



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

TENTO PROJEKT JE SPOLUFINANCOVÁN EVROPSKÝM SOCIÁLNÍM FONDĚM
A STÁTNÍM ROZPOČTEM ČESKÉ REPUBLIKY

anketa

- **Proteinové interakce – 22.10.**
 - jak spolu proteiny interagují?
 - interaktom
- **Proteinové komplexy – 31.10.**
 - protein-proteinové interakce a komplexy
 - komplexom, architektura a funkce komplexů

doc. Jan Paleček
jpalecek@sci.muni.cz

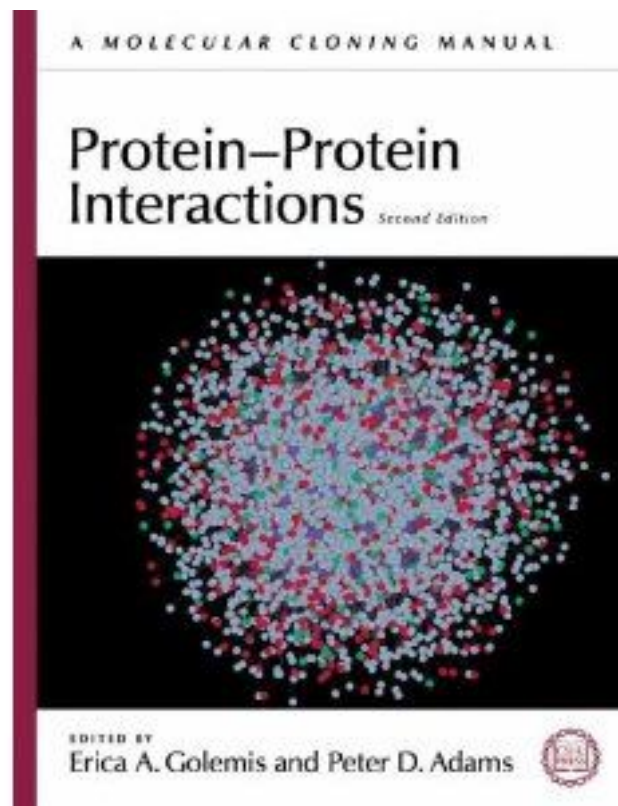
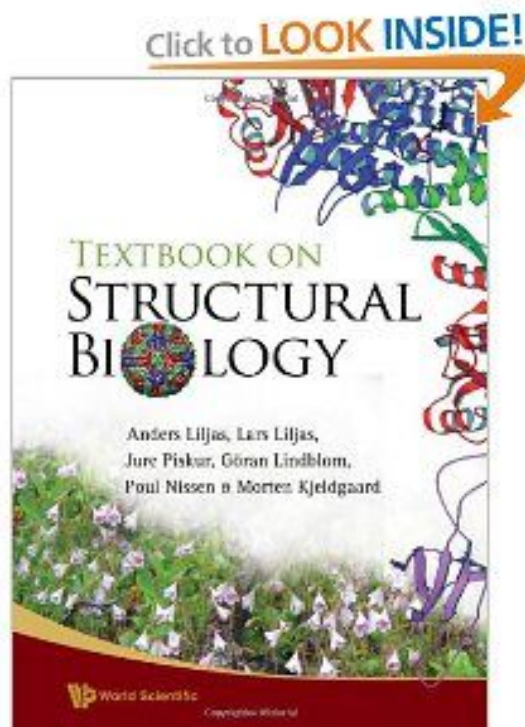
CG030 – Struktura a funkce proteinových komplexů
CG031 – cvičení z modelování proteinových komplexů
(jarní semestr)

Informační zdroje

Alberts a spol: Molecular biology of the Cell (2008 ...)

Liljas a spol: Structural biology (2009) ...

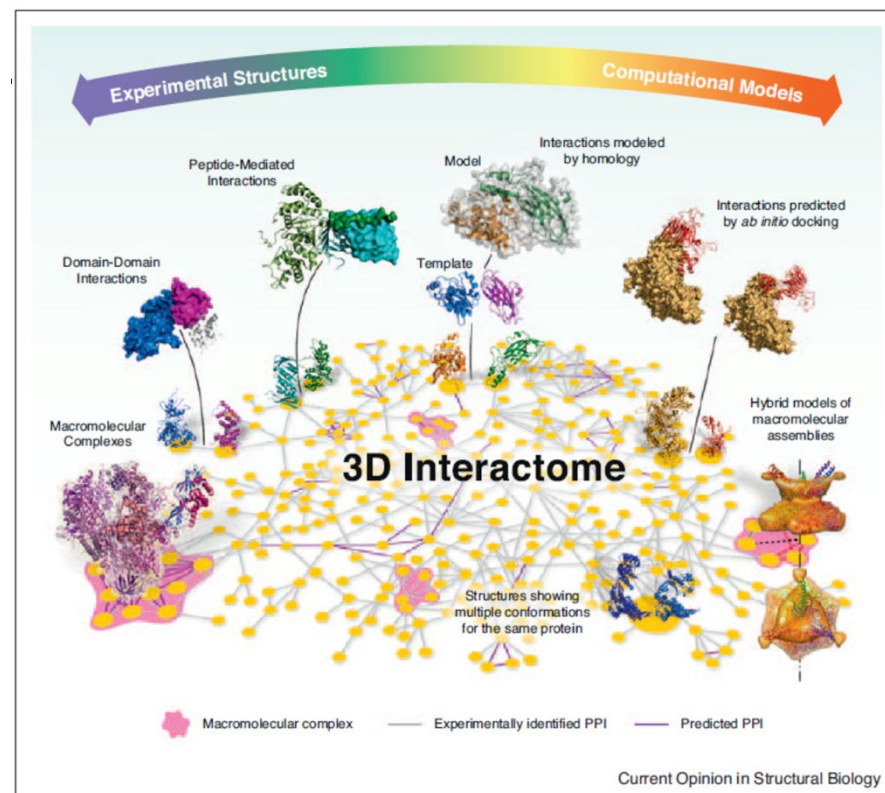
... nejnovější články z časopisů Cell, Nature, Science, PLoS ...



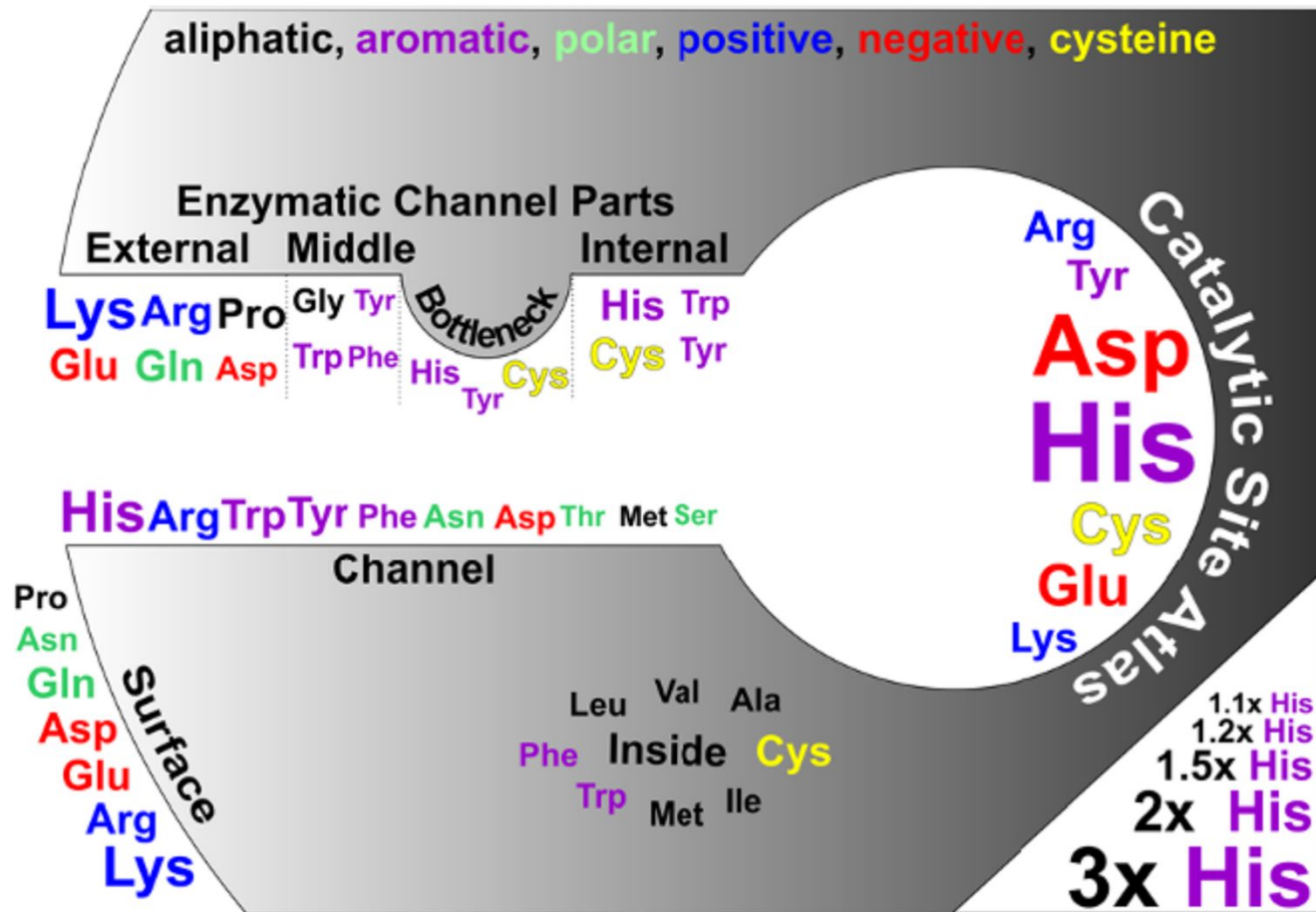
Databáze proteinových struktur: <http://www.rcsb.org/pdb/home/home.do>,
<http://www.ebi.ac.uk/pdbsum/>

Database protein-proteinových interakcí: <http://string-db.org/newstring.cgi> ...
<http://www.ebi.ac.uk/intact/?conversationContext=1>

- Proteinové interakce – 22.10.
 - Interakce: od primární po terciární strukturu
 - Typy vazeb: pohled chemický, geometrický ...
 - Informatika:
 - databáze struktur, interakcí ...
 - docking ...
 - motivy, evoluční aspekty ...
 - nástroje ...
 - interaktom ...

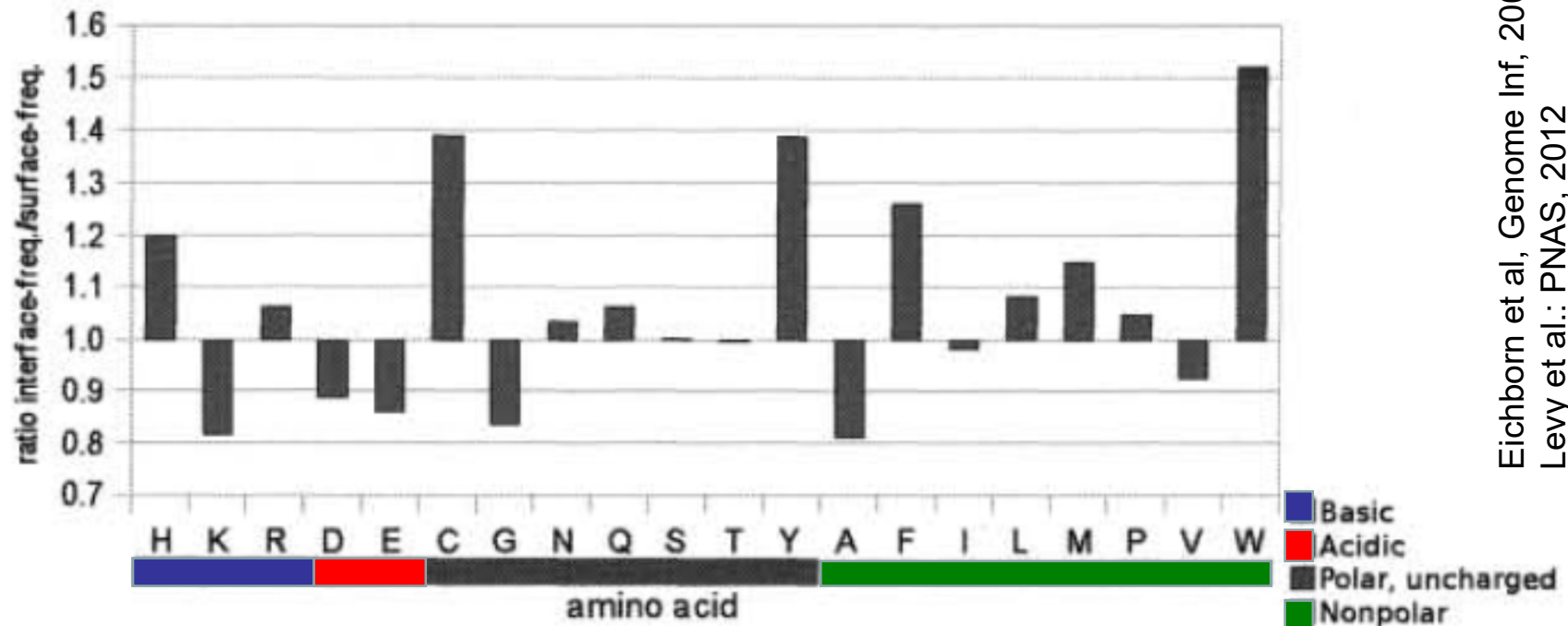


Podíl AMK (primární struktury) na proteinových interakcích

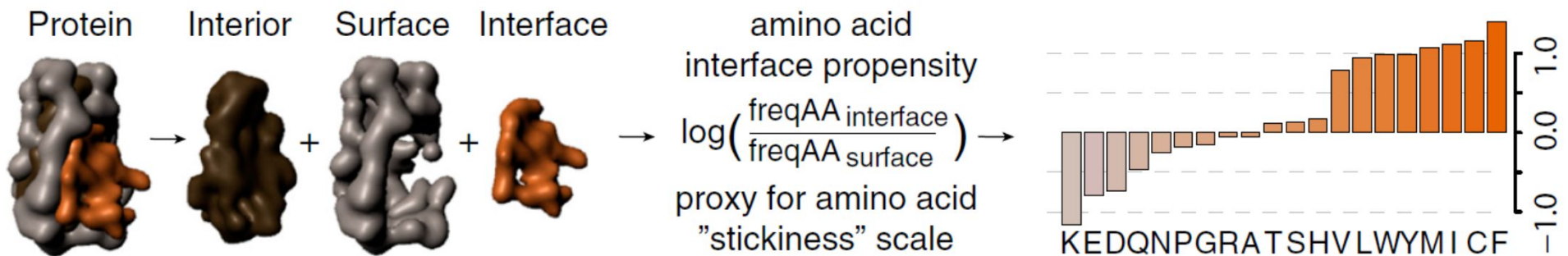


- uvnitř hydrofobní, povrch polární/nabitý (do solventu/vody),
 ale katalytická centra (tunely) jsou také polární a nabité
 (katalýza biochemické reakce)

PPI od primární struktury ...

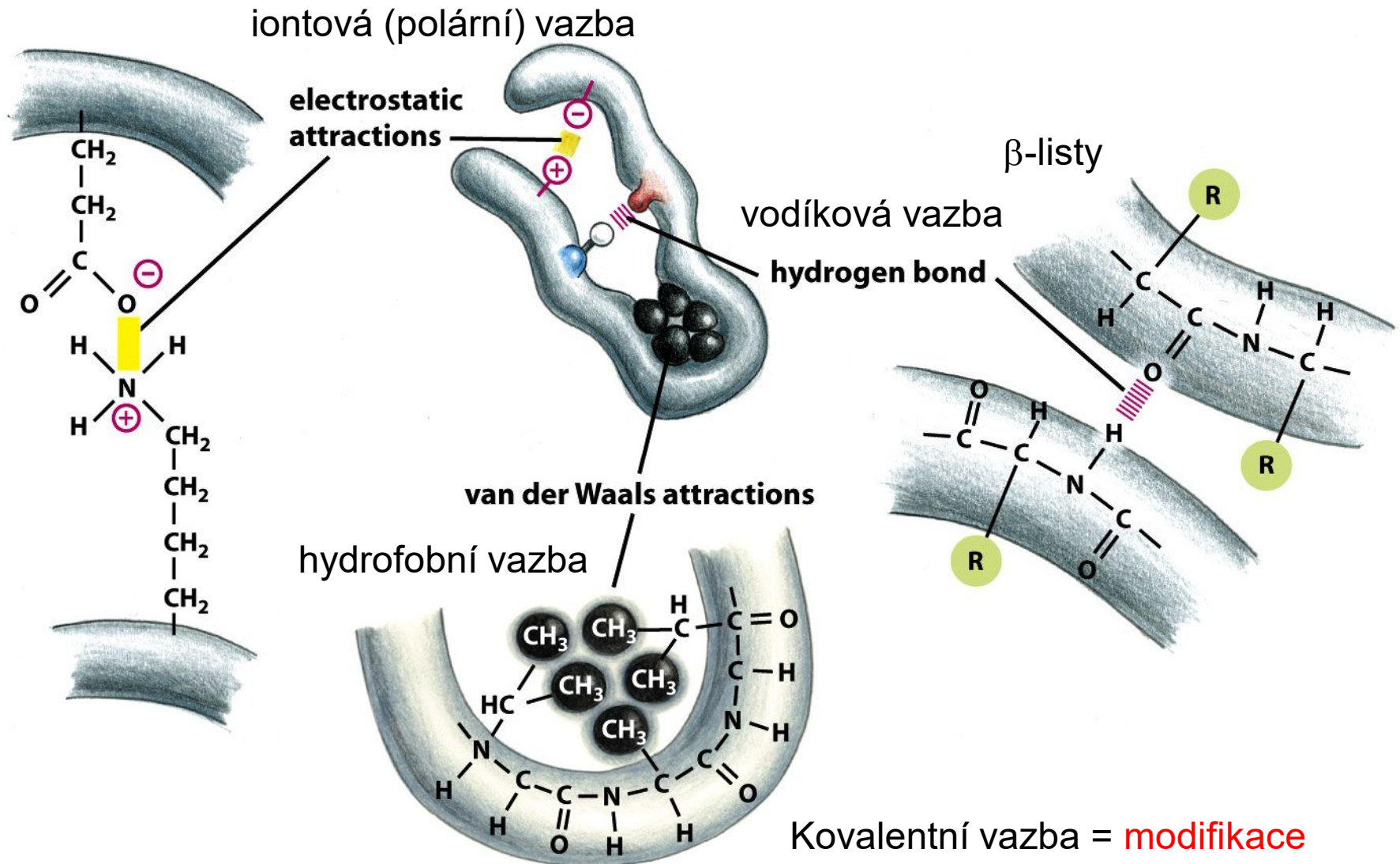


Eichborn et al, Genome Inf, 2009
Levy et al.: PNAS, 2012



poměr mezi výskytem AMK na „intaktním“ povrchu a interakčním povrchu – polární a nabité do solventu tj. povrchu - hydrofobní na povrchu nejčastěji vytváří protein-proteinové interakce

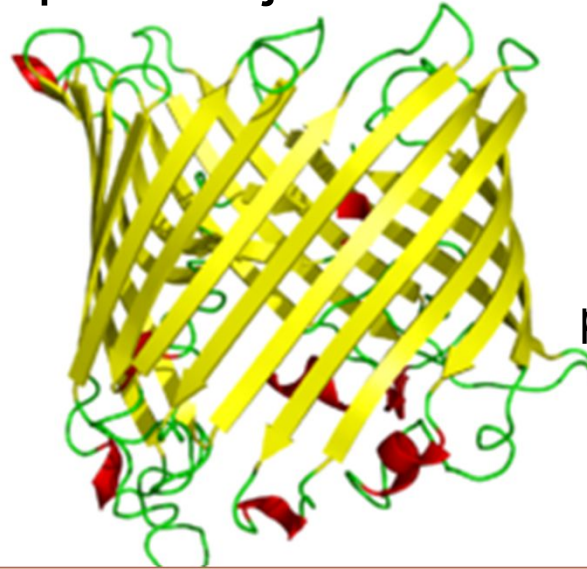
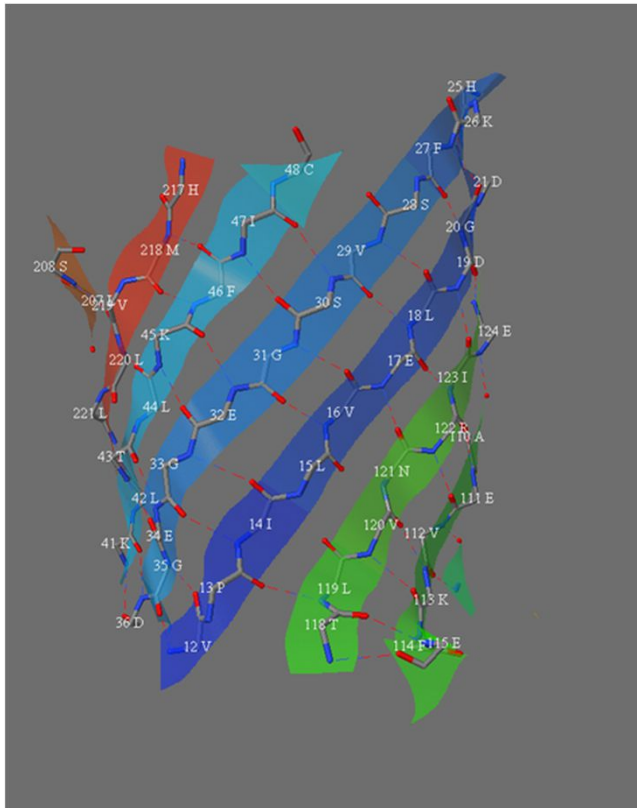
Typy vazeb v PPI



Kovalentní vazba = **modifikace**
 vyjimečně např. disulfidické můstky
 nebo jiné posttranslační modifikace
 (ubikvitinace, SUMOylace)

... sekundární struktury ...

