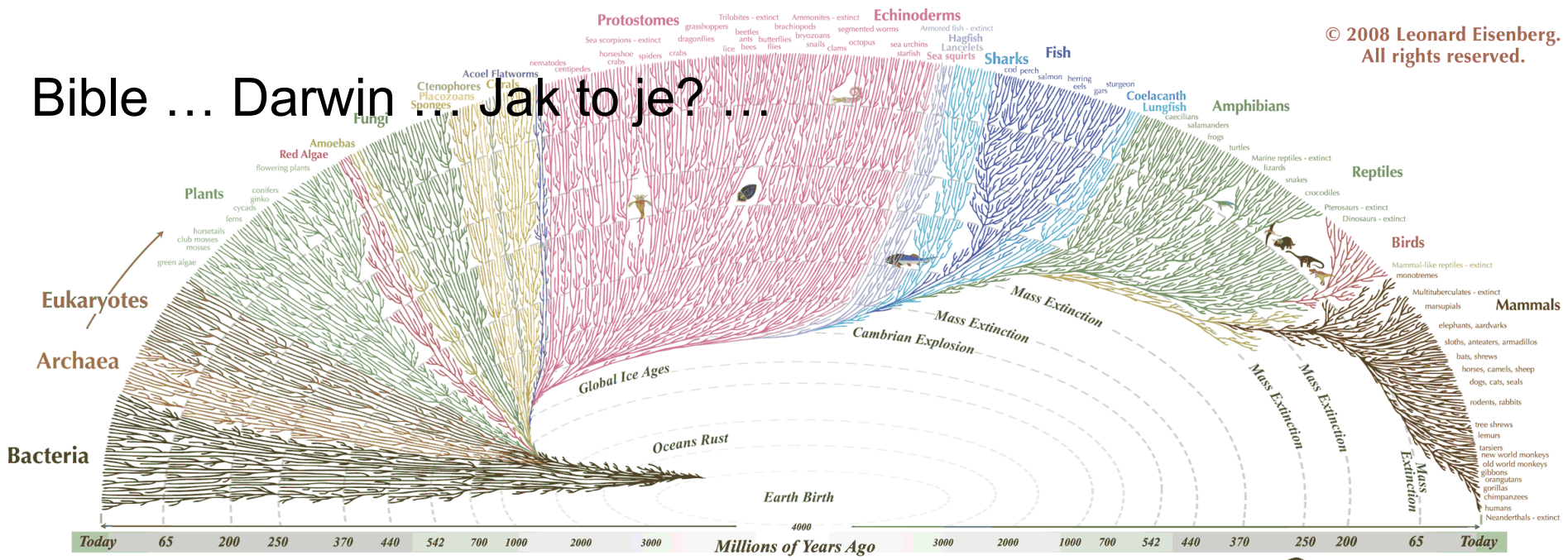



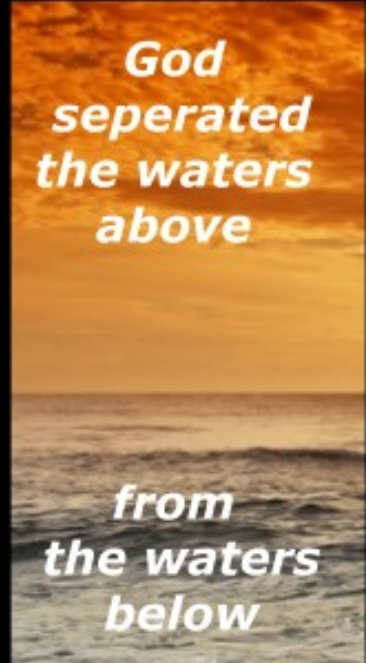
- Úvod - Analýza proteinu
 - Domény
 - fold-struktura (ss, 3D – homolog v PDB? – I-TASSER)
 - Interakce (IntAct, BioGRID)
 - Komplexy
 - Funkce
 - Lokalizace
 - evoluce
- Konkrétní nová data – článek
- V PyMolu připravit 3D movie

Ujasnit si souvislosti, rozšířit si znalosti, aplikovat poznatky z přednášek ...

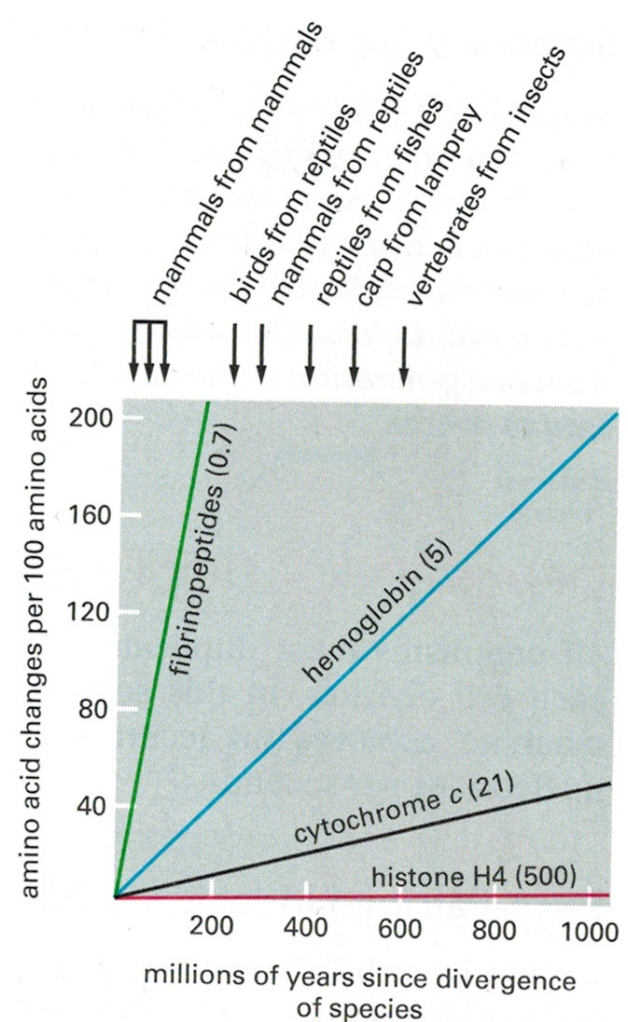
Bible ... Darwin ... Jak to je? ...



All the major and many of the minor living branches of life are shown on this diagram, but only a few of those that have gone extinct are shown. Example: Dinosaurs - extinct 

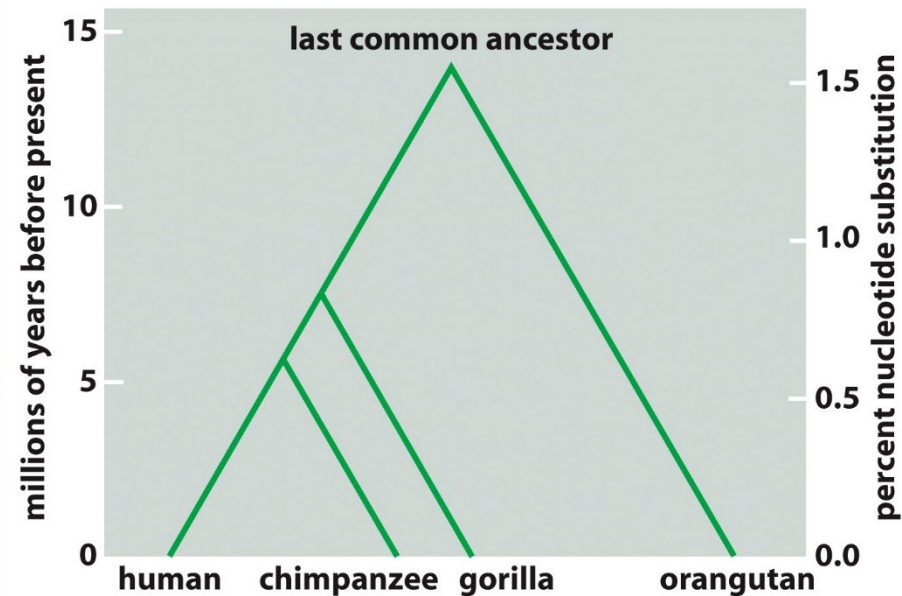
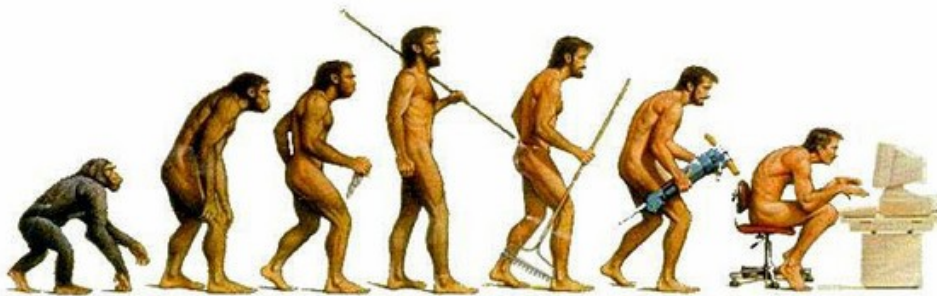


- srovnání genomů/proteomů ukazuje (podobně jako morfologie) na změnu („evoluci“) proteinů v čase
- divergence druhů koreluje do určité míry s konzervací/divergencí DNA/proteinových sekvencí
- DNA je replikována s relativně vysokou přesností (1 změna na 10^9 nukleotidů – cca 4000 TNR_{12} písmen na A4 stranu – 8000 TNR_{12} na list – 500 listů/balík – cca 250 balíčků)



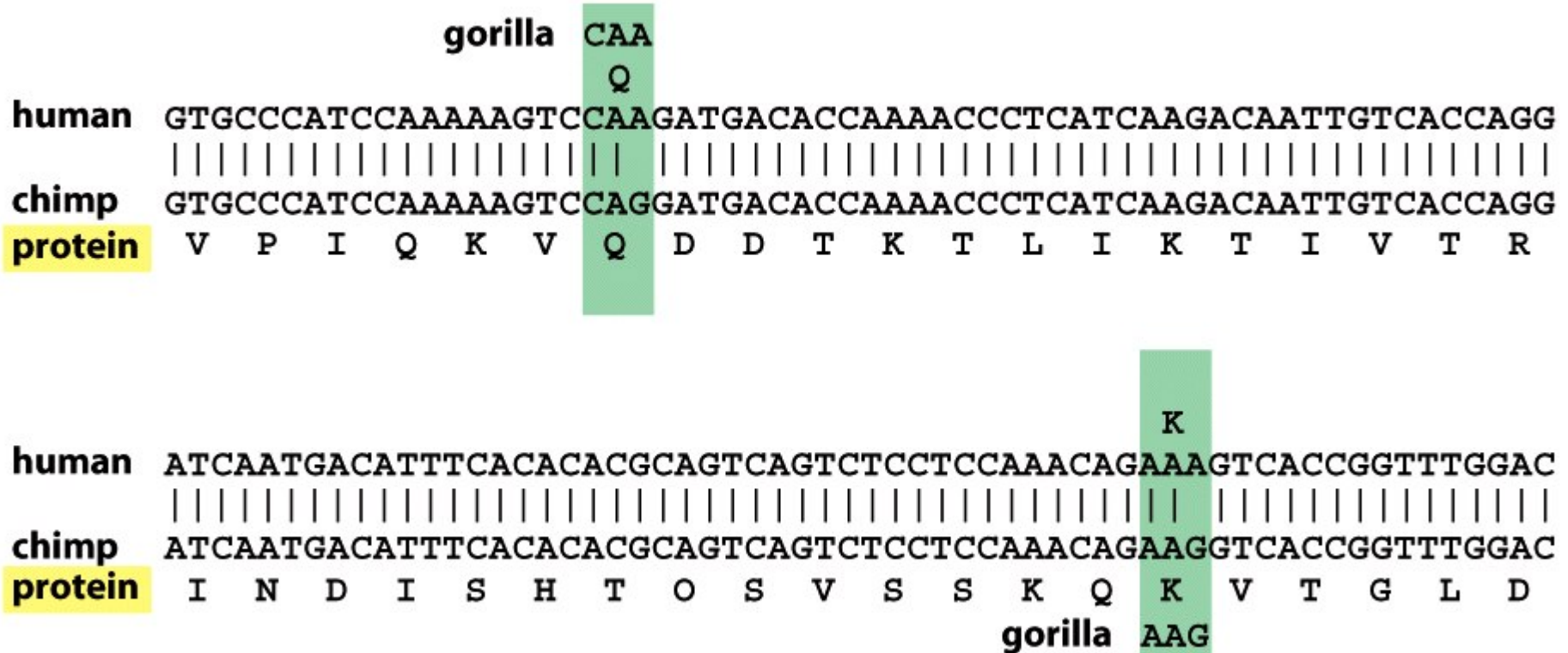
- frekvence mutací DNA je +/- stejná, ale různé proteiny jsou různě změněné – (např. 6 ze 7 změn v cytochromu C jej poškodí; kvasinkový a lidský ubiquitin se liší třemi AMK)
- takto konzervované proteiny jsou lehce identifikovatelné v různých organismech a mají „homologní“ (**ortologní**) funkci

- velmi příbuzné sekvence DNA/proteinů mezi člověkem a šimpanzem (nejbližší člověku) díky „krátké“ době, po kterou mohlo k mutacím docházet – jsou zachovány i nukleotidy ve 3. pozici synonymních kodonů (kodony specifikující stejnou AMK)



- přesný fylogenetický vztah (strom) mezi člověkem a primáty bylo možné určit až dle sekvenací (ne podle morfologických znaků apod.)

- velmi příbuzné sekvence DNA/proteinů mezi člověkem a šimpanzem (nejbližší člověku) díky „krátké“ době, po kterou mohlo k mutacím docházet (jsou zachovány i nukleotidy ve 3. pozici synonymních kodonů - odhad 1/400AMK protein za 200,000 let)



- velkou druhovou rozdílnost nezpůsobuje mnoho mutací (nemohlo jich tolik vzniknout)

- pro vzdálenější organismy (člověk a myš) je sekvence odlišnější (DNA je odlišnější než proteiny – pro 1AMK více kodonů; intron je odlišnější než exon – intr. nekóduje protein)

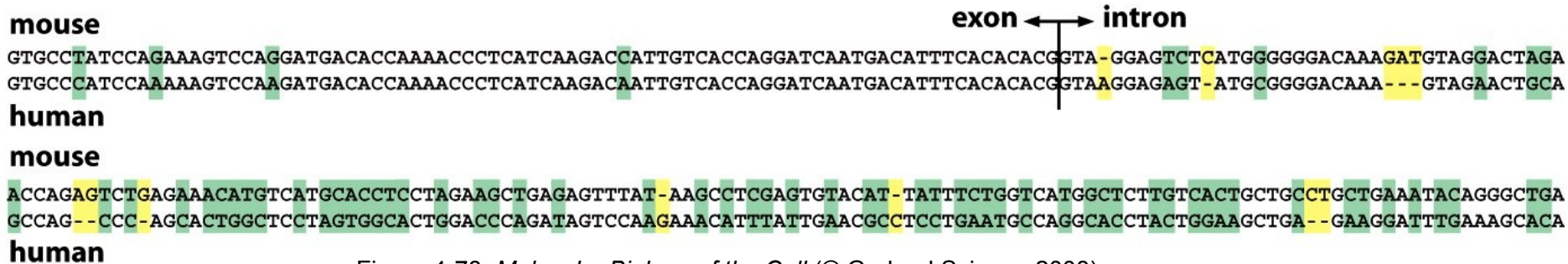


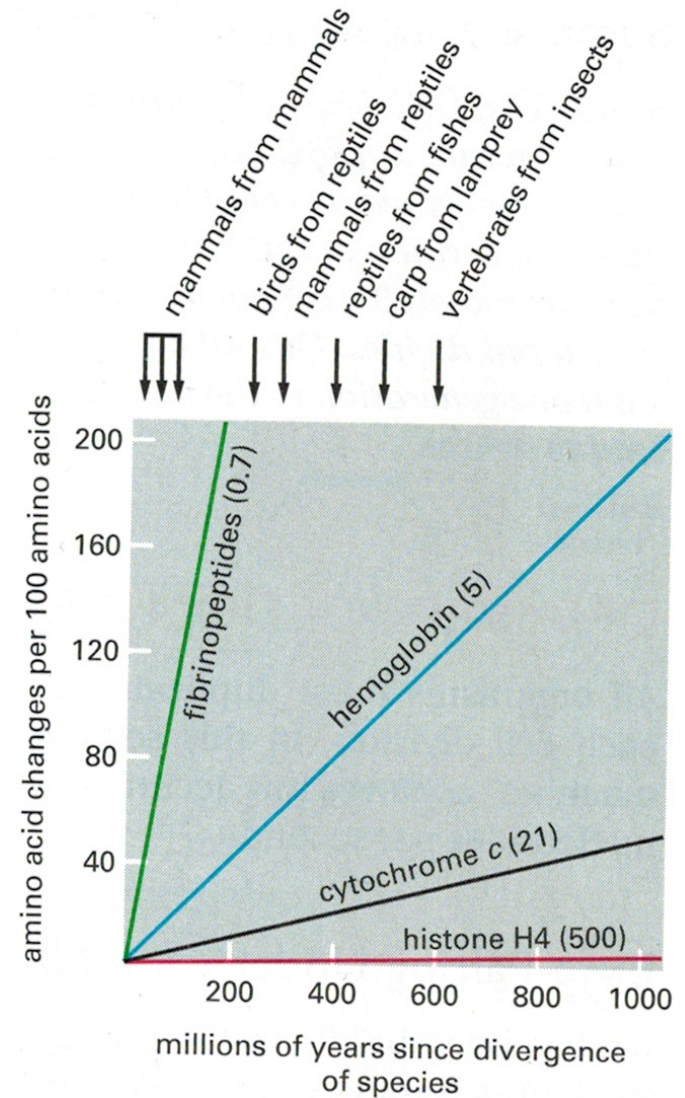
Figure 4-78 *Molecular Biology of the Cell* (© Garland Science 2008)

- čím důležitější funkce proteinu pro organismus (od nejzákladnějších po specializované) tím více je konzervovaná sekvence proteinu respektive jeho domén (velká podobnost i mezi odlišnými organismy)
- konzervace je důsledkem „purifikační selekce“ (eliminace jedinců s mutacemi v esenciálních/důležitých funkcích – většinou proteinech)

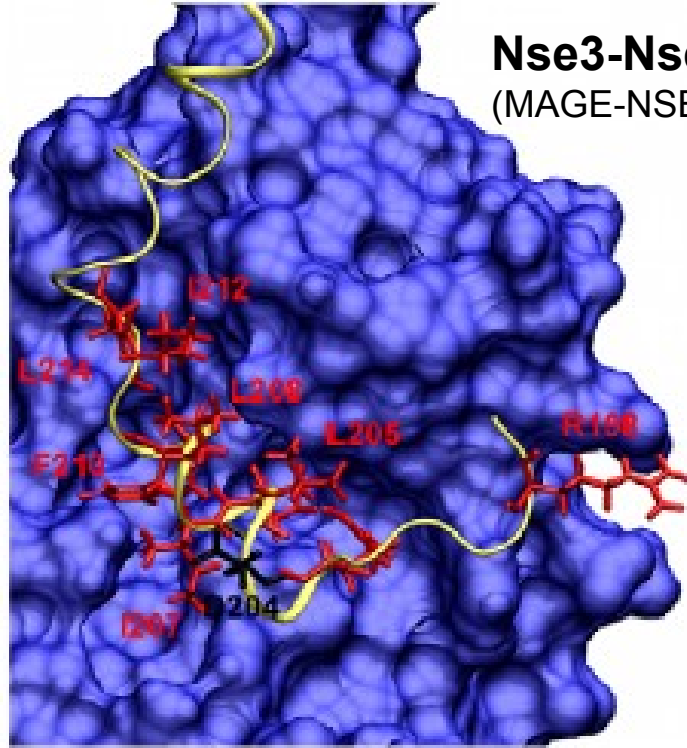
- nejpomaleji se mění proteiny, které jsou zapojeny do nejvíce interakcí s dalšími proteiny (limitována, jak struktura, tak povrch)
- není příliš prostoru pro změny - např. ubikvitin, DNA polymerázy, histony, ribosomální proteiny, ... („drží“ základní systém)

