Pokročilé metody v genomice a proteomice: analýza proteinových komplexů

doc. Jan Paleček
jpalecek@sci.muni.cz

laboratoř Strukturních proteinů eukaryotních chromosomů
(http://www.ncbr.muni.cz/SPEC/)

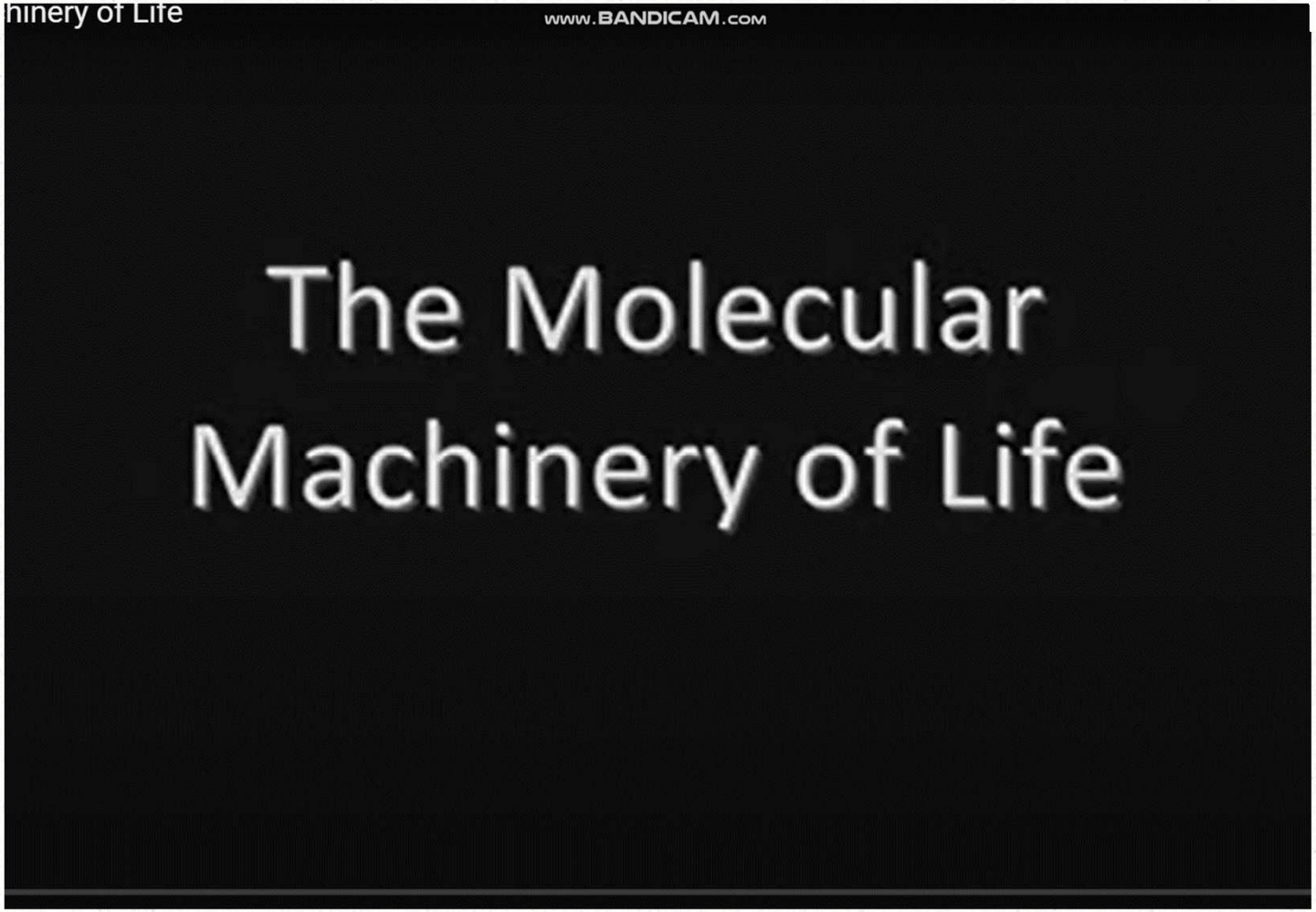
Financováno Evropskou unií
NextGenerationEU

Národní plán obnovy

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

NPO_MUNI_MSMT-16606/2022

Malý průlet světem buňky



většina … proteinové komplexy …

… chromosomy

https://www.youtube.com/watch?v=FJ4N0iSeR8U

… a o čem uslyšíte

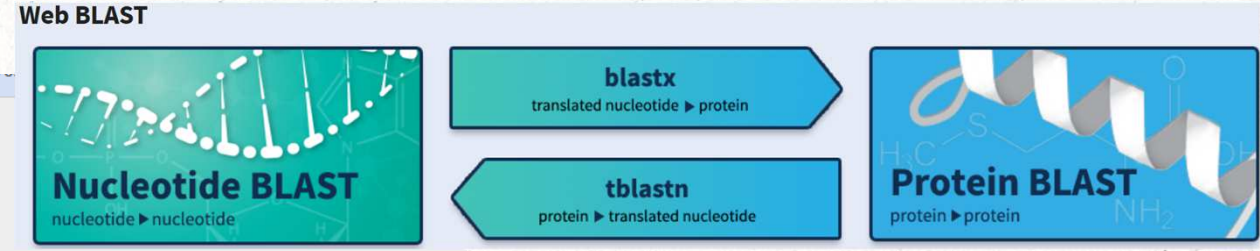
- Co všechno se dozvíte o genu/proteinu z databází
  - 3D (terciární) struktura proteinů
  - Proteinové interakce
  - Proteinové komplexy (kvarterní struktura)
- Jak získat nové informace o proteinových komplexech experimentálně
  - Analýza protein-proteinových interakcí
  - Analýza komplexů

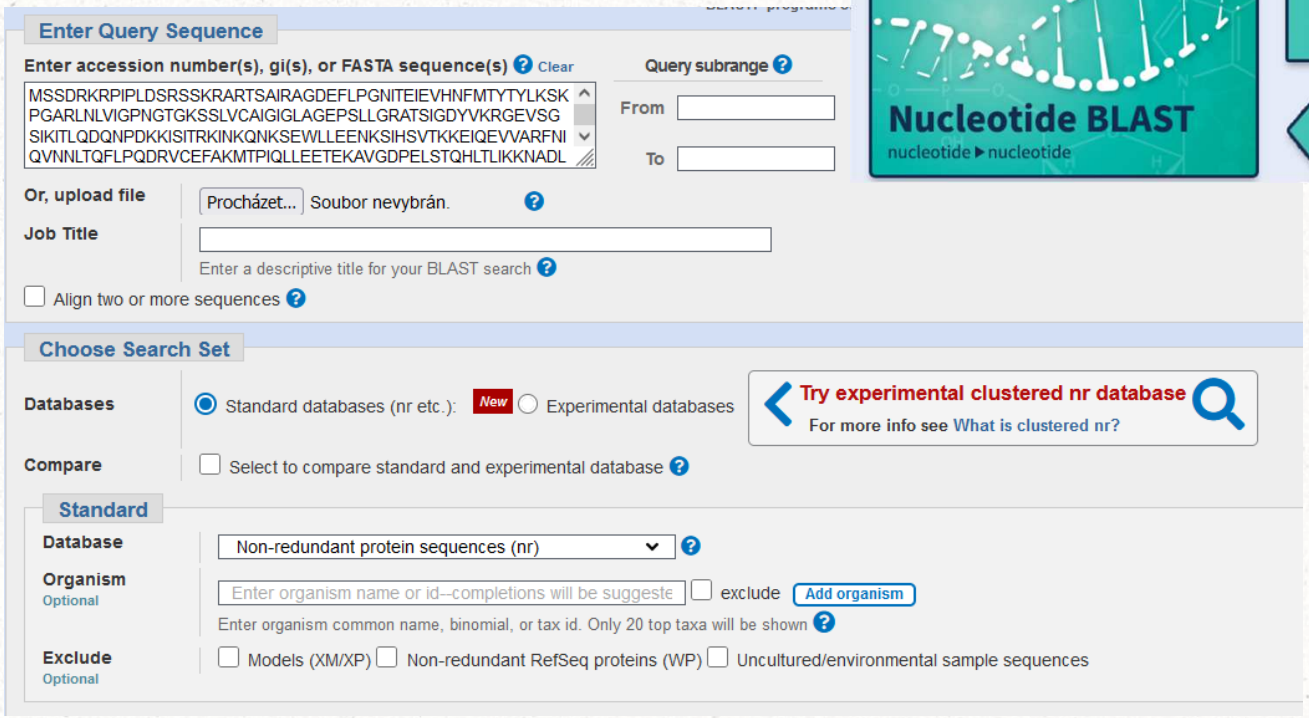
NGS přináší nevídané množství dat

- rychlé sekvenování genomů poskytuje informace o proteinech v různých organismech
- proteinové databáze UNIPROT (SWISSPROT…): https://www.uniprot.org/
- potvrzuje představu evoluce proteinů/organismů – ukazuje na velkou konzervovanost většiny proteinů
- tyto podobnosti umožňují modelování proteinů i z málo charakterizovaných organismů (a jejich anotaci) – spolehlivé modelování pro homologie >30%

Základní alignment

- pokud pracujete s novou sekvencí (např. není v UNIPROT): BLAST – https://blast.ncbi.nlm.nih.gov/
- hledání příbuzných sekvencí vám napoví nakolik je protein evolučně konzervovaný, jaké má domény … jakou mají funkci jemu podobné proteiny (např. UNIPROT databáze)





UniProt a jeho možnosti

Komplexní, vysoce kvalitní a volně přístupný zdroj sekvenčních a funkčních informací o proteinech

- Popis a funkce
- Zjištění přítomnosti domén daného proteinu
- Zjištění interakčních partnerů
- Struktura proteinu – PDB, Alphafold … komplexy

Hledání proteinu

Souvislosti
Funkce
Interaktom
Mutace

Vizualizace
Alphafold
Lokalizace proteinu

Stahování dat

UniProt – vyhledávání

https://www.uniprot.org/

Histon H3

https://www.uniprot.org/uniprotkb?query=H3+cerevisiae

UniProt BLAST Align Peptide search ID mapping SPARQL UniProtKB H3 cerevisiae Advanced | List Search

Status

Reviewed (Swiss-Prot) (244)

Unreviewed (TrEMBL) (308)

Popular organisms

S. cerevisiae (176)

Human (16)

Rat (7)

UniProtKB 552 results

BLAST Align Map IDs Download Add View: Cards ○ Table ◉ Customize columns Share

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
P61830	H3_YEAST	**Histone H3**	**HHT1**, YBR010W, YBR0201, HHT2, SIN2, YNL031C, N2749	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast)	136 AA
Q757N1	H3_ASHGO	**Histone H3**	**HHT1**, ADL202C, HHT2, AER013W	Ashbya gossypii (strain ATCC 10895 / CBS 109.51 / FGSC 9923 / NRRL Y-1056) (Yeast) (Eremothecium gossypii)	136 AA

... i lidský

Informace o proteinu – UNPROT

Histon H3

- Function
- Names & Taxonomy
- Subcellular Location
- Phenotypes & Variants
- PTM/Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequence
- Similar Proteins

Function

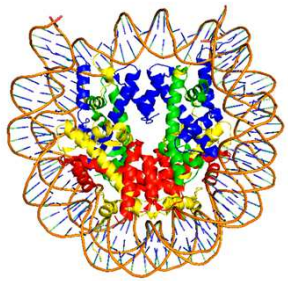
Core component of nucleosome. Nucleosomes wrap and compact DNA into chromatin, limiting DNA accessibility to the cellular machineries which require DNA as a template. Histones thereby play a central role in transcription regulation, DNA repair, DNA replication and chromosomal stability. DNA accessibility is regulated via a complex set of post-translational modifications of histones, also called histone code, and nucleosome remodeling. Component of the UAF (upstream activation factor) complex which interacts with the upstream element of the RNA polymerase I promoter and forms a stable preinitiation complex.

Subcellular Location

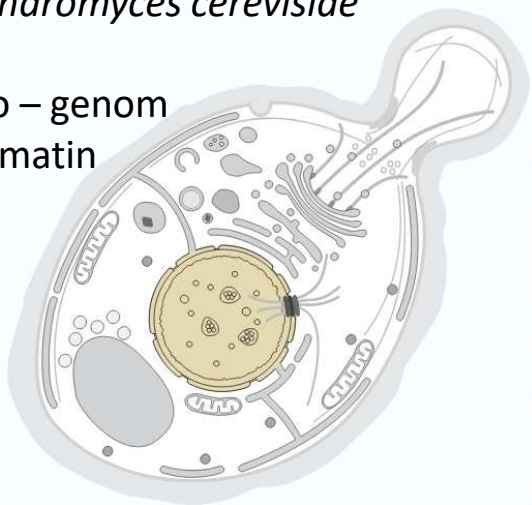
UniProt Annotation | GO Annotation

Saccharomyces cerevisiae

Nucleus [Curated]
Chromosome [Curated]



jádro – genom
chromatin



Genová ontologie

ASPECT	TERM		
Cellular Component	CENP-A containing nucleosome	Source:SGD	1 Publication
Cellular Component	nucleosome	Source:ComplexPortal	1 Publication
Cellular Component	nucleus	Source:SGD	1 Publication
Cellular Component	replication fork protection complex	Source:SGD	1 Publication
Cellular Component	RNA polymerase I upstream activating factor complex	Source:ComplexPortal	1 Publication
Molecular Function	DNA binding	Source:SGD	1 Publication
Molecular Function	protein heterodimerization activity	Source:InterPro	
Molecular Function	structural constituent of chromatin	Source:InterPro	
Biological Process	chromatin organization	Source:SGD	1 Publication
Biological Process	global genome nucleotide-excision repair	Source:SGD	1 Publication
Biological Process	nucleolar large rRNA transcription by RNA polymerase I	Source:ComplexPortal	1 Publication

upozornění: informace nebývají kompletní (i chybné) – dohledat v článcích

Informace o proteinu – UNPROT

Histon H3

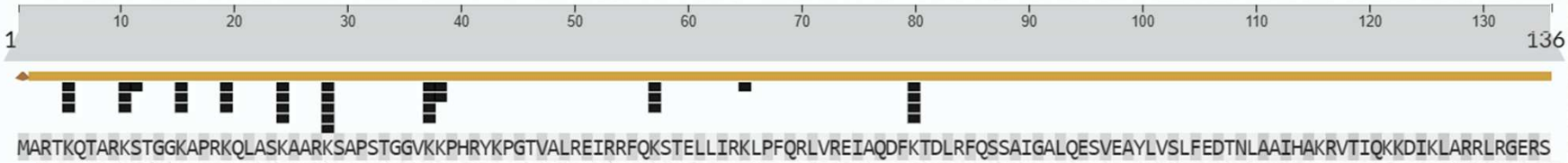
- Function
- Names & Taxonomy
- Subcellular Location
- Phenotypes & Variants
- PTM/Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequence
- Similar Proteins

Function

Core component of nucleosome. Nucleosomes wrap and compact DNA into chromatin, limiting DNA accessibility to the cellular machineries which require DNA as a template. Histones thereby play a central role in transcription regulation, DNA repair, DNA replication and chromosomal stability. DNA accessibility is regulated via a complex set of post-translational modifications of histones, also called **histone code**, and nucleosome remodeling. Component of the UAF (upstream activation factor) complex which interacts with the upstream element of the RNA polymerase I promoter and forms a stable preinitiation complex.

PTM

#Acetylation
#Methylation
#Phosphoprotein



Acetylation of histone H3 leads to transcriptional activation. H3K14ac formation by GCN5, a component of the SAGA complex, is promoted by H3S10ph. Further acetylated by GCN5 to form **H3K9ac**, H3K18ac, H3K23ac, H3K27ac and H3K36ac. H3K14ac can also be formed by ESA1, a component of the NuA4 histone acetyltransferase (HAT) complex. **H3K56ac** formation occurs predominantly in newly synthesized H3 molecules during G1, S and G2/M of the cell cycle and may be involved in DNA repair. 13 Publications

Informace o proteinu – UNPROT

Histon H3

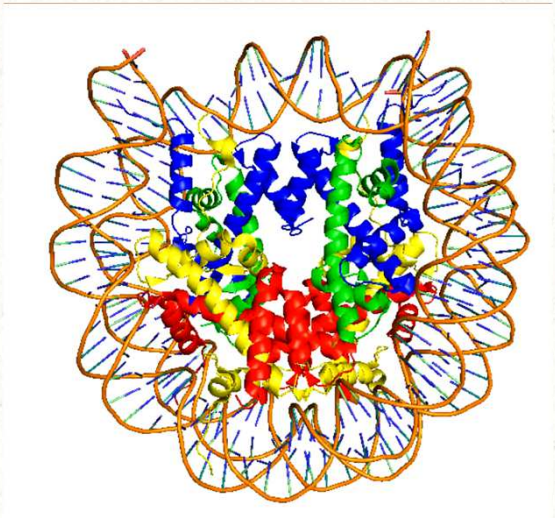
- Function
- Names & Taxonomy
- Subcellular Location
- Phenotypes & Variants
- PTM/Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequence
- Similar Proteins

Function

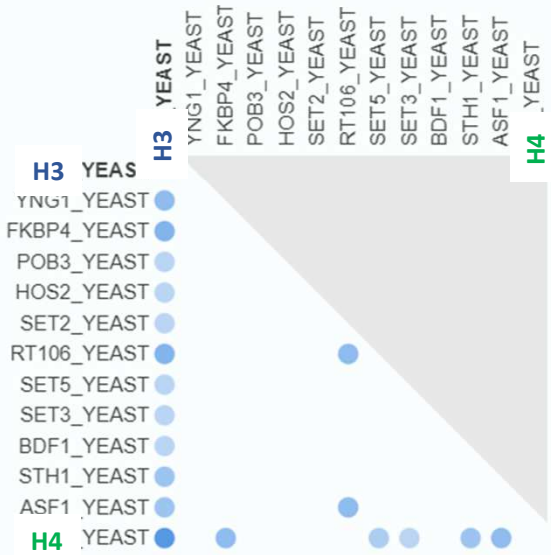
Core component of **nucleosome**. Nucleosomes wrap and compact DNA into chromatin, limiting DNA accessibility to the cellular machineries which require DNA as a template. Histones thereby play a central role in transcription regulation, DNA repair, DNA replication and chromosomal stability. DNA accessibility is regulated via a complex set of post-translational modifications of histones, also called histone code, and nucleosome remodeling. Component of the UAF (upstream activation factor) complex which interacts with the upstream element of the RNA polymerase I promoter and forms a stable preinitiation complex.