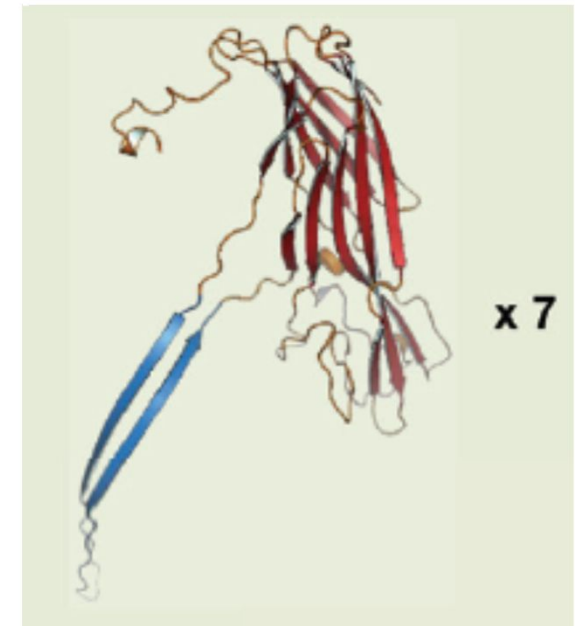
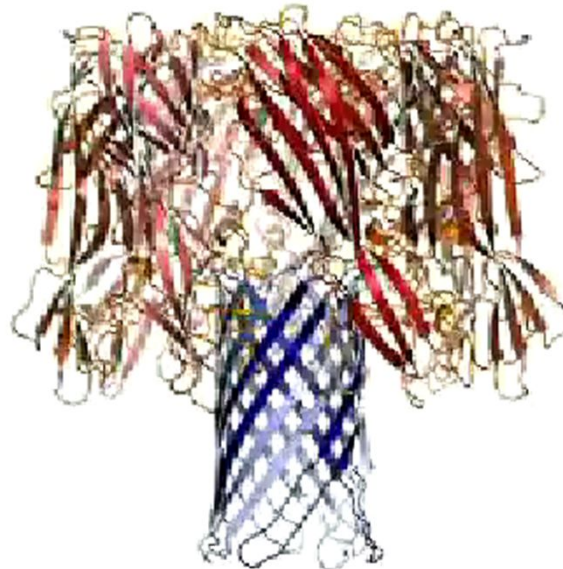
V interakcích beta-listů převažují vodíkové vazby (peptidového řetězce)



Porin
(1 ORF - polypeptid
prostup mitochondriální
membrány)

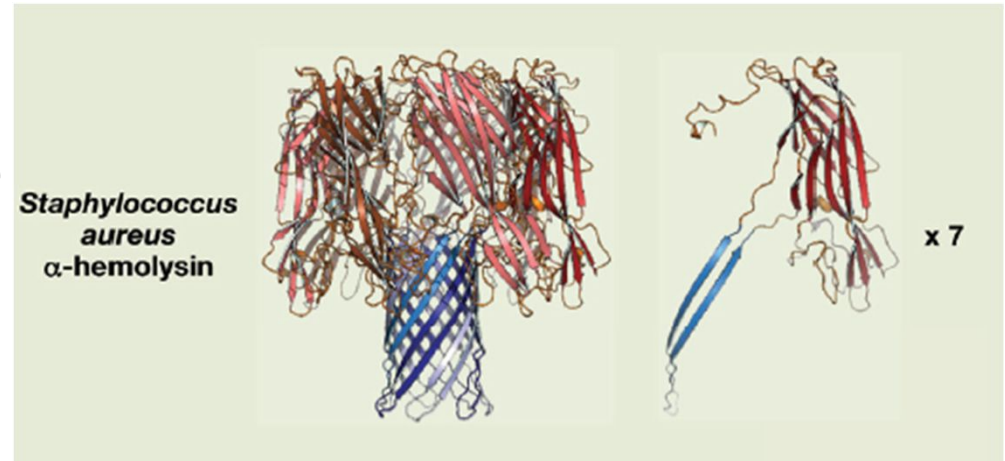
Podobný „pór“ vzniká
interakcí 7 podjednotek

Mueller & Ban, Cell, 2010
Los a spol, MMBR, 2013

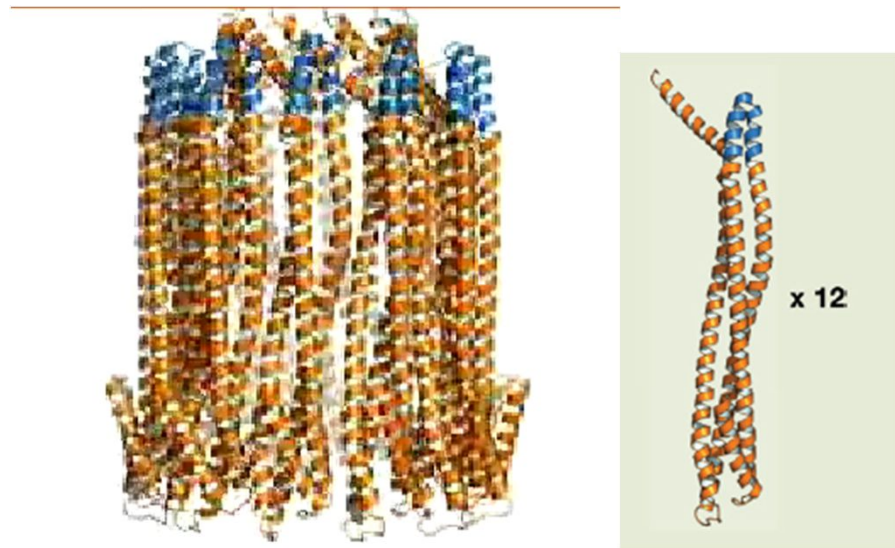
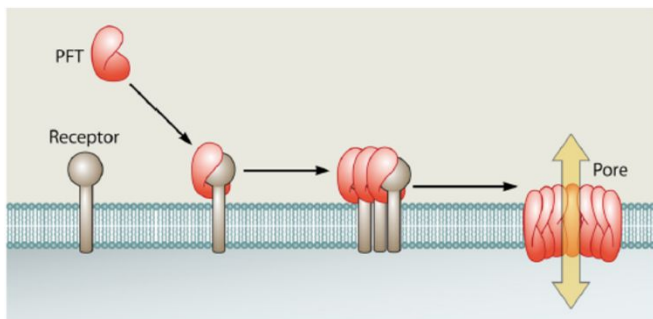


... sekundární struktury ...

- **listy, šroubovice, smyčky**
... se podílí na protein-proteinových interakcích (PPI) podobným způsobem jako při skládání proteinu do 3D – podobné sterické faktory (listy vůči sobě, šroubovice vůči sobě)



- **folding-skládání** ... struktura některých „disordered“ proteinů se utváří až v rámci interakce s druhým proteinem



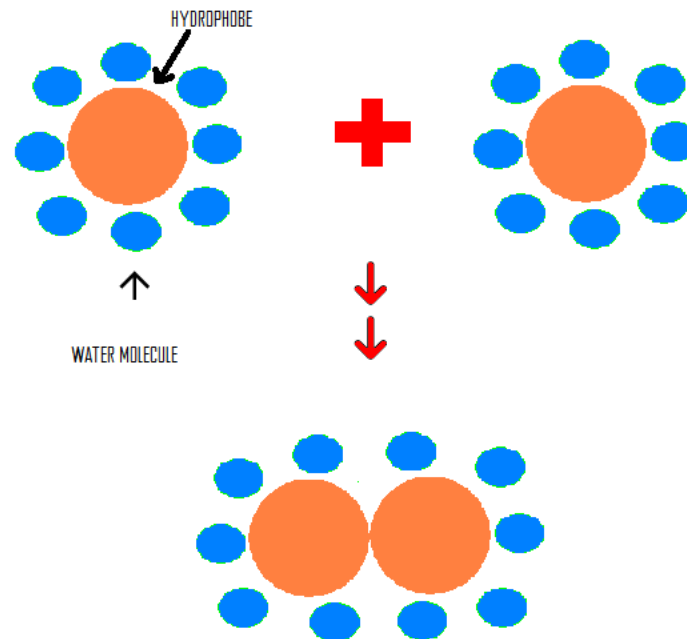
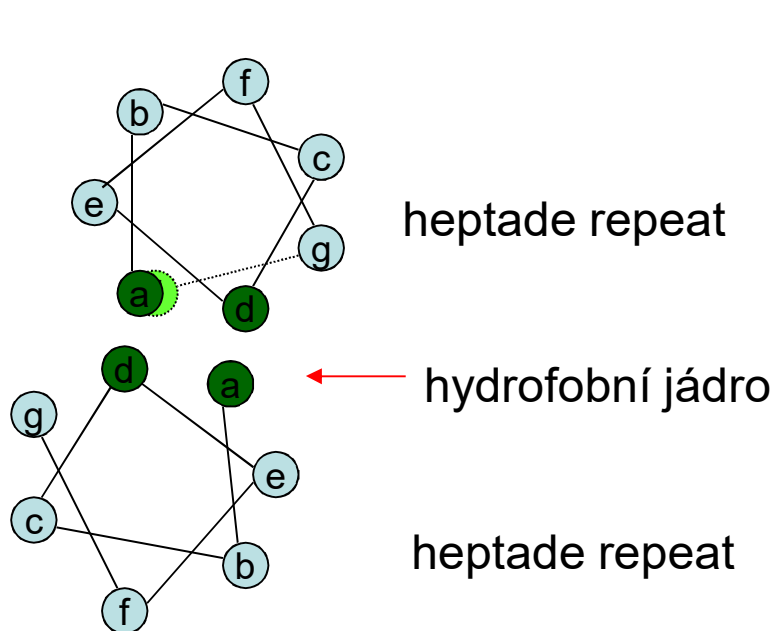
Toxiny – podjednotky se skládají tj. vytváří pór až v místě působení (neublíží původní buňce)

... sekundární struktury ...

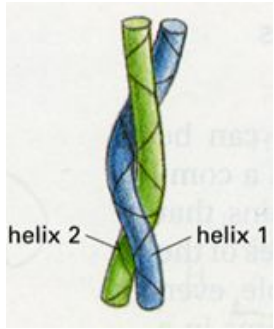
- šroubovice se vůči sobě orientují různým způsobem
- skládání slabých vazeb ovlivňuje sílu a specifitu celkové vazby

coiled-coil struktura

- dvě šroubovice s tzv. heptádovou repeticí (hxxhxxx – hydrofobní zbytky vytváří rozsáhlý povrch a tedy silnou vazbu)

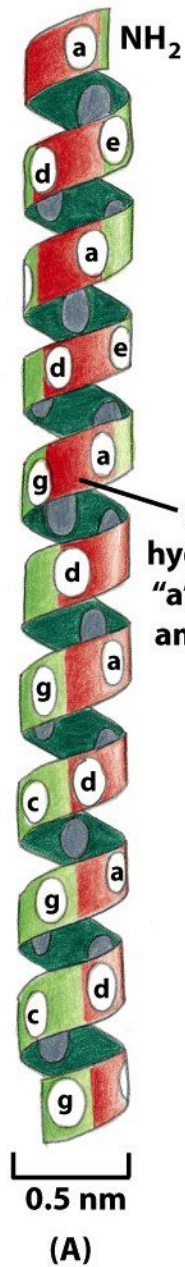


...LKSLHNQLRDLEESLTH...

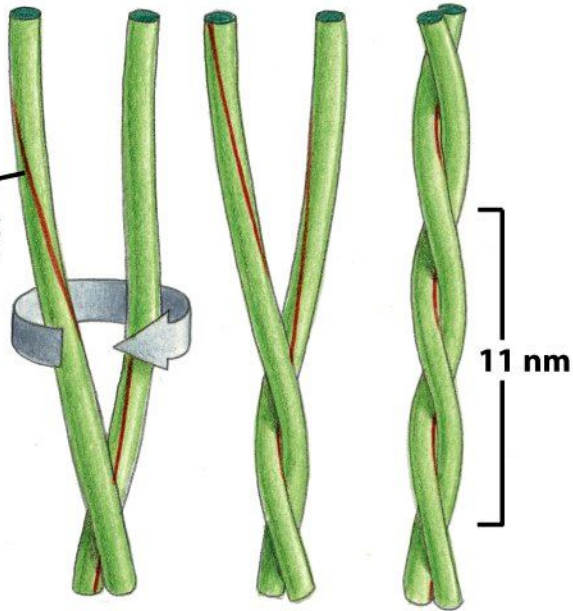


coiled-coil struktura

- dvě šroubovice s tzv. heptádovou repeticí (hxxhxxx – hydrofobní zbytky vytváří rozsáhlý povrch)



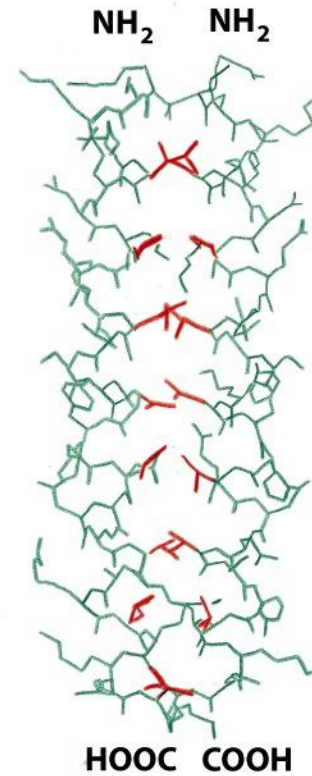
stripe of hydrophobic "a" and "d" amino acids



hydrofobní jádro

...LKSLHNQLRDL~~EE~~SLTH...

(B)



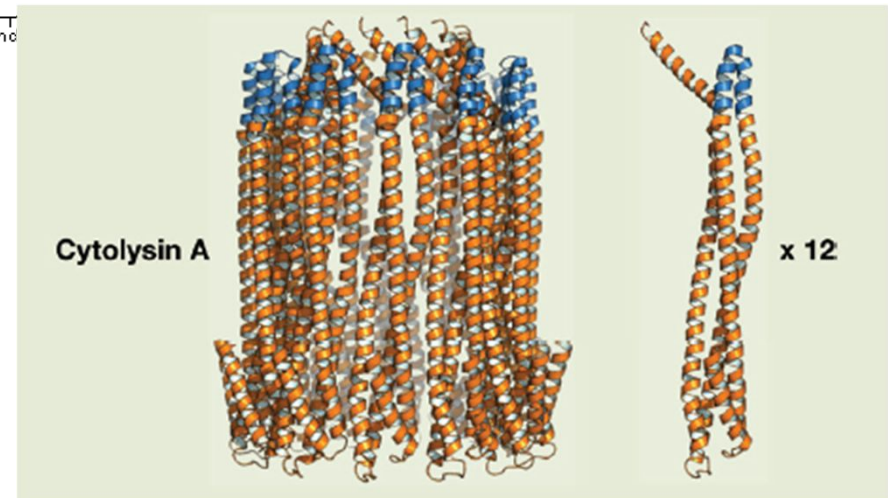
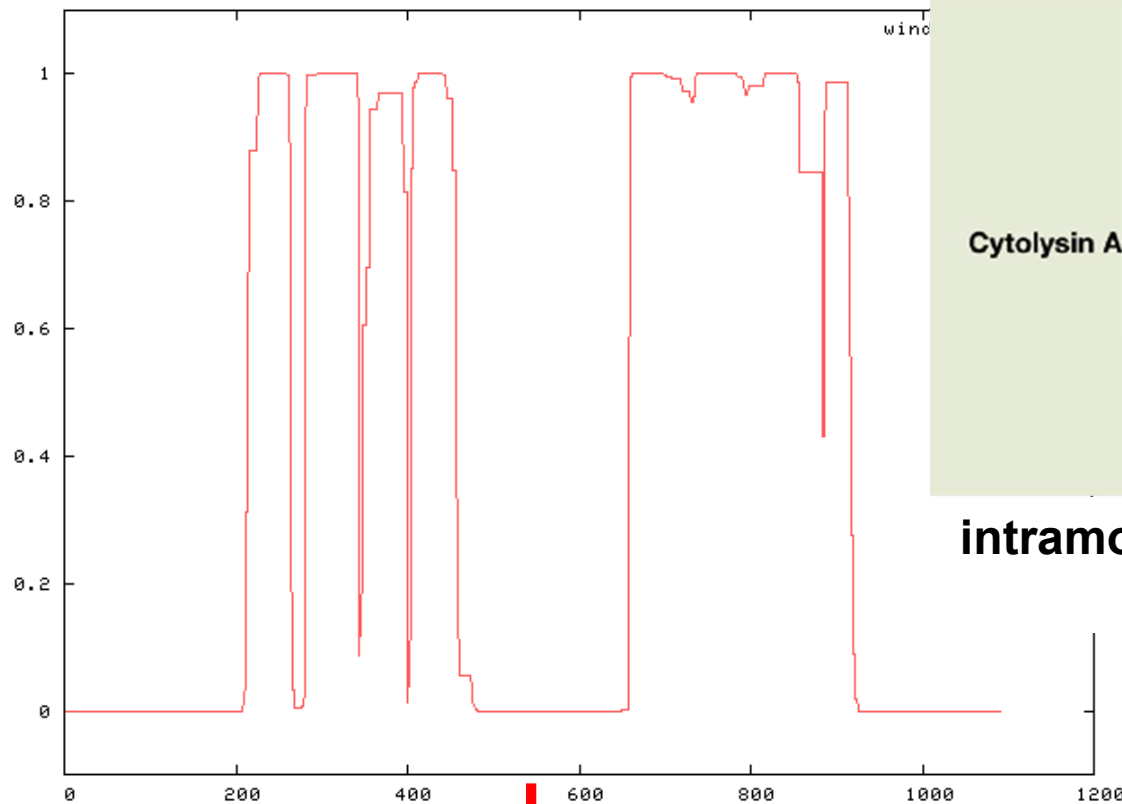
paralelní šroubovice

(C)

coiled-coil struktura

- program COIL: http://www.ch.embnet.org/software/COILS_form.html

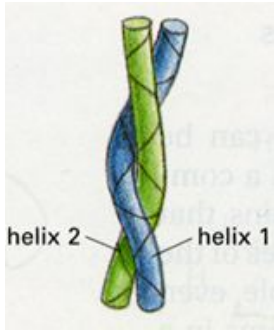
profil SMC6



**intramolekulární vs intermolekulární
folding vs PPI**

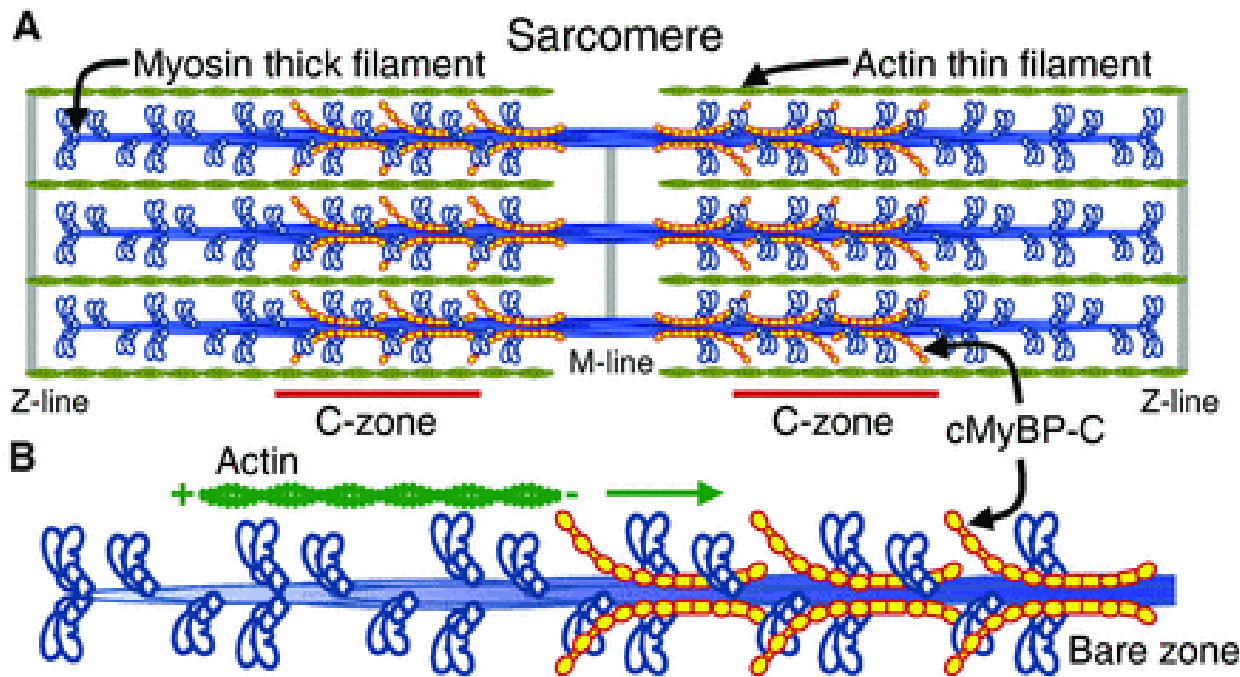


- CC v SMC proteinech jsou intramolekulární (antiparalelní)

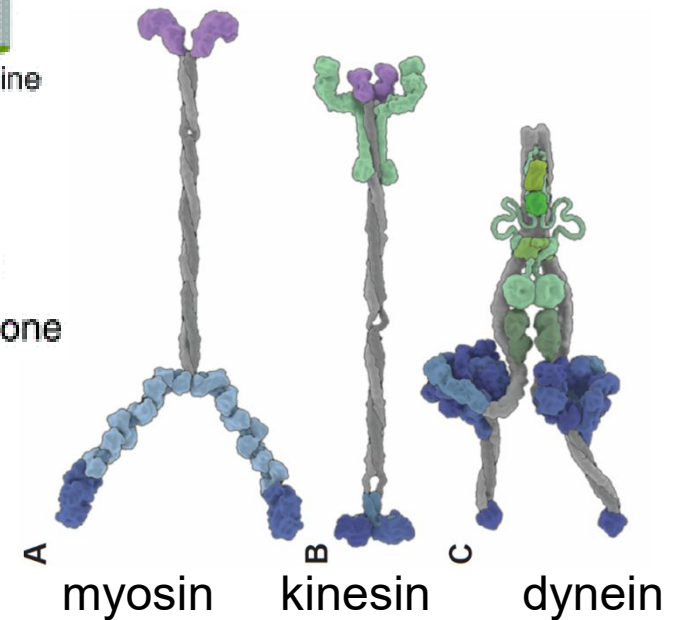


coiled-coil struktura

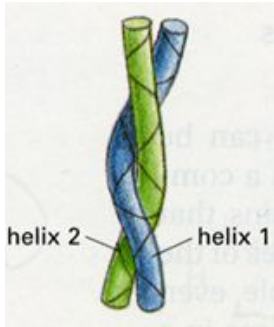
-dlouhé CC (>100AMK) vytváří vláknité struktury (myosin tvoří vlákna - svaly)



