

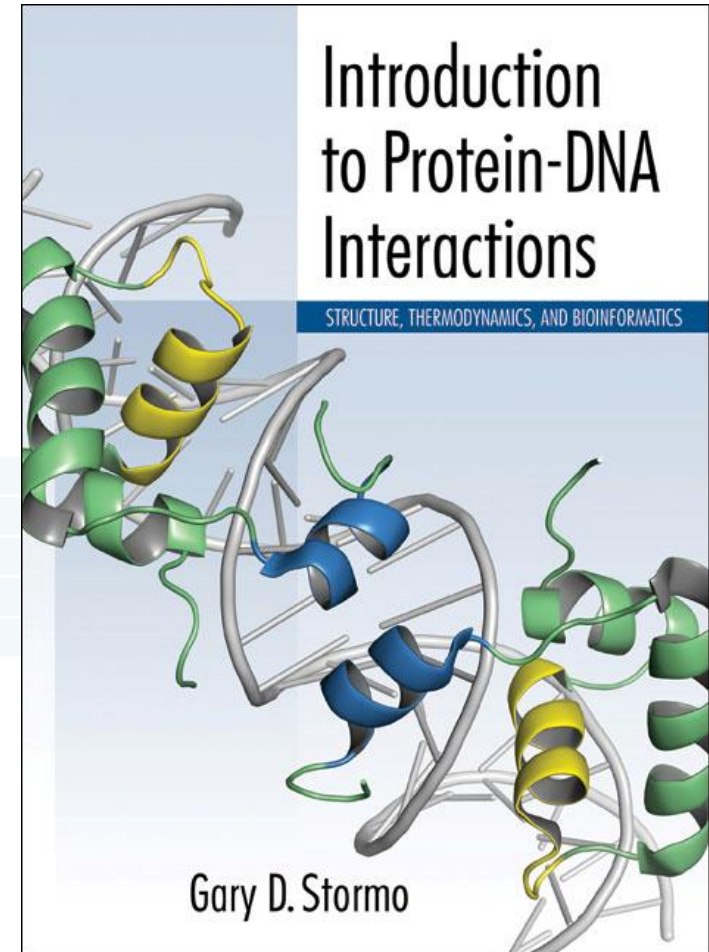
# DNA-proteinové komplexy

Komplexy spojené s transkripcí (až 5% genomu)

Komplexy podílející se na opravě DNA

Chromatinové strukturní komplexy

doc. Paleček	DNA-proteinové interakce, vazebné motivy
doc. Paleček	DNA-proteinové interakce, transkripční komplexy
Dr. Blažek	Cyclin/CDK komplexy v buněčném cyklu a transkripci
Dr. Špirek	Oprava DNA, homologní rekombinace
doc. Paleček	Chromatinové komplexy



Bi7015 - Chemické vlastnosti, struktura a interakce nukleových kyselin (doc. Fojta)

# DNA-proteinové komplexy

Komplexy spojené s transkripcí (až 5% genomu)

Komplexy podílející se na opravě DNA

Chromatinové strukturní komplexy

2500 struktur v PDB (v roce 2014)

**NPIDB** # Home

Database of structures of nucleic acid - protein complexes

Home Browse Download Help About Us

**Search**

Search | PDB:

Search

PDB  Pfam  SCOP  GO terms  Fuzznuc  BLAST

**NPIDB**

The resource NPIDB (Nucleic acid – Protein Interaction DataBase) includes a collection of files in the PDB format containing structural information on DNA-protein and RNA-protein complexes, and a number of online tools for analysis of the complexes. The tools are: an original program CluD for analysis of hydrophobic clusters on interfaces, program for detecting potential hydrogen bonds and water bridges, visualization of structures with [Jmol](#). Information on SCOP and Pfam domains detected in protein chains is presented.

**Reference:** Kirsanov et al. [NPIDB: nucleic acid–protein interaction database](#). Nucleic Acid Research, Volume 41 Issue D1 D517-D523 (January 2013)

**List of complexes** □

Structures of protein-nucleic acid complexes are extracted from Protein Data Bank (PDB) as files in the PDB format.

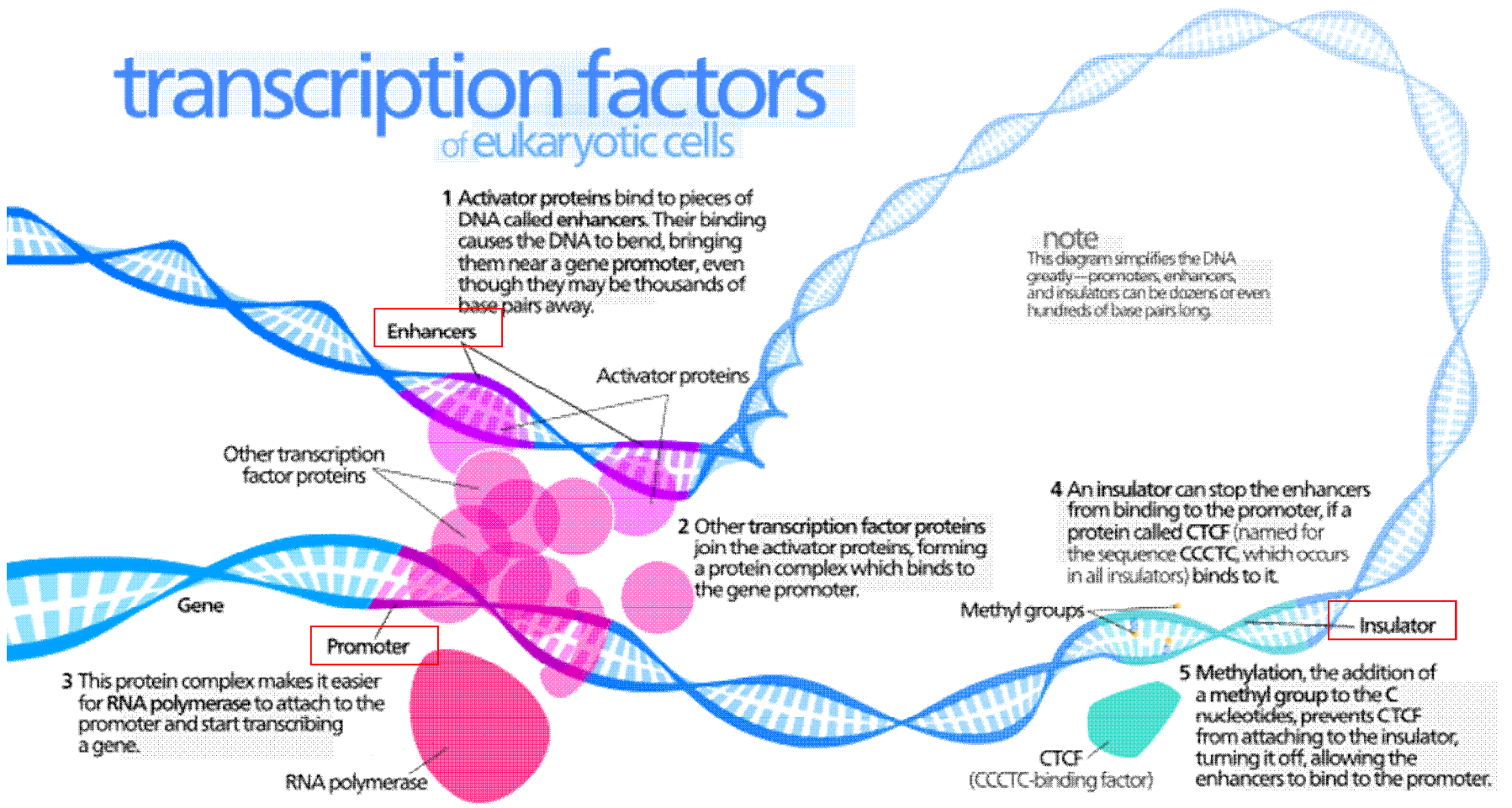
As of **27.11.2013** there are **4482** structures.

Each individual complex has its own web page, containing general information, links to other resources (e.g., PDBsum and BIPA), a table describing biological units or models, tables describing Pfam and SCOP domains in protein chains, and the list of available actions (including Jmol visualization)

<http://npidb.belozersky.msu.ru/>

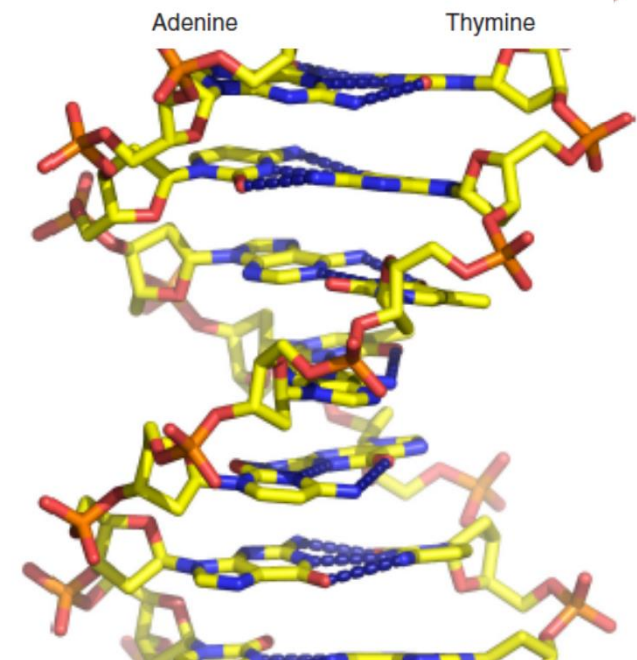
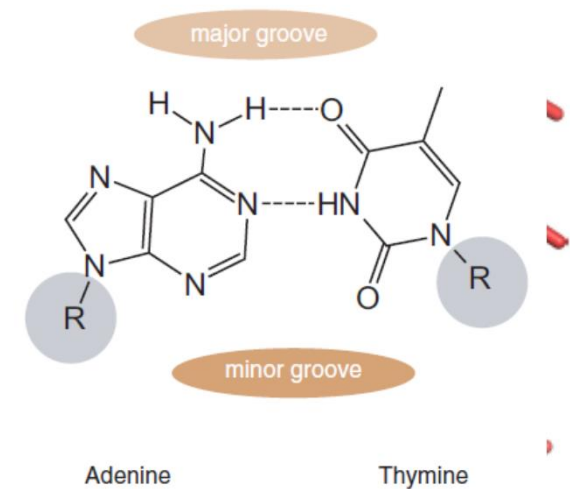
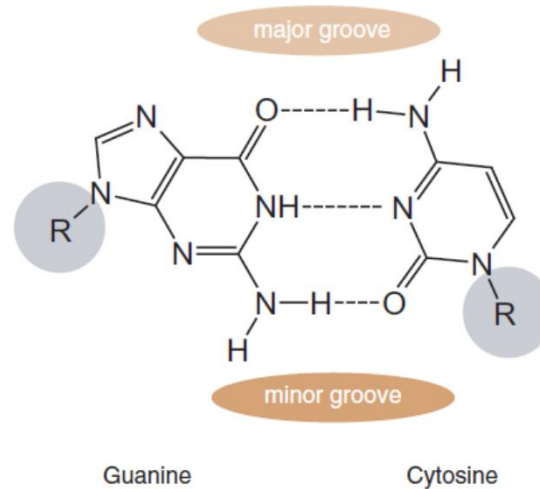
# Komplexy spojené s transkripcí

DNA-vazebné motívy špecifických transkripčných faktorov (enhanceosom)  
Obecné TFII komplexy a proces transkripce