Subunit

The **nucleosome** is a histone octamer containing two molecules each of H2A, H2B, H3 and H4 assembled in one H3-H4 heterotetramer and two H2A-H2B heterodimers. The octamer wraps approximately 147 bp of DNA. Histone H3 is a component of the UAF (upstream activation factor) complex, which consists of UAF30, RRN5, RRN9, RRN10, and histones H3 and H4. 1 Publication



Binary interactions



Informace o proteinu – IntAct

Histon H3

https://www.ebi.ac.uk/intact

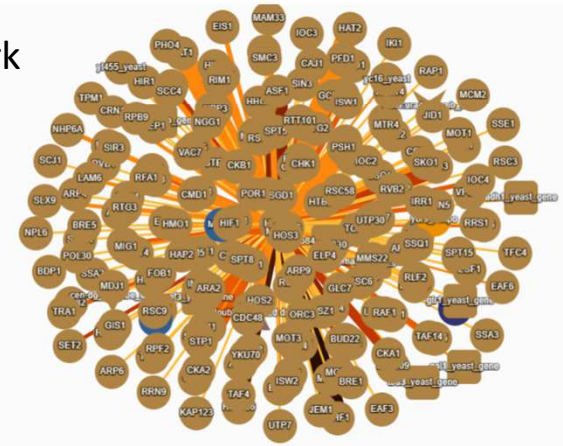
Protein-protein interaction databases

BioGRID	32711 918 interactors 35796 705 interactors
ComplexPortal	CPX-1101 RNA polymerase I upstream activating factor complex CPX-1610 Nucleosome, variant HTA2-HTB2
DIP	DIP-417N
IntAct	P61830 153 interactors
MINT	P61830
STRING	4932.YBR010W

Select	Molecule A	Molecule B	Identifier A	Identifier B	Type A	Type B	Species A	Species B	Host Organism	Positive interaction	Detection Method	Publication IDs	Interaction Type
☐	H4	H3	UniProt P02309	UniProt P61830	protein	protein				✓	tap	16554755 10.1038/nature04670	physical associatio

… i nepřímé interakce

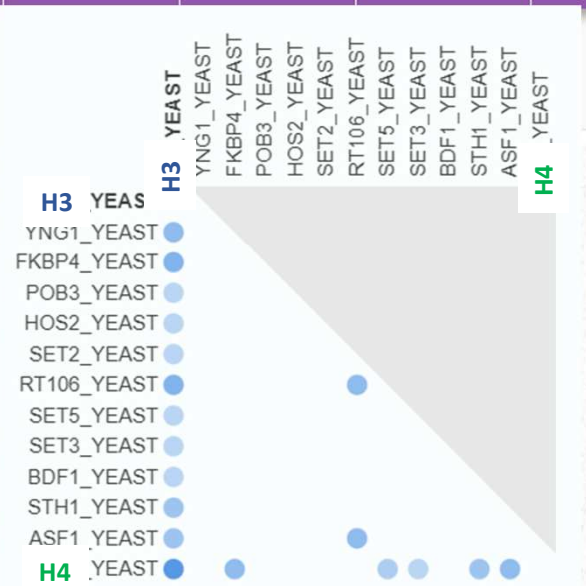
network
síť

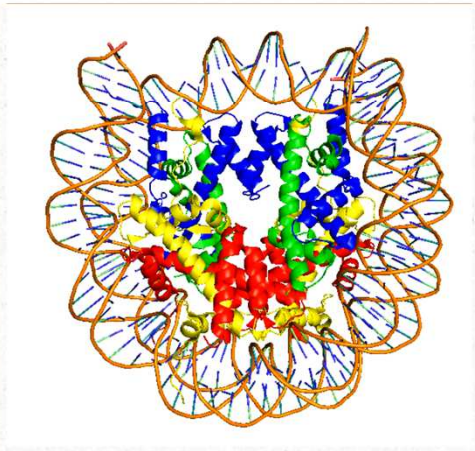


chaperon
deacetylasa
metyltransferasa
chaperon
metyltransferasa

bromodoména

chaperon
histon H4





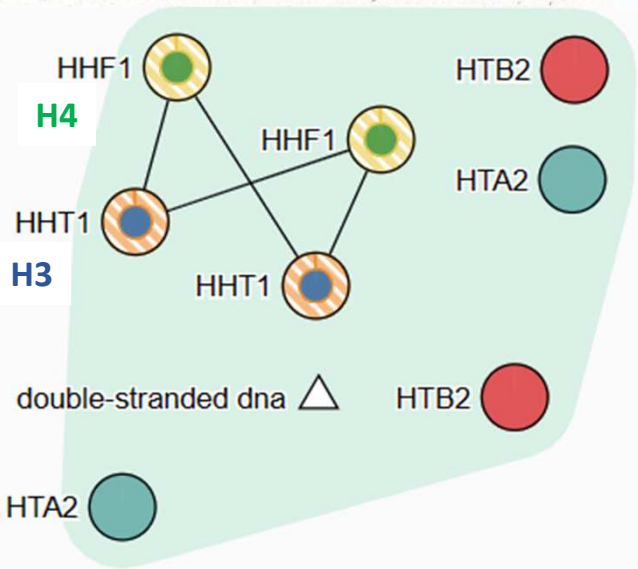
Complex Portal

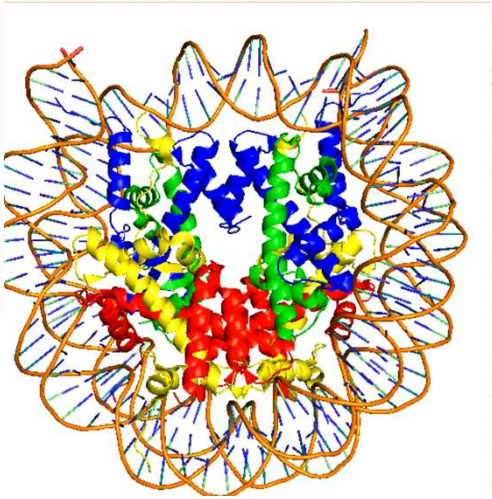
https://www.ebi.ac.uk/complexportal/complex

Histon H3

Protein-protein interaction databases

BioGRID	32711 918 interactors 35796 705 interactors	DIP	DIP-417N
		IntAct	P61830 153 interactors
ComplexPortal	CPX-1101 RNA polymerase I upstream activating factor complex CPX-1610 Nucleosome, CPX-1611 Nucleosome, CPX-1612 Nucleosome, CPX-1613 Nucleosome, More ComplexPortal links	MINT	P61830





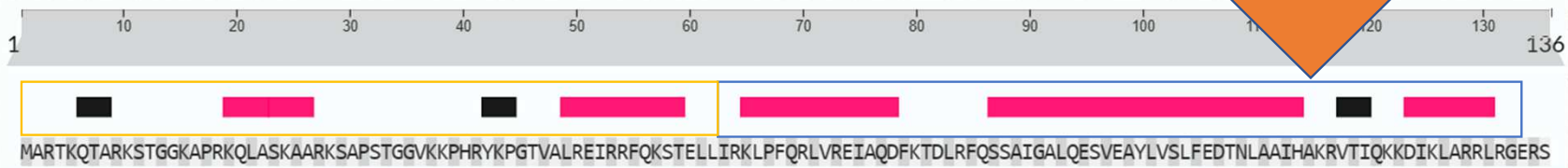
Legend	Description		Stoichiometry
●	**protein - HTA2 (unspecified role)** P04912 Histone H2A.2	H2A	2
●	**protein - HTB2 (unspecified role)** P02294 Histone H2B.2	H2B	2
●	**protein - HHT1 (unspecified role)** P61830 Histone H3	H3	2
●	**protein - HHF1 (unspecified role)** P02309 Histone H4	H4	2
△	**small molecule - double-stranded dna (unspecified role)** CHEBI:4705 double-stranded DNA		1

Informace o proteinu – UNPROT

Histon H3

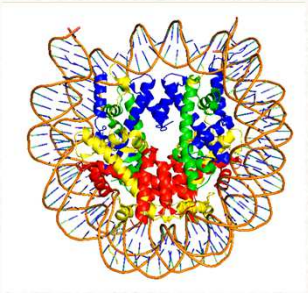
sekundární struktury

- Function
- Names & Taxonomy
- Subcellular Location
- Disease & Variants
- PTM/Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequence & Isoforms
- Similar Proteins



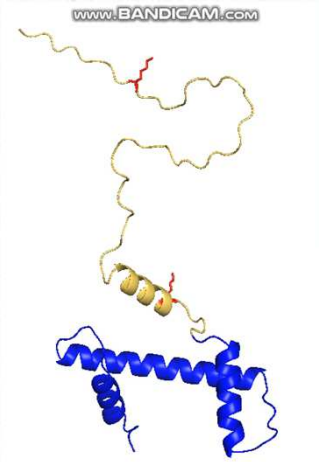
Family and domain databases

Gene3D	1.10.20.10 Histone, subunit A 1 hit
IDEAL	IID50143
InterPro	View protein in InterPro IPR009072 Histone-fold IPR007125 Histone_H2A/H2B/H3 IPR000164 Histone_H3/CENP-A
PROSITE	View protein in PROSITE PS00322 HISTONE_H3_1 1 hit PS00959 HISTONE_H3_2 1 hit
Pfam	View protein in Pfam PF00125 Histone 1 hit



Model Confidence:

- Very high (pLDDT > 90)
- Confident (90 > pLDDT > 70)
- Low (70 > pLDDT > 50)
- Very low (pLDDT < 50)



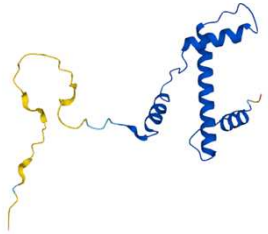
terciární a kvarterní struktura

https://alphafold.ebi.ac.uk/

Structure

PDB	7XAY	X-ray	3.30 Å	D	2-136	PDBe · RCSB-PDB · PDBj · PDBsum
PDB	7Z0O	EM	2.80 Å	A/C	1-136	PDBe · RCSB-PDB · PDBj · PDBsum
AlphaFold	AF-P61830-F1	Predicted			1-136	AlphaFold

Proteinové domény



Protein Classification

histone H3(domain architecture ID 10794185)
histone H3 is a core component of the nucleosome that wraps and compacts DNA into chromatin, limiting DNA accessibility to the cellular machineries which require DNA as a template

Graphical summary

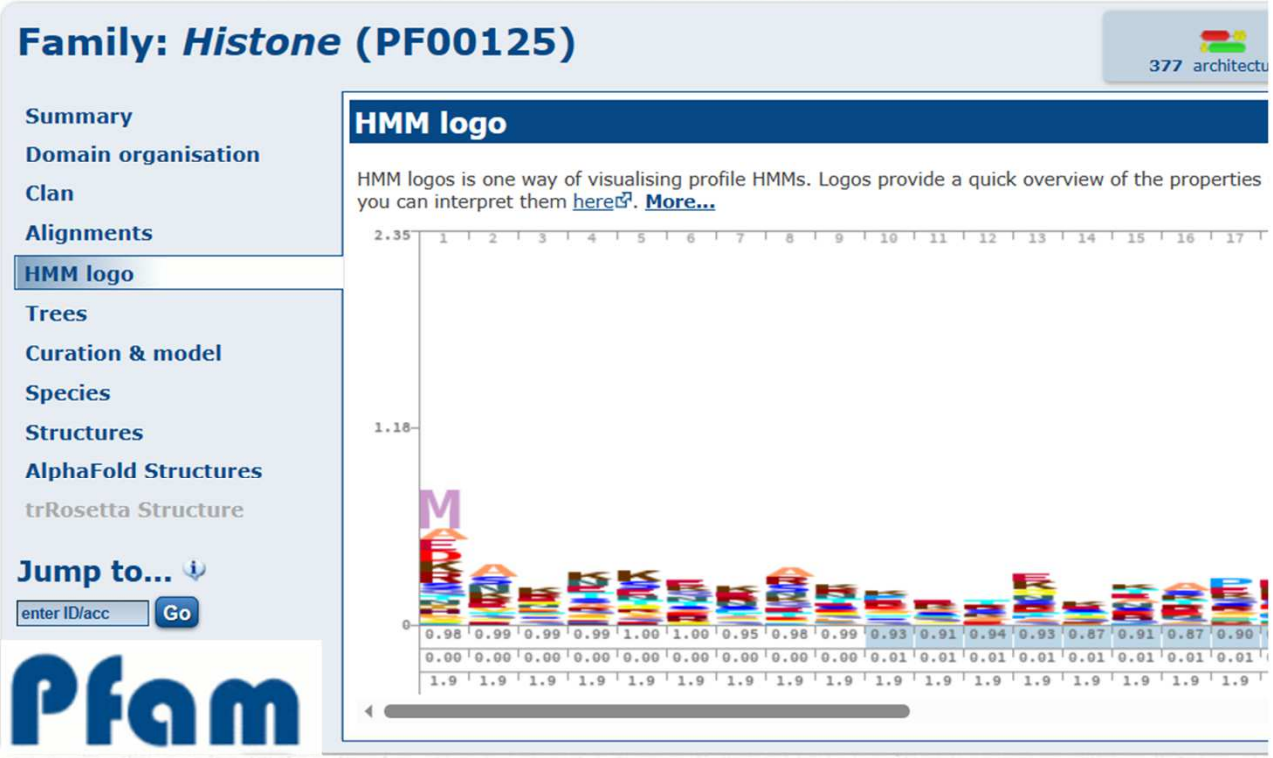
Zoom to residue level show extra options »

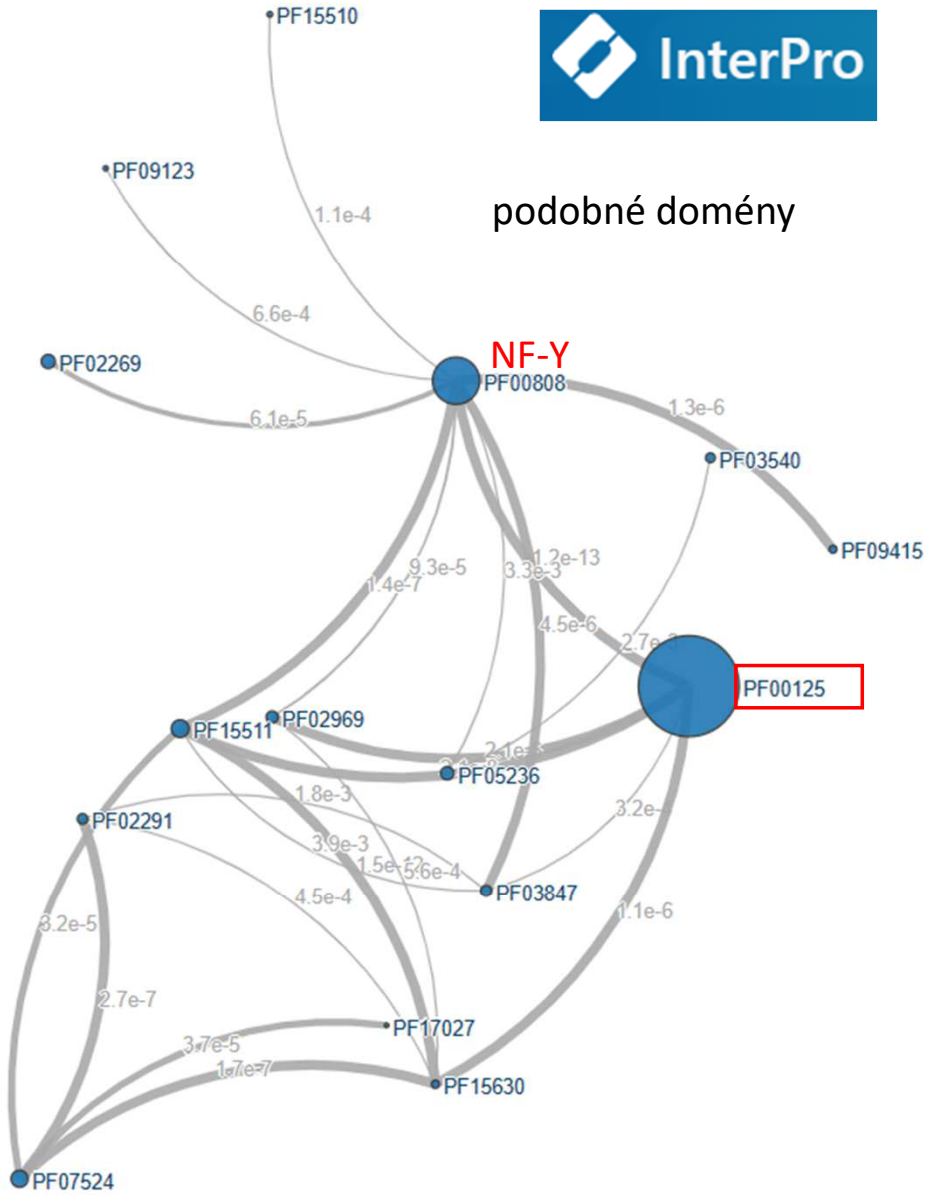
List of domain hits

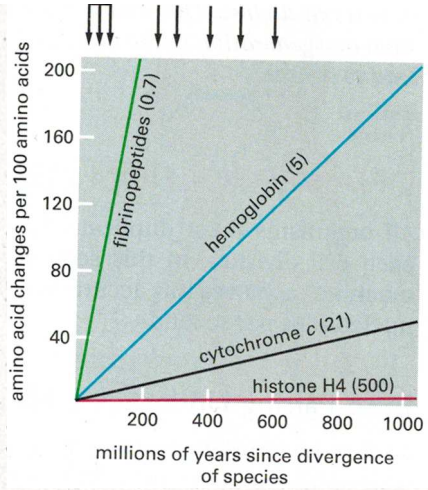
Name	Accession	Description	Interval	E-value
[+] PTZ00018	PTZ00018	histone H3; Provisional	1-135	1.16e-85
[+] H3	smart00428	Histone H3;	37-135	5.99e-58
[+] Histone	pfam00125	Core histone H2A/H2B/H3/H4;	1-131	5.89e-49
[+] HHT1	COG2036	Archaeal histone H3/H4 [Chromatin structure and dynamics];	43-134	8.80e-30

