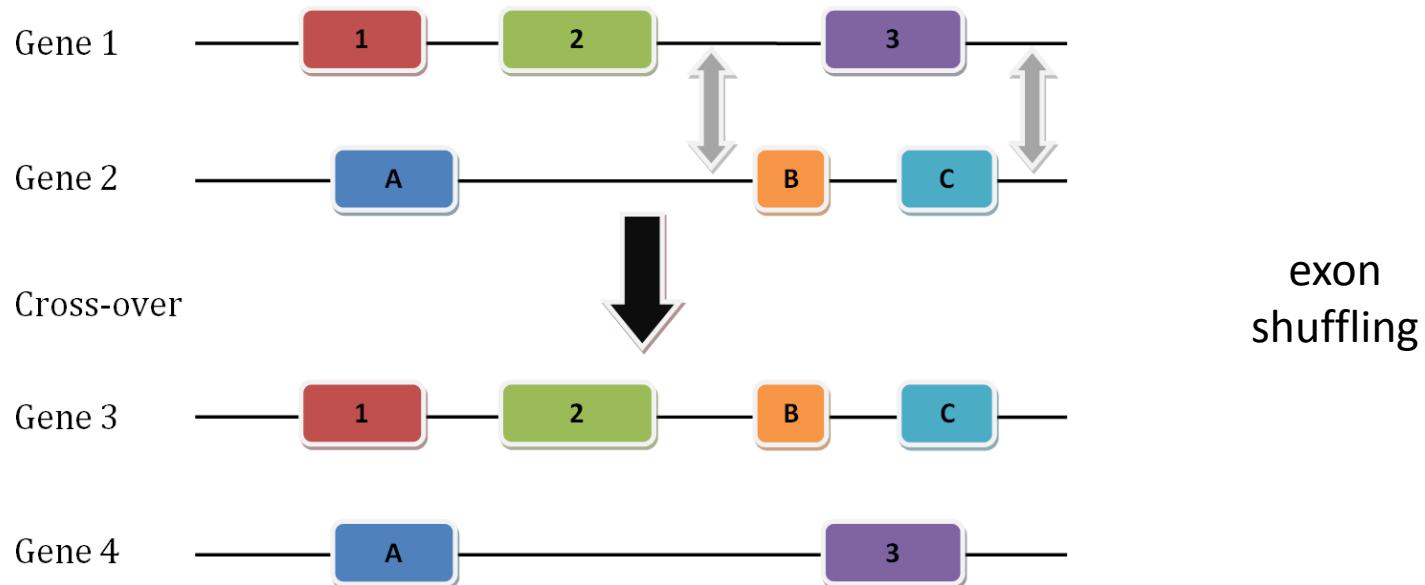


Evolucija večdomenskih proteinov

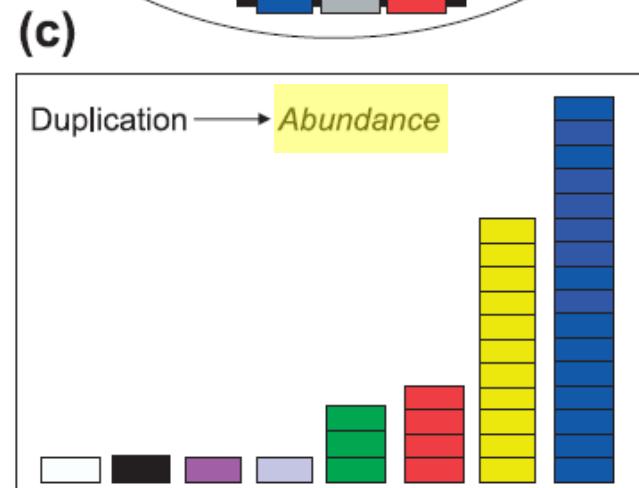
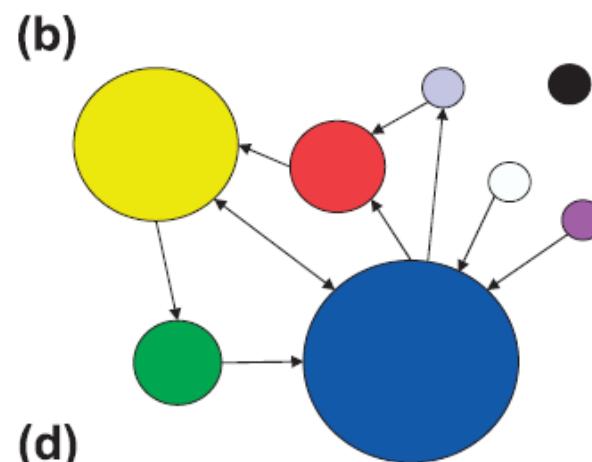
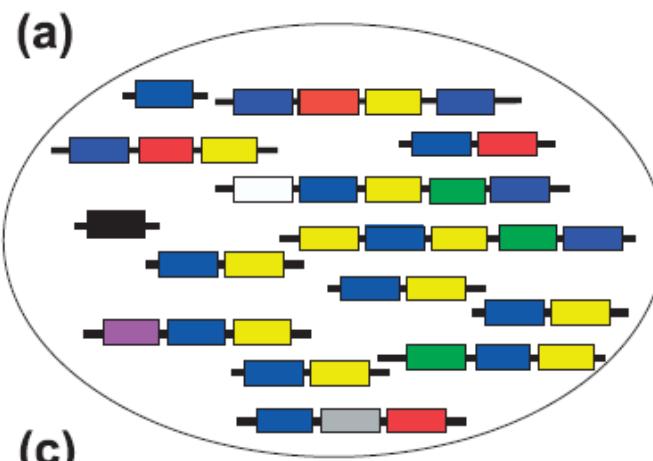
podvojitve
domen



