

# PROTEOMIKA

doc. Jan Paleček  
jpalecek@sci.muni.cz

- Proteinové interakce – 31.10.
  - jak spolu proteiny interagují?
  - interaktom
- Proteinové komplexy – 7.11.
  - protein-proteinové interakce a komplexy
  - komplexom, architektura a funkce komplexů

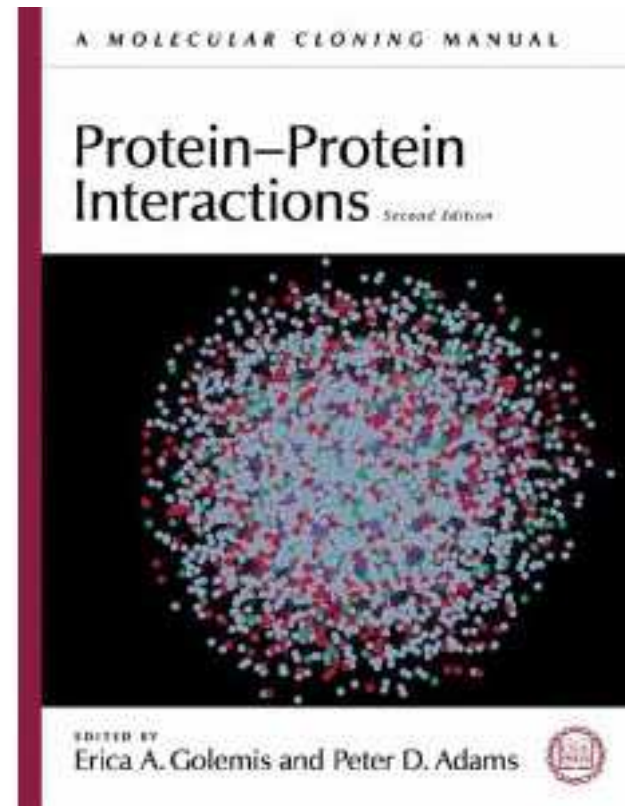
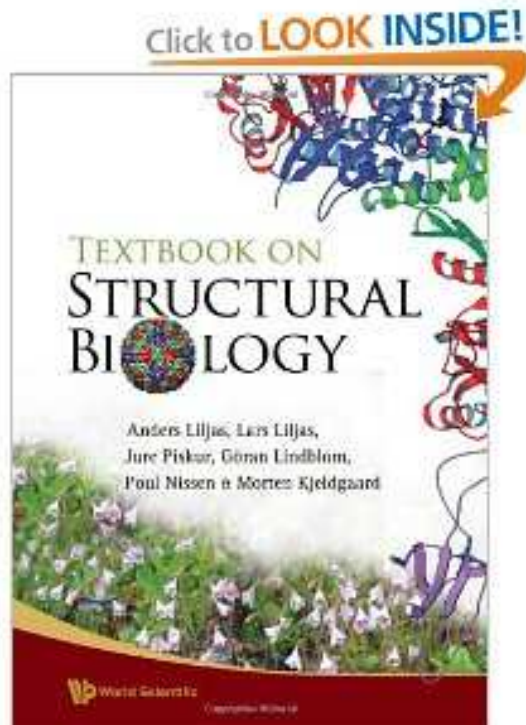
**CG030 – Struktura a funkce proteinových komplexů**

# Informační zdroje

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

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

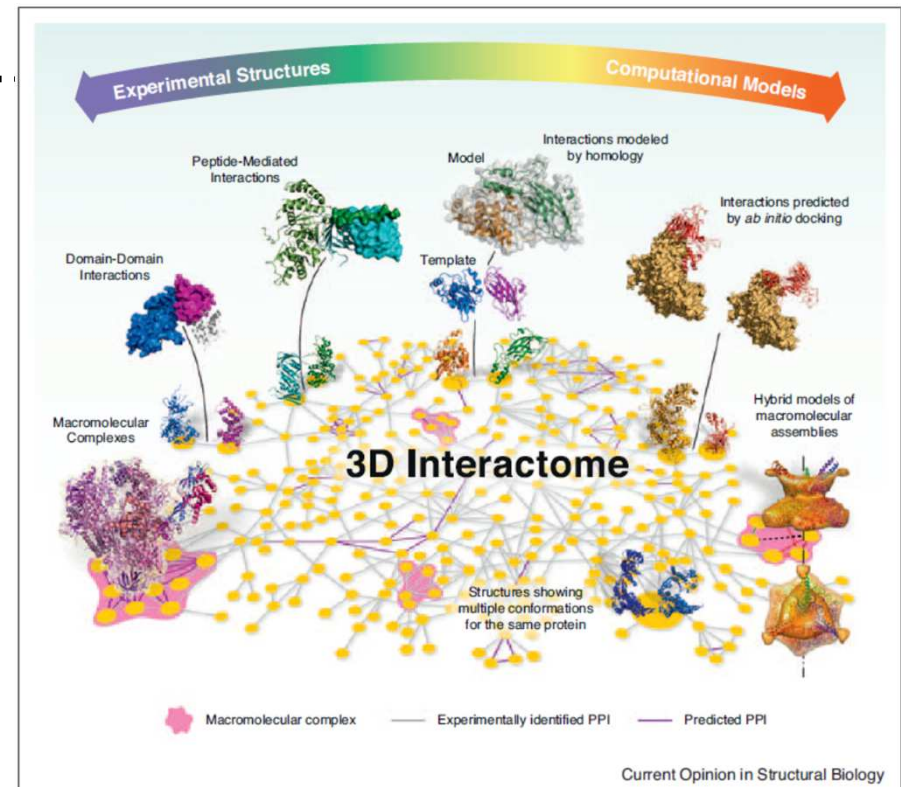
**... nejnovější články z časopisů Nature, ...**

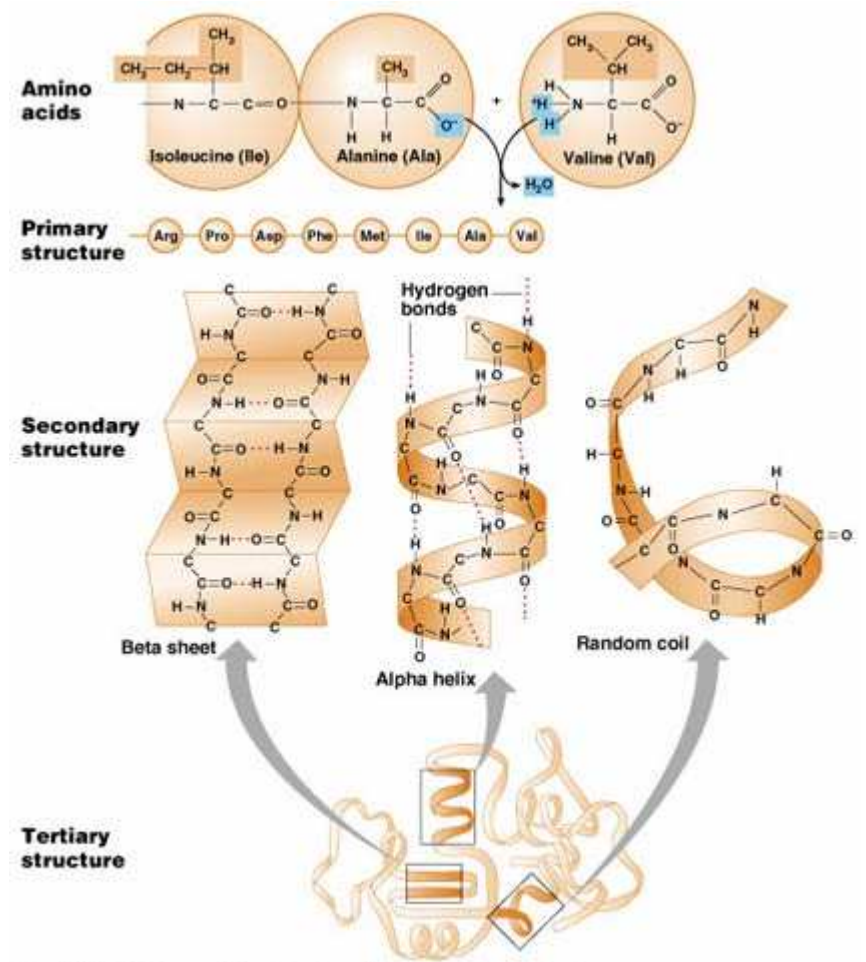
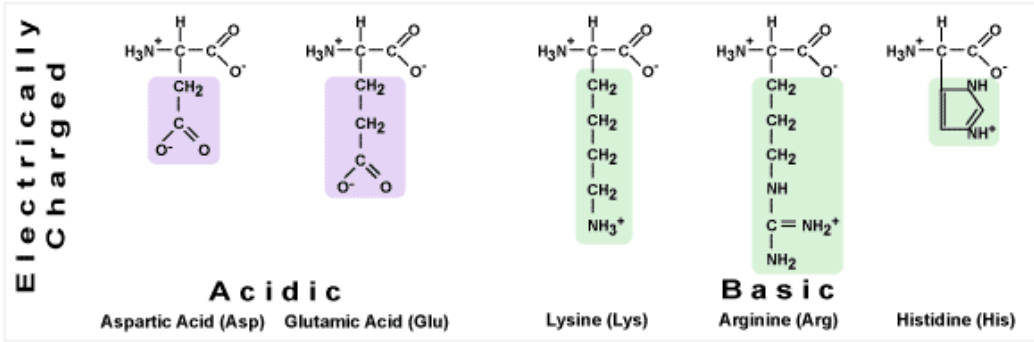
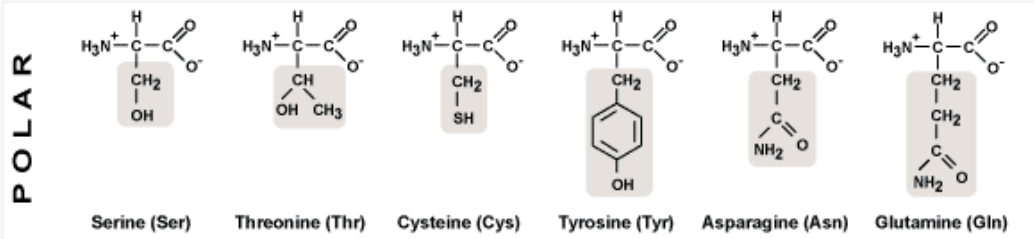
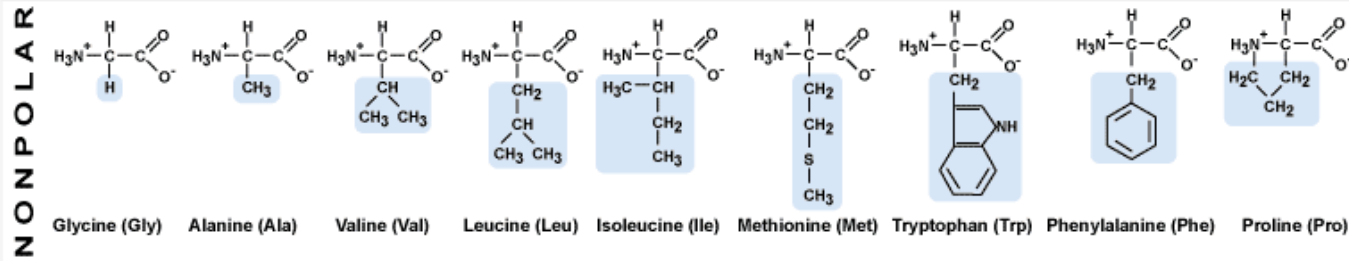


Databáze proteinových struktur: <http://www.rcsb.org/pdb/home/home.do>,  
<http://www.ebi.ac.uk/pdbsum/>

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

- Proteinové interakce – 31.10.
  - Interakce: od primární po terciární strukturu
  - Typy vazeb ...
  - Informatika:
    - databáze struktur, interakcí ...
    - docking ...
    - motivy, evoluční aspekty ...
    - nástroje ...
  - Interaktom ...



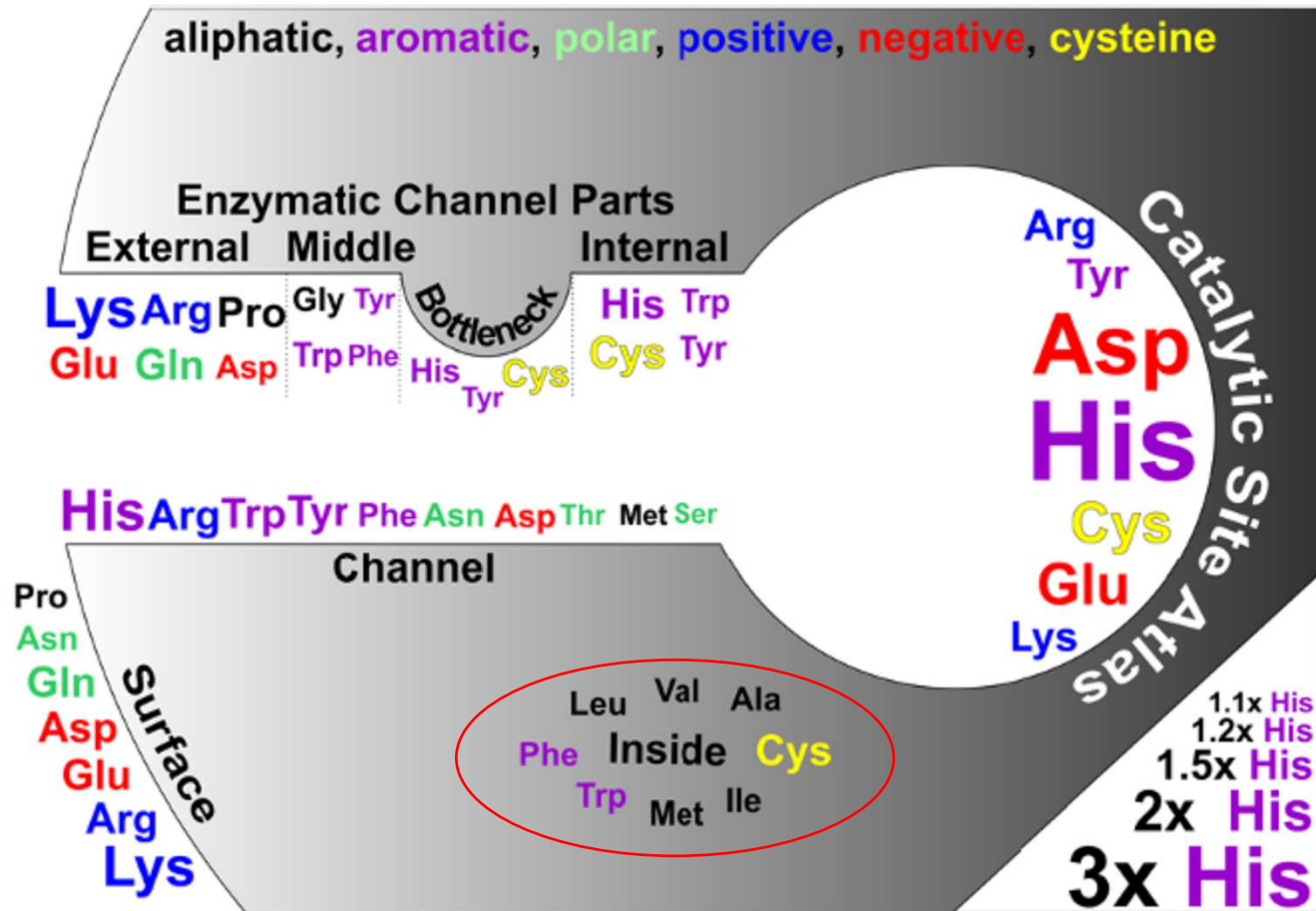


Copyright © 2003 Pearson Education, Inc., publishing as Benjamin Cummings.

# Základní proteinové charakteristiky

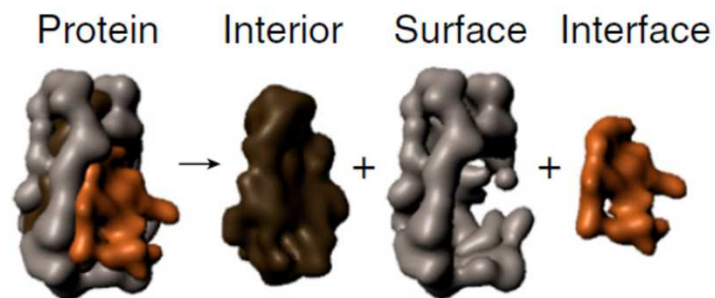
- Primární
- Sekundární
- Terciární
- Kvarterní – dva proteiny a více ...

# Podíl AMK (primární struktury) na proteinových interakcích

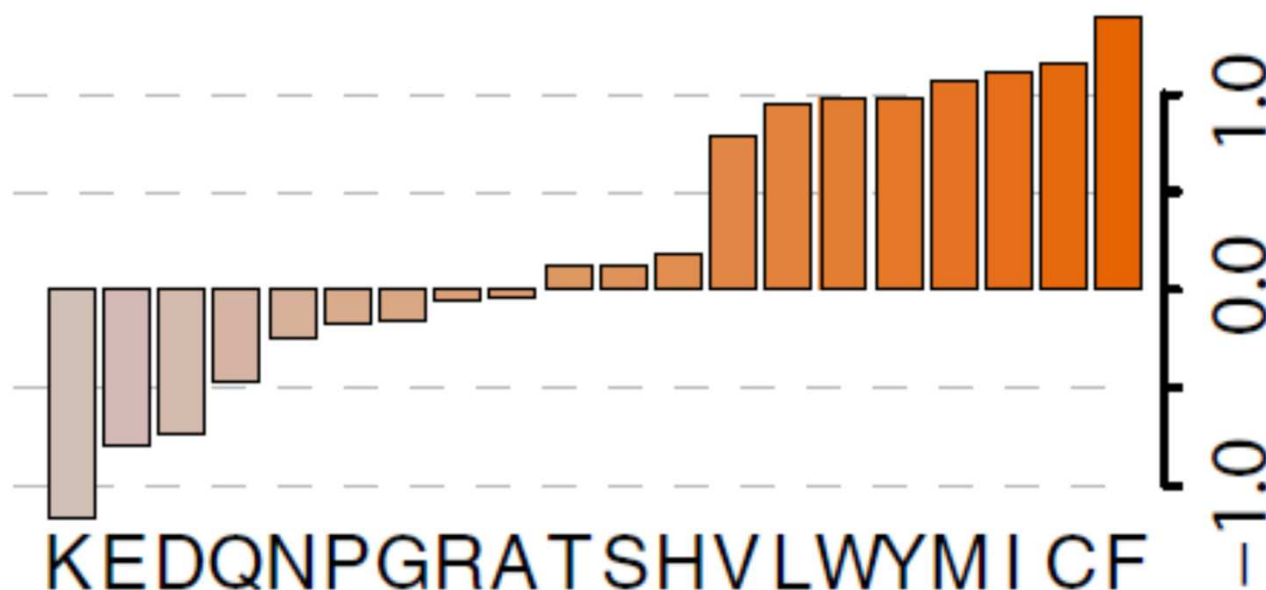


- uvnitř hydrofobní, povrch polární/nabitý (do solventu/vody), ale katalytická centra (tunely) jsou také polární a nabité (katalýza biochemické reakce)