- CC v myosinu je intermolekulární (paralelní)

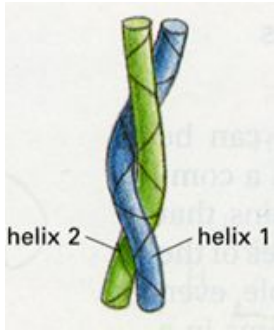


Kinesiny



members (alias) [orthologues]	Directionality and properties	Structural organization	Length (nm)
Kinesin-6 for example, KIF23 (MKLP1) and KIF20A (MKLP2) [ZEN-4 and Pavarotti]	<ul style="list-style-type: none"> • Plus end • Non-processive • Oligomerizes • Crosslinks 	<p>Motor domain Predicted coiled-coil Globular domain Microtubule-binding site</p>	0 25 50 75 100 125 150 175 200
Kinesin-12 for example, KIF15 (hKLP2) [xklp2]	<ul style="list-style-type: none"> • Plus end • Crosslinks with TPX2 	<p>N C</p>	
Kinesin-7 for example, KIF10 (CENP-E) [CENP-meta and CENP-ana]	<ul style="list-style-type: none"> • Plus end • Processive • $F_{stall} \sim 6$ pN 	<p>N C Kinetochole-binding site Microtubule-binding site</p>	
Kinesin-5 KIF11 in humans (EG5) [Cin8, Klp61F and Cut7]	<ul style="list-style-type: none"> • Plus and minus ends • Poor processivity • $F_{stall} \sim 4$ pN • Tetrameric crosslinker 	<p>N C Microtubule-binding site Microtubule-binding site</p>	
Kinesin-14 KIFC1 in mammals (HSET) [Ncd, Kar3, Klp2 and xctk2]	<ul style="list-style-type: none"> • Minus end • Non-processive • Dimeric crosslinker 	<p>N C Microtubule-binding site</p>	
Kinesin-8 for example, KIF18A [Kip3, Klp5, Klp6 and Klp67a]	<ul style="list-style-type: none"> • Plus end • Processive • $F_{stall} \sim 1$ pN • Dampens microtubule dynamics 	<p>N C Microtubule-binding site</p>	
Kinesin-4 for example, KIF4 in mammals [xKlp1 and Klp3a]	<ul style="list-style-type: none"> • Plus end 	<p>N C DNA-binding site</p>	

Cross, NRMCB, 2014

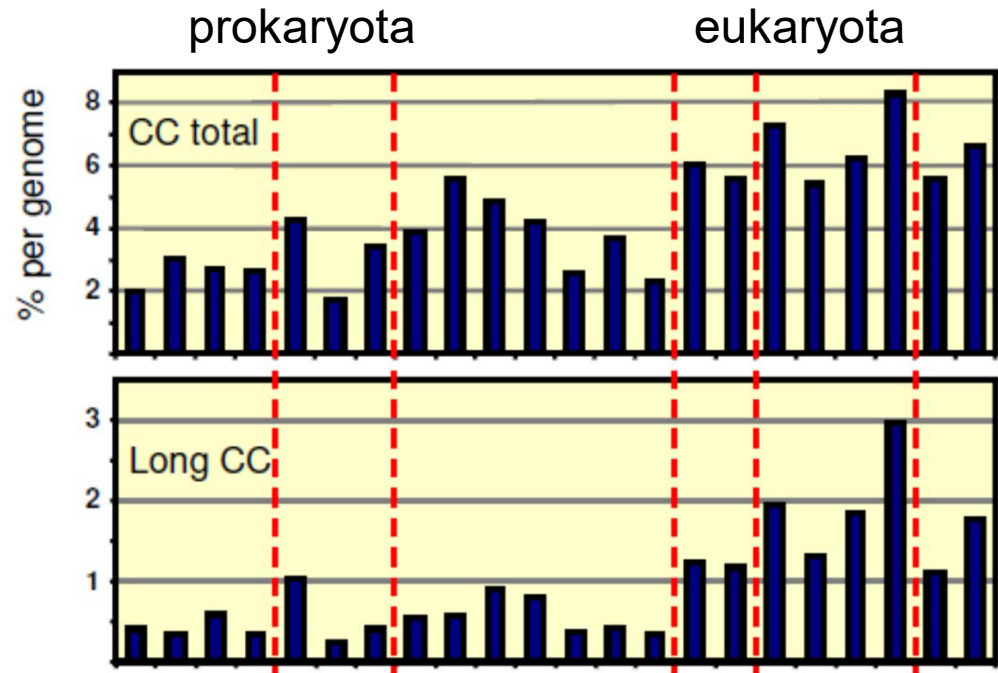


Coiled-coil
doména je
významným
dimerizačním
modulem u mnoha
proteinů (GCN4,
Max ...)

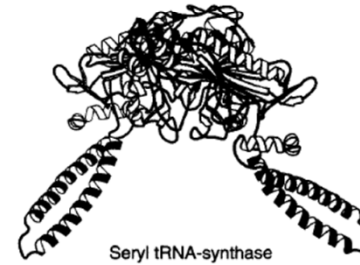
Intermolekulární -
homo- či
heterodimery
(oligomery)



Tropomyosin



DNA polymerase I



Seryl tRNA-synthase



Transcript cleavage factor GreA



Influenza hemagglutinin HA2 (pH7)



Influenza hemagglutinin HA2 (pH4)

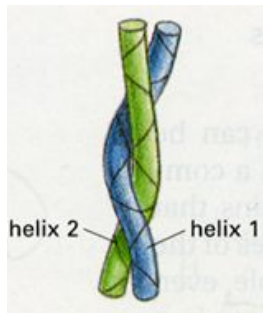


Mannose-binding protein



ROP

Interakce šroubovic

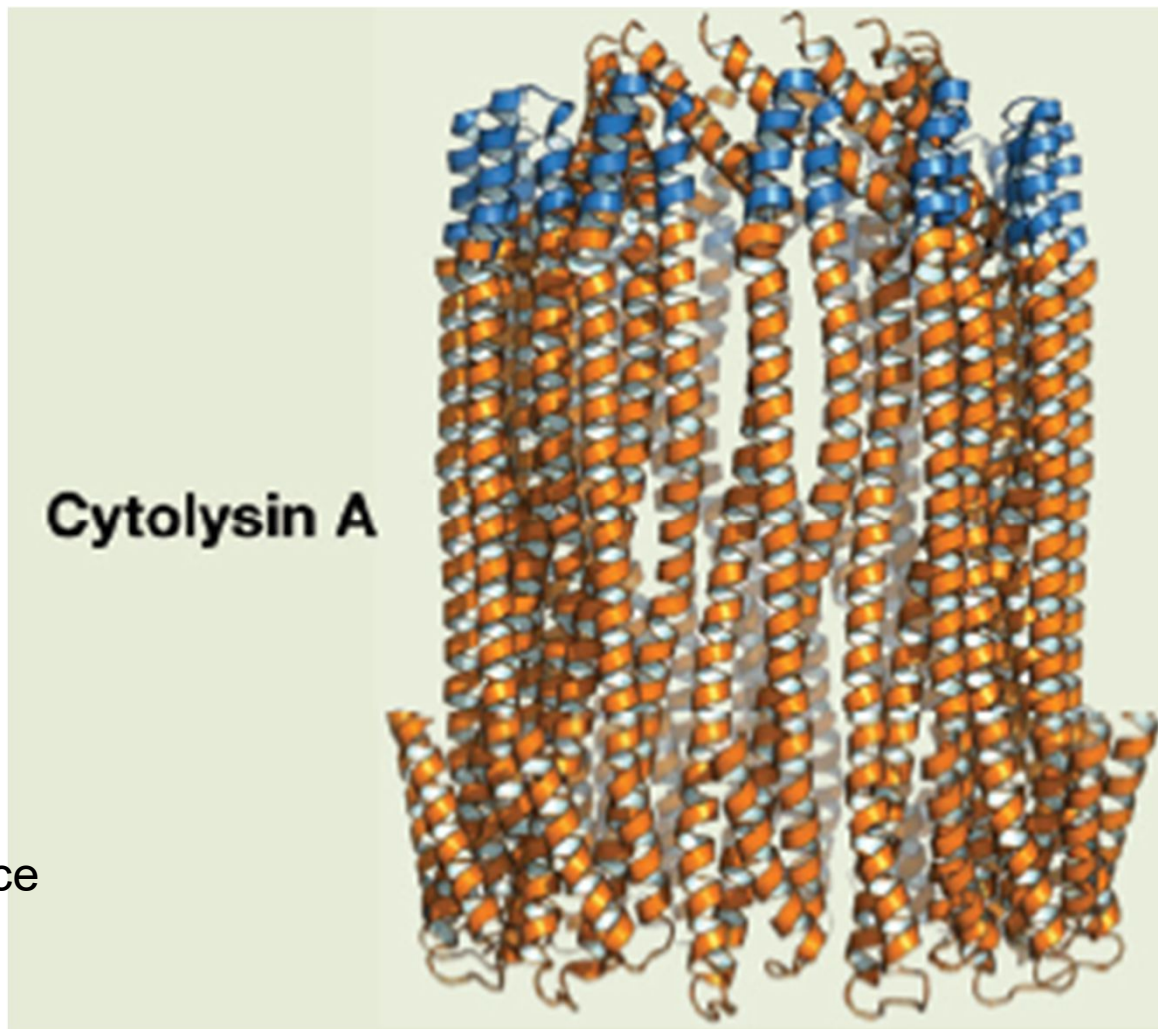


3 šroubovice



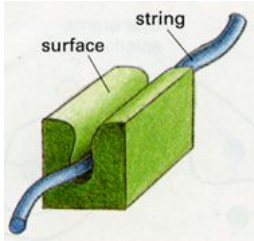
Influenza hemagglutinin

4 šroubovice



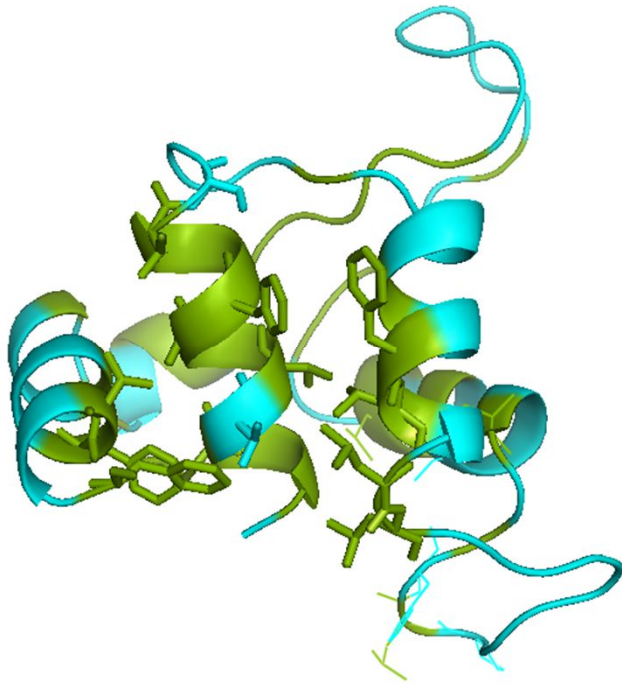
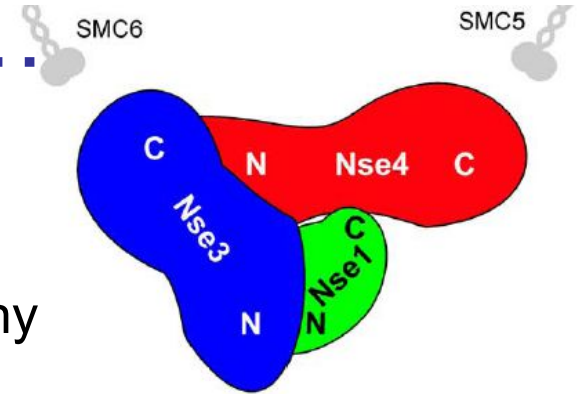
Cytolysin vytváří póry v membránách cizích buněk

Šroubovice se pod určitým úhlem dotýkají - obtáčejí Mueller & Ban, Cell, 2010; 1QOY, 2WCD

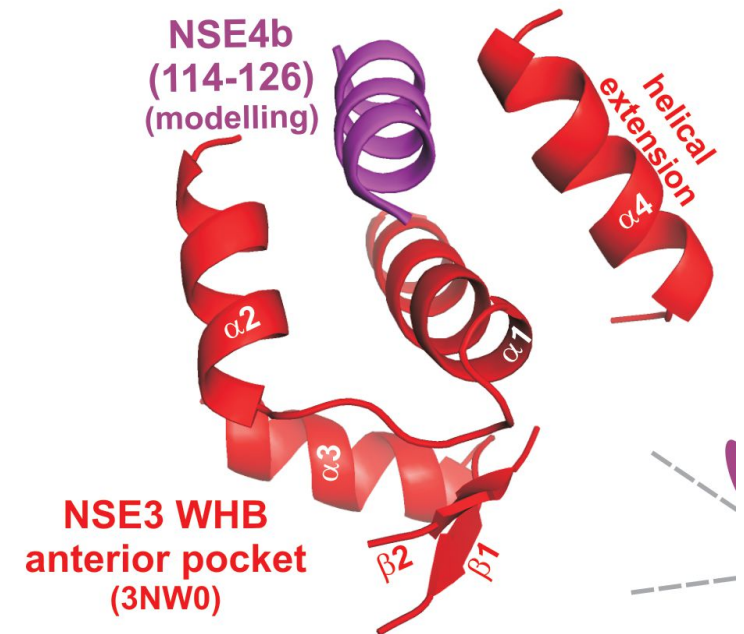


... terciární struktura ... kapsa-peptid

sekundární struktury (šroubovice, beta-listy)
interagují pod různými úhly a vytváří různé povrchy



hlubší prohlubně na povrchu
mohou tvořit kapsy pro vazbu
partnera (šroubovice, peptid)



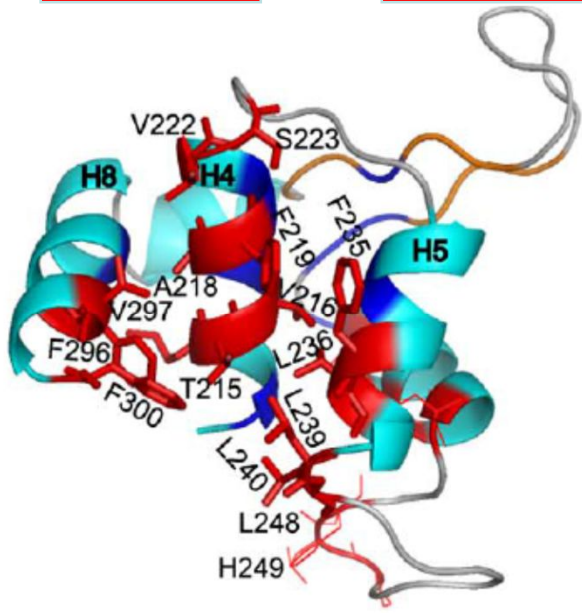
hydrofobní interakce mezi Nse3 a Nse4

Tato hydrofobní šroubovice není transmembránová, ale podílí se na protein-proteinové interakci (NSE3-NSE4)

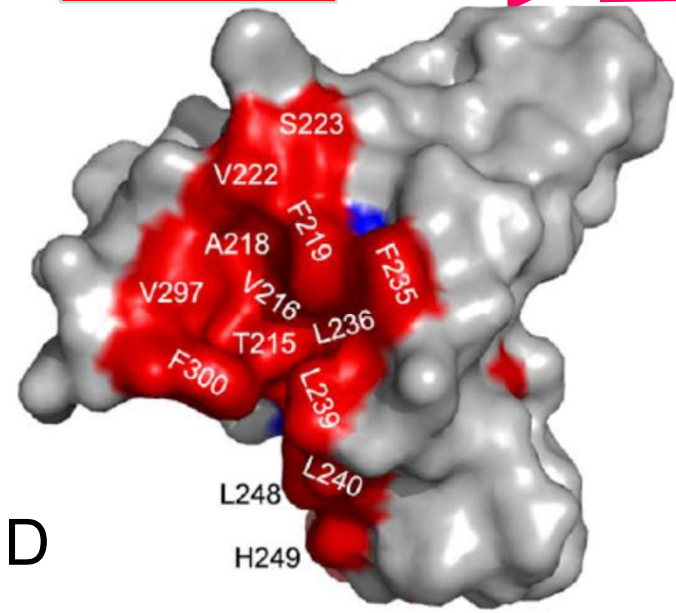
mut. -m-mmmmmmm-mm--m--m--m--mm-mmm-----mmm--m--m--m-mmm-mm-----m-----m-mm-----m--m--mm--mm

Nse4

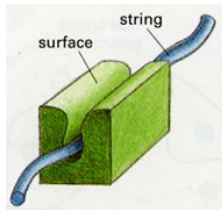
S.p.	GFLMTVIAFIAVSHCSVG-HSELQSFLLQELLT--EEETTPHLHDITRSLSLLVRCGYL--DFVKIDTHNQFVYYI-GSRAVTIISIEGLKSFVTEFF
A.n.	GLYTFIIAVILLNGGTLQ-EOKLDRYLSRMNA---EQFTPVVER-TIHLQLRLCKEGYL--VKNRIM-x-IEEYMV-GPRGKIVGARGVAGLVREYV
N.f.	GLYSFIIAVILLNGGSLP-EOKLERYLKRNTA---DTYTPVUR-TERFLQRLCKEGYL--IRNRIM-x-IEEYMV-GPRGKIVGVQGVAGLVREYV
A.t.	GLYTFIIALILLNGGSLP-EOKLERYLQRTNT---DTYTPIDR-TERFLQRLCKDGYL--VFNRRM-x-VIEYMV-GPRGKIVGTQGVAGLVREYV
A.c.	GLYSFIIAVITLNGGSLP-EOKLERYLKRNTA---DTFTPIDR-TERFLQRLCKEGYL--LRTIRM-x-VVEYLV-GPRGKIVGVQGVAGLVREYV
N.c.	GLYTMLIAIITLGGGELS-EPRLLRYLRLNAAxPNNENAPSEK-TELVLQRMTKCGYL--VFVAEN-x-AITWHV-GPRGKIVDNEATAAVVREYV
M.g.	GLYSMIVTIIQLNRGELS-EPKLKRYLQRLNA---ETNTPVEK-TDLLLQRLIRONYI--VKTVR-x-AITWRV-GPRAKQILTDEAMASIVRDVY
A.o.	GLYSFIIAVIMLNGGSLP-EOKLDRYLARTNA---DTYIPIUR-TERLLQRLCKEGYL--VNRTRM-x-LVEYMV-GPRGKIVGTAGVAGLVREYV
S.c.	GVLSVILCIVFFSKNNIL-HOELIKFLETFCIPSDGSKIAILNITIEDLKSLEKREYI--VFELEK-x-VISYRI-GRRTQALGLESLEKLVQOIM
D.r.	GLLFVILSVIFMKGTTIK-EINLVWNTLKKLRDPGEGKHDFGIV-KKVVTVEFVRQKYLEYCKIPHTEPVEYEFRW-GLRAEKIVSKLKLLEFVGBLF
X.t.	GLLMVILSLIFMKGNTAK-EHAVWEMLRRLRTEPAEKHSDFGIV-KKLITTEFVRQKYLEYSKVLHTDPVEYEFRW-GQRAFKIVTSKMQVLEFVSKIQ
G.g.	GLLIVILSFIKMGNSAK-EHAVWEFLRRLRVHPGKHEVFGIV-KKLVMEFVRQKYLEITPIPLTDPPEFNFQW-GPRAAKIVTSKKDILSFVAKMQ
O.a.	GLLMVILSLIFMKGSATN-ESEVIWETLRKLKRVDTREHEVFGIV-KKLVTEFVRQKYLEYRRIIPHTEPVEFEFQW-GARATKIVTTKMQVILNFVAKGP
M.d.	GLLMVILSLIFMKGNSAR-ESELVWDVLKKLRVDPEKRHKTFGIV-KKLVKDEFVRQKYLEYIRVPHSEPPPEYEFVW-GPRAAHIVTSKMQVLRVAKIQ
D.n.	GLLMIILGLIFMKGNCIK-ESELWRFLRRLGVYPTKKHLVFGIV-KKLITGEFVRQRYLKYQRIIPHTEPVDYELW-GPRSHLIVTSKMKALKFVAKIH
C.f.	GLLMIVLGLIFMKGNTIK-ESEVWDFLRLRGVYPTKKHLVFGIV-KKLITTEFVRQRYLEYRRIIPHTEPVDYEFQW-GPRTNLIVTSKMKVLFVAKVH
M.m.	GLLMIVLGLIFMKGNTIT-ESEVWDFLRLRGVYPTKKHLIFGIV-KKLITTEFVRQRYLEYRRIIPHTEPVDYELQW-GPRTNLIVTSKMKVLFVAKVH
H.s.	GLLMIVLGLIFMKGNTIK-ESEAWDFLRLRGVYPTKKHLIFGIV-KKLITTEFVRQRYLEYRRIIPHTEPVDYEFQW-GPRTNLIVTSKMKVLFVAKVH