## Enhanceosom

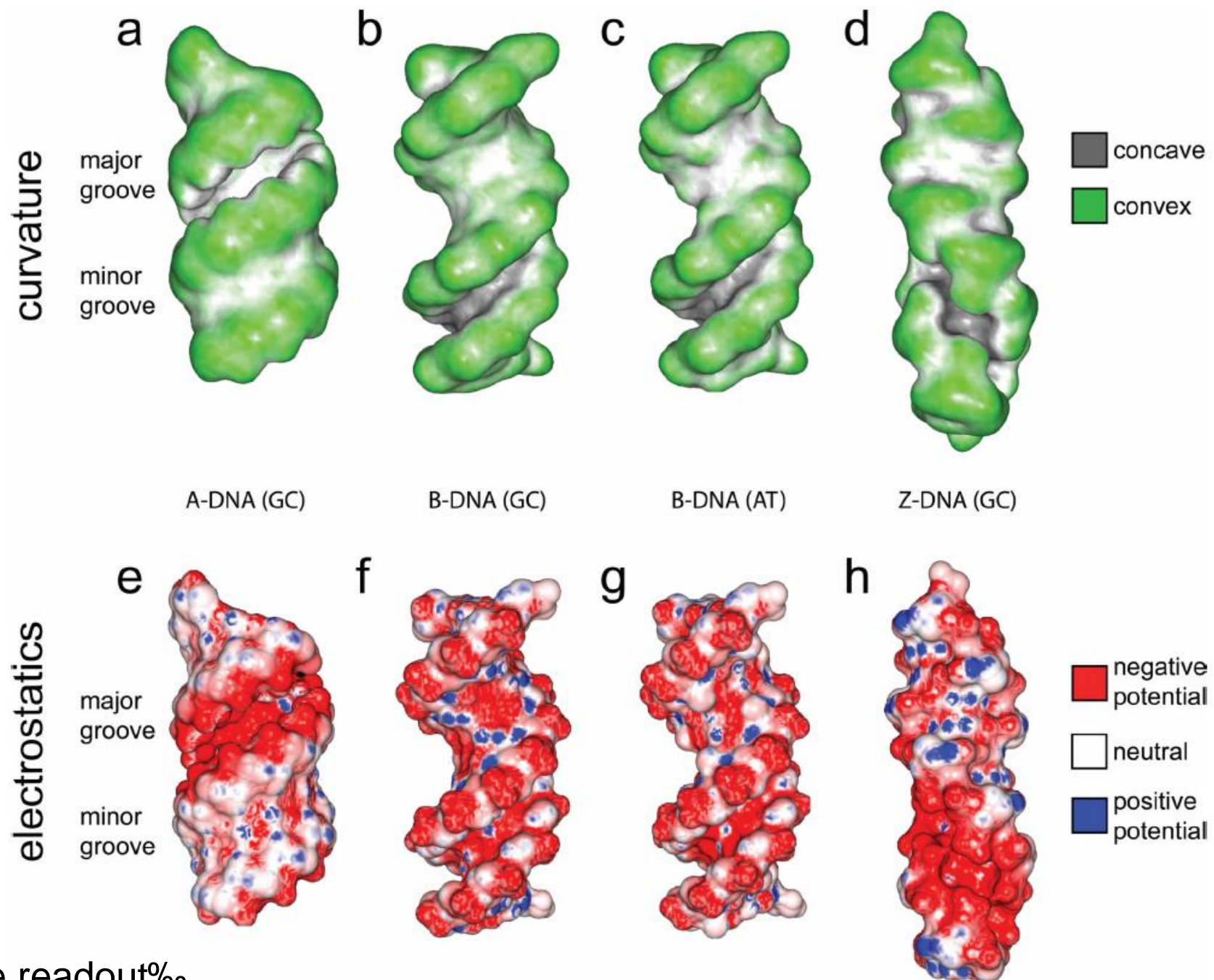
Tvarová a nábojová specifita DNA determinuje typy DNA-vazebných domén (oproti velké rozmanitosti protein-proteinových interakcí domén)



sshape readout%

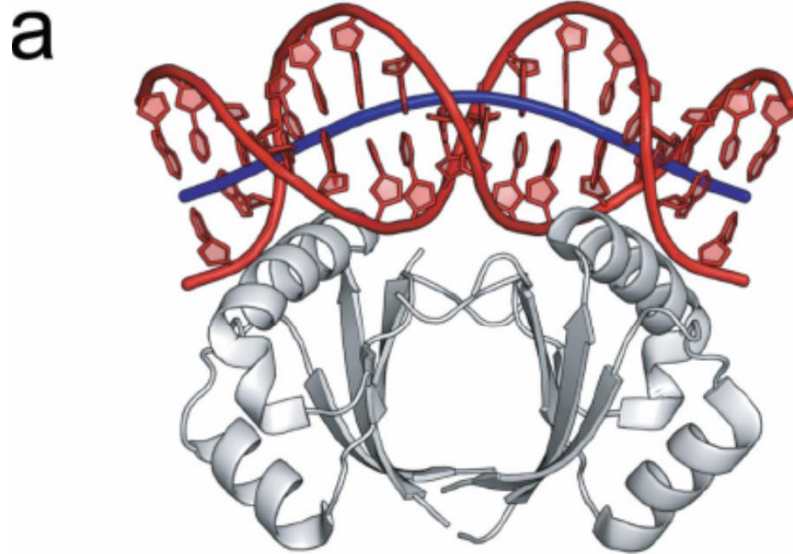
zak ivení kostry souvisí se sekvencí

- proteiny interagují s cukrfofátovou kostrou (fosfát) nebo přes žlábký s bazemi
- Interakce sekvenčně nespecifické (kostra . histony; strukturně specifické - HMG proteiny) nebo sekvenčně specifické (kostra+žlábký – kombinace: *Bgl*II (AGATCT) a *Bam*HI (GGATCC) kontaktují stejné báze a s tou%zak ivení okolní DNA ò )



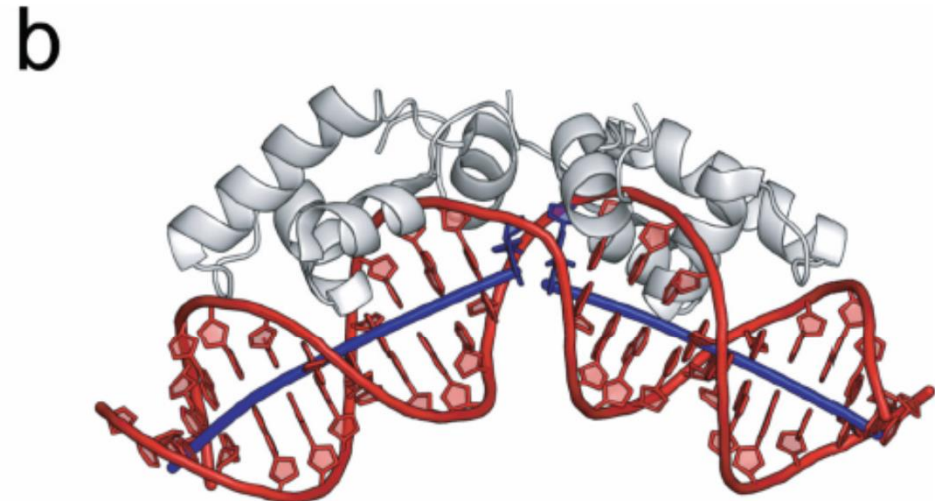
ssshape readout%  
 zak ivení kostry - souvisí se sekvencí a prost edím

# Vazba DNA-protein může indukovat změny



Bend

1jj4 (õ nap . histony)



Rohs et al, Annu Rev Bioch, 2010

Kink

2kei, Lac represor  
(Leu do malého Olábku)

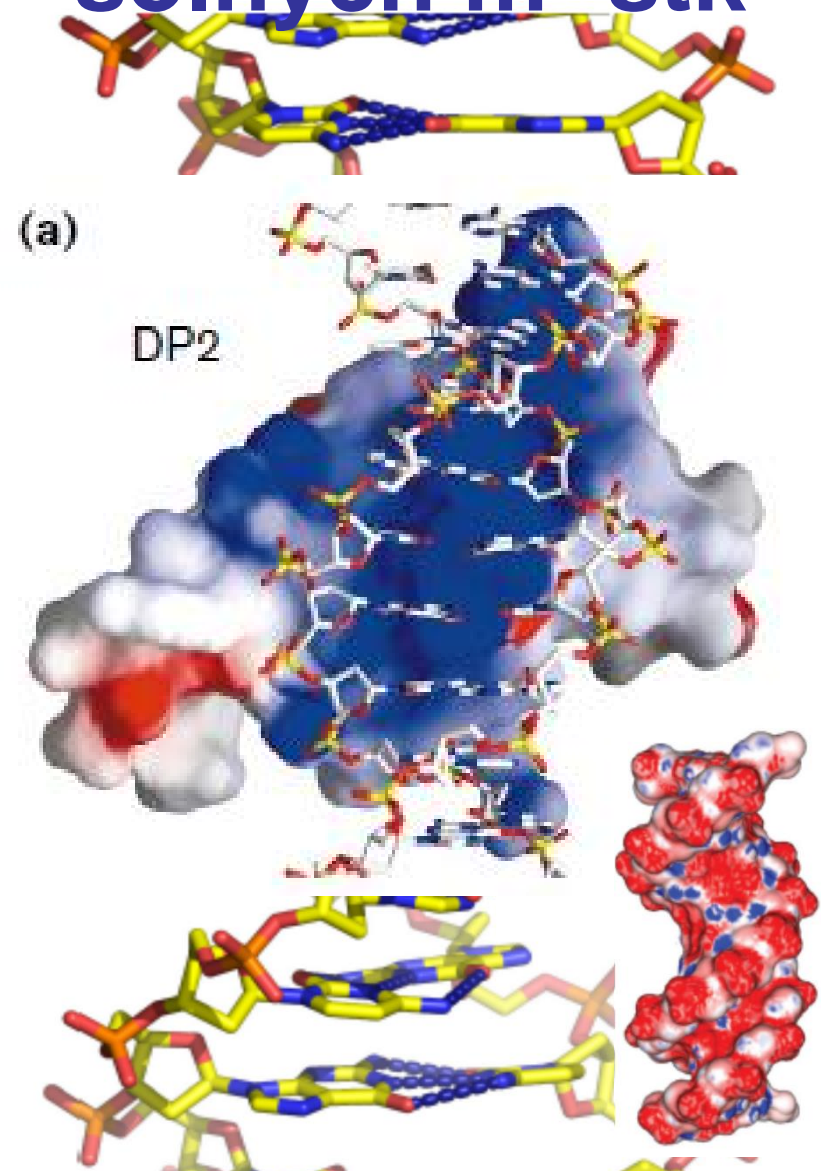
- vazba proteinu může indukovat změny ve struktuře DNA
- vazba DNA na protein často indukuje změny v jeho struktuře
- například u nestrukturovaných proteinů strukturu indukuje (c-Jun/c-Fos = zroubovice až po navázání dimeru na DNA)

# Vazba protein s DNA prost ednictvím solných můstek

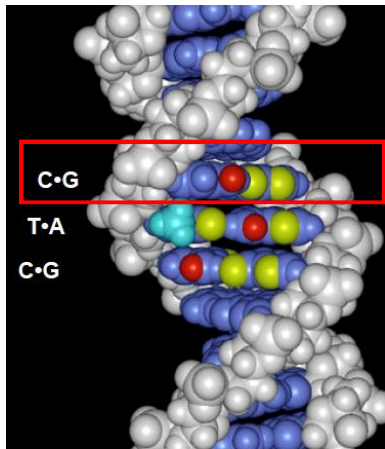
- fosfáty mohou interagovat s Arg a Lys . solné můstky/salt bridges (pozitivní náboje Arg a Lys vytváří vazbu s negativním nábojem fosfátové skupiny)
- Elektrostatický náboj/povrch naznačuje vazebné schopnosti proteinu

Table 2. The simplified relative charge set defined from CHARMM (10) used in the calculation of the electrostatic potential of atoms in the DNA-binding proteins

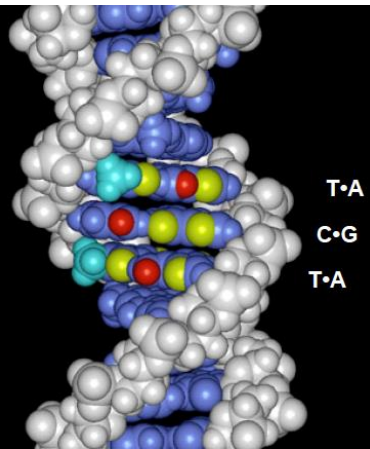
Atom type (PDB entry)	Residue	Relative charge
NZ	Lys	1.00
NH1	Arg	0.50
NH2	Arg	0.50
OE1	Glu	-0.50
OE2	Glu	-0.50
OD1	Asp	-0.50
OD2	Asp	-0.50
OXT	All residues	-1.00
N	All residues	-0.10
CA	All residues	0.10
C	All residues	0.55
O	All residues	-0.55