# PPI od primární struktury ...



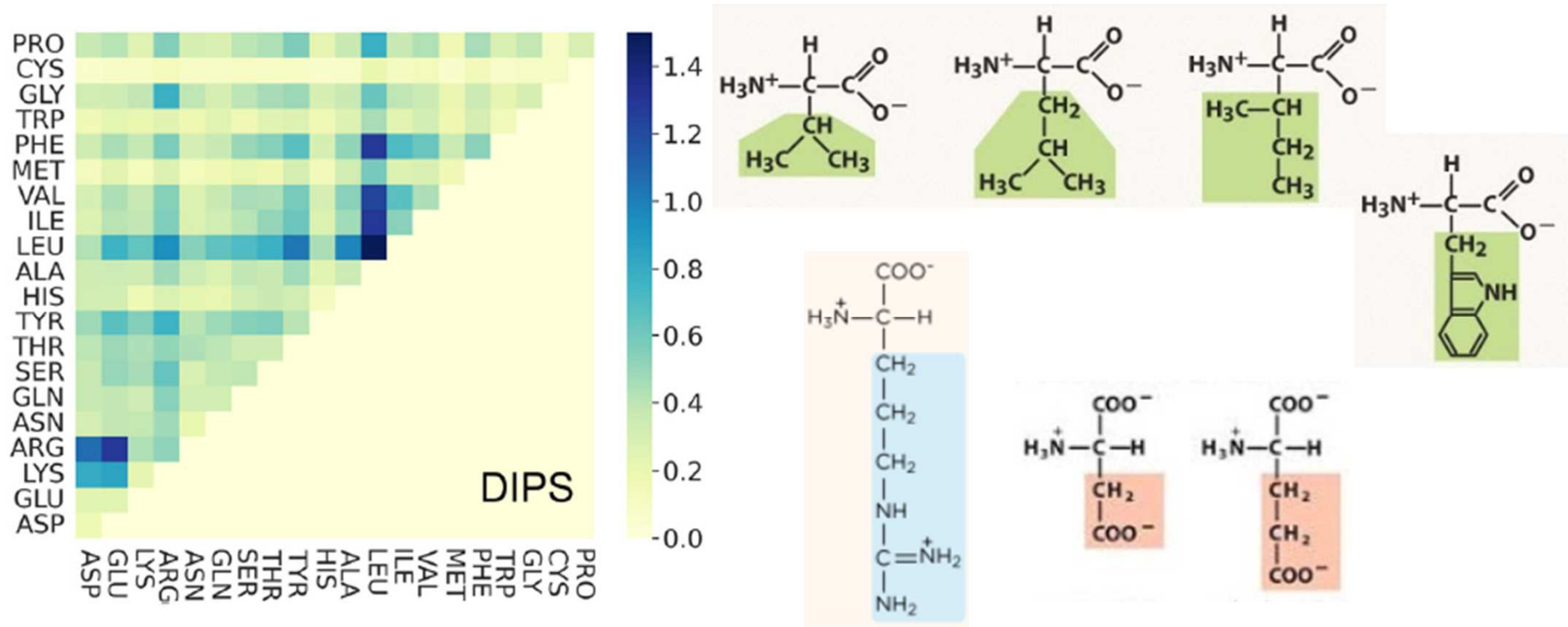
amino acid interface propensity  
 $\log\left(\frac{\text{freqAA}_{\text{interface}}}{\text{freqAA}_{\text{surface}}}\right)$   
 proxy for amino acid "stickiness" scale



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

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

# PPI od primární struktury ...

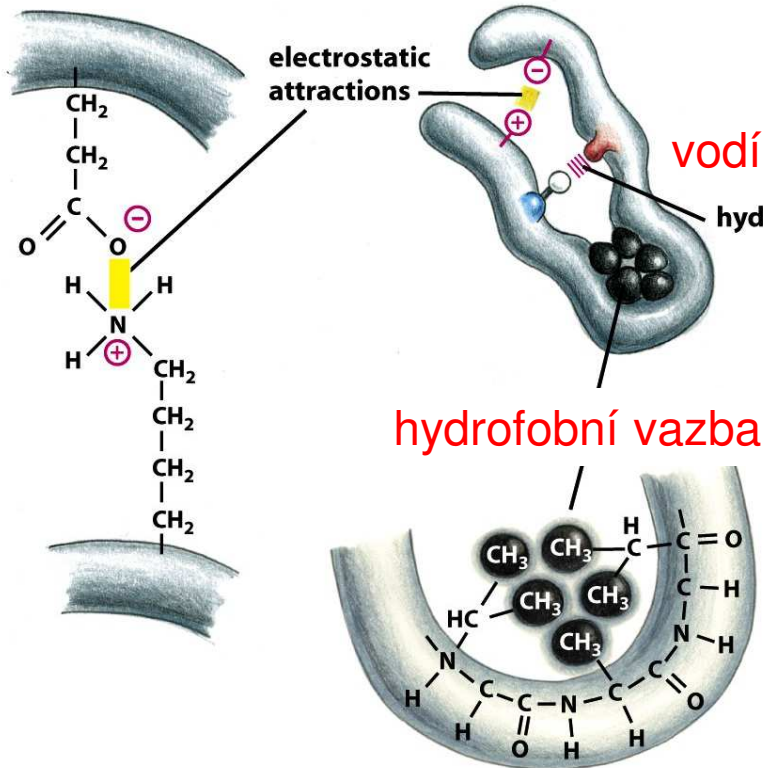


Leu s Ile, Leu, Val a Phe = hydrofobní AMK

Arg(+) s Asp(-) a Glu(-) = nabité AMK

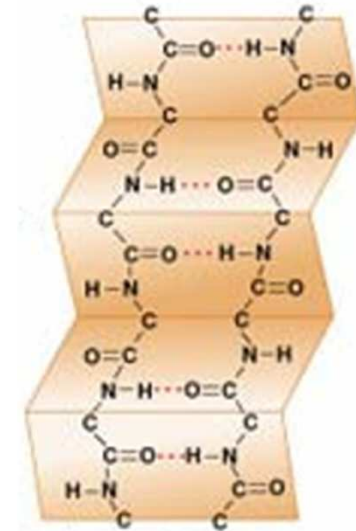
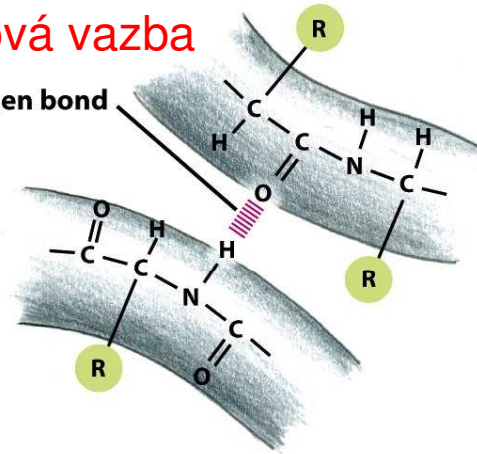
# Typy vazeb v PPI

iontová (polární) vazba



vodíková vazba

hydrogen bond



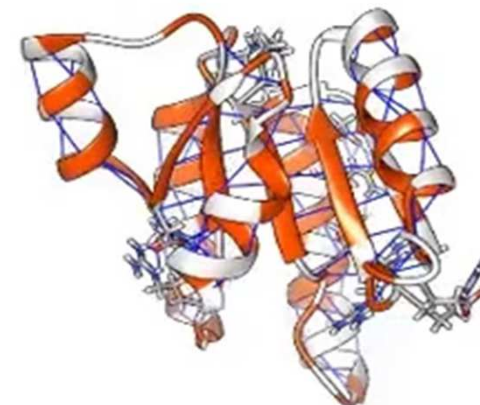
$\beta$ -listy

● H-bonds  
● salt bridges  
● disulfide bonds

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

Která z těchto vazeb převažuje v PPI?

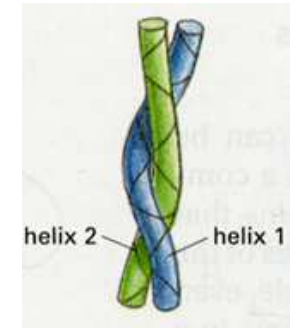
Kovalentní vazba = modifikace (nikoli PPI)  
vyjimečně např. disulfidické můstky  
postranslační modifikace (ubikvitinylace,  
SUMOylace)





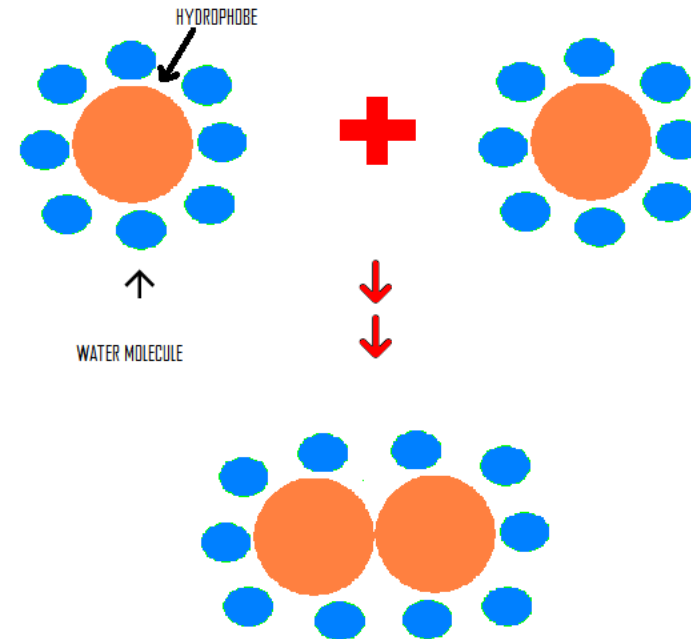
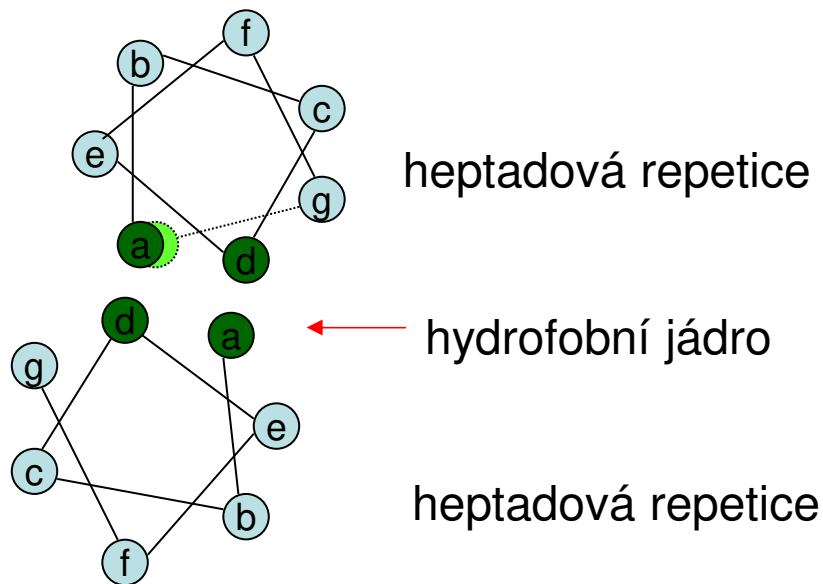
# ... sekundární struktury ...

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

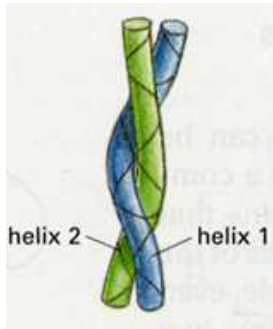


## coiled-coil struktura

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

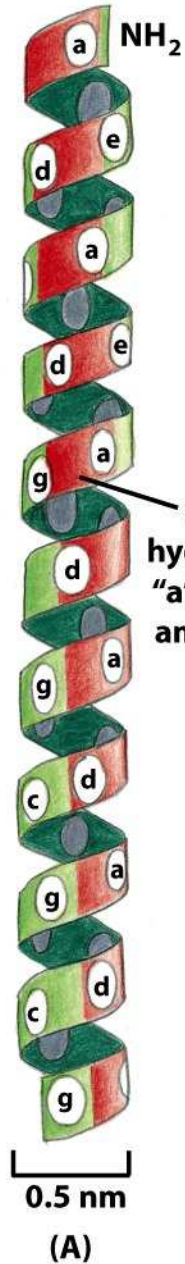


...LKS**L**HNQ**L**RD**L**EES**L**TH...

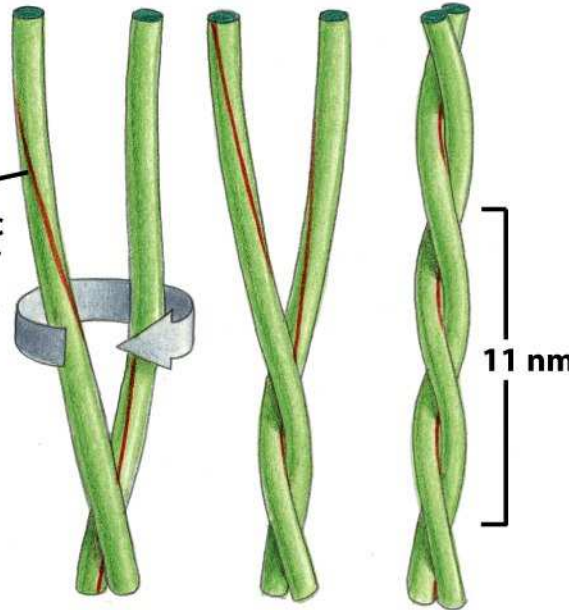


# coiled-coil struktura

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



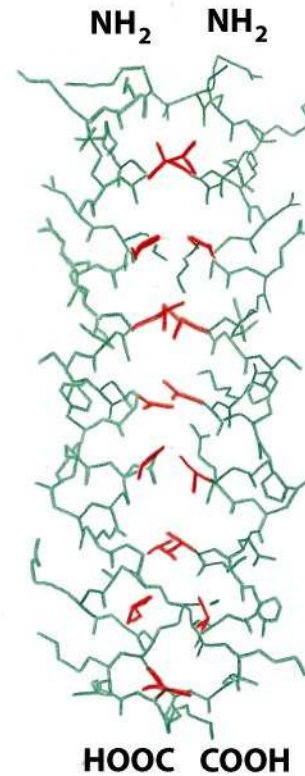
stripe of hydrophobic "a" and "d" amino acids



hydrofobní jádro

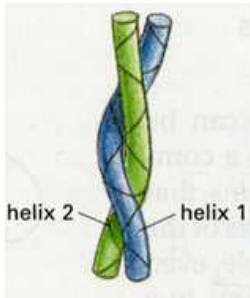
...LKS**L**HNQ**L**RD**L**EES**L**TH...

(B)



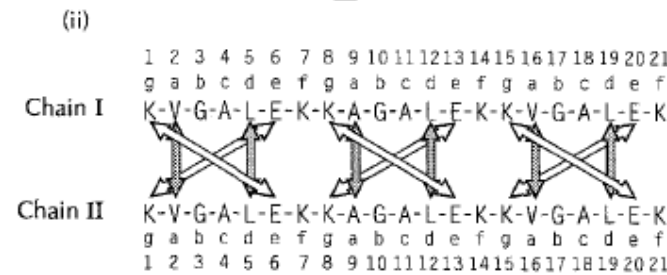
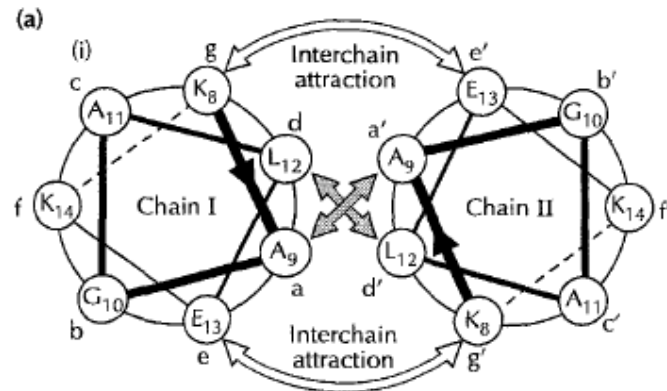
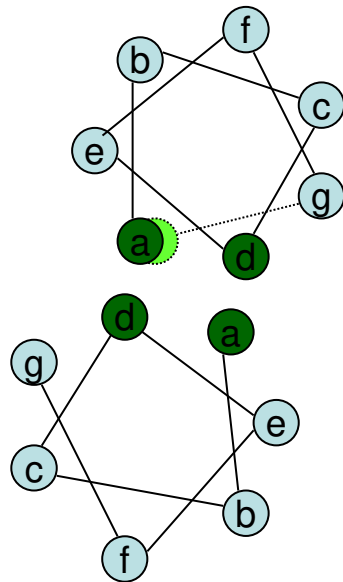
paralelní šroubovice

(C)

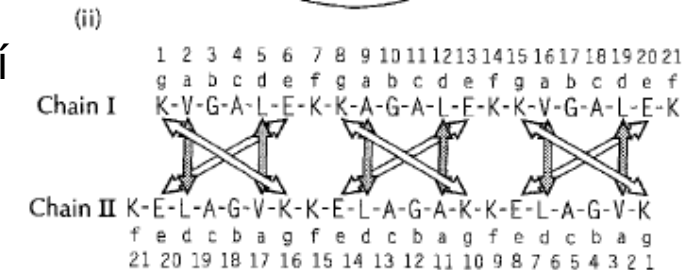
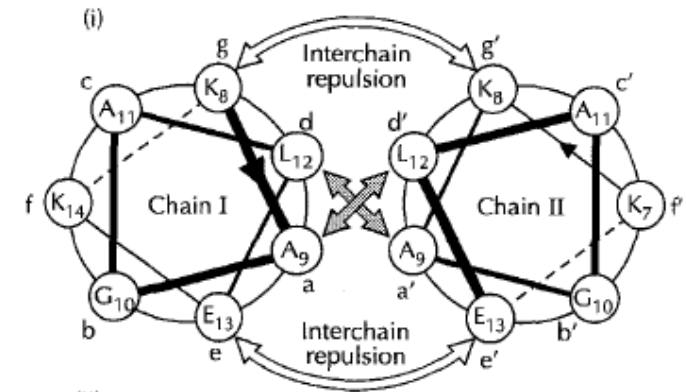


# coiled-coil struktura

Síla interakce může být ovlivněna sousedními AMK

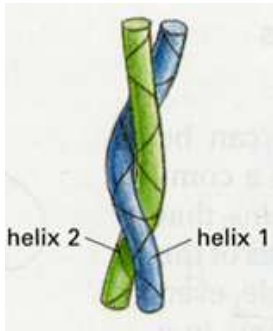


Sousední AMK stabilizují interakce šroubovic



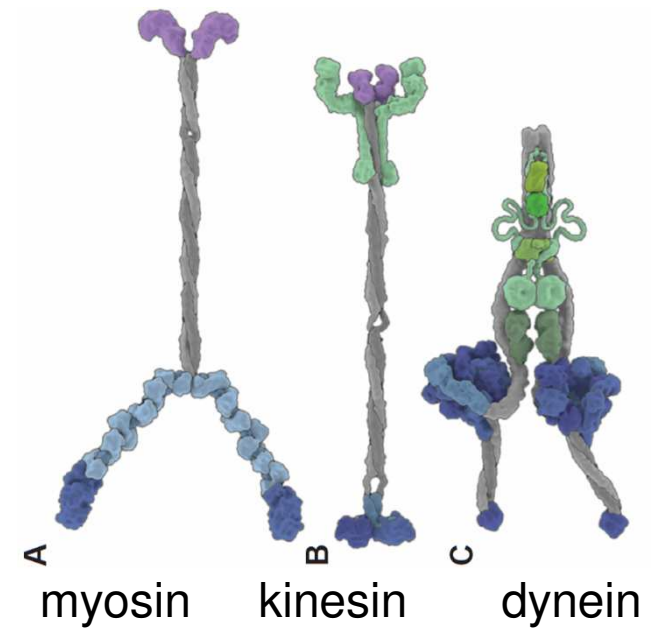
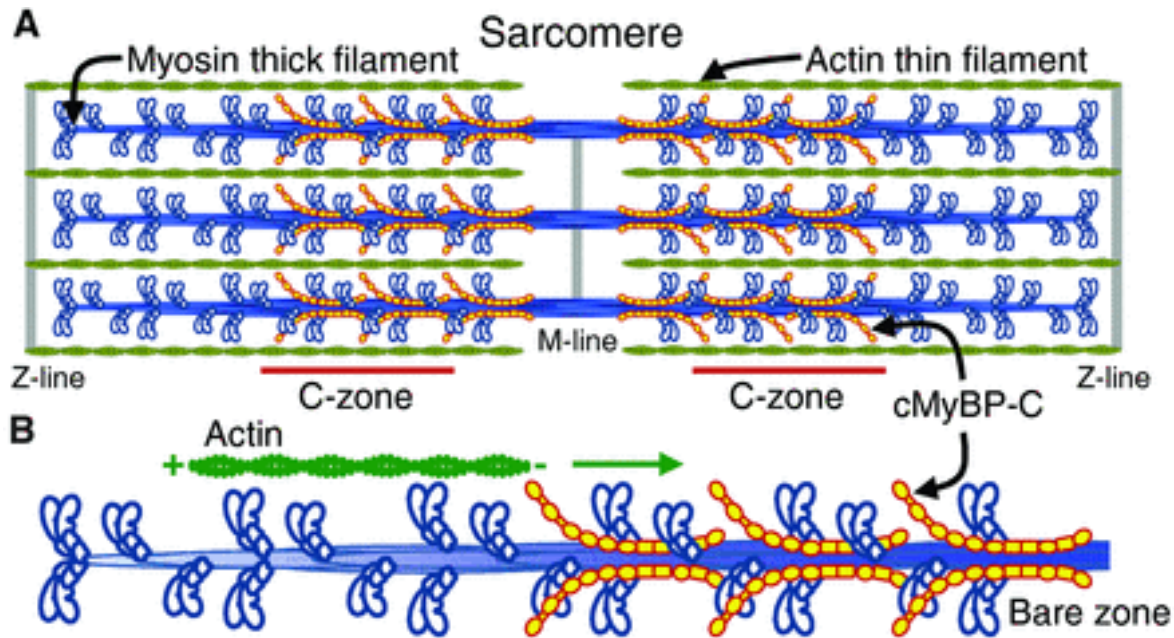
Sousední AMK destabilizují interakce šroubovic