https://prosite.expasy.org/, https://www.ebi.ac.uk/interpro/entry/pfam

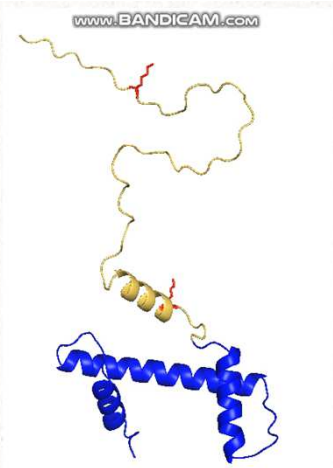
Proteinové domény

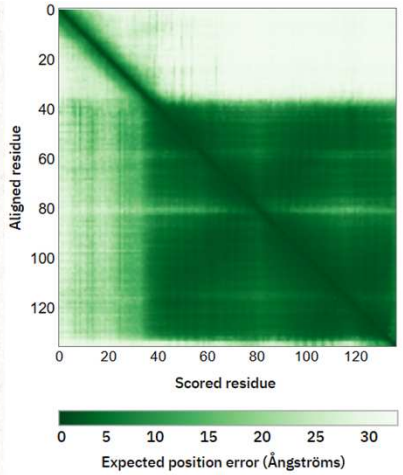


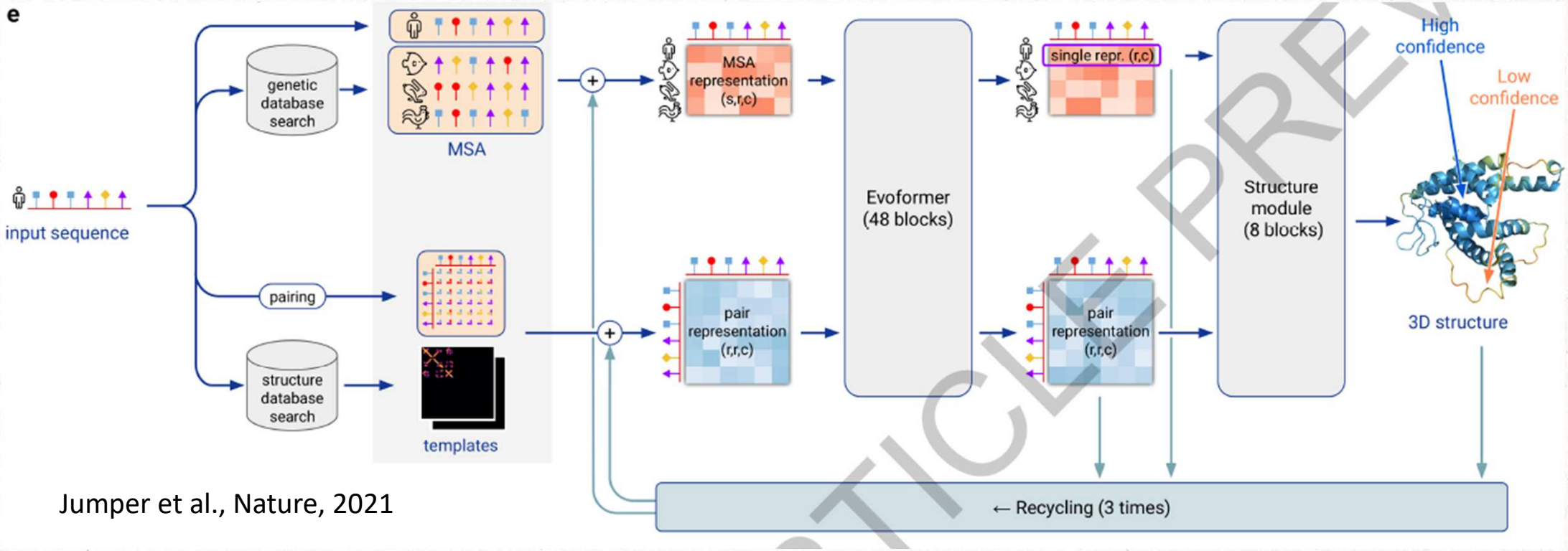




AlphaFold



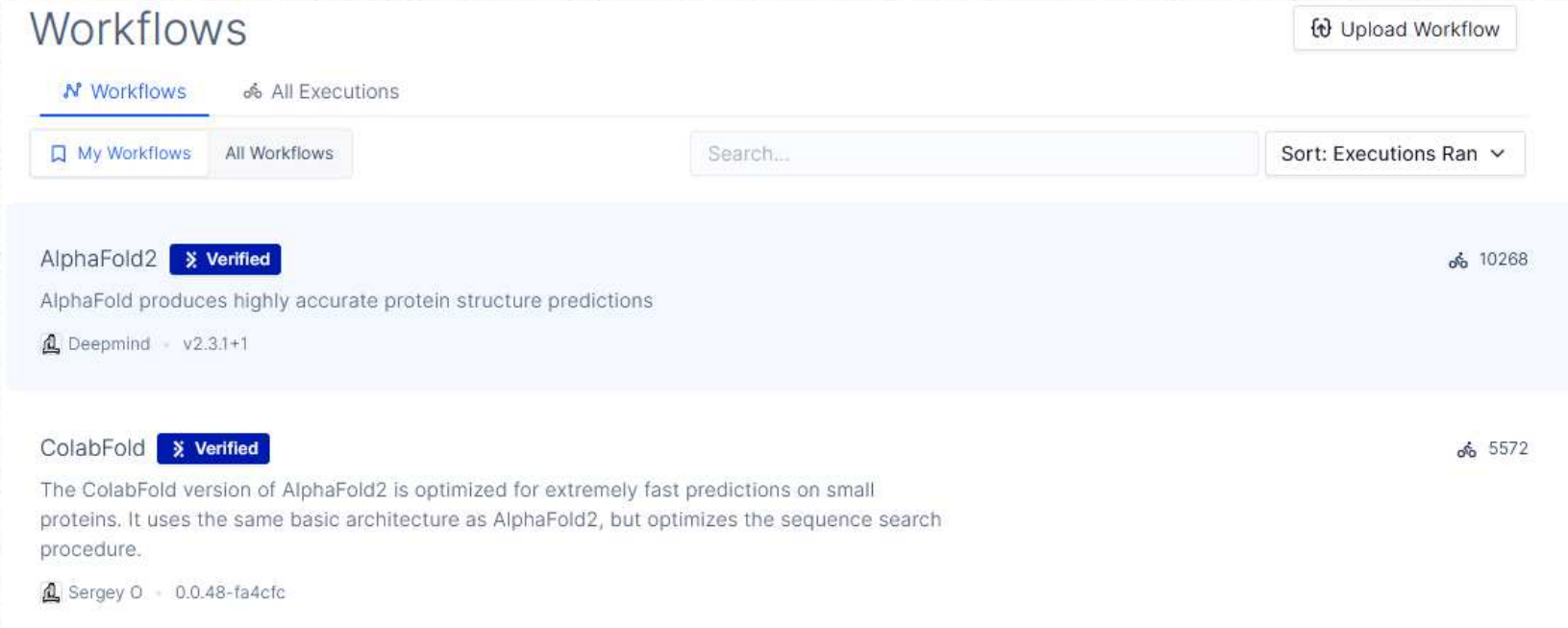




Jumper et al., Nature, 2021

AlphaFold – predikce 3D struktury proteinu i komplexů

- Využití online predikce Alphafold
  - https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=G4yBrceuFbf3
  - LatchBio: https://console.latch.bio/workflows

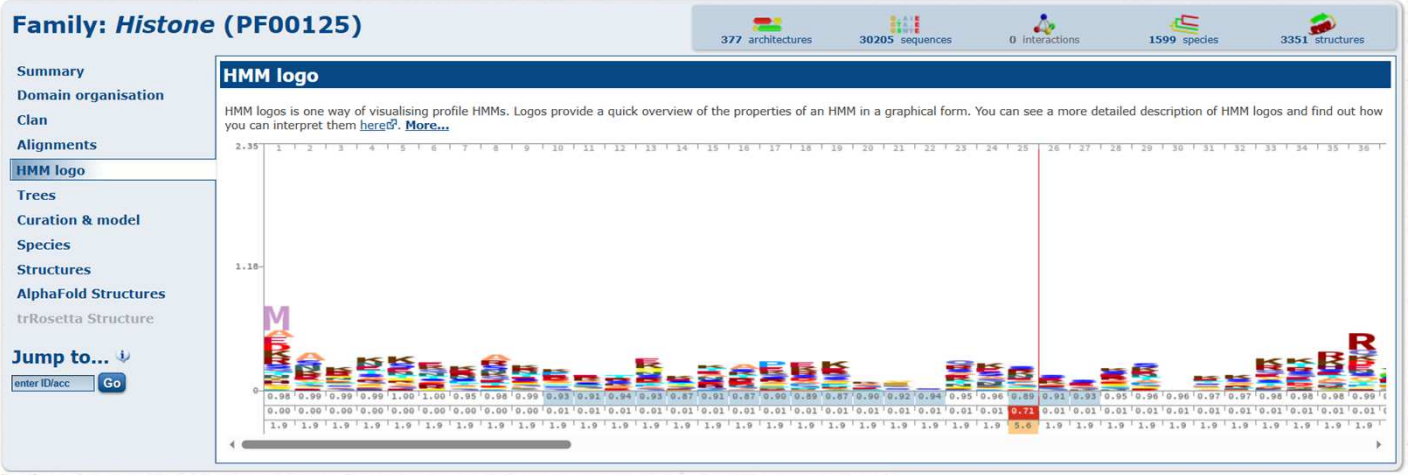


- Umožňuje i predikci 3D struktury proteinového komplexu (dimeru dvou proteinů – díky jejich **koevoluci**)

Informace o proteinu – UNPROT

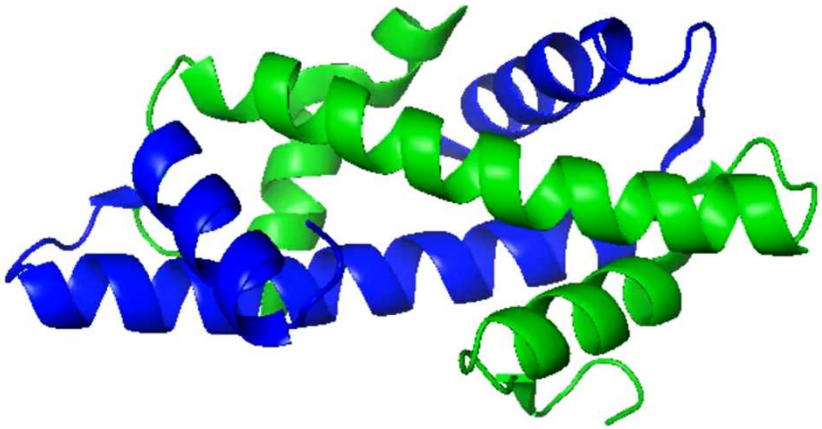
kvarterní struktura

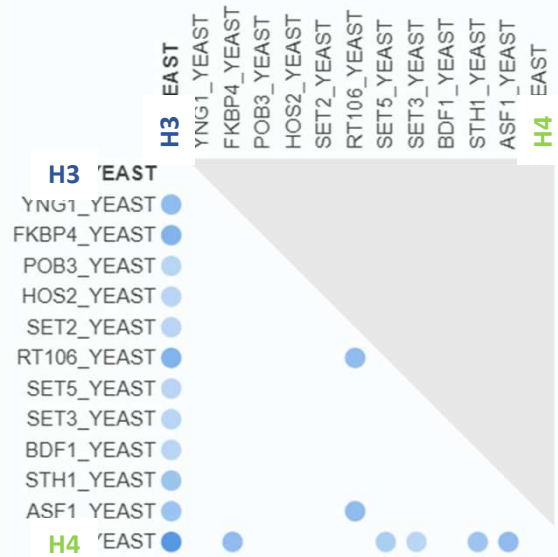
- Function
- Names & Taxonomy
- Subcellular Location
- Disease & Variants
- PTM/Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequence & Isoforms
- Similar Proteins



Histon fold – typicky P-P vazba … vazba na DNA

H3-H4 dimer





Informace o proteinu – UNPROT

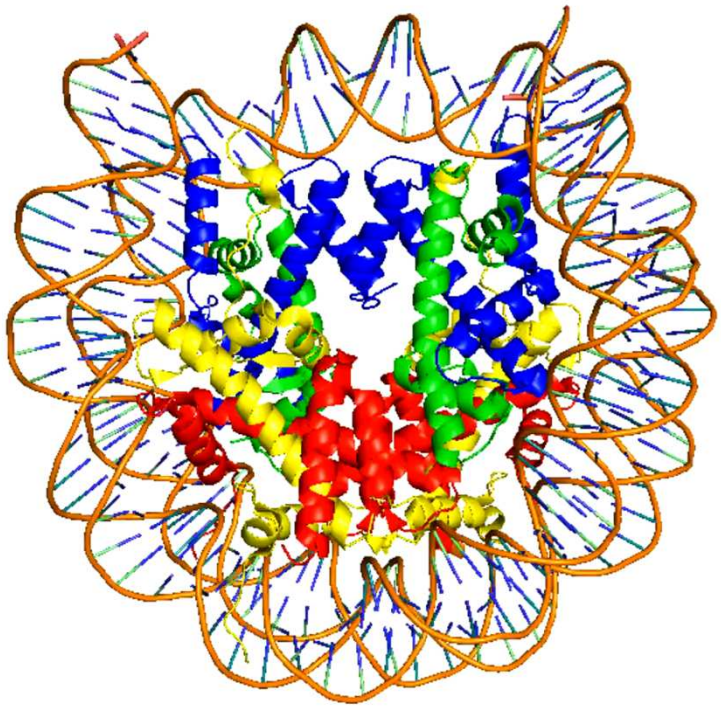
Histon H3

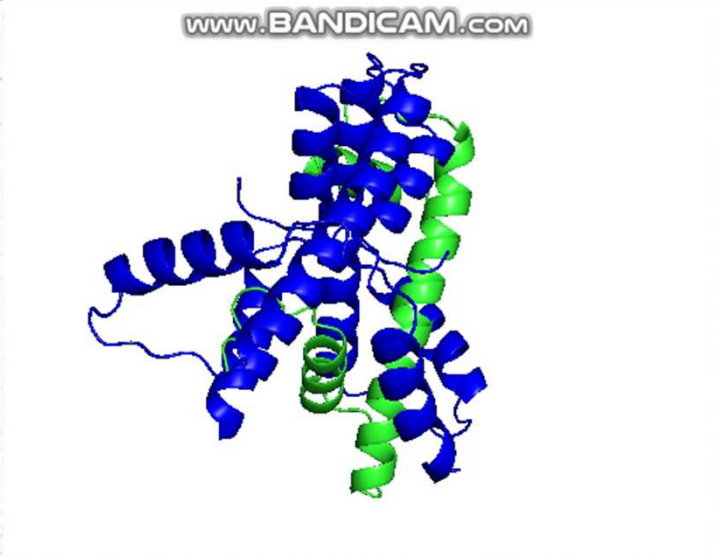
- Function
- Names & Taxonomy
- Subcellular Location
- Disease & Variants
- PTM/Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequence & Isoforms
- Similar Proteins

Structure

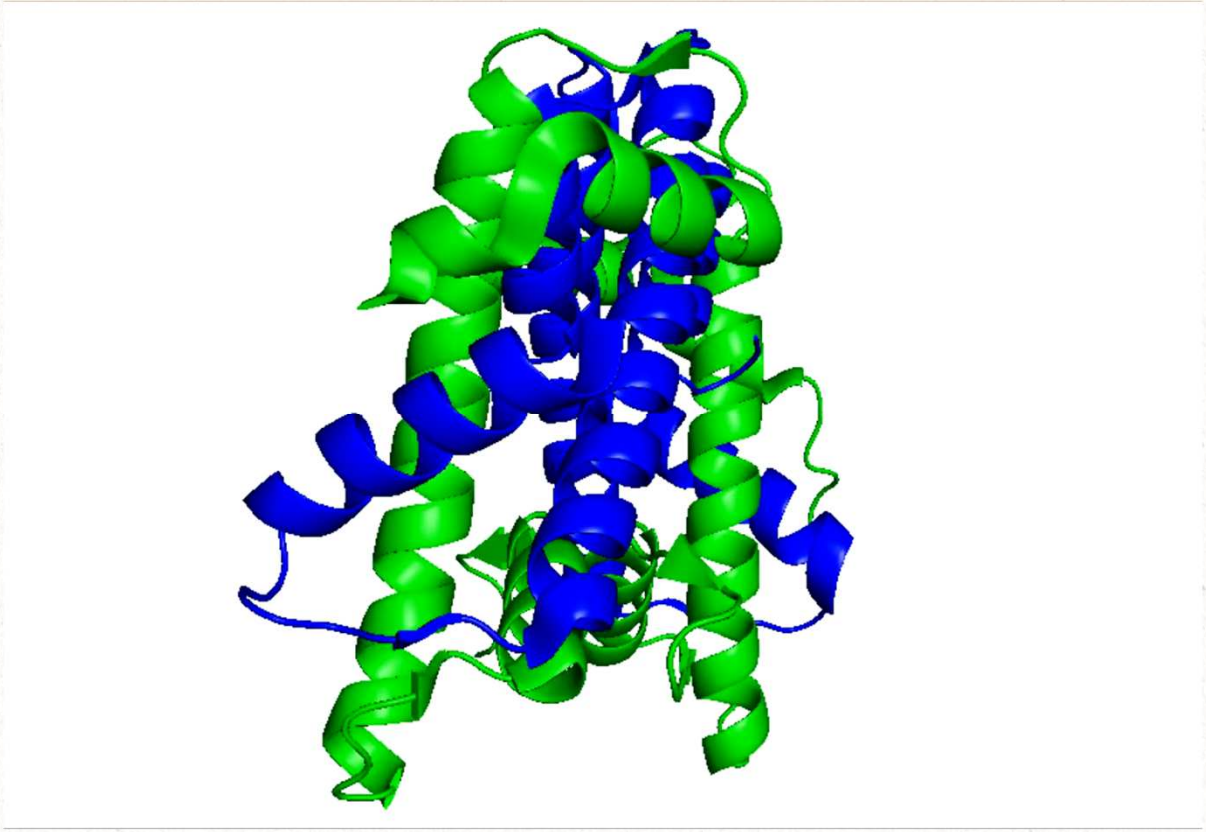
kvarterní struktura

PDB	7XAY	X-ray	3.30 Å	D	2-136	PDBe · RCSB-PDB · PDBj · PDBsum
PDB	7Z0O	EM	2.80 Å	A/C	1-136	PDBe · RCSB-PDB · PDBj · PDBsum
AlphaFold	AF-P61830-F1	Predicted			1-136	AlphaFold



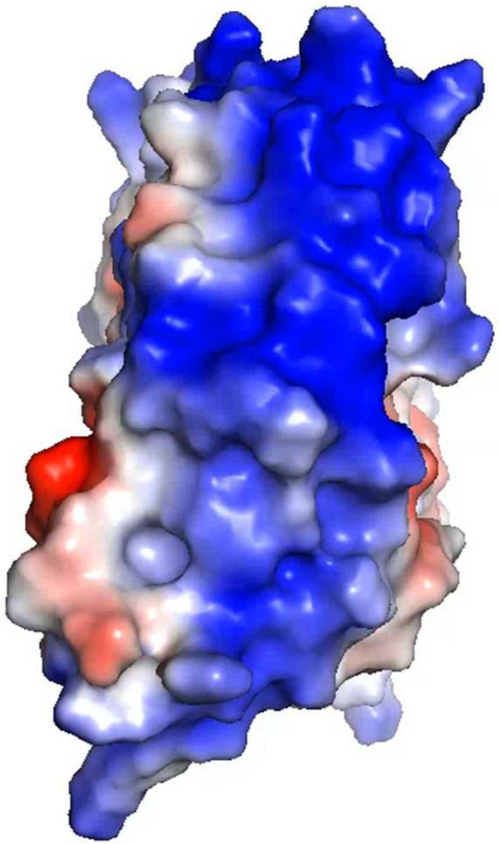


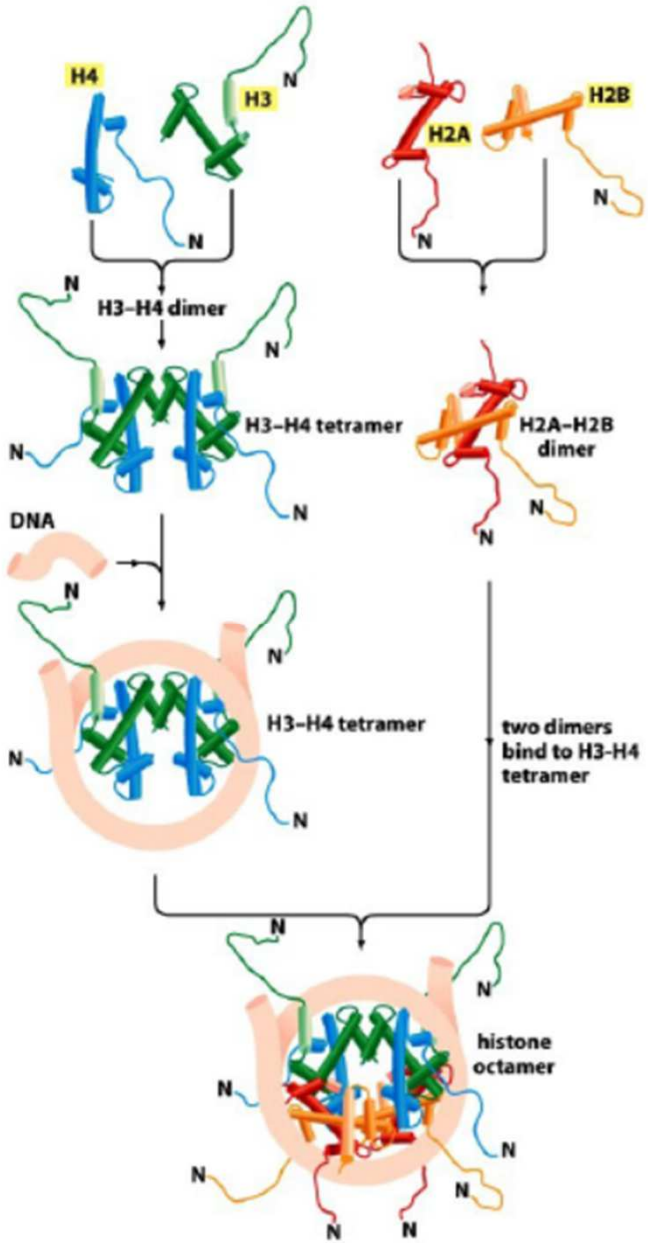
H3-H3 dimer vygeneroval alphafold jako H3-H4 dimer – až po použití H3-H4+H3 umístil H3-H3 dimer správně