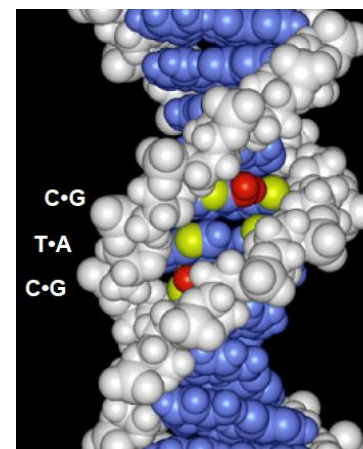
Major groove



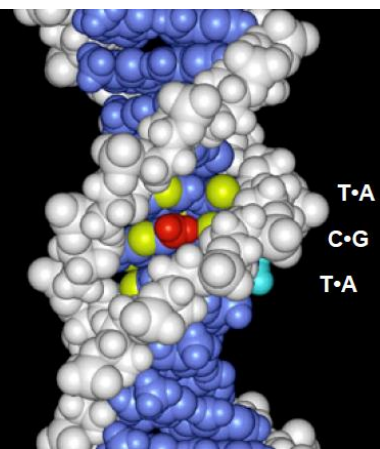
Major groove



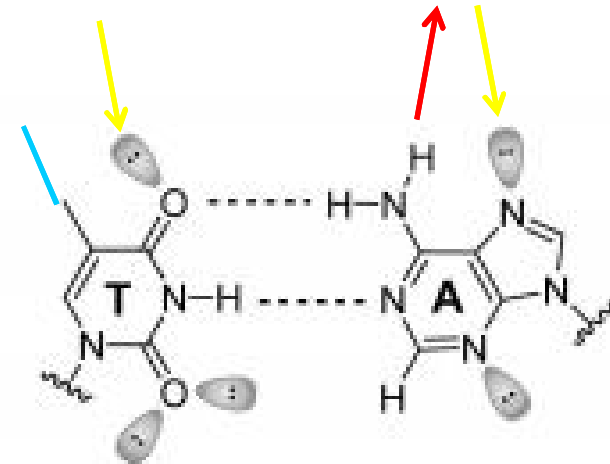
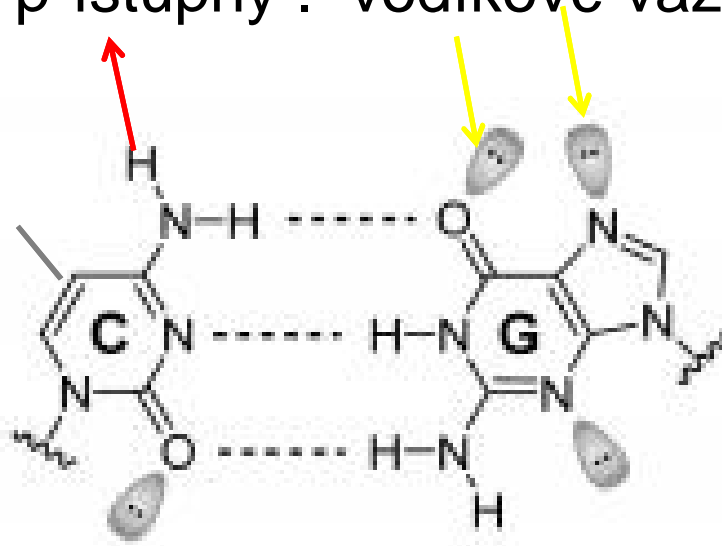
Minor groove



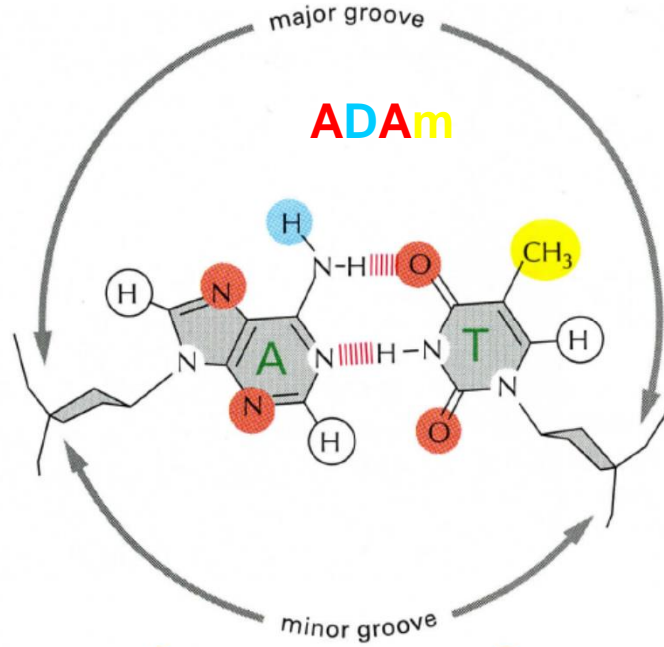
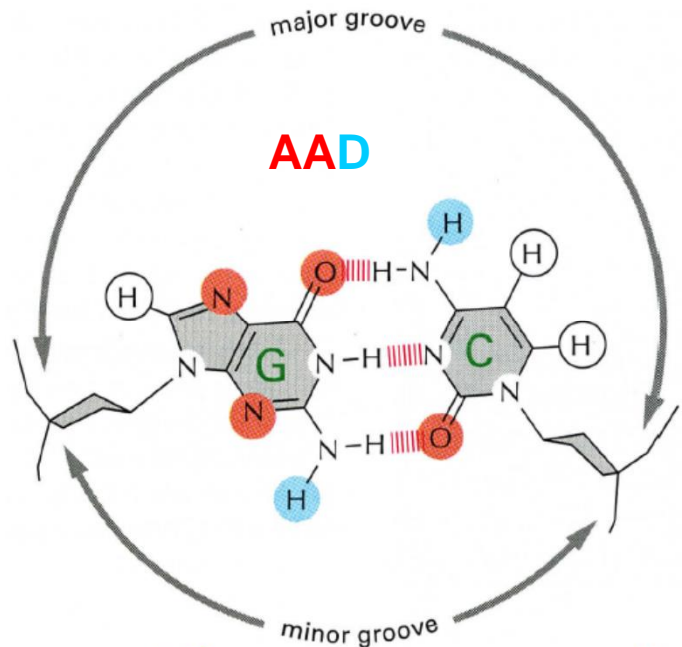
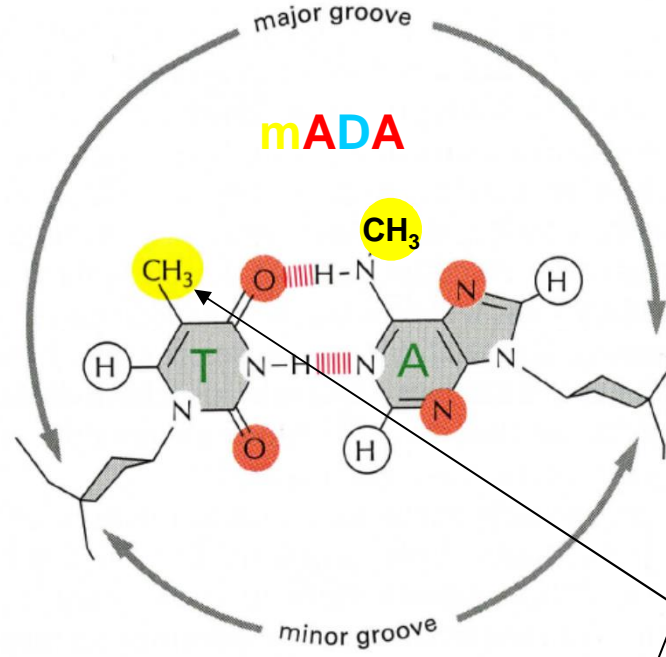
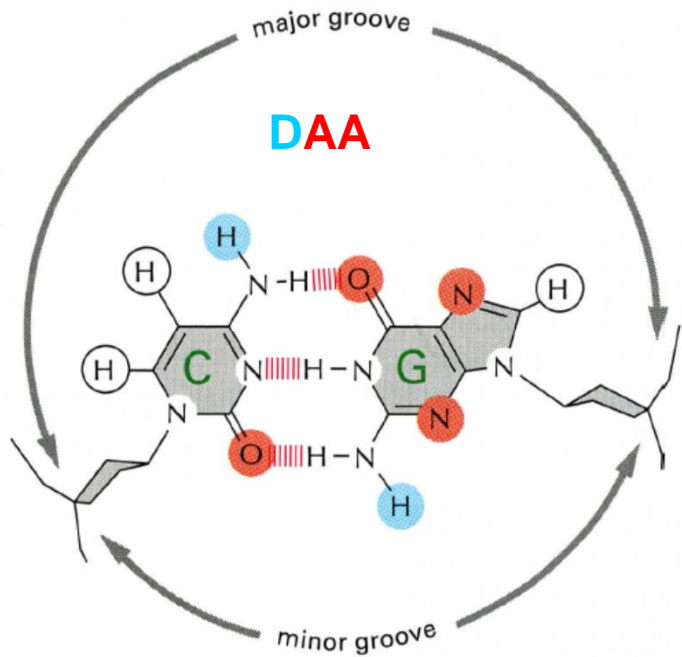
Minor groove



**sekven n** -specifický protein kontaktuje báze (direct readout) . skrze velký nebo malý Olábek . **velký ýlábek** je lépe p ístupný . vodíkové vazby (donor vs akceptor elektronu)











Jak odliší  
protein r zné  
páry bází?

**Í base readoutí**

Pozice donor vs  
akceptor  
+ metyl skupina

Metylace Ade  
(C6NH<sub>2</sub>) u bakterií  
zm na!

KEY:

-  = H-bond acceptor
-  = H-bond donor
-  = hydrogen atom
-  = methyl group

# Vazba proteinů s DNA prostřednictvím vodíkových vazeb

- Velký Olábek má velikost odpovídající rozměru  $\alpha$ -zrůbce a má exponované H-vazebné skupiny
- Ade zbytky C-6(NH<sub>2</sub>) a N-7 mohou tvořit specifické vodíkové vazby s **Gln** a **Asn**
- Guanosinové tvoří specifické vodíkové vazby s **Arg**