Adamson et al.: CO in Biotech, 1993  
Ivanov et al., PLoS One, 2017



# coiled-coil struktura

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

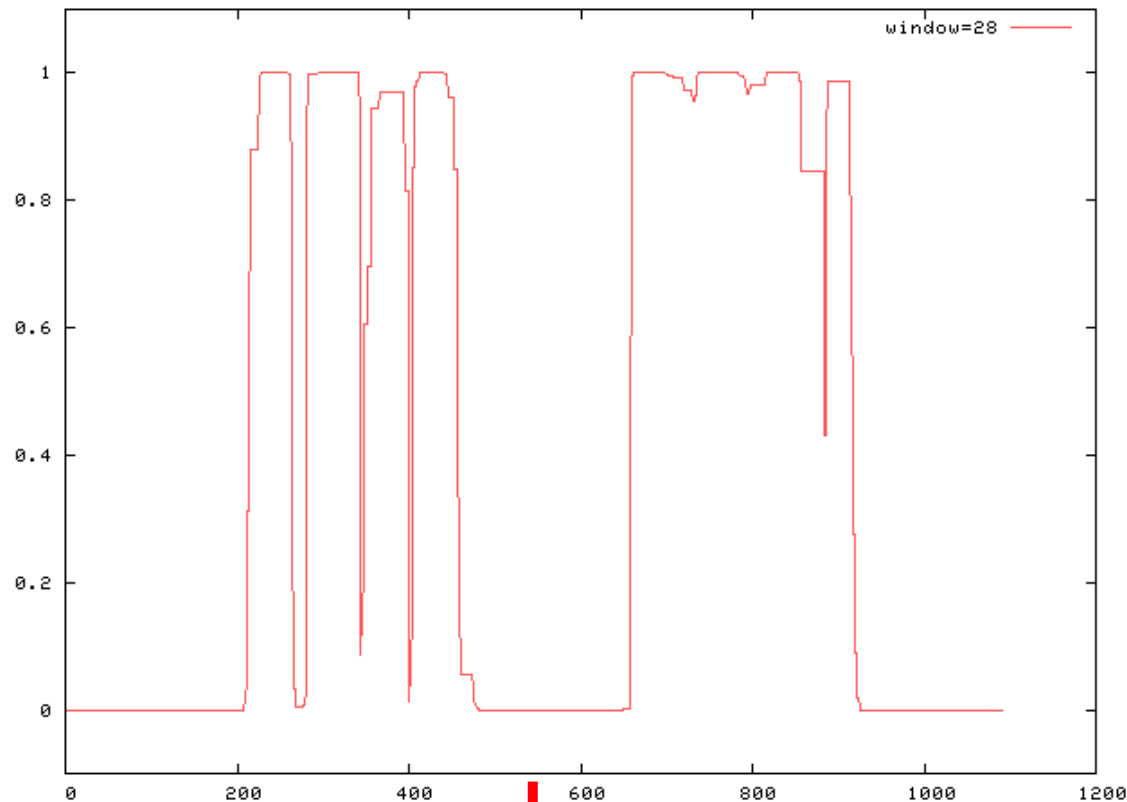


- CC v myosinu je intermolekulární (paralelní)

# coiled-coil struktura

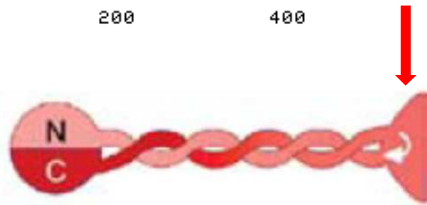
- program COIL: <https://toolkit.tuebingen.mpg.de/tools/deepcoil>

profil SMC6

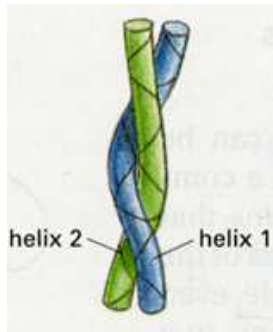


Ludwiczek et al, Bioinformatics, 2019

- CC jako jedna ze sekundárních struktur (viz. Dr. Klumpler)

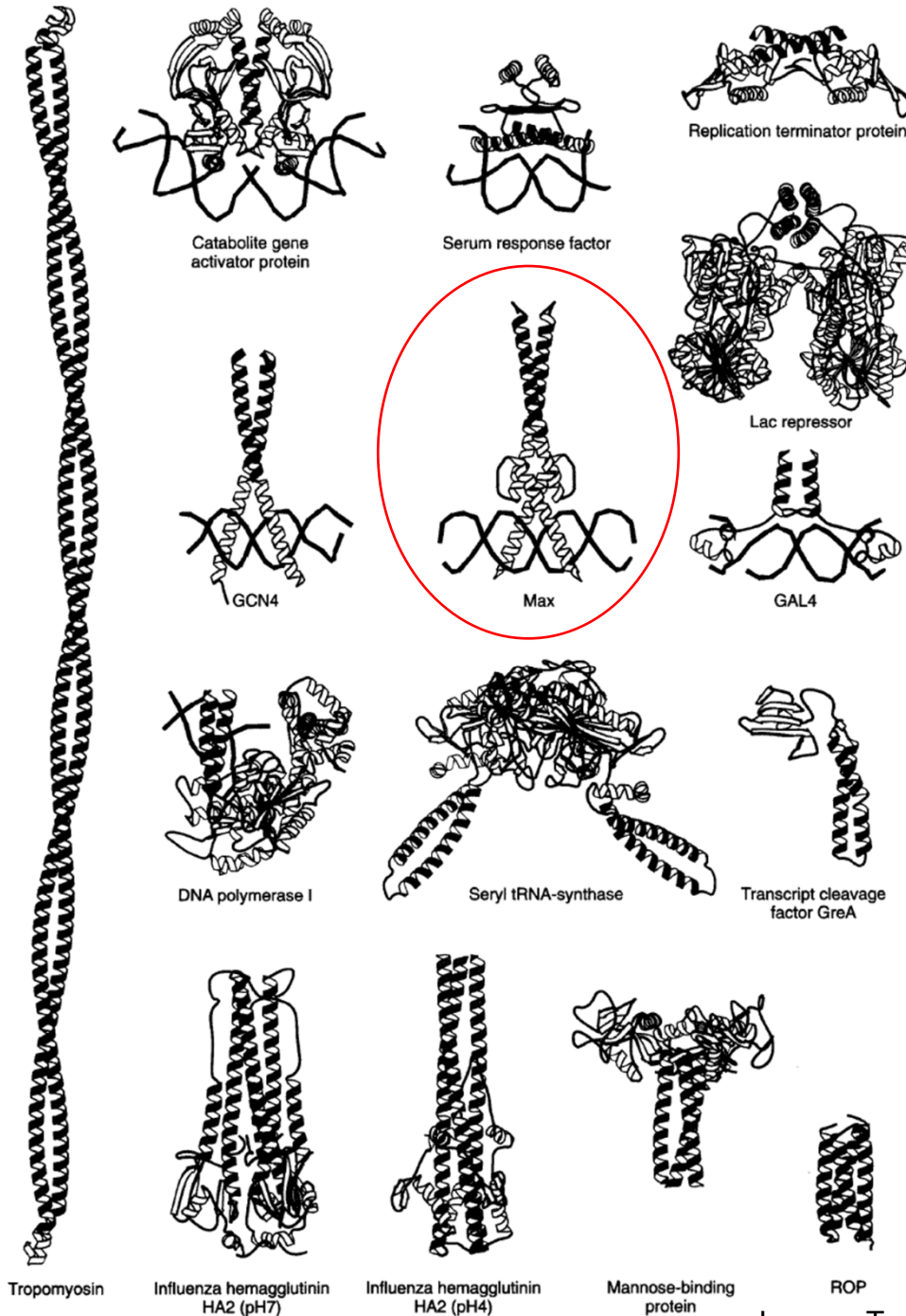


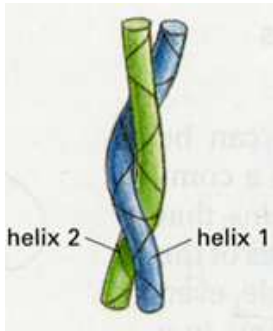
- CC v SMC proteinech jsou intramolekulární (antiparalelní)



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

Intermolekulární - homo- či heterodimery (oligomery)

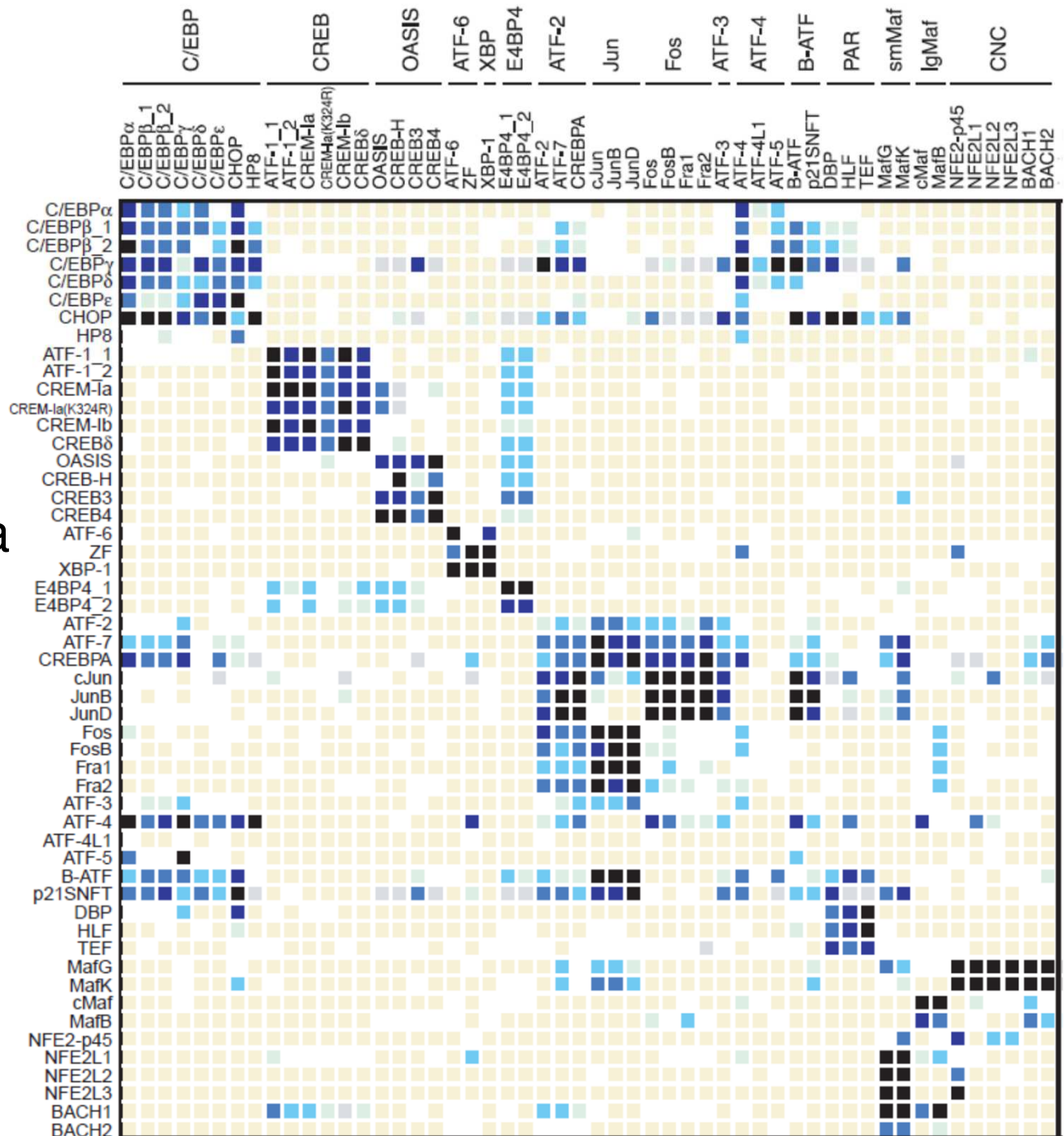




**Coiled-coil**  
doména je  
významným  
**dimerizačním**  
modulem u mnoha  
proteinů: bZIP  
transkripční  
faktory vytváří  
homo- i  
heterodimery

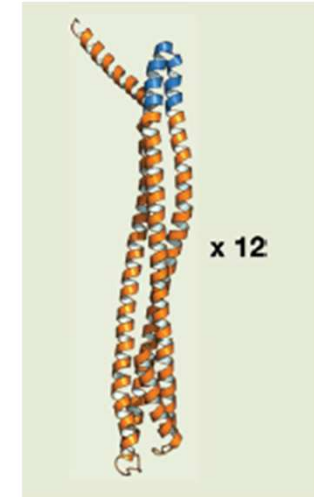
**variabilita**

Newman, Science, 2003

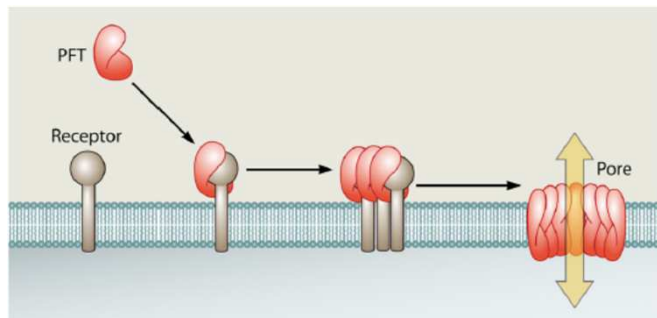
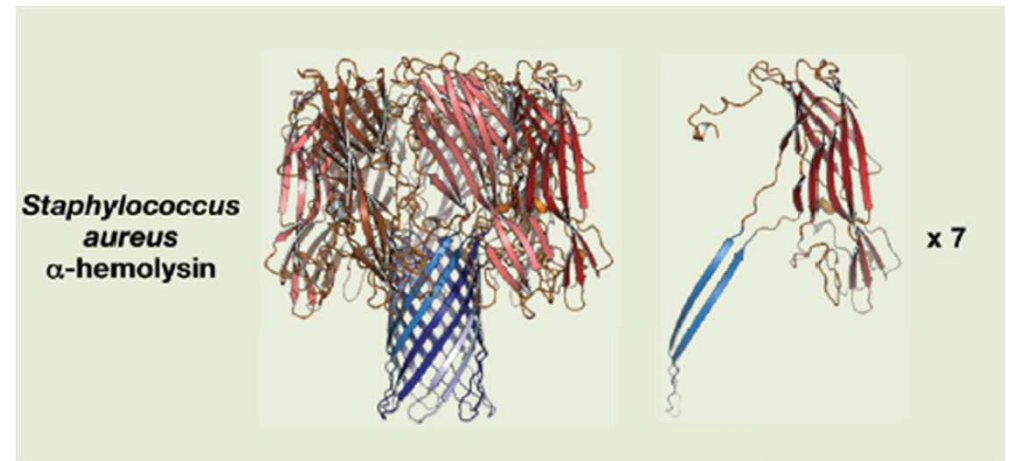


## ... sekundární struktury ...

- **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

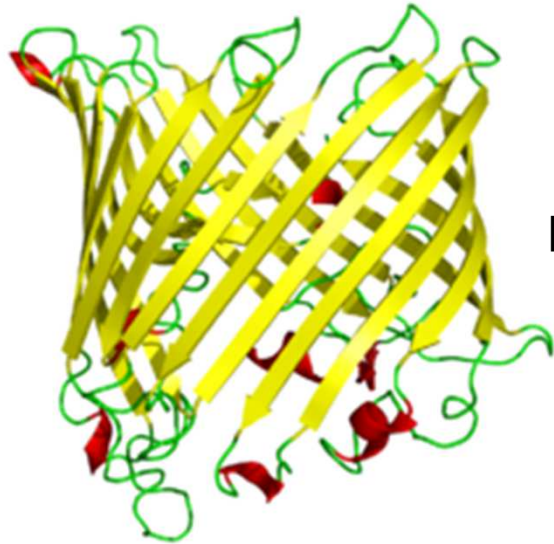


**Toxiny** – podjednotky se skládají tj. vytváří pór až v místě působení (neublíží původní buňce) – lyzuje červené krvinky

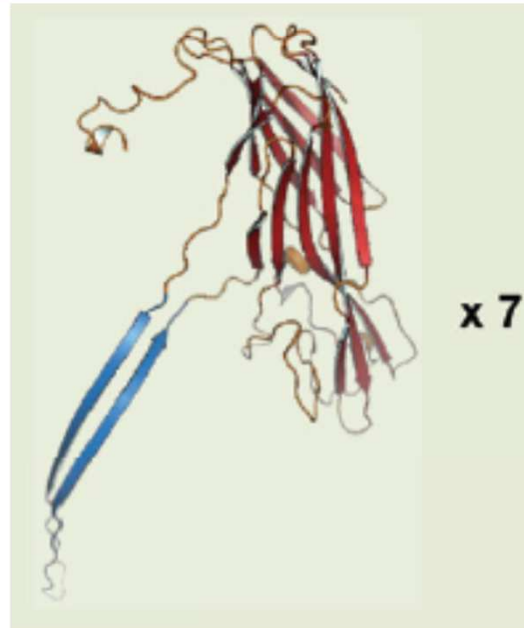
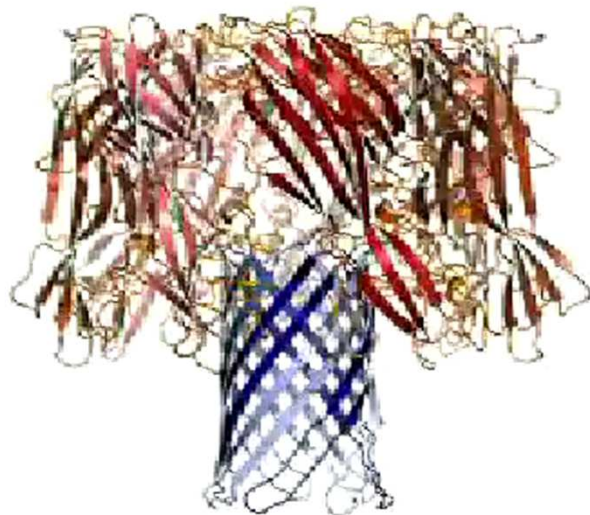
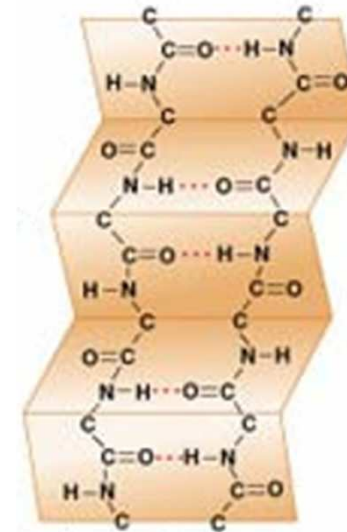


# ... sekundární struktury ...

v interakcích  $\beta$ -listů převažují vodíkové vazby (peptidového řetězce)



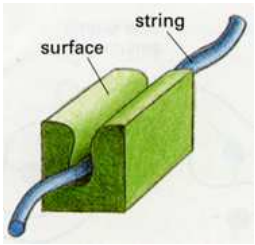
**Porin**  
(1 ORF - polypeptid  
prostup mitochondriální  
membrány)



tento „pór“ vzniká interakcí  
7 podjednotek

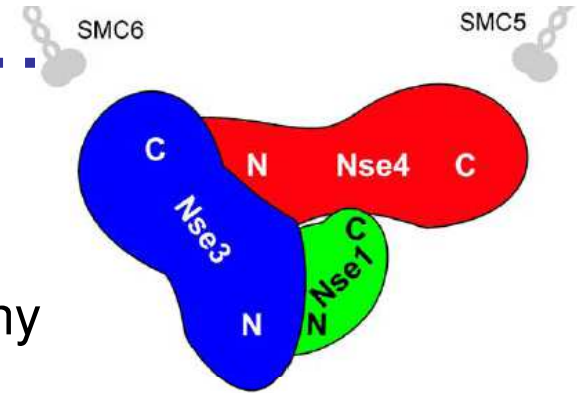
por se používá v NGS  
(nanopore) technologii