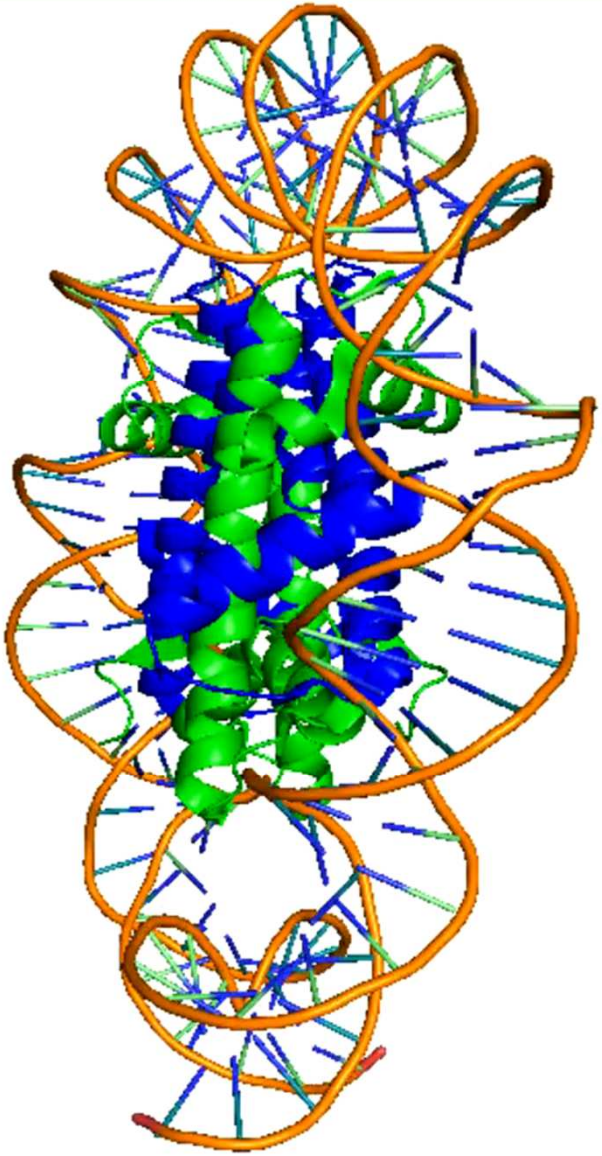


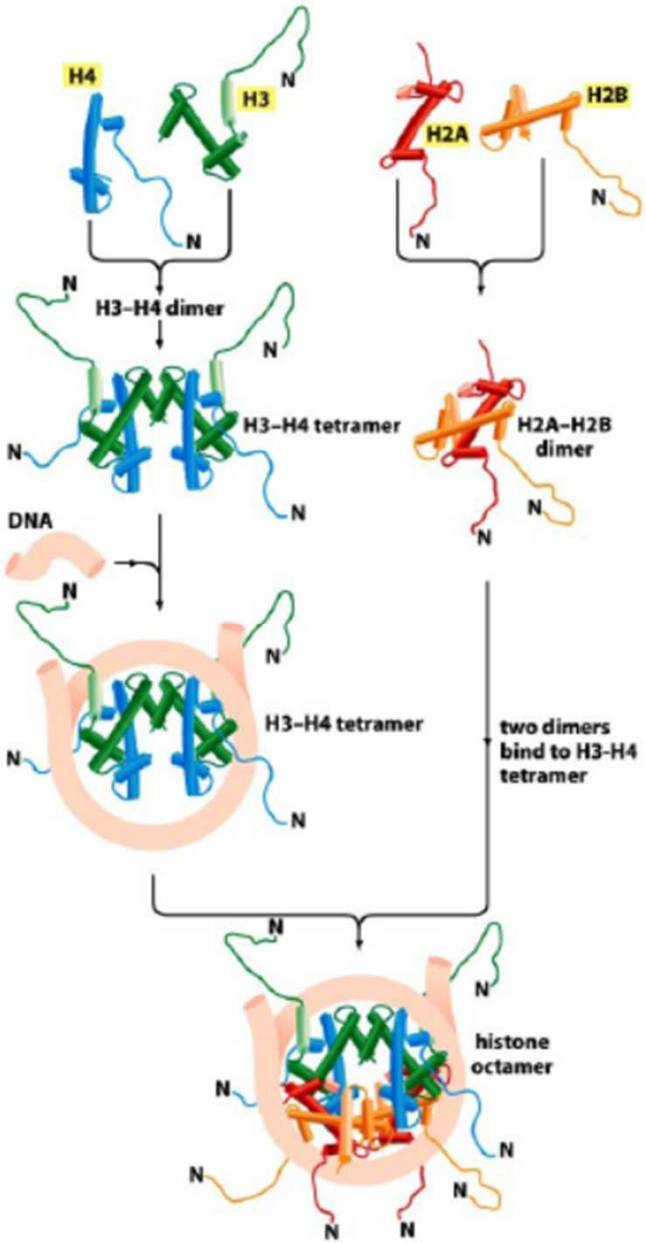
nukleosom se skládá na „začátku“ (replikace DNA – vznik nové DNA …), ale i v průběhu života buňky (oprava DNA, transkripce) – odbalení a sbalení nukleosomu regulují chaperony, remodelační komplexy …

- H3-H3 interakce = $(H3\text{-}H4)_2$ tetramer

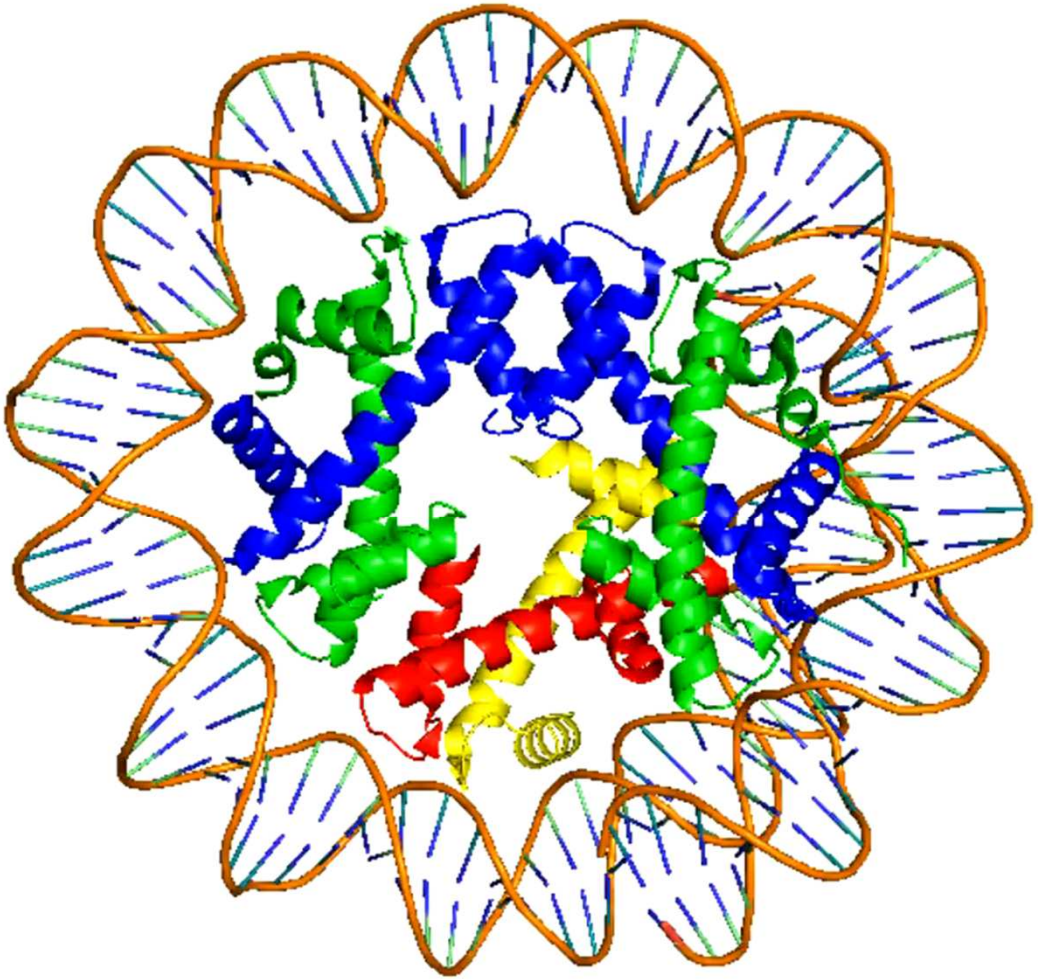


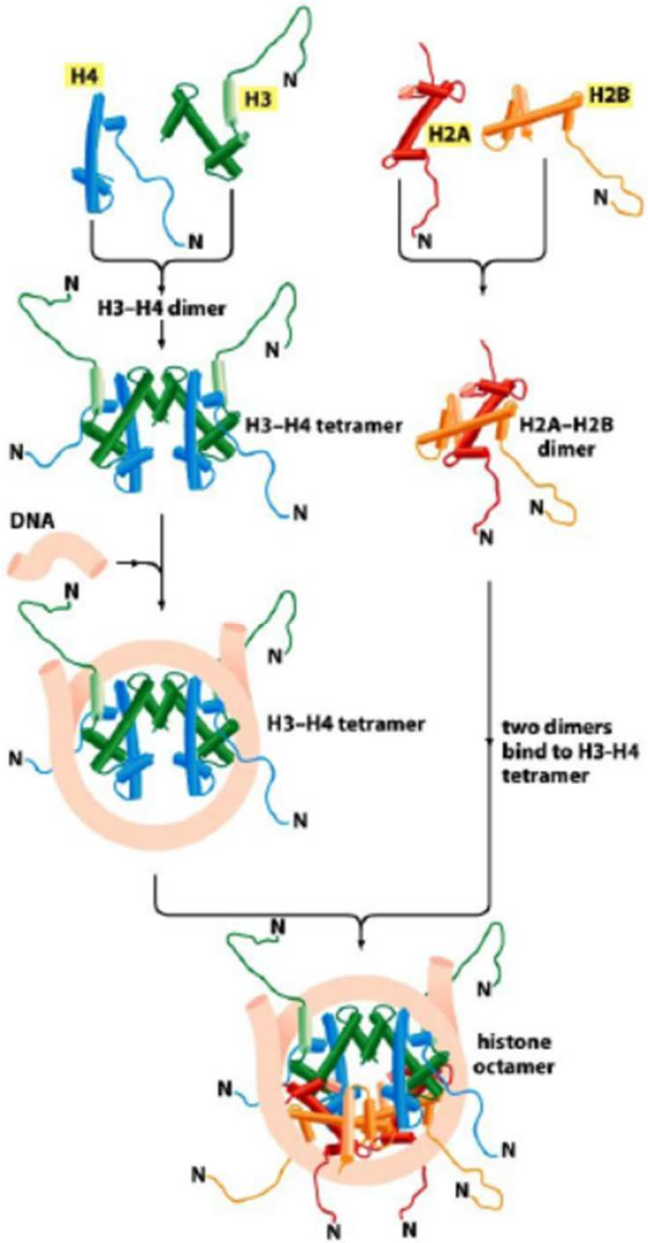

H4
H3
H2A
H2B
N
H3–H4 dimer
H3–H4 tetramer
H2A–H2B dimer
DNA
H3–H4 tetramer
two dimers bind to H3-H4 tetramer
histone octamer



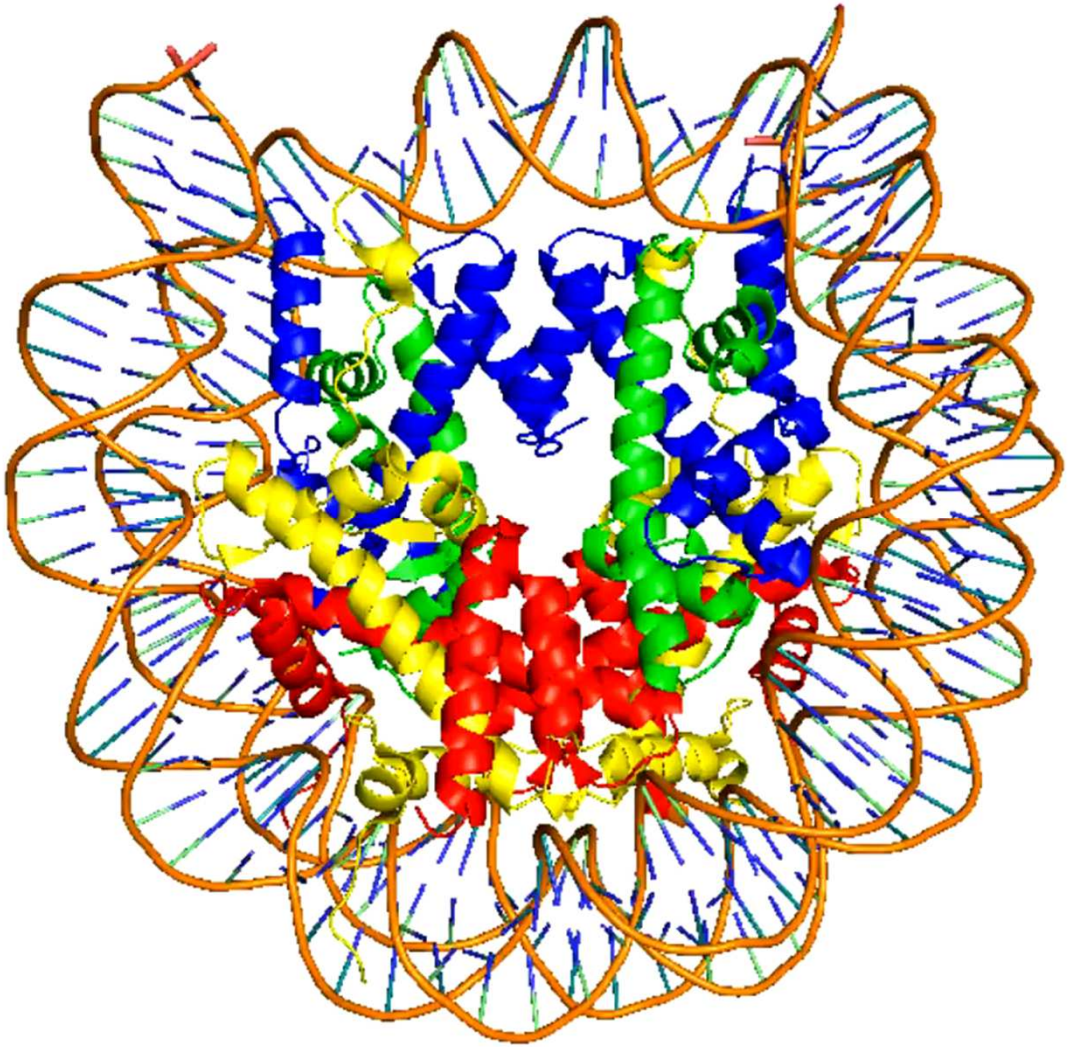


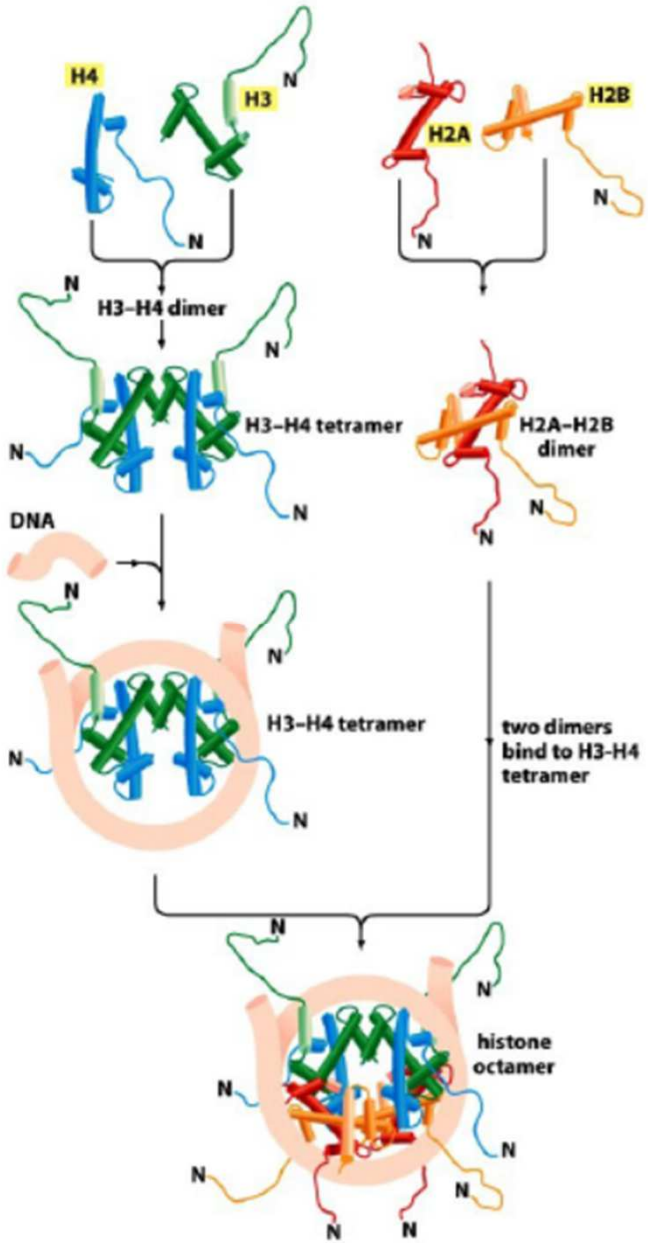
- H3-H3 interakce = $(H3\text{-}H4)_2$ tetramer – vazba na DNA





- H3-H3 interakce = $(H3\text{-}H4)_2$ tetramer
- dimery H2A-H2B (H2B – váže H4)



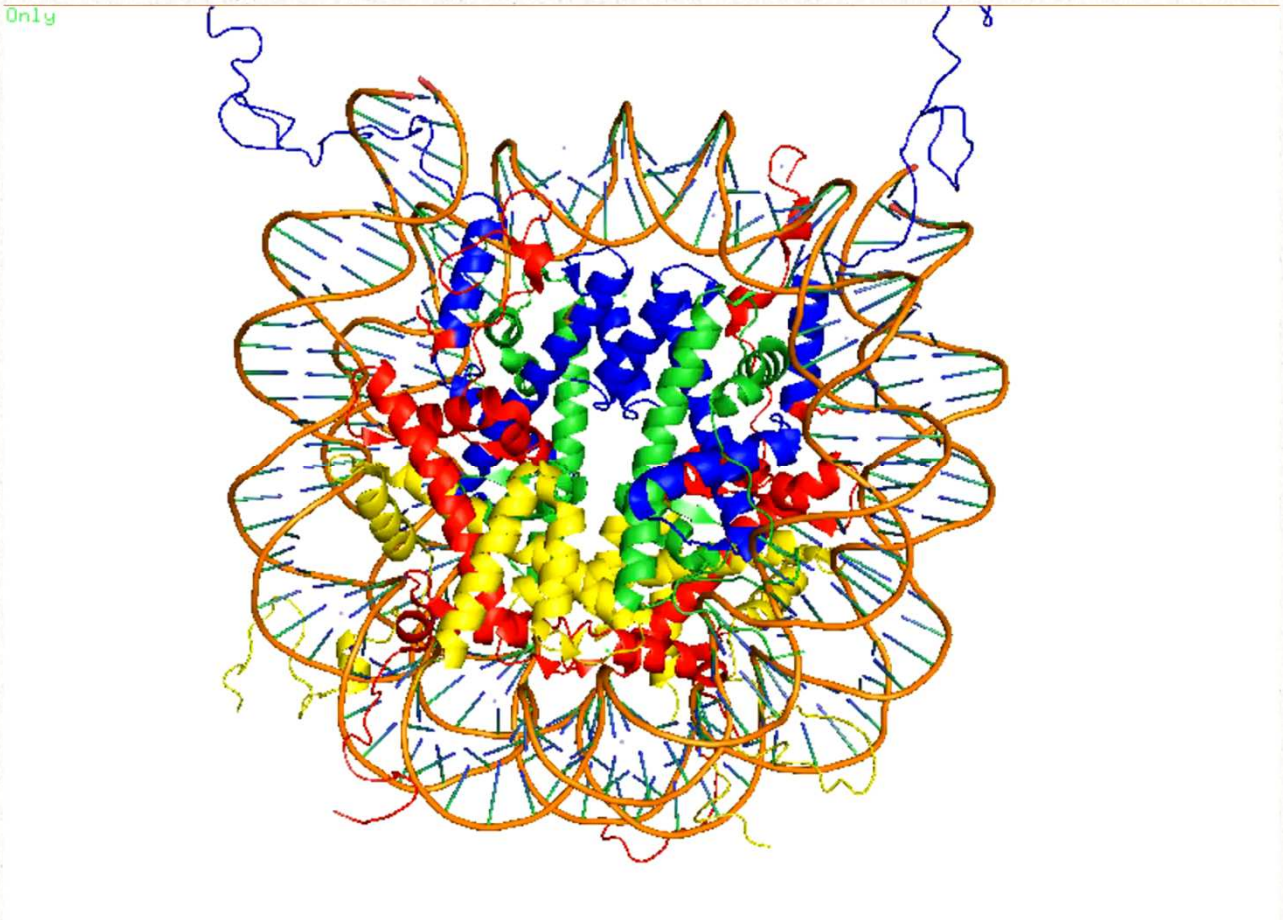


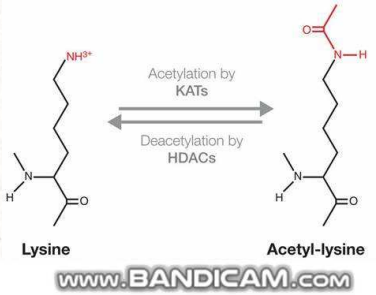
- H3-H3 interakce = $(H3\text{-}H4)_2$ tetramer
- dimery H2A-H2B (H2B – váže H4