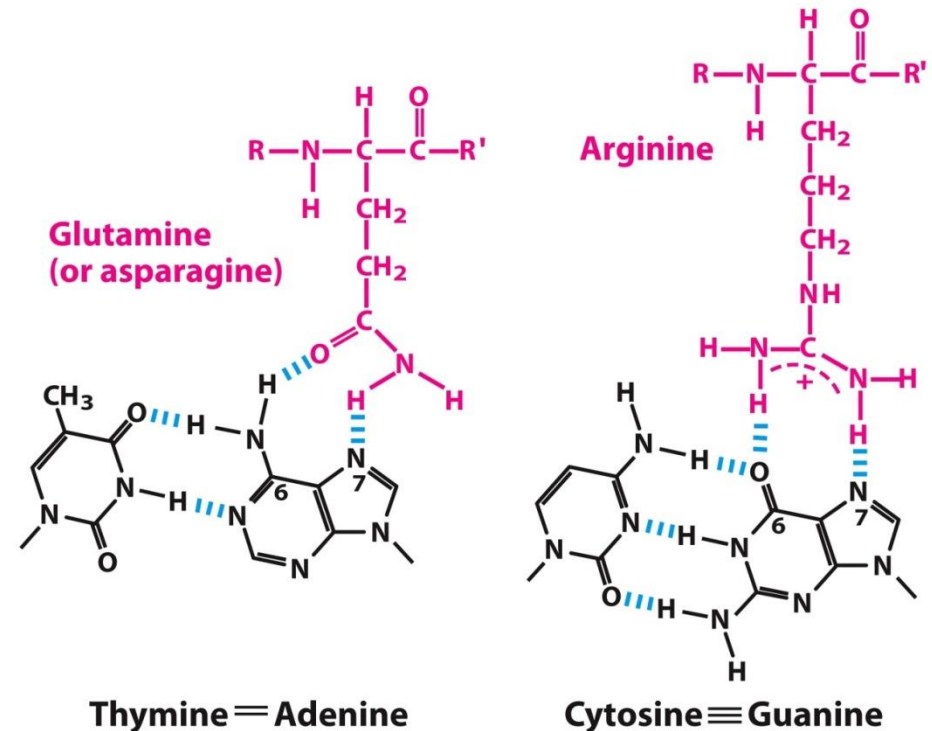
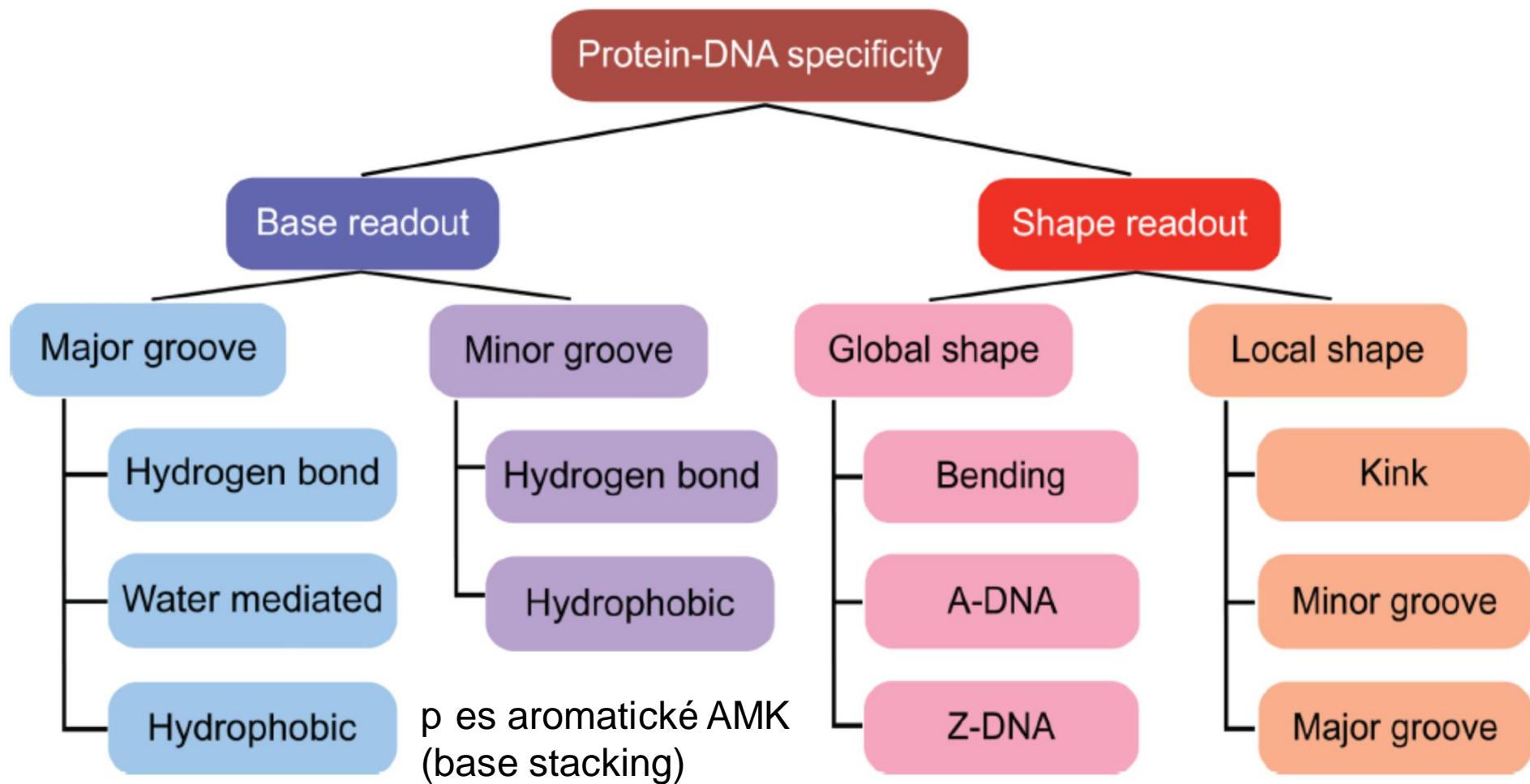


Figure 28-10  
Lehninger Principles of Biochemistry, Sixth Edition  
© 2013 W. H. Freeman and Company

Silná vazba, sekvence specifická - afinita nM .  $\mu$ M  
Slabá vazba, strukturně specifická - afinita  $\mu$ M . mM



- více jak 70 SCOP superrodin (strukturních motiv )
- dle sekundárních struktur .  $\alpha$ -zroubovice (17),  $\beta$ -listy (7), smízené  $\alpha/\beta$  motivy (48)

# Motivy DNA vazebných domén

## “ Zipper typ

- . Leucinový zip
- . Helix-loop-helix

## “ Helix-otá ka-helix

- . HTH
- . Winged helix
- . TALE

## “ Zinkový prst

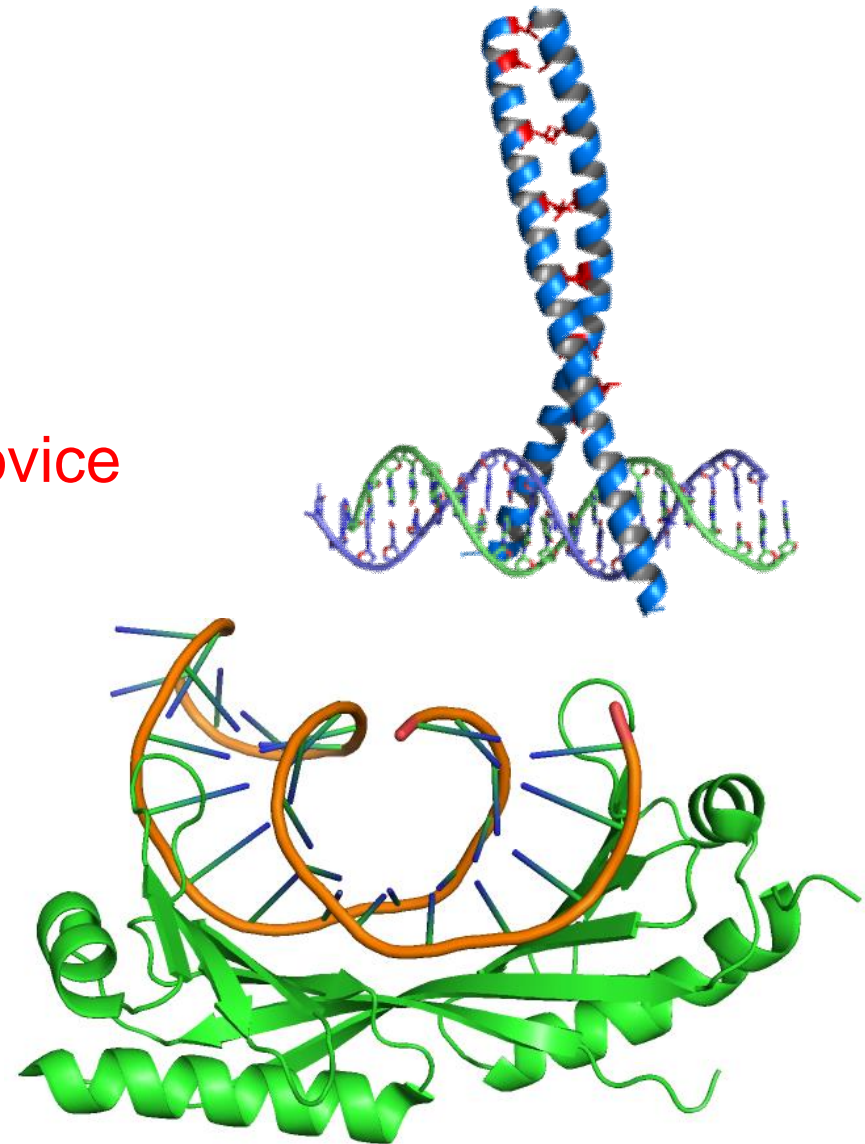
- .  $\beta\beta\alpha$  zinc-finger
- . Hormon-receptor
- . Loop-sheet-helix
- . Gal4

## “ Histon, HMG-box

## “ $\beta$ -sheet motivy

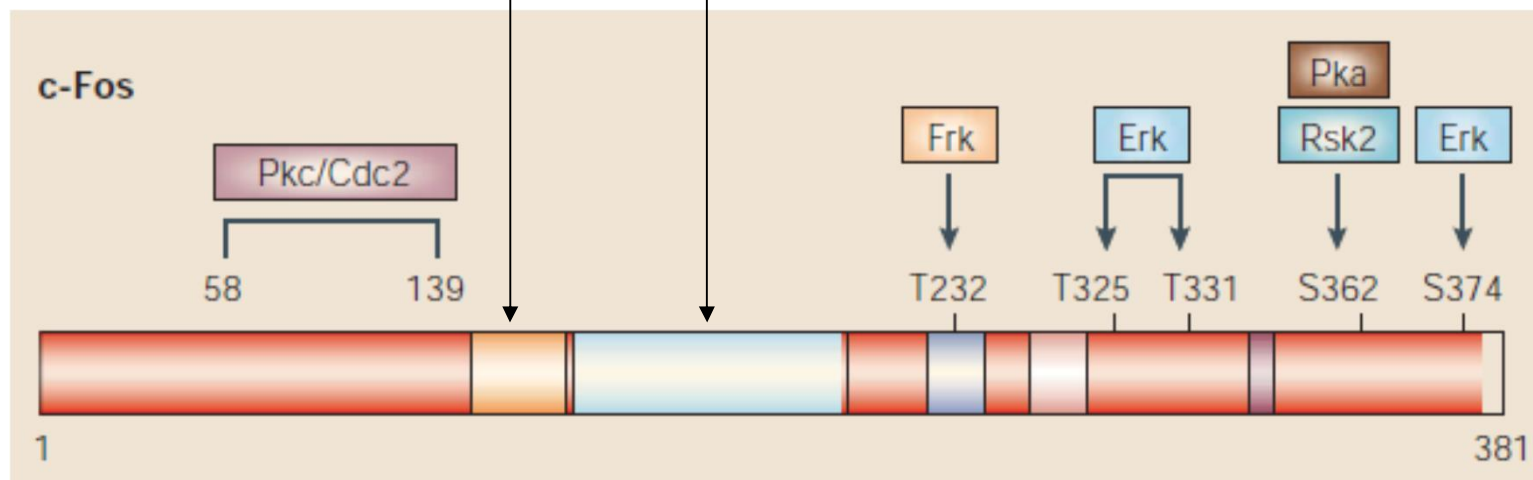
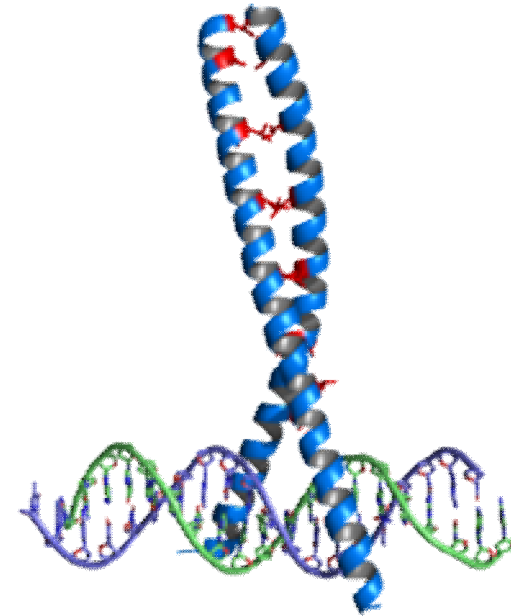
$\alpha$ -zroubovice