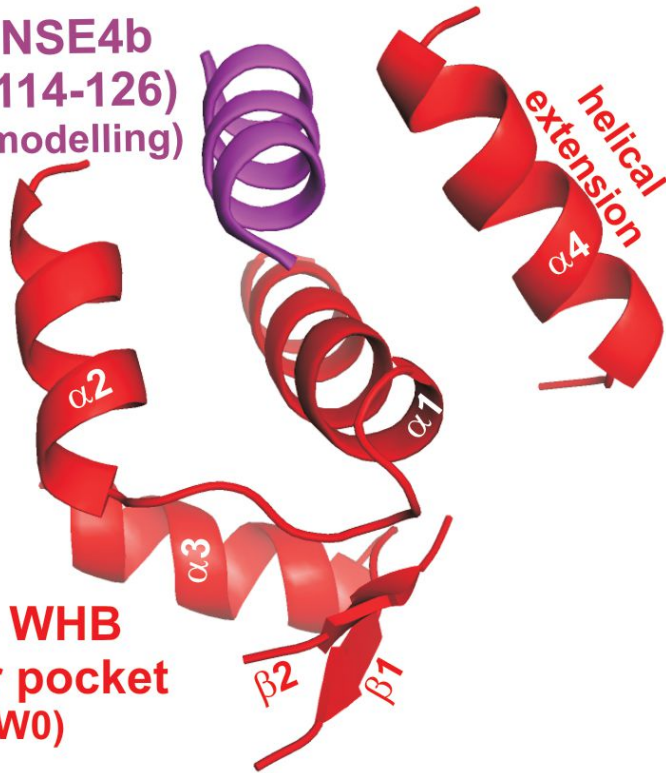
WHD



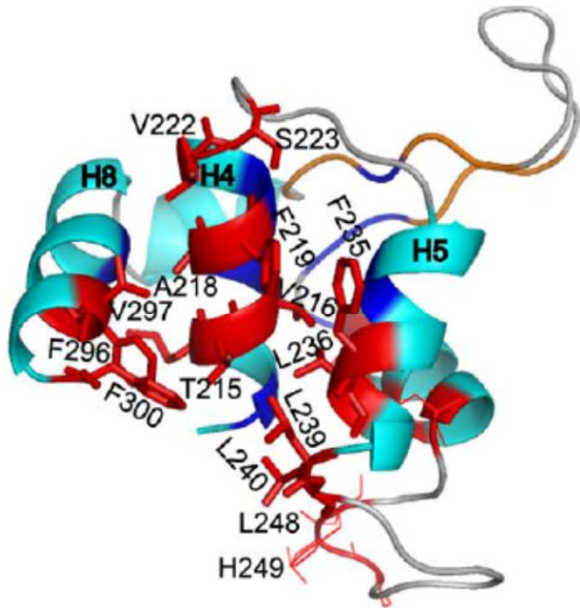
Interakce mapována mutagenézí



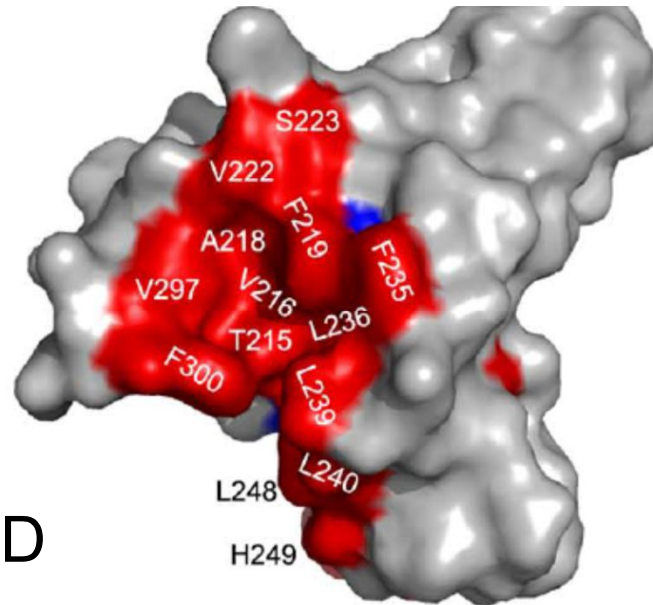
**NSE4b
(114-126)
(modelling)**



**NSE3 WHB
anterior pocket
(3NW0)**



WHD



**Interakce
mapována
mutagenezí**

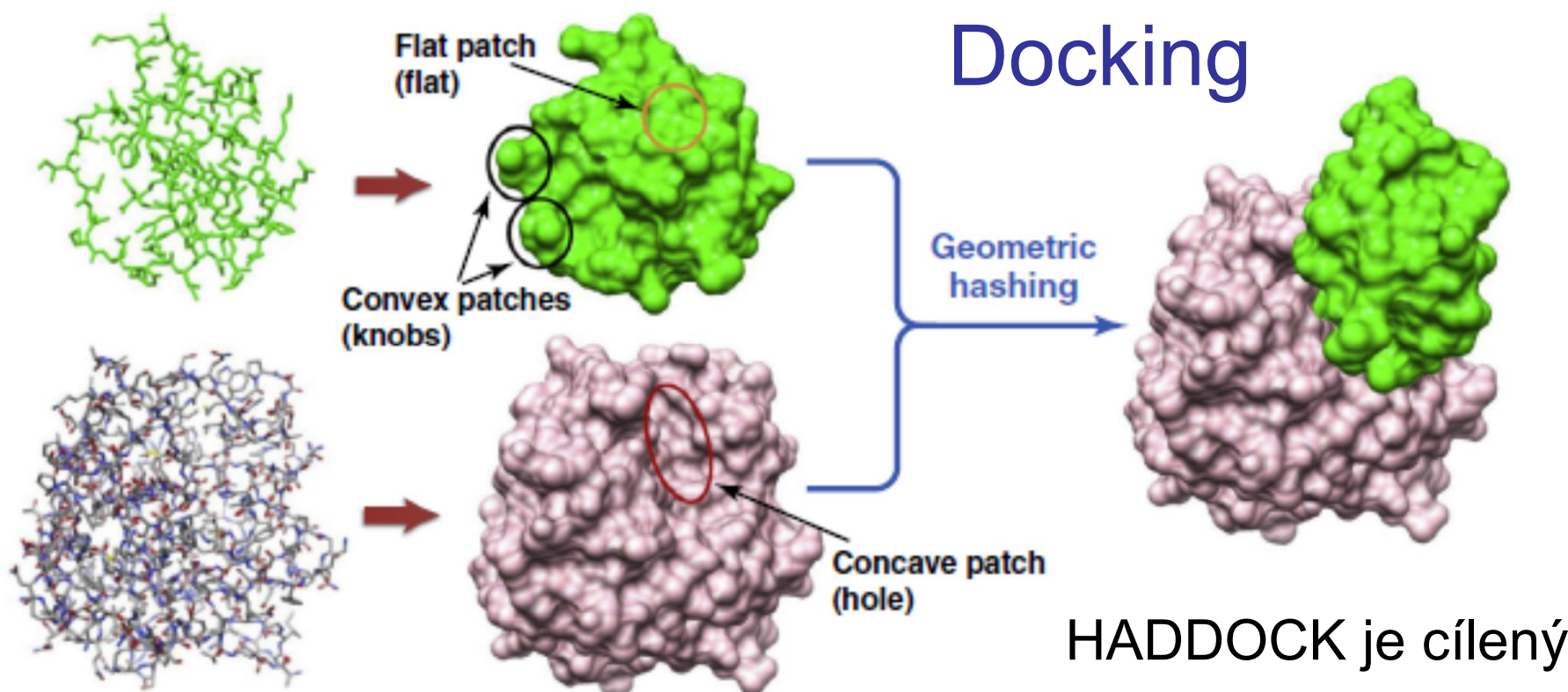
The screenshot shows the HADDOCK Software web portal. The header features the HADDOCK logo and the text "Software web portal". The navigation menu includes: Home, HADDOCK, Whisky, CPDRT, DNA, SQUEEZE, Publications, HADDOCK, and Contact. Below the navigation menu, there is a welcome message: "WELCOME TO THE UTRECHT BIOMOLECULAR INTERACTION WEB PORTAL >>". The main content area includes a description of the portal and a section titled "HADDOCK WEB DOCKING" with a small cartoon character icon. On the right side, there is a "PROFILE >>" section with the Universiteit Utrecht logo and a small image of a protein structure.

Docking

TABLE 1

Search strategies in protein-protein docking

Search algorithms	Examples of docking programs	Refs
Exhaustive global search		
FFT-based search	FTDock, GRAMM, DOT, ZDOCK, MolFit, PIPER, F2DOCK, SDOCK, ASPDock, Cell-Dock	[25–41]
Spherical Fourier transform-based search	HEX, FRODOCK	[45–47]
Direct search in Cartesian space	SOFTDOCK, BIGGER, SKE-DOCK	[49–51]
Local shape feature matching		
Distance geometry algorithm	DOCK	[52]
Geometric hashing	PatchDock, SymmDock, LZerD	[53–56]
Genetic algorithm	GAPDOCK	[57]
Randomized search		
Monte Carlo search	RosettaDock, ICM-DISCO, ATTRACT, HADDOCK	[61–71]
Particle swarm optimization	SwarmDock	[72]
Genetic algorithm	AutoDock	[73]
Post-docking approach		
Using advanced scoring functions	RPScore, ZRANK, PyDock, EMPIRE, DARS, DECK, SIPPER, PIE, MDockPP, etc.	[81–94]
Considering protein flexibility	MultiDock, SmoothDock, RDOCK, FireDock, FiberDock, EigenHex, etc.	[95–104]
Other ranking protocols	SDU, CyClus, CONSRANK, etc.	[105–111]



v PDBsum můžete hledat kapsy (povrchy vhodné pro vazbu partnera) – musí mít **komplementární tvar a charakter** (terciární)

MAGEA4

PDBsum

Go to PDB code:

Top page Protein **Clefts** Tunnels Links

Cleft analysis for: 2wa0 PDB id

WHD

View options

- Binding-site(s)
- Binding-surface(s)

Coloured by

- cleft (as in table below)
- closest atom type
- residue type
- residue conservation

Clefts	Volume	R1 ratio	Accessible vertices	Buried vertices	Average depth	Residue type	Ligands									
1	2370.52	0.98	65.15	1	10.55	1	10.50	1	8	6	5	15	4	4	1	

největší kapsa

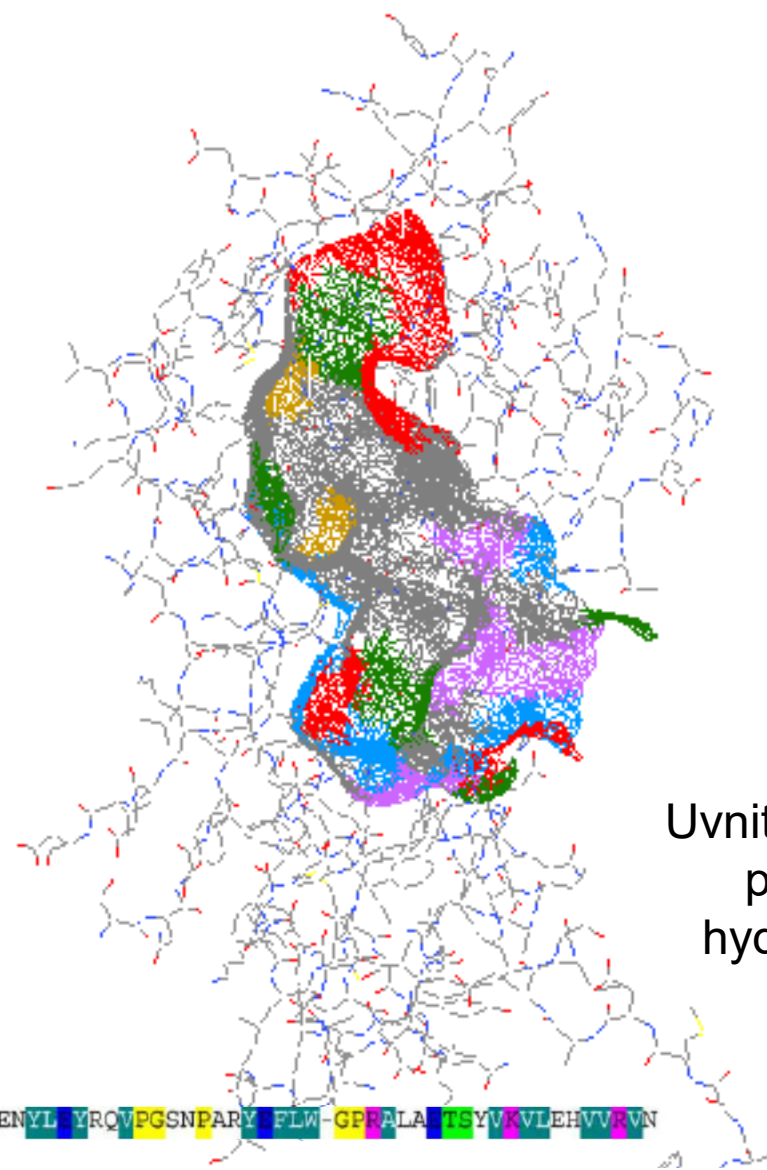
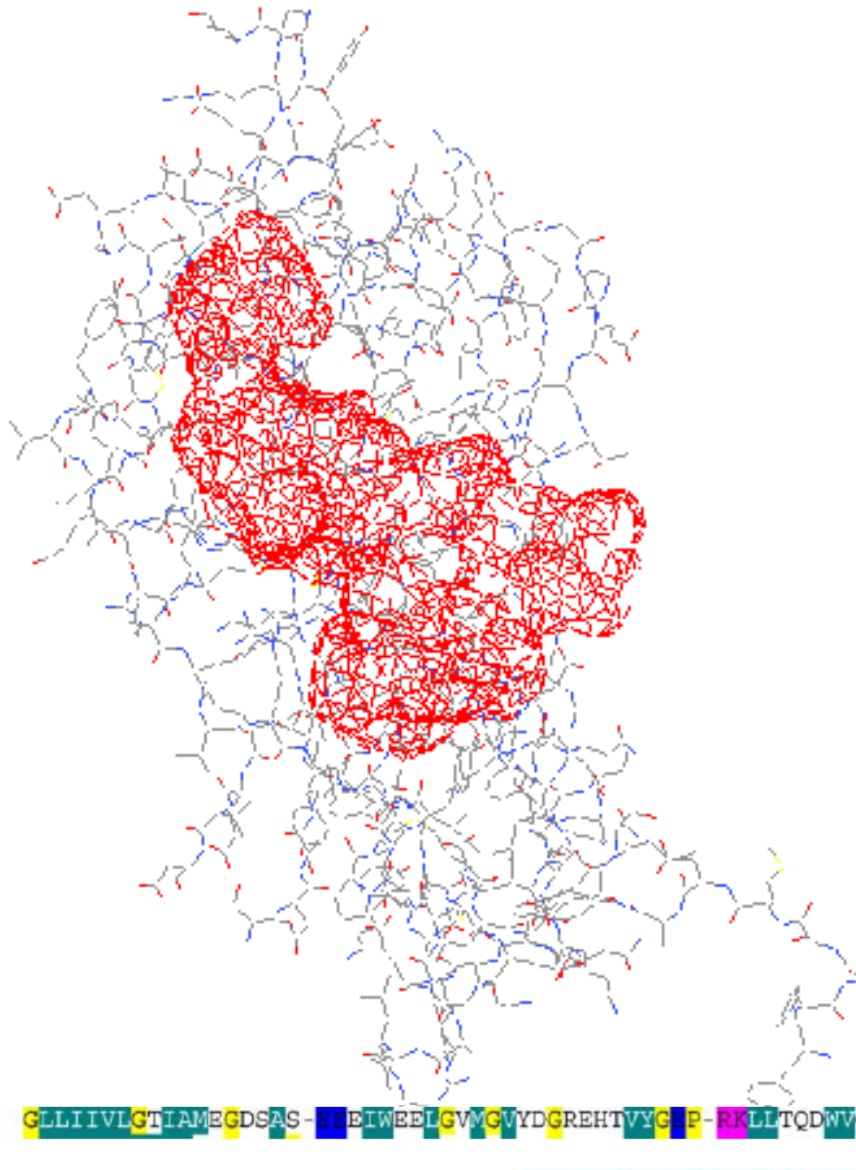
<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/>

Laskowski et al.: Prot Sci, 2018

Residue-type colouring						
Positive	Negative	Neutral	Aliphatic	Aromatic	Pro & Gly	Cysteine
H,K,R	D,E	S,T,N,Q	A,V,L,I,M	F,Y,W	P,G	C

Binding site

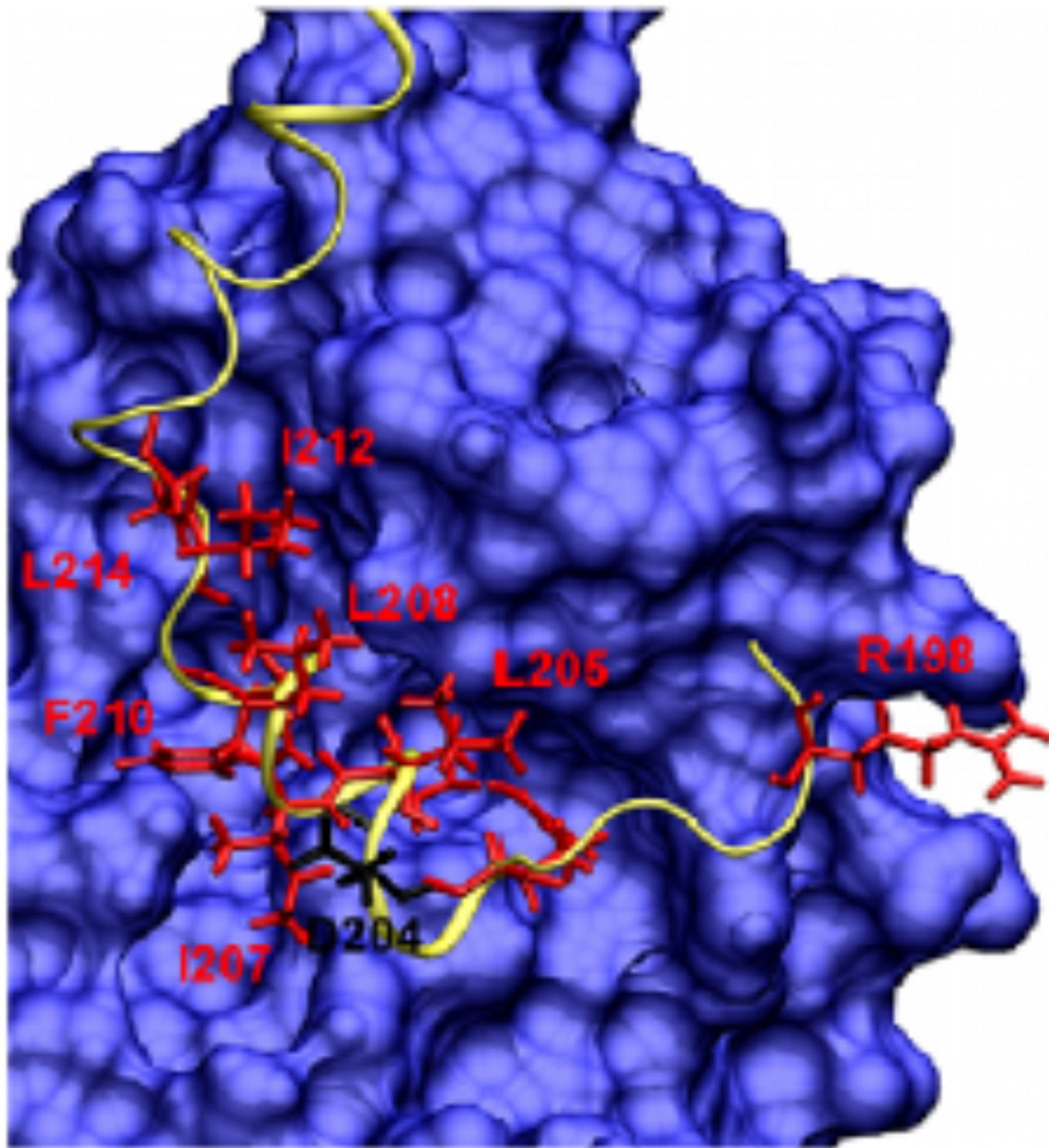
Binding surface



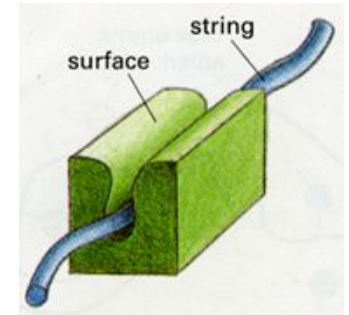
Uvnitř kapsy
převládá
hydrofobní
povrch

A4 GLLIIVLGIAMEGDSAS-ETIWEELGVVGVYDGREHTVYGC-PFKLLTQDHWQENYL YRQVPGSNPARY FLW-GPPALA TSYVAVLEHVVAVN

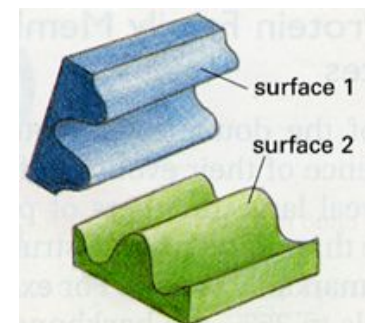
Residue-type colouring						
Positive	Negative	Neutral	Aliphatic	Aromatic	Pro & Gly	Cysteine
H,K,R	D,E	S,T,N,Q	A,V,L,I,M	F,Y,W	P,G	C



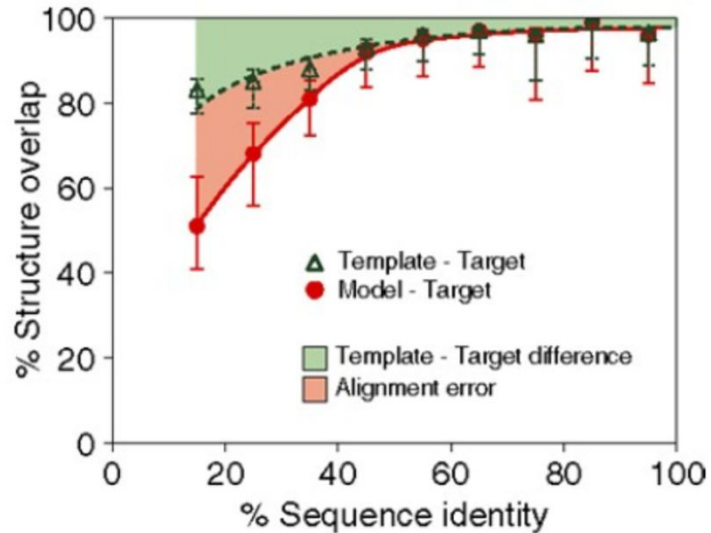
Guerineau et al.: PLoS One, 2012



de novo docking partnera (HEX docking a molekulární dynamika): do hydrofobní kapsy proteinu byl nadockován „jednoduchý“ peptid (*de novo* docking větších povrchů je nespolehlivý)



Modelování proteinů na základě homologií (*de novo* obtížné)



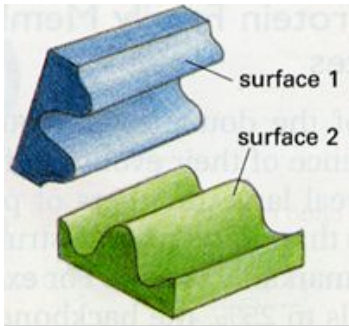
mít struktury všech proteinů (proteom)
nestačí – struktury interagujících
proteinů/komplexů (interaktom/komplexom)
přináší informaci o molekulárních detailech
buněčných procesů