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

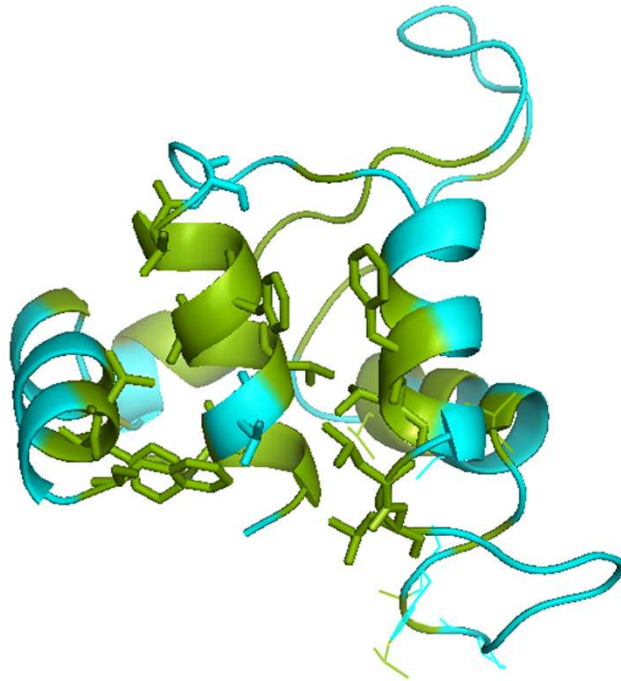


# ... terciární struktura ... kapsa-peptid

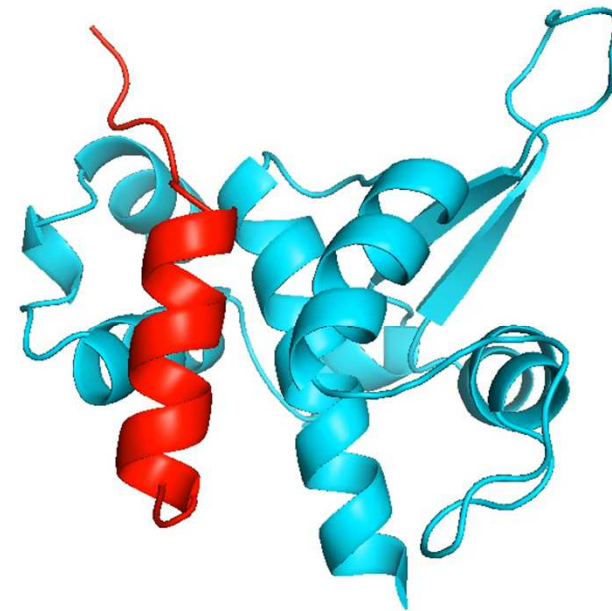
sekundární struktury (šroubovice, beta-listy)  
interagují pod různými úhly a vytváří různé povrchy



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

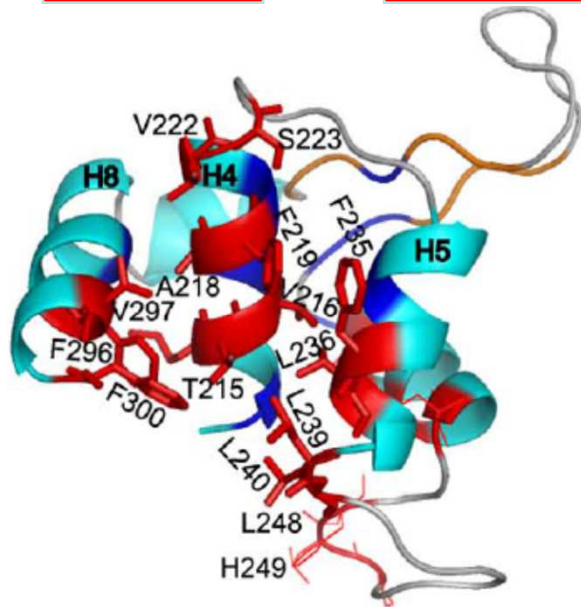
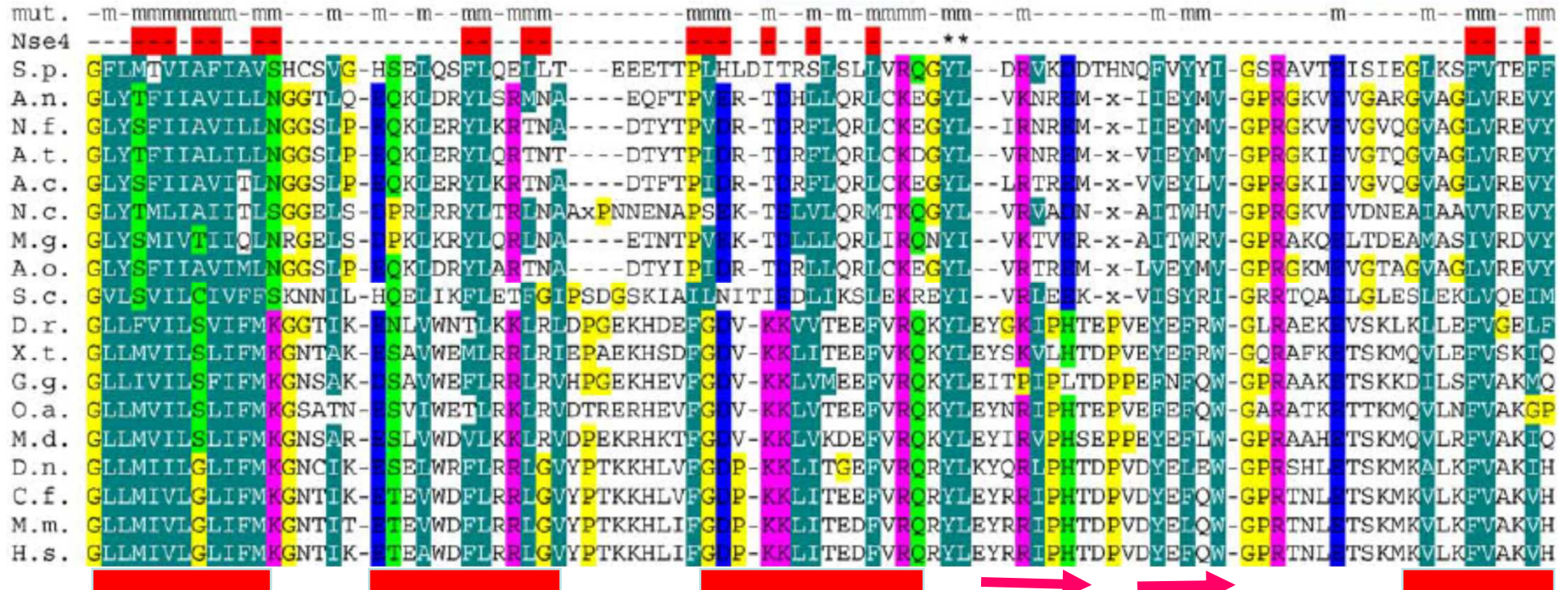


161 166 171 176 181 186 191 196 201 206 211 216 221 226 231 236 241 246 /D/107 111 116 1'  
'ESAIWEMLRRLRIEFGEMHSEFGDVKLVTEEFVKQKYLEYN...KIQGKPKSMTTQYKDAQE HMTVFDPTSFTADLL'  
(ining)

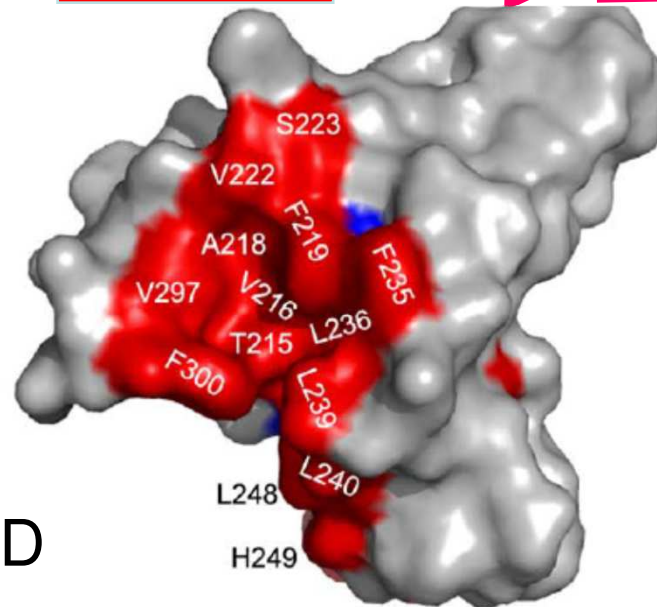


hydrofobní interakce mezi Nse3 a Nse4

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



WHD



Interakce  
mapována  
mutagenezí

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

## MAGEA4 – podobný Nse3

**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

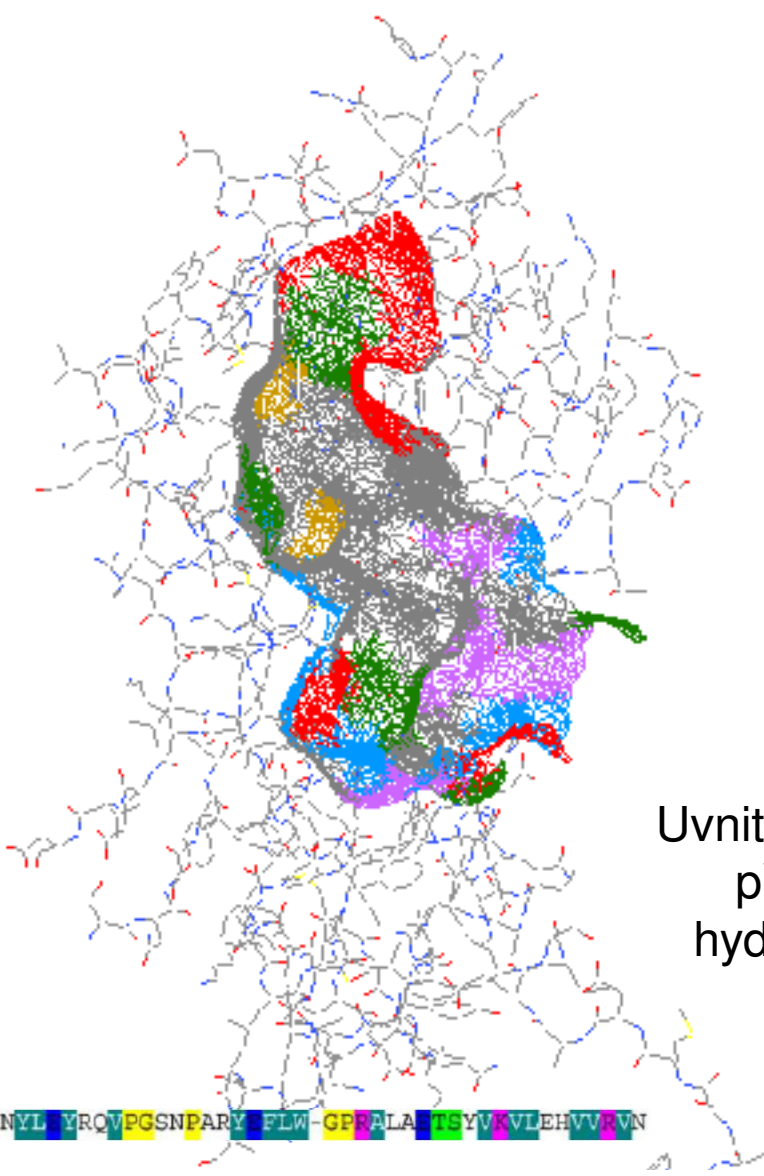
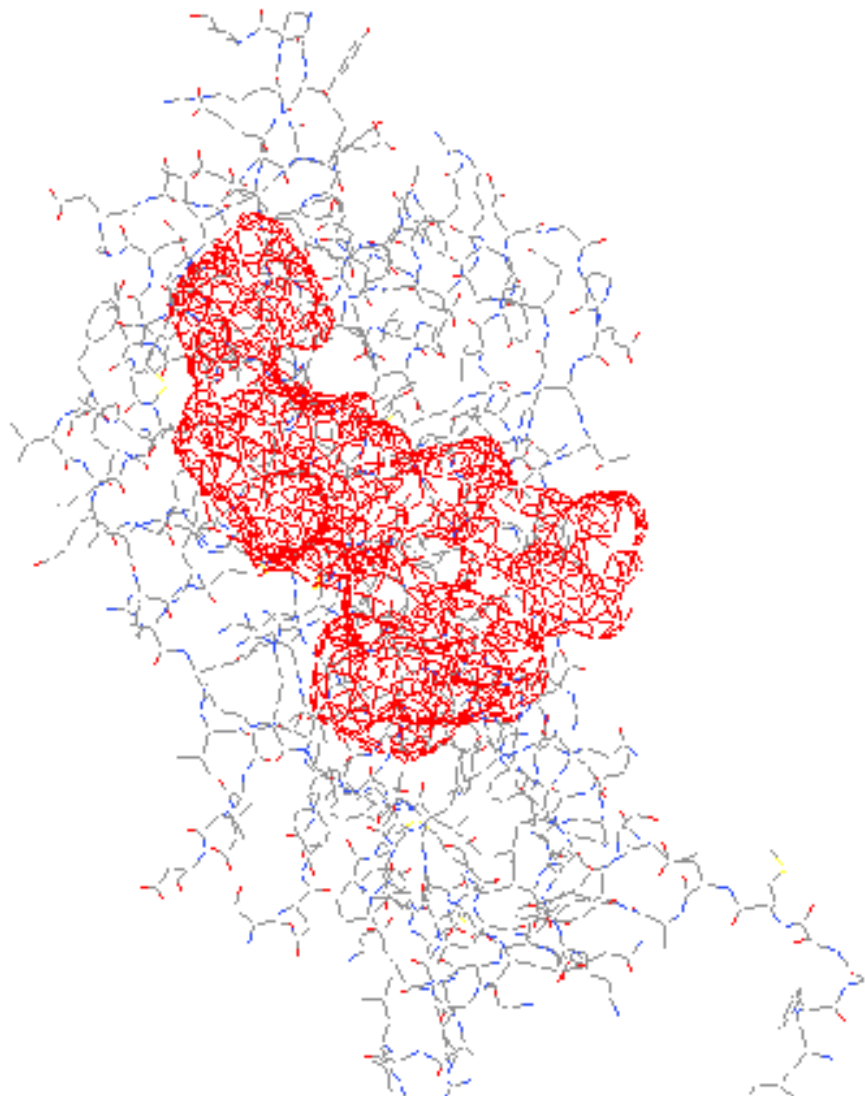
<https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html> /

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

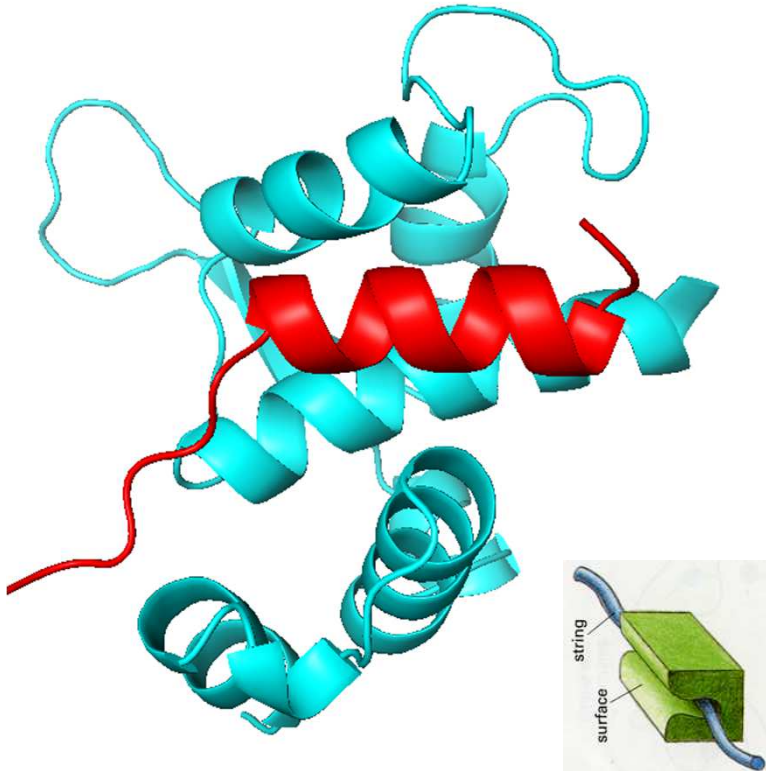
Binding surface



Uvnitř kapsy  
převládá  
hydrofobní  
povrch

A4 GLLIIVLGTIAMEGDSAS-EEIWEELGVVGVYDGREHTVYCF-RKLLTQDWVQENYLVRQVPGSNPARYFLW-GPRALANTSYWVLEHVVRVN

| 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        |



**HADDOCK**  
Software web portal

10111  
01001010010010  
10001010010100101  
01001010010100101  
01001010010100101  
01001010010100101  
01001010010100101

Home HADDOCK Whispy CPORT DNA SQUEEZE Publications HADDOCK Contact

WELCOME TO THE UTRECHT BIOMOLECULAR INTERACTION WEB PORTAL >>

**Universiteit Utrecht**

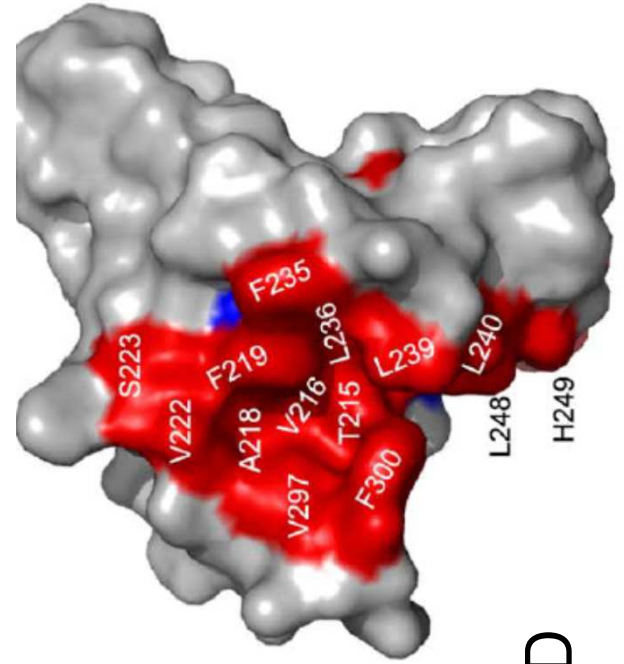
PROFILE >>

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.

