



## INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

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

- **Proteinové interakce – 24.10.**
  - 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**

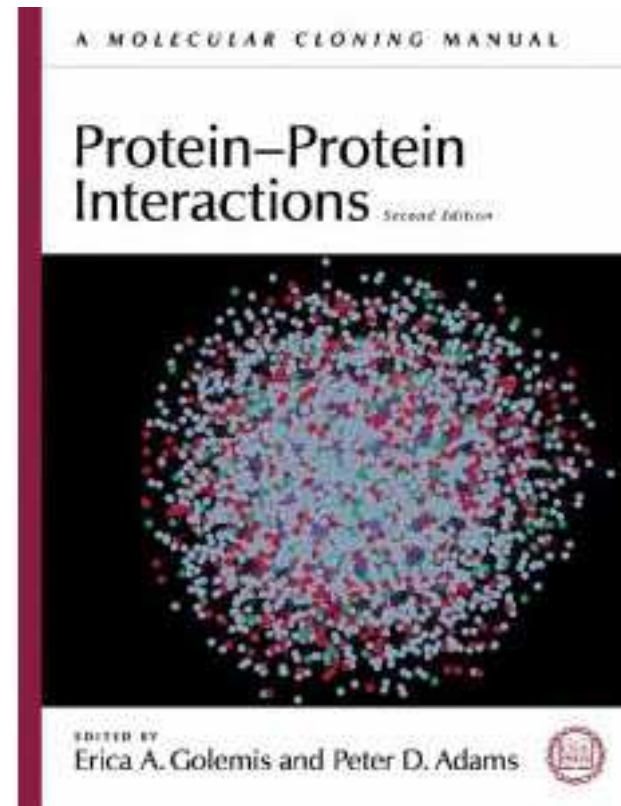
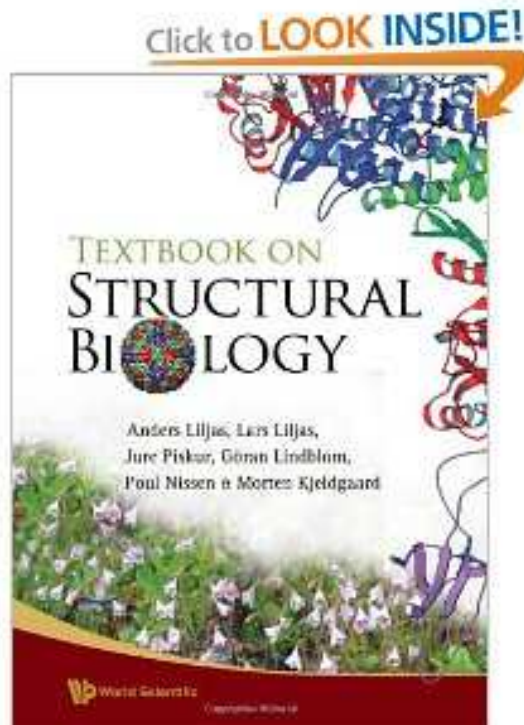
**CG031 – cvičení z modelování proteinových komplexů  
(jarní semestr)**

# 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 ...**

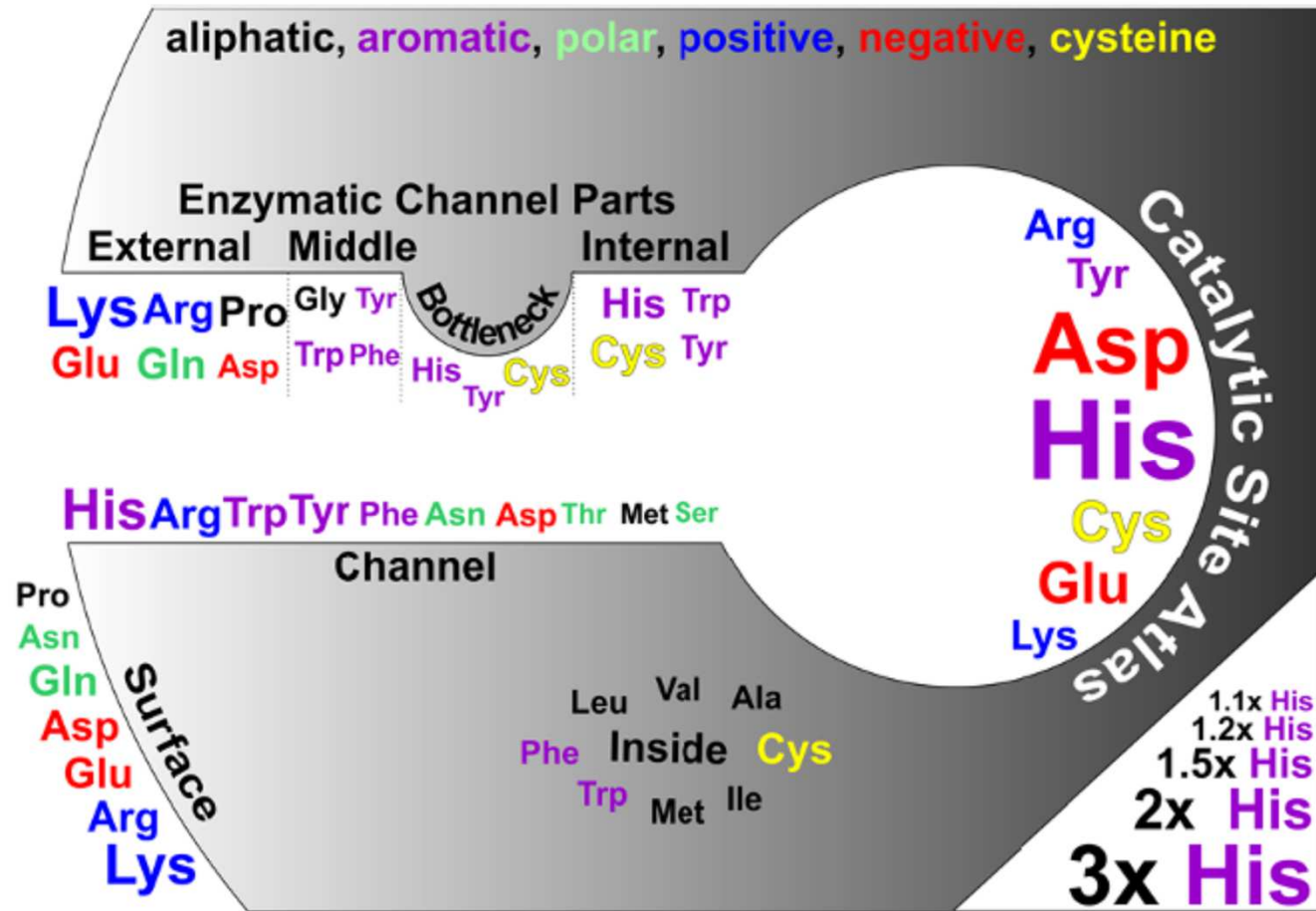


Databáze proteinových struktur: <http://www.rcsb.org/pdb/home/home.do>,  
<http://www.ebi.ac.uk/pdbsum/>

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

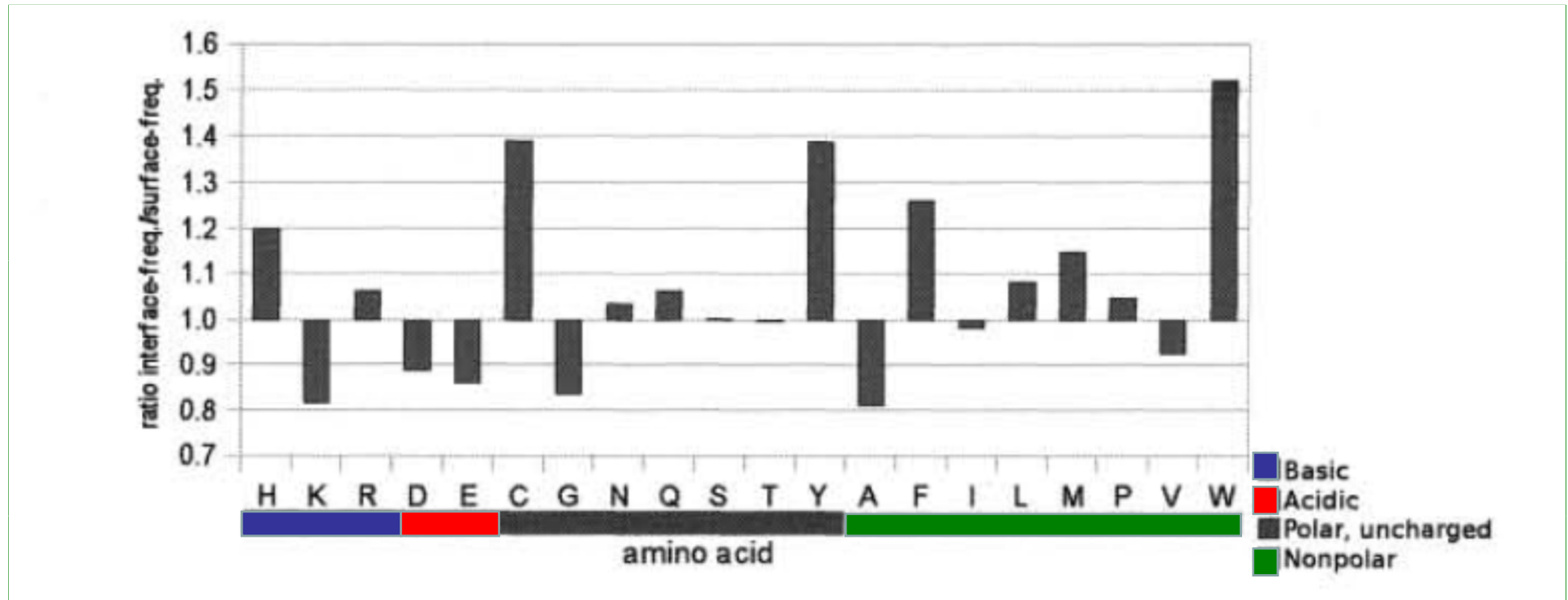


# Podíl AMK (primární struktury) na proteinových interakcích



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

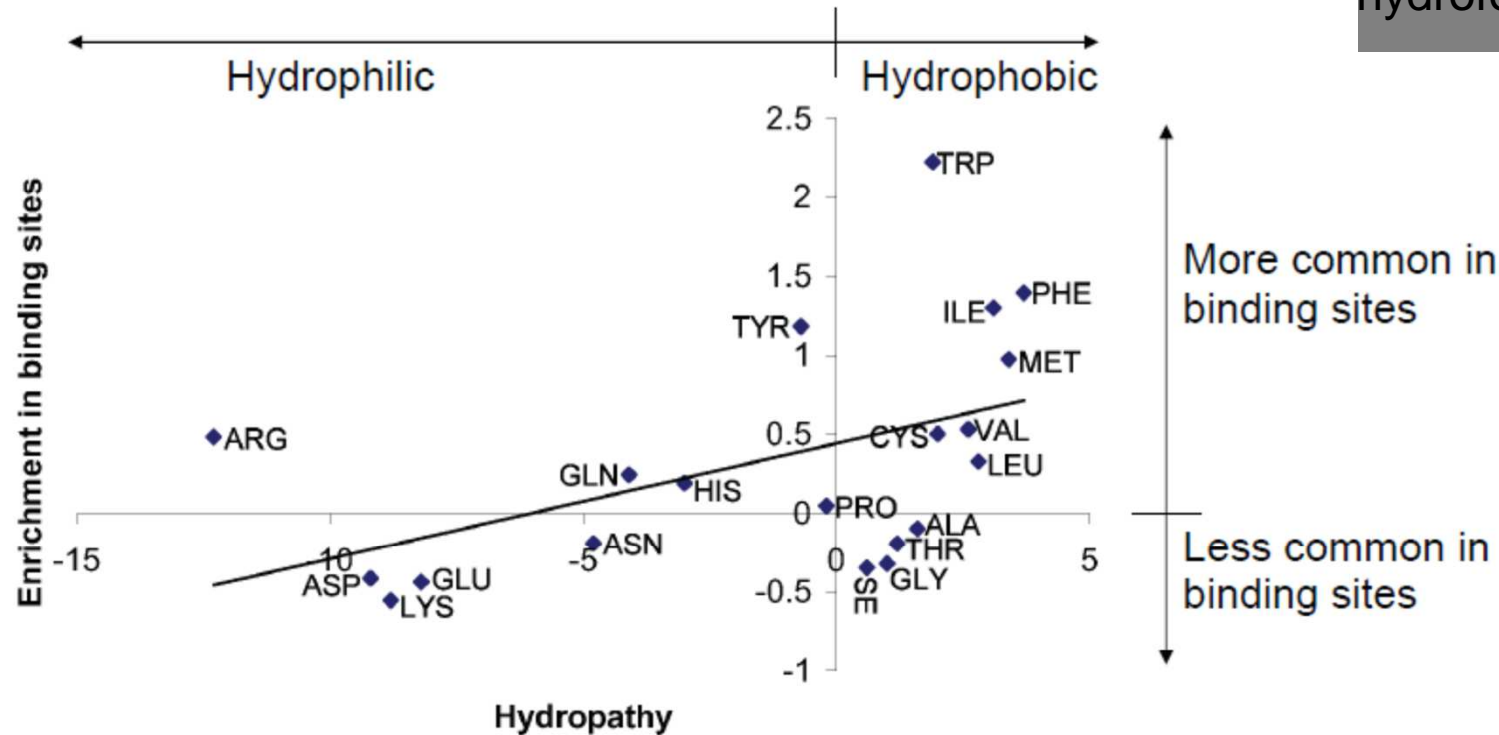
# Protein-proteinové interakce od primární struktury ...



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



# Protein-proteinové interakce od primární struktury ...



- **hydrofobní zbytky** jsou tlačeny dovnitř proteinu (ve vodném prostředí se chovají jako „mastnota“ ve vodě) – pro proteiny s hydrofobním povrchem je tedy „výhodnější“ se přes takový povrch navázat na stejně „mastného“ partnera

# Typy vazeb v PPI

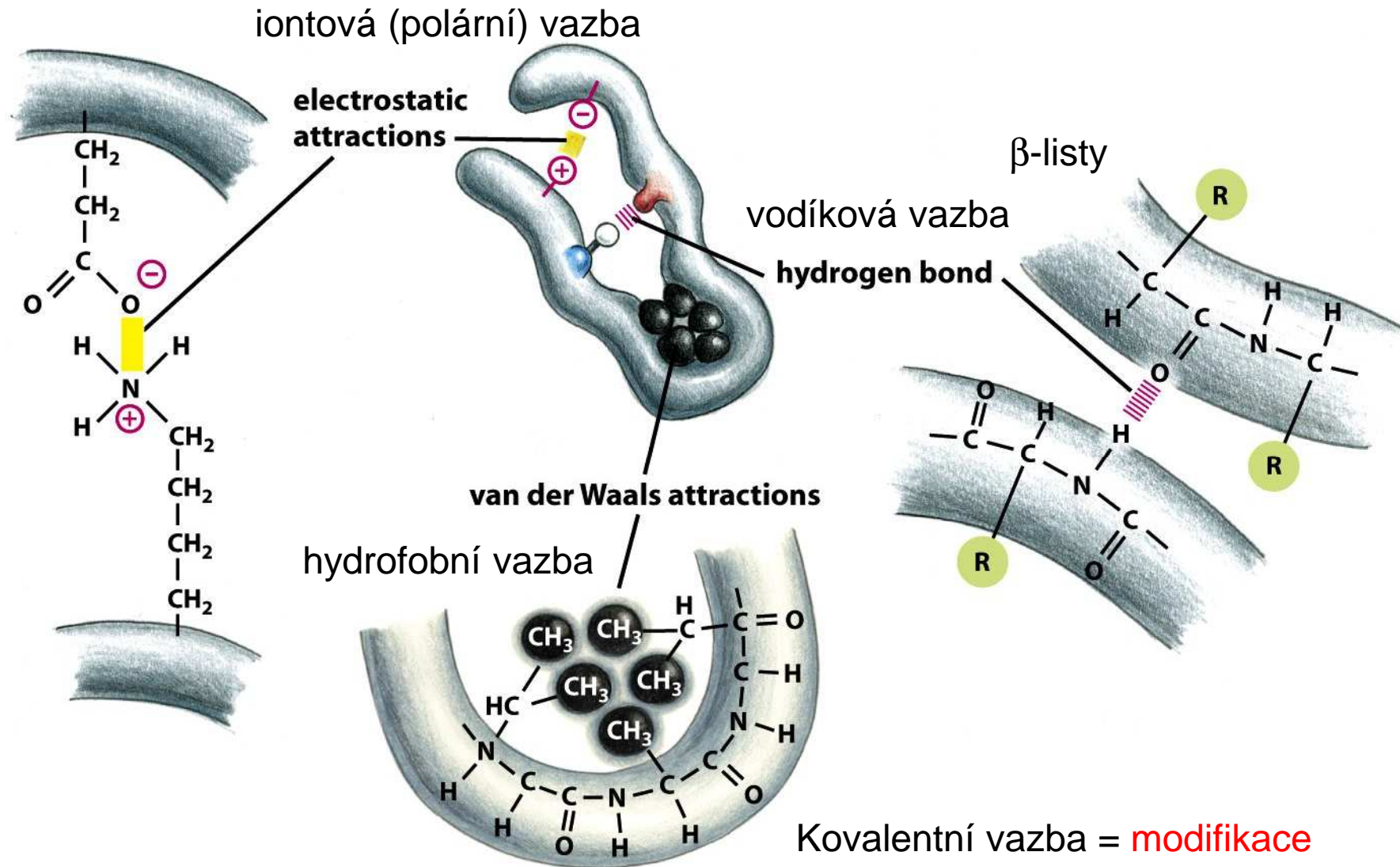


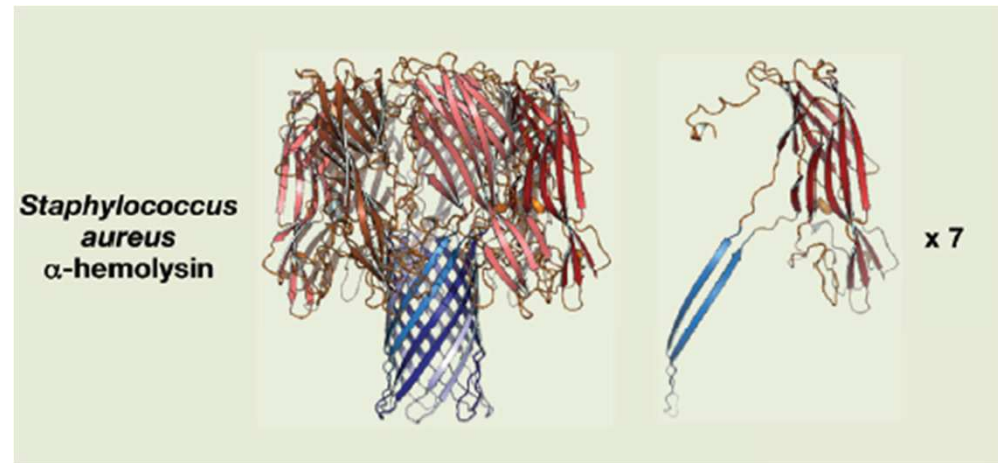
Figure 3-4 *Molecular Biology of the Cell* (© Garland Science 2008)