Modifikace histonů – H3K56ac

- modifikace (např. acetylace) mohou: vázat jiné proteiny (změní tvar povrchu – váže specificky bromodoména) nebo oslabit interakce (změní náboj povrchu – oslabí vazbu s DNA a rozvolní nukleosom) – pozice PTM

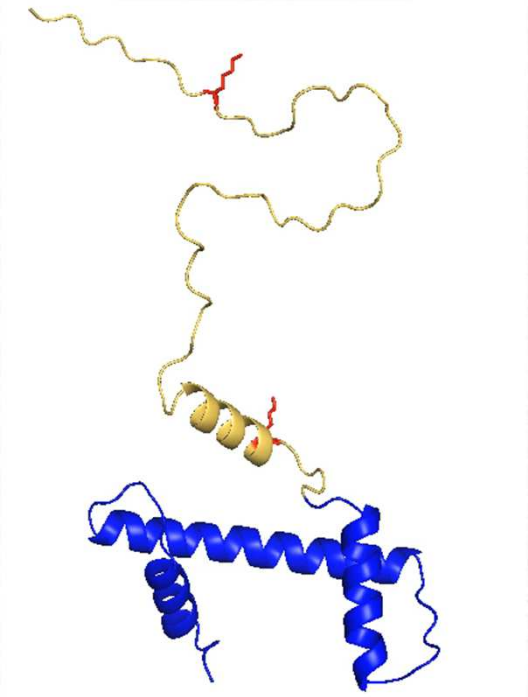
1KX5





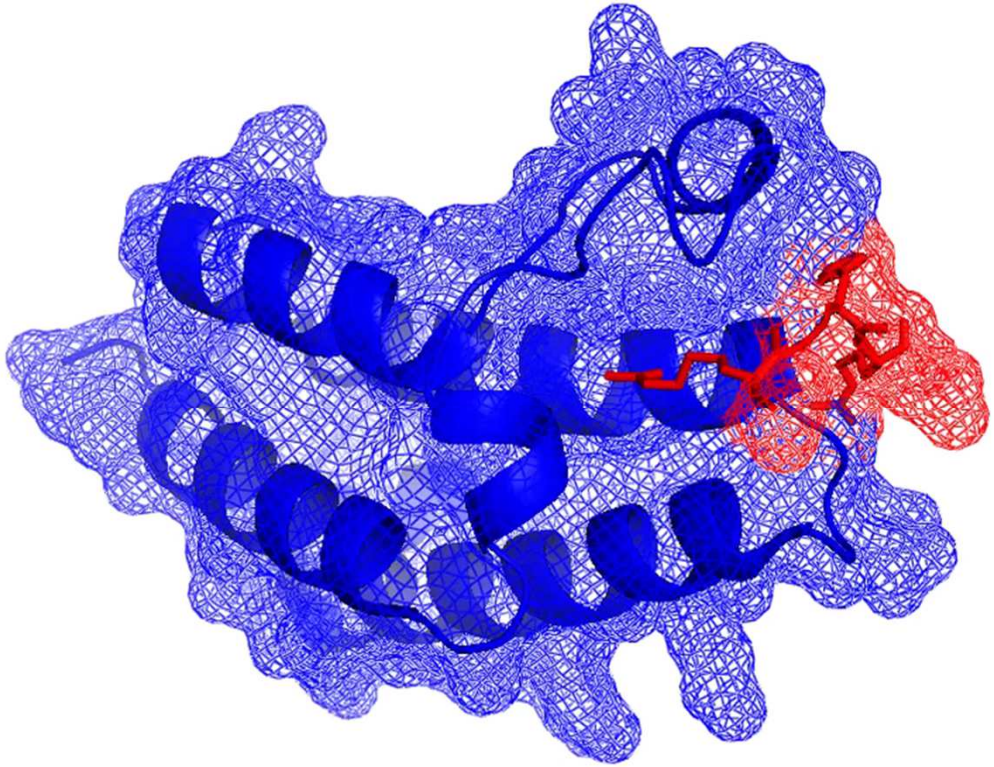
H3K9ac

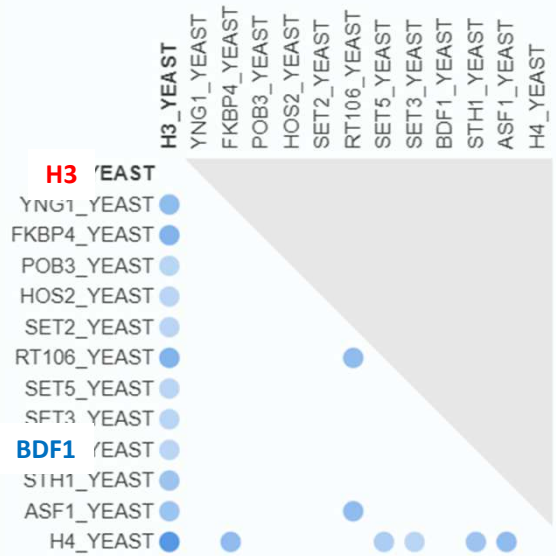
H3K56ac



Modifikace histonů – H3K9ac

- modifikace (např. acetylace) mohou: vázat jiné proteiny (změní tvar povrchu – váže specificky bromodoména) nebo oslabit interakce (změní náboj povrchu – oslabí vazbu s DNA a rozvolní nukleosom) – pozice PTM





bromodoména váže acetylovaný histon

Docking - hra

Bioblox 2½D Game on the Topic of Protein Docking

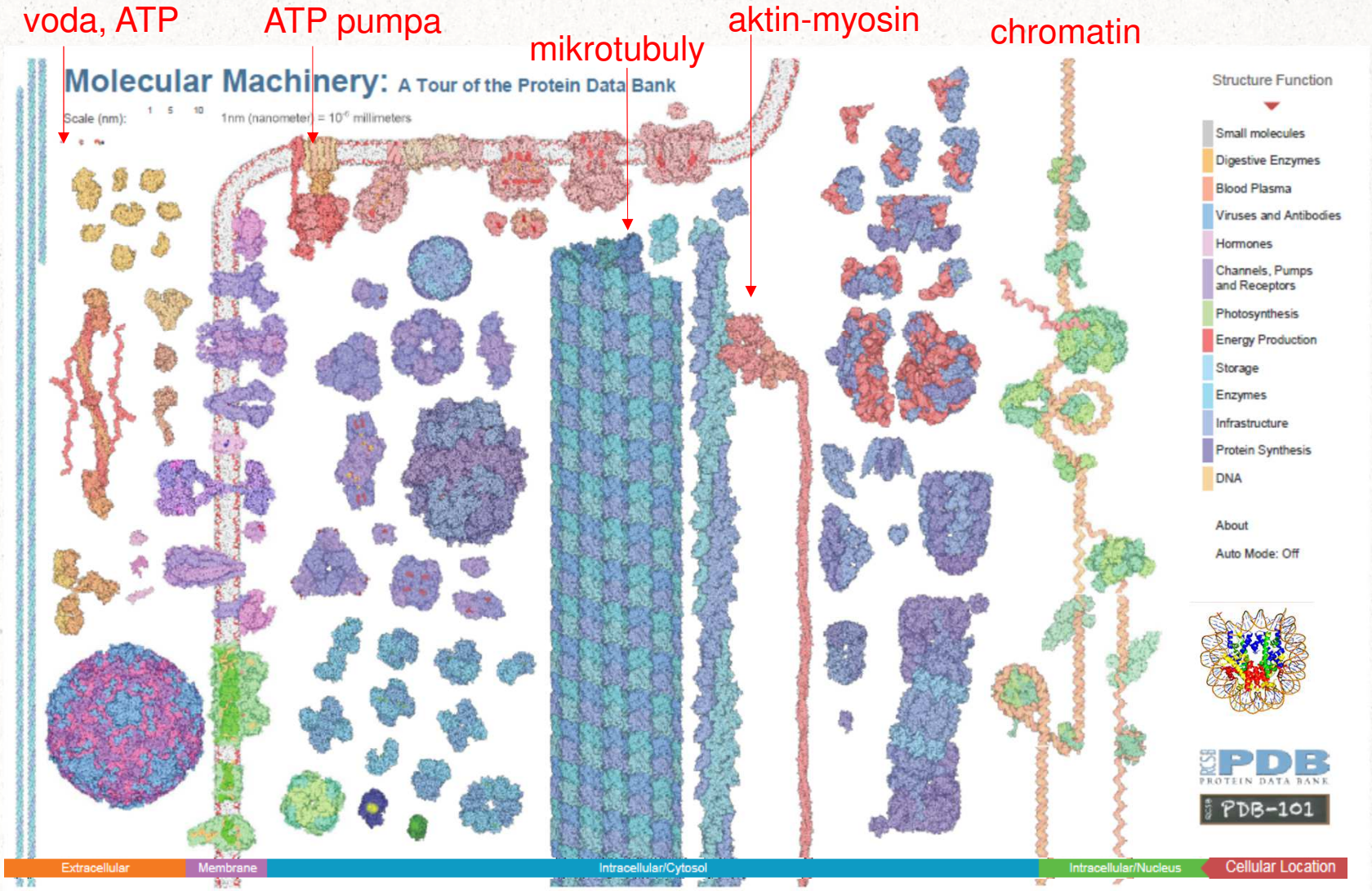
https://www.youtube.com/watch?v=u_-8JyCWpEQ&t=7s

Bioblox 2½D is a free mobile game on the Topic of Protein Docking. Play the Proteins Docking game. Learn about the fascinating world of bio-molecules and their interactions. Drag, Rotate, Swipe and fit the chains together like the components of a mechanism.

Get it on Google play or PLAY ONLINE

https://www.doc.gold.ac.uk/bioblox/

Primárním zdrojem strukturních informací = PDB

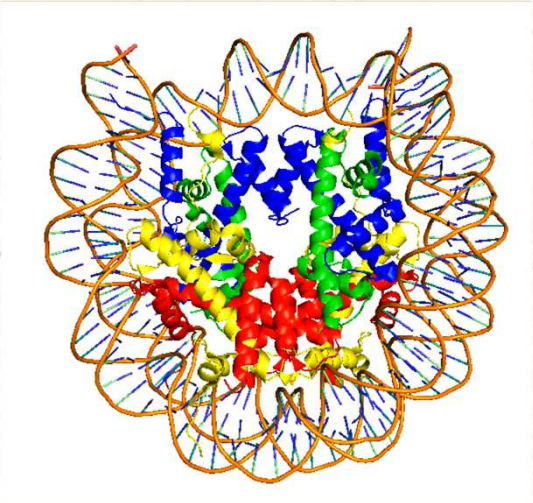


Interaktivní web PDB-101 (pro učitele): https://pdb101.rcsb.org/

přes 200 000 struktur

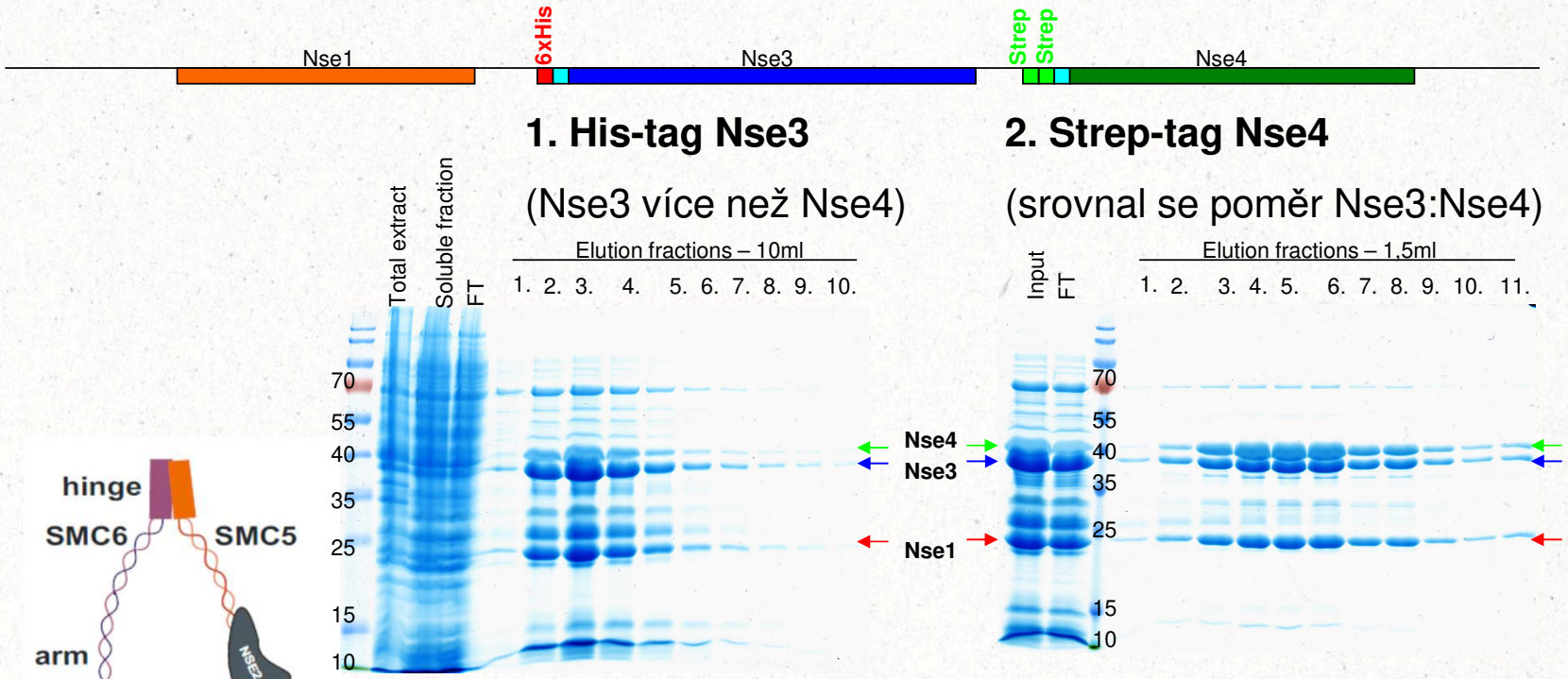
Pokročilé metody v analýze proteinových komplexů

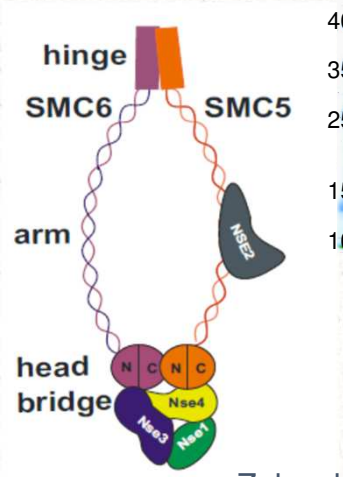
1. - identifikace genů/proteinů … jejich partnerů (PPI) … izolace komplexu
   - funkce podjednotek (genetická analýza, lokalizace v buňce …)
2. - charakterizace komplexu
   - vzájemné PPI podjednotek – architektura komplexu
3. - rekonstituce, struktura a analýza aktivit celého komplexu *in vitro*
   - struktura: crosslink MS, X-ray, (cryo) elektronová mikroskopie, modelování, integrativní analýza



Ko-purifikace - ověření

silné interakce/komplexy – proteiny lze ko-exprimovat a následně ko-purifikovat

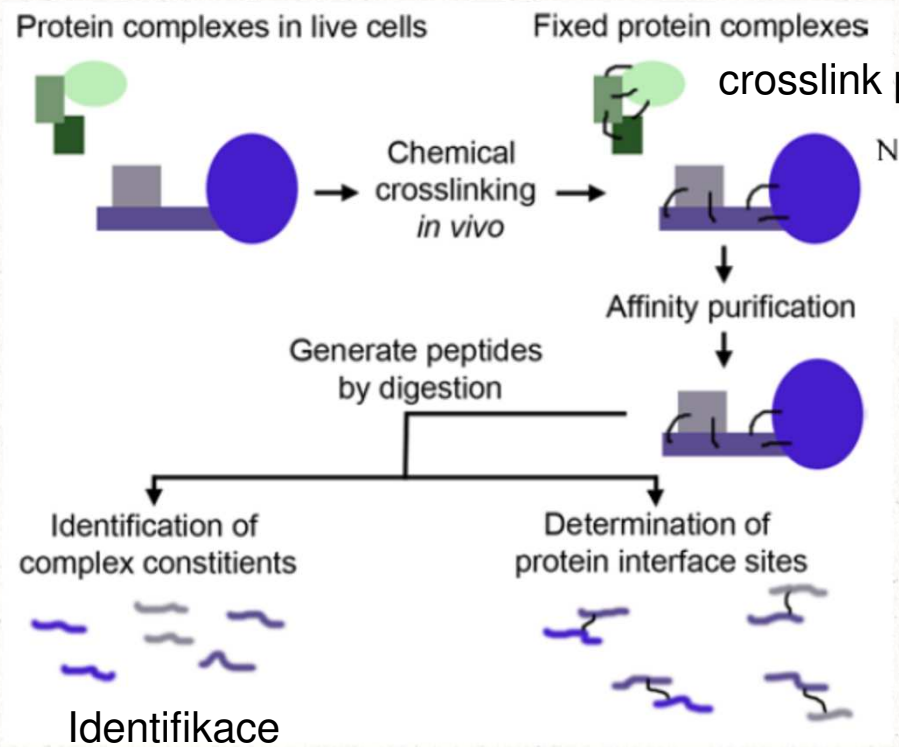




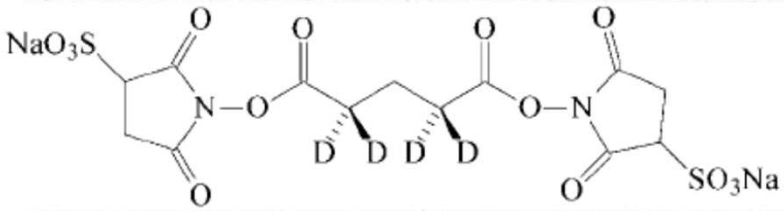
Značky na různých podjednotkách komplexu – postupná purifikace => kompletní komplex (přirozený poměr podjednotek v komplexu)