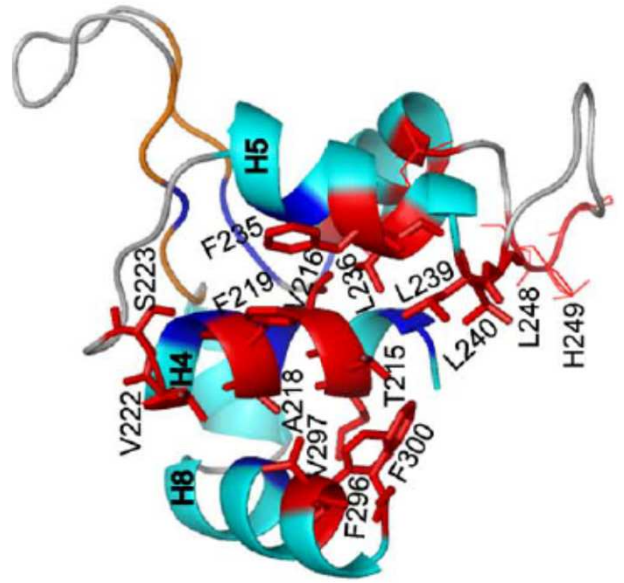
**HADDOCK WEB DOCKING**

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

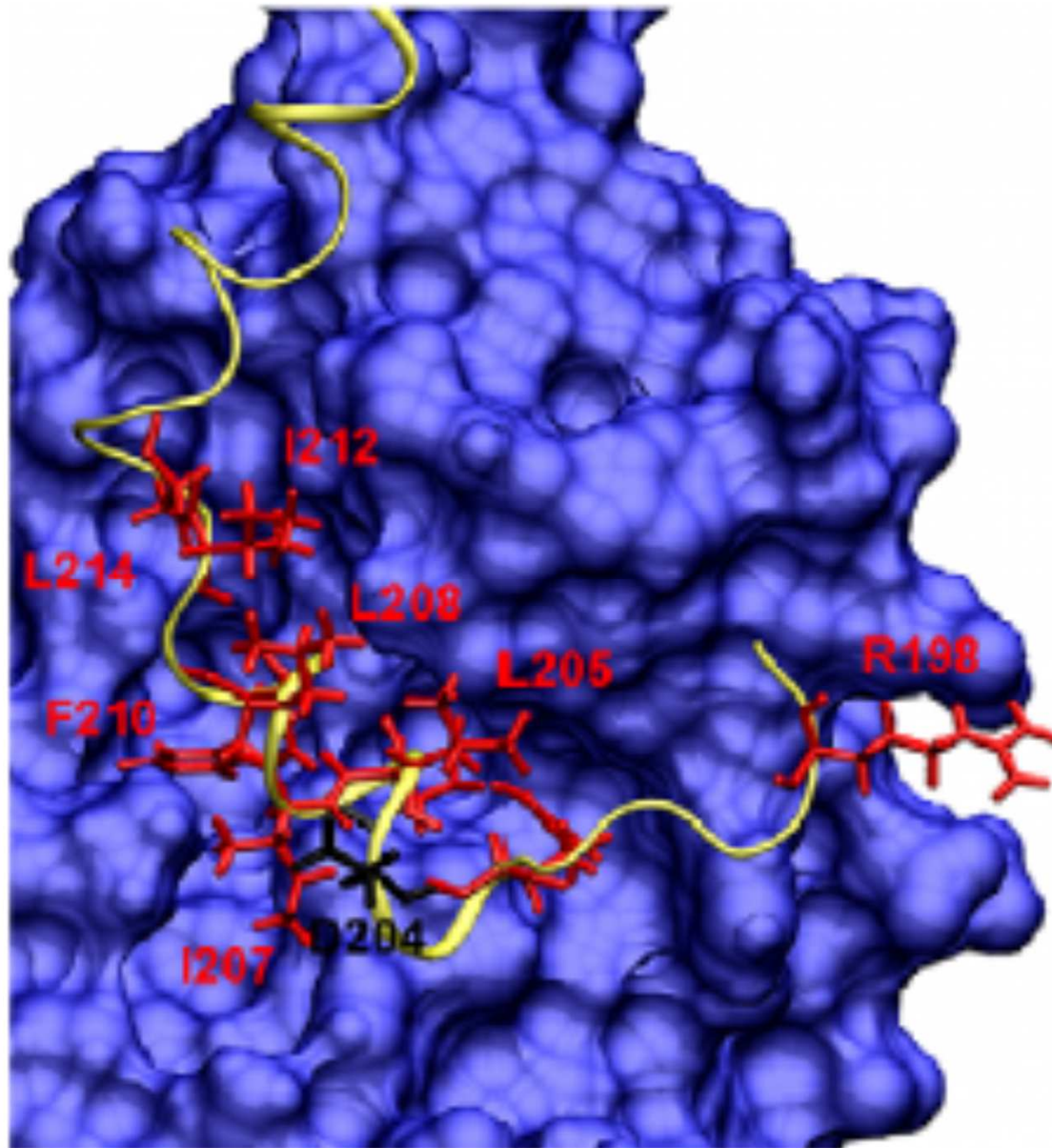
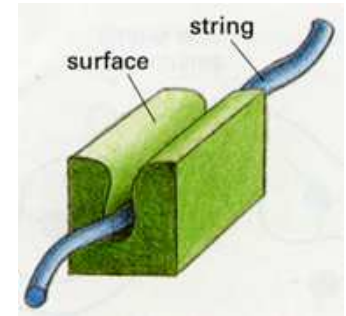


# WHD



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

Interakce  
mapována  
mutagenezí



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

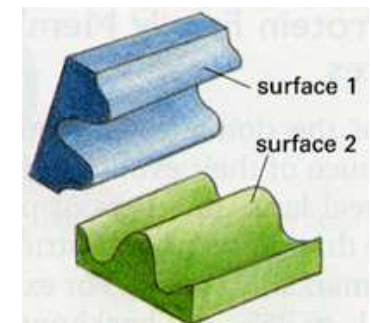
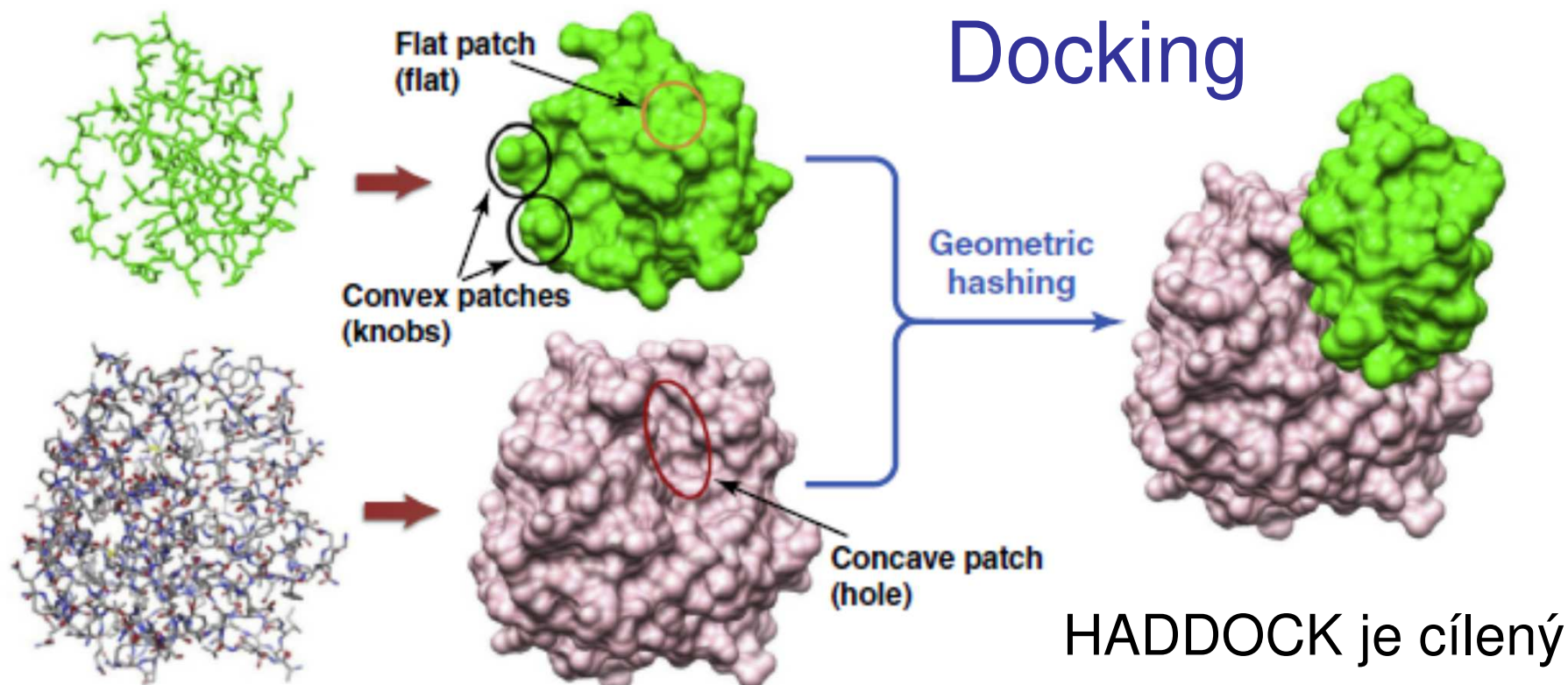


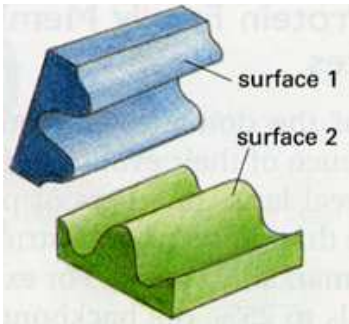
TABLE 1

## 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] |

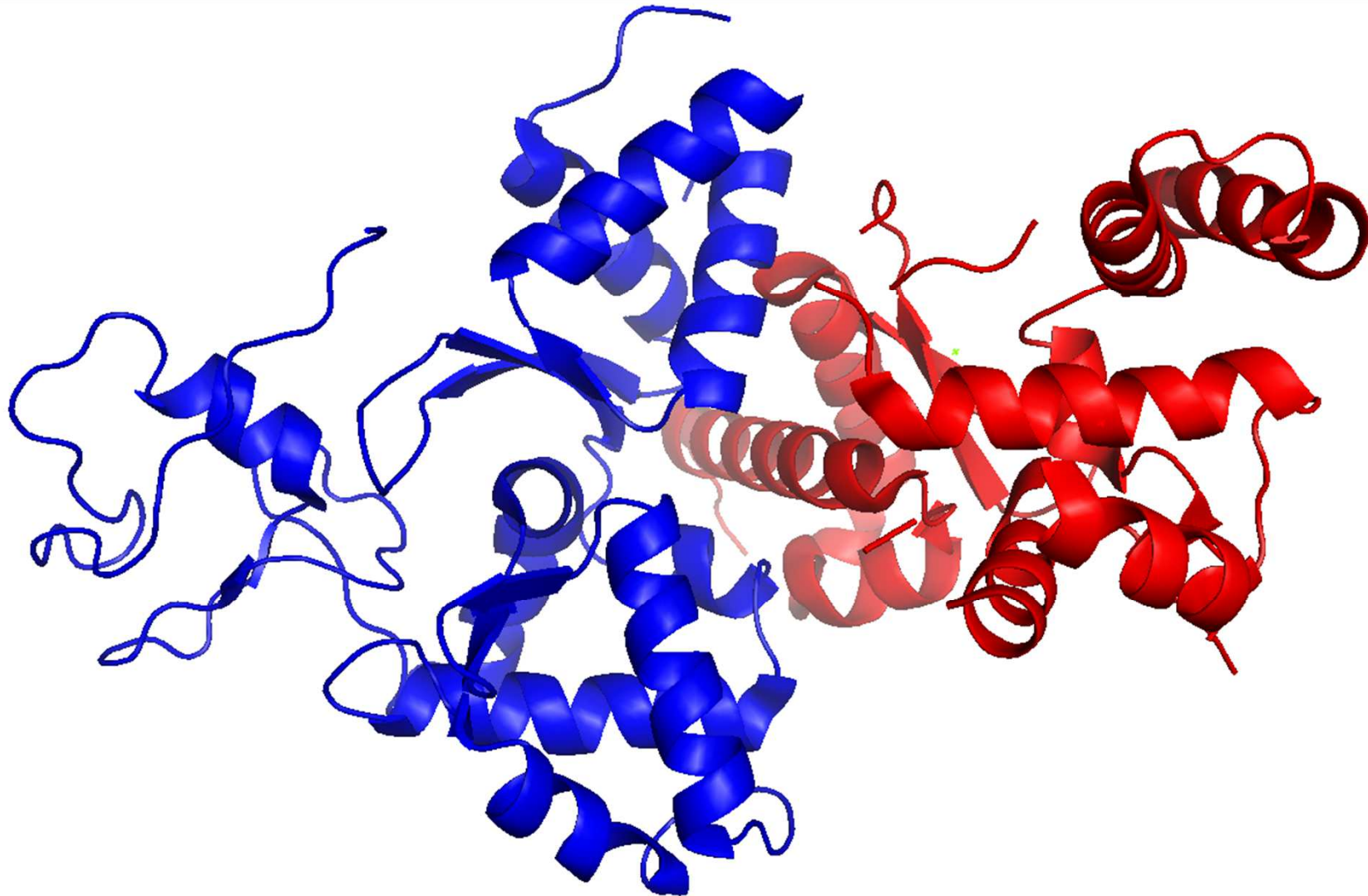


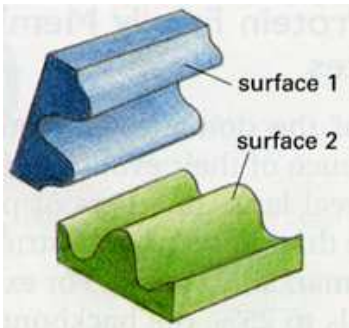




DOMÉNY - šroubovice,  $\beta$ -listy ... interagují pod různými úhly a vytváří různé vazebné motivy s rozsáhlými vazebnými povrchy

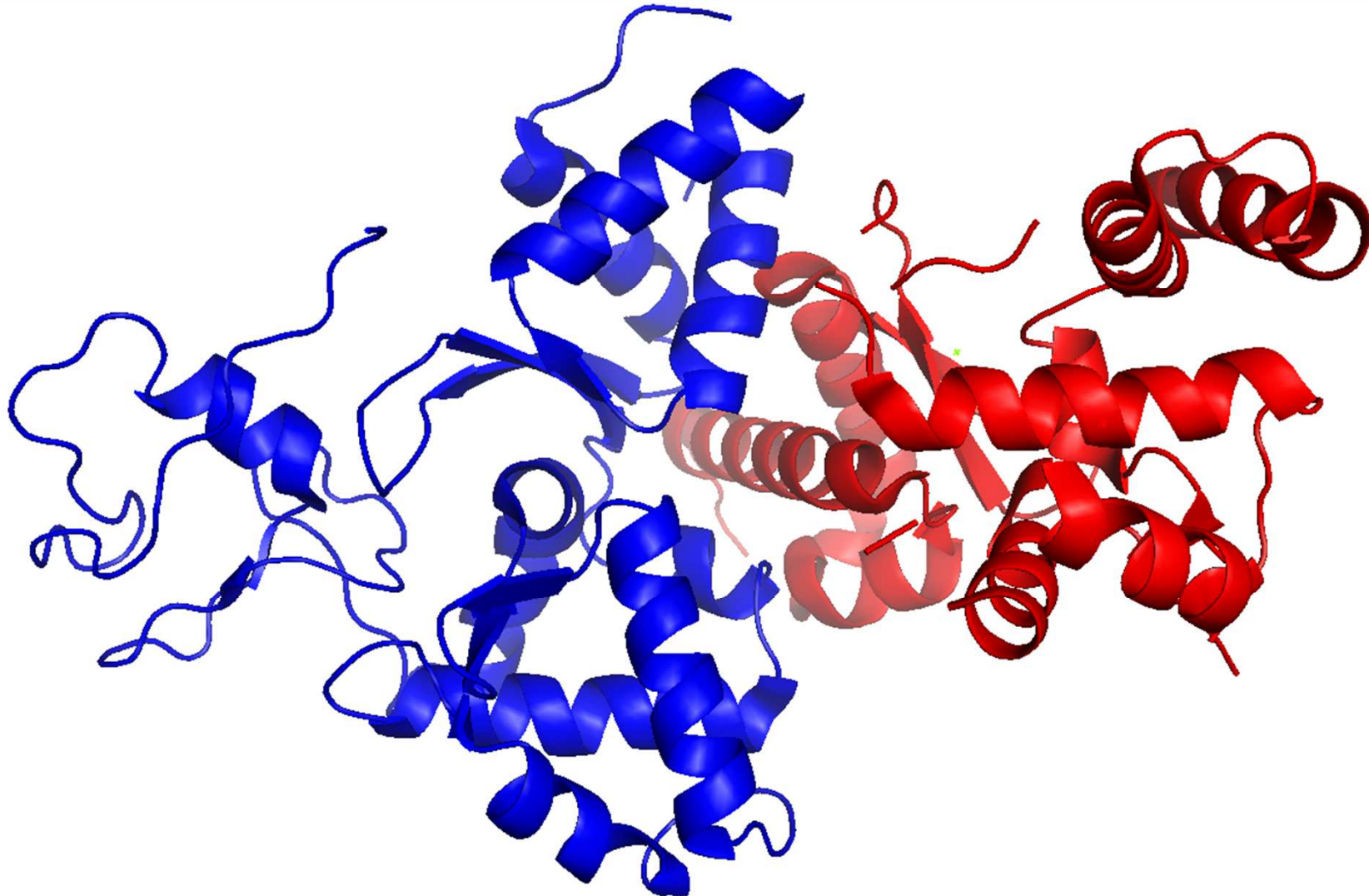
(kokrystal NSE1-NSE3 proteinů)





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

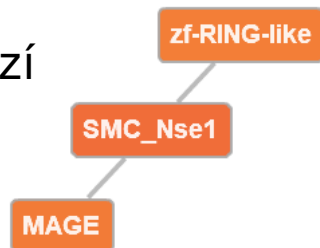
(kokrystal NSE1-NSE3 proteinů)



## 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**

### 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

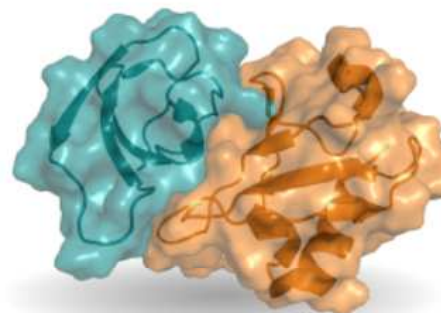
Search motif by name or keyword: ?

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

Search motif



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



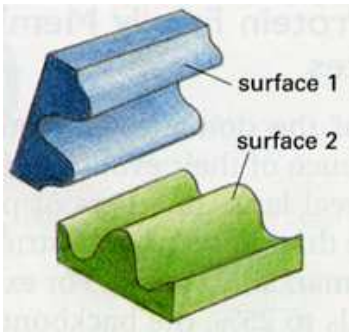
- > **Overview** General information on 3did
- > **Getting Started** Help for new users
- > **Technical Information** Linking to 3did
- > **Download** data files or MySQL tables

### 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
[Clefs](#)
[Tunnels](#)
[Links](#)
PDB id 3nw0

Protein-protein interface: A}B

Protein-protein interface: A}B  
NSE1-NSE3

Chains A and B highlighted  
(click to view)

Chain A
Chain B

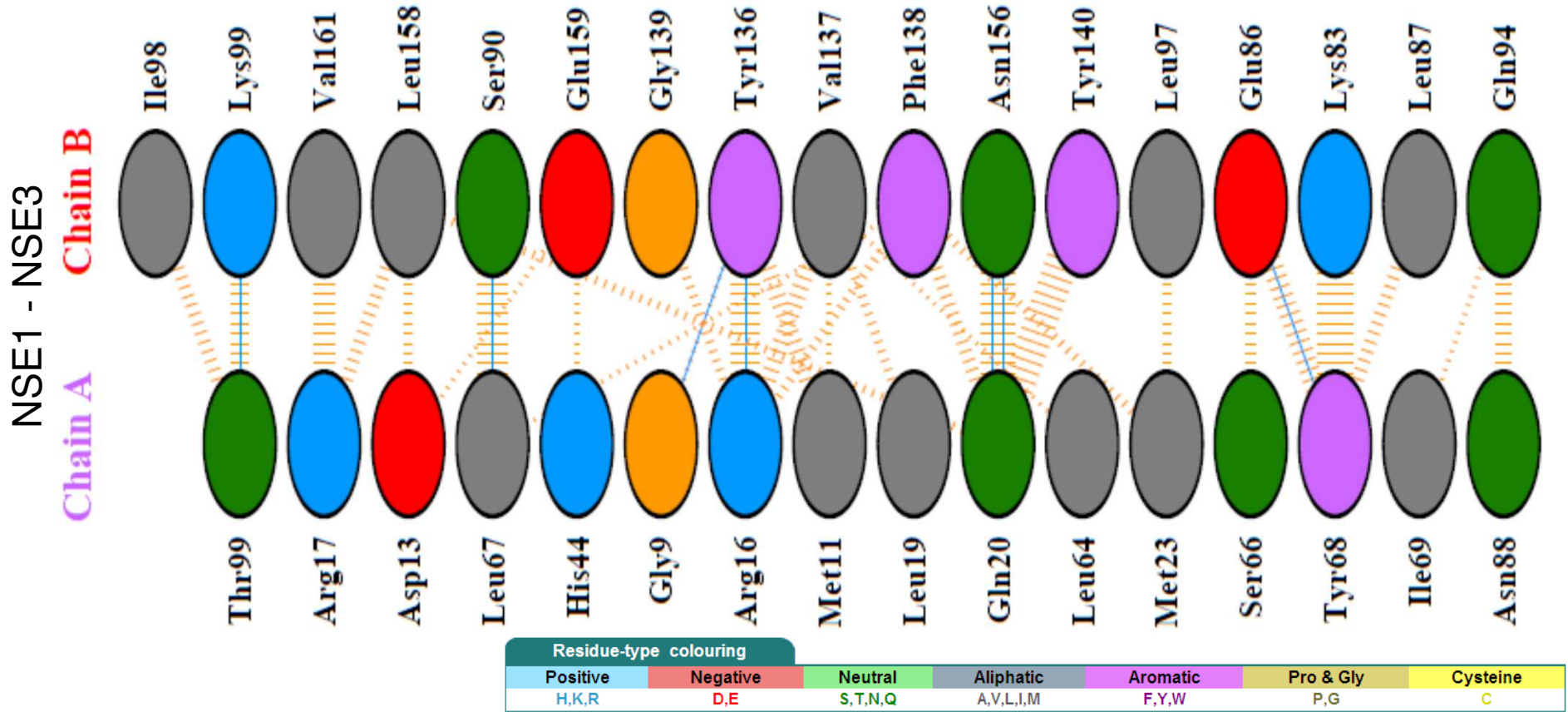
Key:

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

### 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                        |

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



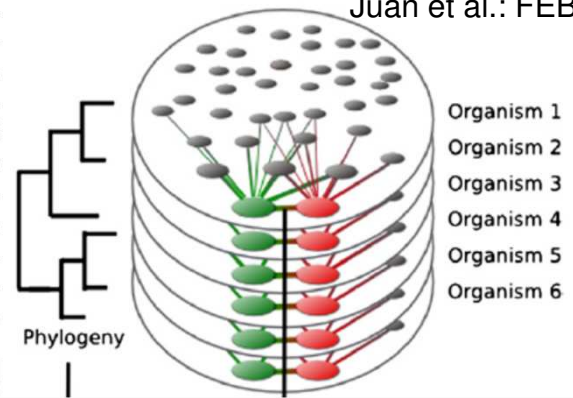
COZOID nástroj: <http://decibel.fi.muni.cz/cozoid/>

```

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.ARSVLTLETAKPILAAS.TFQ.....
U1GD89_ENDPU/11-202   NRAFLQAFM.ARGTLTYETSKPLLASIF.TVH.....
C5GY37_AJEDR/11-207   HRAFLQAFM.ARSTMTYEQAKPVLAAIF.SAR.....
C6H5E2_AJECH/11-203   HRALLOAFM.ARSTMTYEOAKPVLAAIF.TAR.....