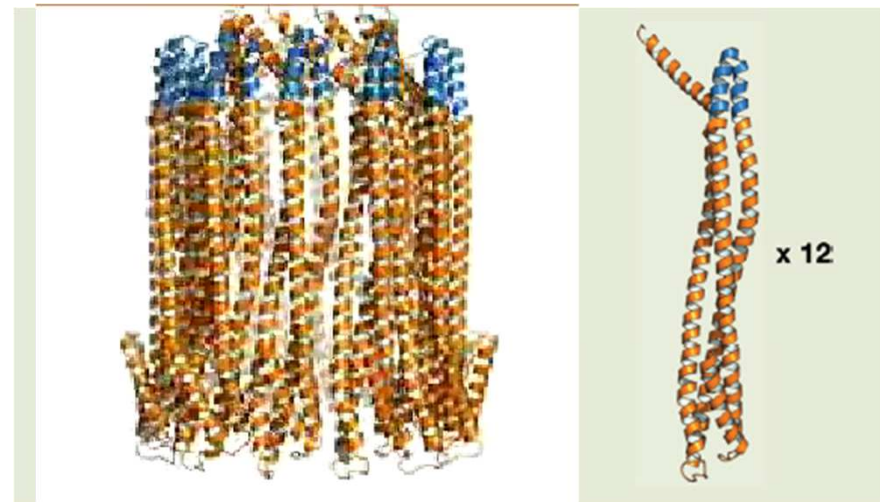
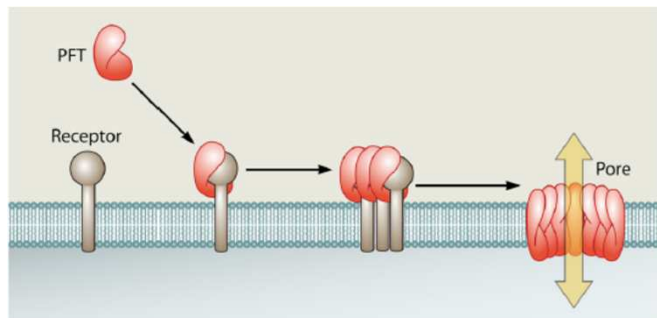


## ... sekundární struktury ...

- **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)



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



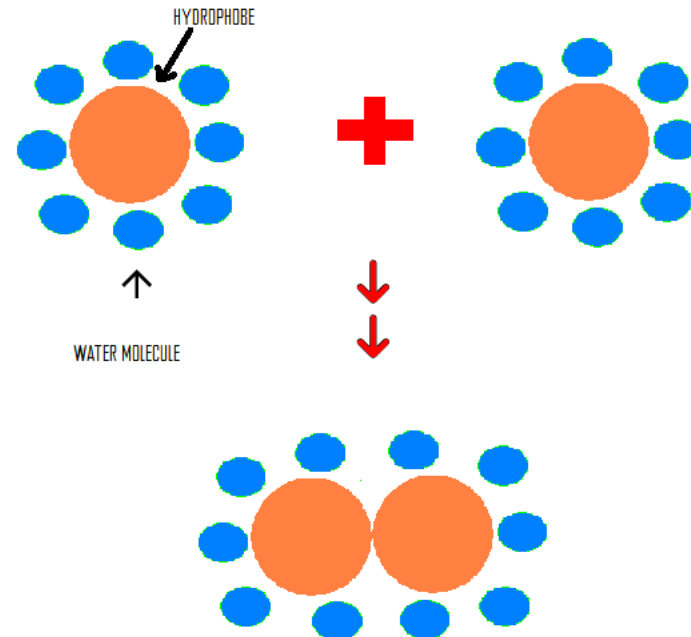
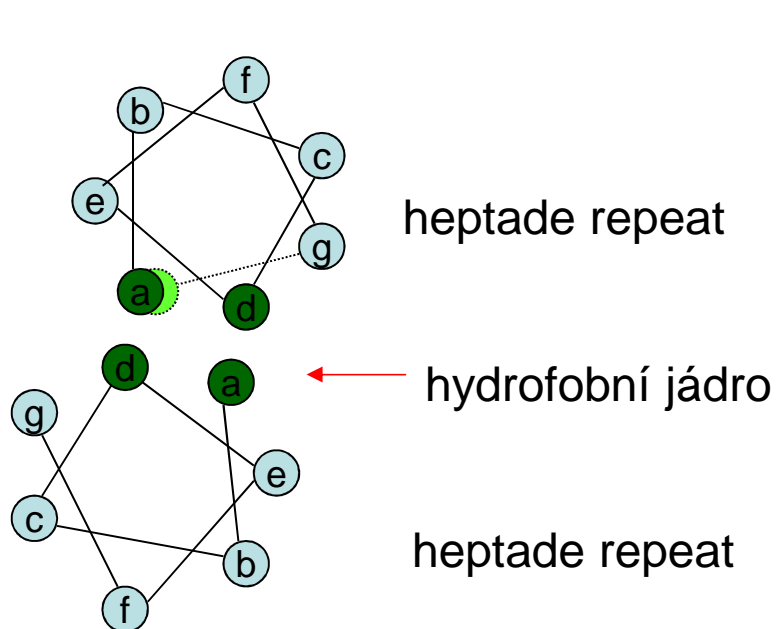
**Toxiny** – podjednotky se skládají tj. vytváří pór až v místě působení (neublíží původní buňce)

# ... sekundární struktury ...

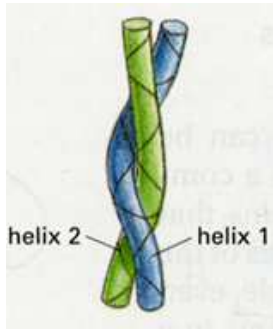
- šroubovice se vůči sobě orientují různým způsobem
- skládání slabých vazeb ovlivňuje sílu a specifitu celkové vazby

## coiled-coil struktura

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

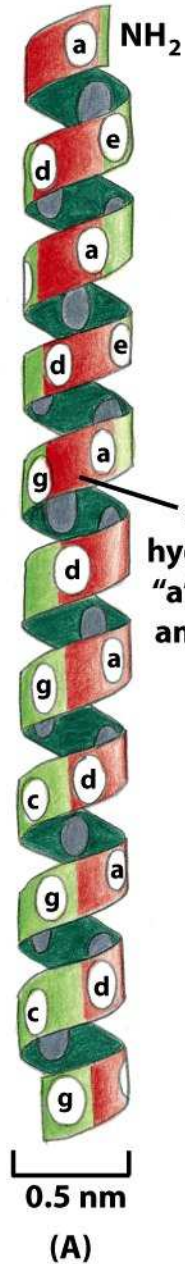


...LKSLHNQLRDLEESLTH...

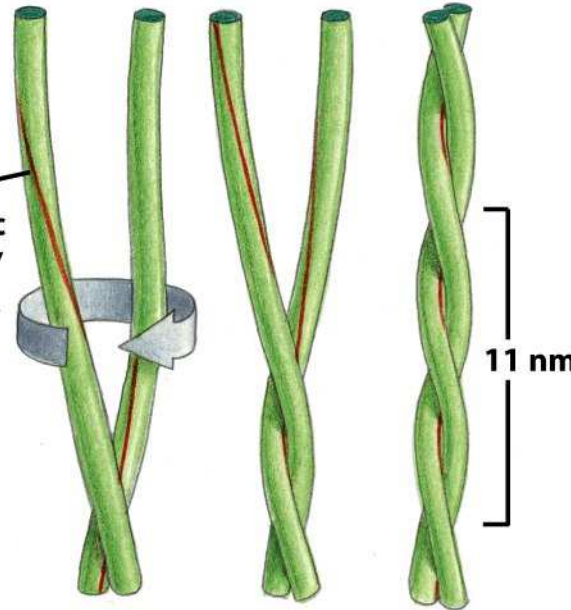


# coiled-coil struktura

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

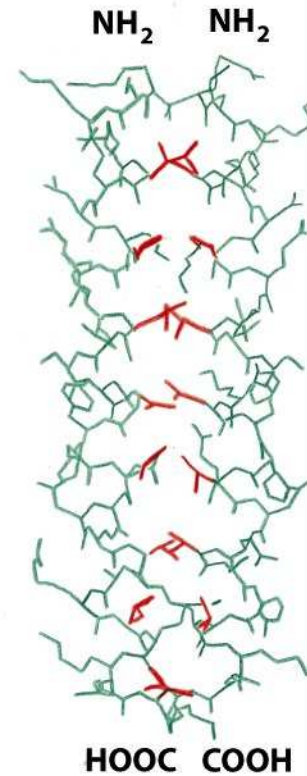


stripe of hydrophobic "a" and "d" amino acids



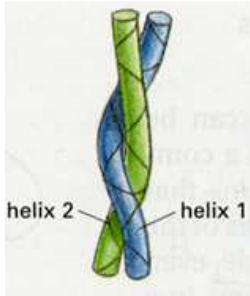
hydrofobní jádro

...LKSLHNQLRDL~~EES~~LTH...



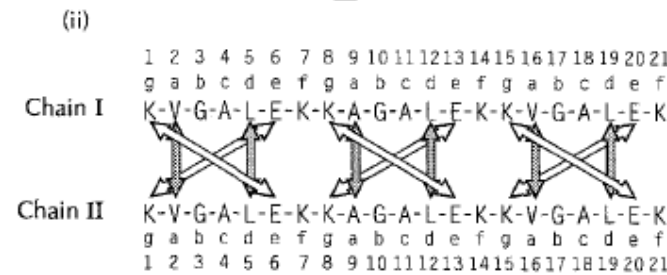
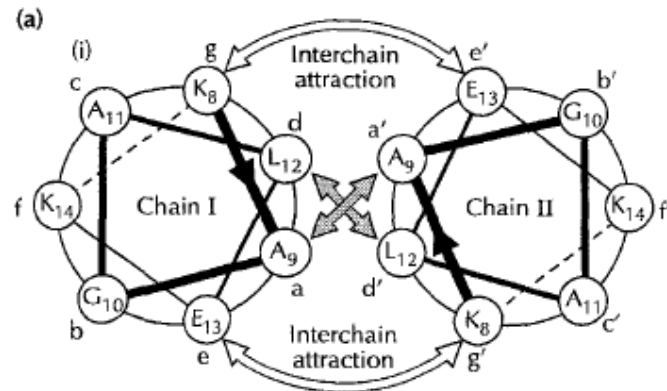
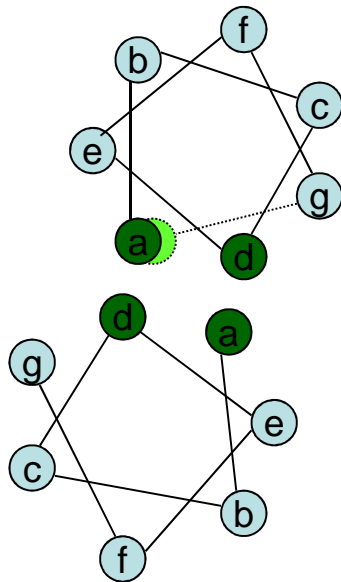
paralelní šroubovice

(C)

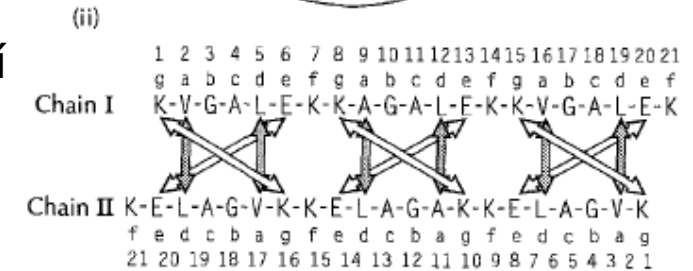
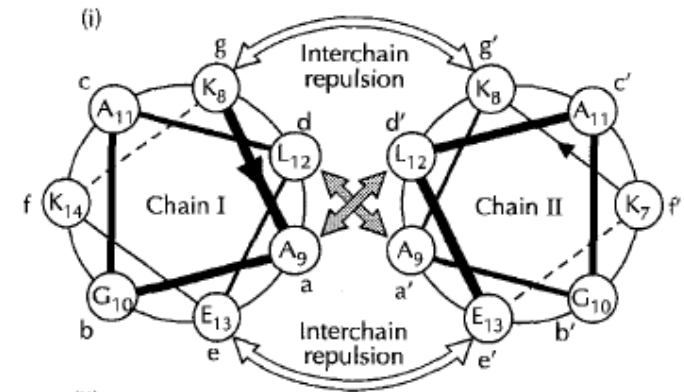


# coiled-coil struktura

Síla interakce může být ovlivněna sousedními AMK



Sousední AMK stabilizují interakce šroubovic



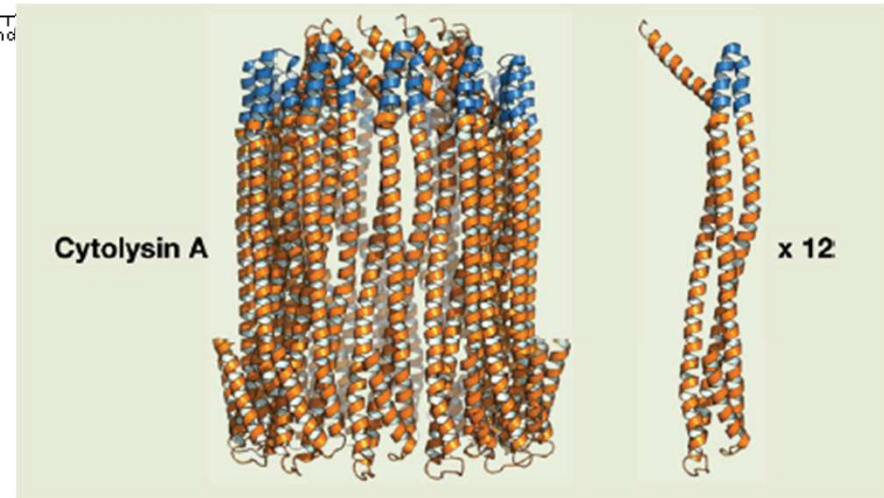
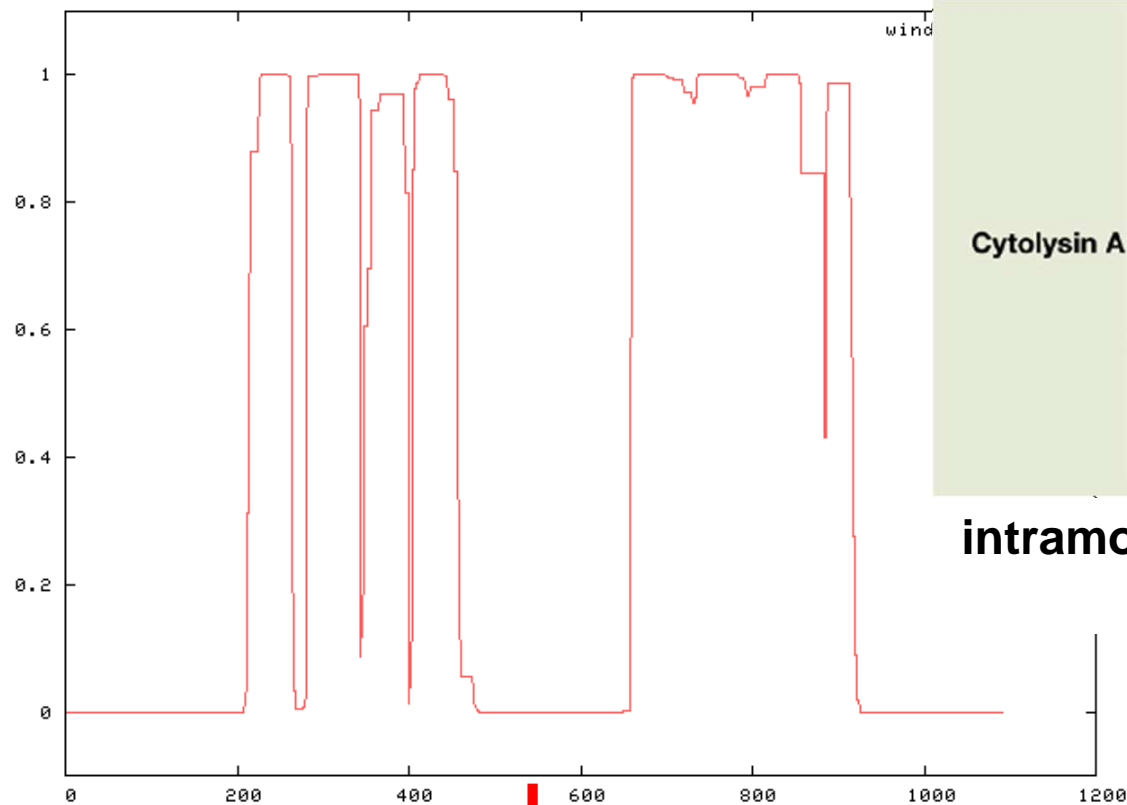
Sousední AMK destabilizují interakce šroubovic