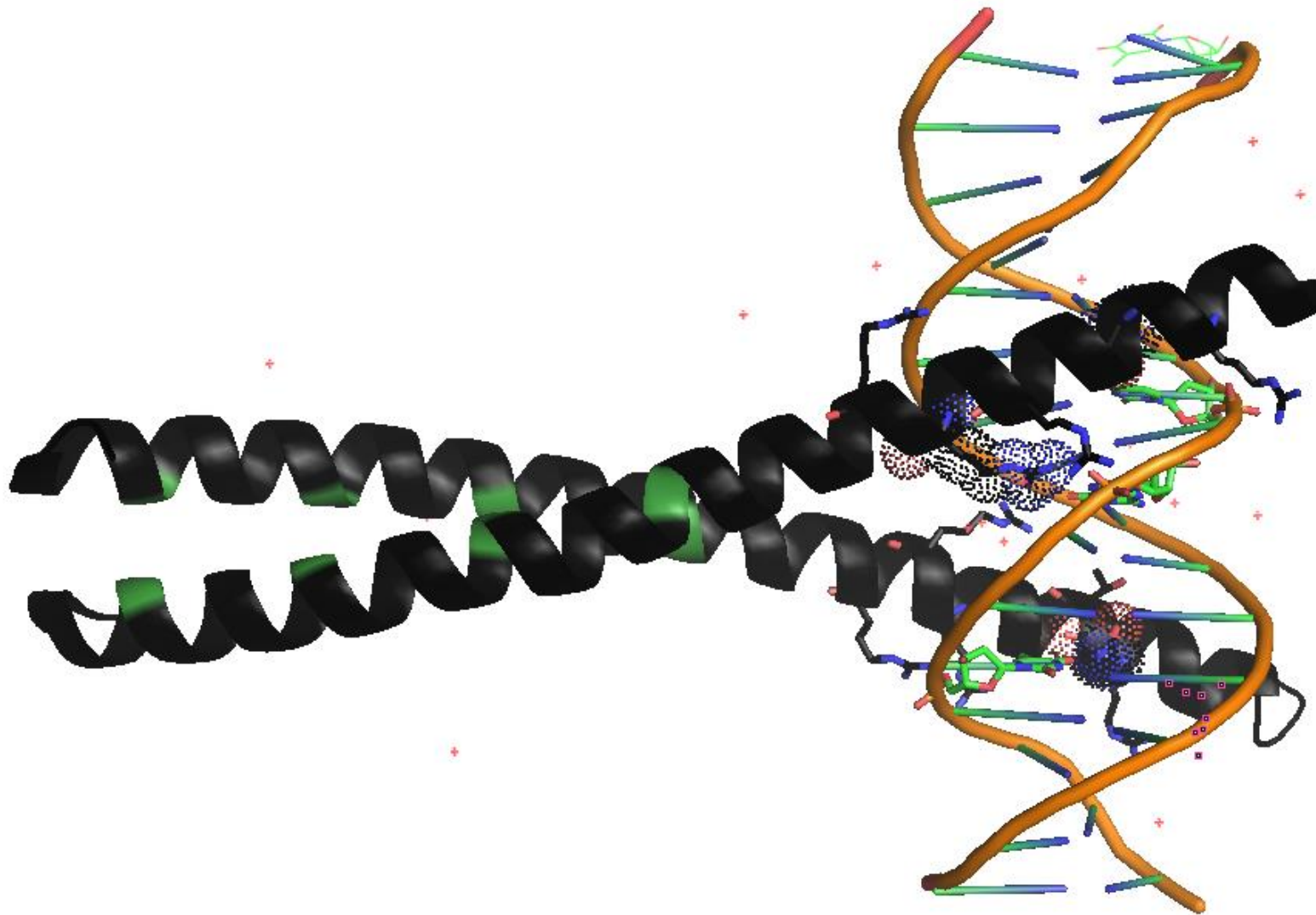
$\beta$ -listy



# Motivy DNA vazebných domén

- “ **Zipper typ** (dle způsobu dimerizace)
  - . **Leucinový zip** (tzv. bZIP = basic)  
(transcr. fact.  $\gamma$ GCN4, c-Jun/c-Fos=AP-1)
    - “ 2  $\alpha$ -helixy (2 x 60 AMK)
    - “ coiled-coil (>30AMK, Leu, C-term)
    - “ bazická část (N-terminus, navazuje na CC)
    - “ bazická zroubovice vázána do VŽ

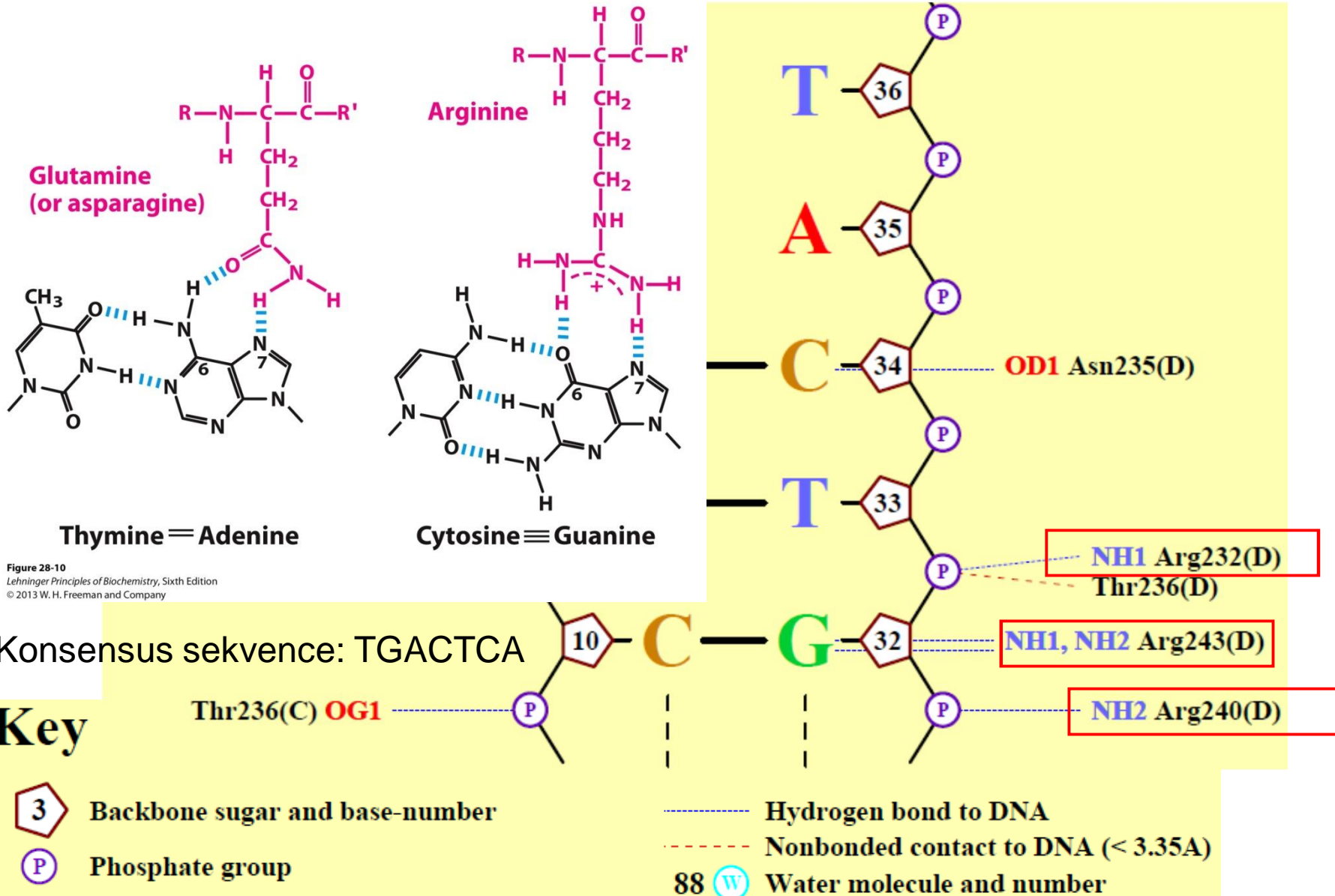




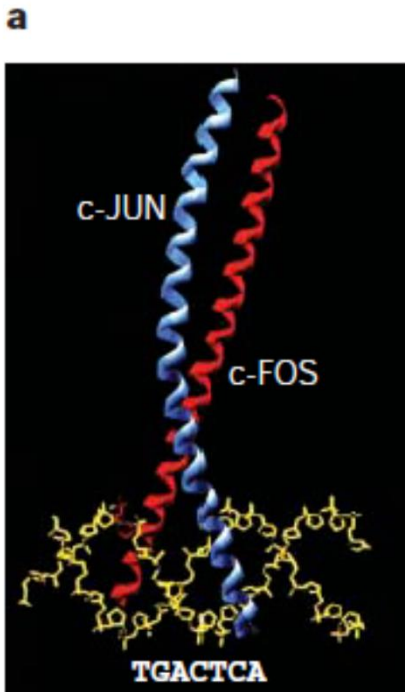
Interakce bazických AMK: Arg(232+240)=PO<sub>4</sub>, Arg(243)=Gua

Konsensus sekvence: TGACTCA

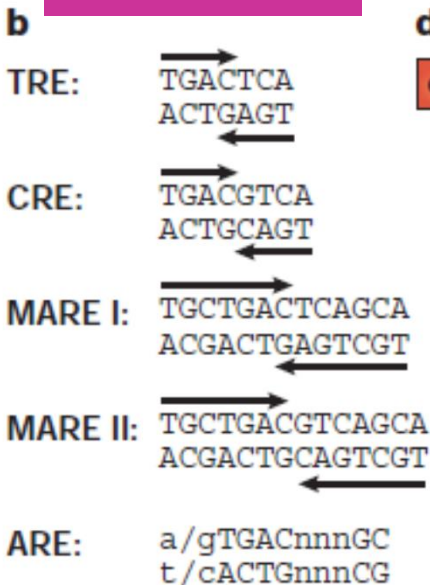
GCN4 . regulace gen pro syntézu AMK



Jones a spol., NAR, 2003



**AP-1**



**c** homo/hetero

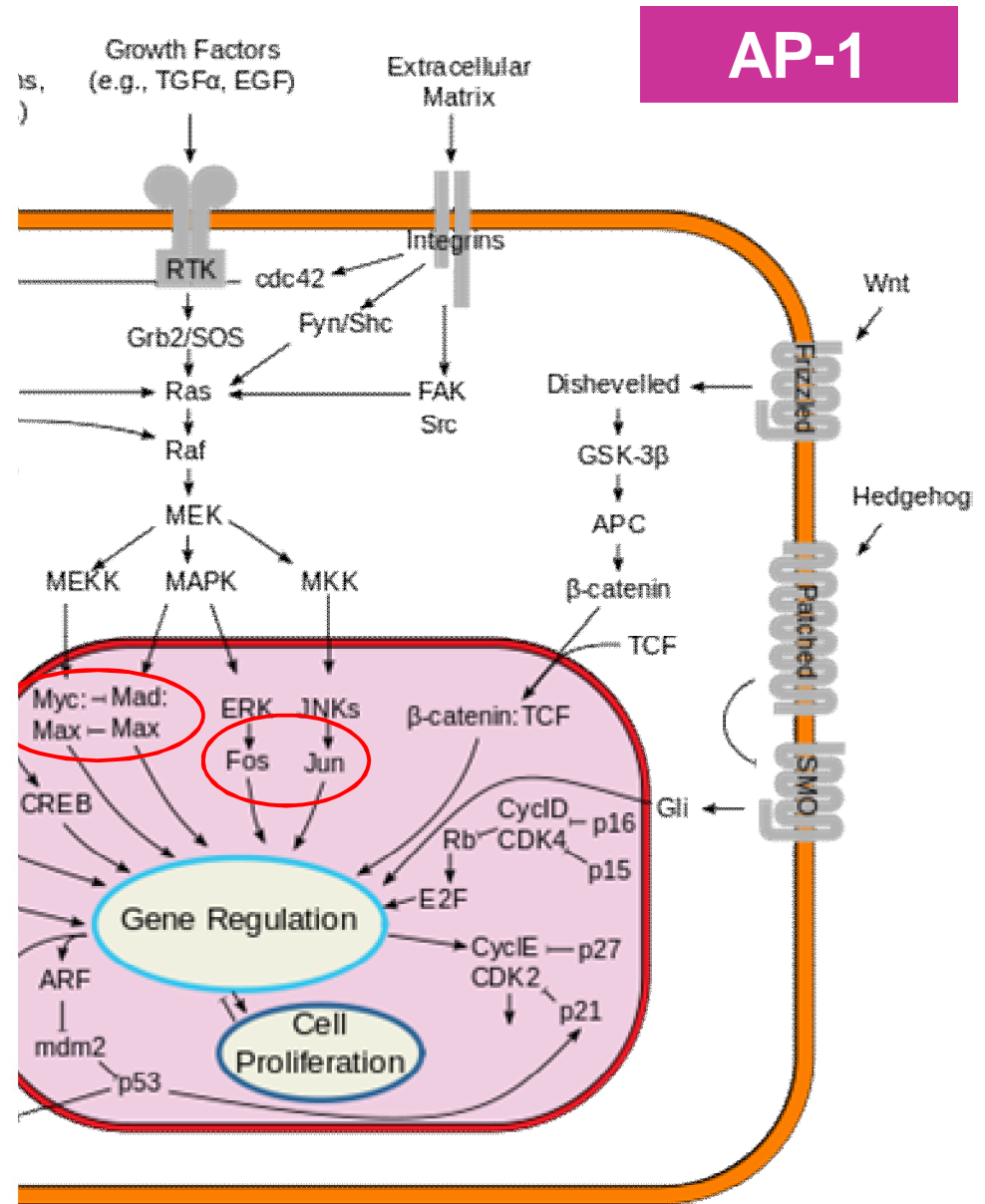
<b>c-JUN</b>	c-JUN	(TRE > CRE)
	JUNB	(TRE > CRE)
	JUND	(TRE > CRE)
	FOSB	(TRE > CRE)
	FRA1	(TRE > CRE)
	FRA2	(TRE > CRE)
	ATFa	(TRE = CRE)
	ATF2	(CRE > TRE)
	ATF3	(CRE > TRE)
	ATF4	(CRE)
	B-ATF	(TRE > CRE)
	c-MAF	(MARE I/II)
	MAFA	(MARE I/II)
	NRL	(TRE-related)
	MAFF/G/K	(MARE I/II)
	NRF1	(ARE)
	NRF2	(ARE)
	NFIL-6	(TRE)