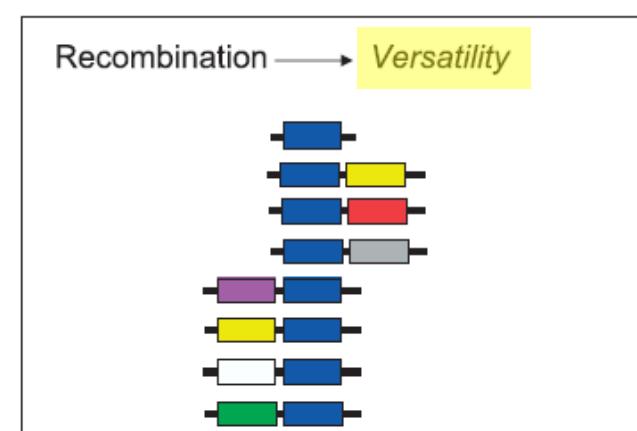
Nature Reviews | Genetics



Večdomenski proteini



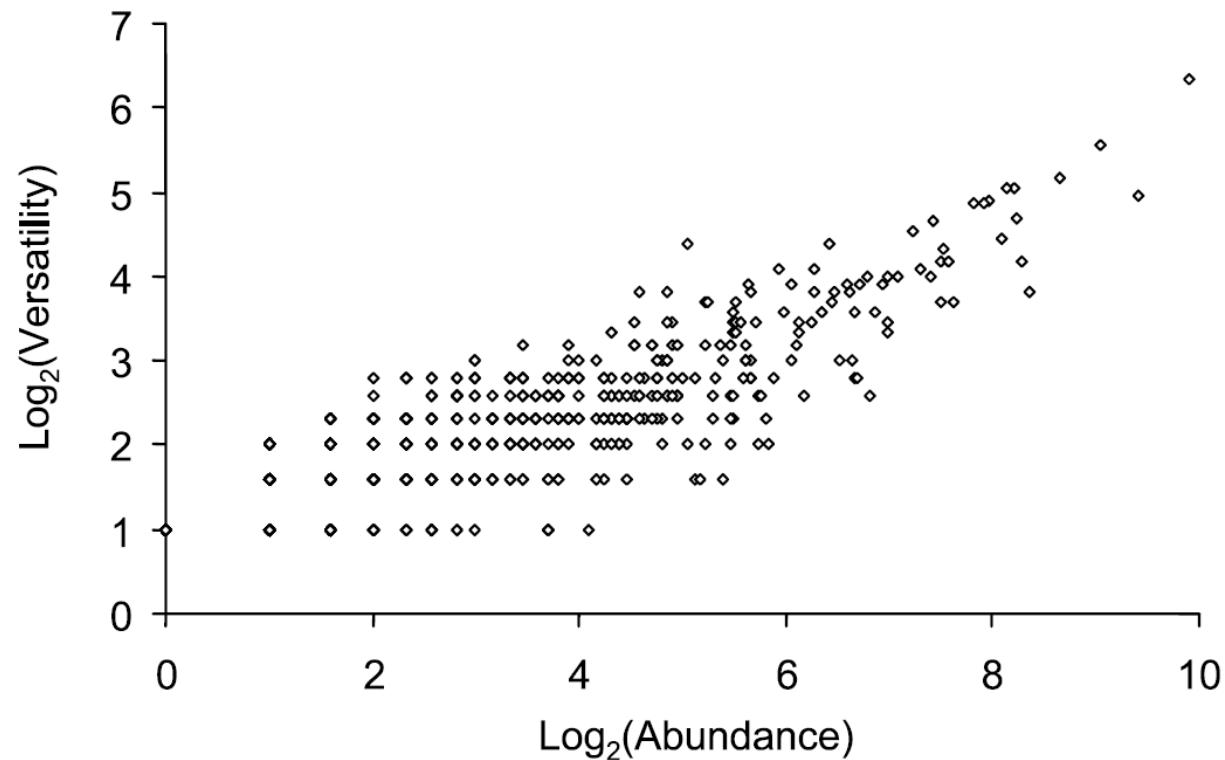
podvojitve vplivajo na