Dr. T. Klumpler

Modelování proteinových komplexů = dockování také na základě podobností (*de novo* ještě obtížnější)

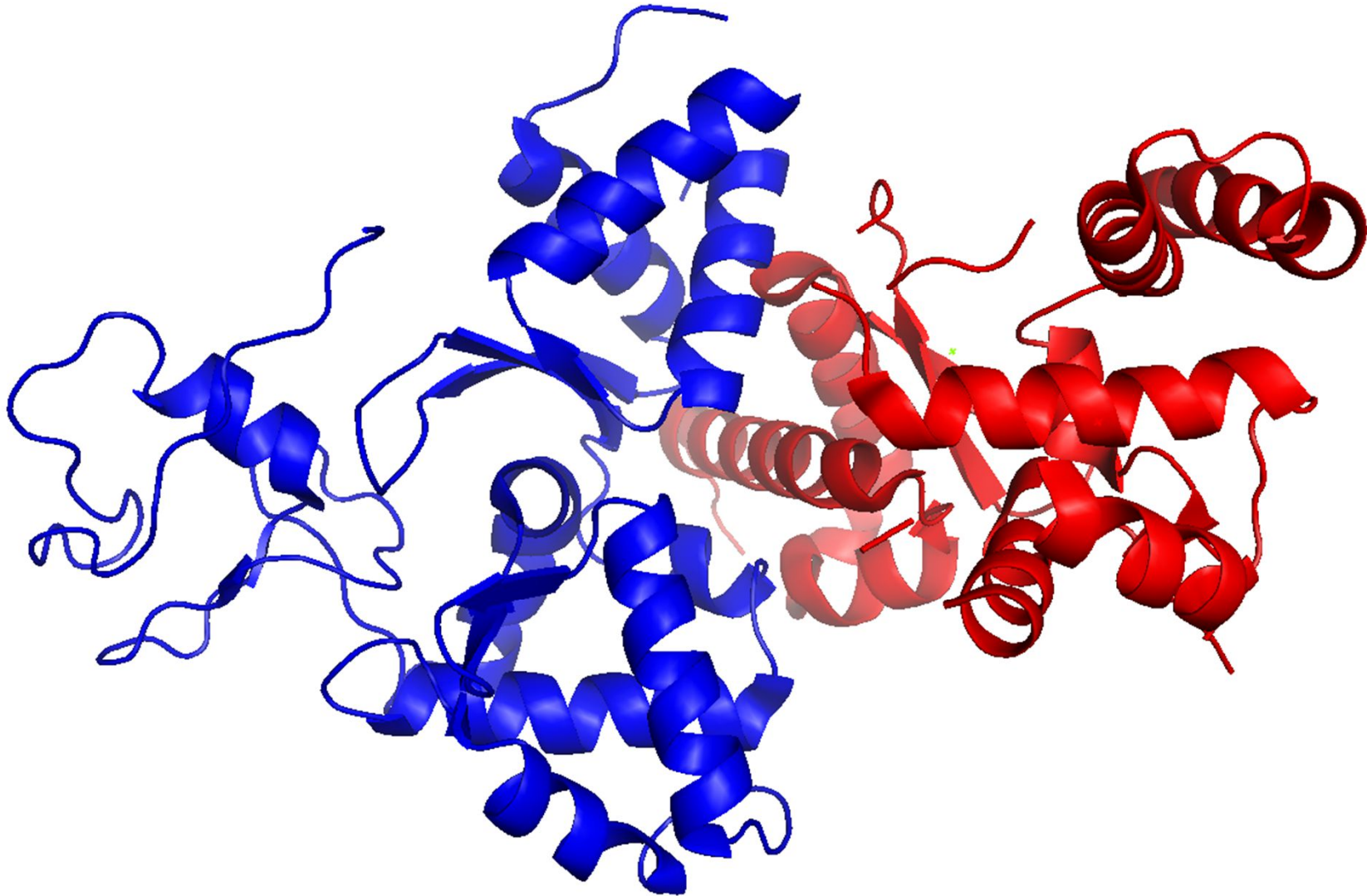
Kde najít další info o interakcích? V kterých organismech spolu dané proteiny interagují?

**CG031 – cvičení z modelování proteinových komplexů
(jarní semestr - Doc. Jan Paleček)**



DOMÉNY - šroubovice, β -listy ... interagují pod různými úhly a vytváří různé vazebné motivy s rozsáhlými vazebnými povrchy

(kokrystal NSE1-NSE3 proteinů)

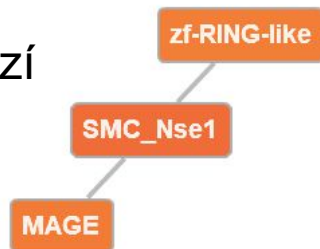


Domain SMC_Nse1 (Pfam: PF07574.11)

GO terms

P GO:0006281 DNA repair C GO:0030915 Smc5-Smc6 complex

integrace PDB,
PFAM a GO databází



D F C P Default color scheme

Interacting domains (2 domains)

MAGE

zf-RING-like

HMM profile interface residues in SMC_Nse1 (2 interfaces)



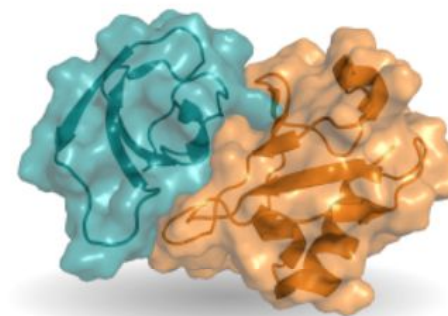
Search motif by name or keyword: ?

Motif name (e.g. SH2_LIG_0)

Search motif



<https://3did.irbbarcelona.org/>

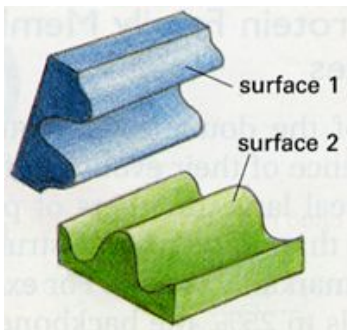


- > **Overview** General information on 3did
- > **Getting Started** Help for new users
- > **Technical Information** Linking to 3did
- > **Download** data files or MySQL tables

Statistics

Pfam version	30.0
PDB version	2017_06
Domain-domain interactions	11200
Motifs in interactions of known 3D structure	702

3DID kategorizuje doména-doména interakce z PDB (06/2017 – cca 10000 doména-doména komplexů/100000 struktur) – topologie ne detaily



PDBsum – detailní info

<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html>

(kokrystal NSE1-NSE3 proteinů)

[Top page](#)
[Protein](#)
[Metals](#)
[Prot-prot](#)
[Clefts](#)
[Tunnels](#)
[Links](#)

Protein-Protein interface: A}{B PDB id 3nw0

Protein-protein interface: A}{B
NSE1-NSE3

Chain A Chain B

Chains A and B highlighted (click to view)

Key:

- Salt bridges
- Disulphide bonds
- Hydrogen bonds
- Non-bonded contacts

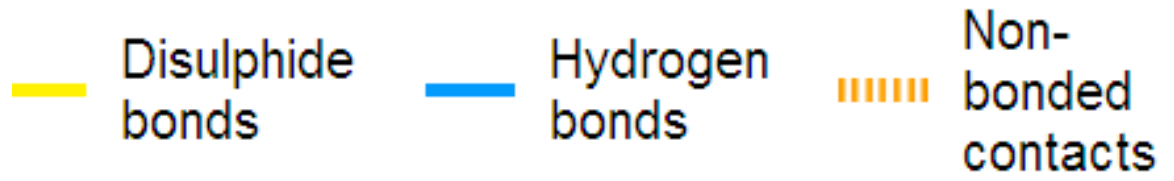
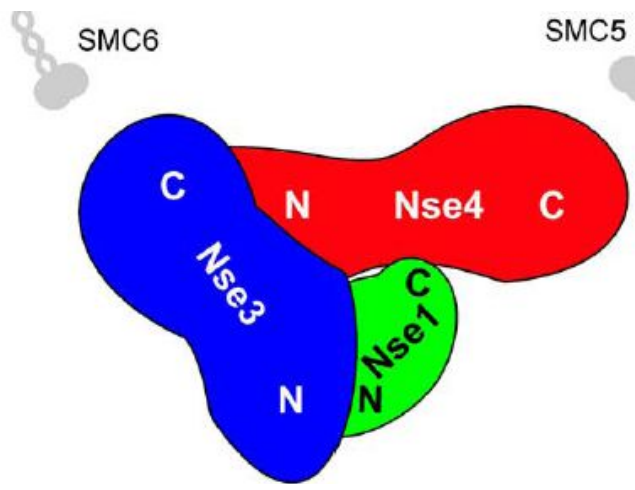
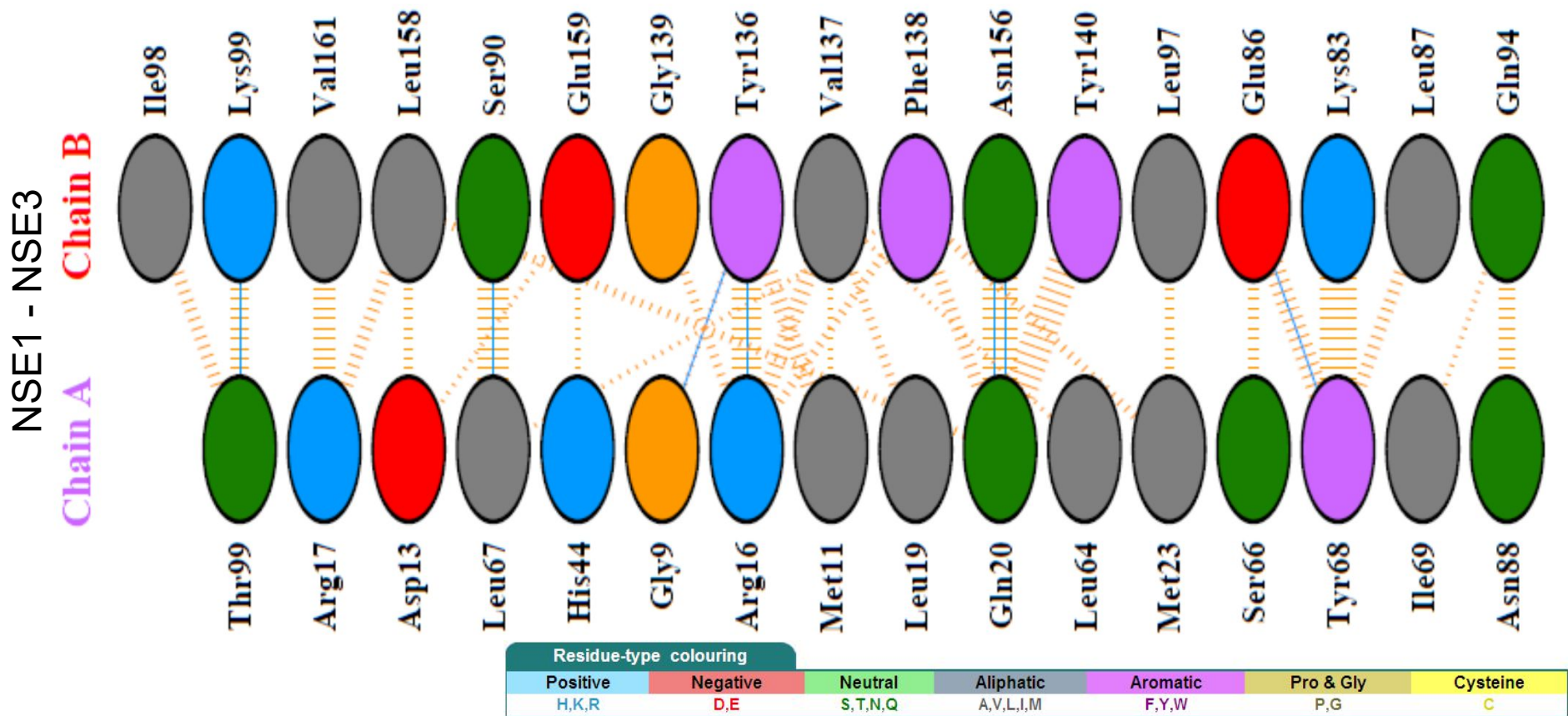
PDF Adobe
Postscript version

Jmol Interfaces
A}{B (16:17 res)

Interface statistics

Chain	No. of interface residues	Interface area (Å ²)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A	16	1015	-	-	7	100
B	17	1003	-	-	7	100

Silná interakce mezi NSE1 (chain A) a NSE3 (chain B)



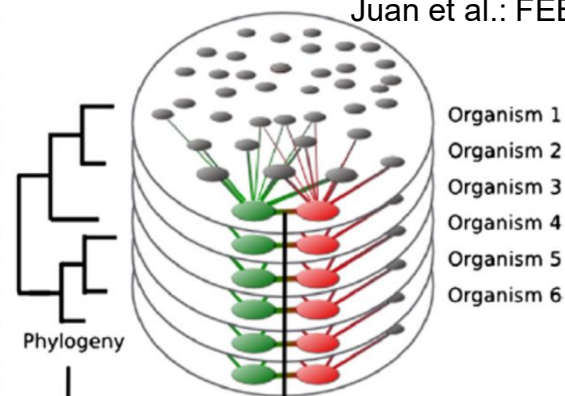
COZOID nástroj: <http://decibel.fi.muni.cz/cozoid/>

```

A1CCH2_ASPCL/14-216   NRAFLQAFM.ARSTMTFAEAKPVLAAIF.SAH.....
A2Q7K6_ASPNC/15-218   NRAFLQAFM.ARSTMTFTQAKPVLAAIF.SIR.....
B8NLA5_ASPFN/14-216   NRAFLQAFM.ARSTMTFAEARPVLAAIF.SVH.....
B6QTR9_TALMQ/14-217   NRAFLQAFM.ARSTMTFDEAKPVLAAIF.SAQ.....
V5FED6_BYSSN/14-236   NRAFLQAFM.ARSTMTFEEAKPVLAAIF.SAHGAQSTIFFDS
S7Z8E9_PENO1/8-209    HRAFLQAFM.ARSTMTFEDAQPVLAAII.SAH.....
B6H9Q9_PENRW/8-210    NRAFLQAFM.ARSCMTFEDAQPILAAIL.TVS.....
H6C926_EXODN/14-207   NRAFLQAFM.ARSVLTLETAKPILAAIS.TFQ.....
U1GD89_ENDPU/11-202   NRAFVQAFM.ARGTLTYETSKPLLASIF.TVH.....
C5GY37_AJEDR/11-207   HRAFLQAFM.ARSTMTYEQAKPVLAAIF.SAR.....
C6H5E2_AJECH/11-203   HRALLOAFM.ARSTMTYEQAKPVLAAIF.TAR.....

```

Juan et al.: FEBS Lett, 2008



Browser address bar: http://pfam.xfam.org/family/PF07574#tabview=tab1



PFAM – databáze proteinových motivů

Nse1 motiv



Family: **SMC_Nse1 (PF07574)**

16 architectures 522 sequences 2 interactions 443 species 1 structure

Summary

Domain organisation

Domain organisation

Below is a listing of the unique domain organisations or architectures in which this domain is found. [More...](#)

There are 393 sequences with the following architecture: SMC_Nse1, zf-RING-like

[W9YTD0_9EURO](#) [Capronia epimyces CBS 606.96] Uncharacterized protein {ECO:0000313|EMBL:EXJ92910.1} (323 residues)



[Show](#) all sequences with this architecture.

There are 102 sequences with the following architecture: SMC_Nse1

[R1GGR5_BOTPV](#) [Botryosphaeria parva (strain UCR-NP2) (Grapevine canker fungus) (Neofusicoccum parvum)] Putative dna repair protein {ECO:0000313|EMBL:EOD47456.1} (255 residues)



[Show](#) all sequences with this architecture.

There are 6 sequences with the following architecture: SMC_Nse1 x 2, zf-RING-like

[NSE1_XENTR](#) [Xenopus tropicalis (Western clawed frog) (Silurana tropicalis)] Non-structural maintenance of chromosomes element 1 homolog EC=6.3.2.- (270 residues)



[Show](#) all sequences with this architecture.

There are 2 sequences with the following architecture: DAO, SMC_Nse1, zf-RING-like

[B8MNY1_TALSN](#) [Talaromyces stipitatus (strain ATCC 10500 / CBS 375.48 / QM 6759 / NRRL 1006) (Penicillium stipitatum)] FAD dependent oxidoreductase superfamily {ECO:0000313|EMBL:FFD14336.1} (744 residues)

A1CCH2 ASPCL/14-216	NRAFLQAFM	ARSTMTFAEAKPVLAALIF.SAH.....EG.....QPVSA...DDVTE
A2Q7K6 ASPNC/15-218	NRAFLQAFM	ARSTMTFTQAKPVLAALIF.SIR.....DD.....EQVSP...EDITE
B8NLA5 ASPFN/14-216	NRAFLQAFM	ARSTMTFAEARPVLAALIF.SVH.....EG.....EPVSA...EDVTE
B6QTR9 TALMQ/14-217	NRAFLQAFM	ARSTMTFDEAKPVLAALIF.SAQ.....EN.....REVLA...EDITQ
V5FED6 BYSSN/14-236	NRAFLQAFM	ARSTMTFEEAKPVLAALIF.SAHGAQSTIFFDSSYVEKEIANLYLATER.....RPVLA...EDITQ
S7Z8E9 PEN01/8-209	HRAFLQAFM	ARSTMTFEDAQPVLAAII.SAH.....EG.....RTVDP...DEVVQ
B6H9Q9 PENRW/8-210	NRAFLQAFM	ARSCMTFEDAQPILAAIL.TVS.....EG.....RTVDP...DEVGE
H6C926 EXODN/14-207	NRAFLQAFM	ARSVLTLETAKPILAAIS.TFQ.....DG.....REVQP...QDMTV
U1GD89 ENDPV/11-202	NRAFLQAFM	ARGTLTYETSKPLLASIF.TVH.....EG.....REILP...NDITE
C5GY37 AJEDR/11-207	HRAFLQAFM	ARSTMTYEQAKPVLAALIF.SAR.....DH.....QDTLP...EDITQ
C6H5E2 AJECH/11-203	HRALLQAFM	ARSTMTYEQAKPVLAALIF.TAR.....DN.....QETLP...EDITQ
F2PT91 TRIEC/10-199	HRAFLQAFM	SRSTMTLEEAKPVLAALIF.TVS.....EG.....REILP...GEDITQ
E9DEJ9 COCPS/13-202	HRAFLQAFM	ARSTMTLNEAKPILAAIL.SVK.....DG.....REVLP...EDVTQ
R7Z157 CONA1/13-202	HRAFLQAFM	ARSVLTFEEAQPILAAIL.TAH.....EG.....RPTLP...ADITT
U4LU38 PYROM/8-205	HRALVQAFM	ARSSMTGEEELGVTAIH.GVE.....NPEEP...TETTL
S8AAF4 DACHA/9-208	HRTFLQALL	IRPFIDIEEGQELLAALIA.SAE.....SG.....TDVPA...NSITV
G1X2Y0 ARTOA/8-207	HRAFLQALL	IRPFIDVQEGRELLAAIK.SAE.....AG.....SDVSI...ESVPP
C5DCF6 LACTC/23-226	BKFLLYVLL	RRGVCSEKALAKALKTL.....ERDG.....EQLEDSETE
C5DQF6 ZYGR/37-240	ARYLLQYLL	CRGICHENMLLVLDKL.....QK.....YTQDPTSQVCS.T
I2H1A9 TETBL/24-252	.RHLLRYIM	SEGICHENMLLLALYAL.....NLDYSGDCQEVLA.....
G8C139 TETPH/22-276	.RYLLQYLL	CGICNESFMLLALMRLH.....LTDGSSSDDFYRLNMVELN
G8ZVJ6 TORDC/16-236	.RIVLQYIL	SRGIVGENILLALMRL.....QGDT..ETND...VQELY
H2ATC7 KAZAF/7-237	SRYLLQYIL	ARGICHENALVVALMRL.....YMDL..GCFDDAWQIDQWL
J5S7B7 SACK1/17-253	TRYLLQYIL	ARGICHENNLILALMKL.....ETDA.....SKWSTEQWT
E7NKI9 YEASO/18-258	AKYLLQYIL	ARGICHENALILALMRL.....ETDA..STLNTEXSIQQWV
J8PZG2 SACAR/18-258	AKYLLQYIL	ARGICHENALILALMRL.....ETDA..LRFDAERSMQQWI
G0W7Y8 NAUDC/22-268	SRYLLQYIL	ARGVCHENALVVALMRL.....KVDS..NTIDPQWTISDWL
G0V5G7 NAUCC/25-272	SRYLLQYIL	ARGVCHENALMLALMRL.....KVDS..HDVNAHWTISDWL
J7S9R2 KAZNA/85-303	.RYLLQYIL	ARGVCHENSLLAALMHL.....SLDLDPHSVSTR.SLDEWS
A7TJ64 VANPO/8-260	SRYLLQYLL	CRGICHENALLLVLIKL.....DKDSVDEEGSPNRTFEDYL

Konzervované AMK svědčí o důležitosti jejich funkce:

- důležité pro proteinovou strukturu
- důležité pro funkci proteinu:
 - enzymy – aktivní centra
 - komplexy – PPI
 - regulační funkce – AMK posttranslačně modifikovaná

A1CCH2_ASPCL/14-216
 A2Q7K6_ASPNC/15-218
 B8NLA5_ASPFN/14-216
 B6QTR9_TALMQ/14-217
 V5FED6_BYSSN/14-236
 S7Z8E9_PEN01/8-209
 B6H9Q9_PENRW/8-210
 H6C926_EXODN/14-207
 U1GD89_ENDPU/11-202
 C5GY37_AJEDR/11-207
 C6H5E2_AJECH/11-203
 F2PT91_TRIEC/10-199
 E9DEJ9_COCPS/13-202
 R7Z157_CONA1/13-202
 U4LU38_PYROM/8-205
 S8AAF4_DACHA/9-208
 G1X2Y0_ARTOA/8-207
 C5DCF6_LACTC/23-226