# coiled-coil struktura

- program COIL: [http://www.ch.embnet.org/software/COILS\\_form.html](http://www.ch.embnet.org/software/COILS_form.html)

profil SMC6



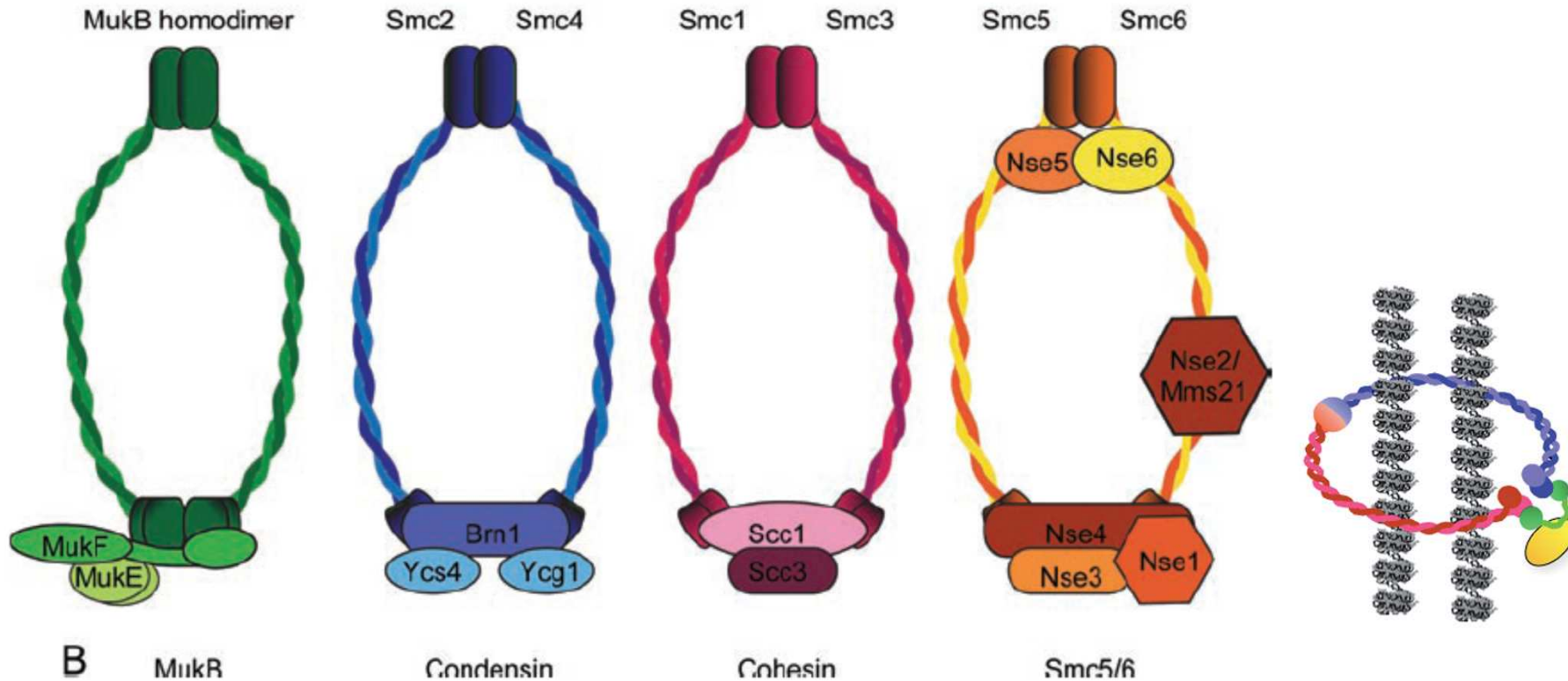
**intramolekulární vs intermolekulární  
folding vs PPI**



- CC v SMC proteinech jsou intramolekulární  
(antiparalelní)

# coiled-coil struktura

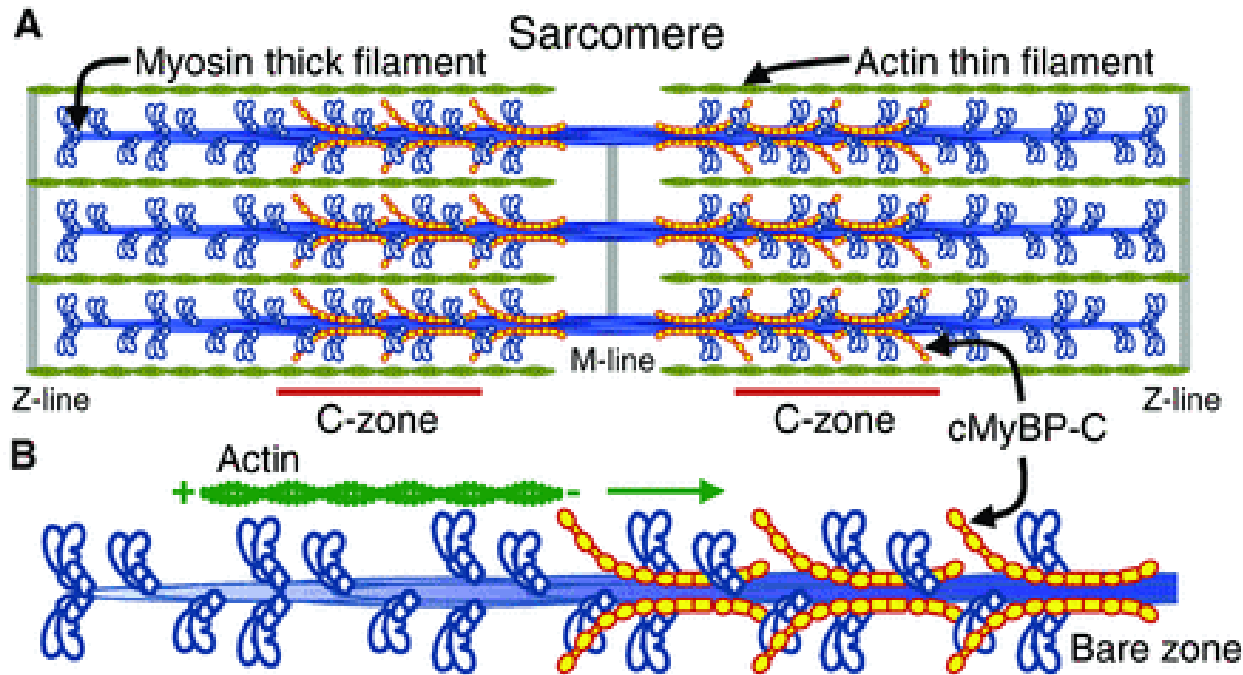
-dlouhé CC (>100AMK) vytváří vláknité struktury (SMC kruhy obejmou DNA)



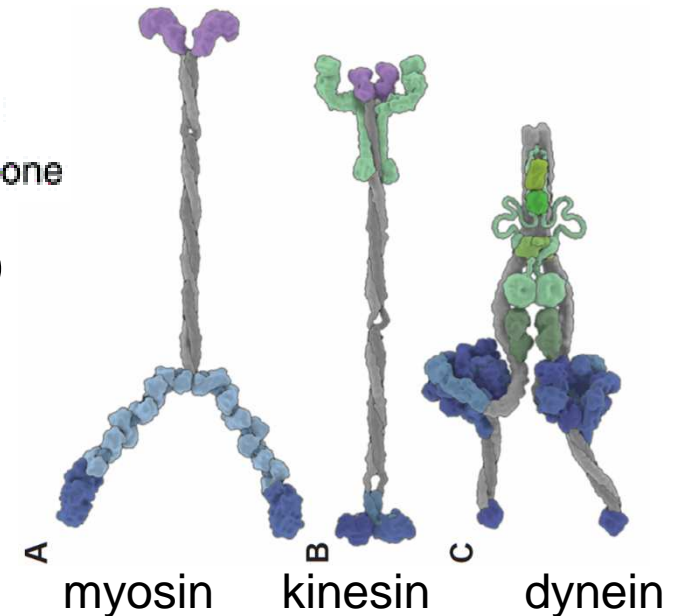
- CC v SMC proteinech jsou intramolekulární (antiparalelní)

# coiled-coil struktura

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



- CC v myosinu je intermolekulární (paralelní)

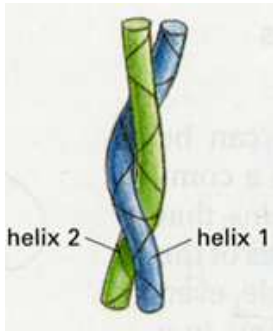


# Kinesiny

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

Cross, NRMCB, 2014





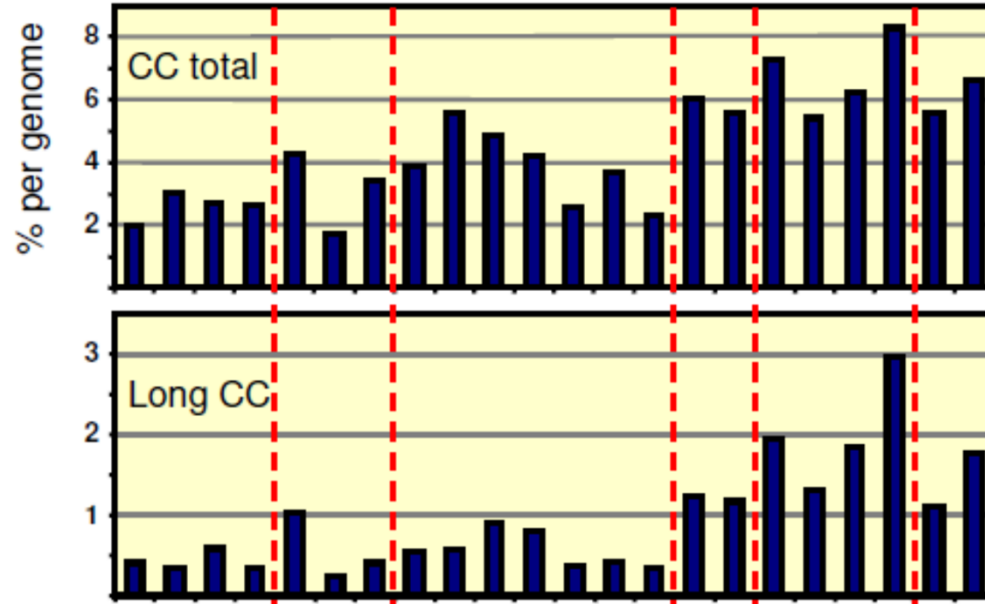
**Coiled-coil**  
doména je  
významným  
**dimerizačním**  
modulem u mnoha  
proteinů (GCN4,  
Max ...)

Intermolekulární -  
homo- či  
heterodimery  
(oligomery)

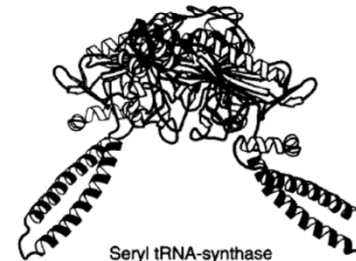


prokaryota

eukaryota



DNA polymerase I



Seryl tRNA-synthase



Transcript cleavage factor GreA



Influenza hemagglutinin HA2 (pH7)



Influenza hemagglutinin HA2 (pH4)



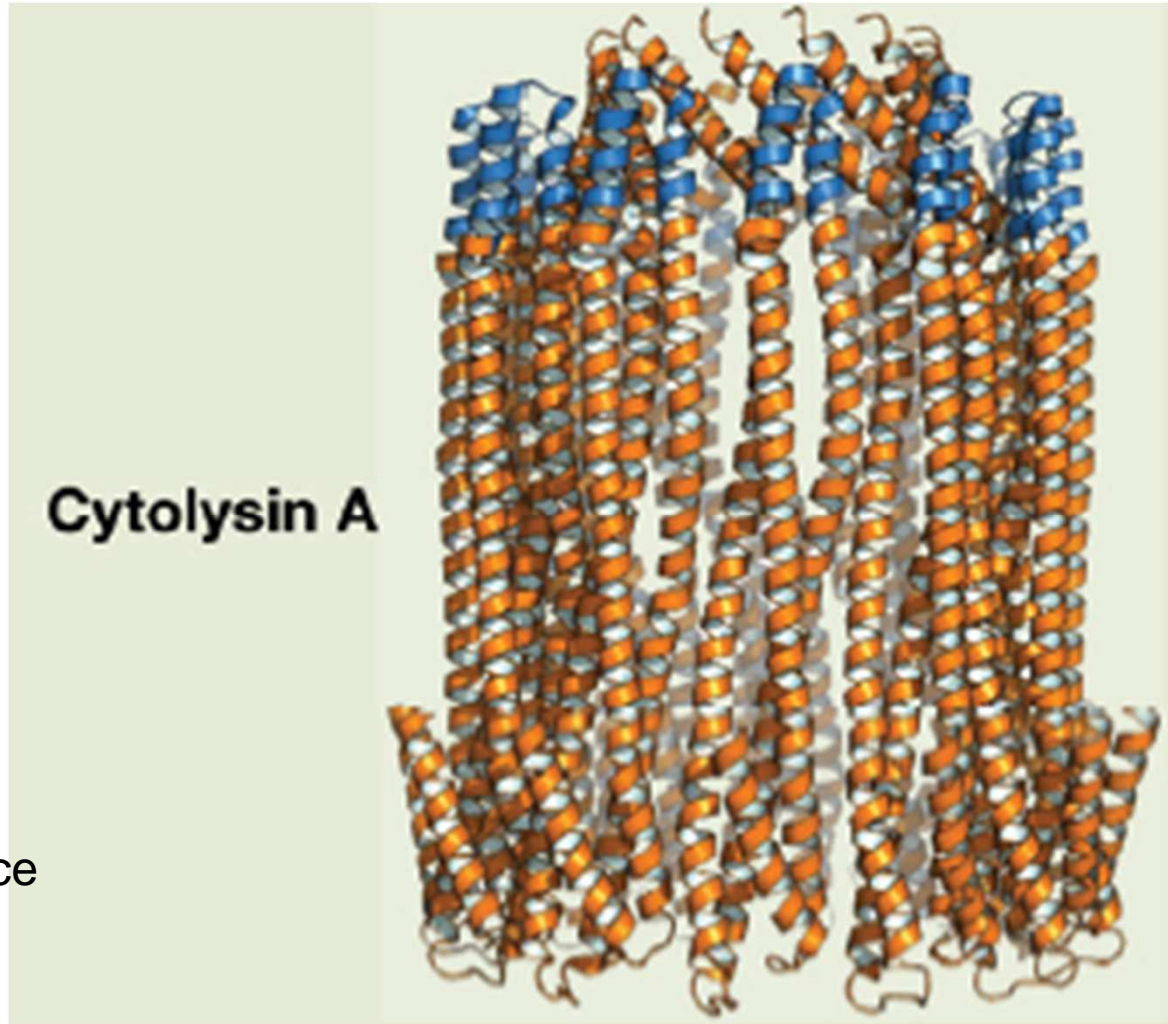
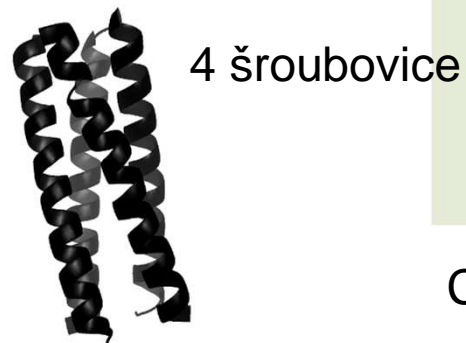
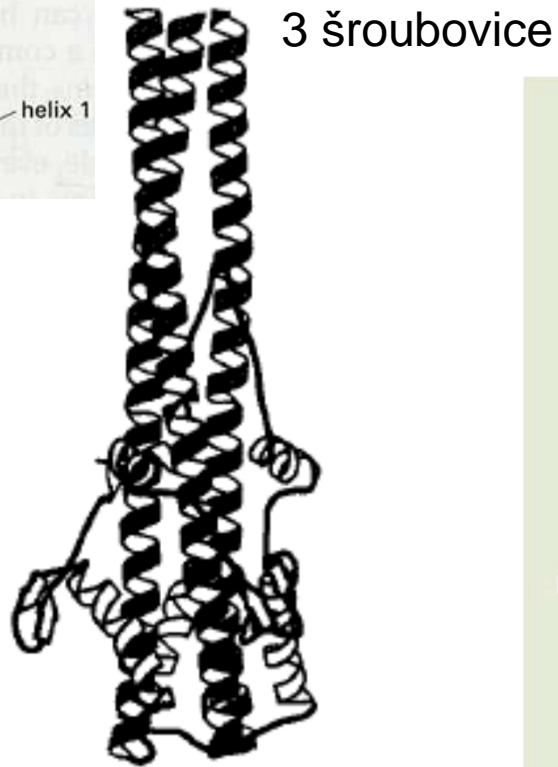
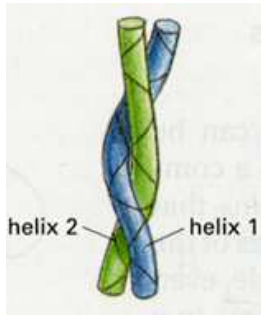
Mannose-binding protein



ROP

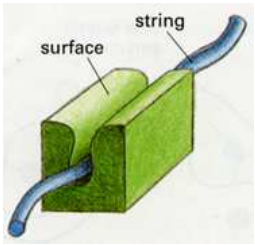
Tropomyosin

# Interakce šroubovic



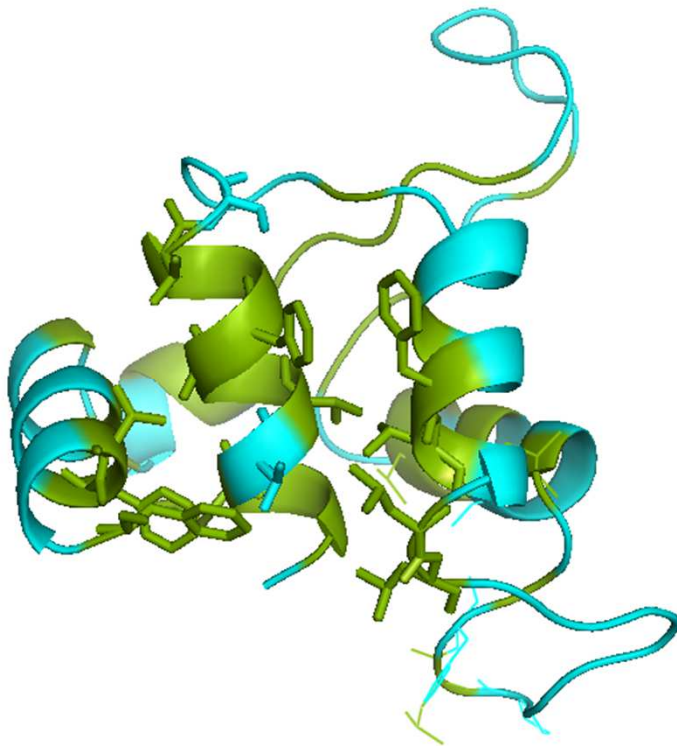
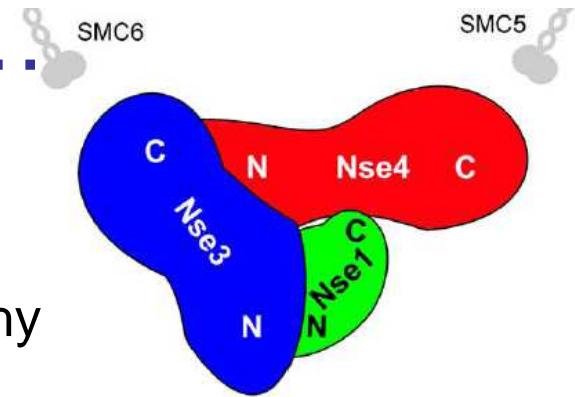
Cytolysin vytváří póry v membránách cizích buněk