izobilje domen



rekombinacija vpliva na
večstranskost domen

Večdomenski proteini

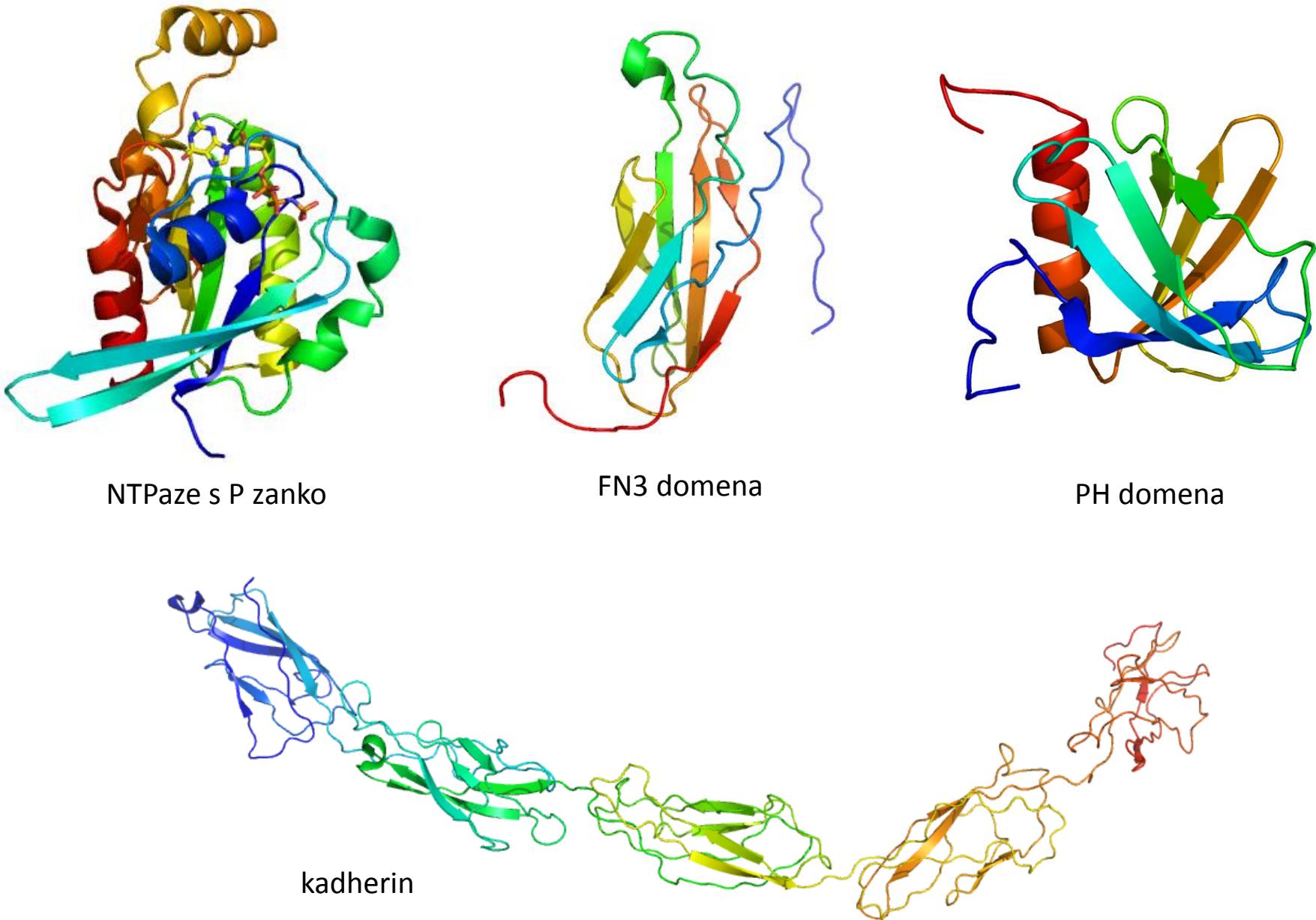
Zveza med izobiljem in večstranskostjo domen:



Večdomenski proteini

Najpogostejše domene	Št. zapisov v človeškem genomu
cinkovi prsti (C2H2 in C2HC)	3693
imunoglobulinske domene	1778
NTPaze s P zanko	1024
z G-proteini sklopljeni receptorji - družina A (receptorji s 7 TM vijačnicami)	824
Fibronektinska domena tipa III.	802
EGF/laminin domene	697
kadherini	686
Proteinske kinaze	539
Domene homologne plekstrinu (PH domene)	491

Večdomenski proteini



Domenski pari

Kombinacije nekaterih domen se ponavljajo v več različnih proteinih v enakem vzorcu. Takim kombinacijam rečemo domenski pari ali **supradomene**.

EF-Tu and eEF1 alpha



eRF3



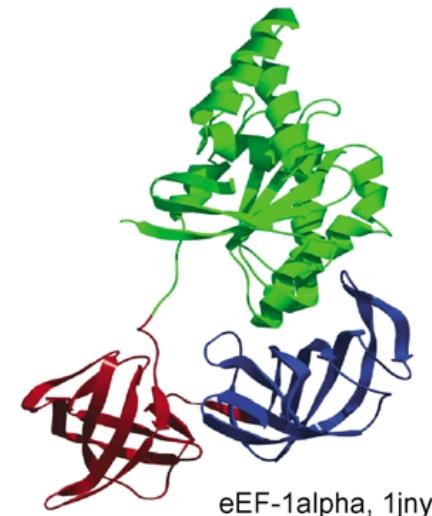
eIF2 gamma



IF2 and eIF5b



EF-G and eEF2



eEF-1alpha, 1jny

— amino acids

■ Translation proteins

■ EF-Tu/eEF-1alpha C-terminal domain

■ Initiation factor IF2/eIF5b, domain 3

■ Ribosomal protein S5 domain 2-like

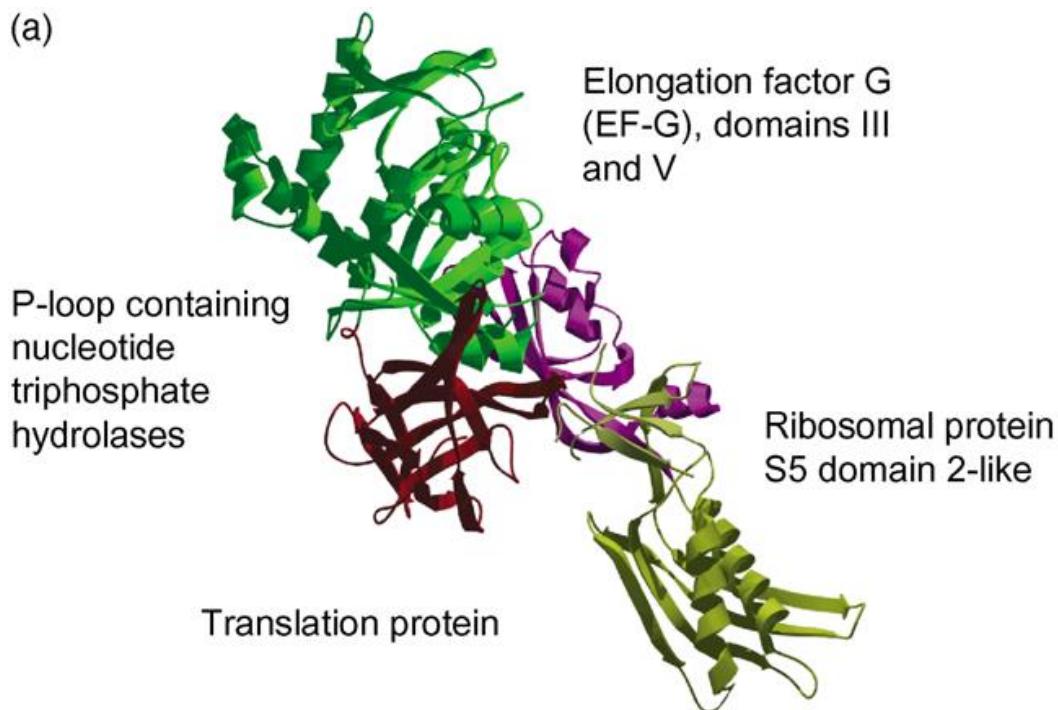
■ Elongation factor G (EF-G) domains III and V

■ P-loop containing nucleotide triphosphate hydrolases

Supradomene

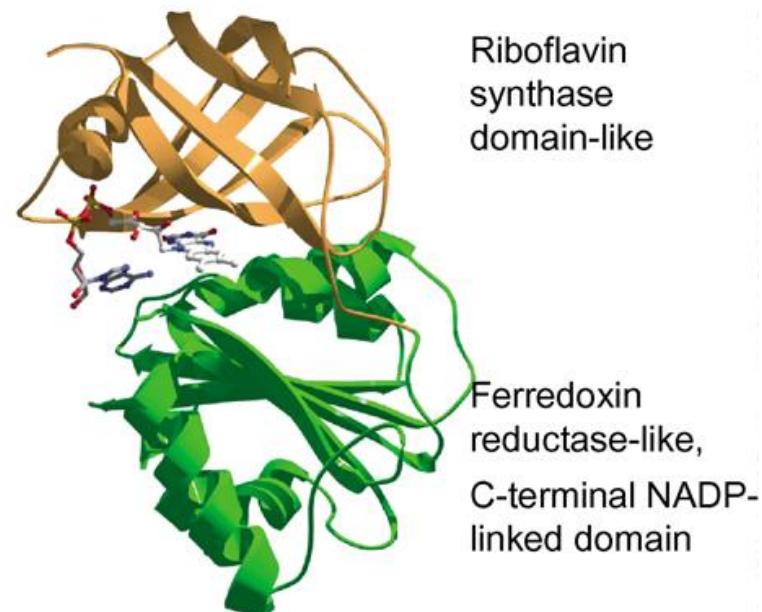
Primera supradomen:

(a)



Elongation factor G (EF-G), 1dar

(b)



Flavodoxin reductase, 1fdr