Zabrady et al, NAR, 2016

Detailní mapování komplexů - crosslinking



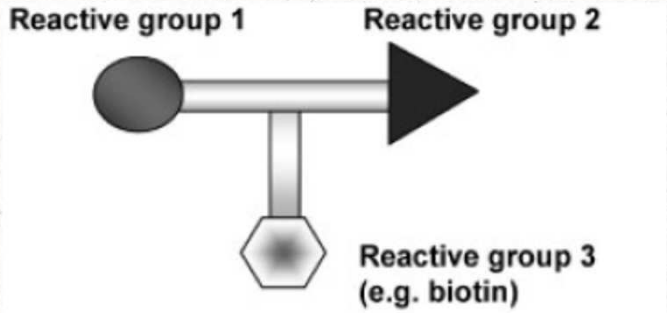
crosslink propojí podjednotky - stabilizuje komplex



- estery reagují s Lys (ε-amin)
- homofunkční - spojí proteiny dohromady v jednom kroku
- heterofunkčních - postupná aktivace

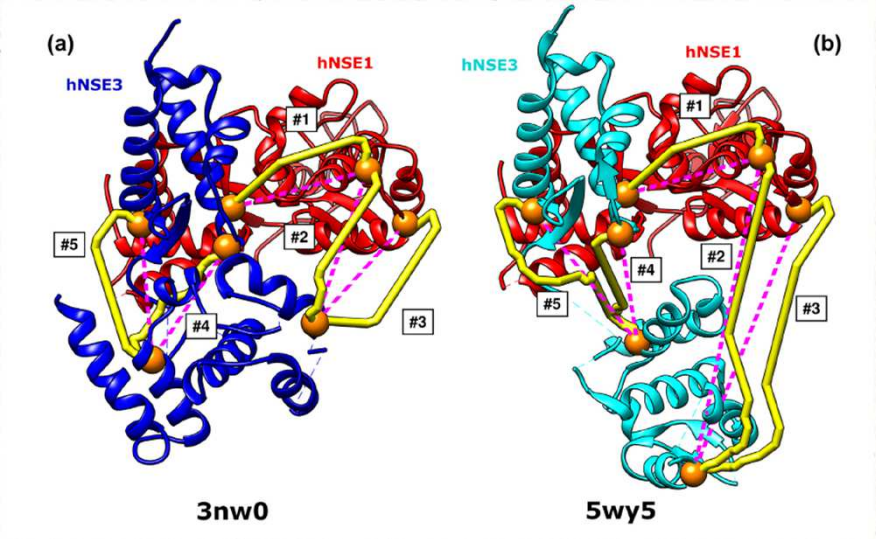
Identifikace podjednotek komplexu

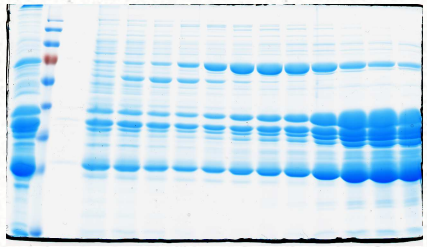
Mapování interakcí podjednotek

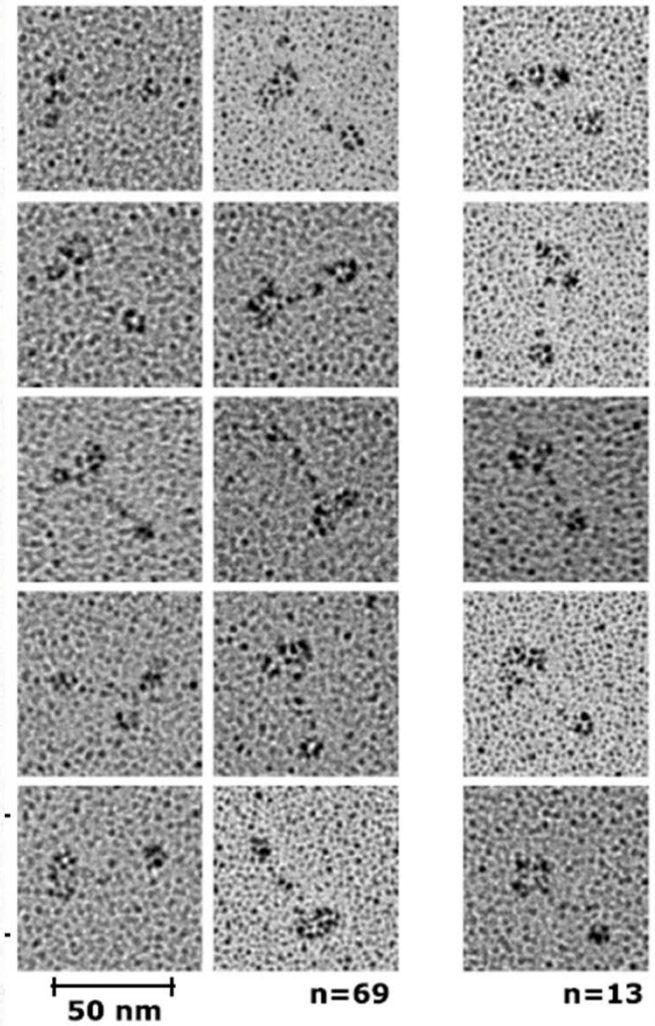


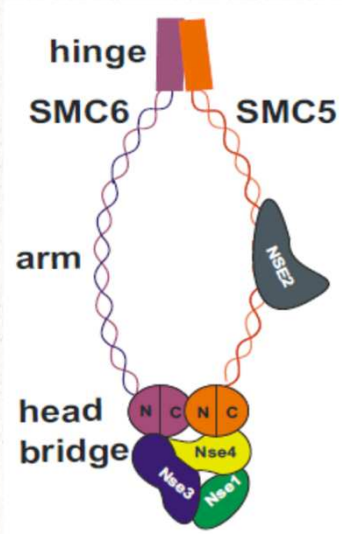
Roux, CMLS, 2013
Sinz, MS Reviews, 2006

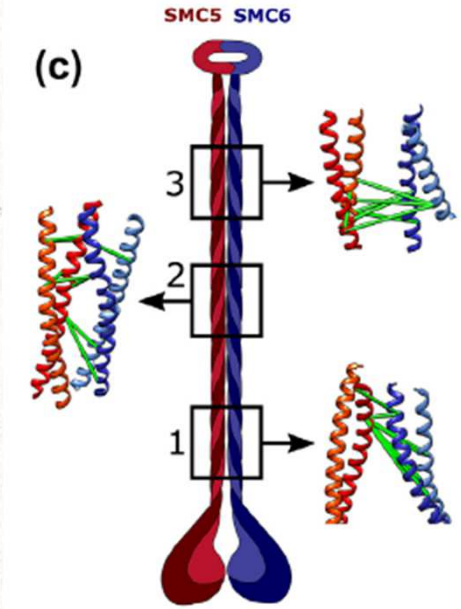
- krosslink pomůže upřesnit strukturní model (Nse1-Nse3)
- ukázal jak jsou ramena Smc5-Smc6 vedle sebe umístěna









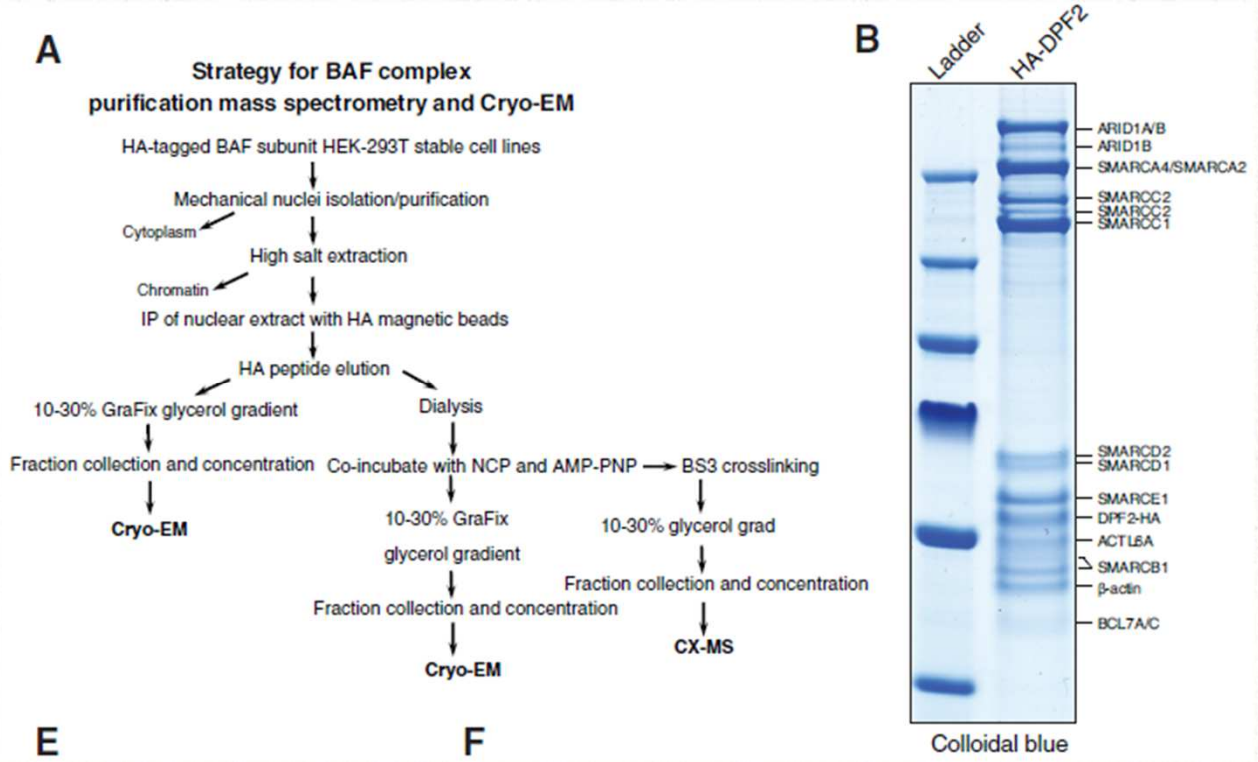


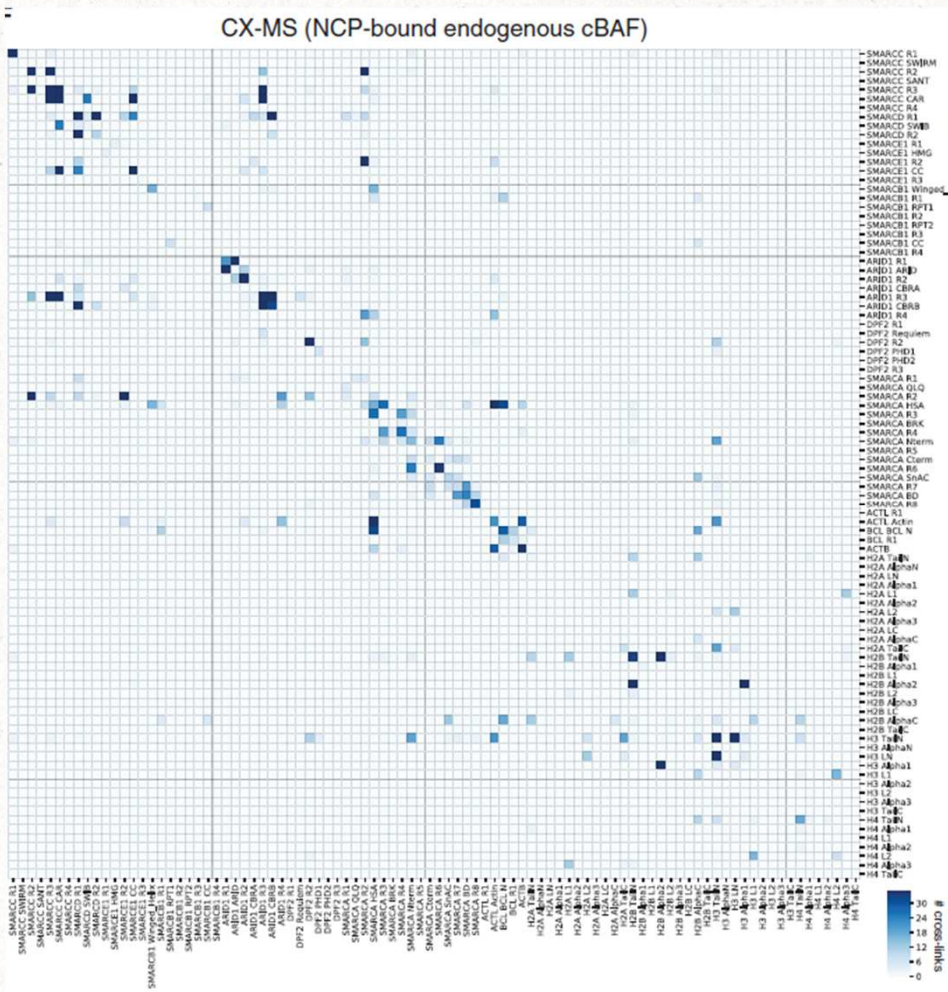
Adamus et al, JMB, 2020

Integrativní modelování

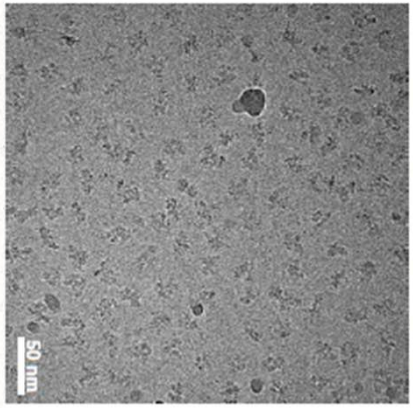
remodelační komplex

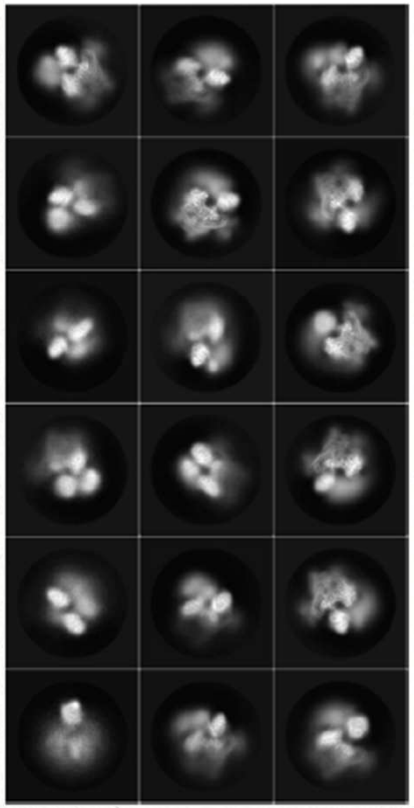
krosslink podjednotek a mapování na připravené modely