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

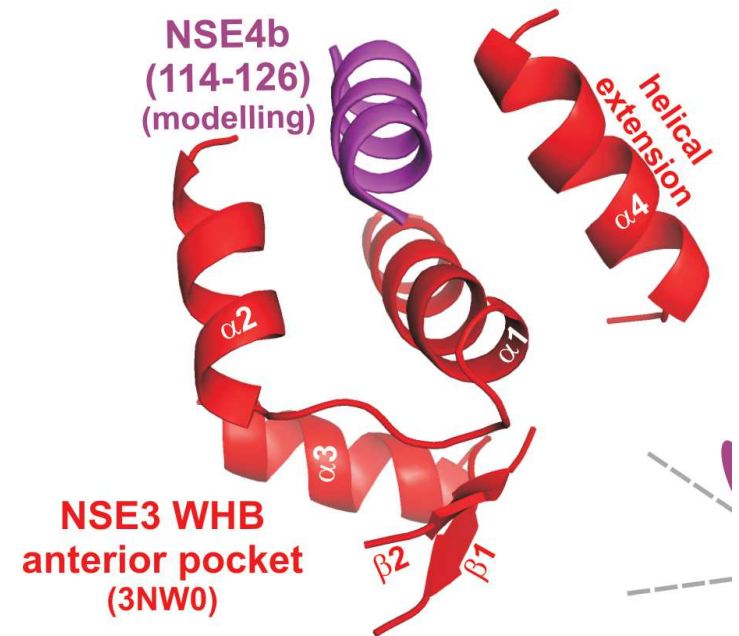


# ... terciární struktura ... kapsa-peptid

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



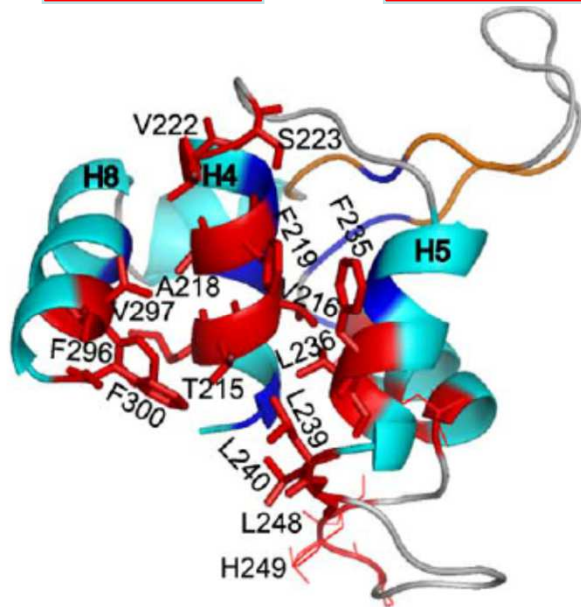
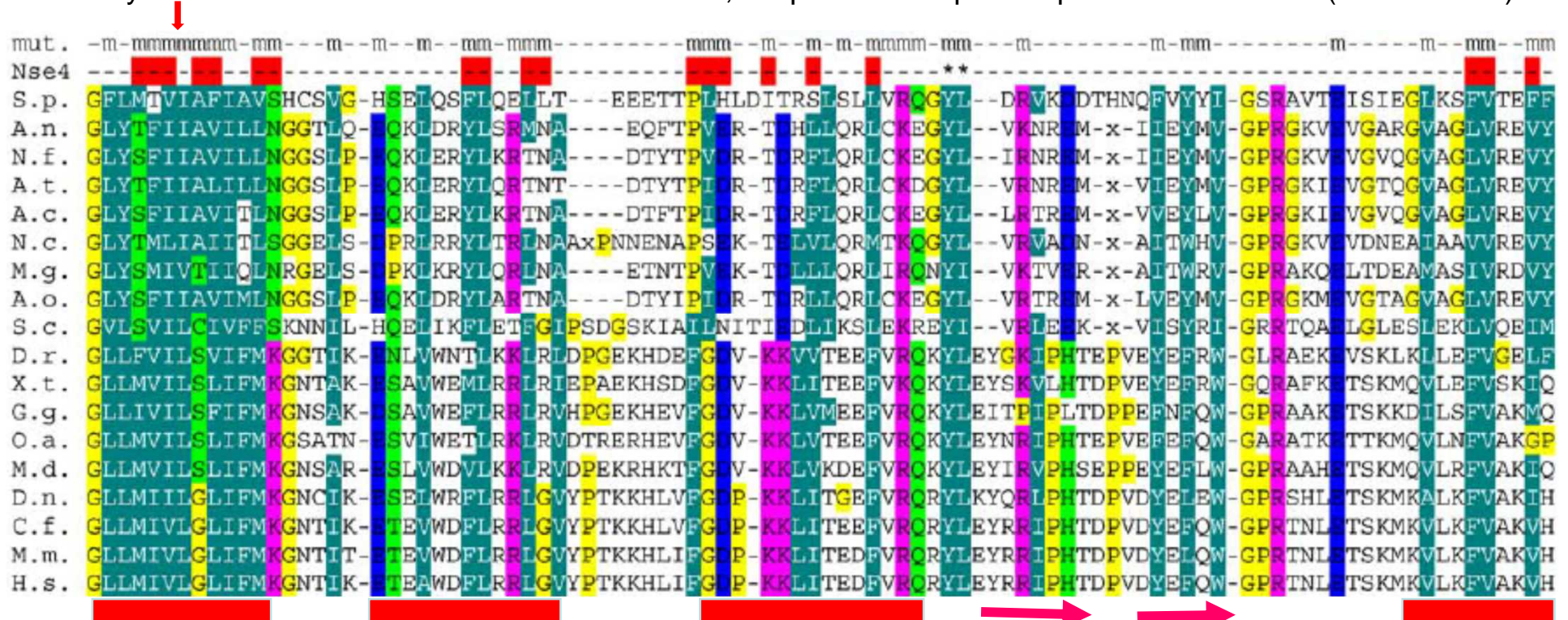
hlubší prohlubně na povrchu  
mohou tvořit kapsy pro vazbu  
partnera (šroubovice, peptid)



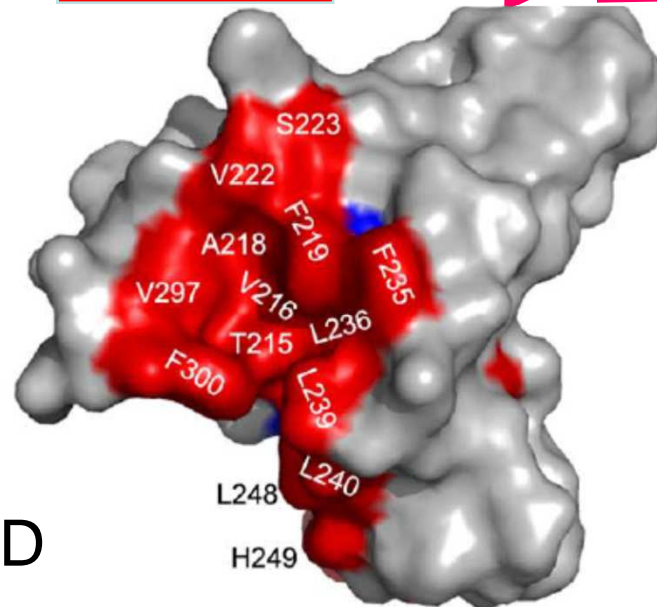
hydrofobní interakce mezi Nse3 a Nse4



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



WHD



Interakce  
mapována  
mutagenezí




**HADDOCK**  
Software web portal

Home HADDOCK Whispy CPDRT DNA SQUEEZE Publications HADDOCK Contact

WELCOME TO THE UTRECHT BIOMOLECULAR INTERACTION WEB PORTAL >>>


PROFILE >>>

**Universiteit Utrecht**

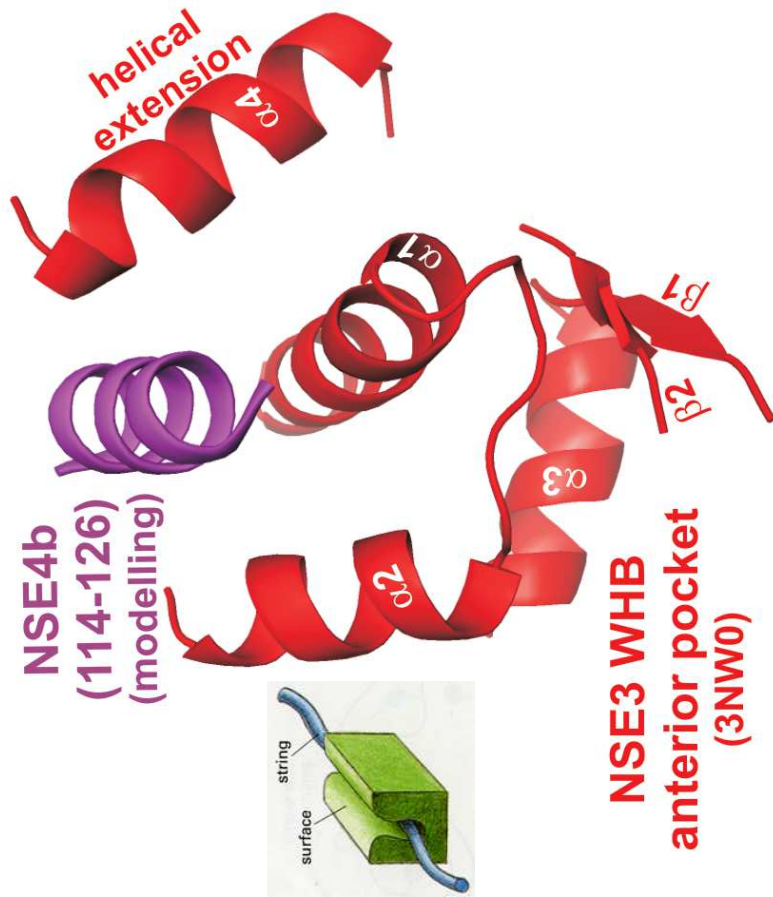


The Utrecht Biomolecular Interactions software portal provides access to software tools developed in the Computational Structural Biology group / NMR Research Group of Utrecht University with a main focus on the characterization of biomolecular interactions. Please note that this site is in active development.

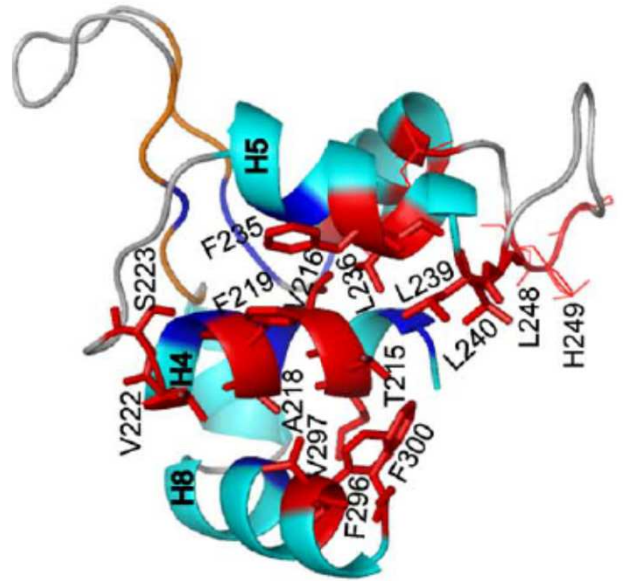
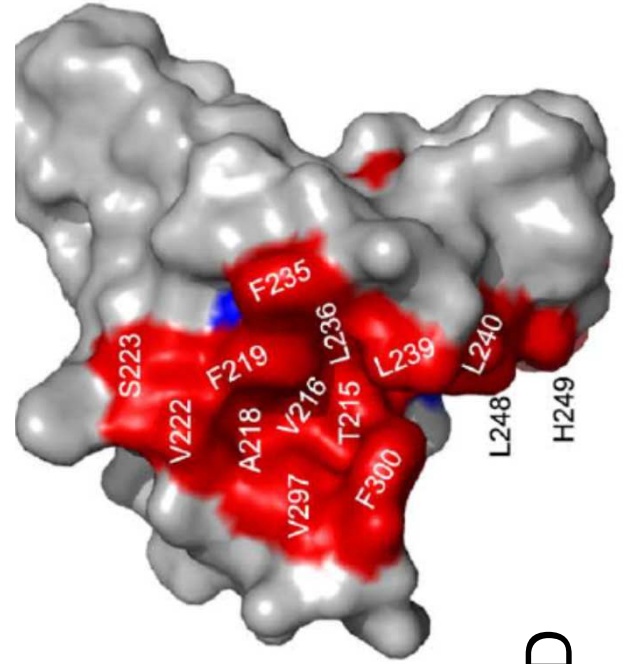
**HADDOCK WEB DOCKING**



HADDOCK (High Ambiguity Driven protein-protein DOCKing) is an information-driven flexible docking approach for the modeling of biomolecular complexes. HADDOCK distinguishes itself from ab-initio docking methods in the fact that it encodes



# Docking



# WHD

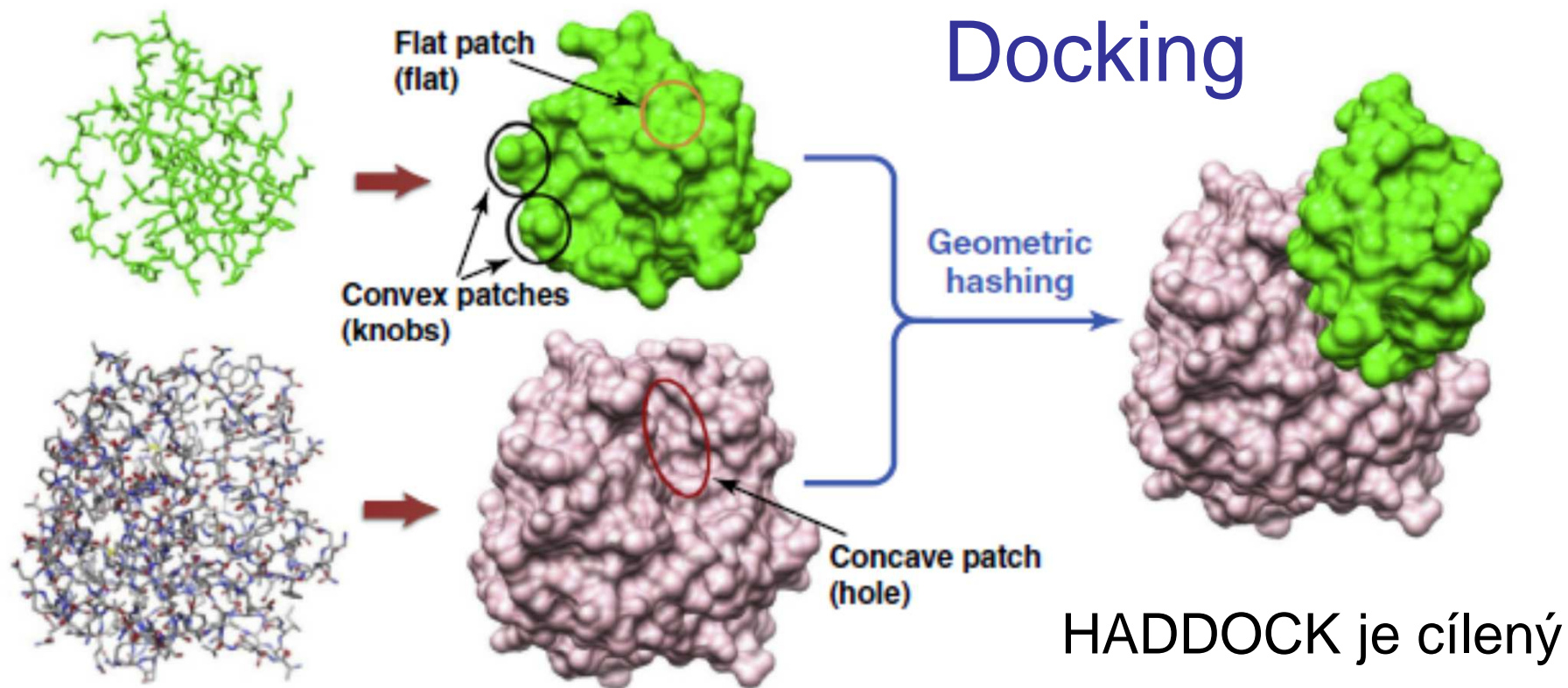
Interakce  
mapována  
mutagenézí

Hudson et al.: PLoS One, 2011  
Guerineau et al.: PLoS One, 2012

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: 2wa0 go

Top page Protein **Clefts** Tunnels Links

Cleft analysis for: 2wa0 PDB id 2wa0

**WHD**

**View options**

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

Coloured by

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

Jmol RasMol

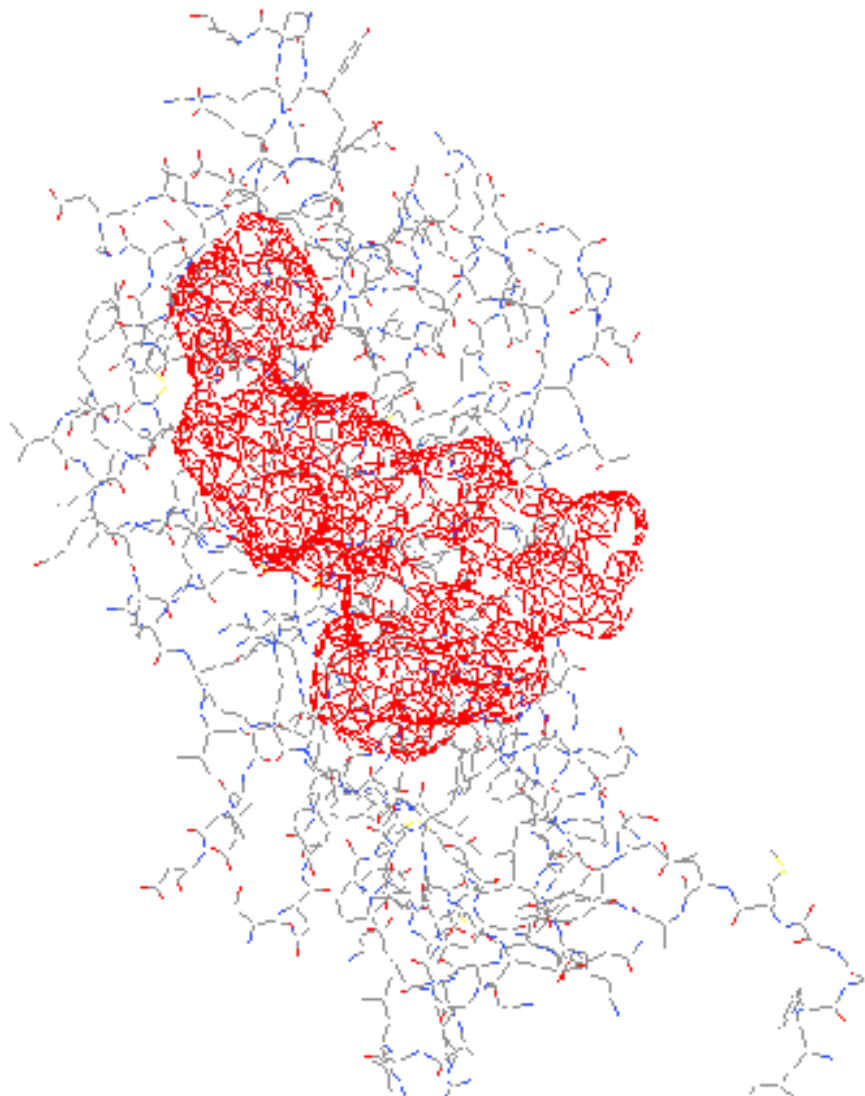
Clefts	Volume	R1 ratio	Accessible vertices	Buried vertices	Average depth	Residue type	Ligands			
1	2370.52	0.98	65.15	1	10.55	1	10.50	1	8 6 5 15 4 4 1	