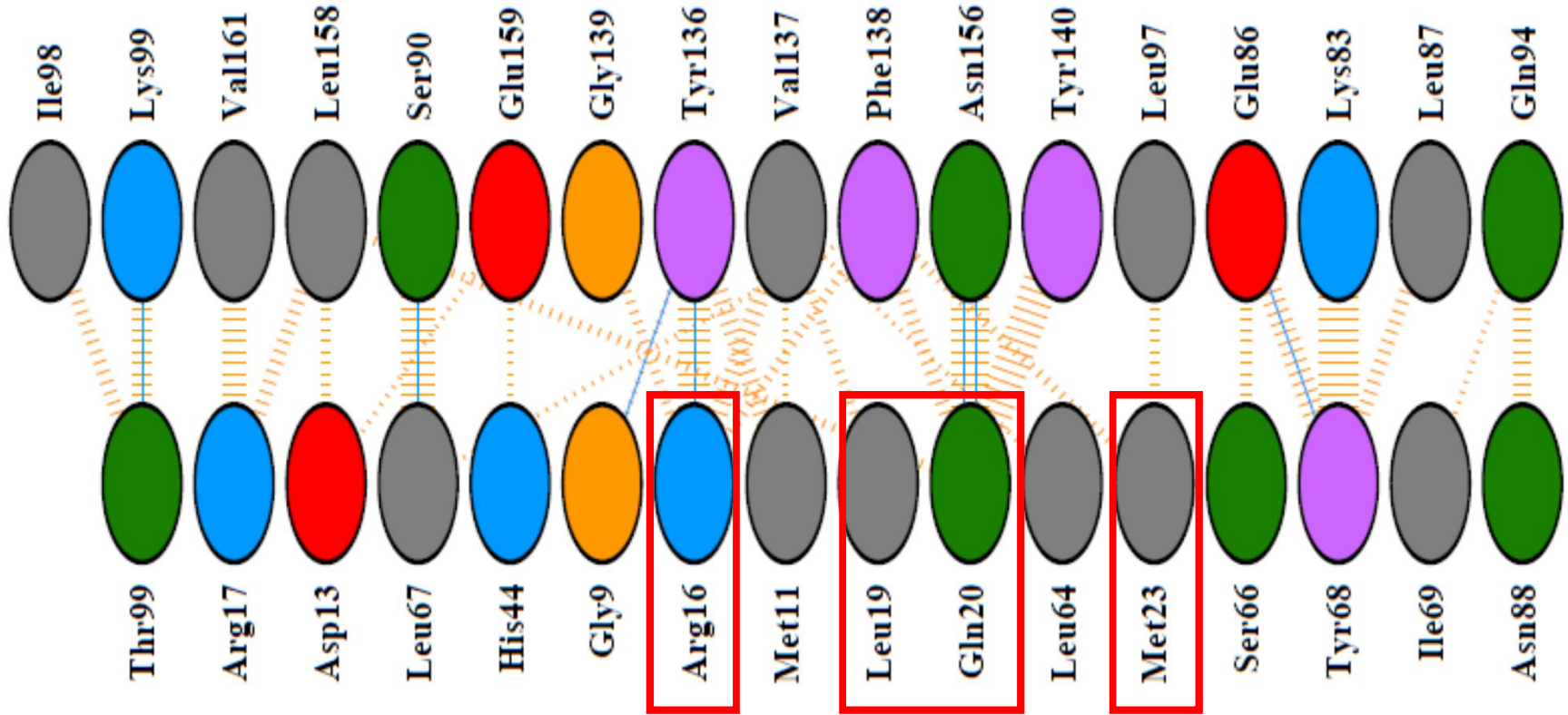
```

VRAFLQAFM ARSTMTFAEAKPVLAALIF.SAH.....EG.....QPVSA...DDVTE
VRAFLQAFM ARSTMTFTQAKPVLAALIF.SIR.....DD.....EQVSP...EDITE
VRAFLQAFM ARSTMTFAEARPVLAALIF.SVH.....EG.....EPVSA...EDVTE
VRAFLQAFM ARSTMTFDEAKPVLAALIF.SAQ.....EN.....REVLA...EDITQ
VRAFLQAFM ARSTMTFEEAKPVLAALIF.SAHGAQSTIFFDSSYVEKEIANLYLATER.....RPVLA...EDITQ
HRAFLQAFM ARSTMTFEDAQPVLAAII.SAH.....EG.....RTVDP...DEVTE
VRAFLQAFM ARSCMTFEDAQPILAAIL.TVS.....EG.....RTVDP...DEVGE
VRAFLQAFM ARSVLTLETAKPILAAIS.TFQ.....DG.....REVQP...QDMTV
VRAFLQAFM ARGTLTYETSKPLLASIF.TVH.....EG.....REILP...NDITE
HRAFLQAFM ARSTMTYEQAKPVLAALIF.SAR.....DH.....QDTLP...EDITQ
HRAFLQAFM ARSTMTYEQAKPVLAALIF.TAR.....DN.....QETLP...EDITQ
HRAFLQAFM SRSTMTLEEAKPVLAALIF.TVS.....EG.....REILP...GEDITQ
HRAFLQAFM ARSTMTLNEAKPILAAIL.SVK.....DG.....REVLP...EDVTQ
VRAFLQAFM ARSVLTFEEAQPILAAIL.TAH.....EG.....RPTLP...ADITT
HRAFLQAFM ARSSMTGEEELGVTVAIH.GVE.....NPEEP...TETTL
HRTFLQALL IRPFIDIEEGQELLAALIA.SAE.....SG.....TDVPA...NSITV
HRAFLQALL IRPFIDVQEGRELLAAIK.SAE.....AG.....SDVSI...ESVPP
GKFLLYVLR RRGVCSEKALAKALKTL.....ERDG.....EQLEDSETE
  
```



Chain B

Chain A



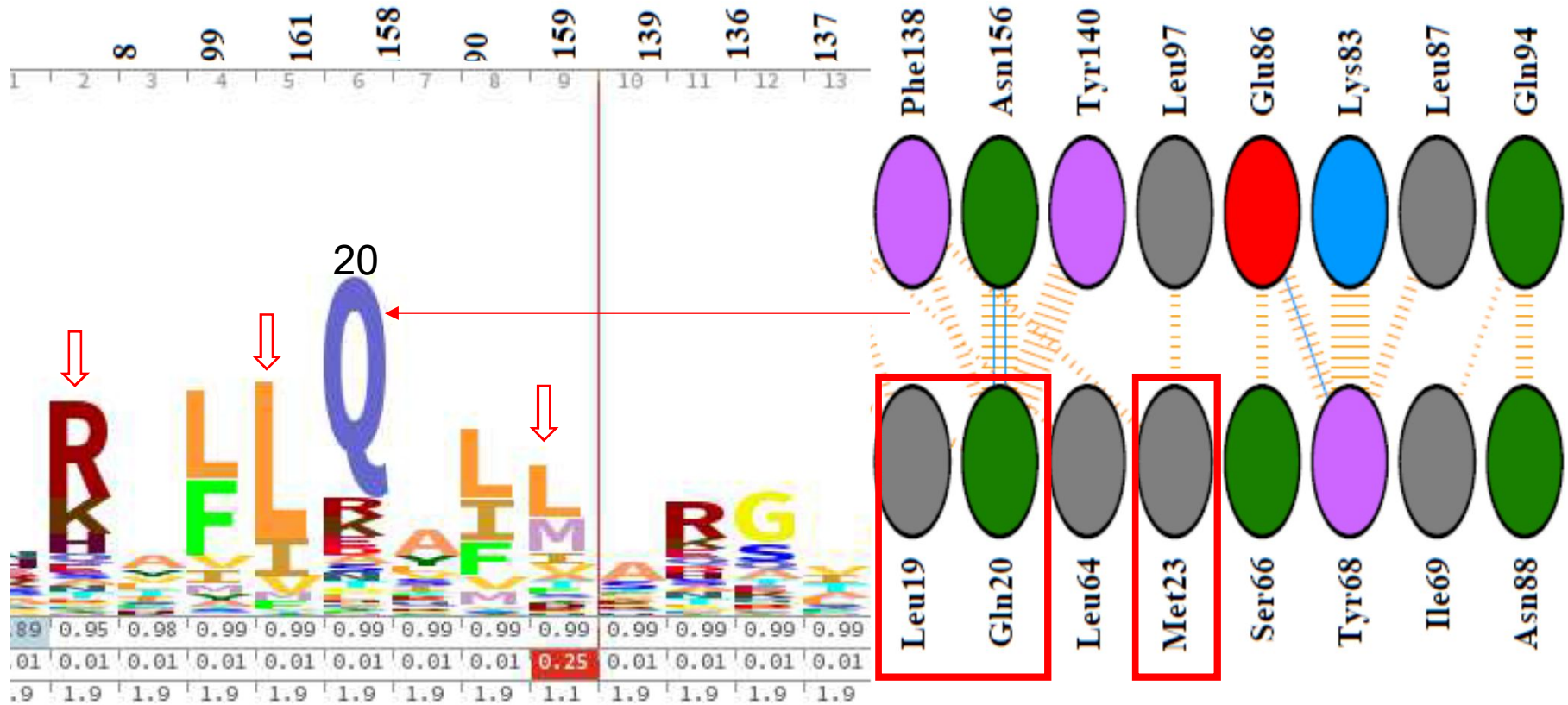
```

A1CCH2 ASPCL/14-216 NRAFLQAFM ARSTMTFAEAKPVLAALIF.SAH.....EG.....QPVSA...DDVTE
A2Q7K6 ASPNC/15-218 NRAFLQAFM ARSTMTFTQAKPVLAALIF.SIR.....DD.....EQVSP...EDITE
B8NLA5 ASPFN/14-216 NRAFLQAFM ARSTMTFAEARPVLAALIF.SVH.....EG.....EPVSA...EDVTE
B6QTR9 TALMQ/14-217 NRAFLQAFM ARSTMTFDEAKPVLAALIF.SAQ.....EN.....REVLA...EDITQ
V5FED6 BYSSN/14-236 NRAFLQAFM ARSTMTFEEAKPVLAALIF.SAHGAQSTIFFDSSYVEKEIANLYLATER.....RPVLA...EDITQ
S7Z8E9 PEN01/8-209 HRAFLQAFM ARSTMTFEDAQPVLAAII.SAH.....EG.....RTVDP...DEVTE
B6H9Q9 PENRW/8-210 NRAFLQAFM ARSCMTFEDAQPILAAIL.TVS.....EG.....RTVDP...DEVEGE
H6C926 EXODN/14-207 NRAFLQAFM ARSVLTLETAKPILAAIS.TFQ.....DG.....REVQP...QDMTV
U1GD89 ENDPV/11-202 NRAFVQAFM ARGTLTYETSKPLLASIF.TVH.....EG.....REILP...NDITE
C5GY37 AJEDR/11-207 HRAFLQAFM ARSTMTYEQAKPVLAALIF.SAR.....DH.....QDTLP...EDITQ
C6H5E2 AJECH/11-203 HRALLQAFM ARSTMTYEQAKPVLAALIF.TAR.....DN.....QETLP...EDITQ
F2PT91 TRIEC/10-199 HRAFLQAFM SRSTMTLEEAKPVLAALIF.TVS.....EG.....REILP...GEDITQ
E9DEJ9 COCPS/13-202 HRAFLQAFM ARSTMTLNEAKPILAAIL.SVK.....DG.....REVLP...EDVTQ
R7Z157 CONA1/13-202 HRAFLQAFM ARSVLTFEEAQPILAAIL.TAH.....EG.....RPTLP...ADITT
U4LU38 PYROM/8-205 HRALVQAFM ARSSMTGEEELGVTVAIH.GVE.....NPEEP...TETTL
S8AAF4 DACHA/9-208 HRTFLQALL IRPFIDIEEGQELLAIA.SAE.....SG.....TDVPA...NSITV
G1X2Y0 ARTOA/8-207 HRAFLQALL IRPFIDVQEGRELLAAIK.SAE.....AG.....SDVSI...ESVPP
C5DCF6 LACTC/23-226 GKFLLOYVL RRGVCSEKALAKAKTL.....ERDG.....EQLEDSETE

```

NRAFLQAFM
ARSTMTFAEAKPVLAALIF.SAH
.....EG
.....QPVSA
...DDVTE

↑
↑
↑





Structures Overview
Structure Selections
3D View Controls

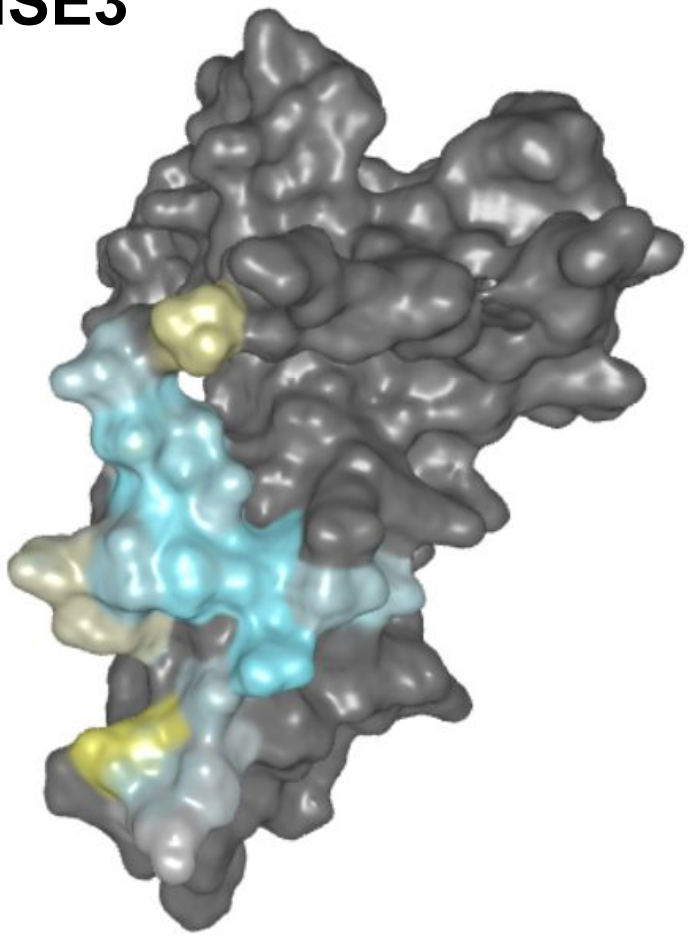
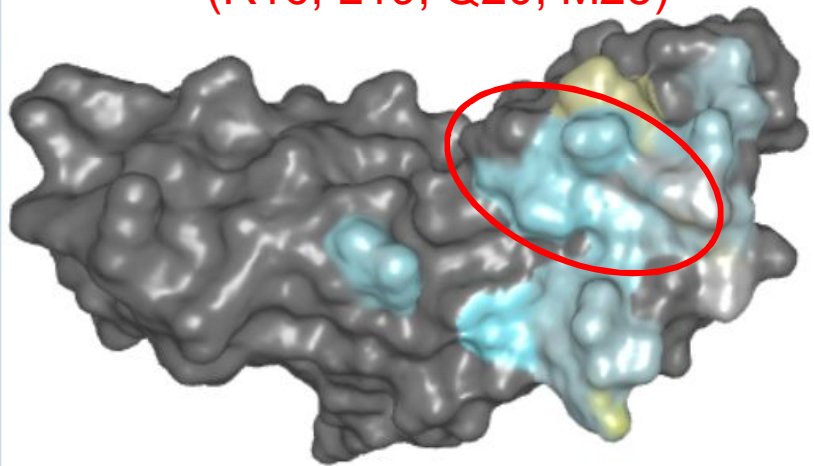


Residue Matrix

Primary Structure: :
Sort by: Conservati

NSE1 - NSE3

Hot spot
(R16, L19, Q20, M23)



N88
Q20
M23
T99
L64
L19
D13
I69
L67
Y68
H44
S66
M11
R16
G9
R17

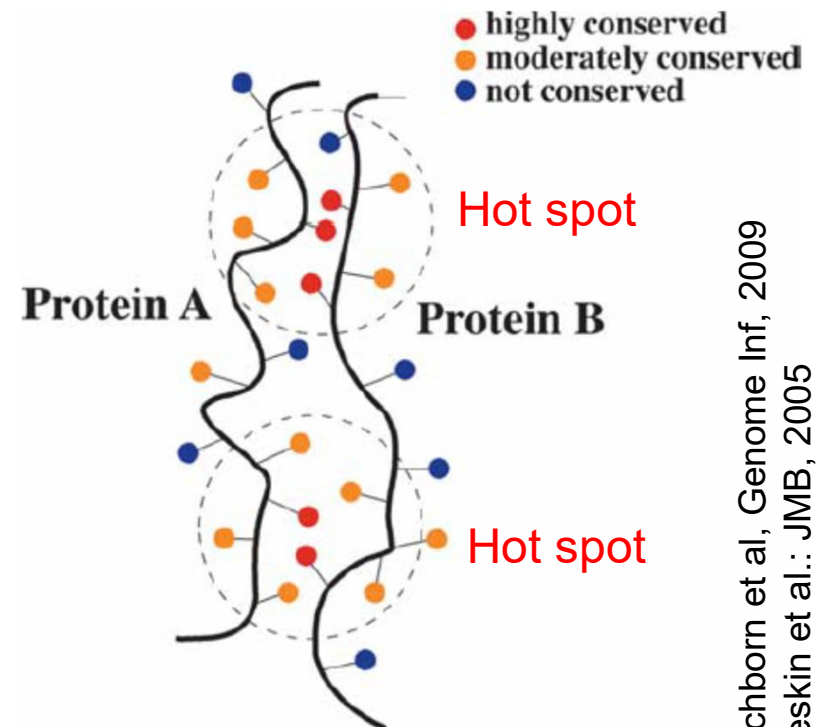
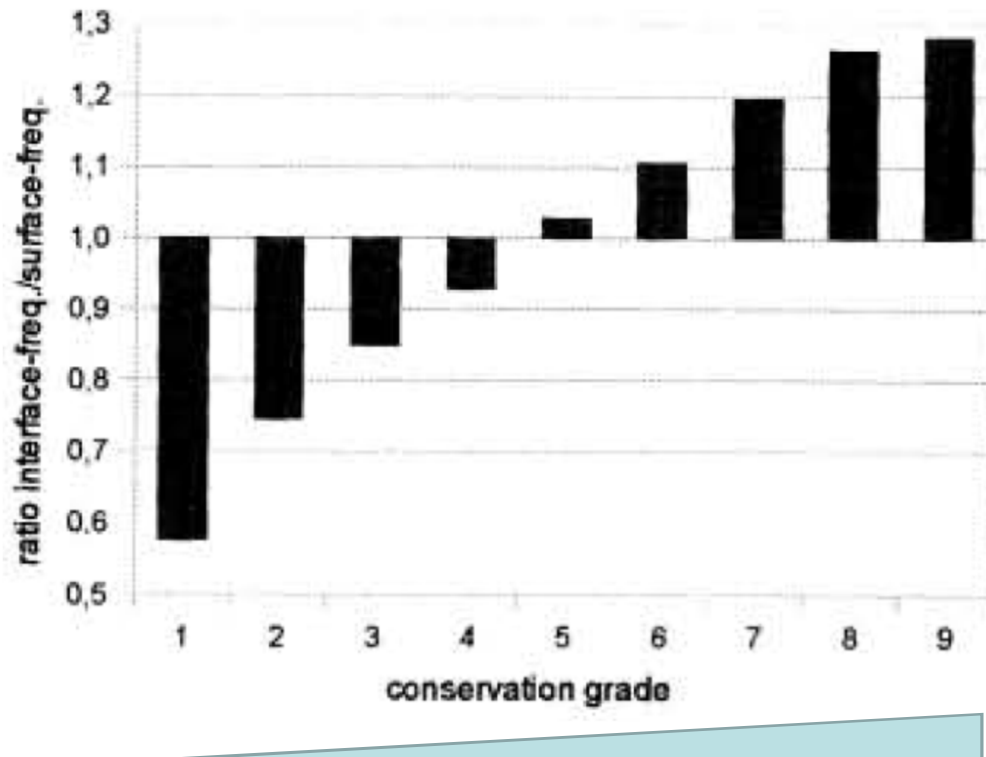


Structure Sequence x

Compact View Selection in All Structures

Silné/důležité interakce (komplexy) jsou evolučně konzervované

- jako jsou proteiny (jejich funkce) evolučně konzervované, tak i jejich interakce jsou evolučně konzervované (zajišťují funkci)
- graf** – povrchové AMK jsou málo konzervované (grade1), zatímco interakční povrchy jsou hodně konzervované (grade9)





Search Results

Gene / Identifier Search

hse3



GO

All Organisms

Your search for **NSE3** produced the following **4** results:Results matching **official symbol / systematic name** - 2 total proteins:**NSE3 (YDR288W)**

Component of the SMC5-SMC6 complex; this complex plays a key role in the removal of X-shaped DNA structures that arise between sister chromatids during DNA replication and repair; protein abundance increases in response to DNA replication stress

PHO*Saccharomyces cerevisiae* (S288c)

407 unique interactors

497 raw interactions

1 post-translational modification

NSE3 (SPCC645.04)

Smc5-6 complex non-SMC subunit Nse3

Schizosaccharomyces pombe (972h)

10 unique interactors

24 raw interactions

NSE1 | YLR007W

Component of the SMC5-SMC6 complex; this complex plays a key role in the removal of X-shaped DNA structures that arise between sister chromatids during DNA replication and repair

5 2

[\[details\]](#)

Experimental Evidence Code	Role	Dataset	Throughput	Curated By	Notes
Affinity Capture-MS	HIT	Hazbun TR (2003)	High	BioGRID	-
Reconstituted Complex	HIT	Hudson JJ (2011)	Low	BioGRID	-
Two-hybrid	HIT	Hazbun TR (2003)	High	BioGRID	-
	HIT	Hu B (2005)	Low	BioGRID	-
	BAIT/HIT	Duan X (2009)	Low	BioGRID	-
Dosage Rescue	HIT	Magtanong L (2011)	High	BioGRID	
Negative Genetic	BAIT/HIT	Costanzo M (2016)	High	BioGRID	

BioGRID – databáze interakcí (včetně genetických) pro různé organismy pučící kvasinky *S. cerevisiae*, pohlivé kvasinky *S. pombe*, octomilky *D. melanogaster*, člověka *H. sapiens* ...

Kde najít další informace o PPI?