```

Juan et al.: FEBS Lett, 2008



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

EMBL-EBI **PFAM – databáze proteinových motivů**

**Nse1 motiv**

keyword search

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

16 architectures | 522 sequences | 2 interactions | 443 species | 1 structure

- 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:EEF14320.1} (744 residues)  
[Show](#) all sequences with this architecture.

|                     |           |                                                                      |
|---------------------|-----------|----------------------------------------------------------------------|
| 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 ENDP/11-202  | NRAFLQAFM | ARSTMTYEQAQPVLAALIF.SAR.....DH.....QDTLP...EDITQ                     |
| C5GY37 AJEDR/11-207 | HRALLQAFM | ARSTMTYEQAQPVLAALIF.TAR.....DN.....QETLP...EDITQ                     |
| C6H5E2 AJECH/11-203 | HRAFLQAFM | SRSTMTLEEAKPVLAALIF.TVS.....EG.....REILP...GDITQ                     |
| F2PT91 TRIEC/10-199 | HRAFLQAFM | ARSTMTLNEAKPILAAIL.SVK.....DG.....REVLP...EDVTQ                      |
| E9DEJ9 COCPS/13-202 | HRAFLQAFM | ARSVLTFEEAQPILAAIL.TAH.....EG.....RPTLP...ADITT                      |
| R7Z157 CONA1/13-202 | HRALVQAFM | ARSSMTGEELLGVVTAIH.GVE.....NPEEP...TETTL                             |
| U4LU38 PYROM/8-205  | HRTFLQALL | IRPFIDIEEGQELLAALIA.SAE.....SG.....TDVPA...NSITV                     |
| S8AAF4 DACHA/9-208  | HRAFLQALL | IRPFIDVQEGRELLAAIK.SAE.....AG.....SDVSI...ESVPP                      |
| G1X2Y0 ARTOA/8-207  | SKFLLQYVL | RRGVCSEKALAKAKLTL.....ERDG.....EQLEDSETE                             |
| C5DCF6 LACTC/23-226 | ARYLLQYLL | CRGICHENMLLVLDKL.....QK.....YTQDPTSQVCS.T                            |
| C5DQF6 ZYGRC/37-240 | RHLLRYIM  | SEGICHENMLLLALYAL.....NLDYSGDCQQEVLA.....                            |
| I2H1A9 TETBL/24-252 | RYLLQYLL  | CGICNESFMLLALMRLH.....LTDGSSSDDFYRLNMVELN                            |
| G8C139 TETPH/22-276 | RIVLQYIL  | SRGIVGENILLALMRL.....QGDT..ETND...VQELY                              |
| G8ZVJ6 TORDC/16-236 | RYLLQYIL  | ARGICHENALVVALMRL.....YMDL..GCFDDAWQIDQWL                            |
| H2ATC7 KAZAF/7-237  | TRYLLQYIL | ARGICHENNLILALMKL.....ETDA.....SKWSTEQWT                             |
| J5S7B7 SACK1/17-253 | AKYLLQYIL | ARGICHENALILALMRL.....ETDA..STLNTEXSIQQWV                            |
| E7NKI9 YEASO/18-258 | AKYLLQYIL | ARGICHENALILALMRL.....ETDA..LRFDAERSMQQWI                            |
| J8PZG2 SACAR/18-258 | RYLLQYIL  | ARGVCHENALVVALMRL.....KVDS..NTIDPQWTISDWL                            |
| G0W7Y8 NAUDC/22-268 | RYLLQYIL  | ARGVCHENALMLALMRL.....KVDS..HDVNAHWTISDWL                            |
| G0V5G7 NAUCC/25-272 | RYLLQYIL  | ARGVCHENSLLAALMHL.....SLDLDPHSVSTR.SLDEWS                            |
| J7S9R2 KAZNA/85-303 | RYLLQYIL  | CRGICHENALLLVLIKL.....DKDSVDEEGSPNRTFEDYL                            |
| A7TJ64 VANPO/8-260  |           |                                                                      |

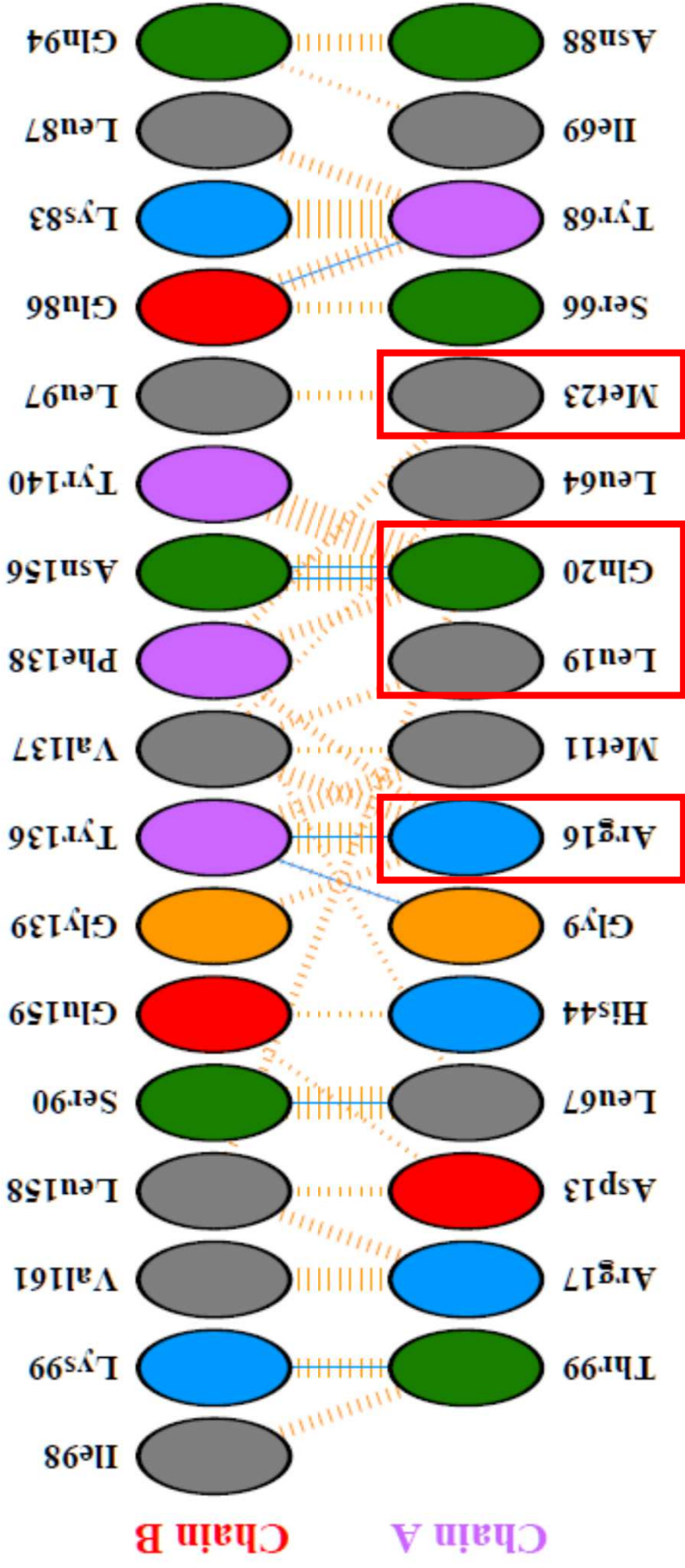
## Konzervované AMK svědčí o důležitosti jejich funkce:

- 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 ARSTMTFAEAKPVLAAIF.SAH.  
 NRAFLOAFM ARSTMTFTQAKPVLAAIF.SIR.  
 NRAFLOAFM ARSTMTFAEARPVLAAIF.SVH.  
 NRAFLOAFM ARSTMTFDEAKPVLAAIF.SAQ.  
 NRAFLOAFM ARSTMTFEEAKPVLAAIF.SAHGAQSTIFFDSSYVEKEIANLYLATER.  
 HRAFLOAFM ARSTMTFEDAQPVLAAII.SAH.  
 NRAFLOAFM ARSCMTFEDAQPILAAIL.TVS.  
 NRAFLOAFM ARSVLTIETAKPILAAIS.TFQ.  
 NRAFVQAFM ARGTLTYETSKPLLASIF.TVH.  
 HRAFLOAFM ARSTMTYEQAKPVLAAIF.SAR.  
 HRALLOAFM ARSTMTYEQAKPVLAAIF.TAR.  
 HRAFLOAFM SRSTMTLEEAKPVLAAIF.TVS.  
 HRAFLOALM ARSTMTLINEAKPILAAIL.SVK.  
 HRAFLOAFM ARSVLTFEEAQPILAAIL.TAH.  
 HRALVQAFM ARSSMTGEEILGVTVAIH.GVE.  
 HRTFLQALL IRPFIDIEEGQELLAAIA.SAE.  
 HRAFLOALL IRPFIDVQEGRELLAAIK.SAE.  
 EKFLLOYVLSRRGVCSEKALAKAIKTLI.

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

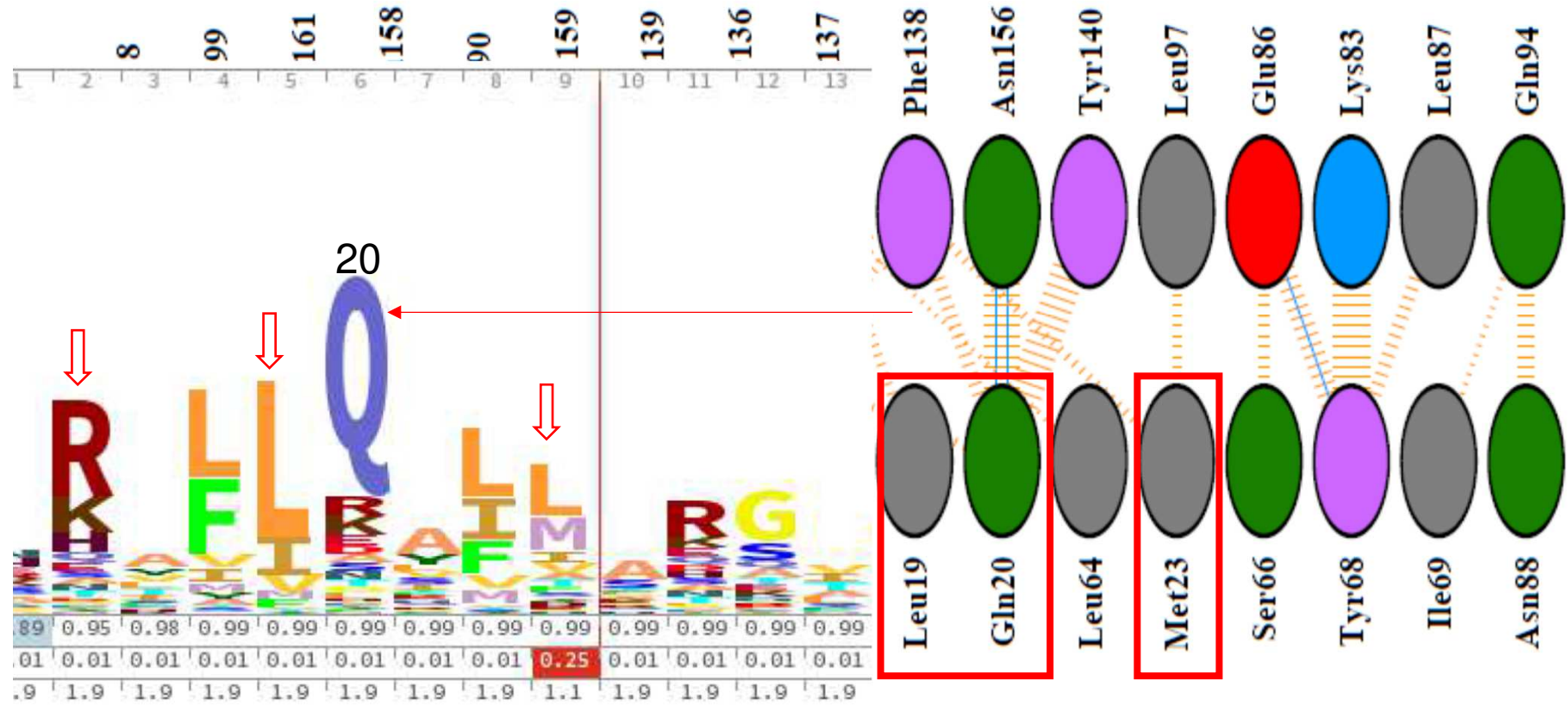
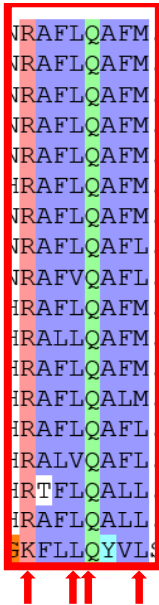




```

A1CCH2_ASPCL/14-216 NRAFLQAFM ARSTMTFAEAKPVLAALF.SAH.....EG.....QPVSA...DDVTE
A2Q7K6_ASPNC/15-218 NRAFLQAFM ARSTMTFTQAKPVLAALF.SIR.....DD.....EQVSP...EDITE
B8NLA5_ASPFN/14-216 NRAFLQAFM ARSTMTFAEARPVLAALF.SVH.....EG.....EPVSA...EDVTE
B6QTR9_TALMQ/14-217 NRAFLQAFM ARSTMTFDEAKPVLAALF.SAQ.....EN.....REVLA...EDITQ
V5FED6_BYSSN/14-236 NRAFLQAFM ARSTMTFEEAKPVLAALF.SAHGAQSTIFFDSSYVEKEIANLYLATER.....RPVLA...EDITQ
S7Z8E9_PENO1/8-209 HRAFLQAFM ARSTMTFEDAQPVLAAII.SAH.....EG.....RTVDP...DEVTE
B6H9Q9_PENRW/8-210 NRAFLQAFM ARSCMTFEDAQPILAAIL.TVS.....EG.....RTVDP...DEVG
H6C926_EXODN/14-207 NRAFLQAFM ARSVLTLETAKPILAALF.TFQ.....DG.....REVQP...QDMTV
U1GD89_ENDPU/11-202 NRAFLQAFM ARSVLTLETAKPILAALF.TVH.....EG.....REILP...NDITE
C5GY37_AJEDR/11-207 HRAFLQAFM ARSTMTYEQAKPVLAALF.SAR.....DH.....QDTLP...EDITQ
C6H5E2_AJECH/11-203 HRALLQAFM ARSTMTYEQAKPVLAALF.TAR.....DN.....QETLP...EDITQ
F2PT91_TRIEC/10-199 HRAFLQAFM SRSTMTLEEAKPVLAALF.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 HRALVQAFM ARSSMTGEELLGVVTAIH.GVE.....NPEEP...TETTL
S8AAF4_DACHA/9-208 HRTFLQALL IRPFIDIEEGQELLAALF.SAE.....SG.....TDVPA...NSITV
G1X2Y0_ARTOA/8-207 HRAFLQALL IRPFIDVQEGRELLAALF.SAE.....AG.....SDVSI...ESVPP
C5DCF6_LACTC/23-226 SKFLLQYVLRRGVCSEKALAKALKTL.....ERDG.....EQLEDSETE

```





Structures Overview  
Structure Selections  
3D View Controls

Visualization



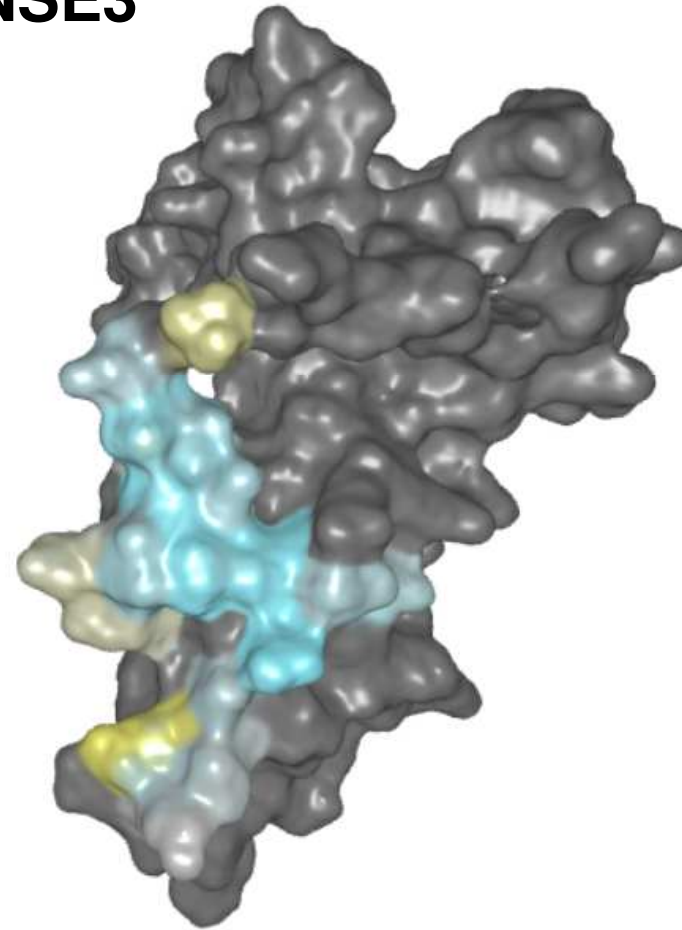
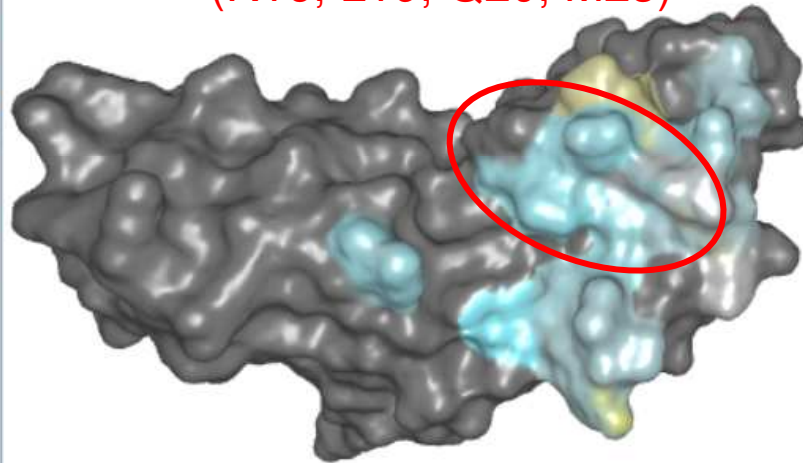
Residue Matrix