největší kapsa

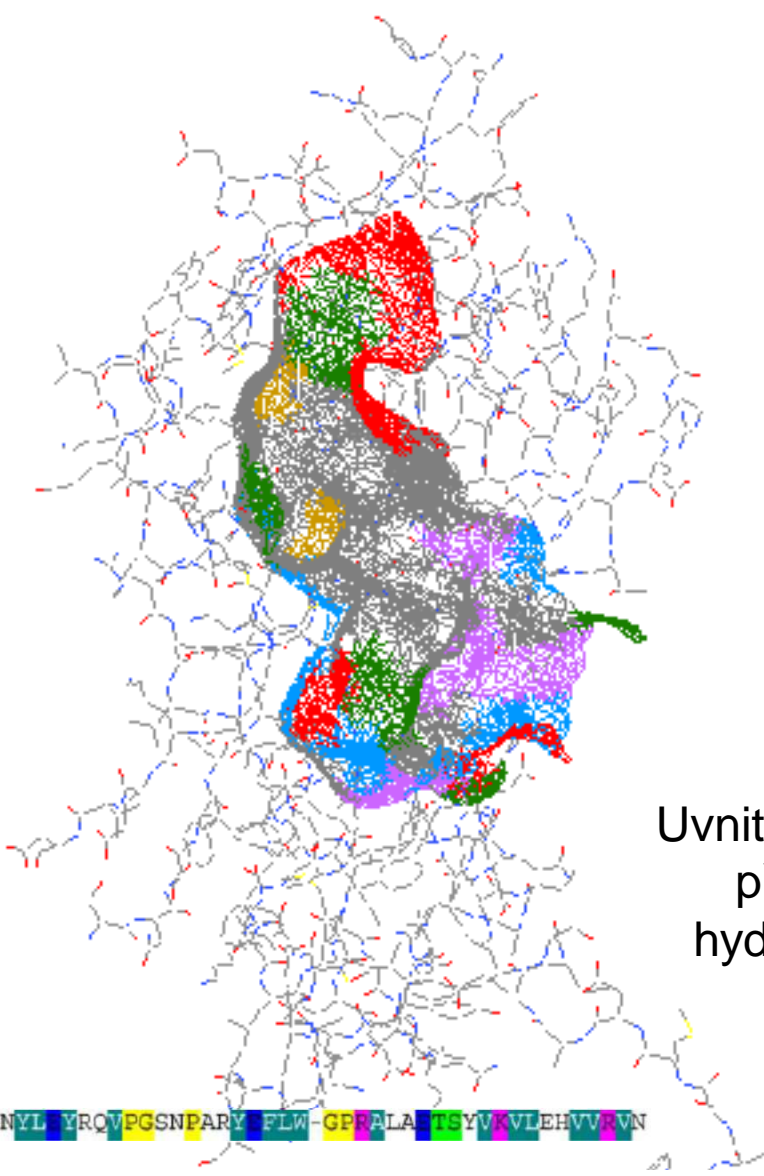
<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/>

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



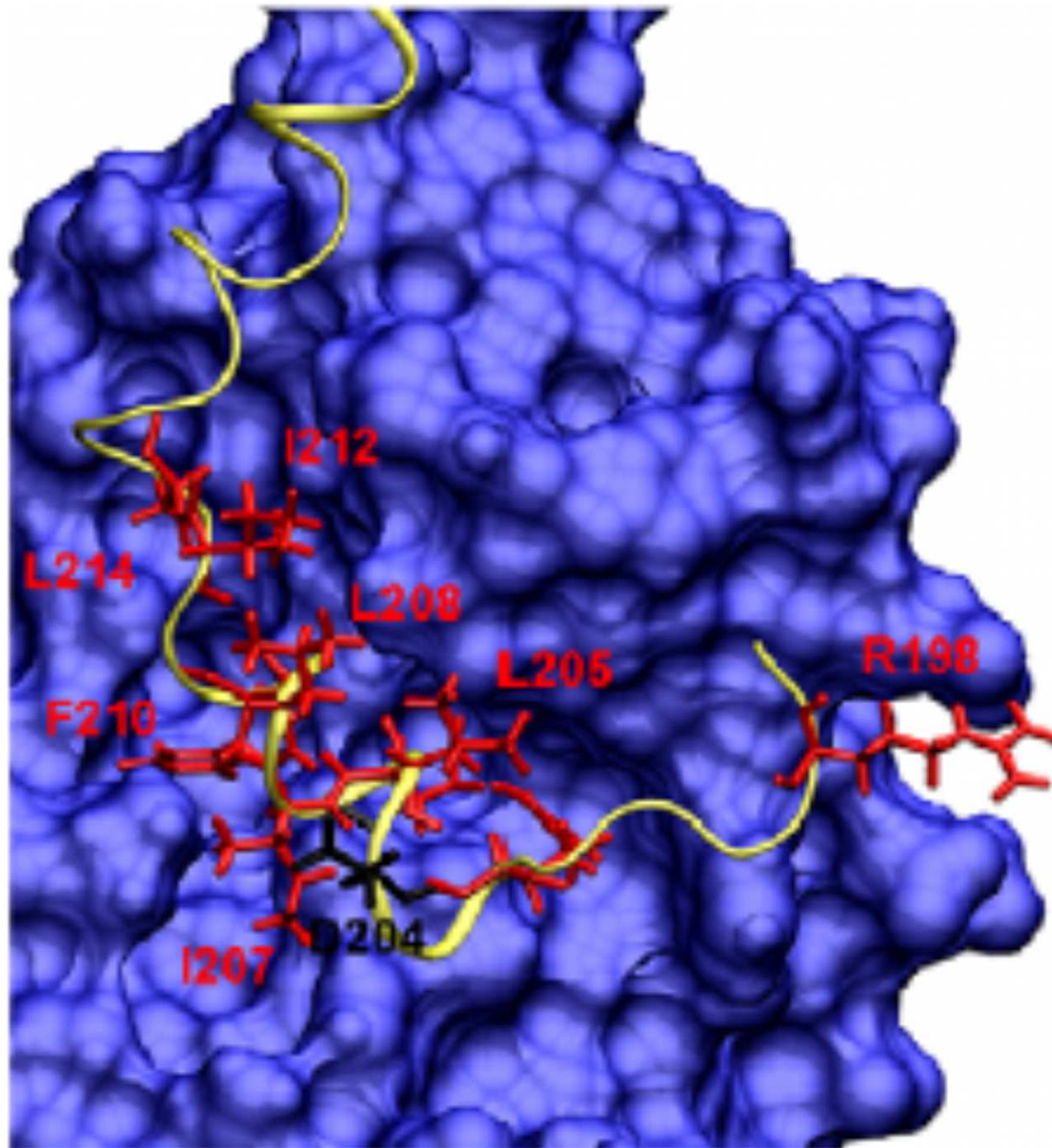
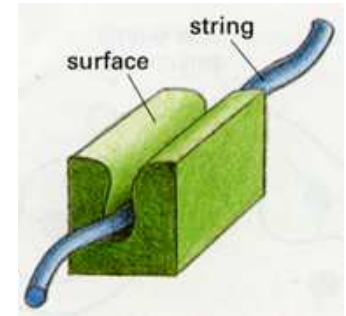
Uvnitř kapsy  
převládá  
hydrofobní  
povrch

A4

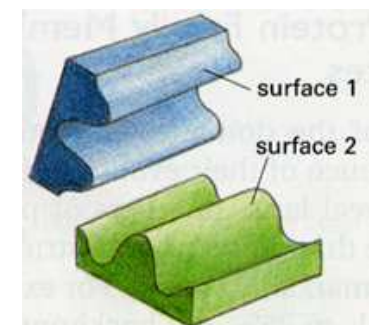
GLLIIVLGTIAMEGDSAS - EIWEEELGVVGVYDGREHTVYGF - RYLLTQDWVQENYL YRQVPGSNPARY FLW - GPRALANTSYVAVLEHVVRVN

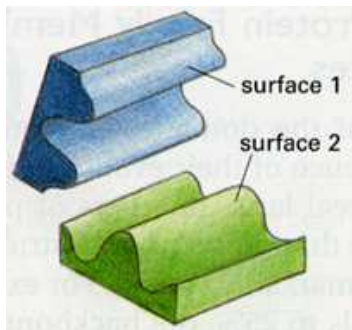
Residue-type colouring						
Positive	Negative	Neutral	Aliphatic	Aromatic	Pro & Gly	Cysteine
H,K,R	D,E	S,T,N,Q	A,V,L,I,M	F,Y,W	P,G	C





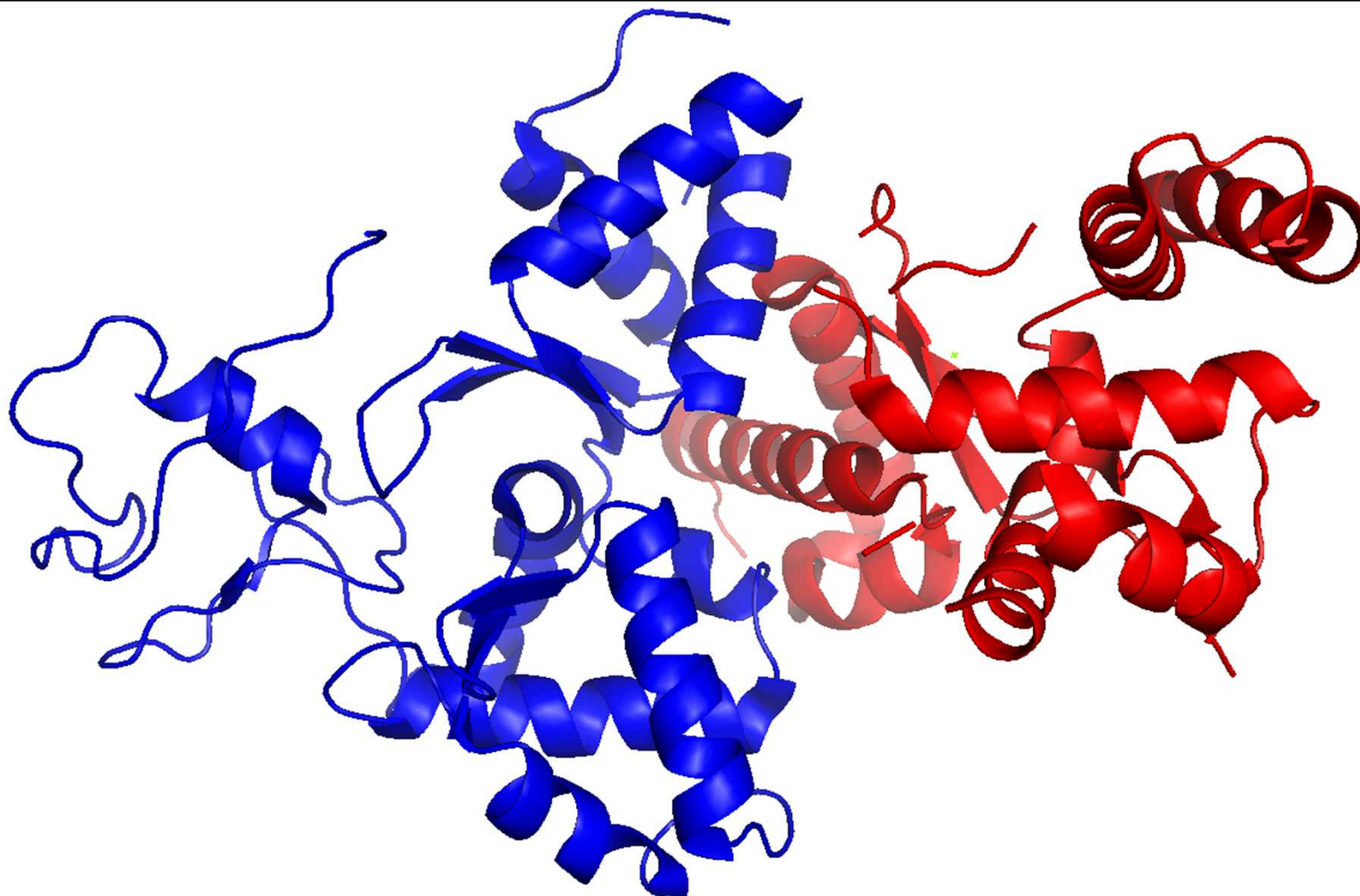
*de novo* docking partnera  
(HEX docking a  
molekulární dynamika):  
do hydrofobní kapsy  
proteinu byl nadockován  
„jednoduchý“ peptid (*de  
novo* docking větších  
povrchů je nespolehlivý)



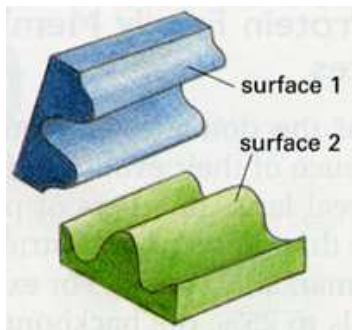


DOMÉNY - šroubovice,  $\beta$ -listy ... interagují pod různými úhly a vytváří různé vazebné motivy s rozsáhlými vazebnými povrchy

(kokrystal NSE1-NSE3 proteinů)

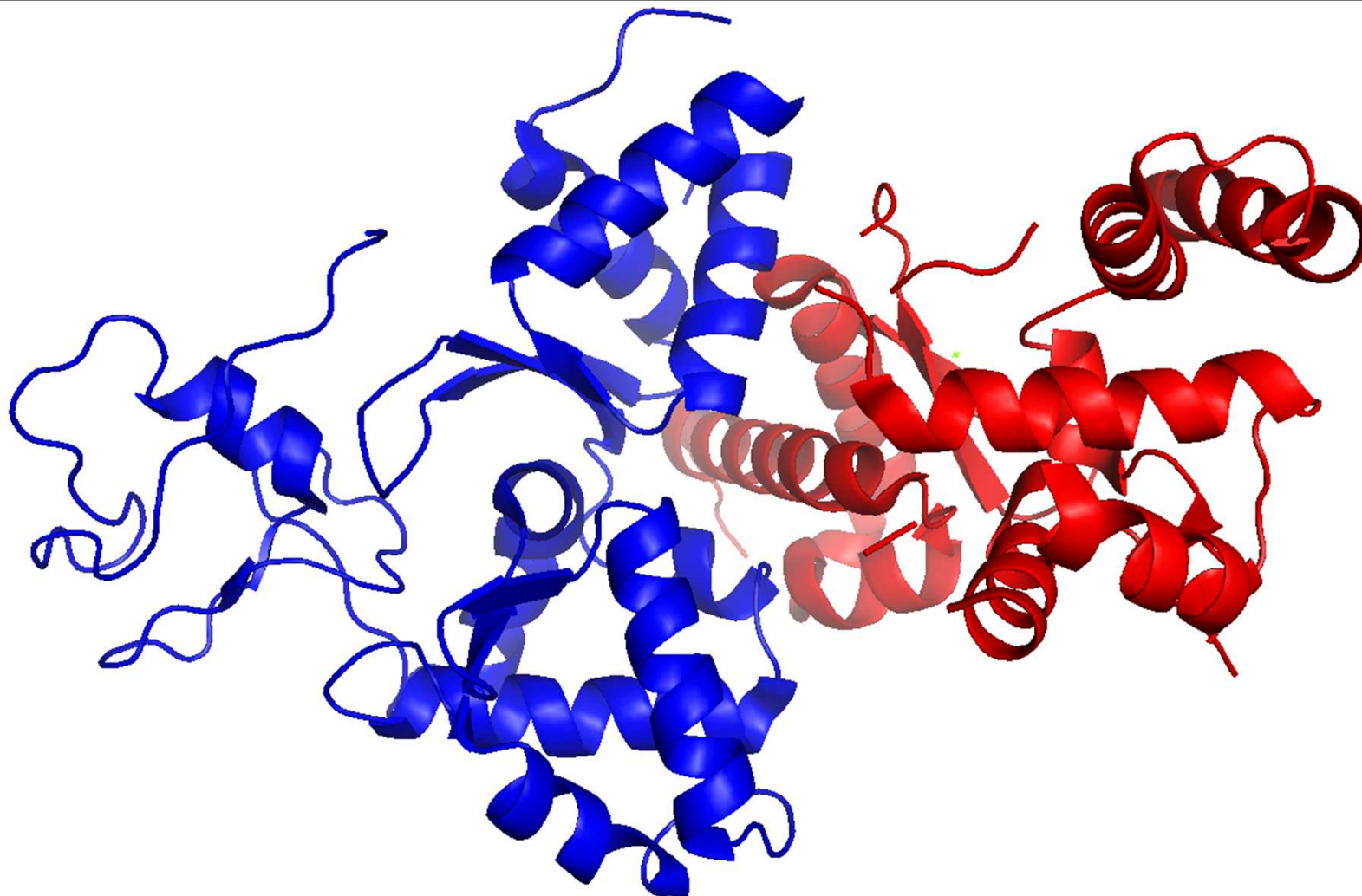






nejlépe lze získat info (vizuální, o typech vazby ...)  
z vyřešených struktur (PDBsum, 3DID - databáze)

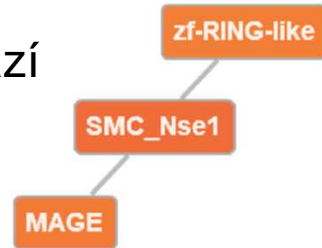
(kokrystal NSE1-NSE3 proteinů)



## 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)

HMM prof. interface res.

Binding partner(s)

MAGE

zf-RING-like

Search motif by name or keyword: ?