The screenshot shows a Windows Internet Explorer browser window displaying the website <http://proteome.wayne.edu/PIDBL.html>. The page is titled "Links to Protein Interaction Databases" and is part of the "Center for Molecular Medicine and Genetics" website. The page content includes:

- Finley Lab Interactions Databases:
 - Drosophila Interactions Database (DroID)***
 - Campylobacter jejuni Interactions Databases***
- Gene or Protein Interactions Databases in the research community:
 - BioGRID** - A Database of Genetic and Physical Interactions
 - DIP** - Database of Interacting Proteins
 - IntAct** - EMBL-EBI Protein Interaction Database
 - MINT** - A Molecular Interactions Database
 - MIPS** - Comprehensive Yeast Protein-Protein interactions
 - Yeast Protein Interactions** - Yeast two-hybrid results
 - BRITE** - Biomolecular Relations in Information Transmission
 - The PIM Database** - by Hybrigenics
 - Mouse Protein-Protein interactions**
 - Human Protein Reference Database**

Na základě PPI v jednom organismu a homologii proteinů v jiných organismech lze odhadnout, zda proteiny interagují i v jiných organismech (lze dovodit i podle genových fúzí)

<http://proteome.wayne.edu/PIDBL.html>

Více Dr. Potěšil

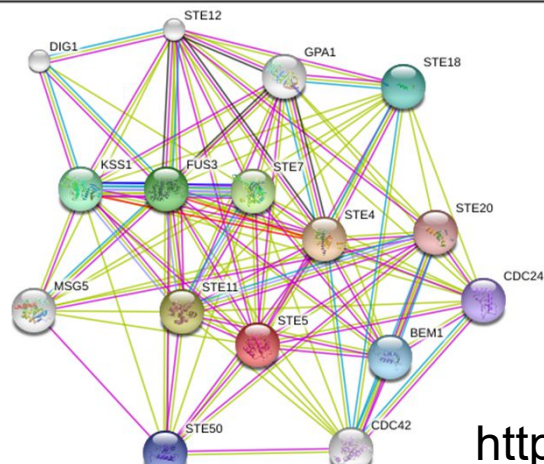
Informační zdroje PPI

Table 2. Databases Available for Searching and/or Downloading Data Related to Protein Interactions

Database	Proteins/Domains	Type	Number of Interactions
DIP ^a , LiveDIP	P	E,S	55,733
BIND ^a	P	E,C,S	83,517
MPact/MIPS ^a	P	E,C,F	15,488 (4,300) ^b
STRING	P ←	E,P,F	730,000 (proteins)
MINT ^a	P	E,C	71,854
IntAct ^a	P ←	E,C	68,165
BioGRID ^a	P ←	E,C	116,000 (30,000) ^b
HPRD	P	E,C	33,710
ProtCom	P,D	S,H	1,770
3did, Interprets	D	S,H	3,304
Pibase, ModBase	D	S,H	2,387
CBM	D	S	2,784
SCOPPI	D	S	3,358
iPfam	D	S	3,019
InterDom	D	P	30,037
DIMA	D	F,S	—
Prolinks	P	F	—

Table 3. URLs and Primary Citations for Protein Interaction-Related Databases

Database	URL/FTP
DIP [102], LiveDIP[103]	http://dip.doe-mbi.ucla.edu
BIND [105]	http://bind.ca
MPact/MIPS [97]	http://mips.gsf.de/services/ppi
STRING [119]	http://string.embl.de
MINT [120]	http://mint.bio.uniroma2.it/mint
IntAct [121]	http://www.ebi.ac.uk/intact
BioGRID [122]	http://www.thebiogrid.org
HPRD [123]	http://www.hprd.org
ProtCom [124]	http://www.ces.clemson.edu/compbio/ProtCom
3did [108], Interprets[125]	http://gatealoy.pcb.ub.es/3did/
Pibase [107], ModBase [126]	http://alto.compbio.ucsf.edu/pibase
CBM [26]	ftp://ftp.ncbi.nlm.nih.gov/pub/cbm
SCOPPI [111]	http://www.scoppi.org/
iPfam [127]	http://www.sanger.ac.uk/Software/Pfam/iPfam
InterDom [128]	http://interdom.lit.org.sg
DIMA [129]	http://mips.gsf.de/genre/proj/dima/index.html
Prolinks [104]	http://prolinks.doe-mbi.ucla.edu/cgi-bin/functionator/pronav/



STRING

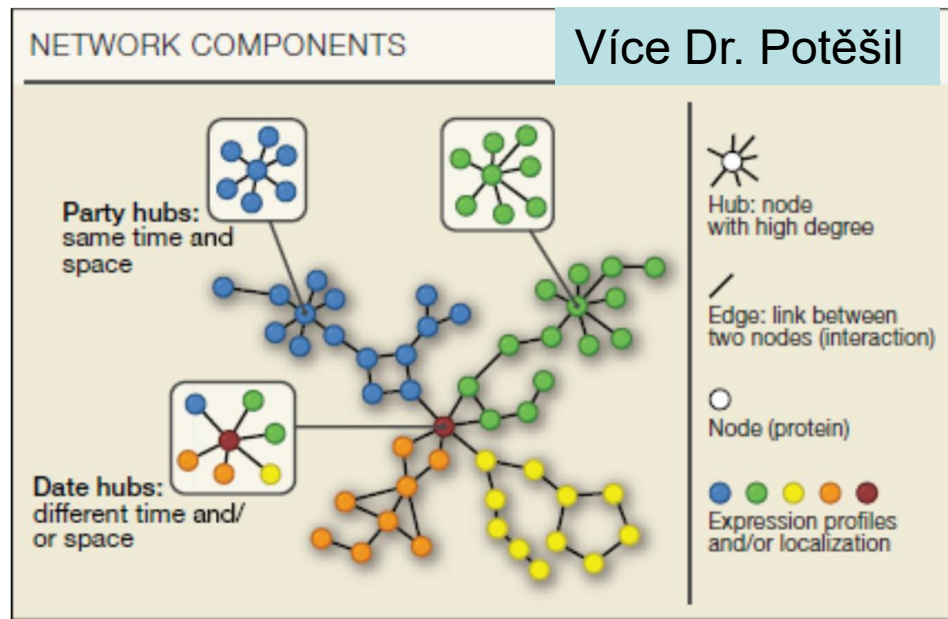
informace o binárních interakcích
v databázích – zobrazeny jako síť
(různé výsledky = různé čáry)

Shoemaker and Panchenko, PLoS Comp Biol, 2007

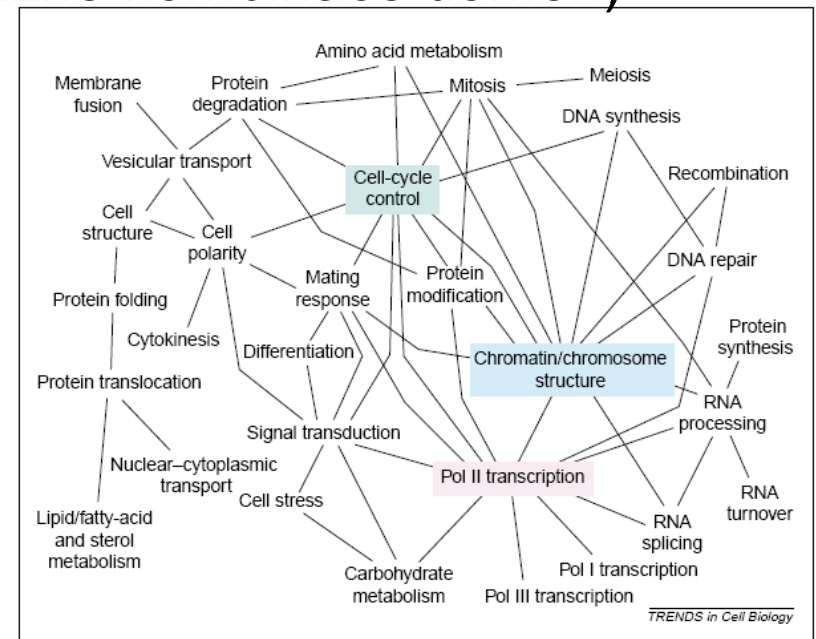
<http://string-db.org> Andreani and Guerois, ABB, 2014

Protein-proteinové interakce

- stabilní (velké plochy, většinou součástí komplexů)
- přechodné/slabé (součást dynamických procesů – předávání signálů, modifikace)
- posttranslační modifikace mohou změnit vazebné vlastnosti povrchu (fosforylace, metylace, hydroxylace, SUMO)
- souhrn proteinových interakcí = **interaktom**
(modularita díky interakcím domén – různé kombinace domén)

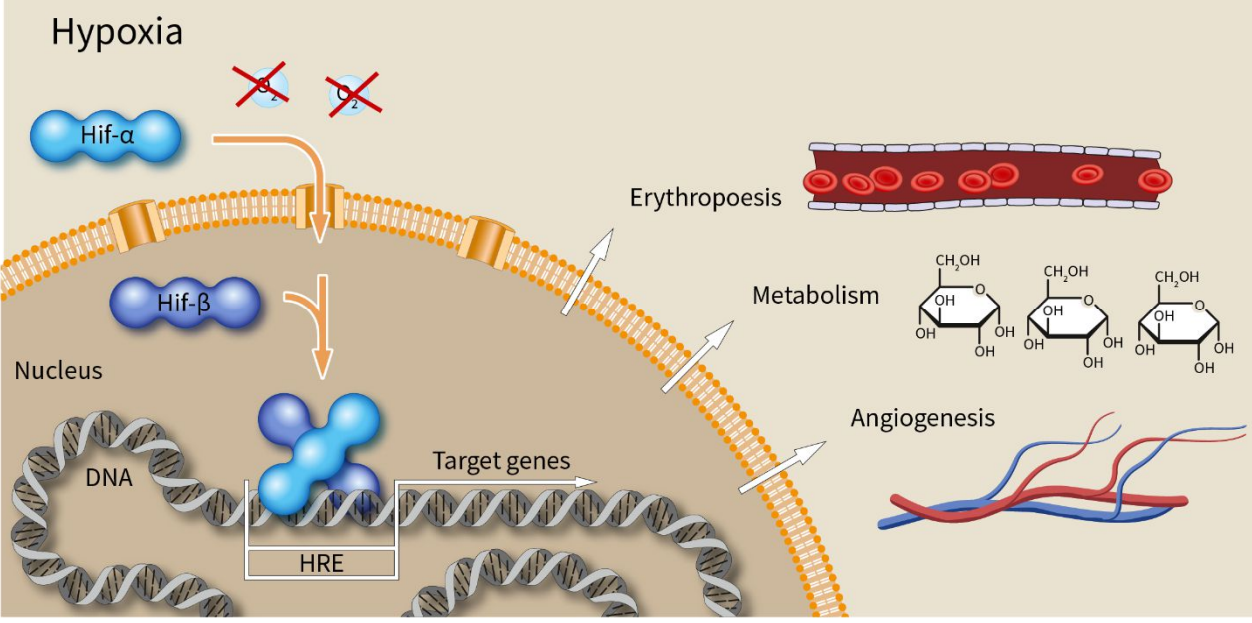
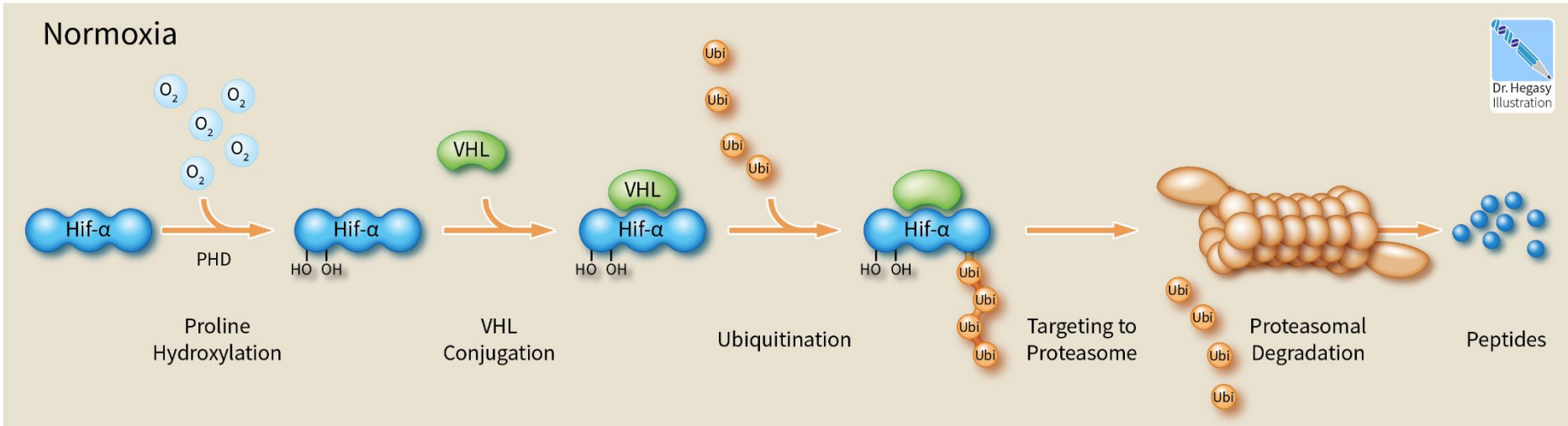


Seebacher & Gavin, Cell (SNAP SHOT), 2011



Network/síť naznačuje funkční vztahy
Tucker et al, TiCB, 2001

© Nobel Prize in Physiology or Medicine 2019: How Cells Sense and Adapt to Oxygen Availability
 Awarded to William G. Kaelin, Sir Peter J. Ratcliffe, and Gregg L. Semenza








-  Hypoxia-inducible factor alpha subunit
-  Hypoxia-inducible factor beta subunit = Aryl Hydrocarbon Receptor Nuclear Translocator, ARNT
- PHD Prolyl Hydroxylase Domain protein
-  von Hippel-Lindau protein
-  Ubiquitin
-  Hypoxia response element

Illustration: www.hegasy.de

Interaktom x komplexom

Figure 3-83 *Molecular Biology of the Cell* (© Garland Science 2008)

Naznačují funkční vztahy (např. buněčný cyklus – struktura chromatinu ... je zprostředkován PPIs)

Modularita – interagují domény (jeden protein více domén – zapojení do více procesů)

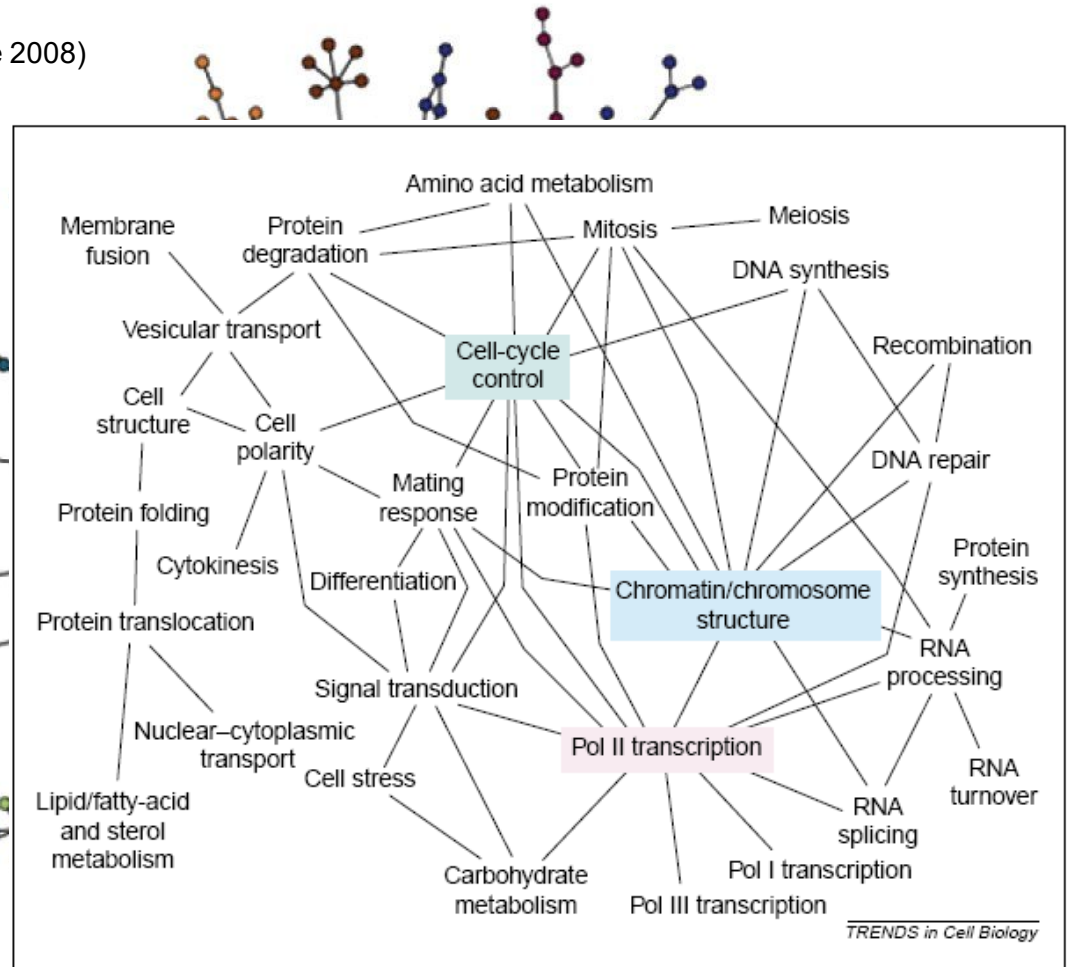
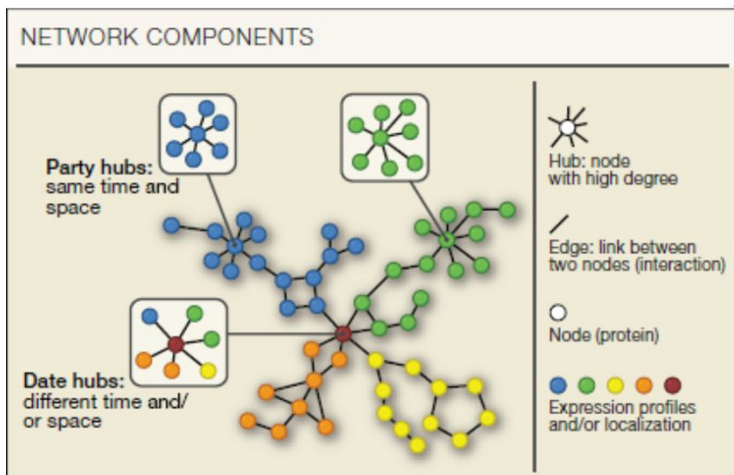
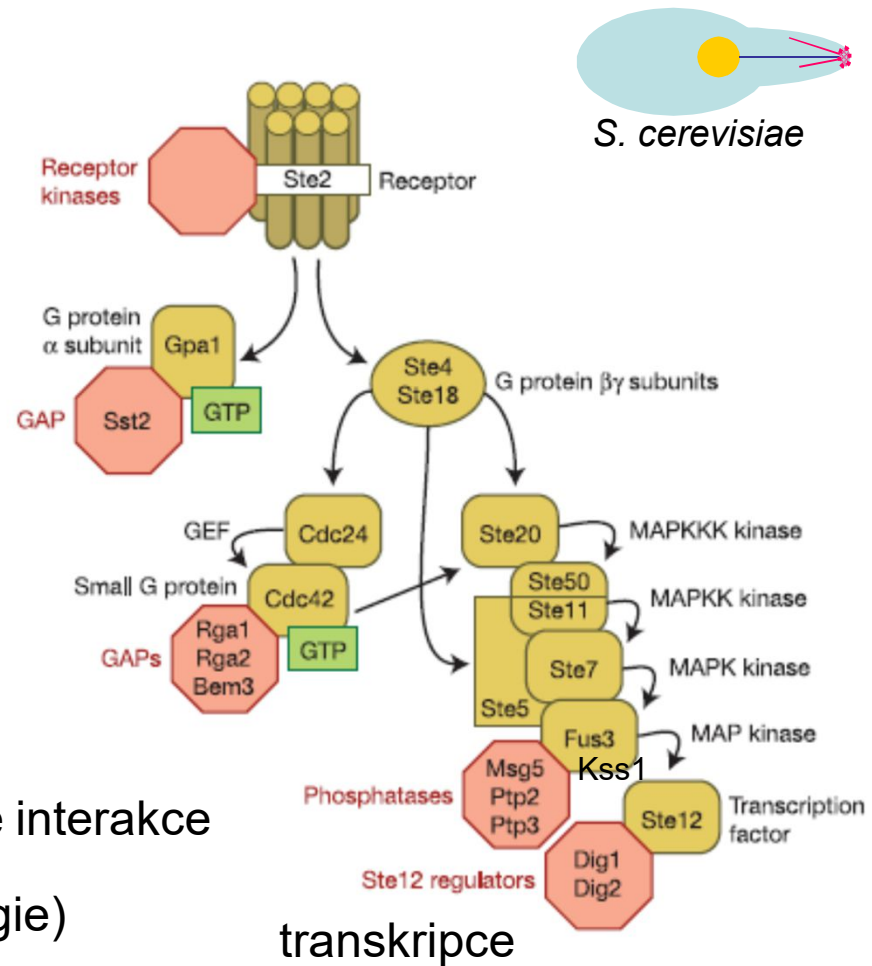
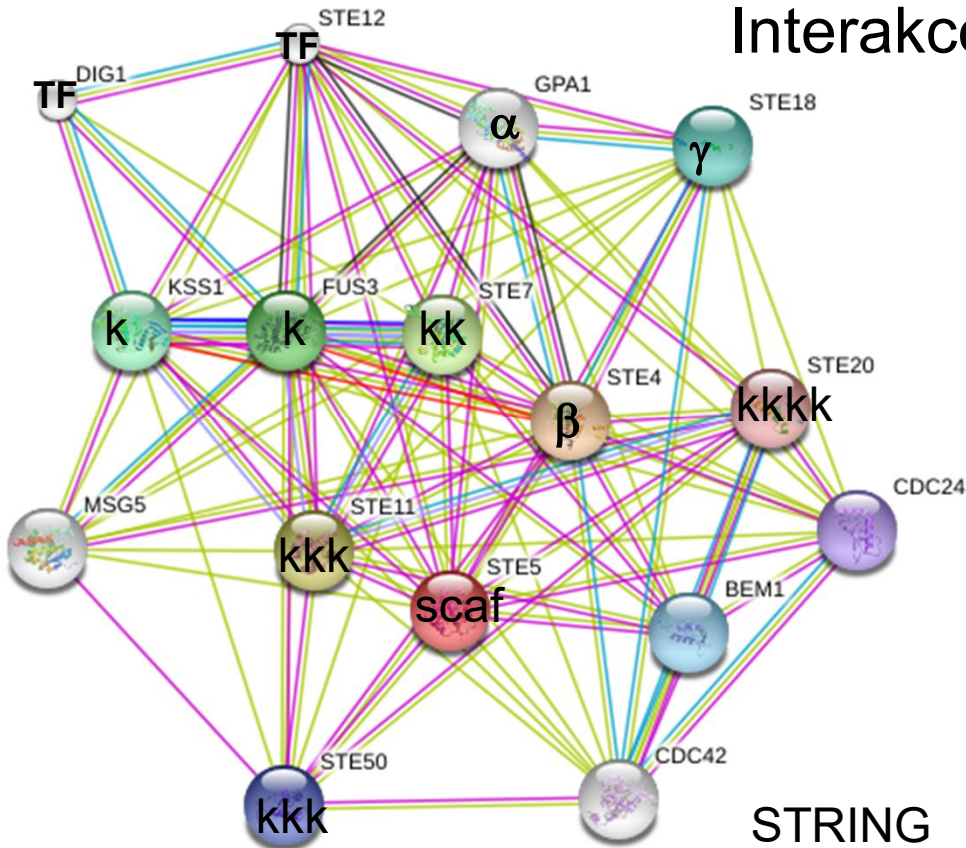