- konzervované proteiny jsou stabilní, optimální pro svoji funkci (enzymatickou aktivitu, pro interakce s partnery ... ko-evoluují celé komplexy)

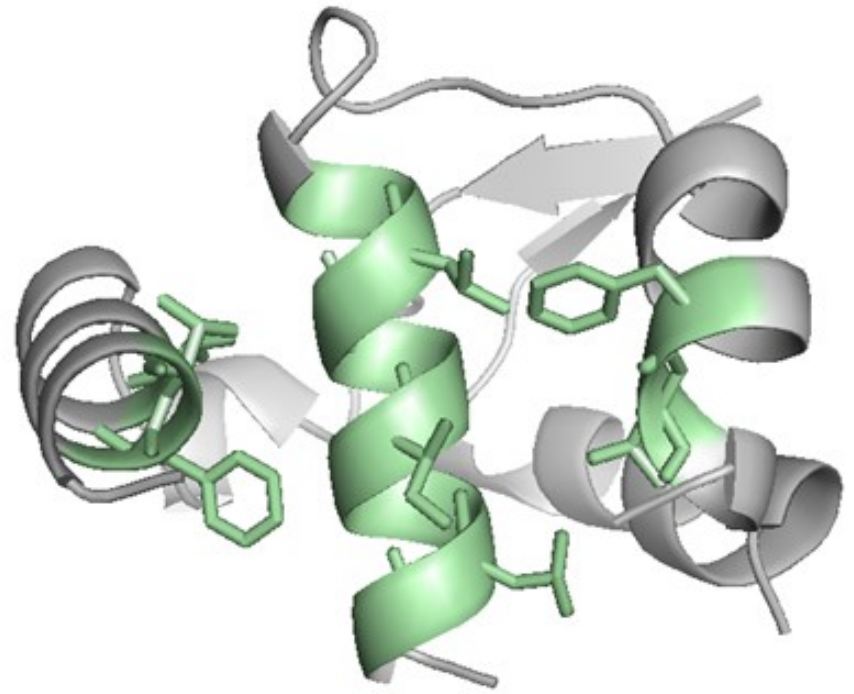
- selekční tlak na stabilitu a funkci - nutnost zachování funkce neposkytuje příliš prostoru pro evoluci/rozvoj nových vlastností (ale neznamená ani selekci na „nejstabilnější“ či „nejaktivnější“ = určitá volnost)



Nse3-Nse4 (MAGE-NSE4)

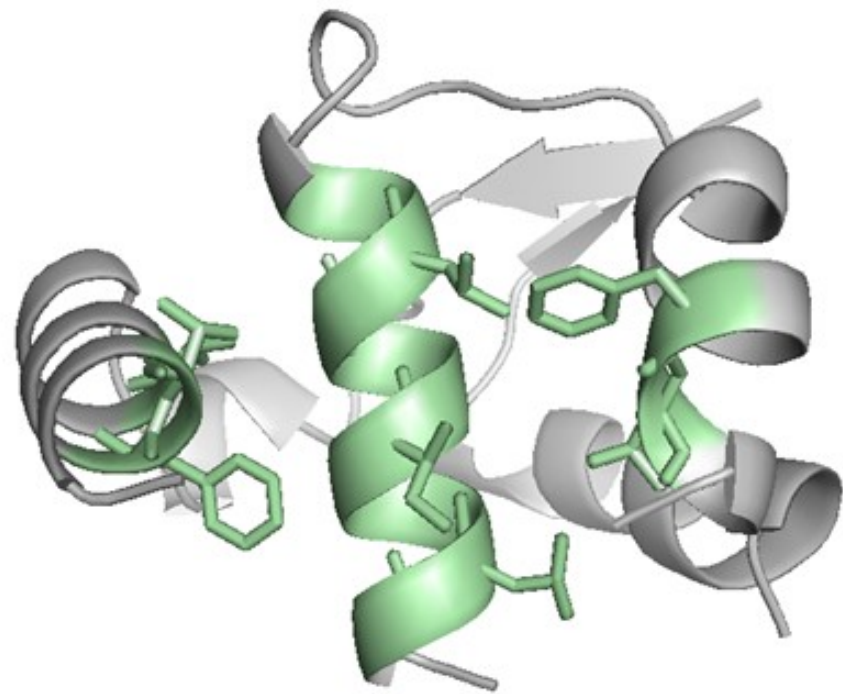
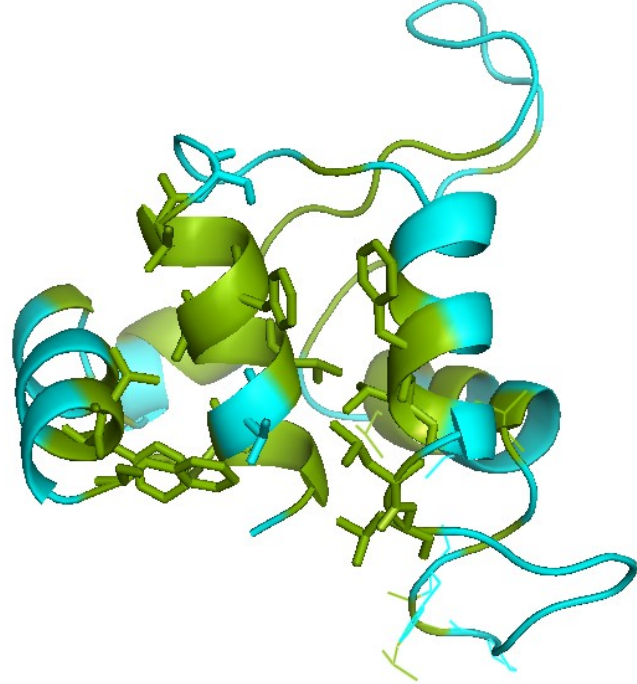


Hudson et al, PLoS One, 2011



...RYBY, OBOŽIVELNÍCI, PTÁCI ...

N.v.	GLLMLVLSVILM	SxxxxN-YTSLW	HFLKMGLEPKKEHEVFGDP	--EKLIAQEFTRQGGYLERRKVTGGE	---	EATFEYSW	GSRSNKE	---	LTKRKVL	EFVS		
D.r.	GLLFLVLSVIFM	KGGTIK	ENLVMNLLKRLRDPGEKHDEF	GDV--KVVVTEEFVVRQKYLEY	GKIPHTE	---	PVEYEFRW	GLRAEKE	---	VSKDKLLE	FVG	
T.n.	GLLFLVLSVIFM	KGGAVR	DSVVMNLLKRLRVQPGERHEPE	GEV--KRVVMEEFVVRQRYL	BCNRI	PHT	ELHEFRW	GQRADTE	---	VSKTKILE	FMA	
X.t.	GLLMVILSLIFM	KGNTAK	ESAVWEMLRRLRIEPAEKHSDF	GDV--KKLITEEFVKQKYLEYS	KVLHTD	---	PVEYEFRW	GQRAFKE	---	TSKMQVLE	FVS	
G.g.	GLLIVILSFIFM	KGNSAK	DSAVWEMLRRLRVHPGEKHEVFGDV	--KKLVMEEFVVRQKYLEIT	PIPLTD	---	PPRFNFQW	GPRAAKE	---	TSKKDIL	SFVA	
O.a.	GLLMVILSLIFM	KGSATN	ESVIWEMLRRLRVDTRERHEVFGDV	--KKLVTEEFVVRQKYLEY	NRI	PHT	PVEYEFQW	GARATKE	---	TTKMQVL	NFVA	
M.d.	GLLMVILSLIFM	KGNSAR	ESLVMNLLKRLRVDPEKRHKTF	GDV--KKLVKDEFVVRQKYLEY	IRVPHSE	---	PEYEFQW	GPRAAHE	---	TSKMQVLR	FVA	
S.e.	GLLMIVLGLIFM	KGNTIK	ETEVMDFLRRLGVYPTKKHLIF	GDV--KKLITEEFVVRQRYL	EYRRIPHT	---	PVDYELQW	GPRTNLE	---	TSKMKVL	KFVA	
C.p.	GLLMIVLGLIFM	KGNTVK	ETEVMDFLRRLGVYPTKKHLV	FGDP--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYELQW	GPRTKLE	---	TSKMKVL	KFVA
M.cr	GLLMIVLGLIFM	KGNTIK	ETEVMDFLRRLGVYPTKKHLV	FGDP--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYEFQW	GPRTNLE	---	TSKMKVL	KFVA
B.t.	GLLMIVLGLIFM	KGNSIK	ETEVMDFLRRLGVHPTKKHLIF	GDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYELQW	GPRTNLE	---	TSKMKVL	KFVA
V.c.	GLLMIVLGLIFM	KGNTIK	ETEVMDFLRRLGVSP	TKKHLIFGDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYEFQW	GPRTNLE	---	TSKMKVL	KFVA
C.f.	GLLMIVLGLIFM	KGNTIK	ETEVMDFLRRLGVYPTKKHLV	FGDP--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYEFQW	GPRTNLE	---	TSKMKVL	KFVA
H.c.	GLLMIVLGLIFM	KGNTIT	ETEVMDFLRRLGVYPTKKHLIF	GDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYELQW	GPRTNLE	---	TSKMKVL	KFVA
Magl	GLLMIVLGLIFM	KGNTIT	ETEVMDFLRRLGVYPTKKHLIF	GDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYELQW	GPRTNLE	---	TSKMKVL	KFVA
MI	GLLMIVLGLIFM	KGNTVK	ETEVMDFLRRLGVYPTKKHF	IFGDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYEFQW	GPRTNLE	---	TSKMKVL	KFVA
P.a.	GLLMIVLGLIFM	KGNTIK	ETEVMDFLRRLGVYPTKKHLIF	GDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYEFQW	GPRTNLE	---	TSKMKVL	KFVA
P.t.	GLLMIVLGLIFM	KGNTIK	ETEVMDFLRRLGVYPTKKHLIF	GDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYEFQW	GPRTNLE	---	TSKMKVL	KFVA
H.s.	GLLMIVLGLIFM	KGNTIK	ETEAWDFLRRLGVYPTKKHLIF	GDV--KKLIT	EDFVVRQRYL	EYRRIPHT	---	PVDYEFQW	GPRTNLE	---	TSKMKVL	KFVA

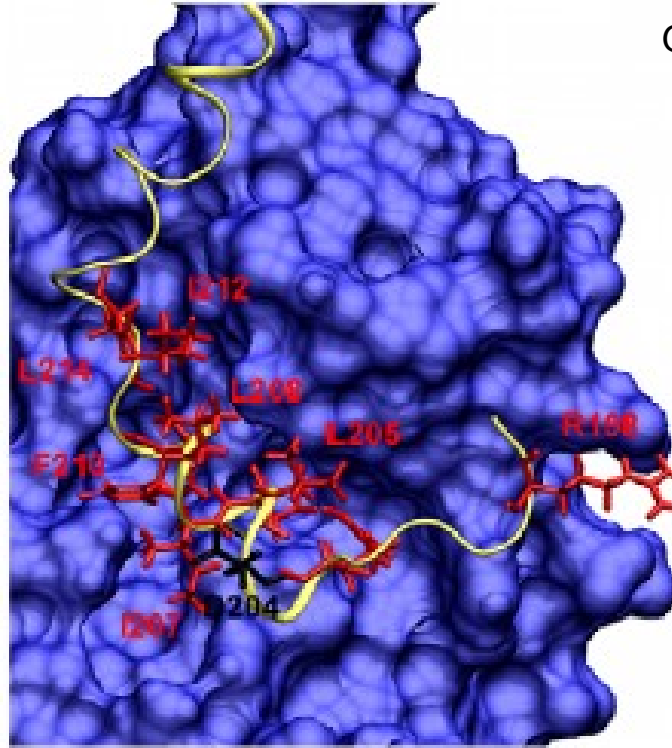


KVASINKY ... určitá volnost v obsazení AMK

S. p.	GFLMTVIAFIIV	S	HCSVG	H	SELQS	F	LQELLT	---	E	ETTP	L	HLD	-	I	TRSL	S	LLVRO	G	YL	--	D	R	V	K	D	D	T	---	H	N	Q	F	V	Y	Y	I	---	G	S	R	A	V	T	E	I	S	-	I	E	G	L	K	S	F	V	T	E	F	F																										
Ani.	GLYTFIIAVILL	N	GGTLQ	E	KLDRYL	S	R	M	N	A	---	E	Q	F	T	P	V	E	R	---	T	H	L	L	Q	R	L	C	K	E	G	Y	L	---	V	K	N	R	E	M	D	G	-	G	D	E	I	E	Y	M	V	---	G	P	R	G	K	V	E	V	G	-	A	R	G	W	A	G	L	V	R	E	V												
Pyr.	ALYTTVIAFIII	S	GGIIP	E	GKIDRAL	R	R	M	N	A	---	D	Q	T	P	L	G	T	---	K	K	T	L	A	A	M	V	K	D	G	Y	I	---	V	K	V	K	D	V	S	G	G	T	E	T	I	D	Y	I	V	---	G	P	R	G	K	V	E	V	G	-	G	E	G	W	A	Q	F	I	R	A	V													
N. f.	GLYSFIIAVILL	N	GGSLP	E	KLERYL	K	R	T	N	A	---	D	T	Y	T	P	V	D	R	---	T	D	R	F	L	Q	R	L	C	K	E	G	Y	L	---	I	R	N	R	E	M	D	G	-	G	E	E	I	E	Y	M	V	---	G	P	R	G	K	V	E	V	G	-	V	Q	G	W	A	G	L	V	R	E	V											
Ate.	GLYTFIIALILL	N	GGSLP	E	KLERYL	Q	R	T	N	T	---	D	T	Y	T	P	I	D	R	---	T	D	R	F	L	Q	R	L	C	K	D	G	Y	L	---	V	R	N	R	E	M	D	S	-	G	E	E	V	E	Y	M	V	---	G	P	R	G	K	I	E	V	G	-	T	Q	G	W	A	G	L	V	R	E	V											
Acl.	GLYSFIIAVITL	N	GGSLP	E	KLERYL	K	R	T	N	A	---	D	T	F	T	P	I	D	R	---	T	D	R	F	L	Q	R	L	C	K	E	G	Y	L	---	L	R	T	R	E	M	D	G	-	G	E	E	V	V	E	Y	L	V	---	G	P	R	G	K	I	E	V	G	-	V	Q	G	W	A	G	L	V	R	E	V										
C. i.	ALYTFIISLITL	S	GGSLA	E	KLDRYL	R	R	V	N	A	---	D	T	Y	T	P	L	D	R	---	T	E	K	L	A	R	L	C	K	D	G	Y	L	---	V	R	N	R	D	V	D	G	-	G	E	E	V	V	E	Y	L	V	---	G	P	R	G	K	I	E	V	G	-	T	E	G	W	A	G	L	T	R	E	V											
P. a.	ALYTTIISLITL	S	GGELS	D	T	R	L	R	R	H	A	R	L	N	A	A	E	Y	M	P	S	M	N	P	N	D	F	x	-	T	D	V	L	Q	R	M	I	K	H	G	Y	L	---	V	R	V	D	N	R	G	x	D	D	D	S	T	T	H	V	---	G	P	R	G	K	A	E	V	P	-	K	E	S	I	A	G	F	V	R	T	I	Y			
N. c.	GLYTMLIAIITL	S	GGELS	D	P	R	L	R	R	Y	L	T	R	L	N	A	A	x	F	M	N	E	N	A	P	S	E	K	---	T	E	L	V	L	Q	R	M	T	K	O	G	Y	L	---	V	R	V	A	D	N	x	A	G	D	D	D	A	I	T	H	V	---	G	P	R	G	K	V	E	V	D	-	N	E	A	I	A	A	V	R	E	V			
M. g.	GLYSMIVTIIQL	N	R	G	E	L	S	-	D	P	K	L	K	R	Y	L	Q	R	L	N	A	---	E	T	N	T	P	V	E	K	---	T	E	L	L	L	Q	R	L	I	R	O	N	Y	I	---	V	K	T	V	E	R	N	x	---	D	D	A	I	T	H	R	V	---	G	P	R	A	K	Q	E	L	T	-	D	E	A	M	A	S	I	V	R	D	V

O
B
R
A
T.

GLLMLVLSVILM	S	x	x	x	x	N	-	Y	T	S	L	W	H	F	L	K	K	M	G	L	E	P	K	K	E	H	E	V	F	G	D	P	---	E	K	L	I	A	Q	E	F	T	R	O	G	Y	D	E	R	R	K	V	T	G	G	E	---	E	A	T	F	E	Y	S	W	---	G	S	R	S	N	K	E	---	L	T	K	R	K	V	L	E	F	V	S
GLLEFVILSVIFM	K	G	G	T	I	K	-	E	N	L	V	W	N	I	L	K	K	L	R	L	D	P	G	E	K	H	D	E	F	G	D	V	---	K	K	V	V	T	E	E	F	V	R	O	K	Y	L	E	Y	G	K	I	P	H	T	E	---	P	V	E	Y	E	F	R	W	---	G	L	R	A	E	K	E	---	V	S	K	L	K	L	E	F	V	G	
GLLEFVILSVIFM	K	G	G	A	V	R	-	D	S	V	W	N	I	L	K	K	L	R	V	Q	P	G	E	R	H	E	F	E	G	V	---	K	R	V	V	M	E	E	F	V	R	O	R	Y	L	E	C	N	R	I	P	H	T	E	---	E	L	E	H	E	F	R	W	---	G	O	R	A	D	T	E	---	V	S	K	T	K	I	L	E	F	M			
GLLMVILSLIFM	K	G	N	T	A	K	-	E	S	A	V	W	E	M	L	R	R	L	R	I	E	P	A	E	K	H	S	D	F	G	D	V	---	K	K	L	I	T	E	E	F	V	K	O	K	Y	L	E	Y	S	K	V	L	H	T	D	---	P	V	E	Y	E	F	R	W	---	G	O	R	A	F	K	E	---	T	S	K	M	Q	V	L	E	F	V	
GLLIVILSFIFM	K	G	N	S	A	K	-	D	S	A	V	W	E	F	L	R	R	L	R	V	H	P	G	E	K	H	E	V	F	G	D	V	---	K	K	L	V	M	E	E	F	V	R	O	K	Y	L	E	I	T	P	I	P	L	T	D	---	P	P	E	F	N	F	Q	W	---	G	P	R	A	A	K	E	---	T	S	K	K	D	I	L	S	F	V	
GLLMVILSLIFM	K	G	S	A	T	N	-	E	S	V	I	W	E	T	L	R	K	L	R	V	D	T	R	E	R	H	E	V	F	G	D	V	---	K	K	L	V	T	E	E	F	V	R	O	K	Y	L	E	Y	N	R	I	P	H	T	E	---	P	V	E	F	E	F	Q	W	---	G	A	R	A	T	K	E	---	T	T	K	M	Q	V	L	N	F	V	
GLLMVILSLIFM	K	G	N	S	A	R	-	E	S	L	W	D	V	L	K	K	L	R	V	D	P	E	K	R	H	K	T	F	G	D	V	---	K	K	L	V	K	D	E	F	V	R	O	K	Y	L	E	Y	I	R	V	P	H	S	E	---	P	P	E	Y	E	F	L	W	---	G	P	R	A	A	H	E	---	T	S	K	M	Q	V	L	R	F	V		
GLLMIVLGLIFM	K	G	N	T	I	K	-	E	T	E	W	D	F	L	R	R	L	G	V	Y	P	T	K	K	H	L	I	F	G	D	P	---	K	K	L	I	T	E	E	F	V	R	O	R	Y	L	E	Y	R	R	I	P	H	T	D	---	P	V	D	Y	E	L	Q	W	---	G	P	R	T	N	L	E	---	T	S	K	M	K	V	L	K	F	V		



- selekční tlak na strukturu, na povrch = interakce
- mutace, které neruší jsou neutrální (protein je částečně modifikován)
- modifikace je kompenzována
- (i později) mutací partnera
- změní se v čase a v budoucnu může přinést novou vlastnost