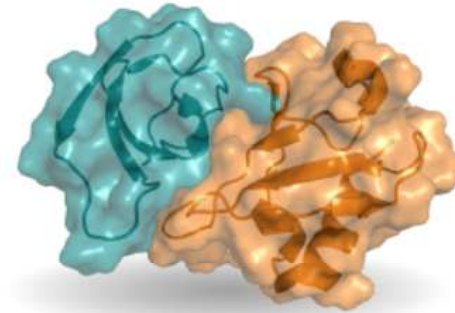
Motif name (e.g. SH2\_LIG\_0)

Search motif

# 3did

References Statistics Help About

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



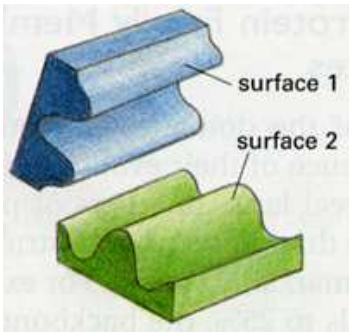
- > **Overview** General information on 3did
- > **Getting Started** Help for new users
- > **Technical Information** Linking to 3did
- > **Download** data files or MySQL tables

### 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





nejlépe lze získat info (vizuální, o typech vazby)  
z vyřešených struktur (**PDBsum – detailní info**)

(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

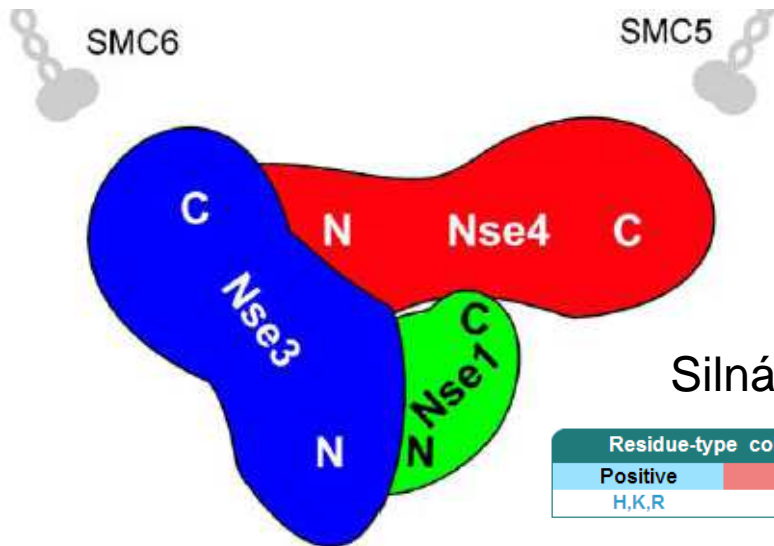
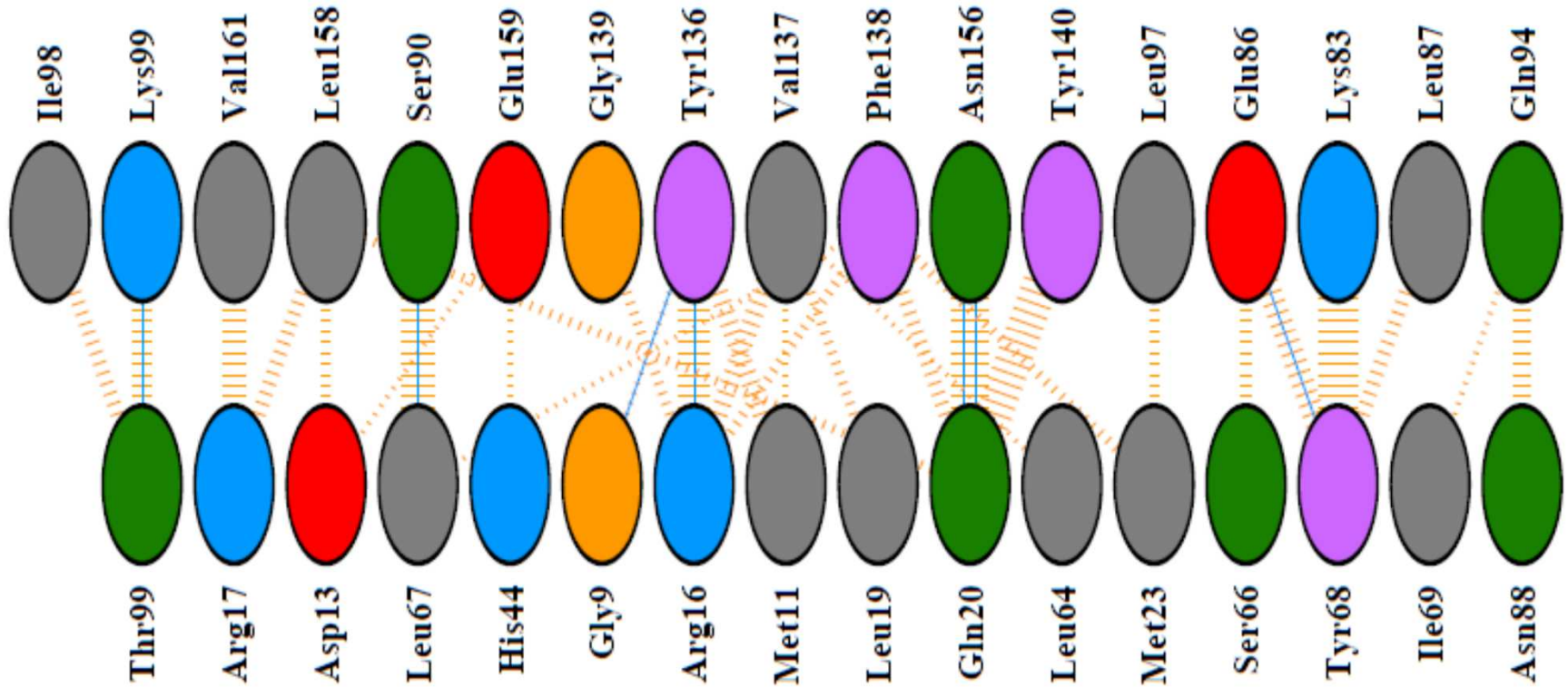
A}B (16:17 res)

### Interface statistics

Chain	No. of interface residues	Interface area (Å <sup>2</sup> )	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

# NSE1 - NSE3

Chain A Chain B



Disulphide bonds



Hydrogen bonds



Non-bonded contacts

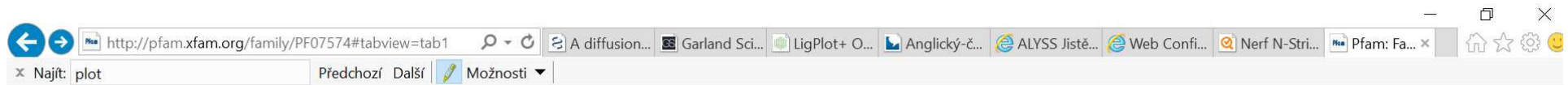
Silná interakce mezi NSE1 (chain A) a NSE3 (chain B)

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

```

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.SAHGAQSTIFFDSSSYVEKEIANLYLATER.....RPVLA...EDITQ
S7Z8E9_PENO1/8-209    HRAFLQAFM.ARSTMTFEDAQPVLAAII.SAH.....EG.....RTVDP...DEVTQ
B6H9Q9_PENRW/8-210    NRAFLQAFM.ARSCMTFEDAQPILAAIL.TVS.....EG.....RTVDP...DEVGE
H6C926_EXODN/14-207   NRAFLQAFM.ARSVLTLETAKPILAAIS.TFQ.....DG.....REVQP...QDMTV
U1GD89_ENDPU/11-202   NRAFVQAFM.ARGTLTYETSKPLLASIF.TVH.....EG.....REILP...NDITE
C5GY37_AJEDR/11-207   HRAFLQAFM.ARSTMTYEQAKPVLAALIF.SAR.....DH.....QDTLP...EDITQ
C6H5E2_AJECH/11-203   HRALLOAFM.ARSTMTYEOAKPVLAALIF.TAR.....DN.....OETLP...EDITO

```







## Family: *SMC\_Nse1* (PF07574)

- Summary
- Domain organisation**
- Clan
- Alignments
- HMM logo
- Trees
- Curation & model
- Species
- Interactions
- Structures
- Jump to...**
- 

### 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:EEF14330.1} (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.SAHGAQSTIFFDSSSYVEKEIANLYLATER.....RPVLA...EDITQ
S7Z8E9_PENO1/8-209	HRAFLQAFM	ARSTMTFEDAQPVLAAII.SAH.....EG.....RTVDP...DEVTE
B6H9Q9_PENRW/8-210	NRAFLQAFM	ARSCMTFEDAQPILAAIL.TVS.....EG.....RTVDP...DEVGE
H6C926_EXODN/14-207	NRAFLQAFM	ARSVLTLETAKPILAALIS.TFQ.....DG.....REVQP...QDMTV
U1GD89_ENDPU/11-202	NRAFLQAFM	ARSTMTYEQAQPVLAALIF.SAR.....DH.....QDTLP...EDITQ
C5GY37_AJEDR/11-207	HRAFLQAFM	ARSTMTYEQAQPVLAALIF.SAR.....DH.....QDTLP...EDITQ
C6H5E2_AJECH/11-203	HRALLQAFM	ARSTMTYEQAQPVLAALIF.TAR.....DN.....QETLP...EDITQ
F2PT91_TRIEC/10-199	HRAFLQAFM	SRSTMTLEEAKPVLAALIF.TVS.....EG.....REILP...GDITQ
E9DEJ9_COCPS/13-202	HRAFLQAFM	ARSTMTLNEAKPILAAIL.SVK.....DG.....REVLP...EDVTQ
R7Z157_CONA1/13-202	HRAFLQAFM	ARSVLTFEEAQPILAAIL.TAH.....EG.....RPTLP...ADITT
U4LU38_PYROM/8-205	HRAFLQAFM	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	SKFLLQYVL	RRGVCSEKALAKALKTL.....ERDG.....EQLEDSETE
C5DQF6_ZYGR/37-240	ARYLLQYLL	CRGICHENMLLVLDKL.....QK.....YTQDPTSQVCS.T
I2H1A9_TETBL/24-252	RHLLRYIM	SEGICHENMLLLALYAL.....NLDYS...DCQEVLA.....
G8C139_TETPH/22-276	RYLLQYLL	CGICNESFMLLALMRLH.....LTDGSSSDDFYRLNMVELN
G8ZVJ6_TORDC/16-236	RIVLQYIL	SRGIVGENILLALMRL.....QGDT...ETND...VQELY
H2ATC7_KAZAF/7-237	RYLLQYIL	ARGICHENALVVALMRL.....YMDL...GCFDDAWQIDQWL
J5S7B7_SACK1/17-253	RYLLQYIL	ARGICHENNLILALMKL.....ETDA.....SKWSTEQWT
E7NKI9_YEASO/18-258	RYLLQYIL	ARGICHENALILALMRL.....ETDA...STLNTXSIQQWV
J8PZG2_SACAR/18-258	RYLLQYIL	ARGICHENALILALMRL.....ETDA...LRFDAERSMQQWI
G0W7Y8_NAUDC/22-268	RYLLQYIL	ARGVCHENALVVALMRL.....KVDS...NTIDPQWTISDWL
G0V5G7_NAUCC/25-272	RYLLQYIL	ARGVCHENALMLALMRL.....KVDS...HDVNAHWTISDWL
J7S9R2_KAZNA/85-303	RYLLQYIL	ARGVCHENSLLAALMHL.....SLDLD...PHSVSTR.SLDEWS
A7TJ64_VANPO/8-260	RYLLQYLL	CRGICHENALLLVLIKL.....DKDSVDEEGSPNRTFEDYL

## Konzervované AMK svědčí o důležitosti jejich funkce:

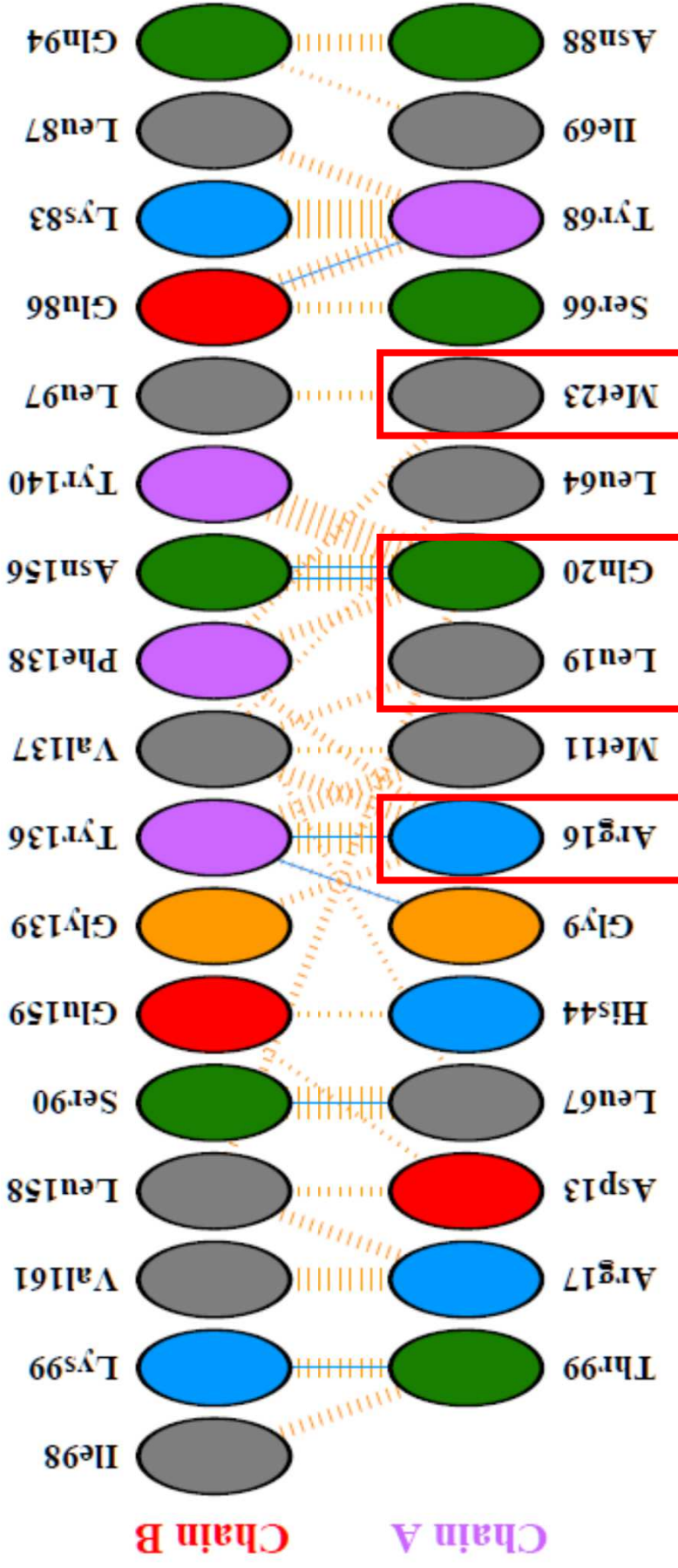
- důležité pro proteinovou strukturu
- důležité pro funkci proteinu:
  - enzymy – aktivní centra
  - **komplexy – PPI**
  - regulační funkce – AMK posttranslačně modifikovaná



A1CCH2\_ASPCL/14-216  
 A2Q7K6\_ASPNC/15-218  
 B8NLA5\_ASPFN/14-216  
 B6QTR9\_TALMQ/14-217  
 V5FED6\_BYSSN/14-236  
 S7Z8E9\_PENO1/8-209  
 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

NRAFLOAFM ARSTMTFAEAKPVLAALF.SAH.  
 NRAFLOAFM ARSTMTFTQAKPVLAALF.SIR.  
 NRAFLOAFM ARSTMTFAEARPVLAALF.SVH.  
 NRAFLOAFM ARSTMTFDEAKPVLAALF.SAQ.  
 NRAFLOAFM ARSTMTFEEAKPVLAALF.SAHGAQSTIFFDSSYVEKEIANLYLATER.  
 HRAFLOAFM ARSTMTFEDAQPVLAAII.SAH.  
 NRAFLOAFM ARSCMTFEDAQPILAAIIL.TVS.  
 NRAFLOAFM ARSVLTLETAKPILAAIS.TFQ.  
 NRAFVQAFM ARGTLTYETSKPLLASIF.TVH.  
 HRAFLOAFM ARSTMTYEQAKPVLAALF.SAR.  
 HRALLOAFM ARSTMTYEQAKPVLAALF.TAR.  
 HRAFLOAFM SRSTMTLEEAKPVLAALF.TVS.  
 HRAFLOAFM ARSTMTLNEAKPILAAIIL.SVK.  
 HRAFLOAFM ARSVLTFEEAQPILAAIIL.TAH.  
 HRALVQAFM ARSSMTGEEILGVTVAIH.GVE.  
 HRTFLQALL IRPFIDIEEGQELLAAIA.SAE.  
 HRAFLOALL IRPFIDVQEGRELLAAIK.SAE.  
 EKFLLOYVLSRRGVCSEKALAKAKITL

EG...QPVSA...DDVTE  
 DD...EQVSP...EDITE  
 EG...EPVSA...EDVTE  
 EN...REVLAA...EDITQ  
 RPVLA...EDITQ  
 RTVDF...DEVTDQ  
 RTVDF...DEVGE  
 REVQP...QDMTV  
 REILP...NDITE  
 QDTLP...EDITQ  
 QETLP...EDITQ  
 REILP...GEDITQ  
 REVLP...EDVTQ  
 RPTLP...ADITT  
 NPEEP...TETTL  
 TDVPA...NSITV  
 SDVSI...ESVPP  
 ERDG...EQLEDESETE



# Pfam consensus sequence alignment for SMC\_Nse1 (PF07574) with binding site information

## KBDock

## Propojení PDB a PFAM

- Note (i) Top sequence is the Pfam family consensus sequence  
 (ii) Binding site residue color-coding scheme: "center" , "core" , "rim"  
 (iii) Position cursor on a residue to display the PDB residue label  
 (iv) [Table of Pfam Consensus Amino Acid Classes](#)