Fig. 2. Functional group interaction map based on Fig. 1 (modified from Ref. 10). Shown are interactions between functional groups of yeast proteins. Each line indicates that there are 15 or more interactions between proteins of the connected groups. Connections with fewer than 15 interactions are not shown because one or a few interactions occur between almost all groups and often tend to be spurious – that is, based on false positives in two-hybrid screens or other assays. Note that only proteins with known function are included and that about one-third of all yeast proteins belong to several classes.



proteinové sítě – chybí info o posloupnosti, síle ... interakcí

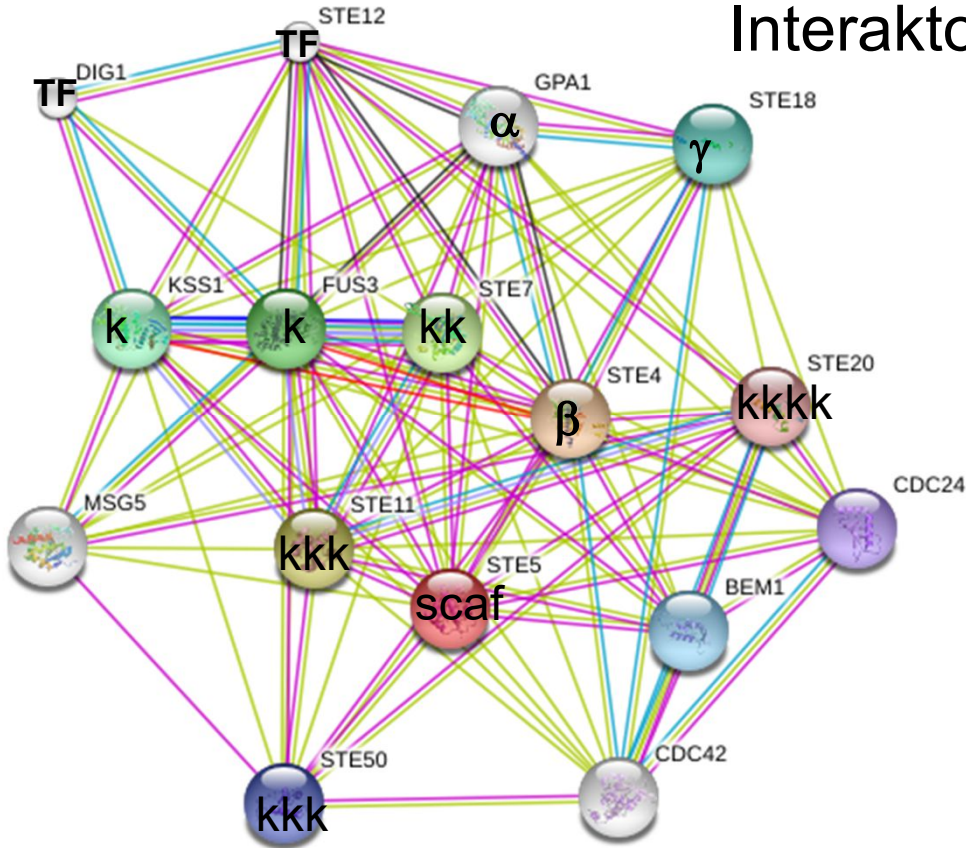
Interakce x signální dráha



- ▬ Experiments Y2H, colP ... genetické interakce
- ▬ Databases Funkční vztahy (ontologie)
- ▬ Textmining
- ▬ Gene Fusion Svědčí o potřebě PPI
- ▬ Coexpression Potřeba výskytu ve stejném okamžiku a společná translace

proteinové sítě – chybí info o lokalizaci, komplexech ...

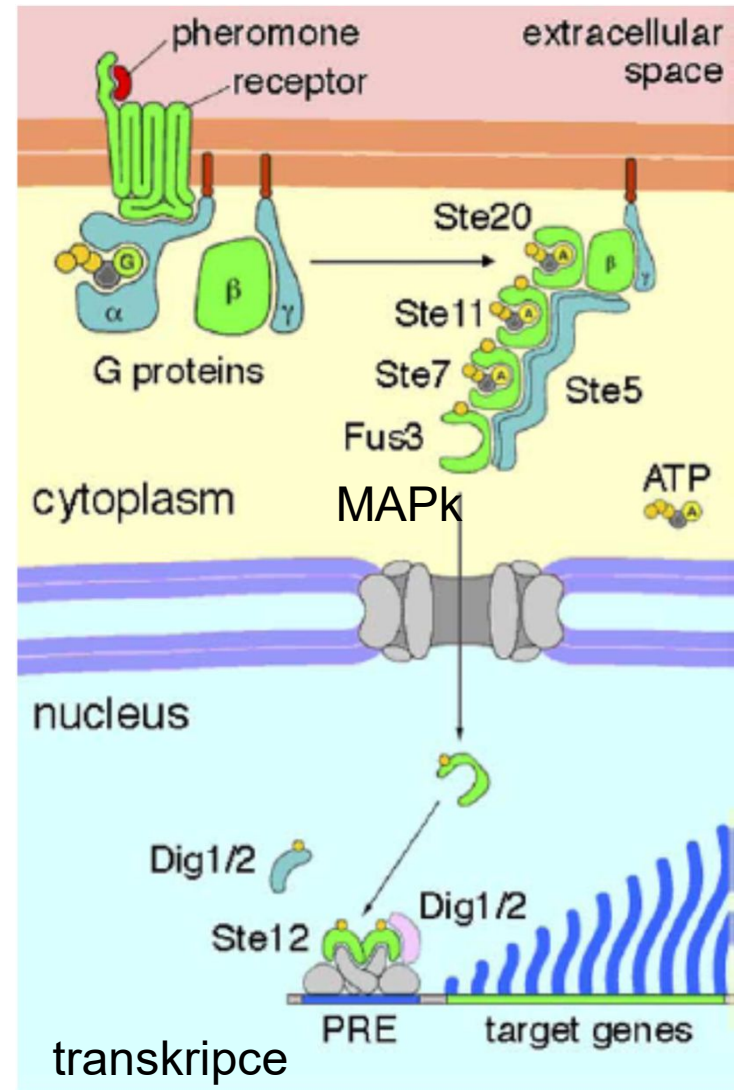
Interaktom x komplexom



- Experiments
- Databases
- Textmining
- Gene Fusion
- Coexpression

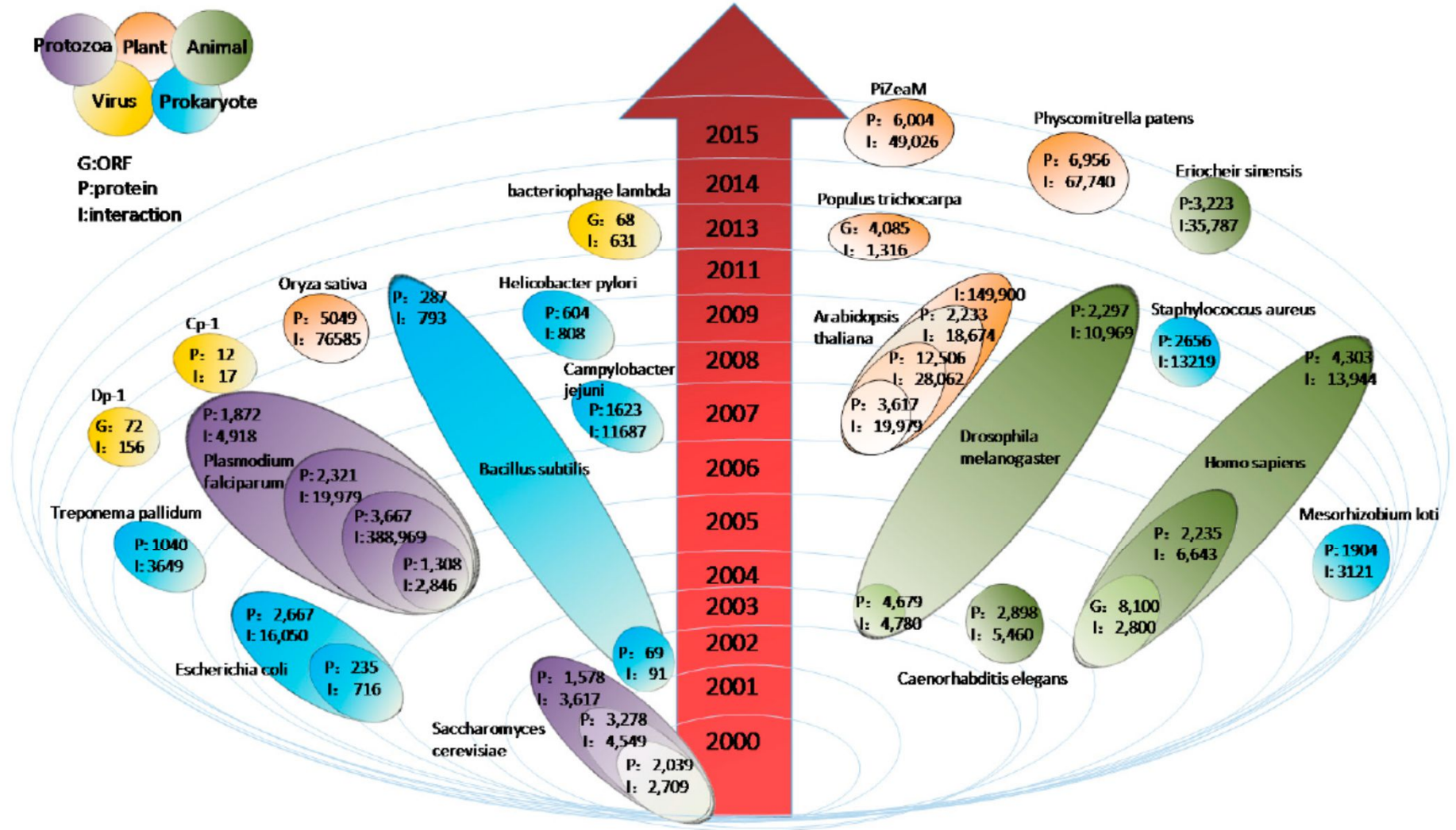
Síť neznamená komplex,
ale vztah

souhrn proteinových
komplexů = **komplexom**



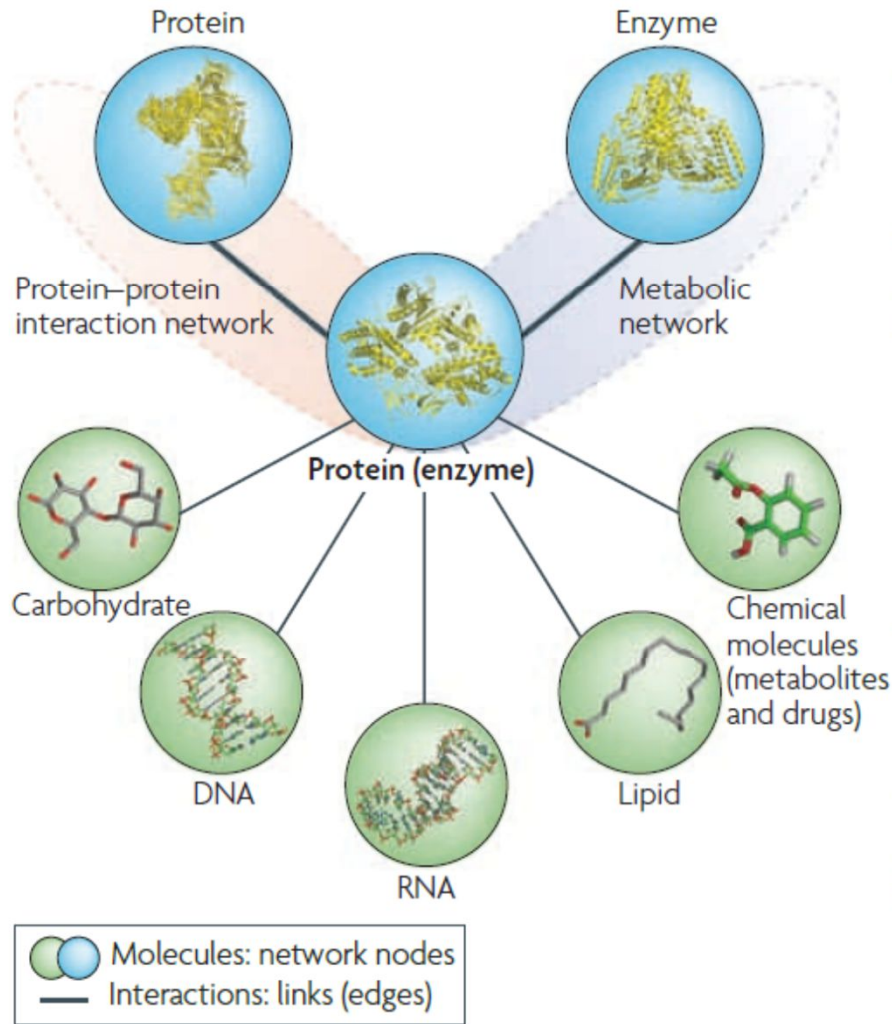
Wang et al., Nature, 2004

High-throughput screens – interaktomy organismů

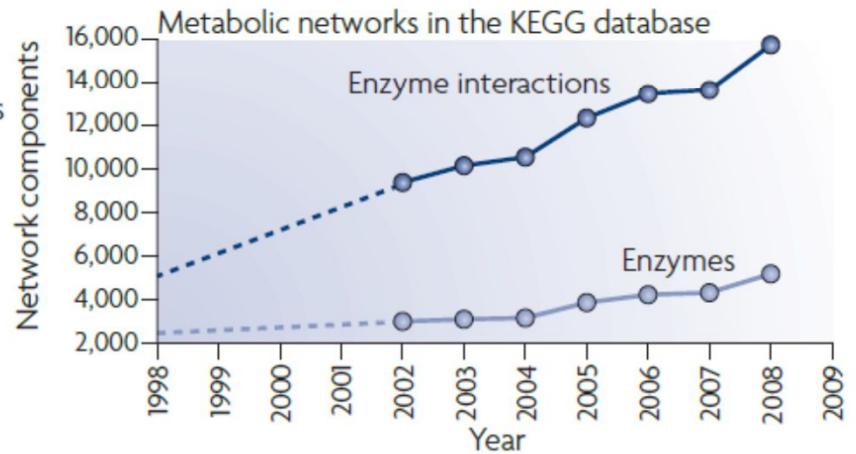
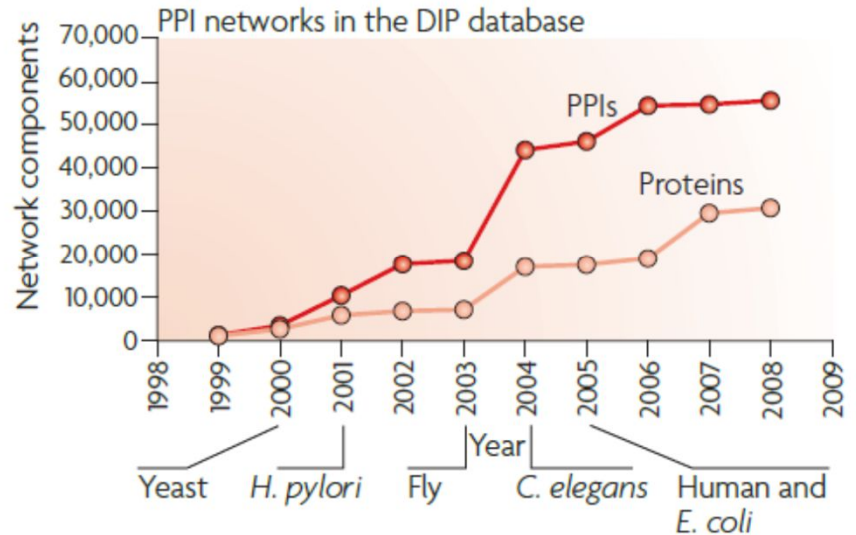


Kvasinkový dvoj-hybridní systém
 TAP-tag s MS analýzou

a Biomolecular network components

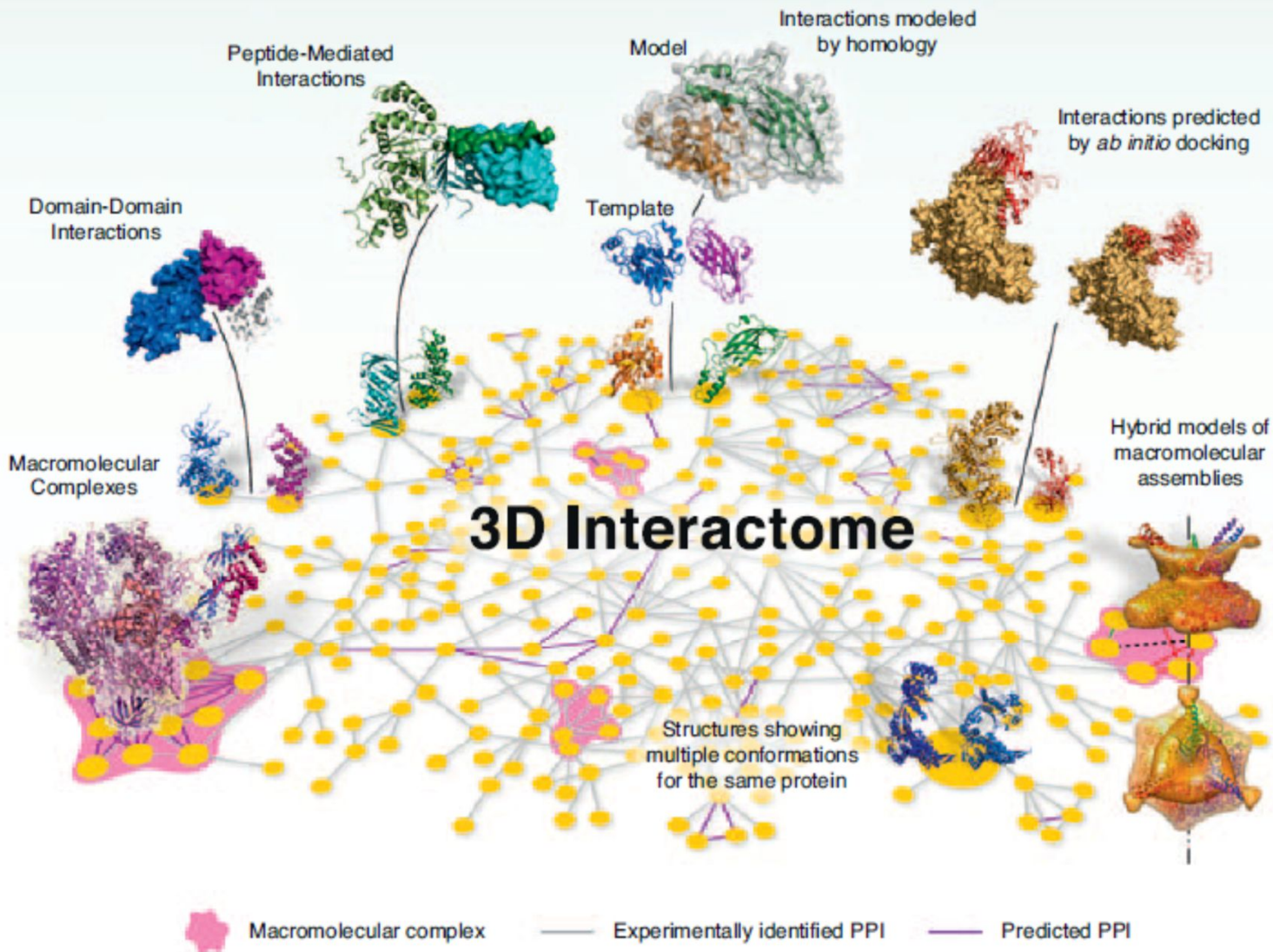


b Accumulation of network components over the past 10 years



Experimental Structures

Computational Models

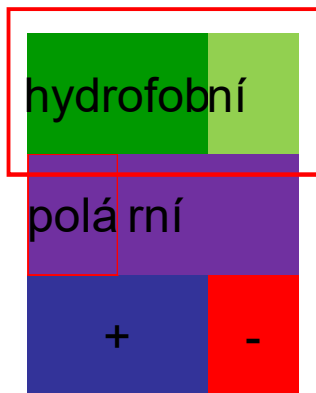


Current Opinion in Structural Biology

Souhrn - protein-proteinové interakce

- proteiny jsou troj-rozměrné - mají různé tvary a více domén => mají více vazebných míst na povrchu => komplexy a "sítě"
- části proteinů/domény/motivy interagují s partnery
 - domény mají určitou strukturu, která do značné míry determinuje tvar jejího povrchu, ale ...
 - charakter (hydrofobicitu, polaritu, náboj) povrchu určují postraní řetězce aminokyselin směřujících do solventu, takže ...
 - interakce proteinu je determinována povrchem, který musí mít tvar i charakter komplementární s interakčním partnerem (typy interakcí: ...)

primární struktura



sekundární a terciární struktura