Nse4

aa75 aa103

Species	Sequence
S.p.	LHIGRPKEIE-LFTKNIKQFLNYPTSHS
A.n.	L-x-SAGIDVDFVSKCISFMRRAPADSQ
N.f.	L-x-SAGIDVDFVSKCISFMRRGSPDP
A.t.	L-x-TARIDVDFVSKCISFMRRAPGDSQ
A.o.	L-x-SAGIDVDFVSKCISFMCRAPEDSQ
A.c.	L-x-SAGIDVDFVSKCISFMRRGSPDPE
D.r.	LHAEGLSSFDLPSAFALHLLSFMDLNRLEDE
X.l.	LHADMTVFDPSTFTADLLSFMGLNRMESP
G.g.	LIRSEITTFDSLTFADLLTFMGINRTETE
M.d.	LHSDMTLFDLPSFAEDLLKFMGLNRLVEE
C.f.	LRSIDLSSFDMLRYVETLLTHMGMVPLEAE
M.m.	LRSIDLSSFDMLRYVETLLTHMGMVPLEAQ
H.s.	LRSIDLSSFDMLRYVETLLTHMGMVPLEAE

Nse3/MAGE-binding domain

NSE4b/EID3

D.n.	LNSDMSSFEVAFCDLFLFVGLNWMEDD
C.f.	LNSDMSSFEVAFCDLFLFVGLNWMEDD
M.m.	LNTDMNFFNLAFCDLLLLFVGFVWVEEE
H.s.	LNSDMNFFNLAFCDLFLFVGLNWMEDD
aa106++ +++++ + + -++ ++ + aa135	

core region

KVASINKY ... určitá volnost v obsazení AMK

S.p.	GF LMTVIAFI AV- SHCSVG -HSELQSFLQELLT---EETTP LHL D-ITRSL S L L V R G G Y L---DRVKDDT---HNQFVYYI--GSRAVTEIS-IEGLKSEVTEFF
Ani.	GLYTFIIAVILL-NGGITLQ-EKLD RYL SRMNA---E Q F T P V E R--T E H L L Q R L C K E G Y L---VKNREMDG-GDEIIEYMV-GPRGKVEVG-ARGWAGLVREVV
Pyr.	ALYTTVIAFI I I-SGGIIP-EGKLD RAL RRMNA---DQTTPLGT--K K T L A A M V K D G Y I---V K V K D V S G G T E E T I D Y I V--G P R G K V E V G-G E G W A Q F I R A V Y
N.f.	GLYSFIIAVILL-NGGSLP-EKLD RYL KRTNA---D T Y T P W D R--T D R F L Q R L C K E G Y L---I R N R E M D G-G E E I I E Y M V-G P R G K V E V G-V O G W A G L V R E V Y
Ate.	GLYTFIIALILL-NGGSLP-EKLD RYL QRTNT---D T Y T P I D R--T D R F L Q R L C K D G Y L---V R N R E M D S-G E E V I E Y M V-G P R G K I E V G-T O G W A G L V R E V Y
Ac1.	GLYSFIIAVITL-NGGSLP-EKLD RYL KRTNA---D T F T P I D R--T D R F L Q R L C K E G Y L---L R T R E M D G-G E E V V E Y L V-G P R G K I E V G-V O G W A G L V R E V Y
C.i.	ALYTFIISLITL-SGGSLA-EKLD RYL RRVNA---D T Y T P L D R--T E K L L A R L C K D G Y L---V R N R D V D G-G E E V V E Y L V-G P R G K I E V G-T E G W A G L T R E V Y
P.a.	ALYTTIISLITI-SGGELS-DTRLRRHLARLNAAEYMP SMNPNDP x-TD V V L Q R M I K H G Y L---V R V V D N R G x D D D S T T W H V-G P R G K A E V P-K E S I A G F V R T I Y
N.c.	GLYTMLIAIITL-SGGELS-DPRLRRYLTRLNAXPMENAPSEK--TELVLQRMTKOGYL---V R V A D N x A G D D D A I T W H V-G P R G K V E V D-N E A I A A V V R E V Y
M.g.	GLYSMIVTIIQL-NRGE L S-D P K L K R Y L Q R L N A---E T N T P V E K--T E L L L Q R L I R Q N Y I---V K T V E R N x--D D A I T W R V-G P R A Q E L T-D E A M A S I V R D V Y

O
B
R
A
T.

GLLMLVLSVILM-SxxxxN-YTSLWHLKKGLEPKKEHEVFGDP--EKLIAQEFTRQGYLERRKVTGGE---EATFEBYSW-GSRNKEL---LTKRQVLEFVS:
GLLEFVLSVIFM-KGGTIK-ENLVWNLKRLRDPGEKHDEFQGDV--KKVVTEEFVRQKYLDEYKIPHTE---PVEYEFQW-GLRAEKE---VSKDKLEFVG:
GLLEFVLSVIFM-KGGAVR-DSVWVNLKRLRVQPERHEPFGEV--KRVMEEFVRQRYLBCNRIHTE---ELHEFQW-GORADTE---VSKTKILEFMA:
GLLMVILSLIFM-KGNTAK-ESAVVEMLRRLRIEPAEKHSDFGDV--KKLITTEFVKQKYLEYSKVLHTD---PVEYEFQW-GORAFKE---TSKMQLVLEFVS:
GLLIVILSIFM-KGNSAK-DSAVVWFLRRLRVHPGEKHEVFGDV--KKLVMEEFVRQKYLEITPIPLTD---PPVFNQW-GPRAAKE---TSKKDILSFVA:
GLLMVILSLIFM-KGSAI N-ESVIMWTLRKL RVDTRE RHEVFGDV--KKLVTEEFVRQKYLEYNRIPHTE---PVEYEFQW-GARATKE---TTKMQVLENFVA:
GLLMVILSLIFM-KGNSAR-ESLWVDV LKRL RVDPEKRHKTFGDV--KKLVKDEFVRQKYLEYIRVPHSE---PPVEYFLW-GPRAAHE---TSKMQVLEFVA:
GLLMIVLGLIFM-KGNTIK-ETEVDVFLRRLGVYPTK KHLIFGDV--KKLITTEFVRQRYLEYSRRIPHTE---PVDYELQW-GPRTNLE---TSKMKVLEKFA:

- selekční tlak na stabilitu a funkci - nutnost zachování funkce neposkytuje příliš prostoru pro evoluci/rozvoj nových vlastností
- pro rozvoj nových vlastností, nových druhů - spíše než druhově specifické mutace proteinů lze vidět **expansi různých genových/proteinových rodin** v různých živočišných druzích

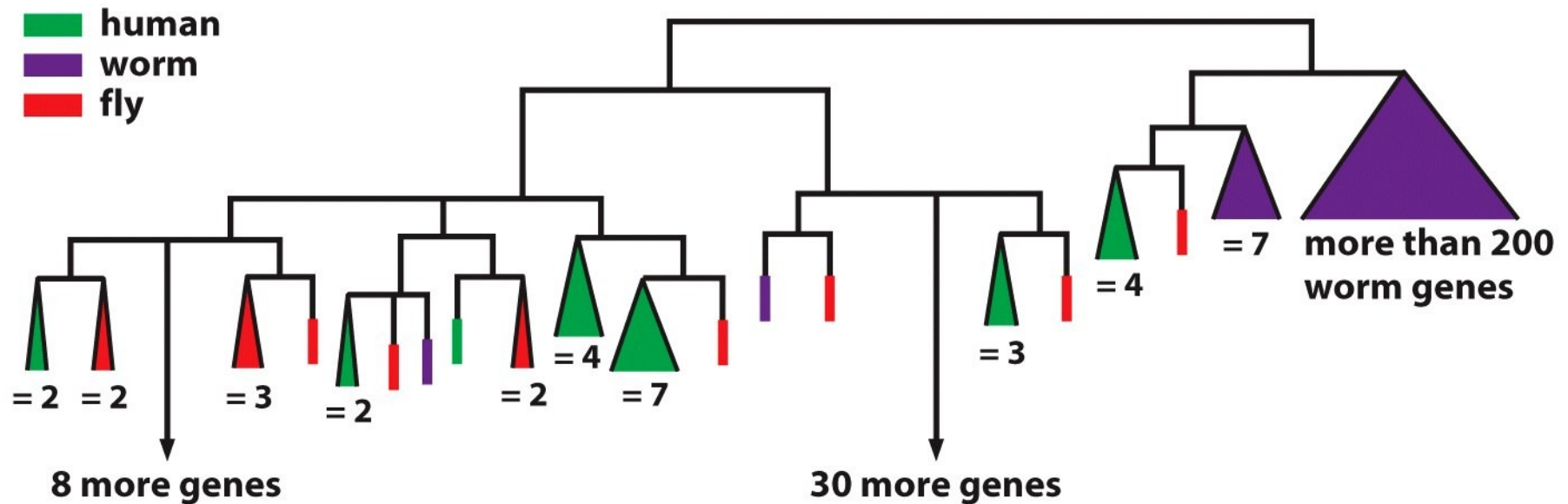
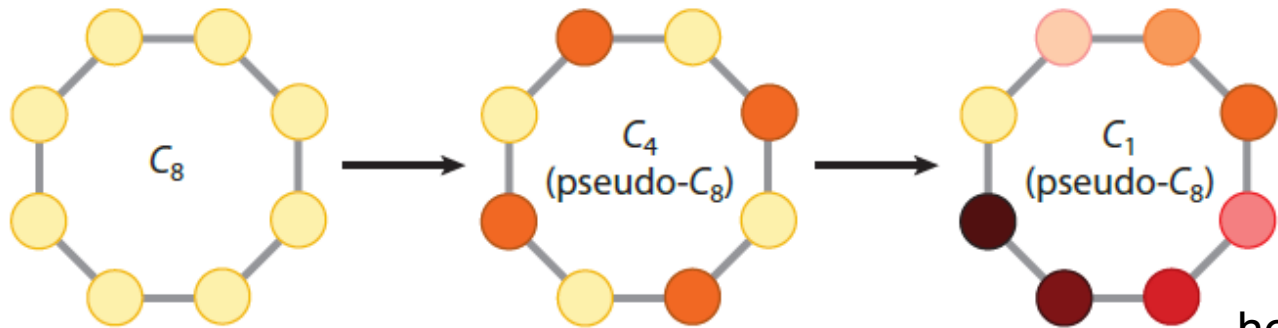


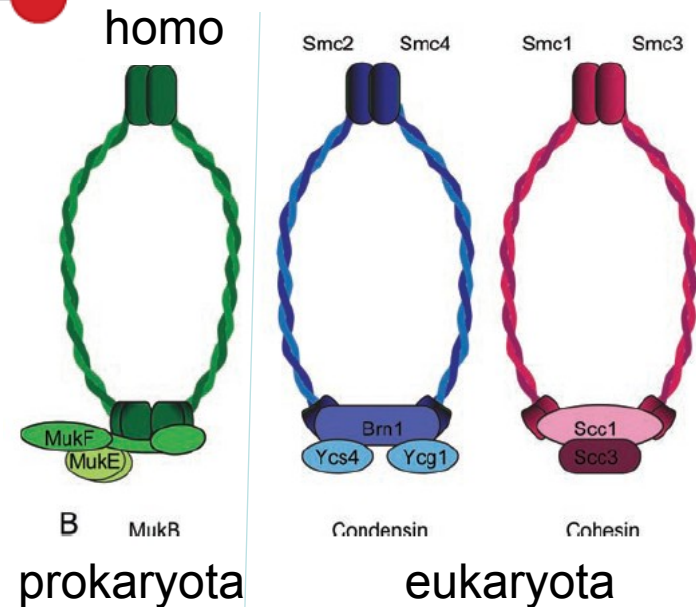
Fig. 7-114

- **duplikace a divergence** jsou hlavními tahouny evolučních procesů (všechny geny/proteiny jsou „potomky“ několika ancestrálních genů/proteinů (foldů), které existovaly v nejčasnějších živých formách (nyní cca 1000 foldů na >100000 struktur v PDB, odhad cca 2000 foldů)

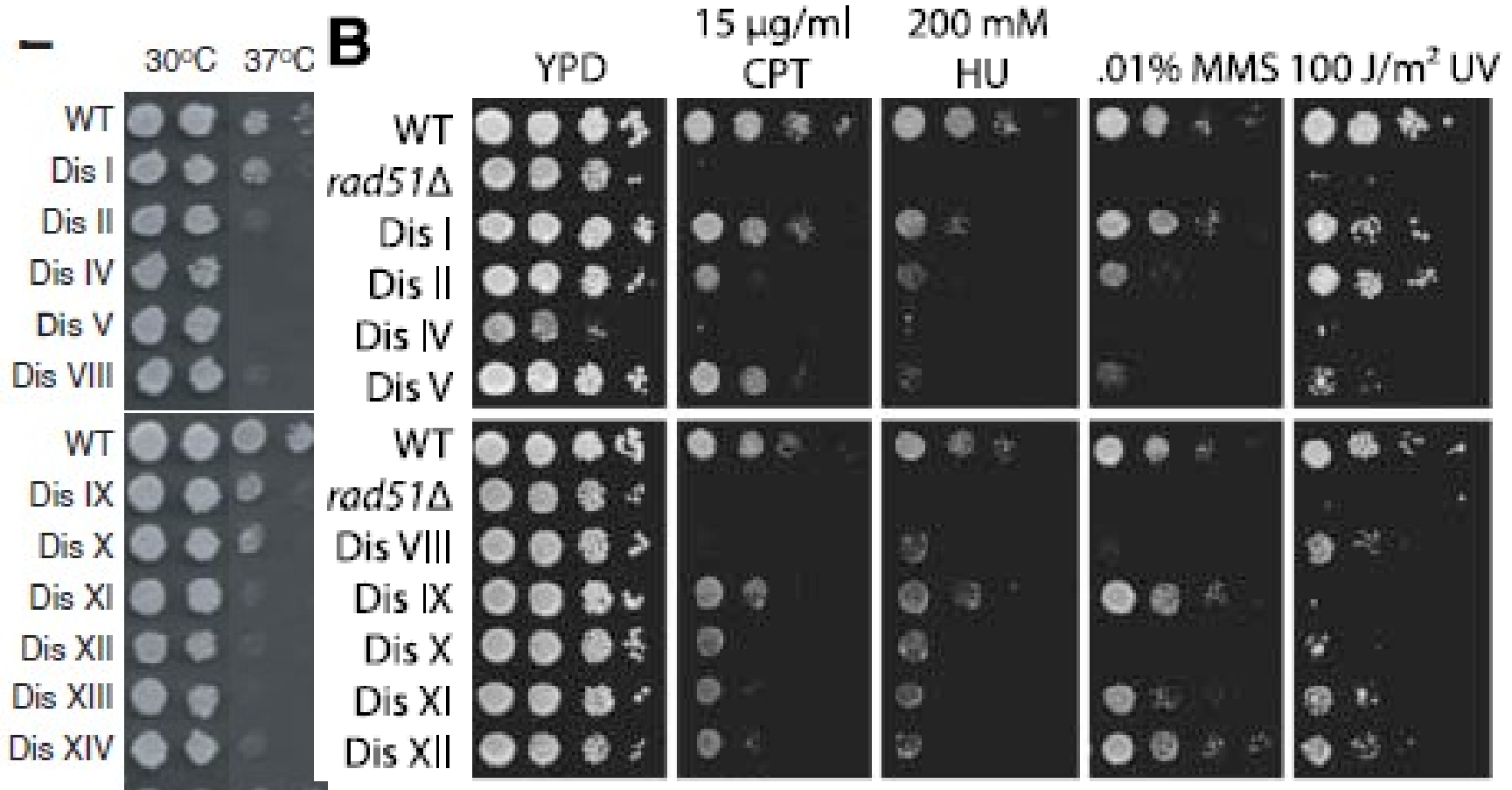
a Gene duplication



- po **duplikaci** jsou oba proteiny stejné a vytváří stejný homomerní komplex – později jeden protein diverguje a vzniká heteromer



- duplikace – na počátku stejné sekvence = stejné funkce
- vyšší hladina proteinu/ů může být toxická (**aneuploidie** – kvasinky s 1 chromosomem navíc, nádorové buňky)

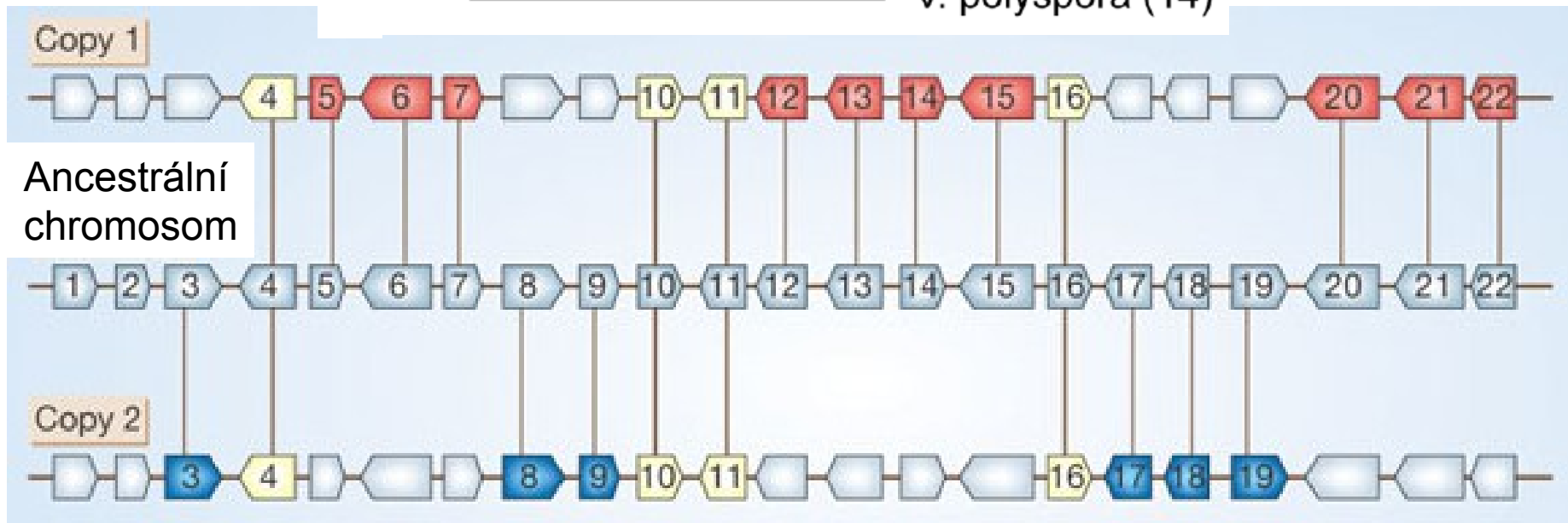
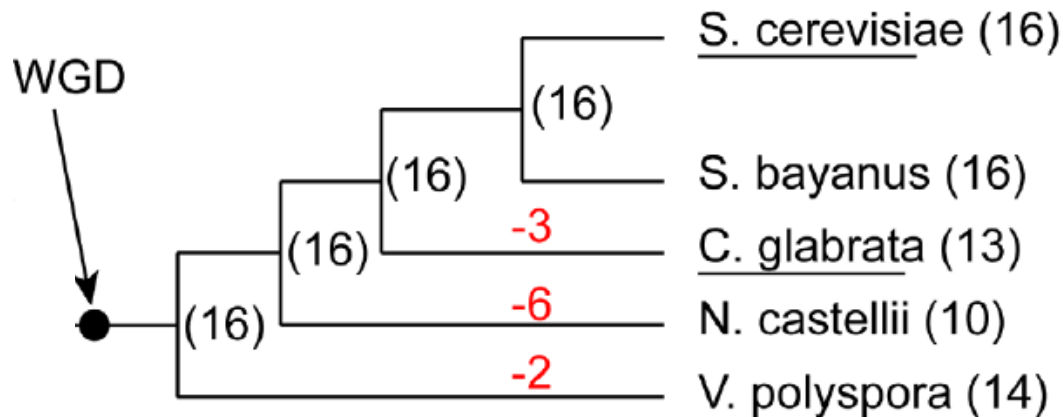


Shelzer et al, Science (2011)