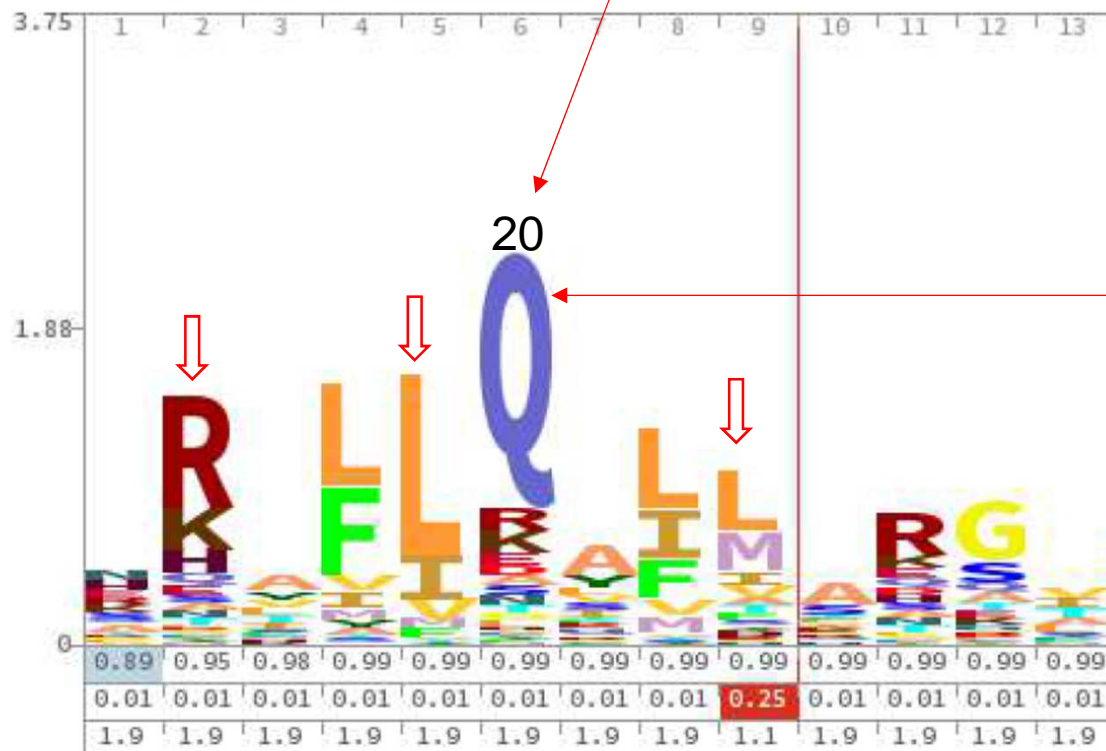
```

Consensus_PF07574          DVHRRFLQLLMTHGVLEEDVKRLQTHCYKVHNATVDKLEDFINNINSVL
PF07574_3nw0_A_13_176    DVHRRFLQLLMTHGVLEEDVKRLQTHCYKVHNATVDKLEDFINNINSVL

Consensus_PF07574          ESLYIEIKRGVTEDDGRPIYALVNLATTSISKMATDFAENELDLFRKALE
PF07574_3nw0_A_13_176    ESLYIEIKRGVTEDDGRPIYALVNLATTSISKMATDFAENELDLFRKALE

Consensus_PF07574          LIIDSETGFASSTNILNLVDQLKGGKMRKKEAEQVLQKFVQNKWLIEKEG
PF07574_3nw0_A_13_176    LIIDSETGFASSTNILNLVDQLKGGKMRKKEAEQVLQKFVQNKWLIEKEG

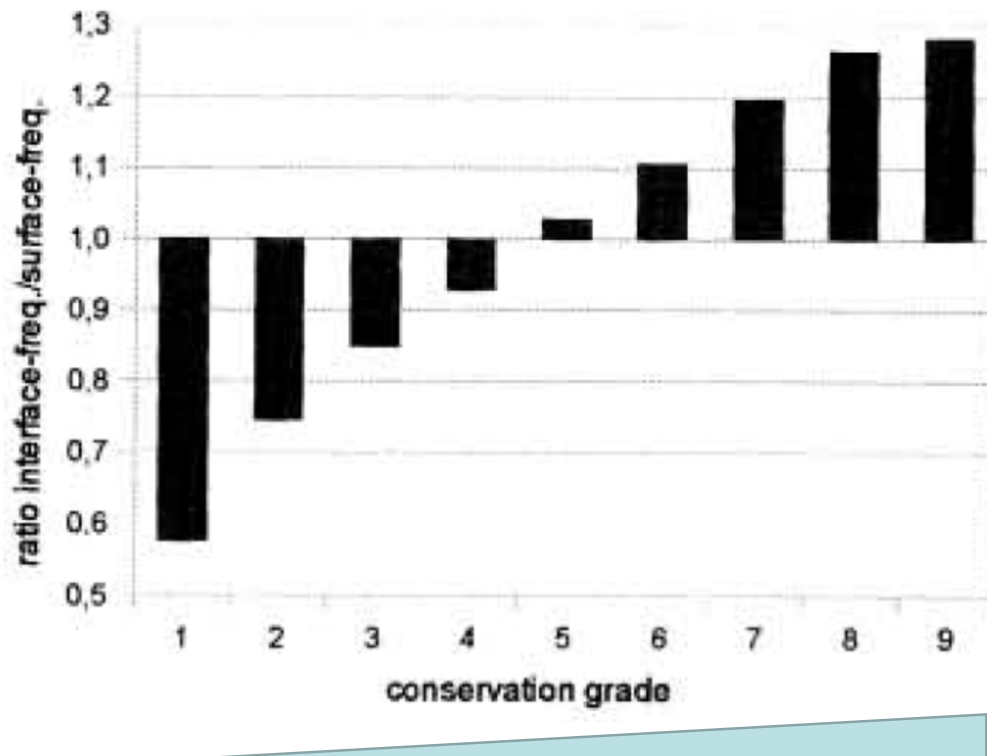
Consensus_PF07574          EFTLHGRAILEM
PF07574_3nw0_A_13_176    EFTLHGRAILEM
    
```



>PF07574_1	3nw0	PF07574	A	13	176	PF01454	B	92	263
13	D	rim	158	L	core				
13	D	rim	159	E	rim				
16	R	rim	136	Y	rim				
16	R	rim	137	V	core				
16	R	rim	138	F	core				
16	R	rim	139	G	core				
17	R	rim	161	V	rim				
17	R	rim	158	L	core				
19	L	core	137	V	core				
20	Q	core	138	F	core				
20	Q	core	140	Y	rim				
20	Q	core	156	N	rim				
20	Q	core	158	L	core				
23	M	core	138	F	core				
23	M	core	97	L	rim				
44	H	rim	159	E	rim				
67	L	rim	133	R	rim				
69	I	core	94	Q	rim				
88	N	rim	94	Q	rim				
96	K	core	94	Q	rim				
96	K	core	98	I	core				
99	T	core	95	F	rim				
99	T	core	98	I	core				
99	T	core	99	K	rim				
100	D	rim	98	I	core				

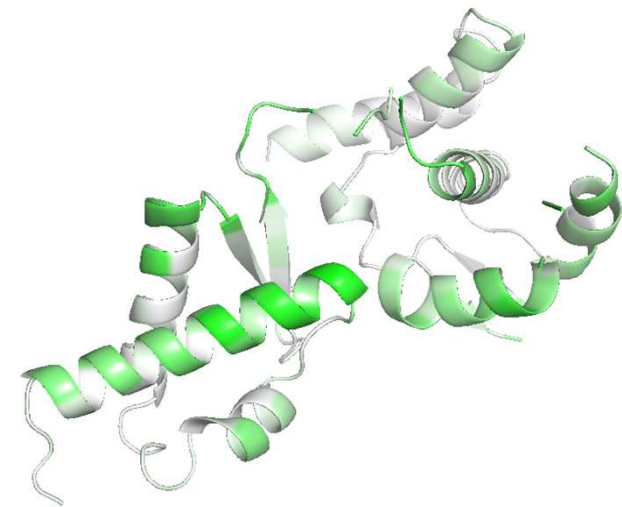
# Silné/důležité interakce (komplexy) jsou evolučně konzervované

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



## InterEvDOCK

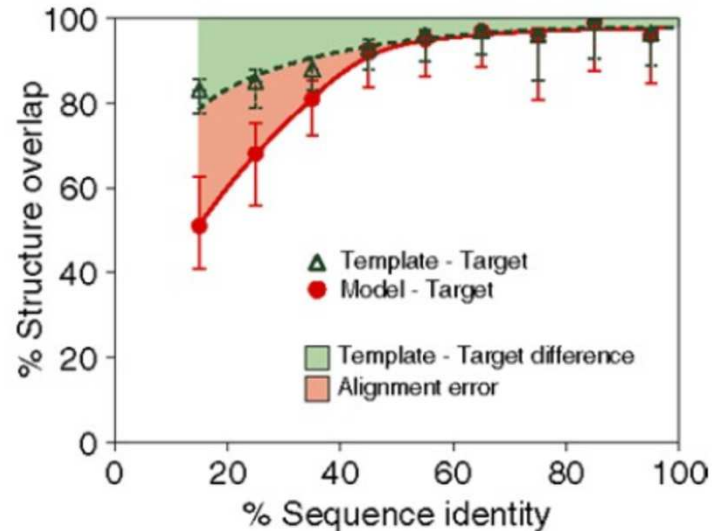
docking na základě templátu a homologie



PYMOL

Eichborn et al, Genome Inf, 2009

## 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* velmi 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)**





## Search Results

Gene / Identifier Search

nse3



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

5 2

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

[\[details\]](#)

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* ...

Experimental Evidence Code	Role	Dataset	Throughput	Curated By	Notes
<b>člověka <i>H. sapiens</i> ...</b> Affinity Capture-MS	HIT	Doyle JM (2010)	Low	BioGRID	-
	BAIT	Taylor EM (2008)	Low	BioGRID	-
	BAIT	Huttlin EL (2015)	High	BioGRID	
	BAIT	Huttlin EL (2017)	High	BioGRID	
<b>Affinity Capture-Western</b>	BAIT	Doyle JM (2010)	Low	BioGRID	-
	BAIT/HIT	Taylor EM (2008)	Low	BioGRID	-
	HIT	Hudson JJ (2011)	Low	BioGRID	-
<b>Co-crystal Structure</b>	HIT	Newman JA (2016)	Low	BioGRID	-
<b>Co-fractionation</b>	BAIT	Taylor EM (2008)	Low	BioGRID	-
	BAIT	Wan C (2015)	High	BioGRID	
<b>Reconstituted Complex</b>	BAIT/HIT	Doyle JM (2010)	Low	BioGRID	-

### NSE1 | YLR007W

## pučící kvasinky *S. cerevisiae*

5 2

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

[\[details\]](#)

Experimental Evidence Code	Role	Dataset	Throughput	Curated By	Notes
<b>Affinity Capture-MS</b>	HIT	Hazbun TR (2003)	High	BioGRID	-
<b>Reconstituted Complex</b>	HIT	Hudson JJ (2011)	Low	BioGRID	-
<b>Two-hybrid</b>	HIT	Hazbun TR (2003)	High	BioGRID	-
	HIT	Hu B (2005)	Low	BioGRID	-
	BAIT/HIT	Duan X (2009)	Low	BioGRID	-

### NSE1 | SPCC550.05

## poltivé kvasinky *S. pombe*

4

Smc5-6 complex ubiquitin-protein ligase E3 subunit 1

[\[details\]](#)

Experimental Evidence Code	Role	Dataset	Throughput	Curated By	Notes
<b>Affinity Capture-Western</b>	BAIT	Pebernard S (2004)	Low	BioGRID	-
<b>Reconstituted Complex</b>	BAIT	Sergeant J (2005)	Low	BioGRID	-
<b>Two-hybrid</b>	BAIT	Pebernard S (2004)	Low	BioGRID	-
	HIT	Hudson JJ (2011)	Low	BioGRID	-

# 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 "Finley Lab" website, which is the "Center for Molecular Medicine and Genetics". The page lists "Finley Lab Interactions Databases" and "Gene or Protein Interactions Databases in the research community".

**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 fúz)

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

Více Dr. Potěšil



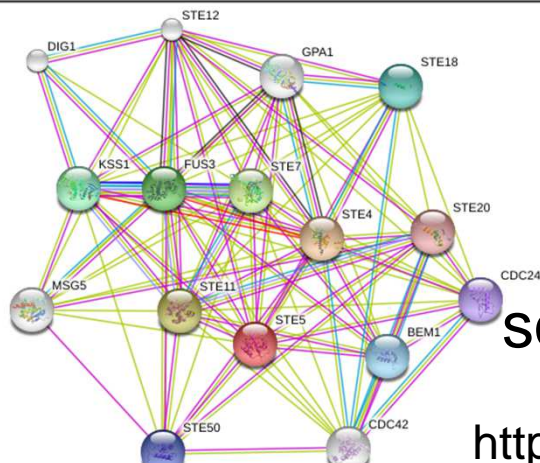
# Informační zdroje PPI

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

Database	Proteins/Domains	Type	Number of Interactions
DIP <sup>a</sup> , LiveDIP	P	E,S	55,733
BIND <sup>a</sup>	P	E,C,S	83,517
MPact/MIPS <sup>a</sup>	P	E,C,F	15,488 (4,300) <sup>b</sup>
STRING	P ←	E,P,F	730,000 (proteins)
MINT <sup>a</sup>	P	E,C	71,854
IntAct <sup>a</sup>	P ←	E,C	68,165
BioGRID <sup>a</sup>	P ←	E,C	116,000 (30,000) <sup>b</sup>
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]	<a href="http://dip.doe-mbi.ucla.edu">http://dip.doe-mbi.ucla.edu</a>
BIND [105]	<a href="http://bind.ca">http://bind.ca</a>
MPact/MIPS [97]	<a href="http://mips.gsf.de/services/ppi">http://mips.gsf.de/services/ppi</a>
STRING [119]	<a href="http://string.embl.de">http://string.embl.de</a>
MINT [120]	<a href="http://mint.bio.uniroma2.it/mint">http://mint.bio.uniroma2.it/mint</a>
IntAct [121]	<a href="http://www.ebi.ac.uk/intact">http://www.ebi.ac.uk/intact</a>
BioGRID [122]	<a href="http://www.thebiogrid.org">http://www.thebiogrid.org</a>
HPRD [123]	<a href="http://www.hprd.org">http://www.hprd.org</a>
ProtCom [124]	<a href="http://www.ces.clemson.edu/compbio/ProtCom">http://www.ces.clemson.edu/compbio/ProtCom</a>
3did [108], Interprets[125]	<a href="http://gatealoy.pcb.ub.es/3did/">http://gatealoy.pcb.ub.es/3did/</a>
Pibase [107], ModBase [126]	<a href="http://alto.compbio.ucsf.edu/pibase">http://alto.compbio.ucsf.edu/pibase</a>
CBM [26]	<a href="ftp://ftp.ncbi.nlm.nih.gov/pub/cbm">ftp://ftp.ncbi.nlm.nih.gov/pub/cbm</a>
SCOPPI [111]	<a href="http://www.scoppi.org/">http://www.scoppi.org/</a>
iPfam [127]	<a href="http://www.sanger.ac.uk/Software/Pfam/iPfam">http://www.sanger.ac.uk/Software/Pfam/iPfam</a>
InterDom [128]	<a href="http://interdom.lit.org.sg">http://interdom.lit.org.sg</a>
DIMA [129]	<a href="http://mips.gsf.de/genre/proj/dima/index.html">http://mips.gsf.de/genre/proj/dima/index.html</a>
Prolinks [104]	<a href="http://prolinks.doe-mbi.ucla.edu/cgi-bin/functionator/pronav/">http://prolinks.doe-mbi.ucla.edu/cgi-bin/functionator/pronav/</a>



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

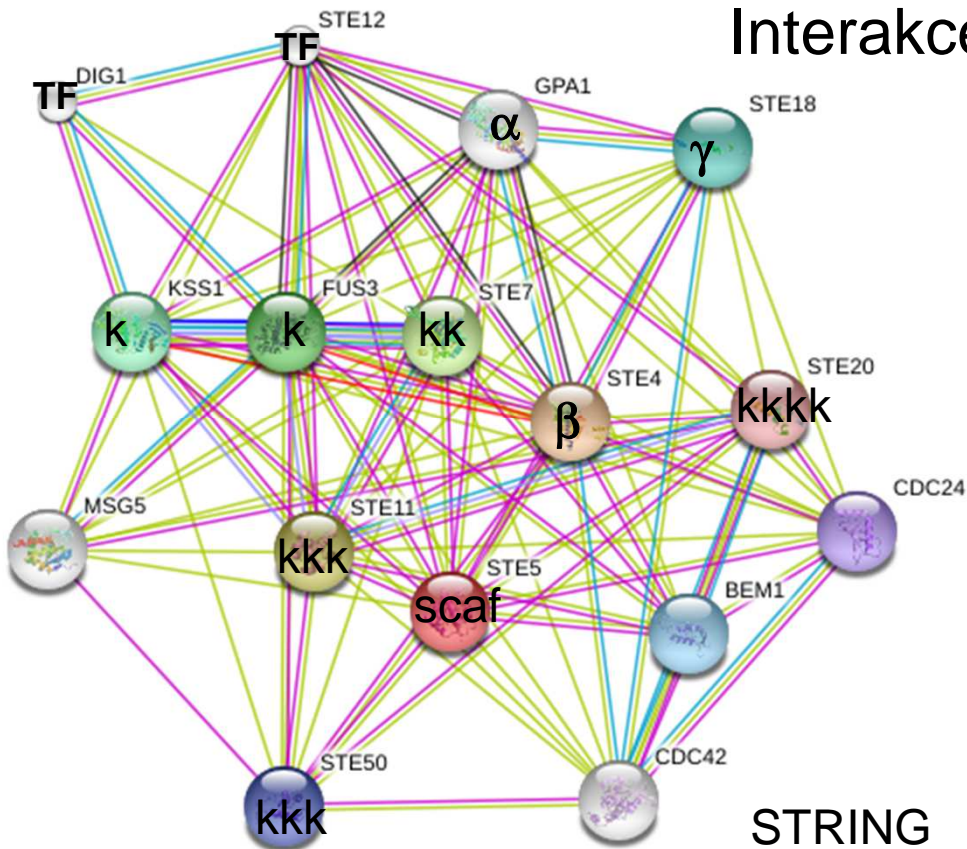
souhrn proteinových interakcí = **interaktom**

Shoemaker and Panchenko, PLoS Comp Biol, 2007






<http://string-db.org>

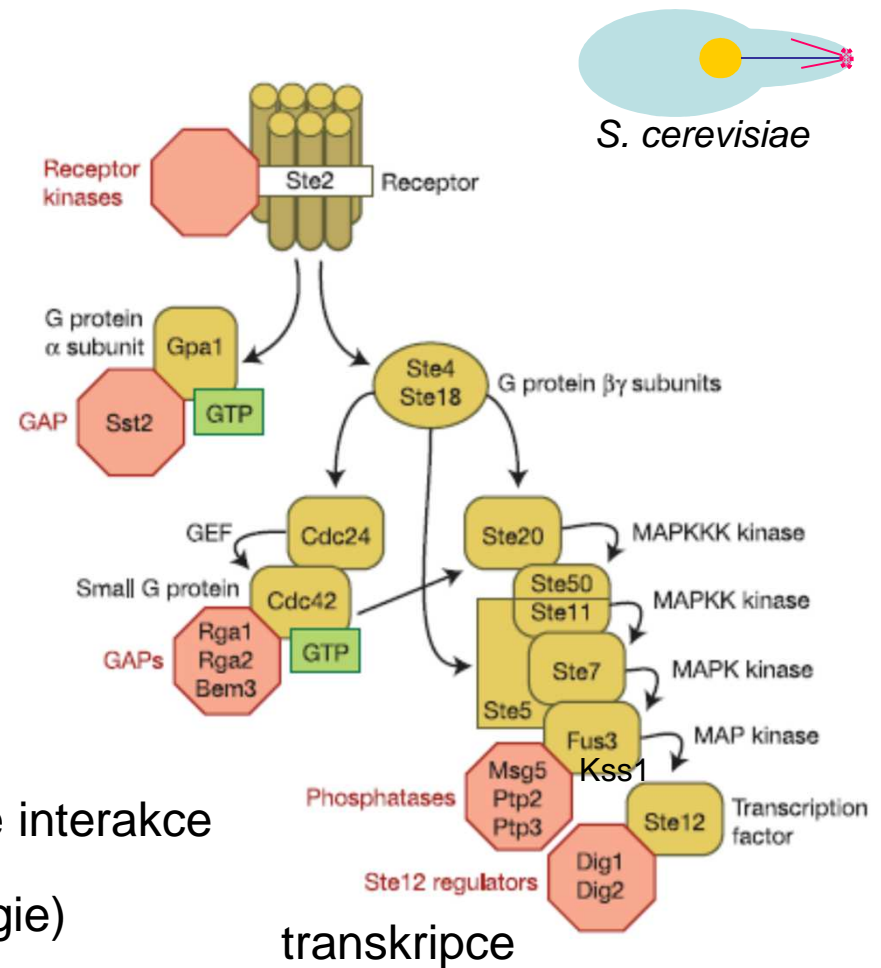
# proteinové sítě – chybí info o posloupnosti, síle ... interakcí