Primary Structure: 3

Sort by: Conservati

## NSE1 - NSE3

Hot spot  
(R16, L19, Q20, M23)



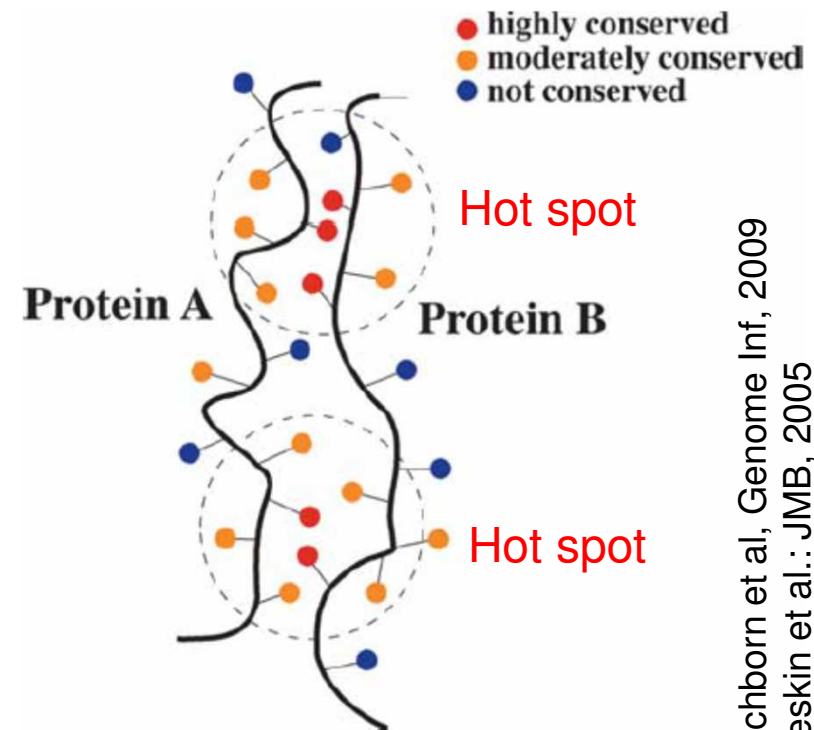
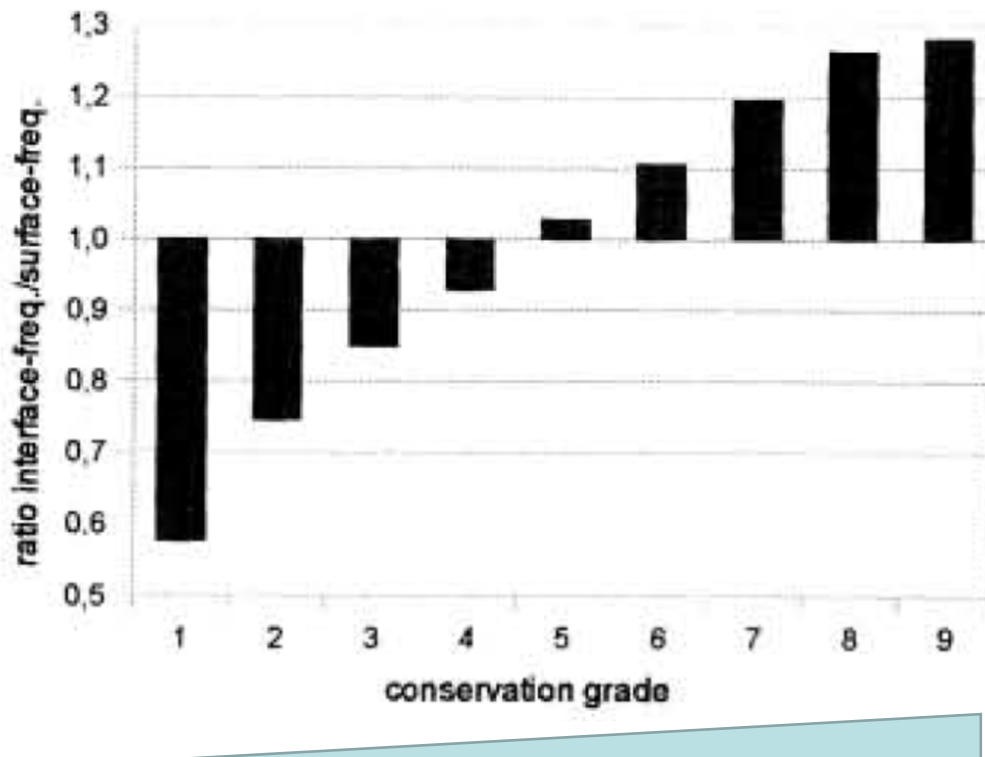
|     |
|-----|
| N88 |
| Q20 |
| M23 |
| T99 |
| L64 |
| L19 |
| D13 |
| I69 |
| L67 |
| Y68 |
| H44 |
| S66 |
| M11 |
| R16 |
| G9  |
| R17 |

Structure Sequence x

Compact View  Selection in All Structures

# Silné/důležité interakce (komplexy) jsou evolučně konzervované

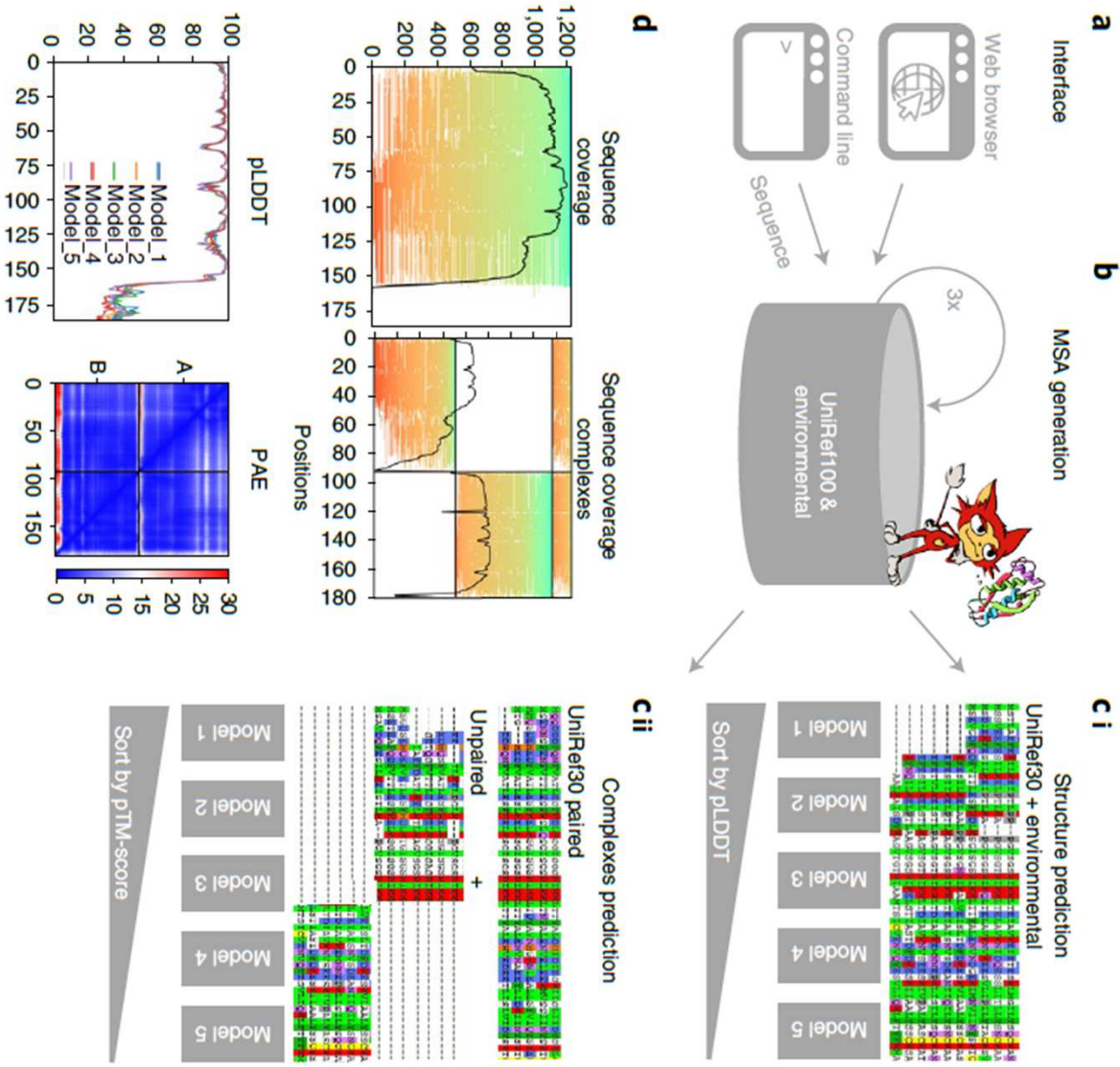
- jako jsou proteiny (jejich funkce) evolučně konzervované, tak i jejich interakce jsou evolučně konzervované (zajišťují funkci)
- graf** – povrchové AMK jsou málo konzervované (grade1), zatímco interakční povrchy jsou hodně konzervované (grade9)



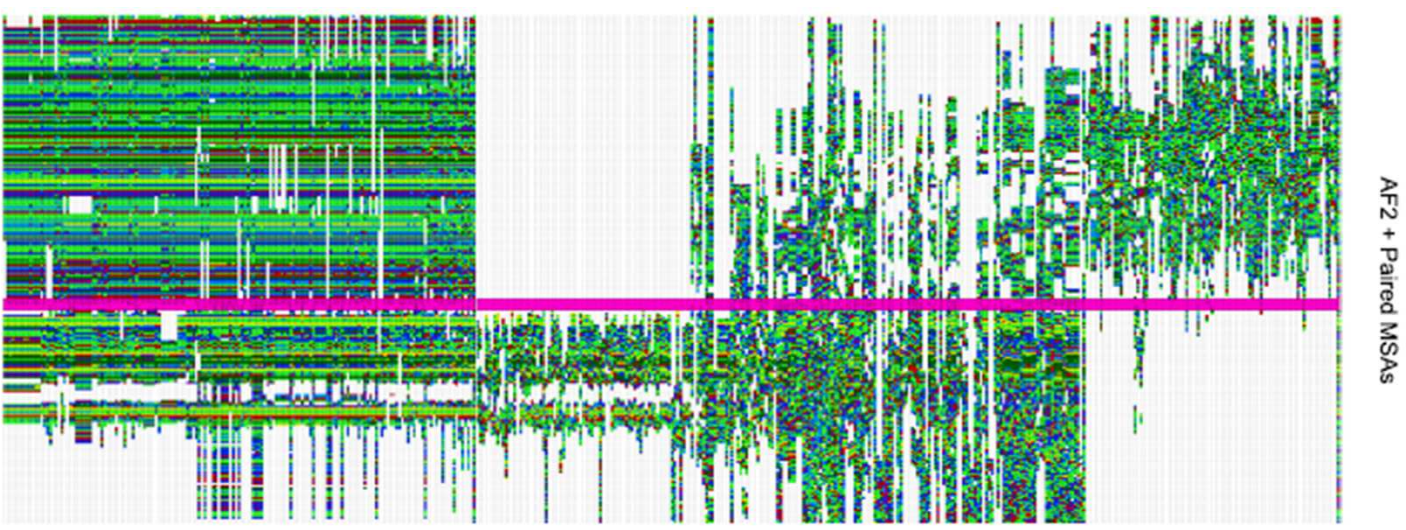
Eichborn et al, Genome Inf, 2009  
Keskin et al.: JMB, 2005

**Koevoluce - potenciál pro AlphaFold?**

# AlphaFold/Colab



Čerpá info z databází ...



Bryant et al., Nat Com, 2022; Mirdita et al., Nat Met, 2022

# 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 fúzí)

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

Více Dr. Potěšil



## 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

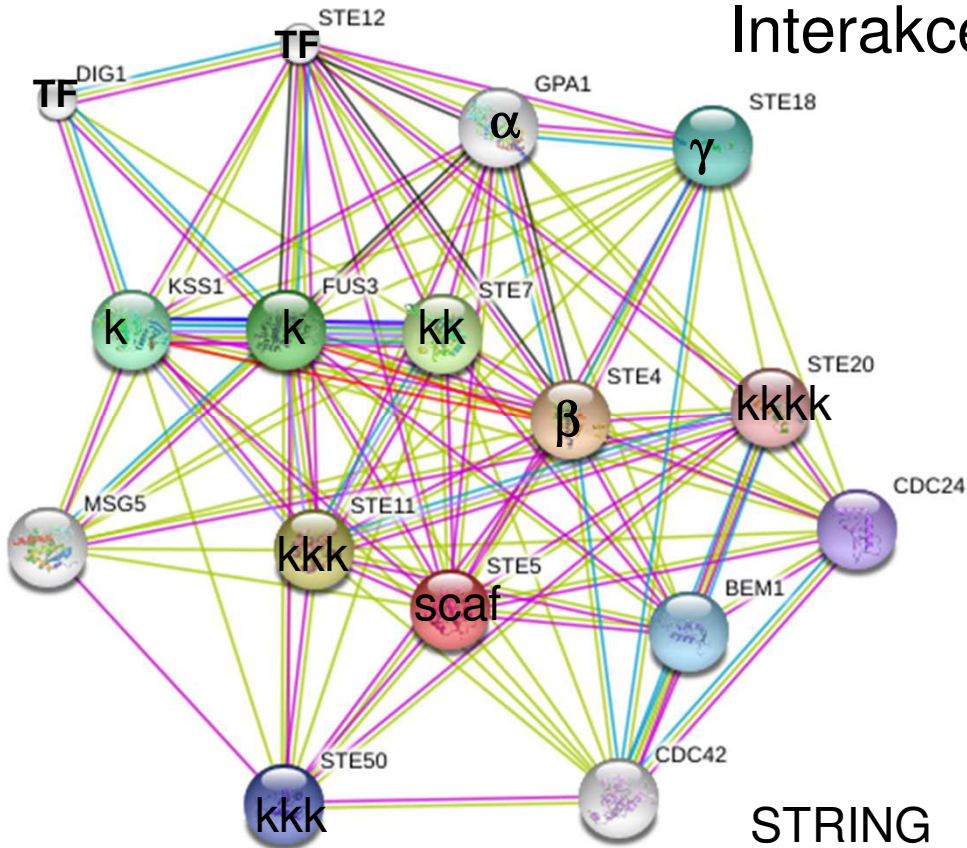
[\[details\]](#)

| 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* ...

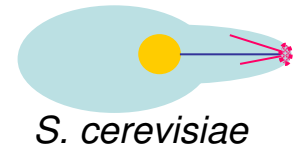
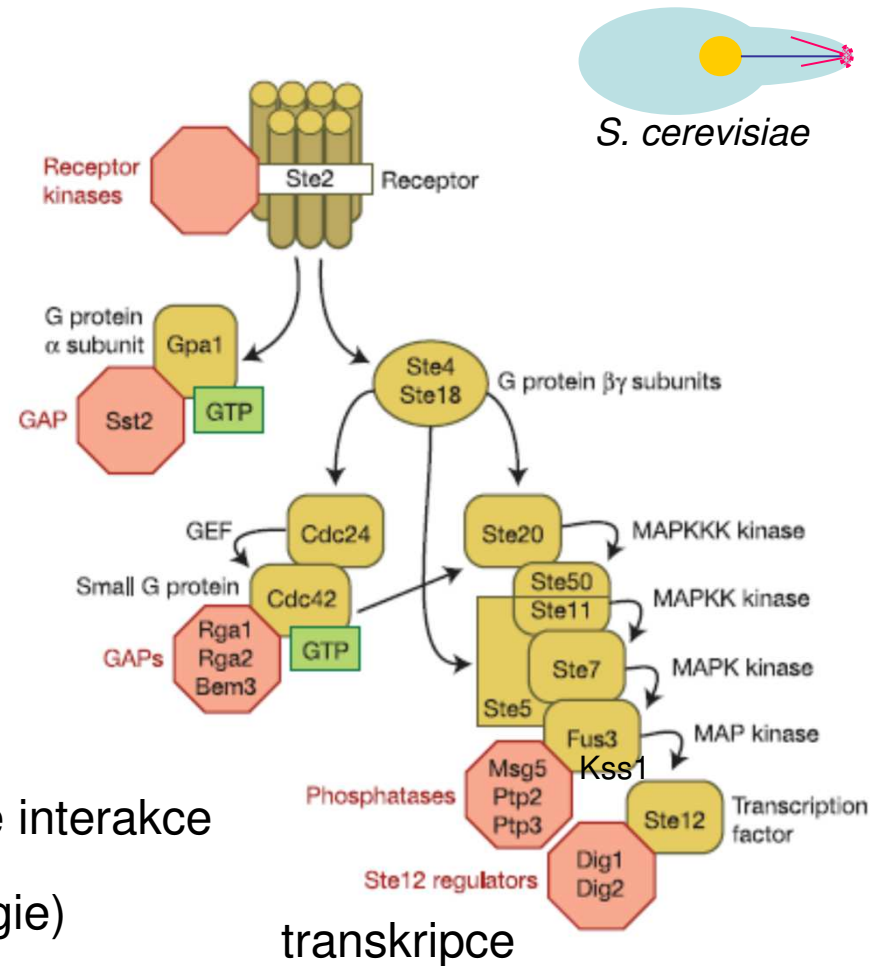
# proteinové sítě – víc než PPI vs dynamika interakcí ...