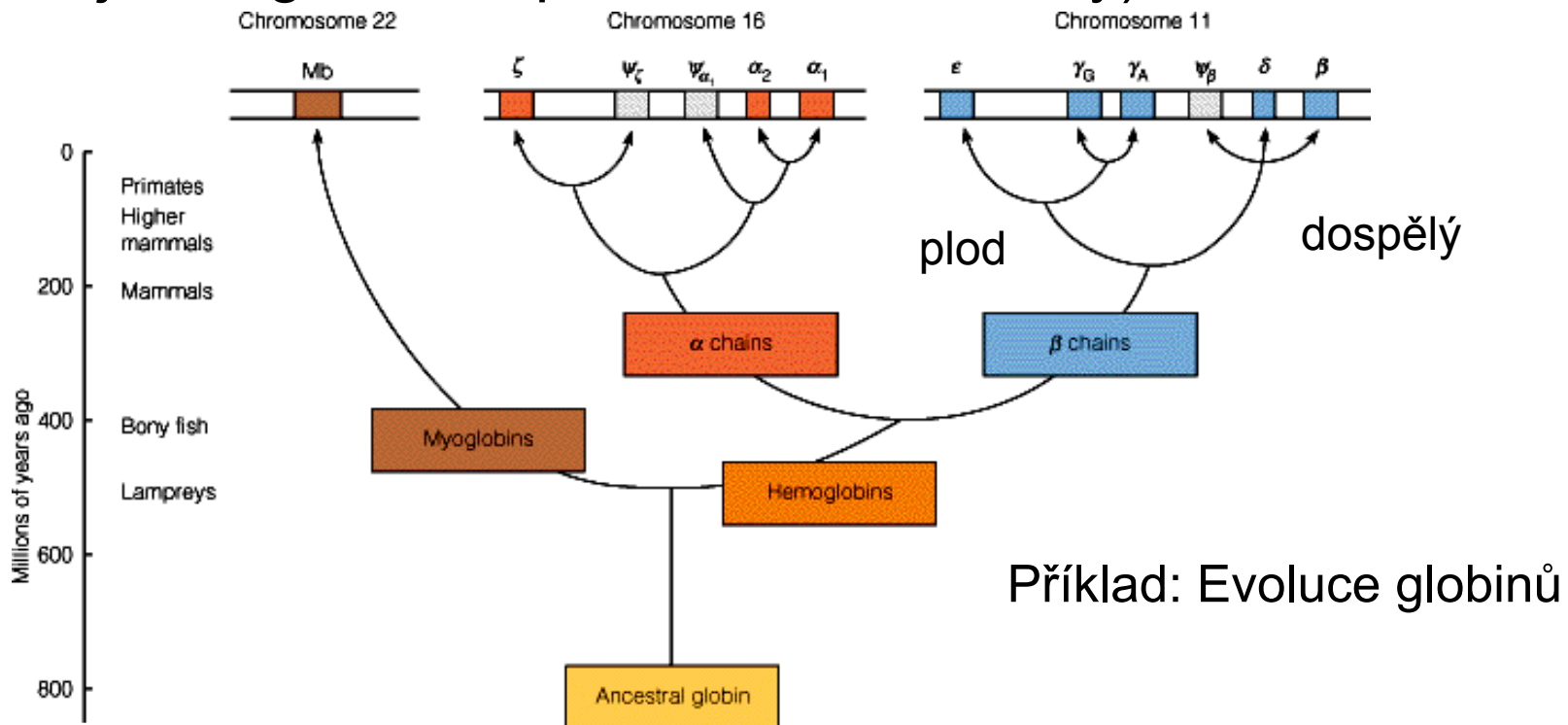
- nadbytek proteinů (např. transkripčních faktorů) nebo nerovnováha podjednotek komplexů znamená disregulaci některých procesů

Celogenomová duplikace u kvasinek

cca 30% genomu *S.c.* vzniklo duplikacemi => došlo k **celogenomové duplikaci** (WGD) => a poté došlo k přeskupování a redukci segmentů (i chromosomů) – polyploidie u kyttek ...



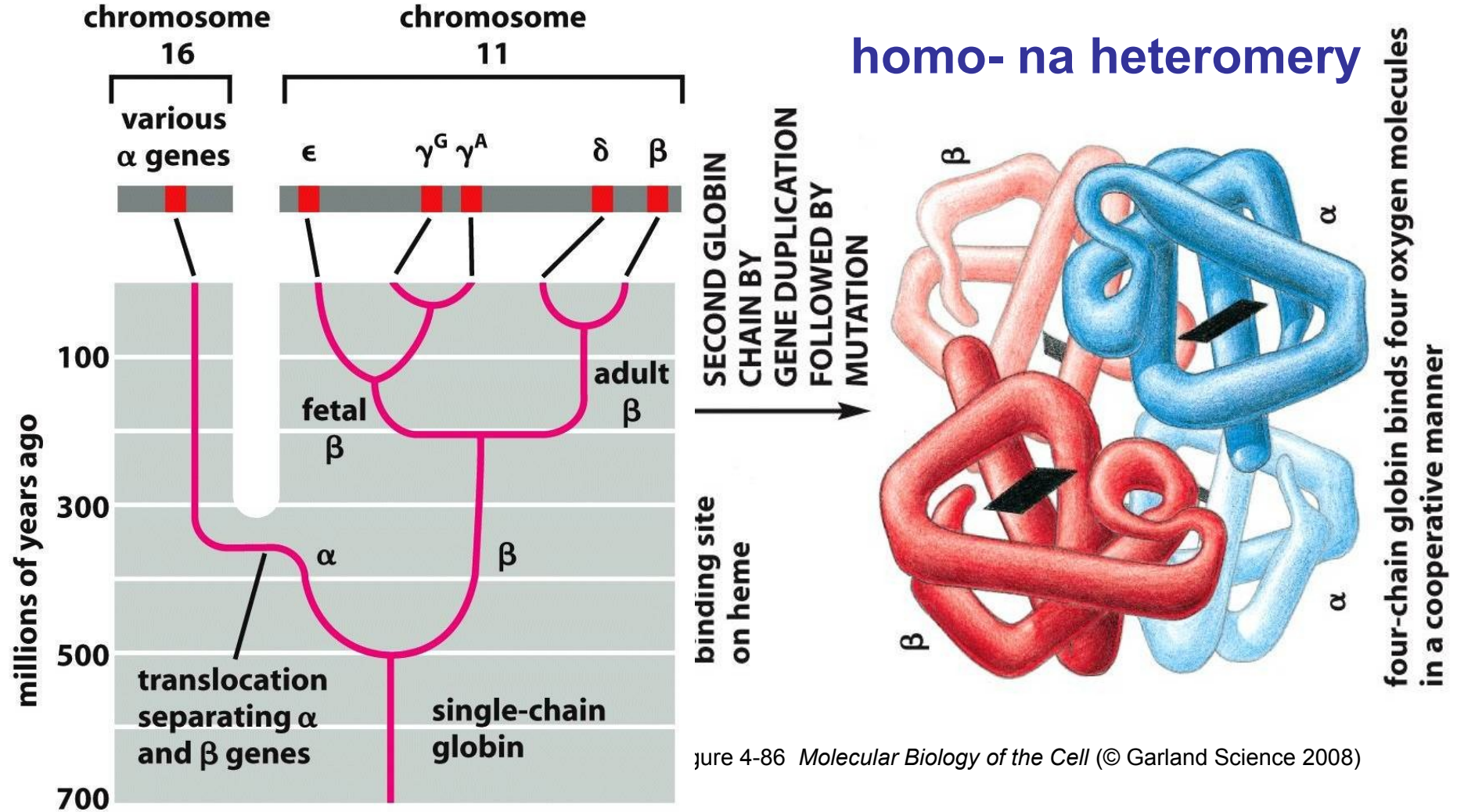
- duplikace – **reverzní transkripcí** a integrací DNA (pouze individuální geny nebo pouze domény = exony)
- na počátku stejné sekvence = stejné funkce (pod jinými **promotory** – jiný lokus tzn. jiné „okolí“ zaintegrované „mRNA“ tzn. jiná regulace exprese – „nové“ buňky)



- následují **mutace** – inaktivující tj. pseudogeny (ustaví hladinu proteinu zpět na původní) nebo nefunkční (zatěžují expresní-chaperonový aparát – snaha odstranit)

Evoluce rodiny globinů (duplikace)

- červy, hmyz a primitivní ryby mají jeden globin (150AMK)

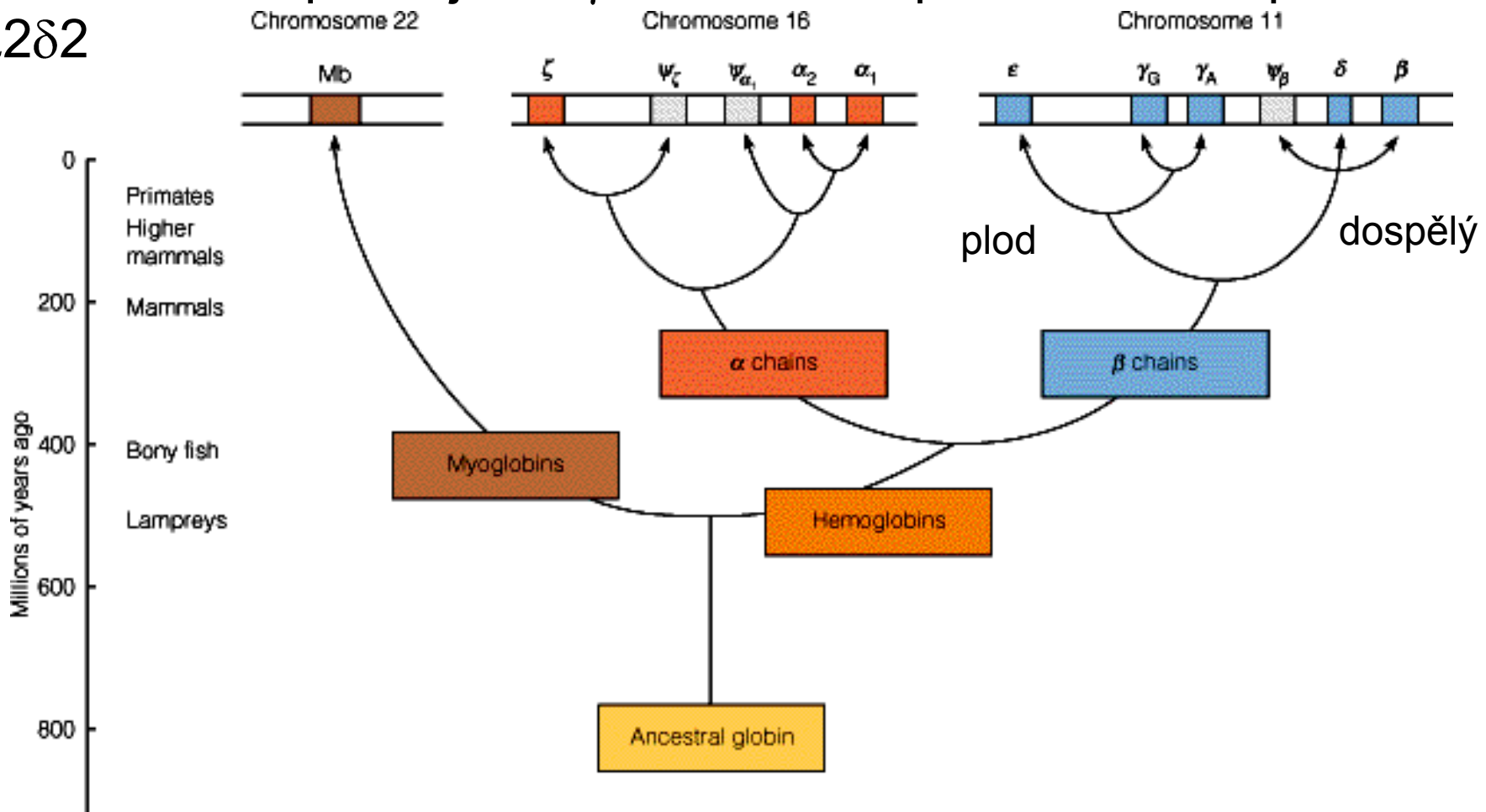


- vyšší obratlovci mají α - a β -globin (tvoří $\alpha_2\beta_2$ komplex – účinnější přenos kyslíku) – vyvinul se u vyšších ryb (500Mya)

Evoluce rodiny globinů (duplikace)

- u savců se dále duplikoval β -globin (**paralog**) - exprimován specificky v embryu – má vyšší afinitu ke kyslíku a napomáhá přenosu kyslíku z krve matky do krve plodu

- dále se duplikoval a specializoval se na časná vývojová stádia $\alpha_2\varepsilon_2$ a pozdější $\alpha_2\gamma_2$ - k další duplikaci došlo u primátů $\alpha_2\delta_2$



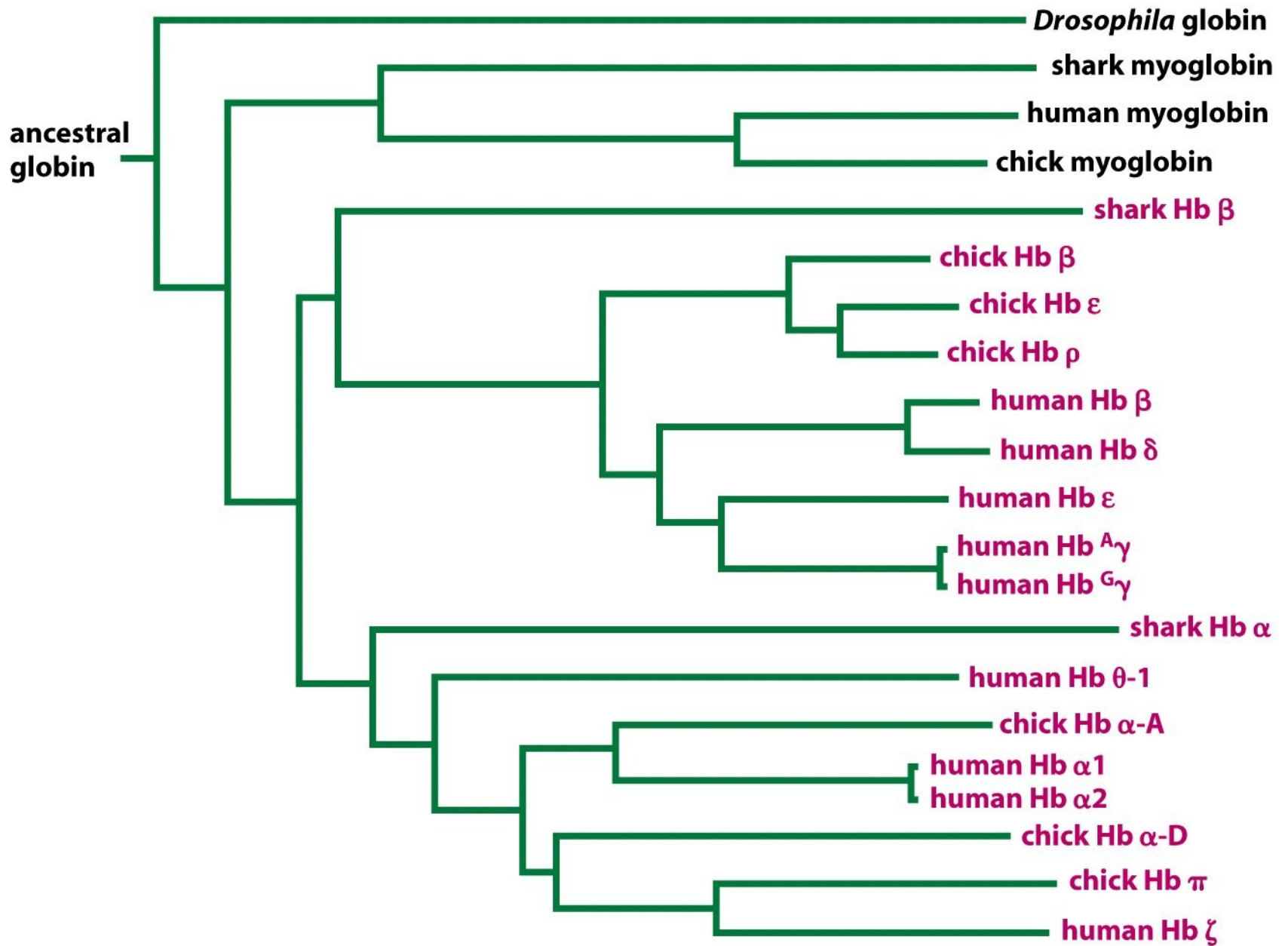
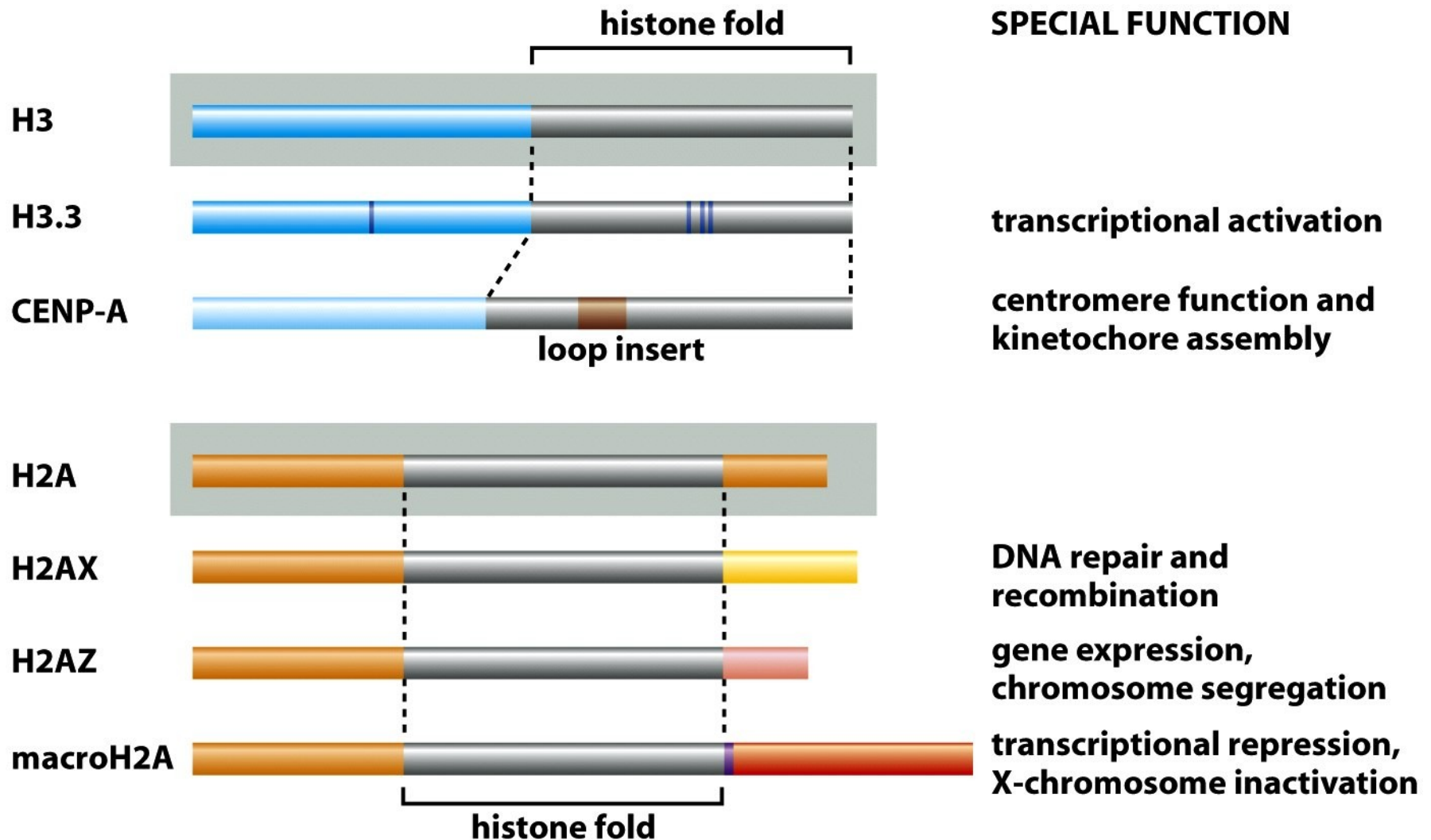