## Interakce x signální dráha



STRING

-  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

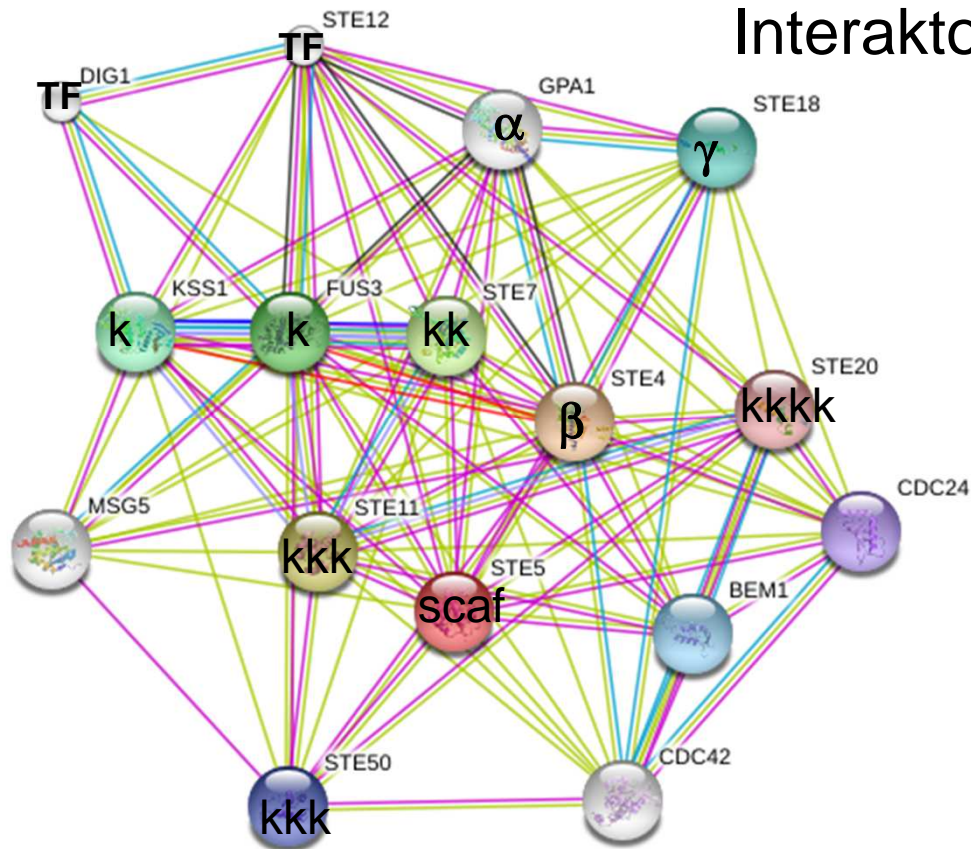


transkripce



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

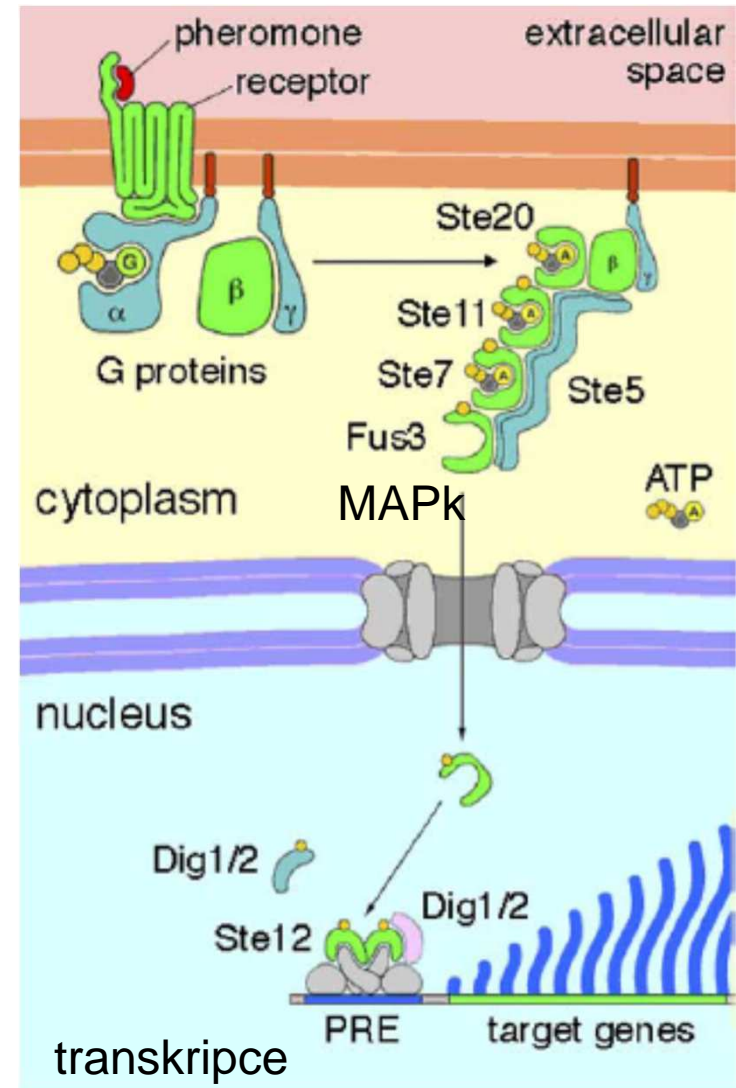
## Interaktom x komplexom



- Experiments
- Databases
- Textmining
- Gene Fusion
- Coexpression

Síť neznamená komplex,  
ale vztah

souhrn proteinových  
komplexů = **komplexom**



Wang et al., Nature, 2004

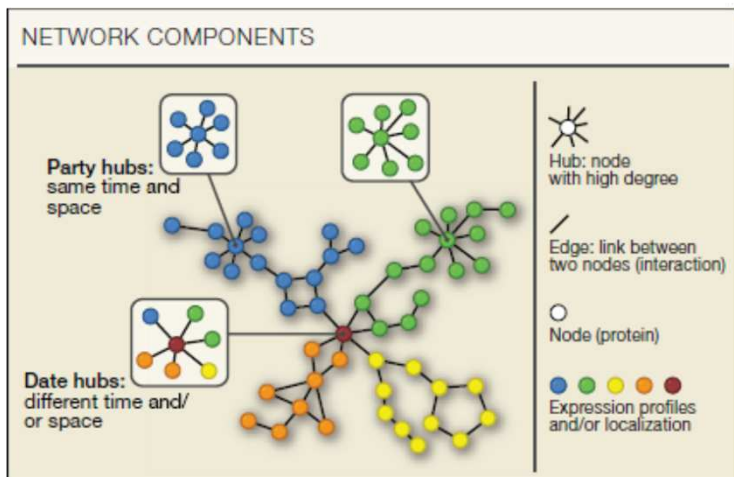
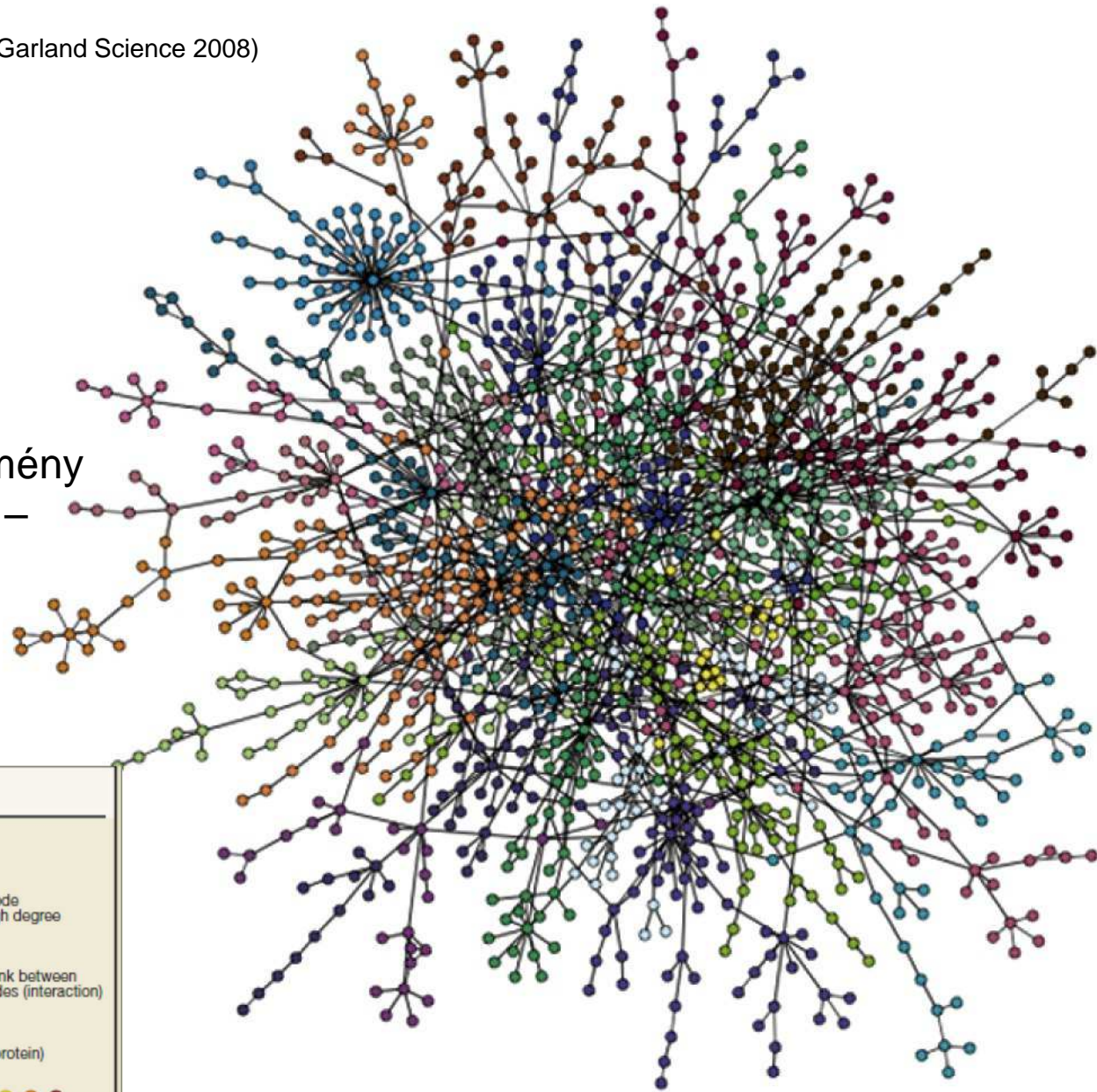


# Interaktom x komplexom

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

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

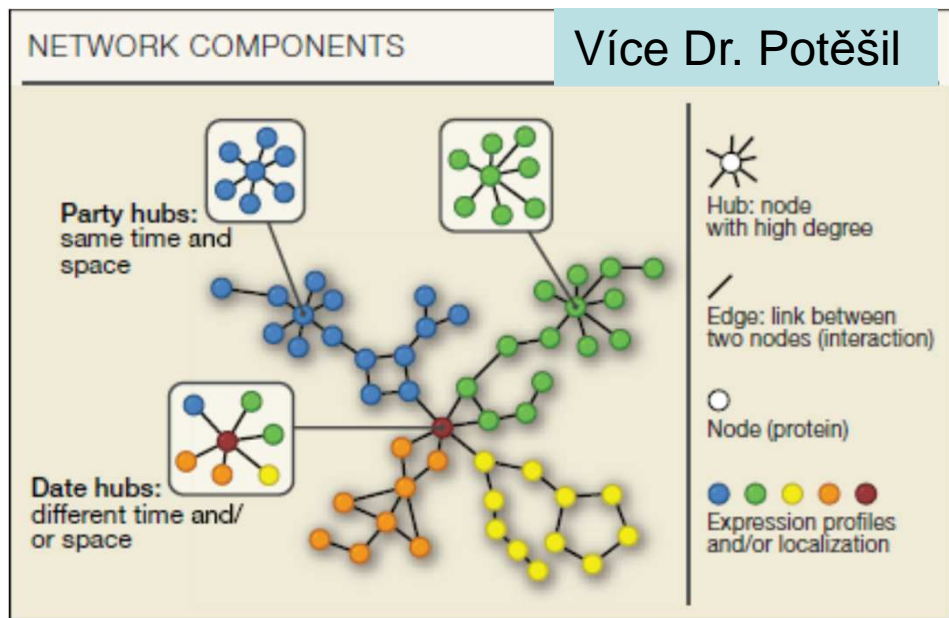
Modularita – interagují domény (jeden protein více domén – zapojení do více procesů)



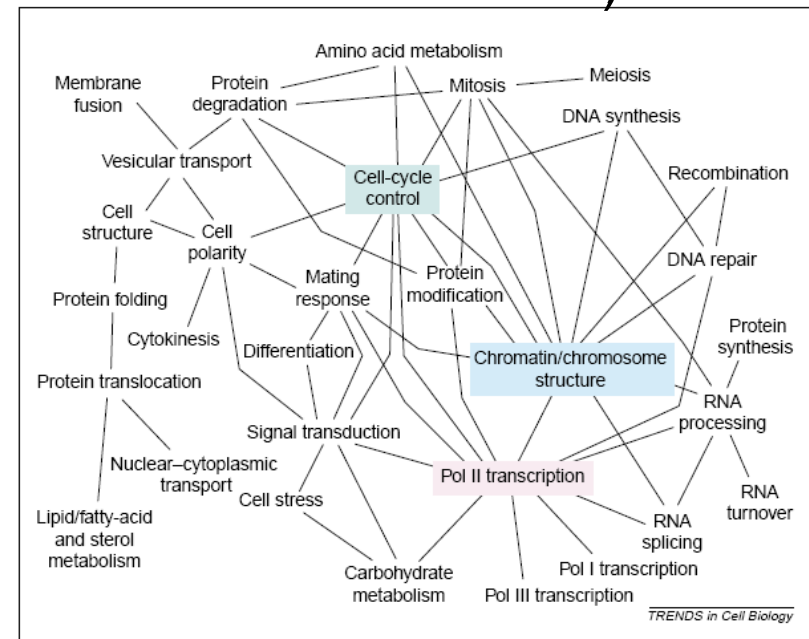
Seebacher & Gavin, Cell (SNAP SHOT), 2011

# 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, SUMO)
- souhrn proteinových interakcí = **interaktom**  
(modularita díky interakcím domén – různé kombinace domén)

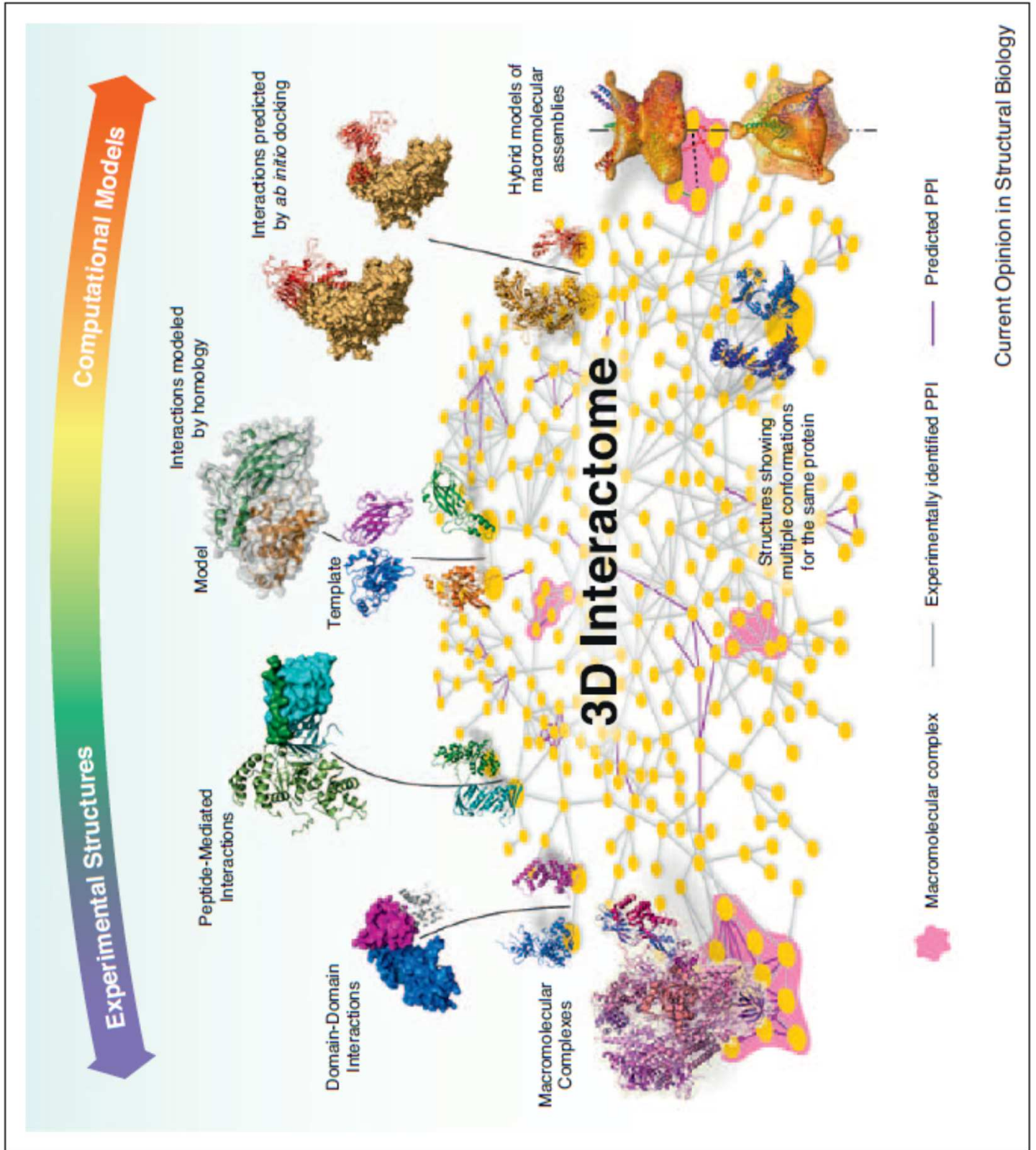


Seebacher & Gavin, Cell (SNAP SHOT), 2011



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







# 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: ...)
  - predikce PPI je obtížná (založená na podobnosti komplexů)