**d** heterodimery

<b>c-FOS</b>	c-JUN	(TRE > CRE)
	JUNB	(TRE > CRE)
	JUND	(TRE > CRE)
	ATFa	(No binding)
	ATF2	(CRE > TRE)
	ATF4	(CRE)
	c-MAF	(MARE I/II)
	MAFA	(MARE I/II)
	MAFB	(MARE I/II)
	NRL	(TRE-related)
	MAFF/G/K	(MARE I/II)
	NRF2	(ARE)
	NFIL6	(TRE)

(e.g. FasL, TNF)

kombinace

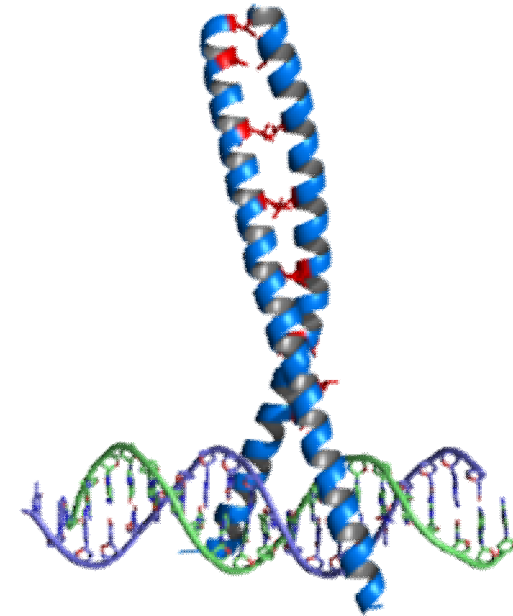


Efferl & Wagner, NRC, 2003  
Wikipedie

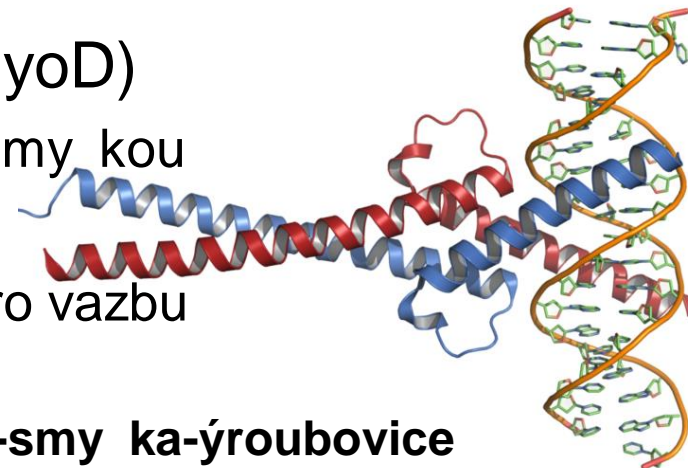


# Motivy DNA vazebných domén

- “ **Zipper typ** (dle způsobu dimerizace)
  - . **Leucinový zip** (tzv. bZIP = basic transcr. fact.  $\gamma$ GCN4, c-Jun/c-Fos)
    - “ 2  $\alpha$ -helixy (2 x 60 AMK)
    - “ coiled-coil (30AMK, Leu, C-term)
    - “ bazická část (N-terminus, navazuje na CC)
    - “ bazická zroubovice vázána do Vž



- . **Helix-loop-helix** (c-Myc/Max, MyoD)
  - “ CC a bazické části jsou odděleny smyčkou
  - “ bazická zroubovice vázána do Vž
  - “ smyčka poskytuje větší flexibilitu pro vazbu



**zroubovice-smyčka-zroubovice**

# Motivy DNA vazebných domén

## “ Zipper typ

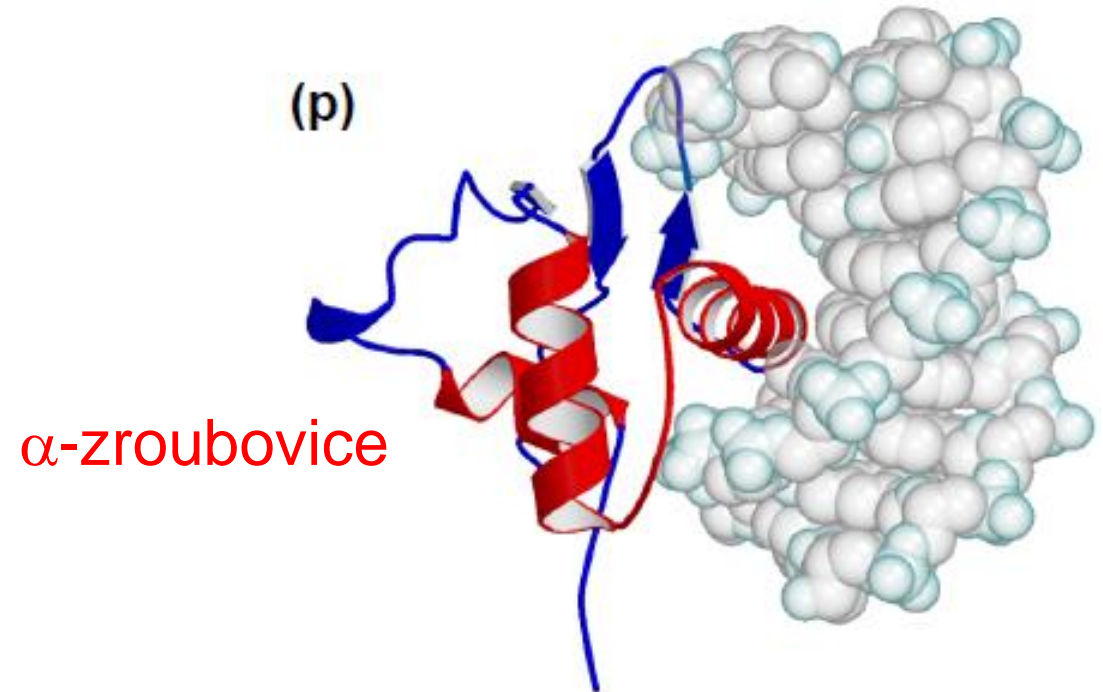
- . Leucinový zip
- . Helix-loop-helix

## “ Helix-otá ka-helix

- . HTH
- . Winged helix
- . TALE

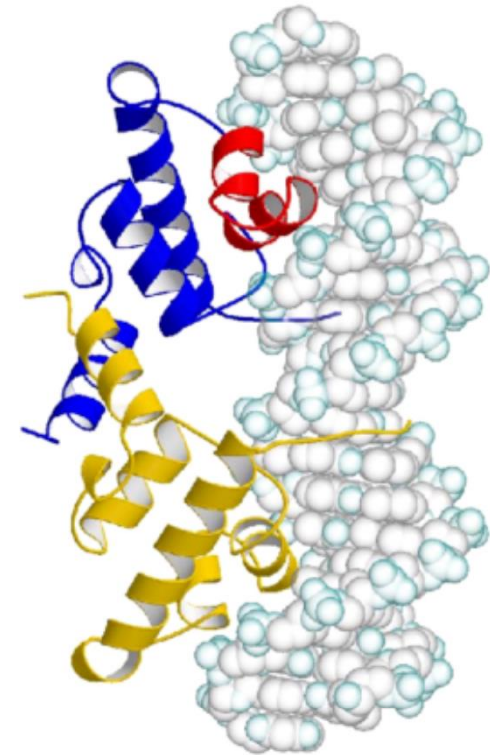
## “ Zinkový prst

- .  $\beta\beta\alpha$  zinc-finger
- . Hormon-receptor
- . Loop-sheet-helix
- . Gal4

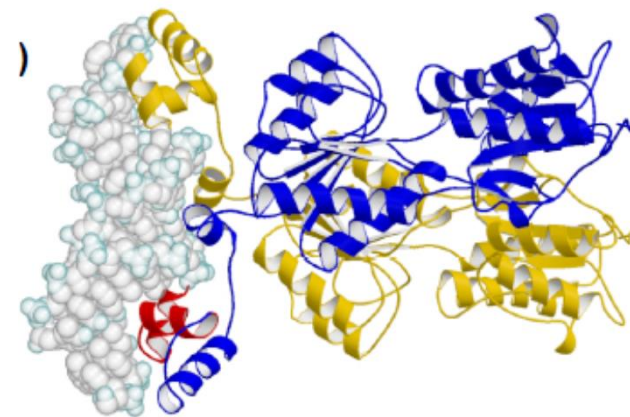


# Helix-turn-helix motiv (HTH)

- “ Obsahuje ~ 20 AMK ve dvou zroubovicích vzájemně kolmých
  - $\alpha$ -helix pro vazbu na DNA (recognition) -  $\beta$ -obrátko - druhá zroubovice
  - Sekvenčně specifická vazba prostřednictvím recognition zroubovice a velkého  $\beta$ lábku
  - nejčastěji motiv u prokaryot - homodimery vázanou palindrom. sekvencí
  - HTH motiv se obvykle vyskytuje ve svazku 3-6 zroubovic (stabilizovaných hydrofobním jádrem)
  - motiv může být buď součástí hlavního proteinu (Cro) nebo z něj může pouze vybíhat (LacI)



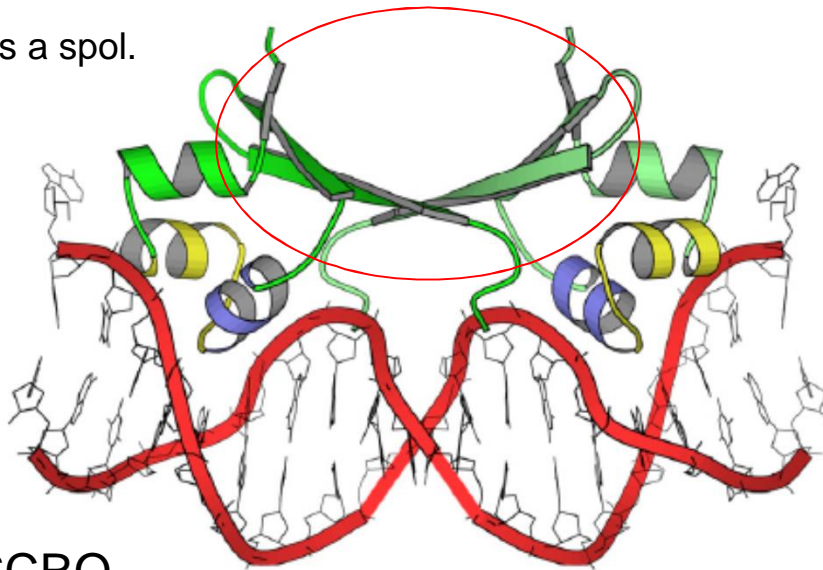
1. Cro and Repressor (1lmb)



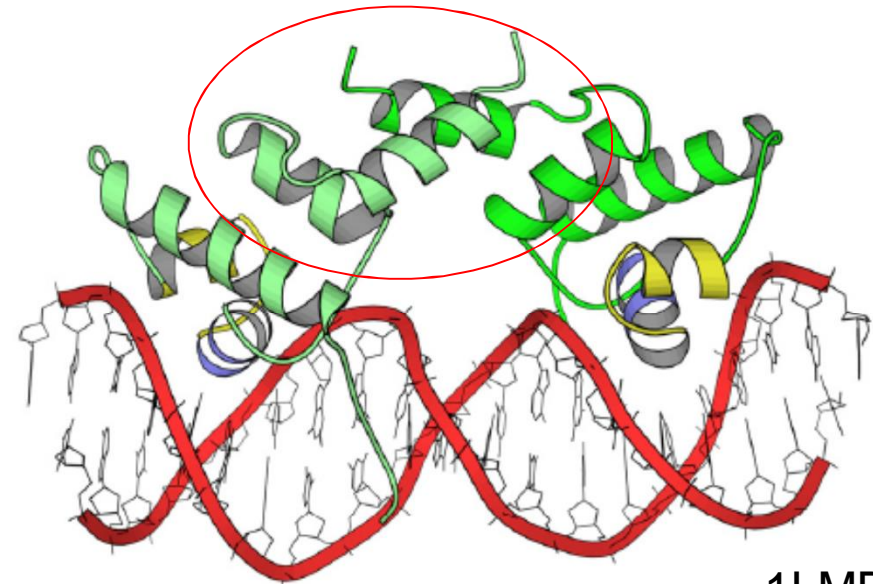
Luscombe et al, Genome Biology, 2000

3. LacI repressor (1wet)

Liljas a spol.