Figure 1-26 *Molecular Biology of the Cell*, Fifth Edition (© Garland Science 2008)

- následují **mutace** – modifikující původní protein na novou variantu (podobná funkce)

Příklad: histony - histon fold + odlišné sekvence (odlišně PTM) odlišné funkce (stále v rámci nukleosomů/chromatinu)



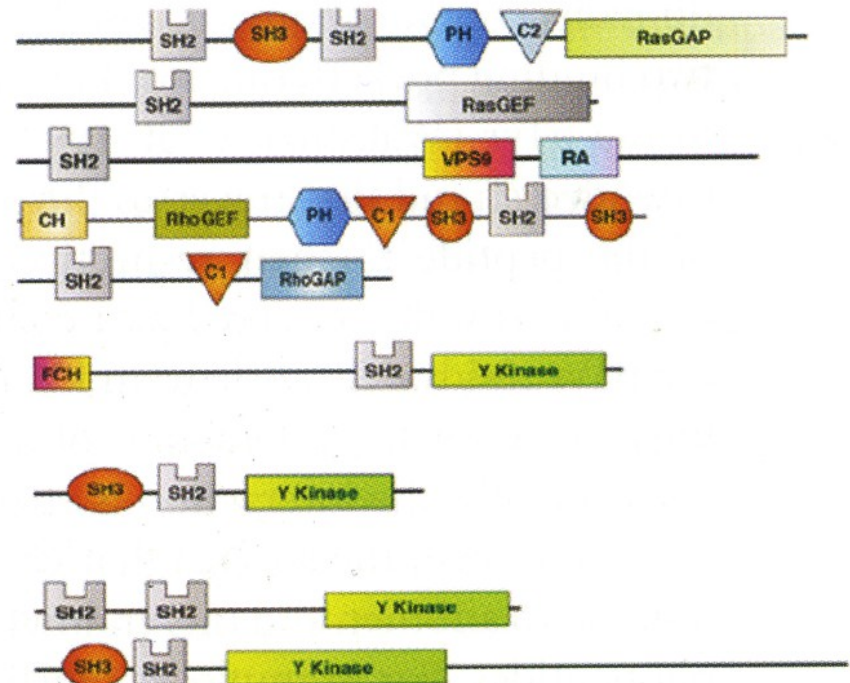
- duplikace celého genu nebo pouze **domény** – většina domén v proteinech obratlovců pochází/existuje v bezobratlých – 7% lidských proteinů/domén je specifických pro obratlovce)
- opakování domén (imunoglobuliny, kolageny apod.) nebo spojování různých domén
- fúze domén naznačuje funkční příbuznost nebo dokonce přítomnost v komplexu (v organismu s nefúzovanými doménami - ChimeraDB)

Small GTPase Signaling

Ras-GAP
Nsp1,2,3
Rin1
Vav1,2,3
Chimerin

Kinases

Fps, Fer
Src, Csk, Ctk/Hyl,
Fgr, Fyn, Yes, Hck,
Lck, Lyn, Blk, Frk,
Brk, DJ697K14.1
Zap70, Syk
c-Abl, Arg/Abl2



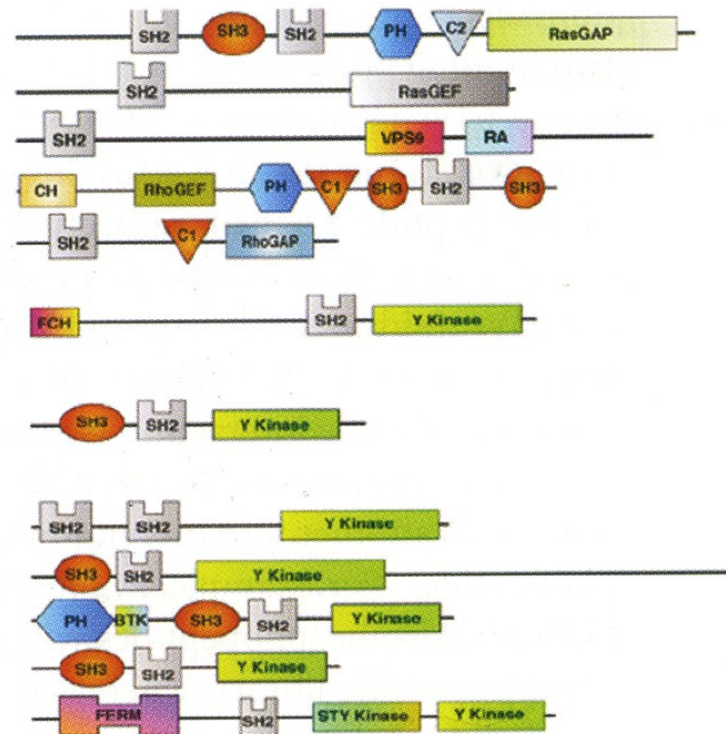
- duplikace **domény** (většinou koreluje s exonem)
- hranice domény jsou většinou kódovány **introny** – bez intronů by bylo obtížnější přesně vybrat pouze „doménovou“ část sekvence
- duplikací a „**shuffling**“ **domén** – poskládají se nové geny/proteiny – vytváří nová funkční/fyzická provázání (interakce => „networks“)

Small GTPase
Signaling

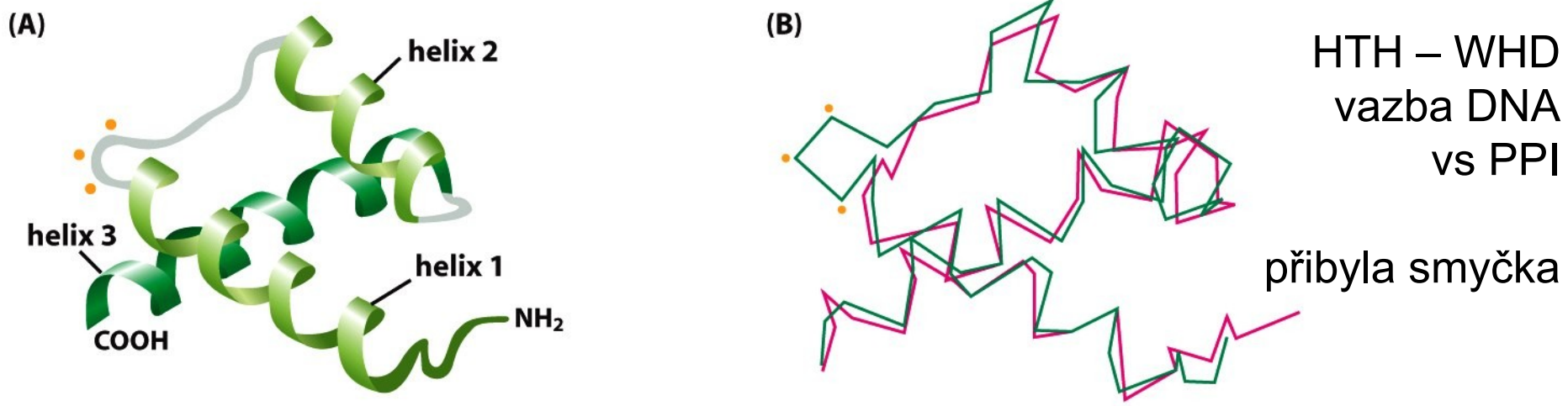
Ras-GAP
Nsp1,2,3
Rin1
Vav1,2,3
Chimerin

Kinases

Fps, Fer
Src, Csk, Ctk/Hyl,
Fgr, Fyn, Yes, Hck,
Lck, Lyn, Bik, Frk,
Brk, DJ697K14.1
Zap70, Syk
c-Abl, Arg/Abl2
Btk, Tec, Itk, Bmx
Txk
Jak1,2,3,Tyk2



- duplikací a „shuffling“ domén – poskládají se nové geny/proteiny
- nová doména (protein) není pod tlakem „funkčním“, ale zůstává pod tlakem „strukturním“ („misfold“ = degradace)



(C) **Hydrofobní – struktura, bazické – vazba DNA**

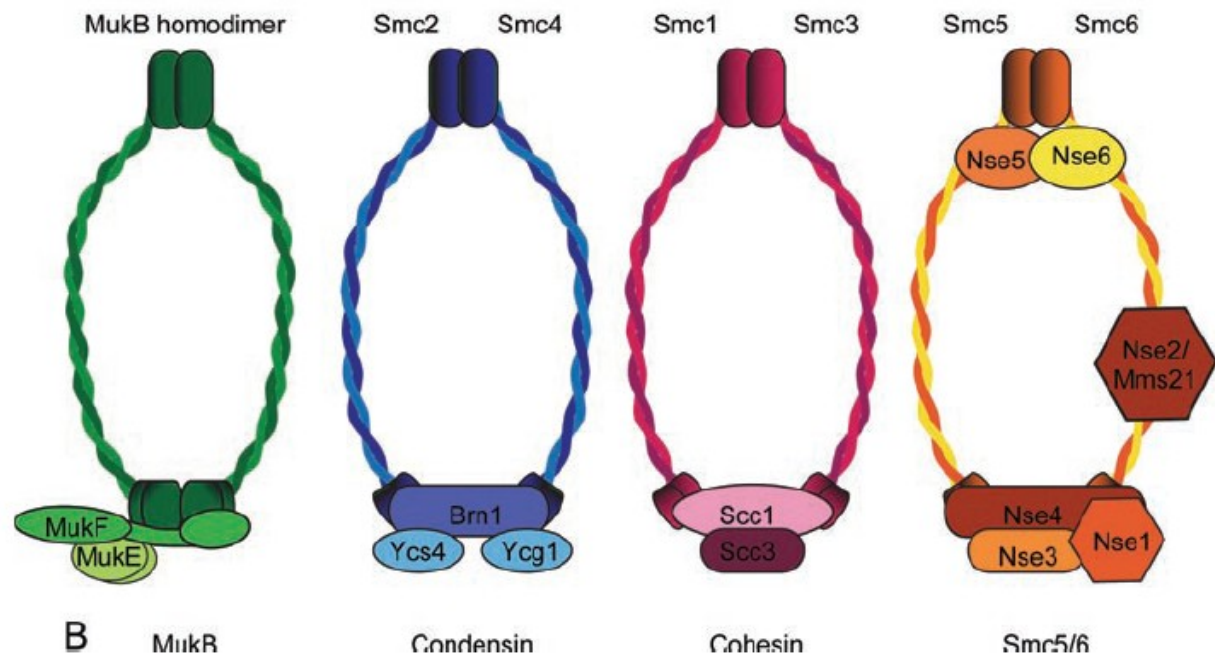
yeast **H₂N** GHRFTKENVRILESWFAKNIENPYLDTKGLENLMKNTSLSRRIQIKNWVSNRRRKEKTI **COOH**

Drosophila RTAFSSÉOLARLKREFNEN - - - RYLTERRRQQLSSELGLNEAQIKIWFQNKRAKIKKS

Figure 3-13 *Molecular Biology of the Cell* (© Garland Science 2008)

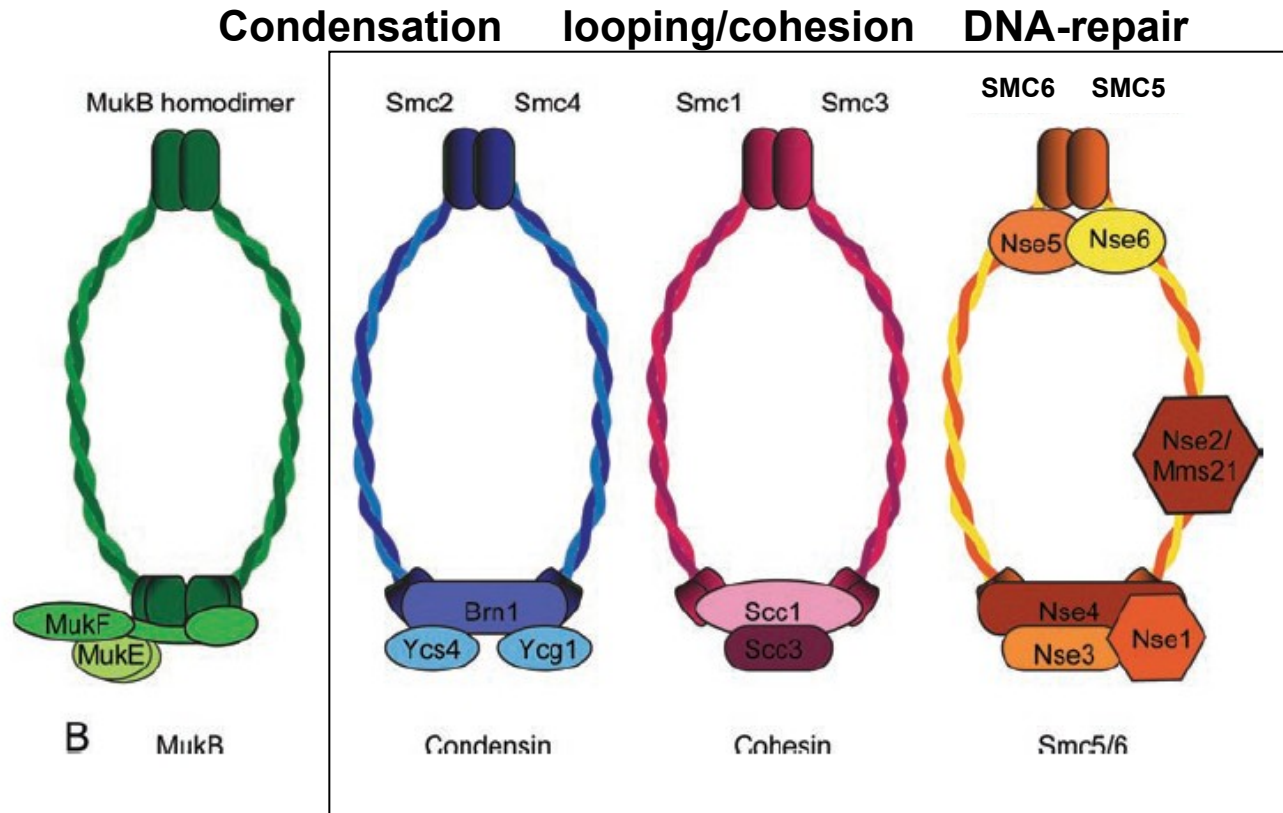
- zůstává hydrofobní profil (hydrofobní AMK jsou uvnitř a drží strukturu domény) – mění se povrch (můžete hledat „paralogy“ i podle 3D struktury, sekvence <25% ident. – funkce ? odlišná)

- „staré“ proteiny jsou konzervovány „strukturně“ a „funkčně“ – funkčně znamená většinou „konzervaci“ interakcí (většinou PPI)
- nová doména (protein) je pod tlakem „strukturním“ („misfold“ = degradace) – je ale i pod tlakem „povrchovým“ tj. tendencí povrchů interagovat (zvláště hydrofobní povrchy – hydrofobní povrch je rozeznáván chaperony a bez interakčního partnera degradován)
- z toho plyne tendence duplikovaných proteinů vytvářet podobné komplexy (podobné interakce)



Evoluce SMC komplexů

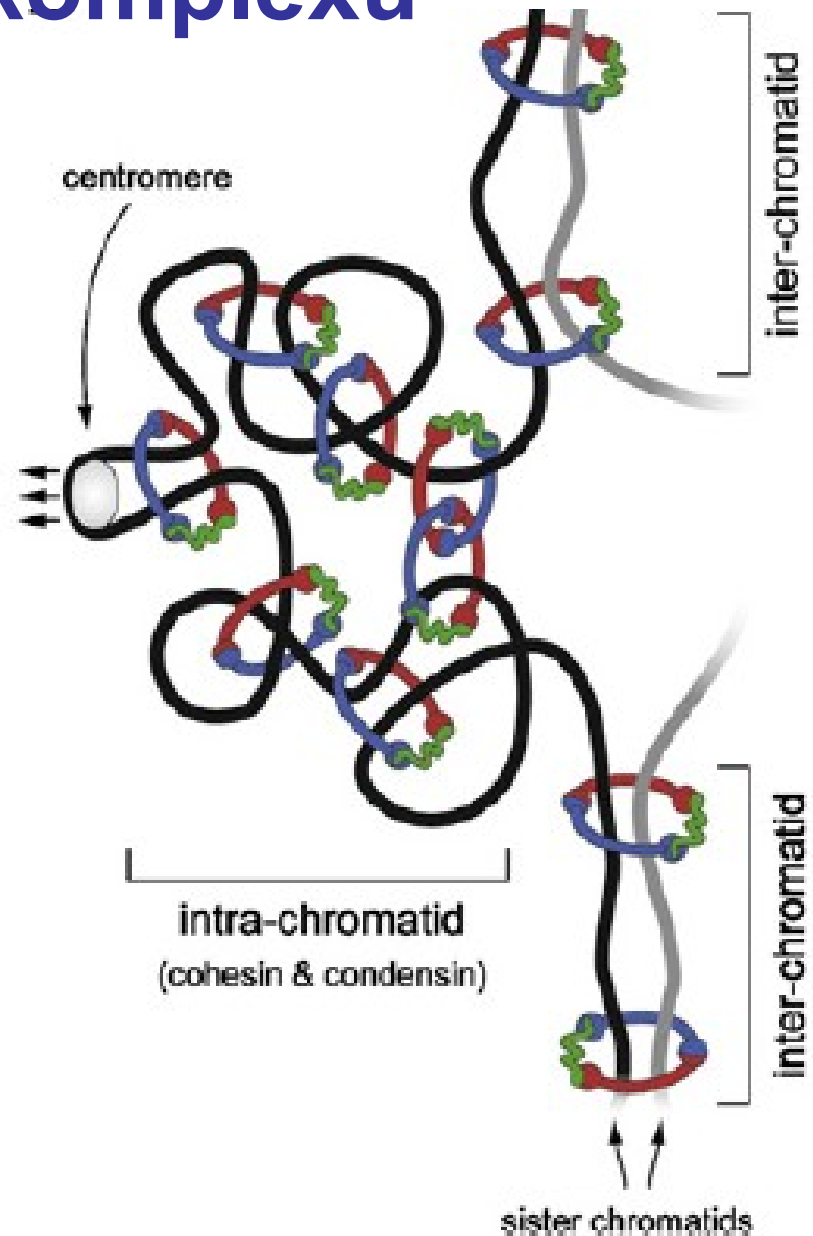
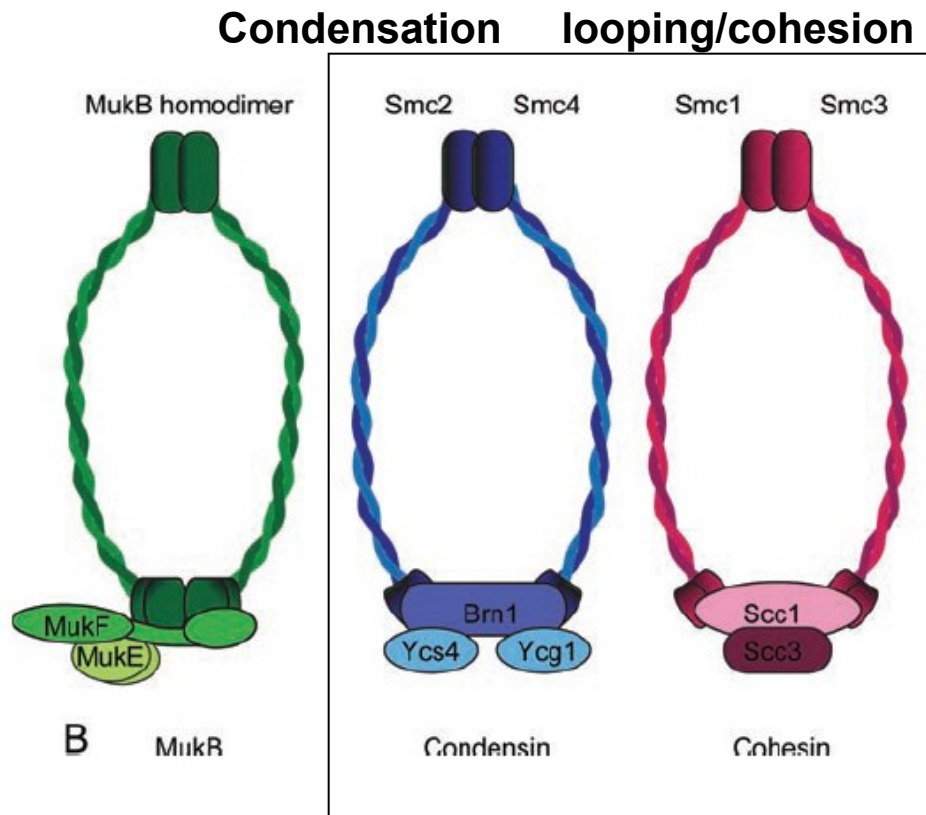
- bakterie mají po jednom komplexu složeném z homodimeru SMC/MukB a 2 Nse