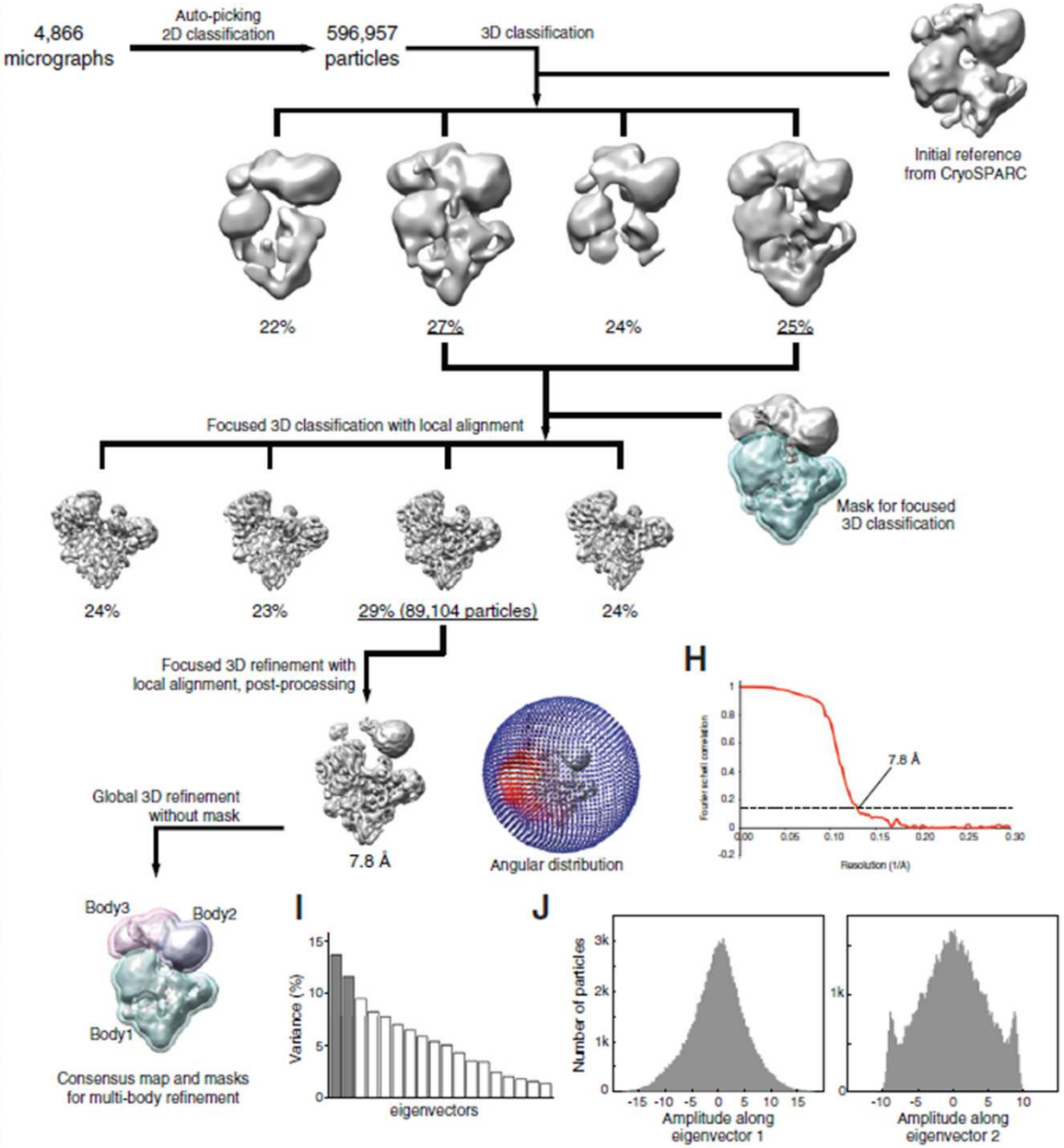




Mashtalir et al, Cell, 2020



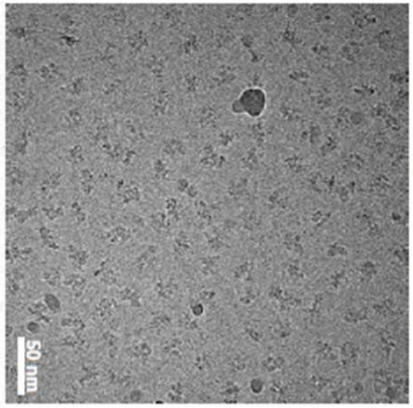


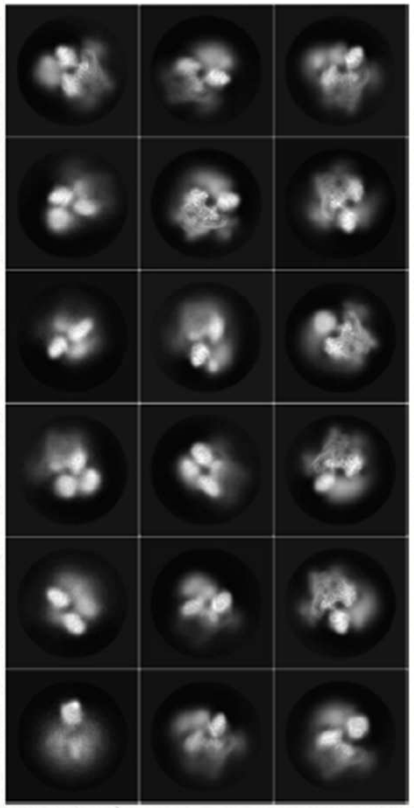


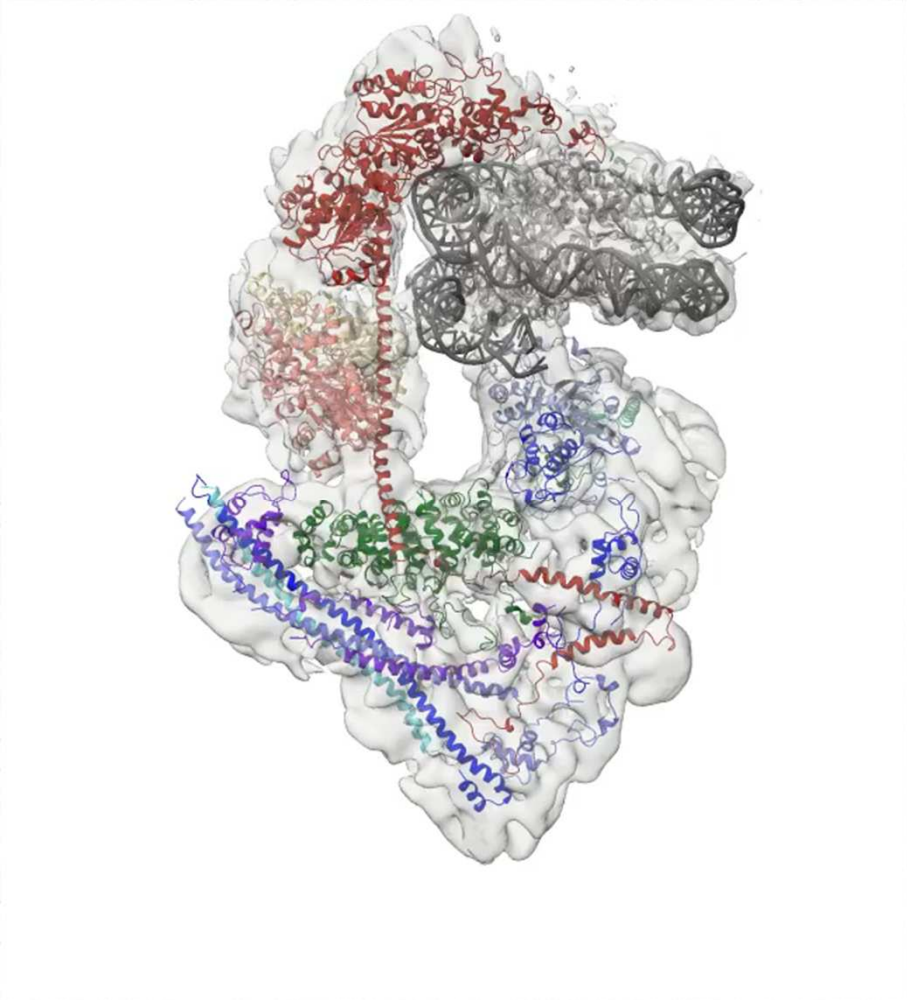
remodelační komplex

sběr velkého počtu částic, klasifikace a rekonstrukce struktury komplexu

Mashtalir et al, Cell, 2020



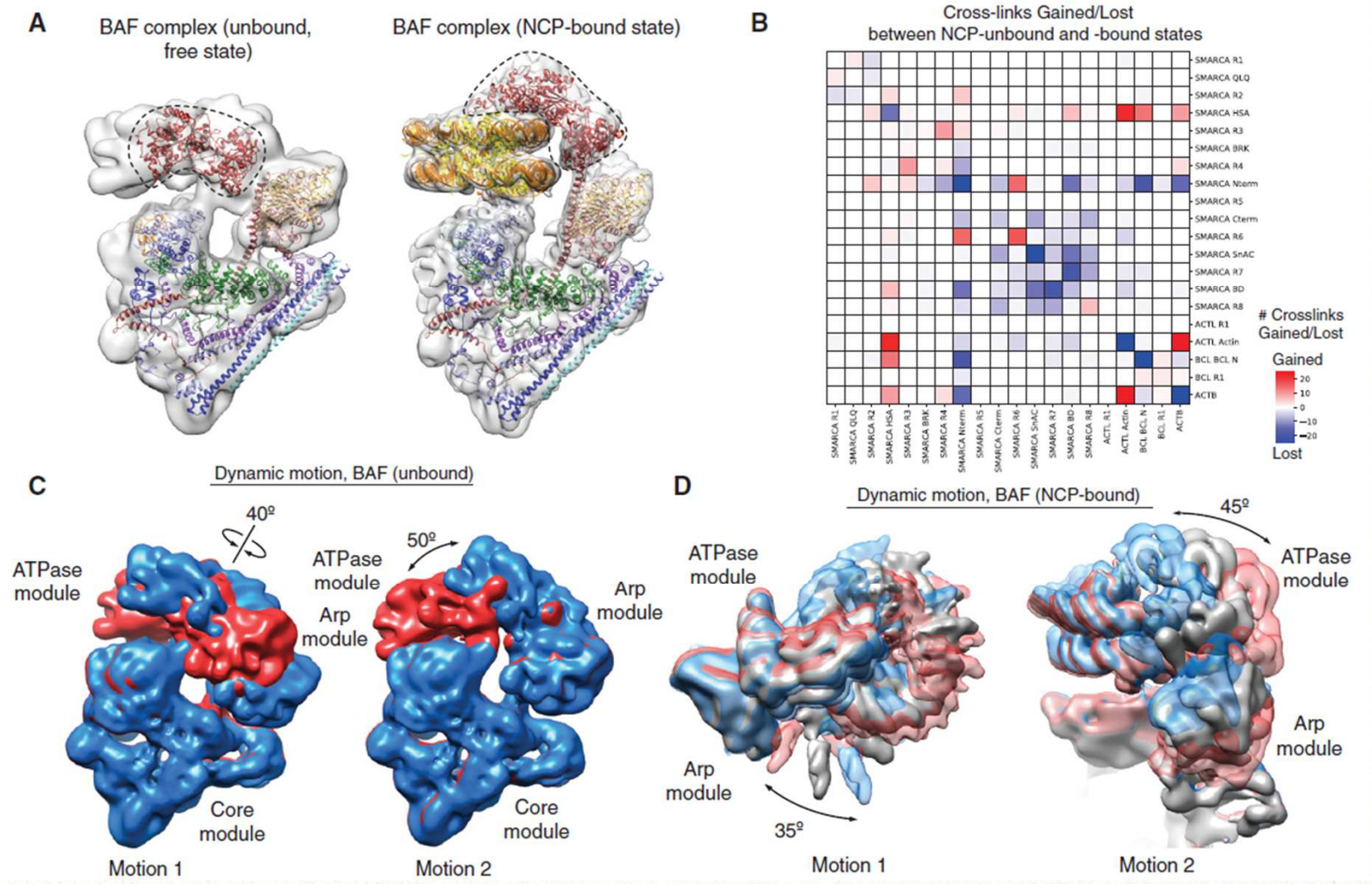




remodelační komplex

sběr velkého počtu částic, klasifikace a rekonstrukce struktury komplexu

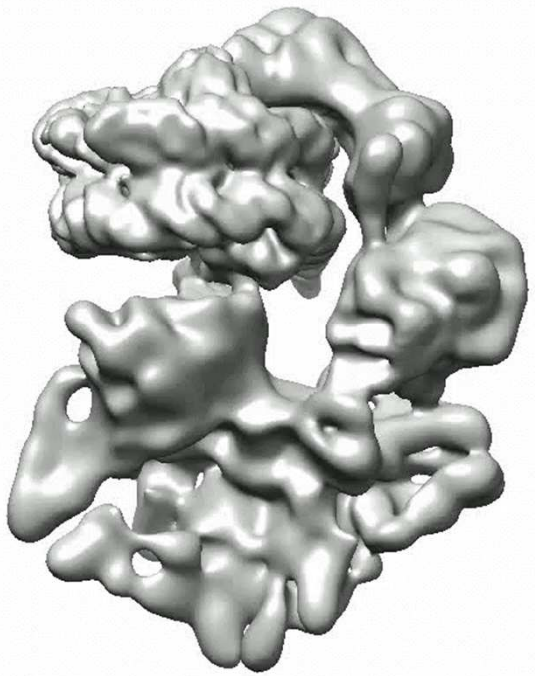
Mashtalir et al, Cell, 2020

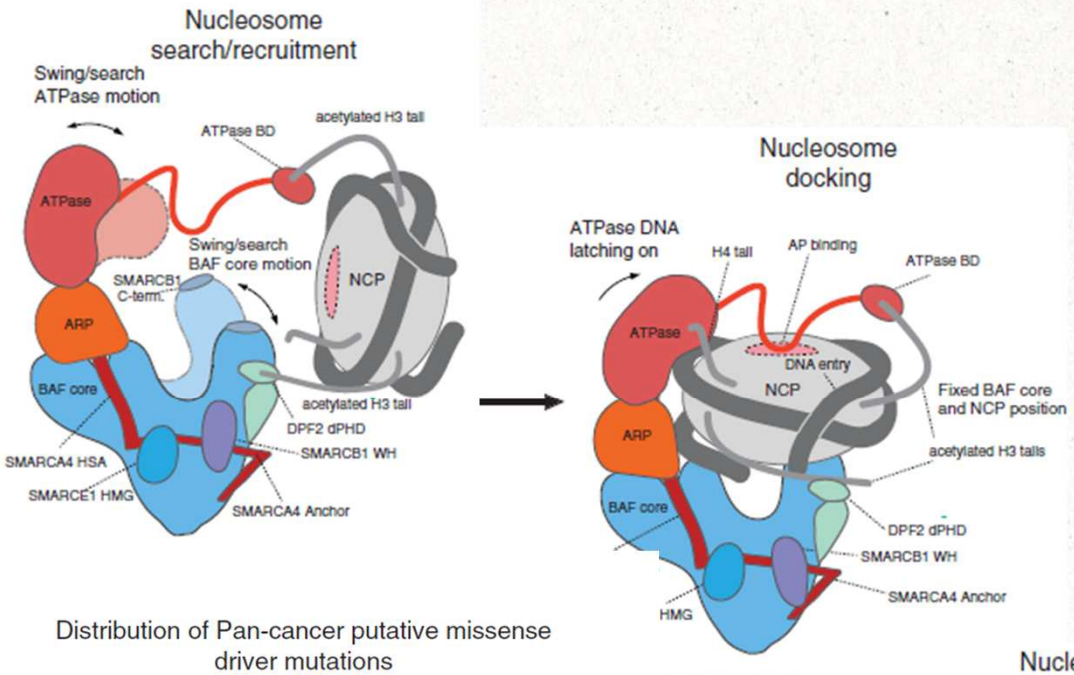


remodelační komplex

zachyceno několik konformací + rozdíly v krosslinku = dynamika

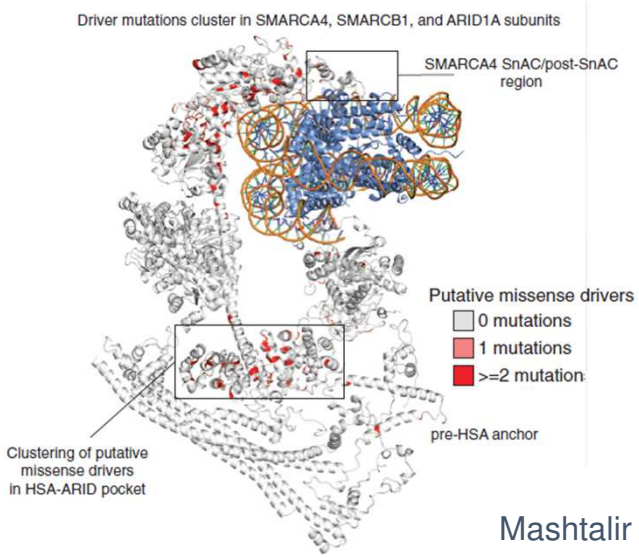
Mashtalir et al, Cell, 2020

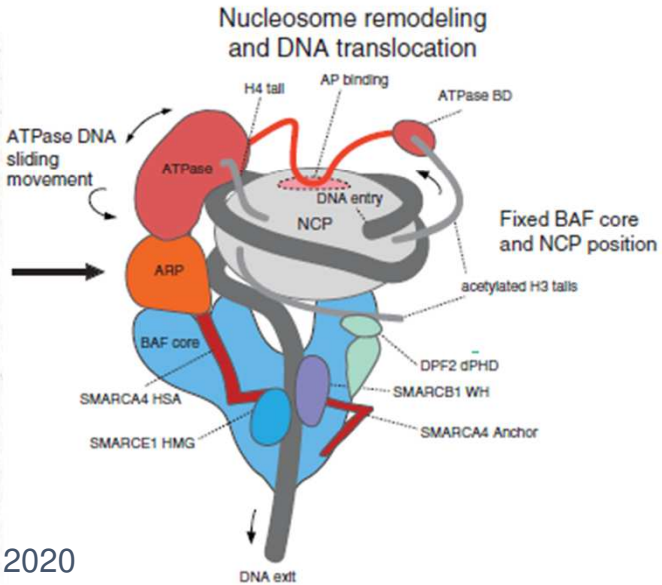




remodelační komplex

pomáhá v transaktivaci genové exprese – váže acetylované H3 (bromodomény) - uvolňuje nukleosom (volnější konce DNA)

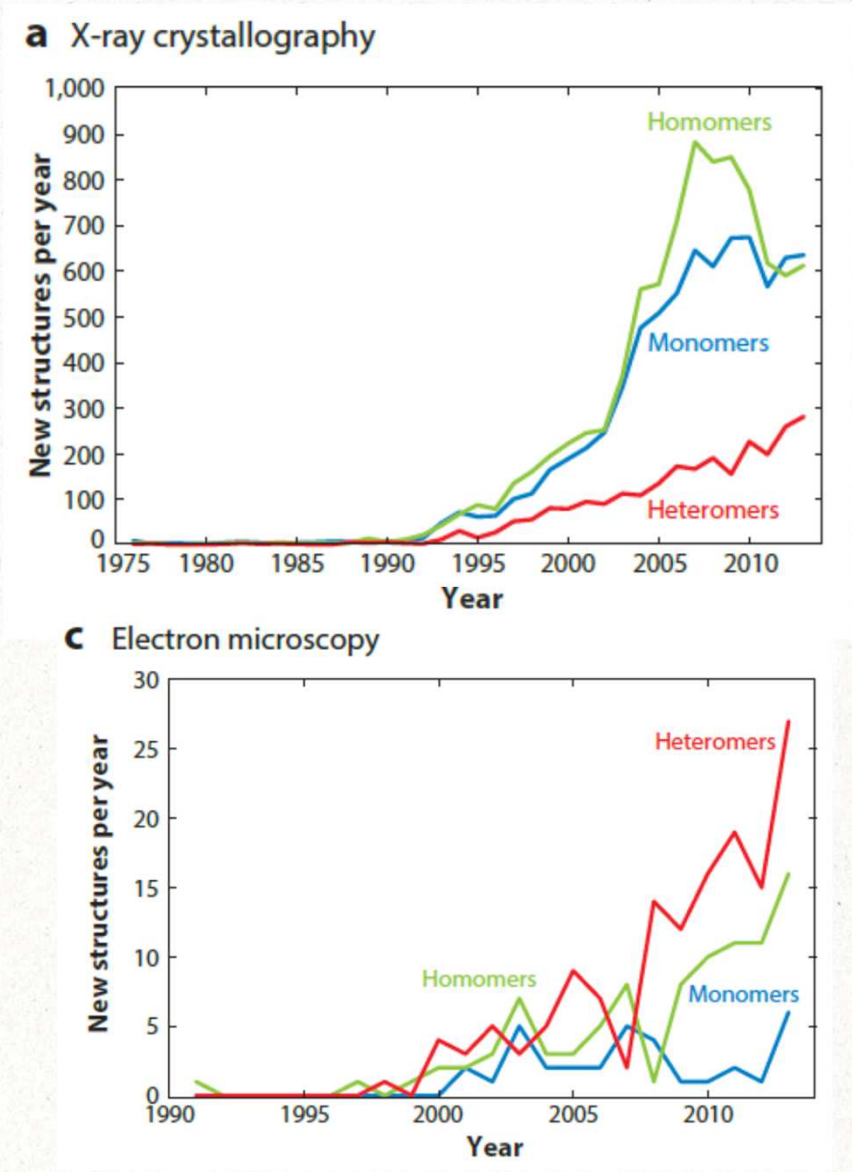




Mashtalir et al, Cell, 2020

Strukturní metody

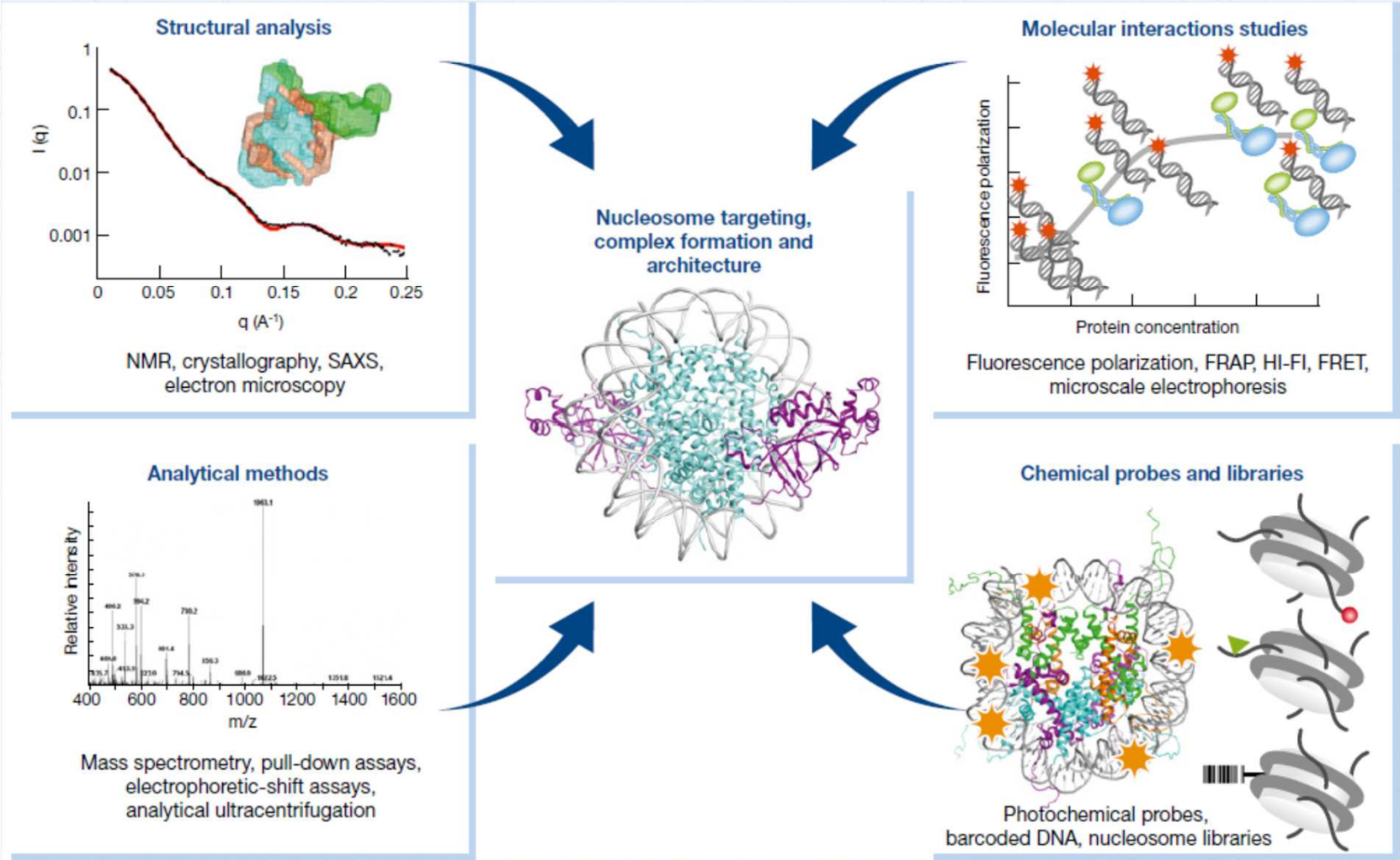
- krystalografie – nejvhodnější (boom v 1. dekádě díky sekvenačním projektům)
- NMR je limitována velikostí
- cryoEM je vhodná pro velké komplexy (boom v současnosti)