6CRO



1LMB

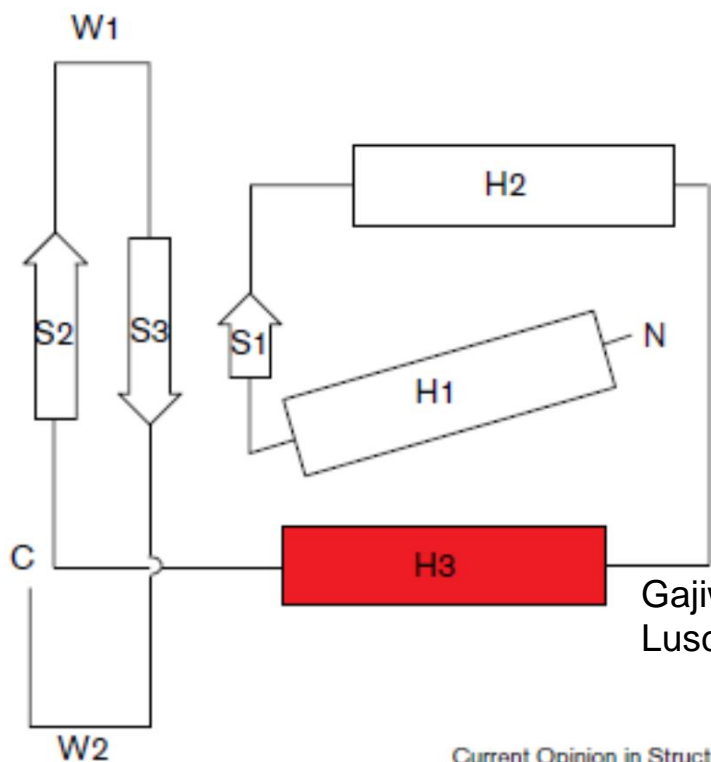
Helix-turn-helix motivy (spojené listy nebo zroubovicemi) .  
odstup HTH ( $34\text{\AA}$ ) odpovídá jedné otá ce B-DNA

Sekven n se r zné HTH p íliz nepodobají

1. variabilita v rozpoznávaných sekvencích DNA
2. variabilita v pozici srecognition%zroubovice ve velkém Olábku (paraleln k rovin bází nebo delzí zroubovice jsou paraleln k cukr-fosfátové kost e)

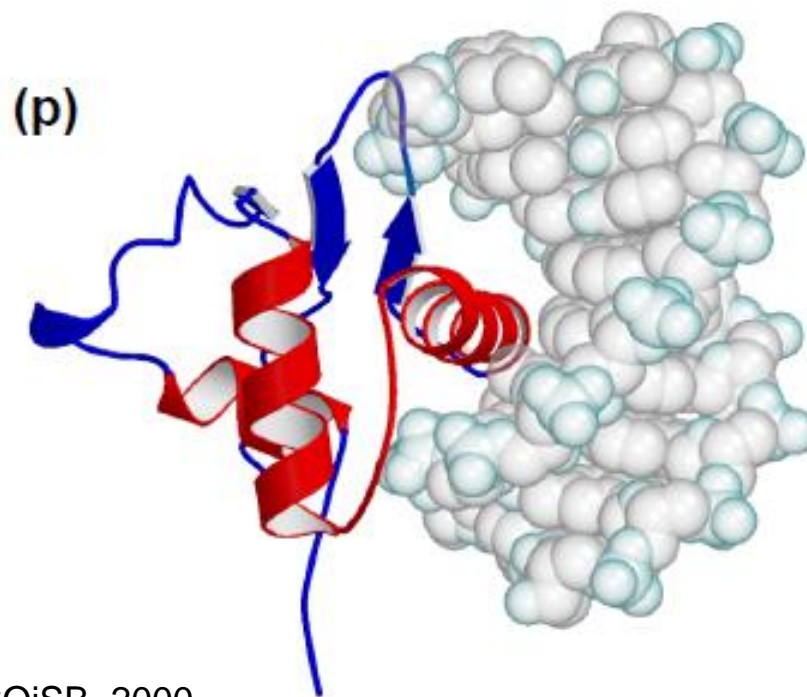
# Winged helix (okřídlená zroubovice)

- winged HTH obsahuje recognition zroubovici (**H3**) a  $\beta$ -listy, které poskytují další kontakty s DNA



Gajiwala & Burley, COiSB, 2000  
Luscombe et al, Genome Biology, 2000

Current Opinion in Structural Biology



16. Ets domain (1bc8)

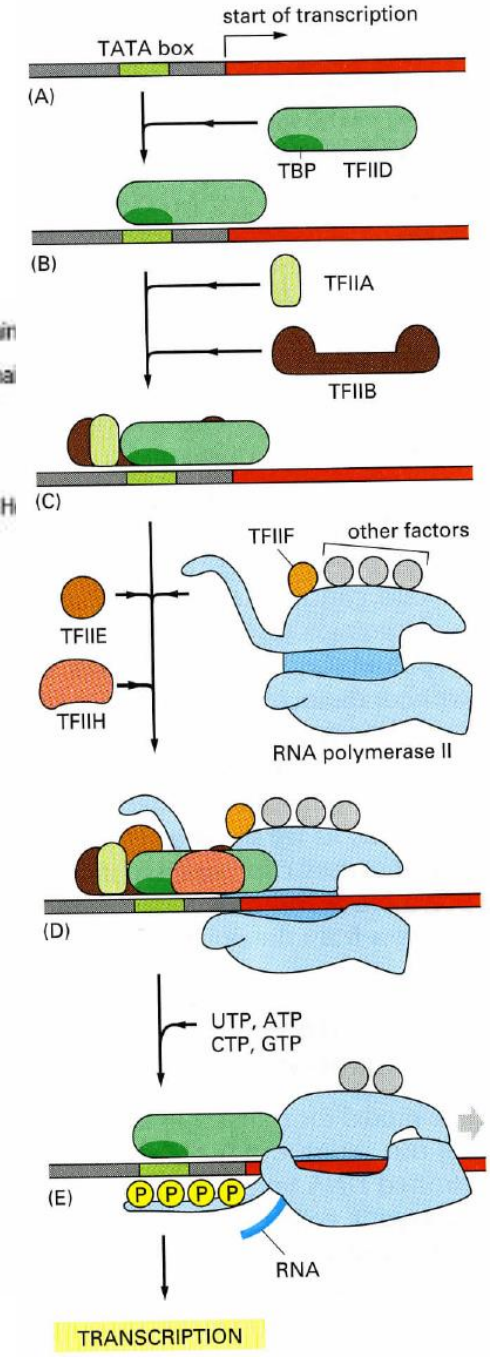
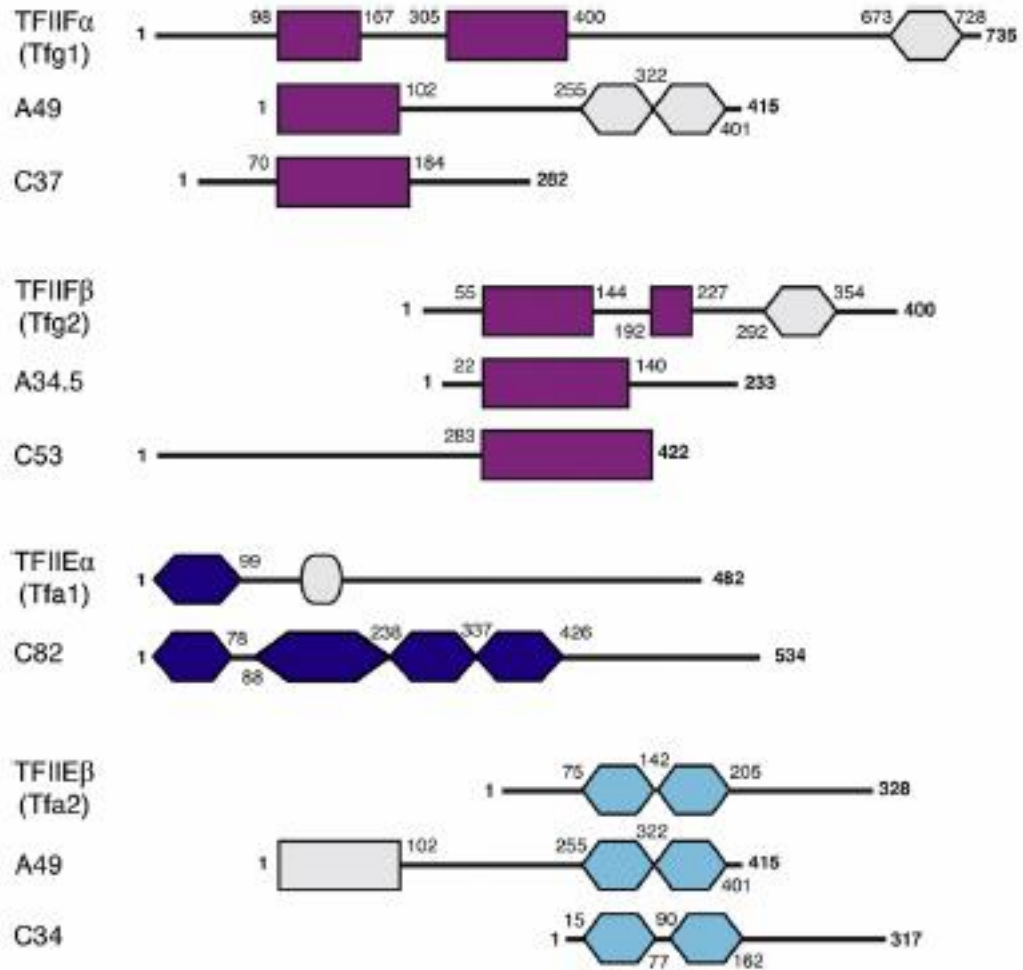
Ménasto kídlo ve VŽ a cukr-fosfátová kostra se zroubovicí (hRFX1)