- eukaryota mají 3 SMC komplexy – SMC heterodimery + kleisin (další podjednotky nepříliš konzervované) - (příklad využití konzervovaného motivu a alterace částí systému/komplexu)

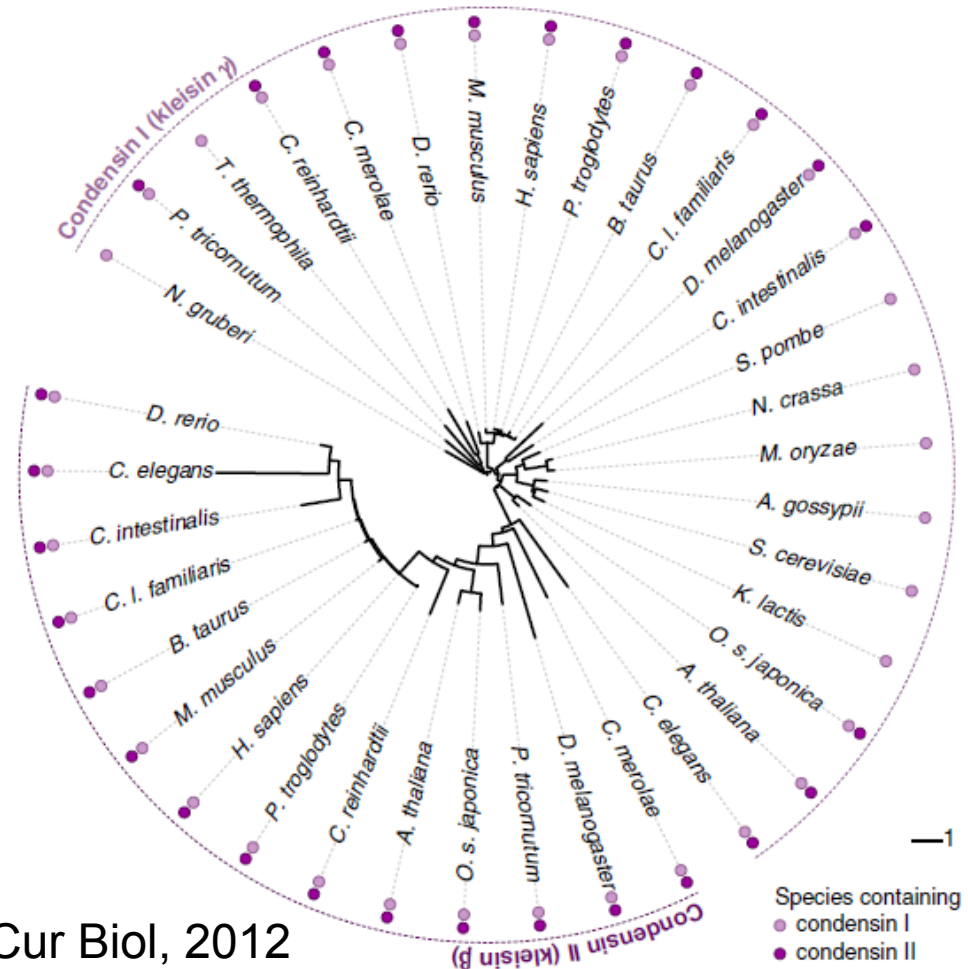
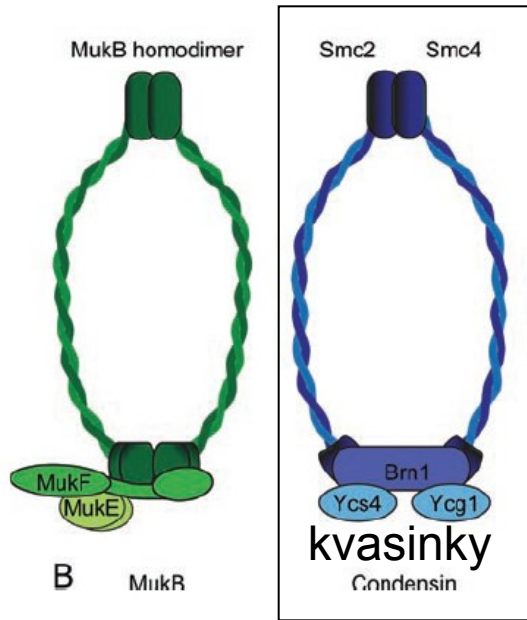
Evoluce SMC komplexů

- konzervovaný motiv (kroužek)
- SMC heterodimery + kleisin



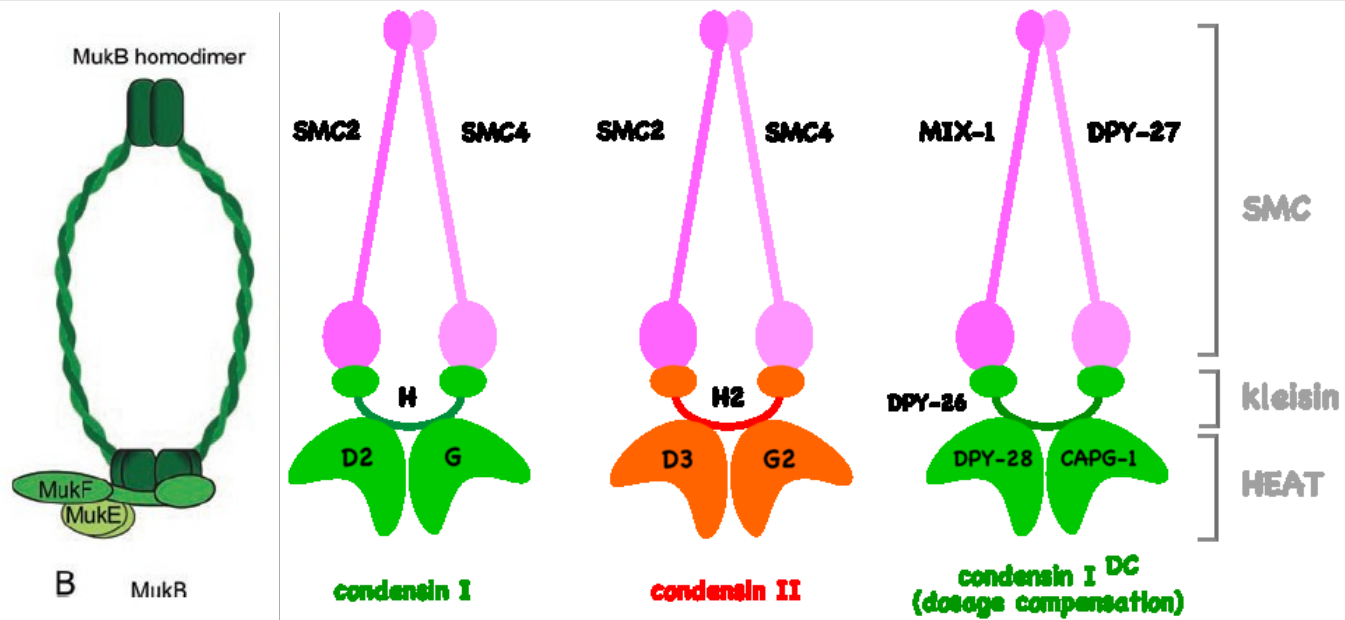
- alterace – kohese x kondensace
- (Nse podjednotky, faktory ovlivňující loading = místo, čas ...)

Name	<i>Saccharomyces cerevisiae</i>	<i>Schizosaccharomyces pombe</i>	<i>Drosophila</i>	Vertebrates
Smc1	Smc1	Psm1	DmSmc1	Smc1
Smc3	Smc3	Psm3	DmSmc3	Smc3
Scs1	Mcd1/Pds3	Rad21	DmRad21	Rad21
Scs3	Scs3	Psc3	DmSA	SA1 and SA2



- eukaryota mají 3 SMC kom
- kohesin = mitotický a meiot
- Scs1xRec8, navíc obratlovci
- kondensin (komplex I a II, I

Name	<i>Saccharomyces cerevisiae</i>	<i>Schizosaccharomyces pombe</i>	<i>Drosophila</i>	Vertebrates
Smc1	Smc1	Psm1	DmSmc1	Smc1
Smc3	Smc3	Psm3	DmSmc3	Smc3
Scs1	Mcd1/Pds3	Rad21	DmRad21	Rad21
Scs3	Scs3	Psc3	DmSA	SA1 and SA2



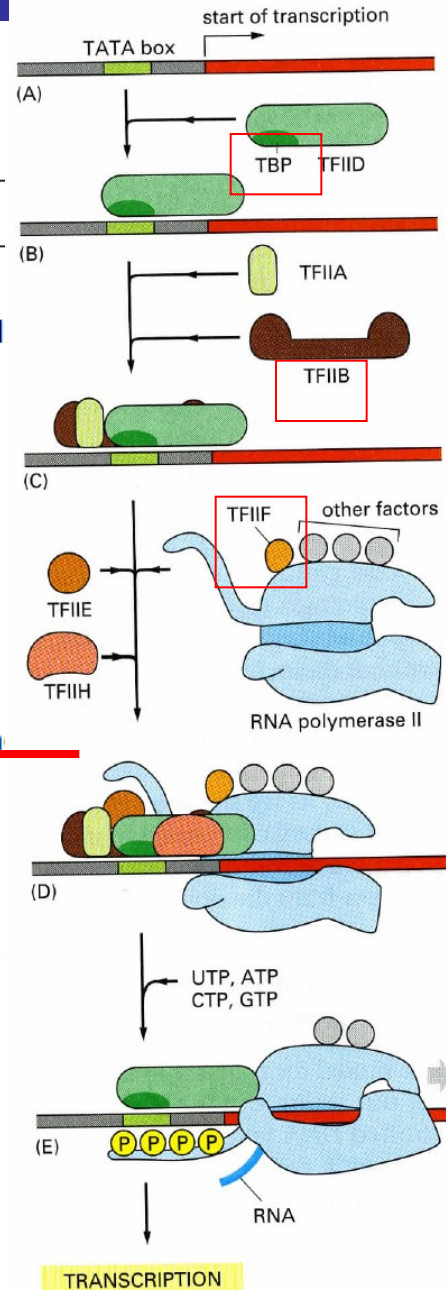
- eukaryota mají 3 SMC komplexy +
- kohesin = mitotický a meiotický (liší se kleisinem = kvasinky Scc1xRec8, navíc obratlovci Scc3 2x = SA1 a SA2)
- kondensin (I = H+D2+G; II = H2+D3+G2)
- SMC5/6 komplex (lidský = NSE4a x NSE4b-testis/meiosa specific)

RNA pol II a podobnosti s pol I a III

TABLE 1 Components of the human general transcription machinery

Factor	Protein composition	Function
TFIIA	p35 (α), p19 (β), and p12 (γ)	Antirepressor; stabilizes TBP-TATA complex; coactivator
TFIIB	p33	Start site selection; stabilize TBP-TATA complex; pol II/TFI
TFIID	TBP + TAFs (TAF1-TAF14)	Core promoter-binding factor Coactivator Protein kinase Ubiquitin-activating/conjugating activity Histone acetyltransferase
TFIIE	p56 (α) and p34 (β)	Recruits TFIIH Facilitates formation of an initiation-competent pol II Involved in promoter clearance
TFIIF	RAP30 and RAP74	<u>Binds pol II and facilitates pol II recruitment to the prom</u> Recruits TFIIE and TFIIH Functions with TFIIB and pol II in start site selection Facilitates pol II promoter escape Enhances the efficiency of pol II elongation
TFIIH	P89/XPB, p80/XPD, p62, p52, p44, p40/CDK7, p38/Cyclin H, p34, p32/MAT1, and p8/TFB5	ATPase activity for transcription initiation and promoter Helicase activity for promoter opening Transcription-coupled nucleotide excision repair Kinase activity for phosphorylating pol II CTD E3 ubiquitin ligase activity
pol II	RPB1-RPB12	<u>Transcription initiation, elongation, termination</u> Recruitment of mRNA capping enzymes Transcription-coupled recruitment of splicing and 3' end CTD phosphorylation, glycosylation, and ubiquitination

TBP-všetchny 3 pol



RNA polymerasy

Pol I = 12 + A49/34.5

Pol II = 12 podjednotek (Tabulka)

Pol III = 12 + C37/53 + C82/34/31

TBP
TFIIB

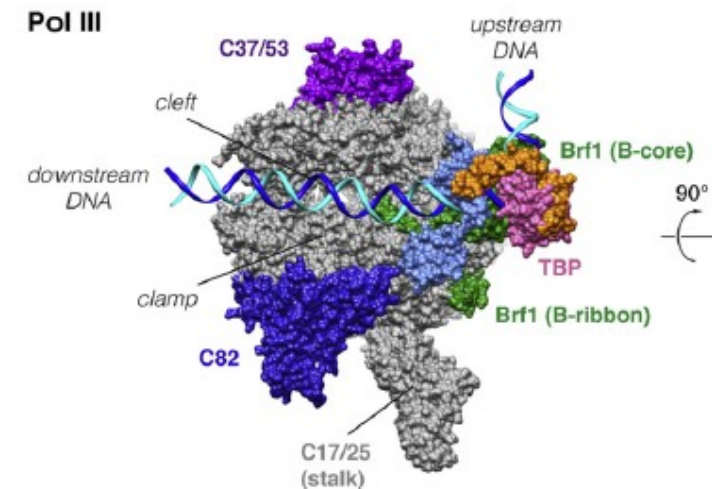
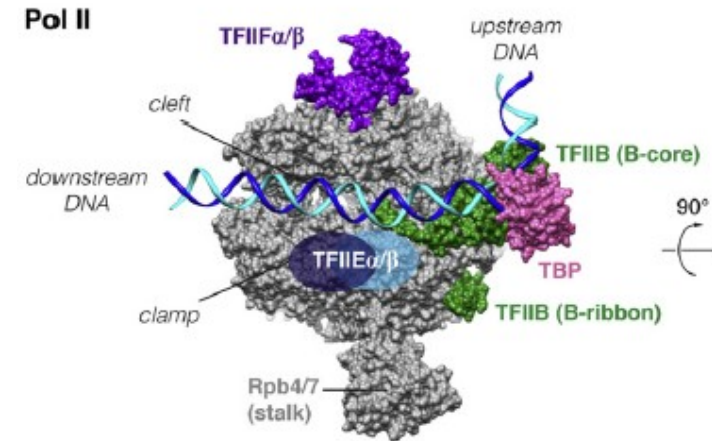
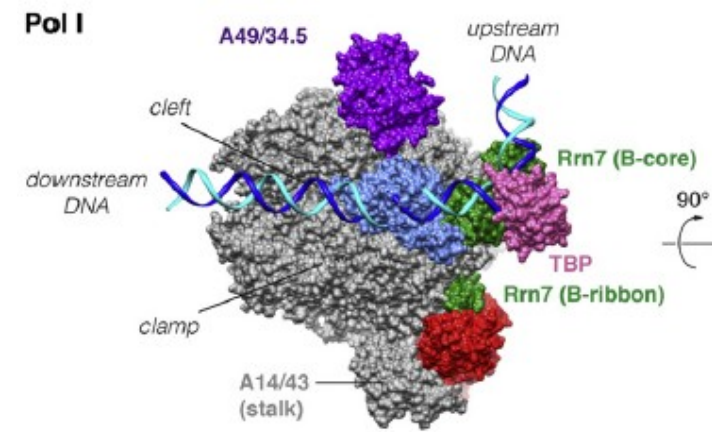
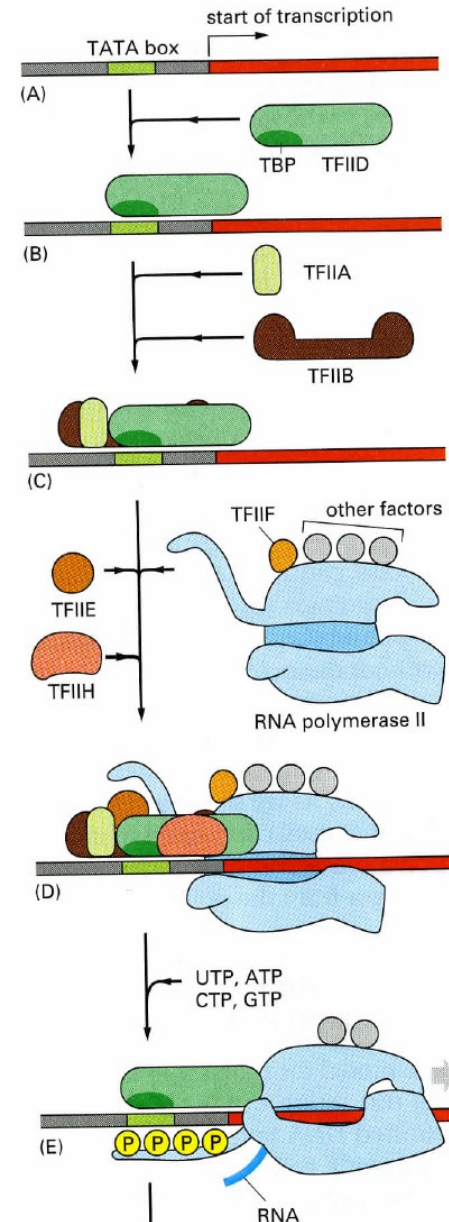
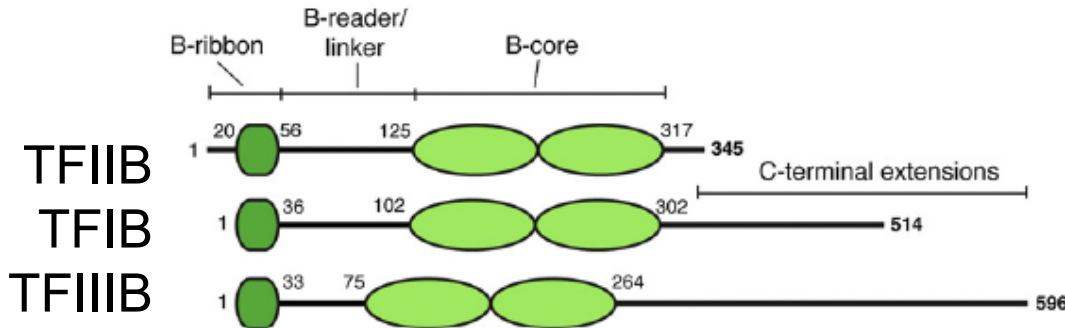
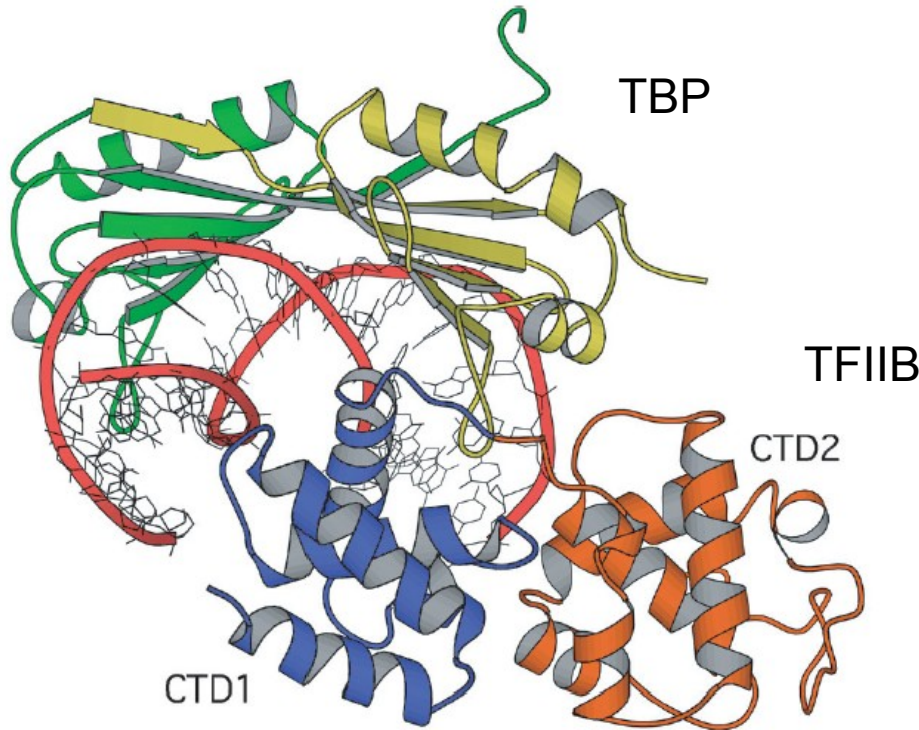
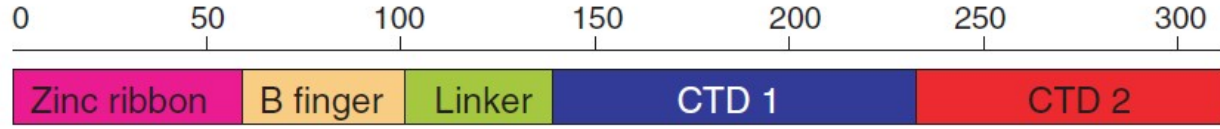
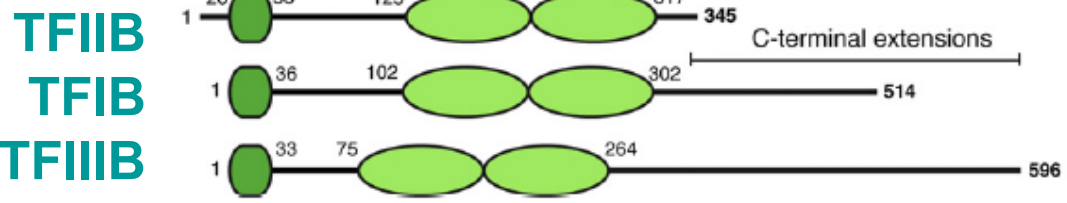
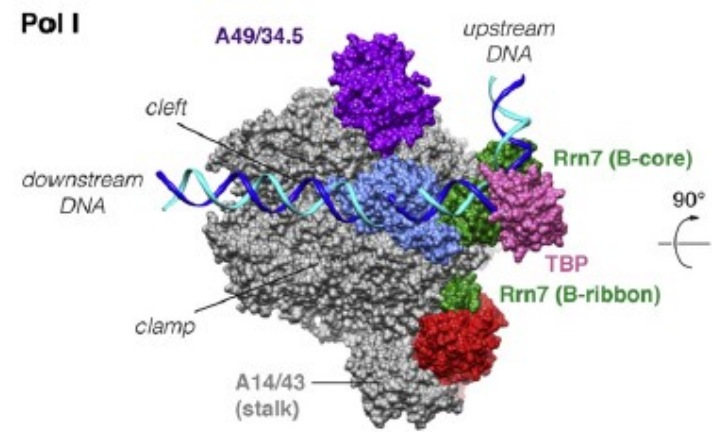