## Interakce x signální dráha



STRING

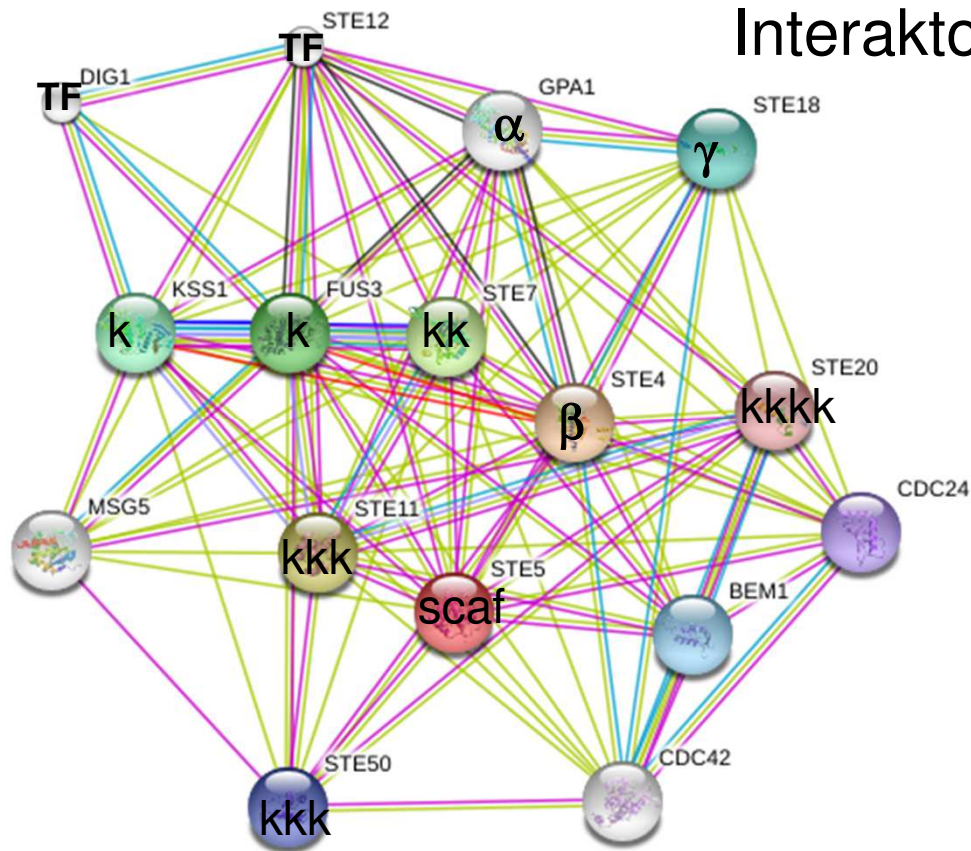
- Experiments Y2H, colP ... genetické interakce
- Databases Funkční vztahy (ontologie)
- Textmining
- Gene Fusion Svědčí o potřebě PPI
- Coexpression Potřeba výskytu ve stejném okamžiku a společná translace



*S. cerevisiae*

transkripce

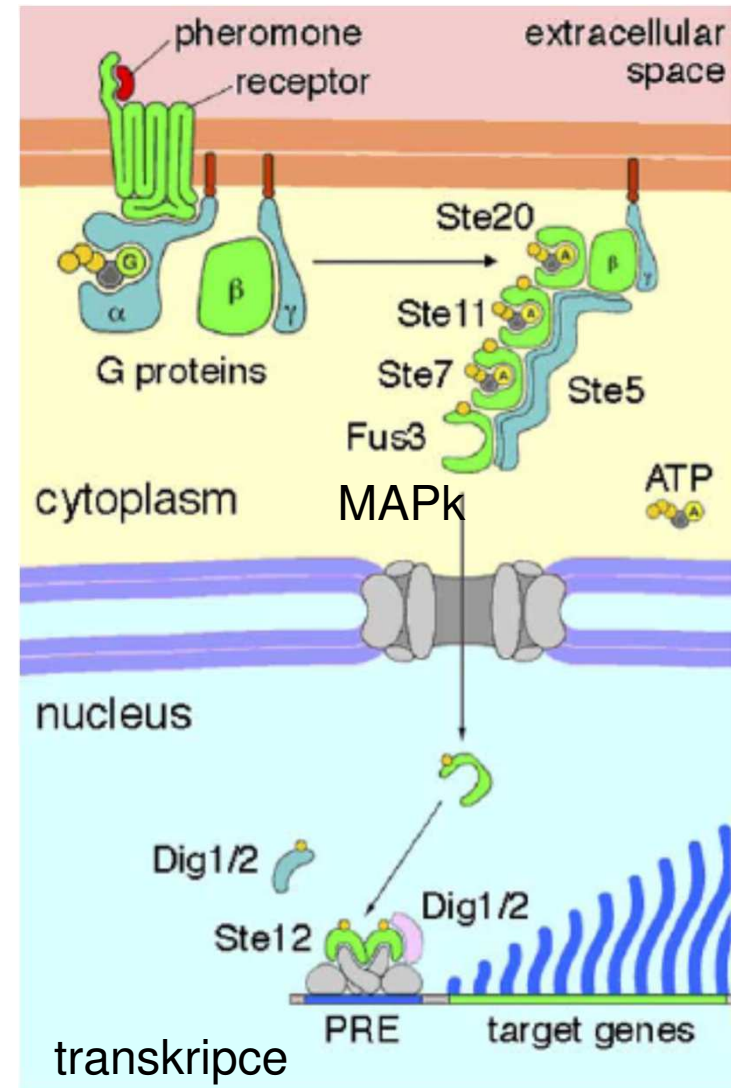
# proteinové sítě – schéma vs realita



Interaktom x komplexom

- Experiments
- Databases
- Textmining
- Gene Fusion
- Coexpression

Síť neznamená komplex,  
ale vztahy  
souhrn proteinových:  
interakcí = **interaktom**  
komplexů = **komplexom**

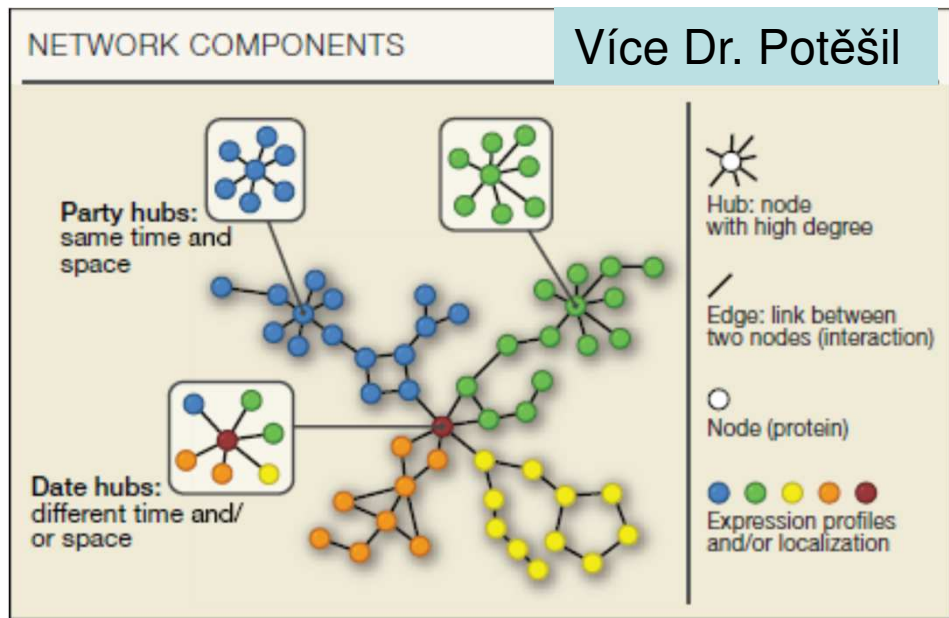


Wang et al., Nature, 2004

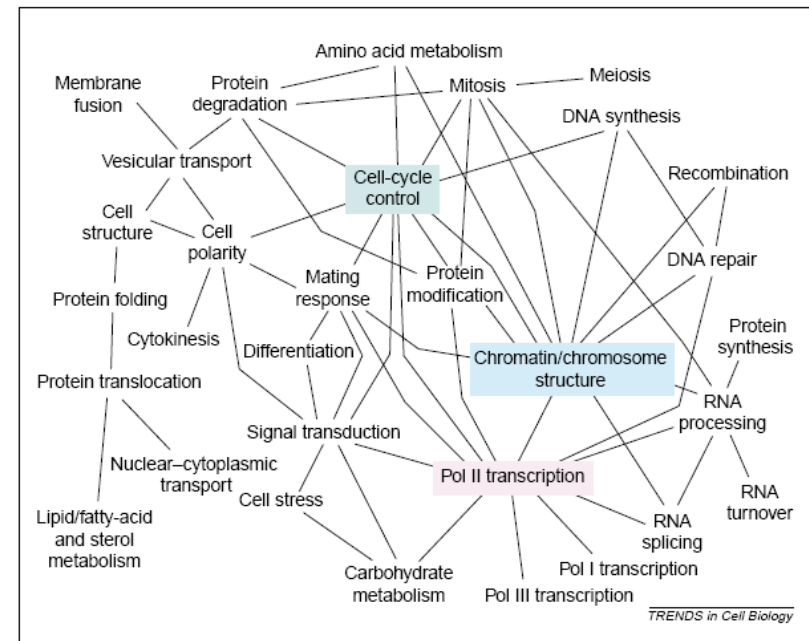


# Protein-proteinové interakce

- stabilní (velké plochy, většinou součástí komplexů)
  - přechodné/slabe (součást dynamických procesů – předávání signálů, modifikace)
  - posttranslační modifikace mohou změnit vazebné vlastnosti povrchu (fosforylace, metylace, SUMO)
- (modularita díky interakcím domén – různé kombinace domén)



Seebacher & Gavin, Cell (SNAP SHOT), 2011



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

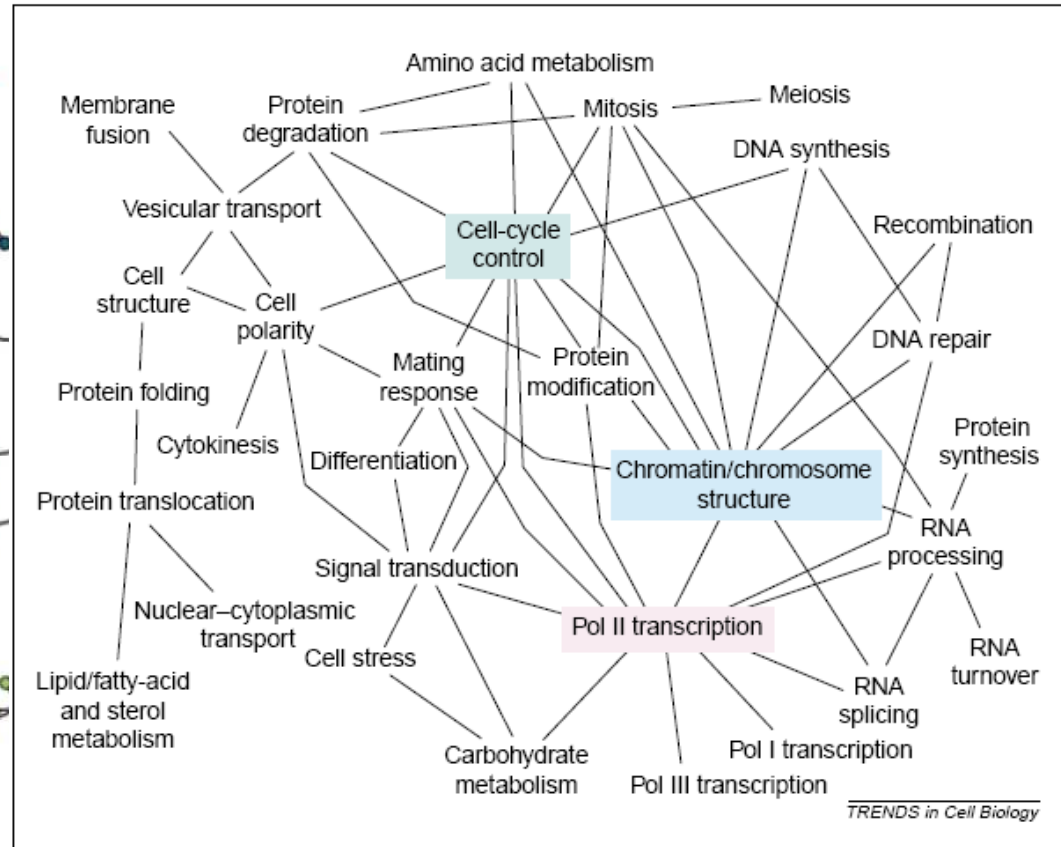
# Interaktom x komplexom

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

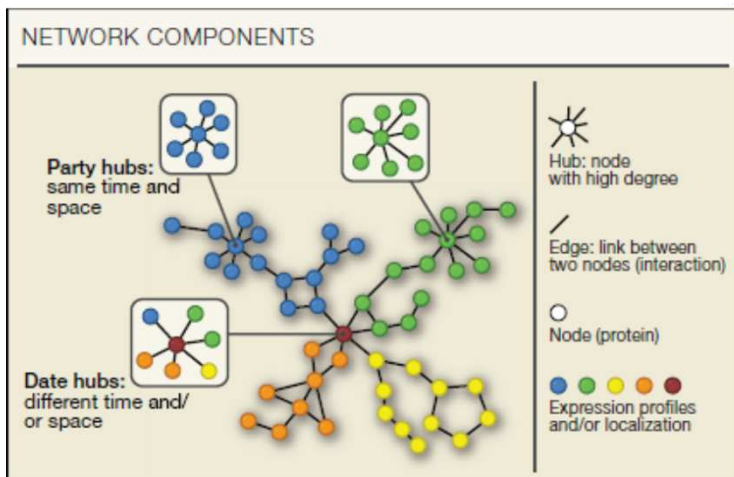
Naznačují funkční **vztahy**  
(např. buněčný cyklus –  
struktura chromatinu ... jsou  
zprostředkovány PPIs)

více interakcí jedné **domény** =  
různé povrchy (3D) nebo v  
různém čase stejný povrch

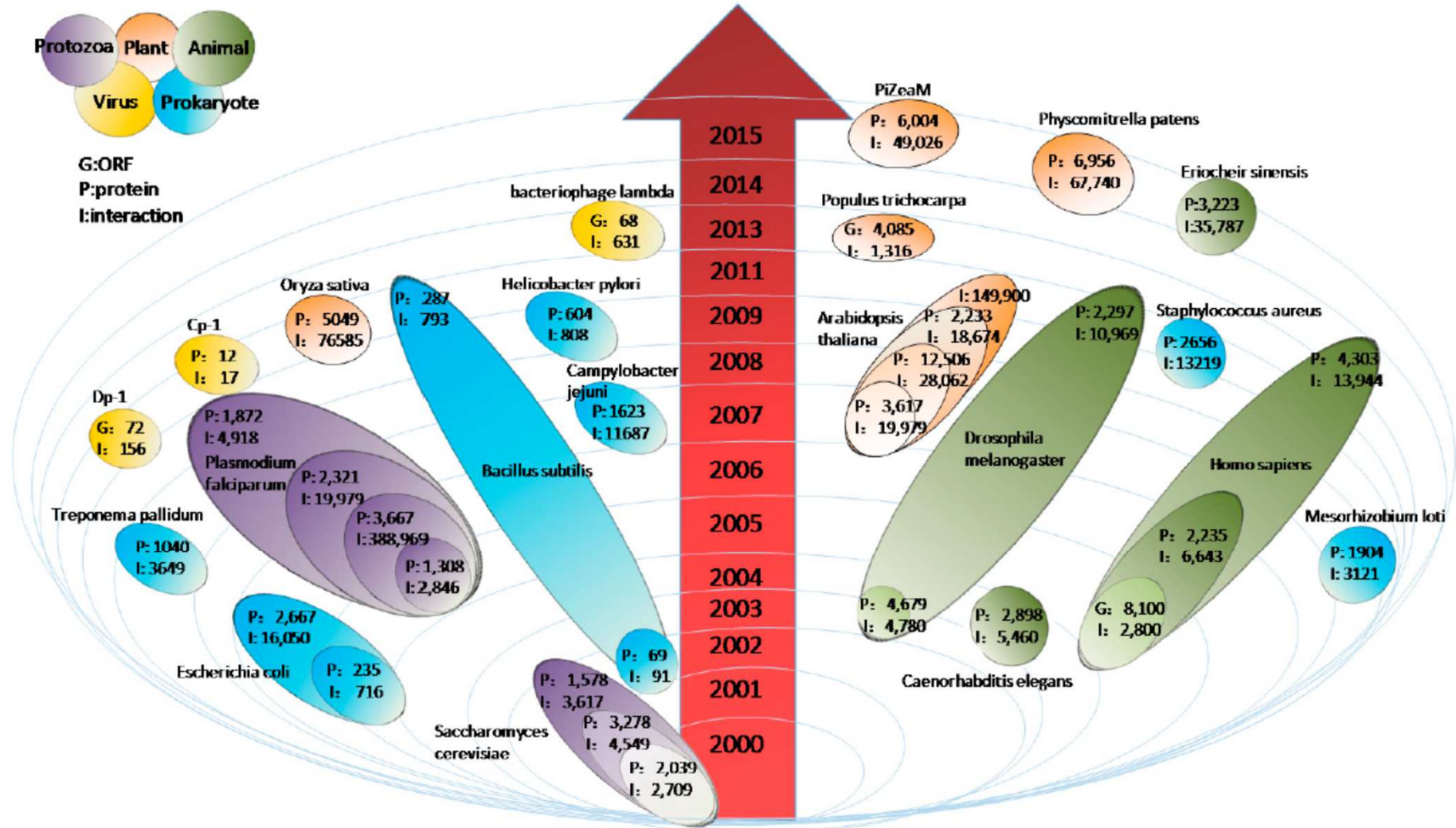
**Modularita** – interagují domény  
(jeden protein více domén –  
zapojení do více procesů)



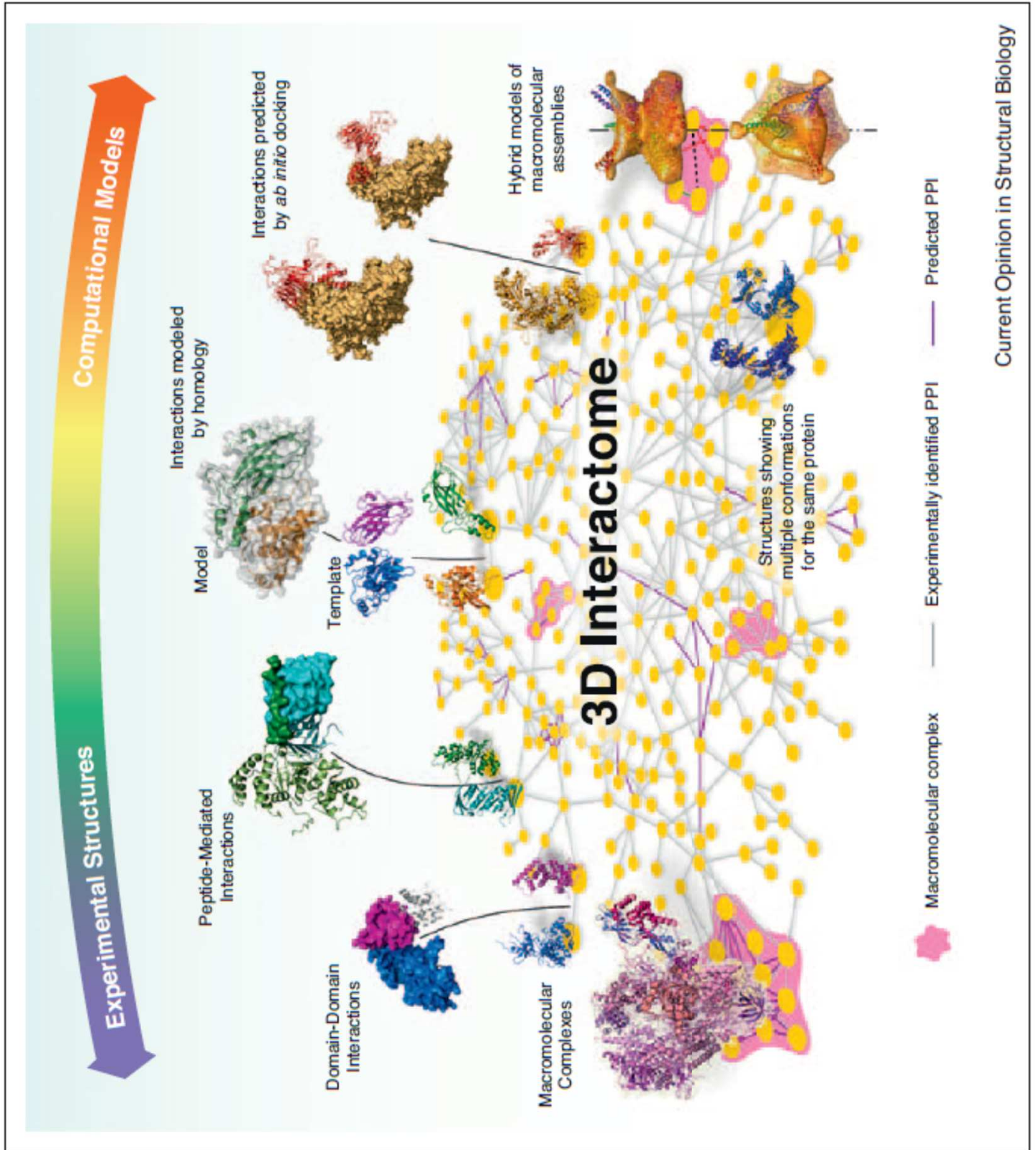
**Fig. 2.** Functional group interaction map based on Fig. 1 (modified from Ref. 10). Shown are interactions between functional groups of yeast proteins. Each line indicates that there are 15 or more interactions between proteins of the connected groups. Connections with fewer than 15 interactions are 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.



# High-throughput screens – interaktomy organismů



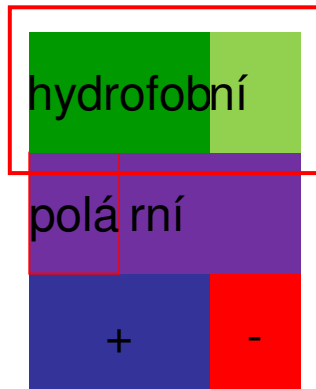
Kvasinkový dvoj-hybridní systém  
 TAP-tag s MS analýzou



# 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