Marsh et al, ARB, 2015

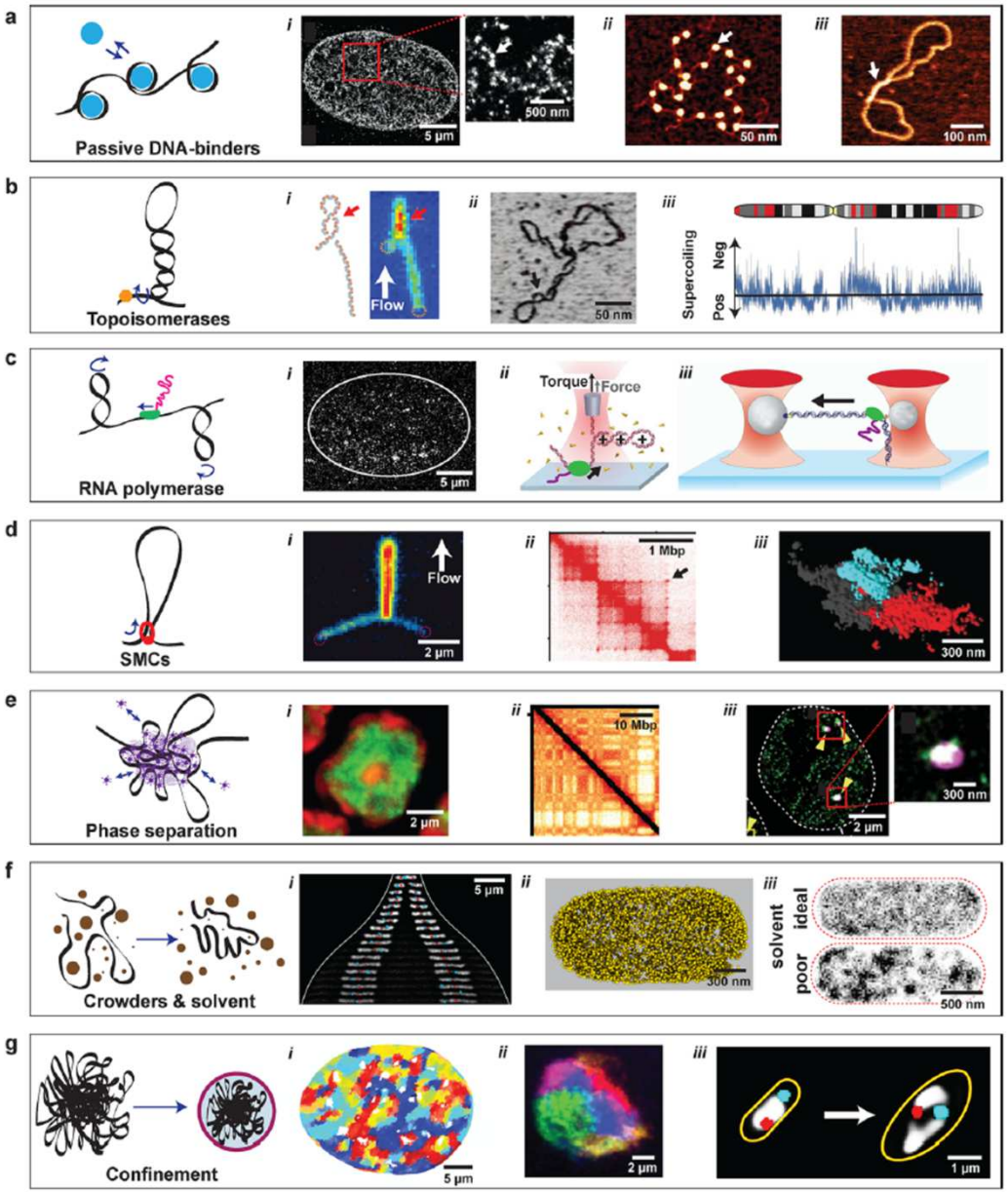
Analýza proteinových komplexů

více doc. Hofr

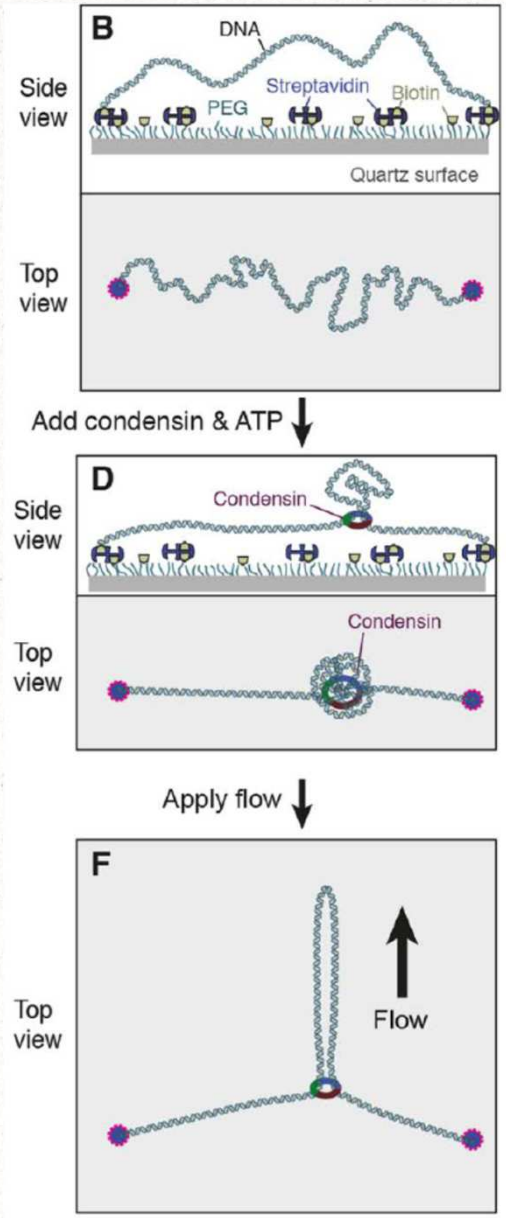


Speranzini et al, EMBO J, 2016

… speciální biofyzikální metody …

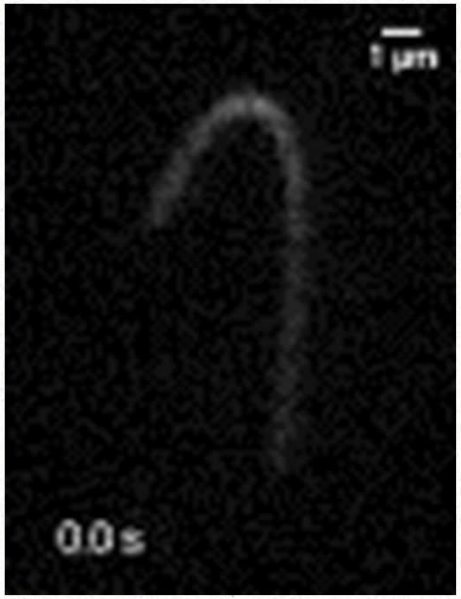


Birnie et al, ACS nano, 2020

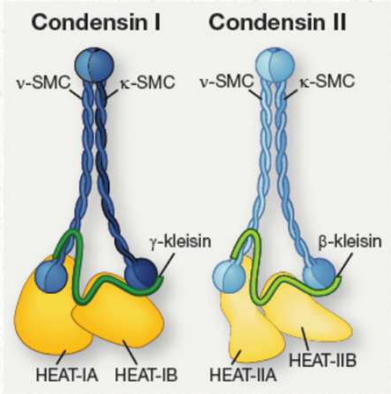


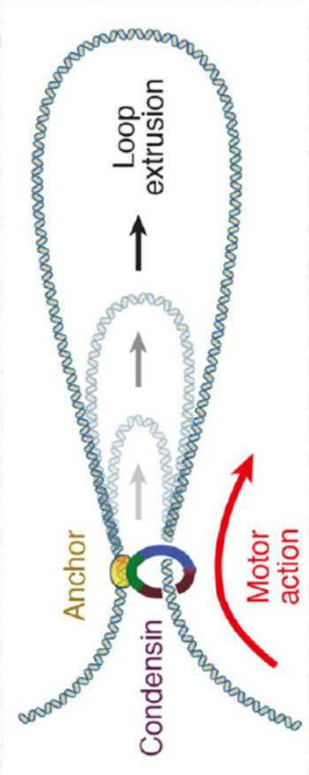
konce molekuly DNA uchyceny

přidán SMC komplex + vazba a hydrolýza ATP – vytvoření smyčky („loop extrusion)“

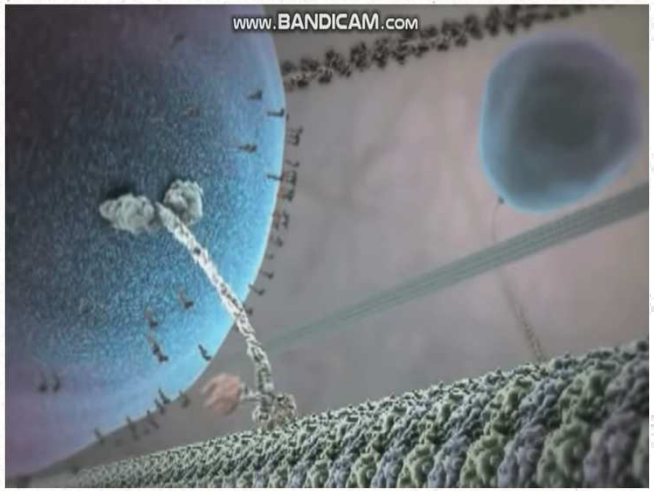


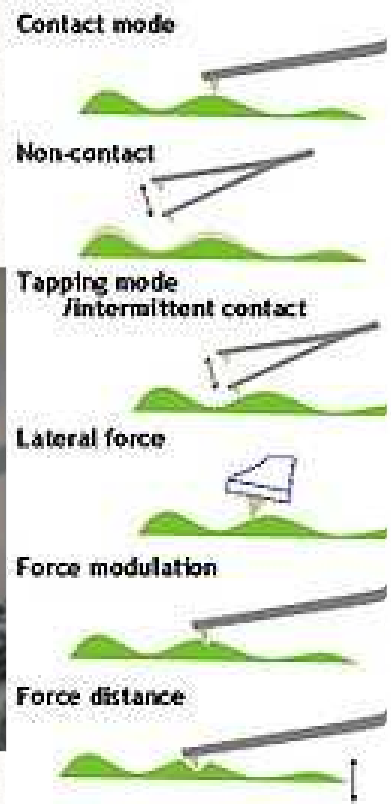
Ganji et al, Science, 2018
Takahashi, CO in CB, 2019

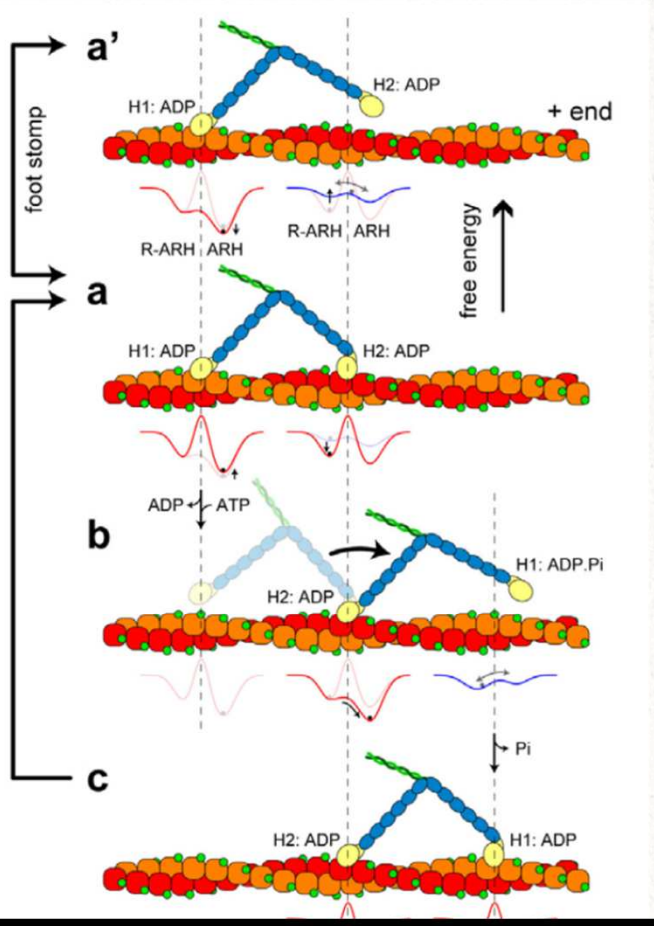




AFM
(atomic force microscopy)

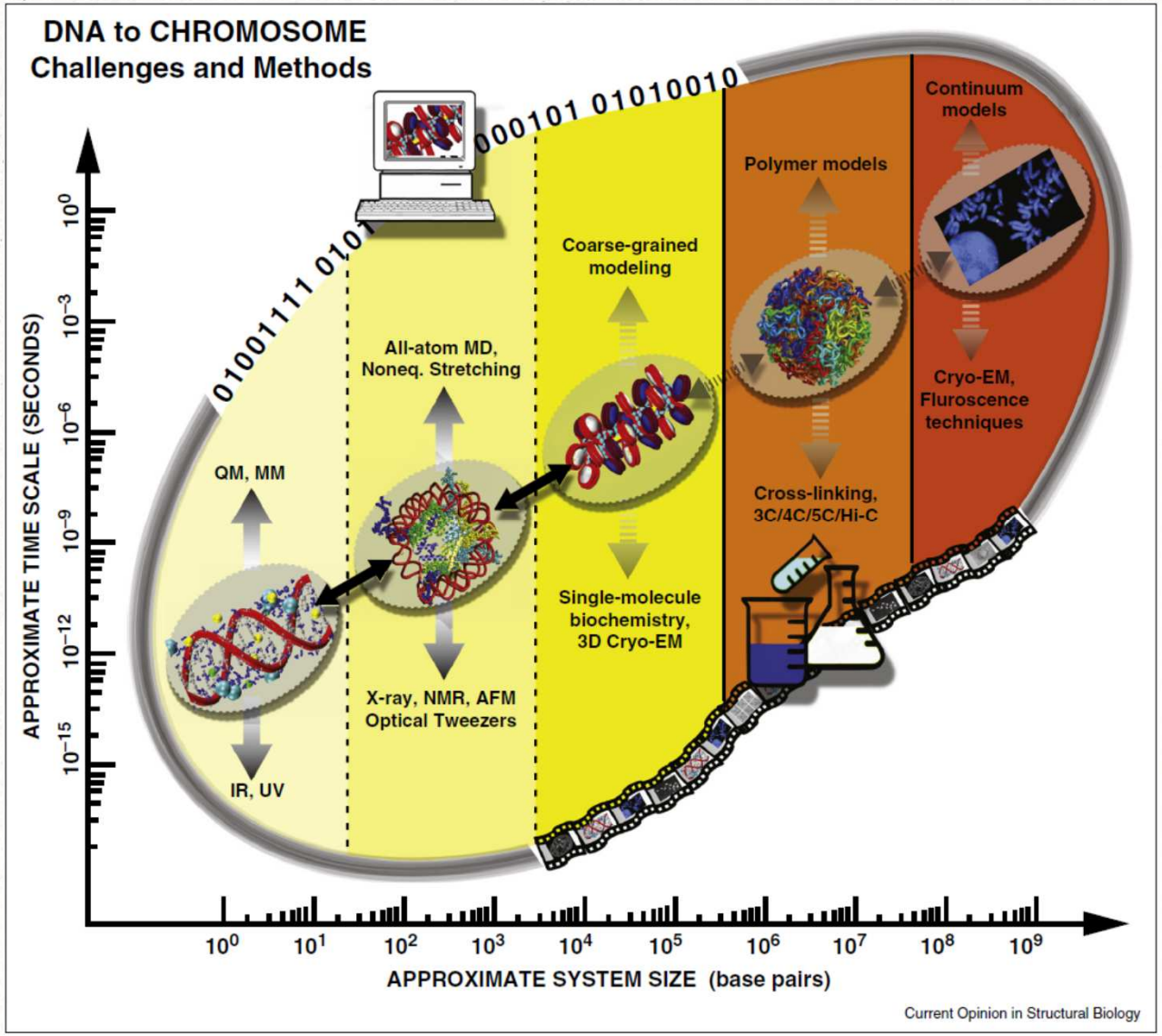






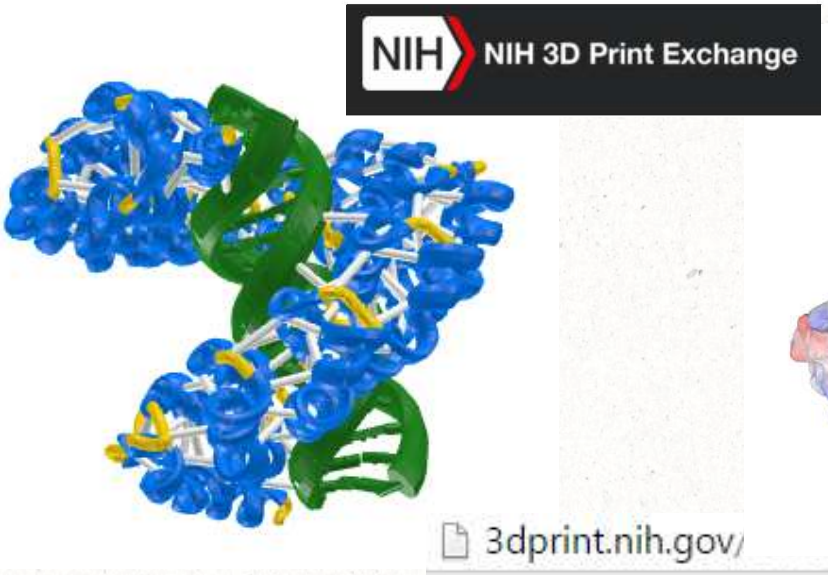
Ando et al, Annu Rev BioPhys, 2013

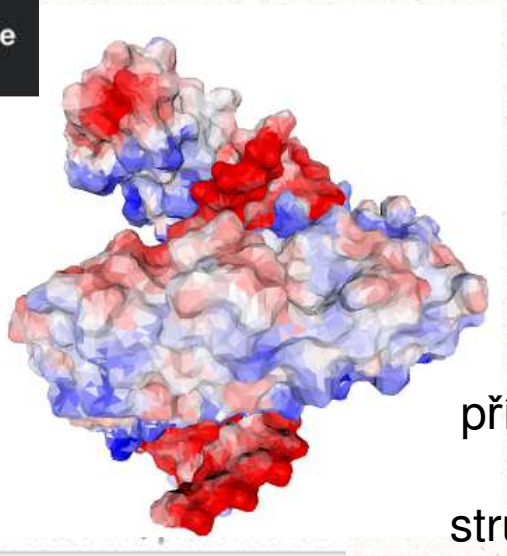
úrovně proteinových komplexů



Ozer et al, CO in SB, 2015

Visualizace proteinových komplexů





Existuje mnoho nástrojů na visualizaci komplexů

od **PyMOL** pro přímou visualizaci krystalových struktur … **3D tisk**