Table 1. Yeast RNA Polymerase Subunits and Initiation Factor Homologies

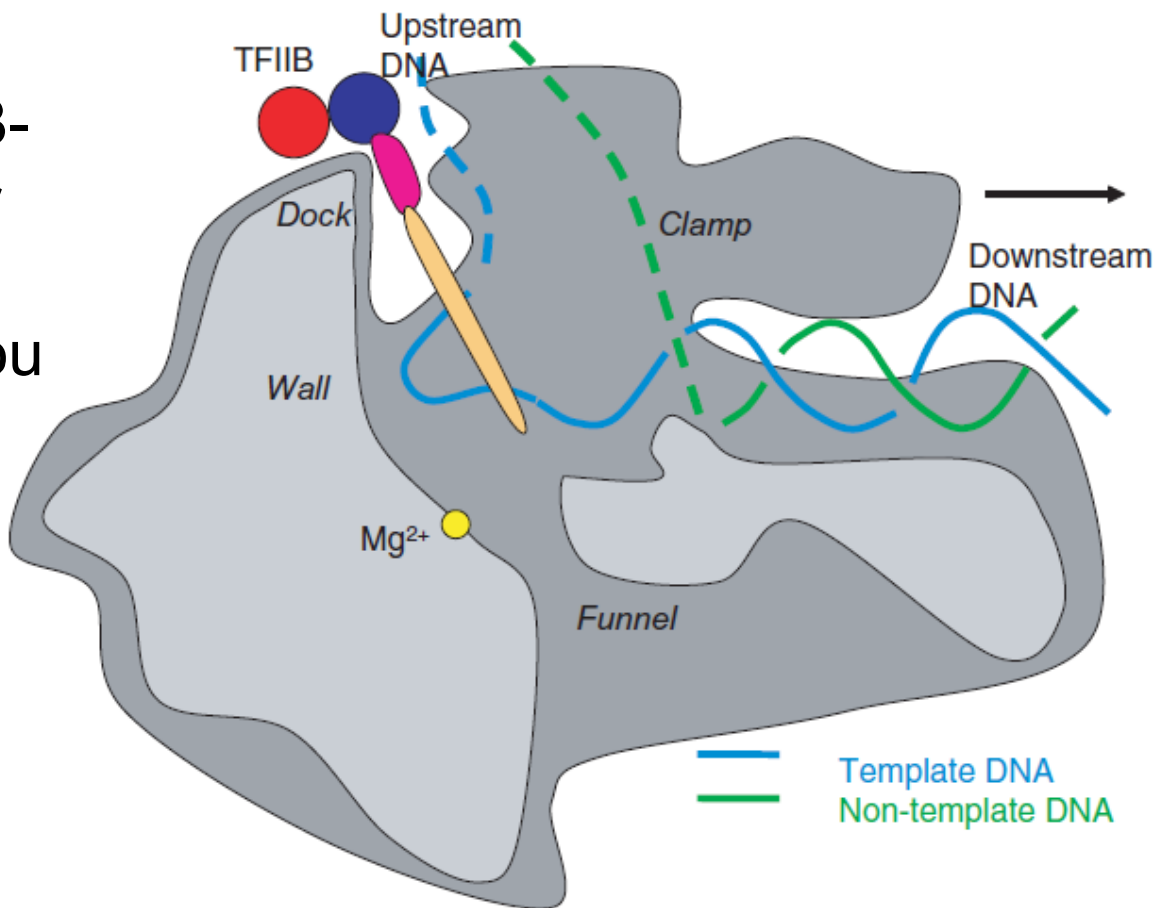
Pol II	Pol I	Pol III	Function
Polymerase Core			
Rpb1	A190	C160	Active center
Rpb2	A135	C128	Active center
Rpb3	AC40	AC40	
Rpb11	AC19	AC19	
Rpb9	A12.2 N ribbon	C11 N ribbon	RNA cleavage
TFIIS C-ribbon ^a	A12.2 C ribbon	C11 C ribbon	RNA cleavage
Rpb5	Rpb5	Rpb5	
Rpb6	Rpb6	Rpb6	
Rpb8	Rpb8	Rpb8	
Rpb10	Rpb10	Rpb10	
Rpb12	Rpb12	Rpb12	
Polymerase Stalk			
Rpb4	A14	C17	Initiation complex formation
Rpb7	A43	C25	Initiation complex formation

- TFIIB (C-konec) váže TBP a 6-7bp up- a downstream od ohybu DNA přes cukrfosfátovou kostru (kolem TATA boxu)

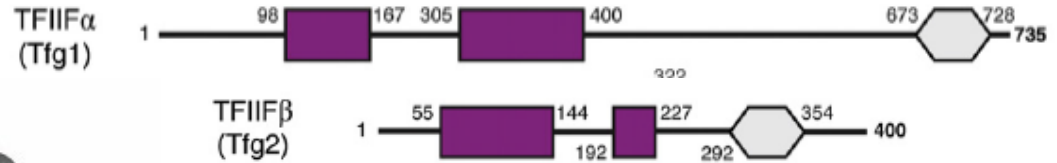
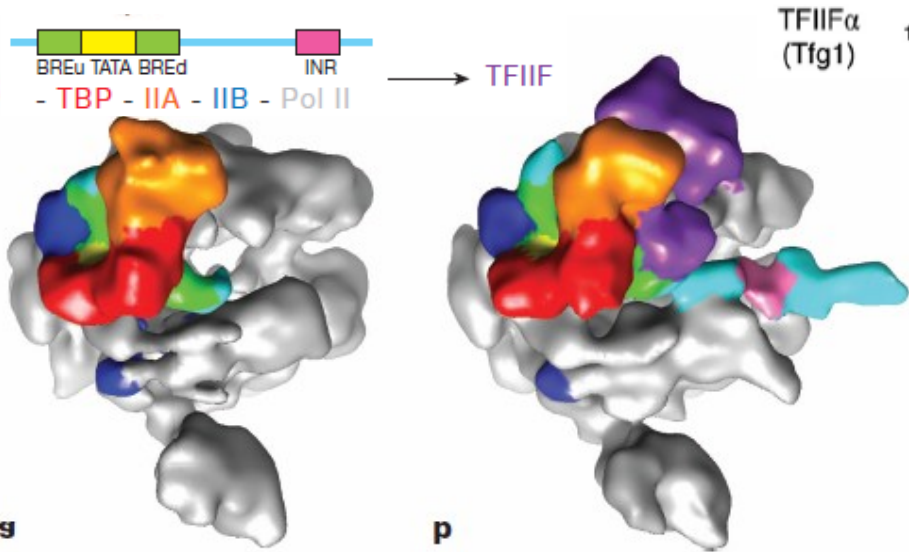




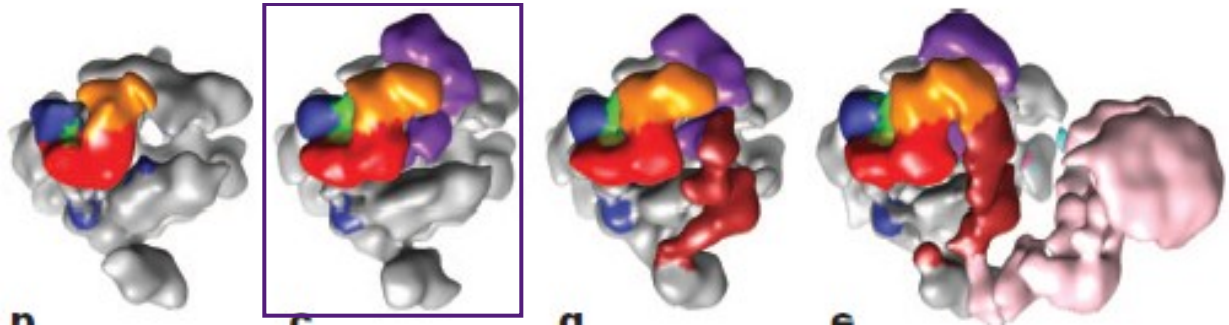
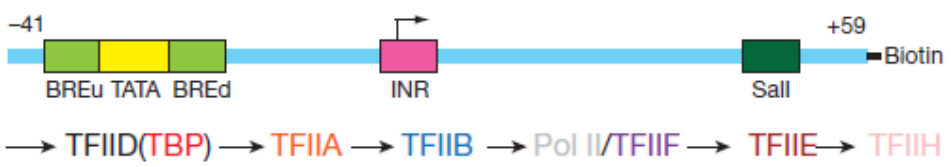
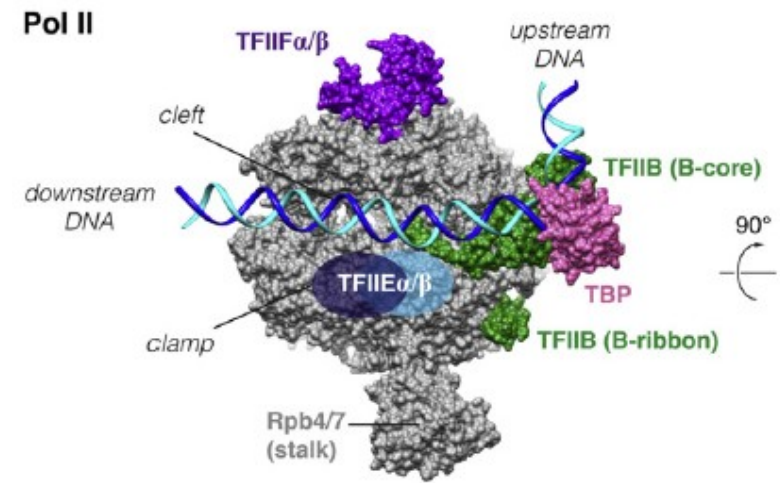
- Zn-ribbon s dock, B-finger/reader RNA tunel, B-core/CTD1 s wall, B-linker s CC clamp
- podobné natolik, že vážou TBP + ribbon a linker mohou být zaměněny pro kteroukoli polymerásu – zatímco reader je specifický pro určitou polymerásu



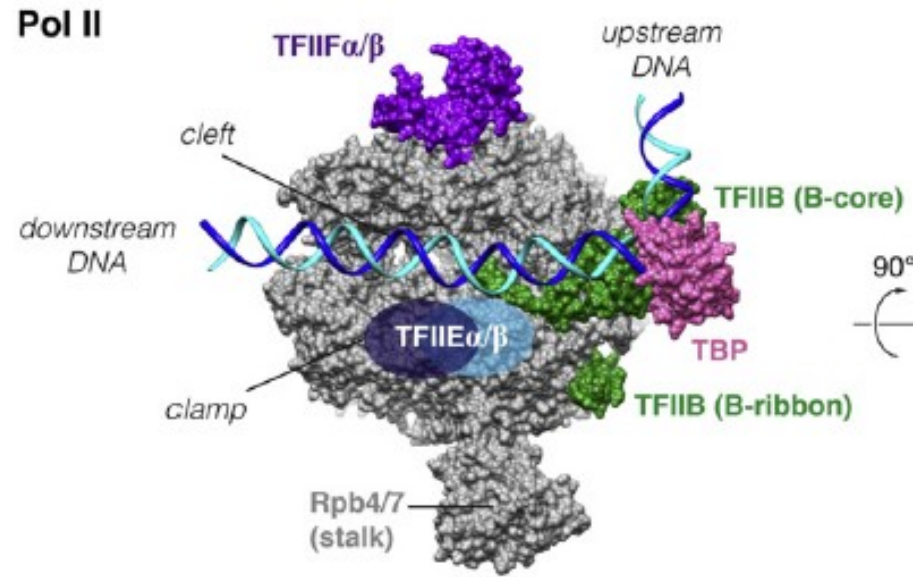
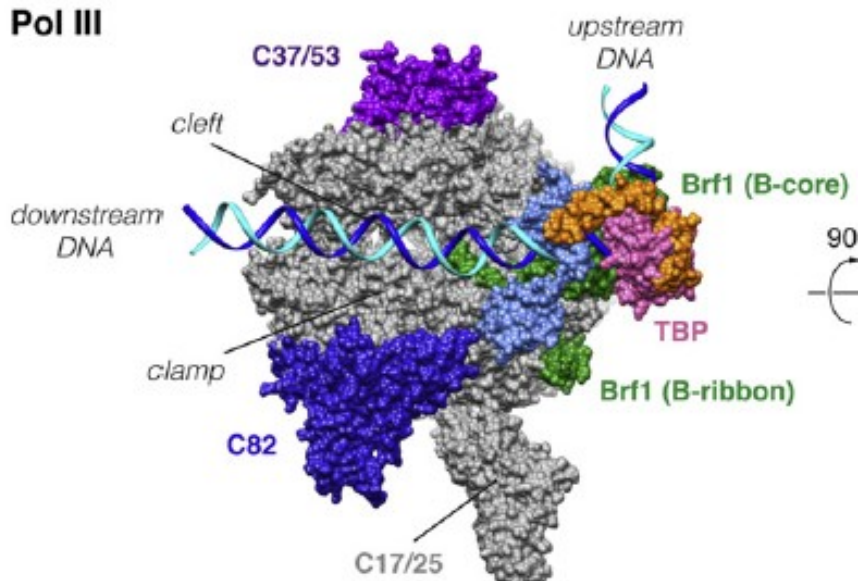
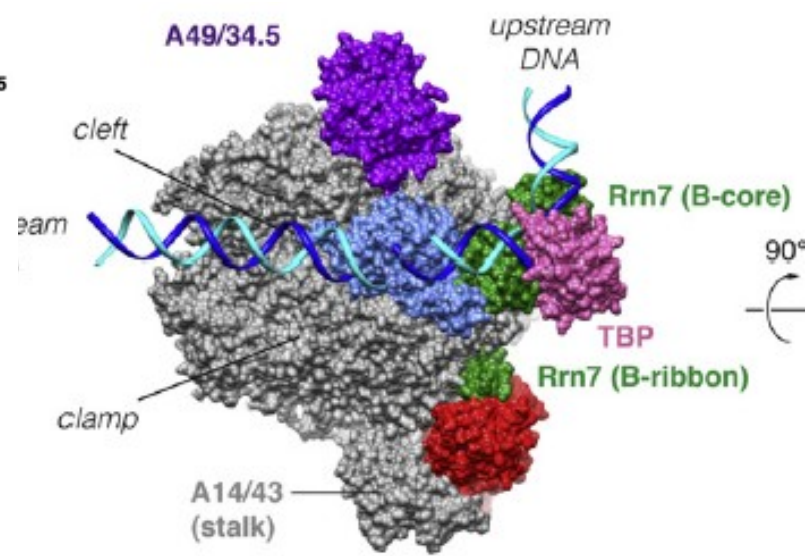
- **TFIIF** se váže na pol II, stabilizuje DNA v prohlubni/cleft RNA pol II a pomáhá TFIIIB s nastavením startu
- dimerizace a WHD (váže přímo DNA: BRE_{downstream})

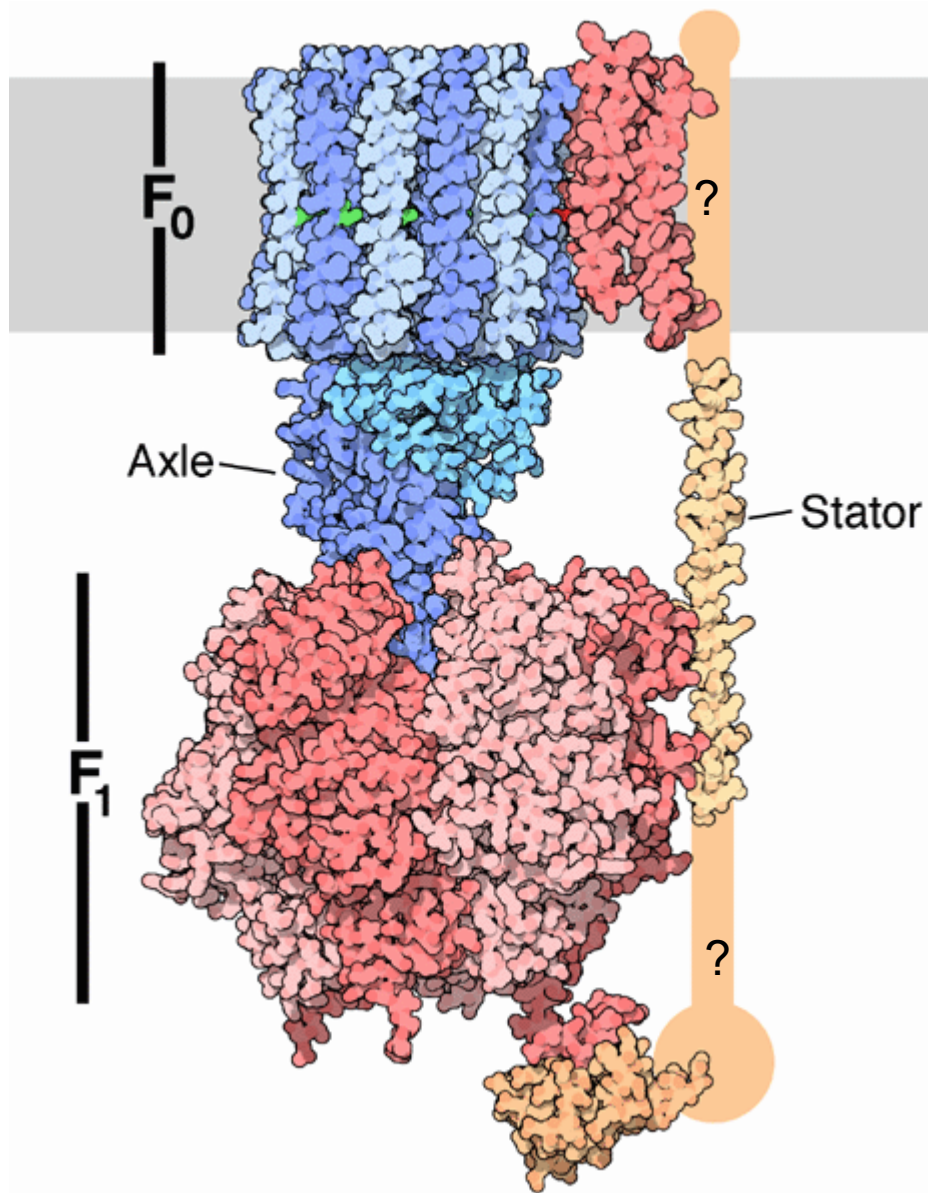


Dimerizace WHD

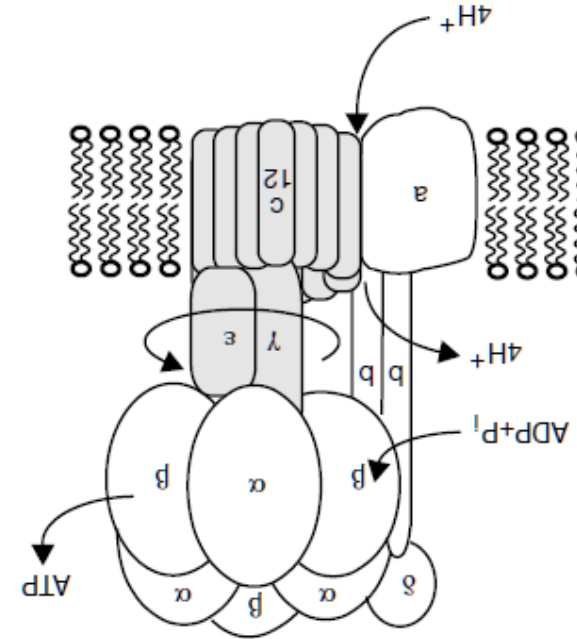


- **TFIIF** se váže na pol II, stabilizuje DNA v prohlubni/cleft RNA pol II a pomáhá TFIIB s nastavením startu
- dimerizace a WHD (váže přímo DNA: BRE_{downstream})



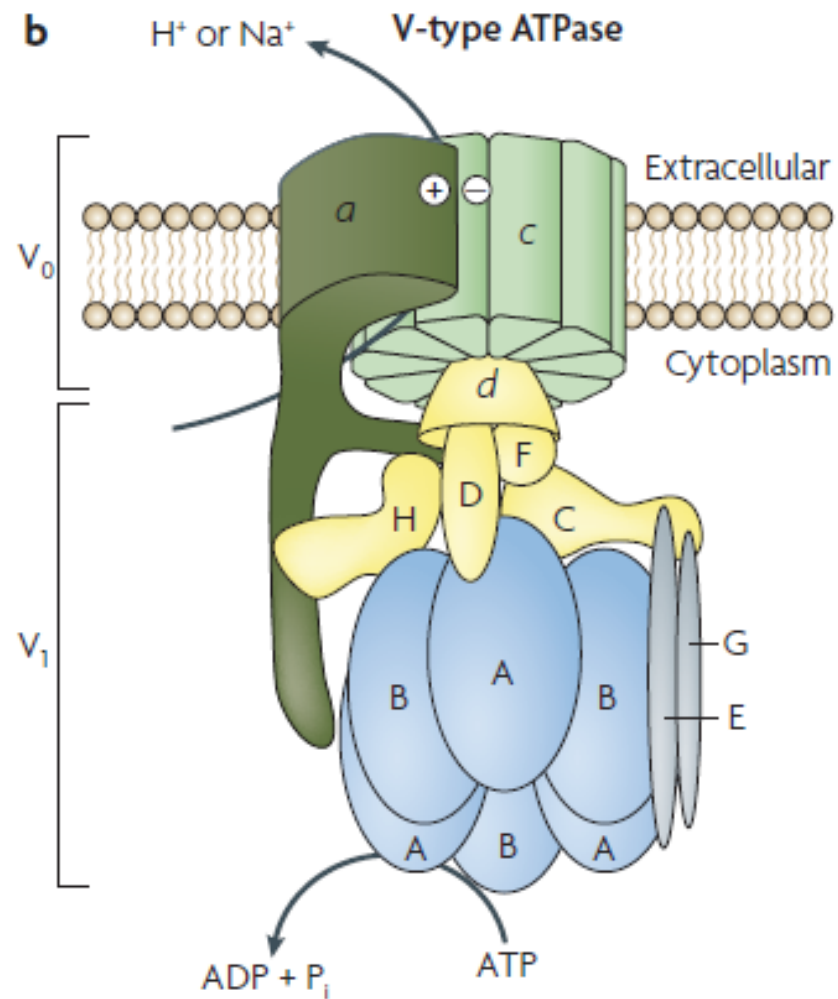
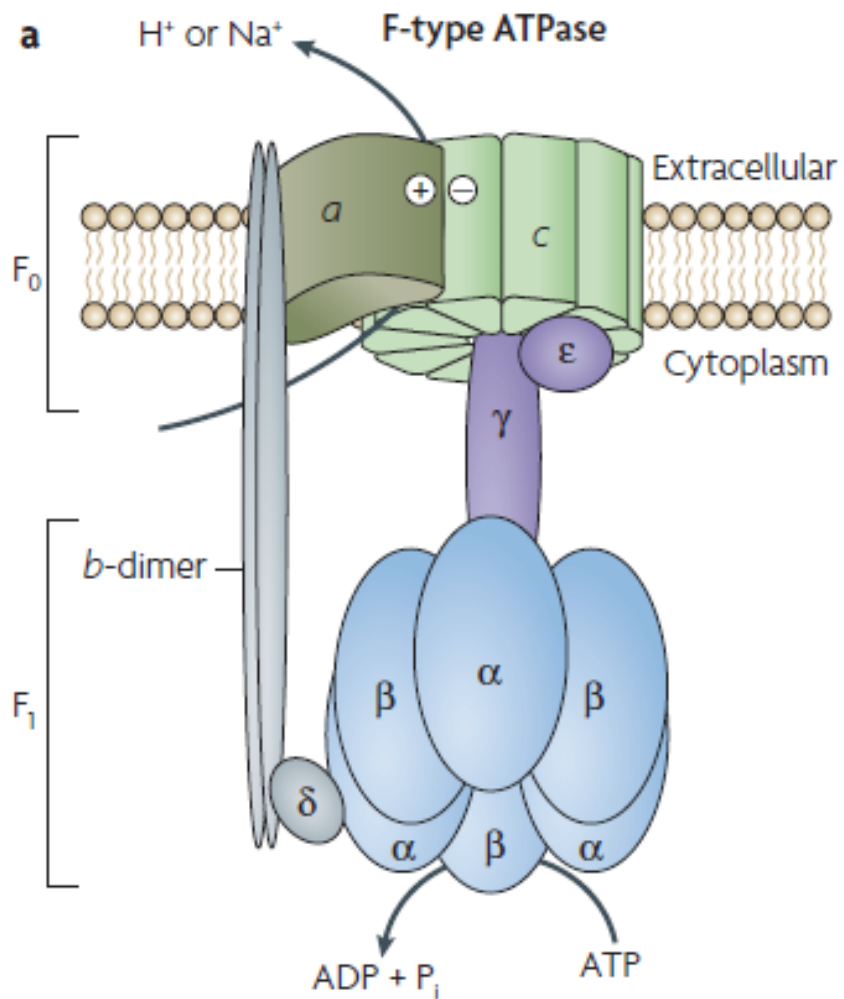


Molekula měsíce v prosinci 2005
 Rastogi & Girvin, Nature, 1999



F₀ je protonový motor (uložen v membráně) poháněný tokem vodíkových iontů (z dýchacího řetězce) přes membránu. Tento rotor je spojen s druhým F₁ chemickým motorem poháněným ATP (nebo vyrábějícím ATP). Oba motory jsou spojeny státorem.

„ATP syntasa je jedním z divů molekulárního světa“. Je to dvojitý molekulární motor – „nanostroj“ – vyrábějící většinu ATP (energie).

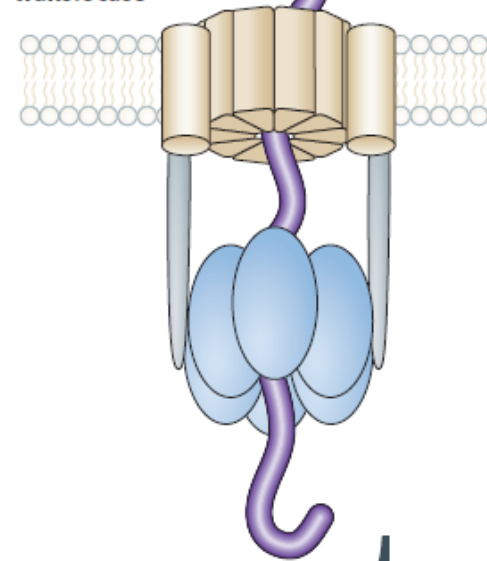
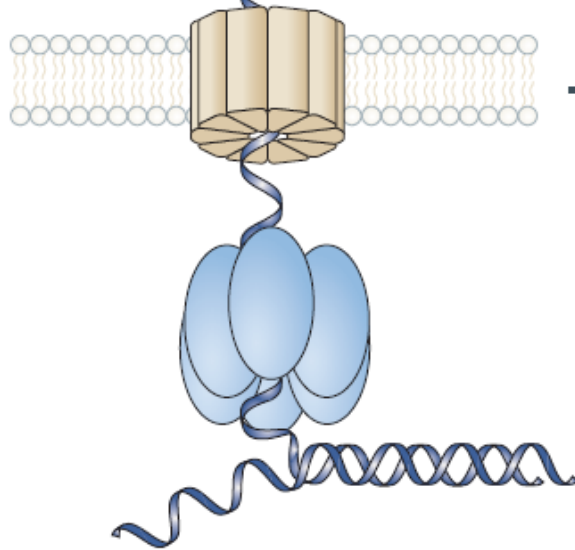
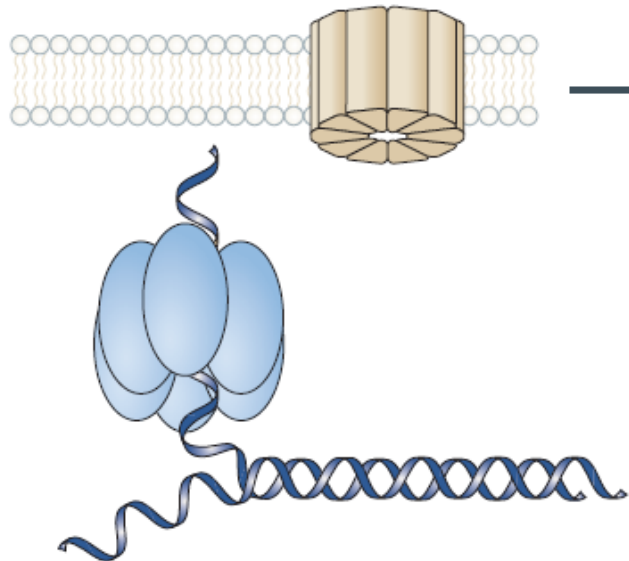


RNA helicase

Membrane channel

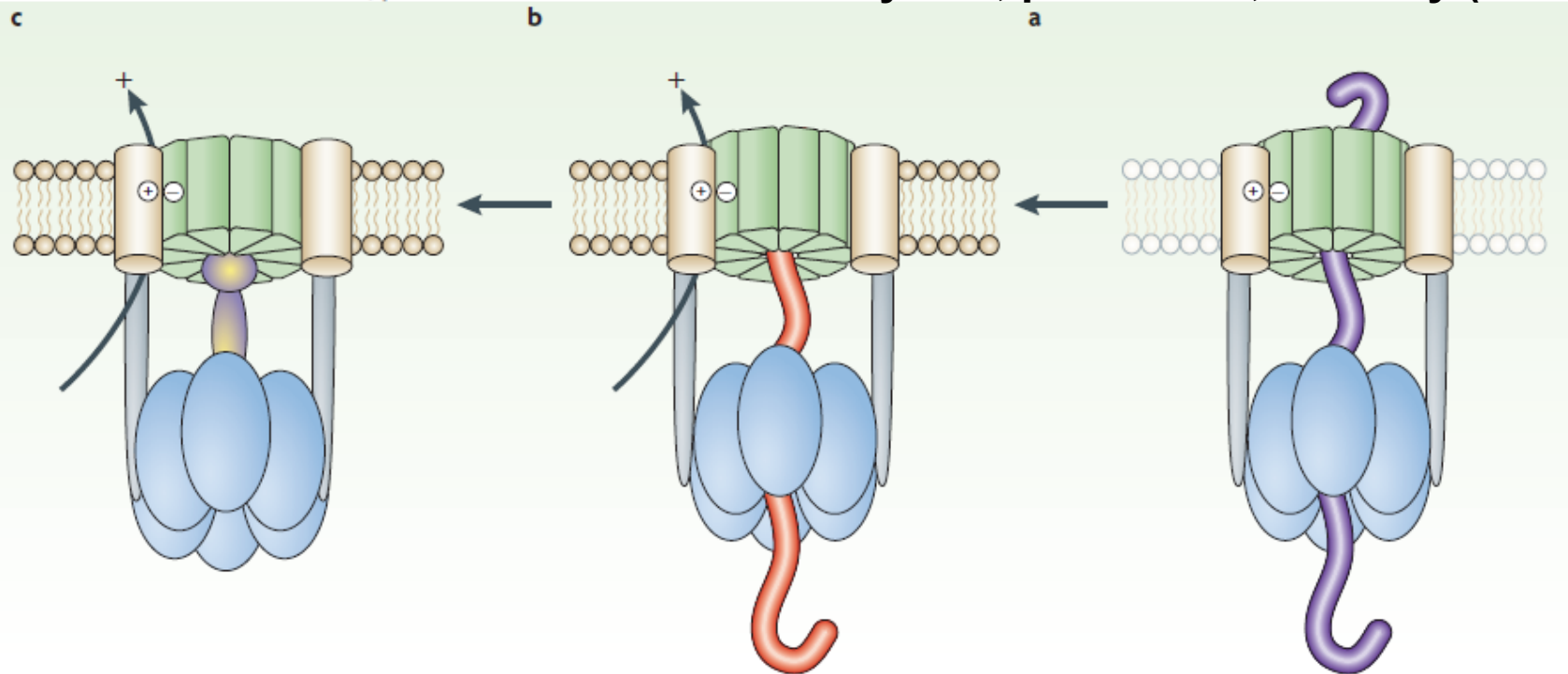
RNA translocase

Protein translocase



Common ancestor of F- and V-type ATPases

AAA+ - dynein, proteasom, helikasy (i PCNA)



I

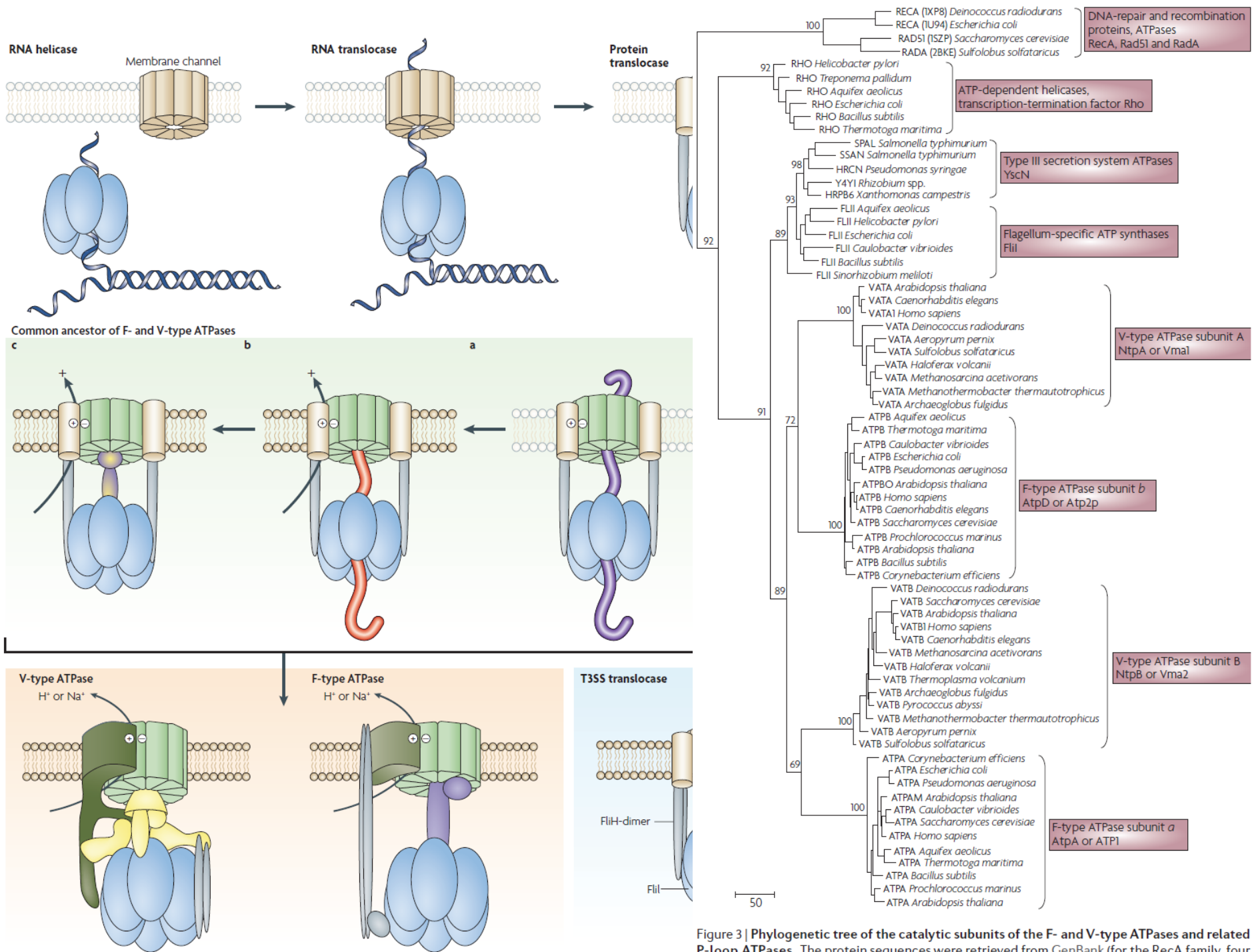


Figure 3 | Phylogenetic tree of the catalytic subunits of the F- and V-type ATPases and related P-loop ATPases. The protein sequences were retrieved from GenBank (for the RecA family, four

Souhrn

Průběh evoluce je ovlivněn mnoha faktory

- Selekcční tlaky udržující funkční a stabilní (dříve vytvořené) komplexy
- Mutace modifikující (především duplikované) proteiny/komplexy
- Neofunkcionalizace je hnací silou vzniku nových komplexů – funkcí – buněk - organismů