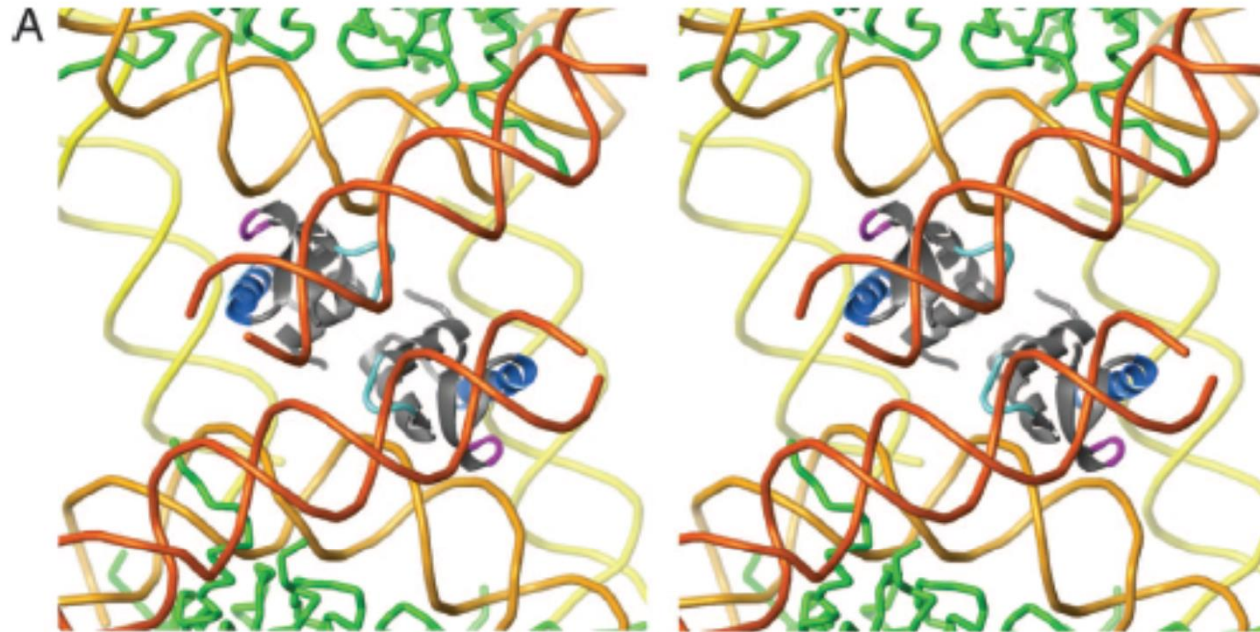
Interakce bazí se zroubovicí (H3) a kídla s cukr-fosfátovou kostrou



Ménásto kídlo ve Vž a cukr-fosfátová kostra se zroubovicí (hRFX1)

- swunged%bHTH v mnoha specifických transkrip ních faktorech, ale také v sgeneral%o TFII faktorech (strukturní úloha)

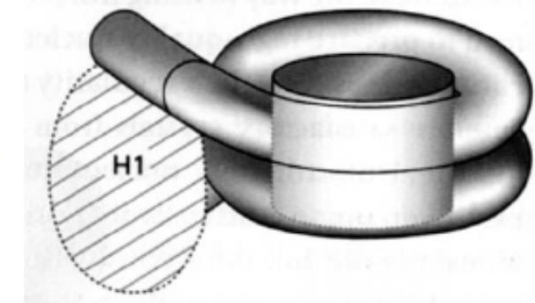
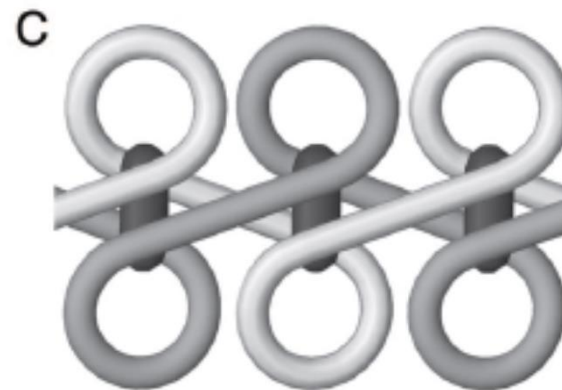
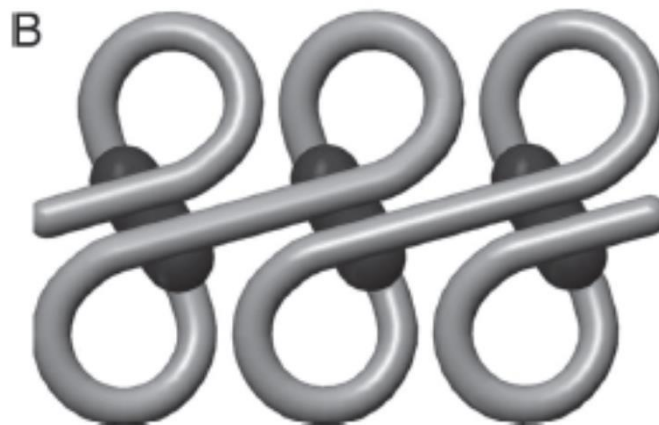




stereoskopický obrázek  
docking (pouze model)

Skládání nukleosom  
do kompaktních  
struktur

Fan & Roberts, PNAS, 2006



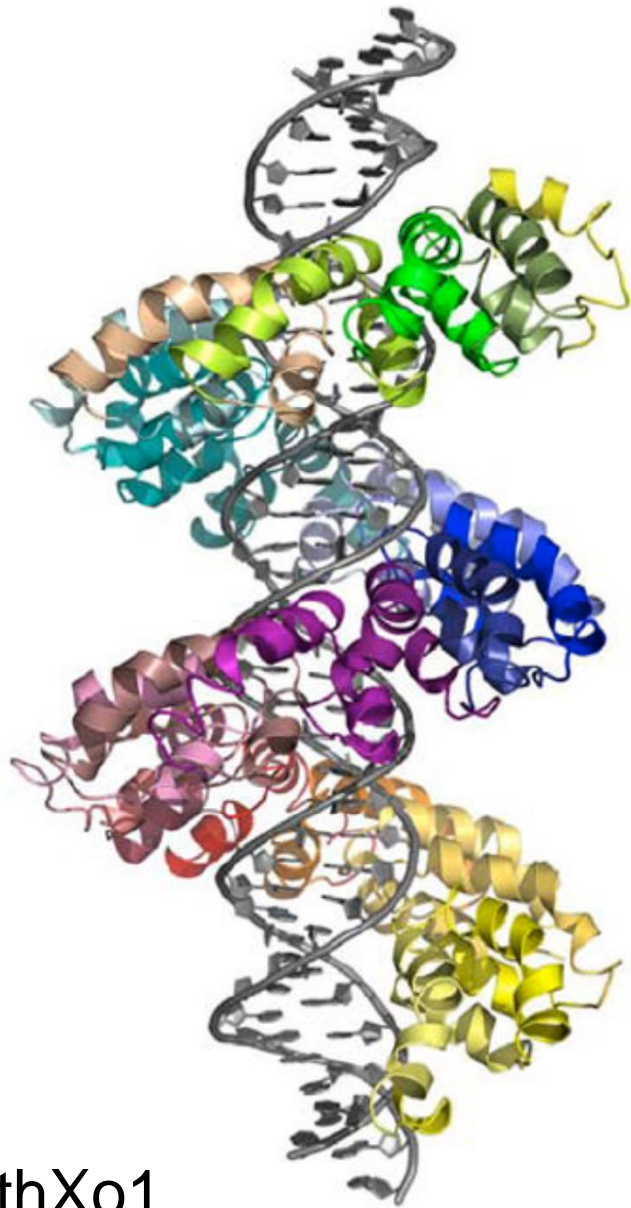
J. Fajkus

Histon H1/H5 interaguje s DNA (nukleosomem) a vytváří dimery (nukl. diady). WHD doména může vytvářet více kontaktů (H3 zroubovice-Mž, wing-cukrfosfátová kostra, protein-protein int.)



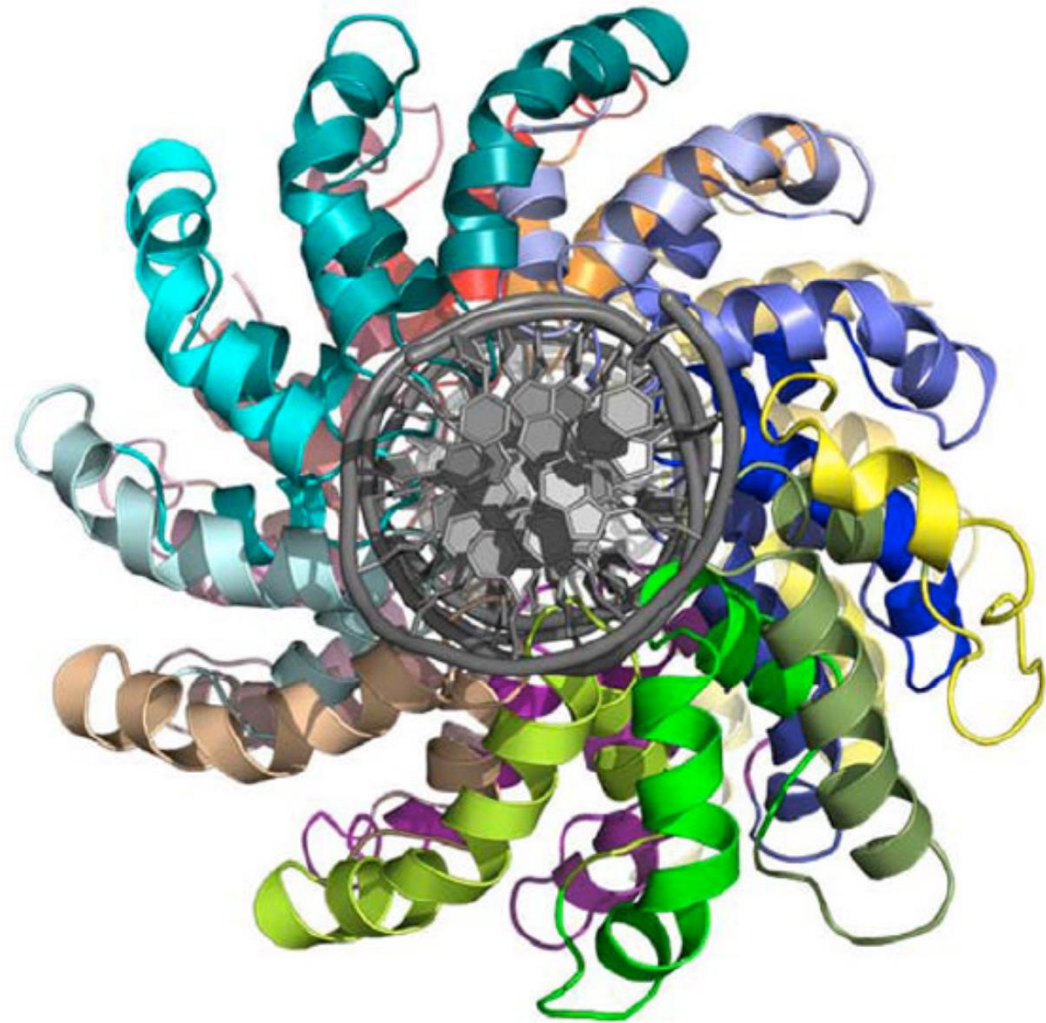
# Transcription activator-like effectors (TALE)

Patogenní bakterie injikují do rostlinných bun k ovlivnění transkripci rostlinných promotor



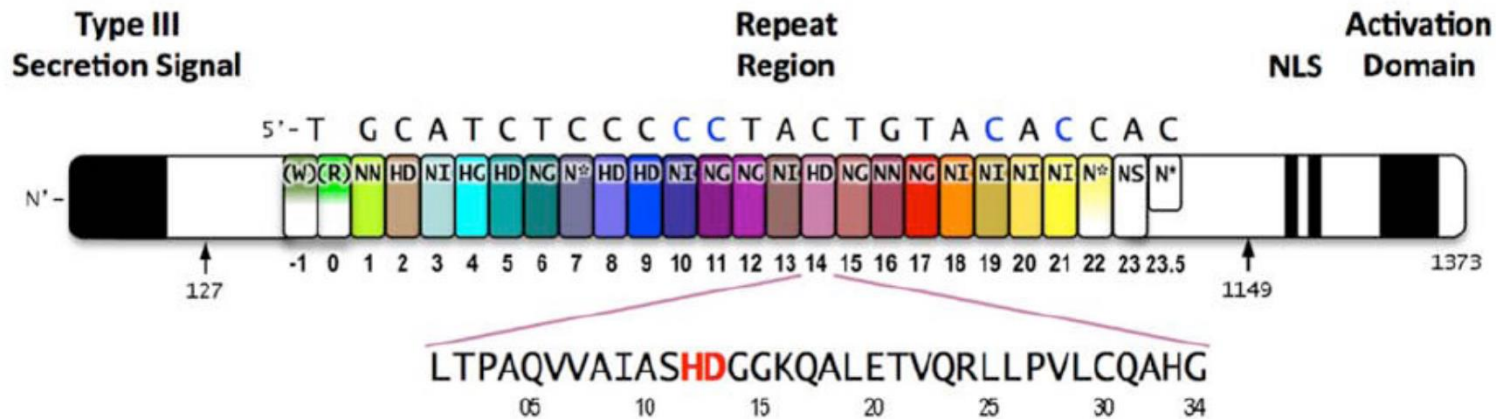
PthXo1

23 repetitivní DNA ve Vž

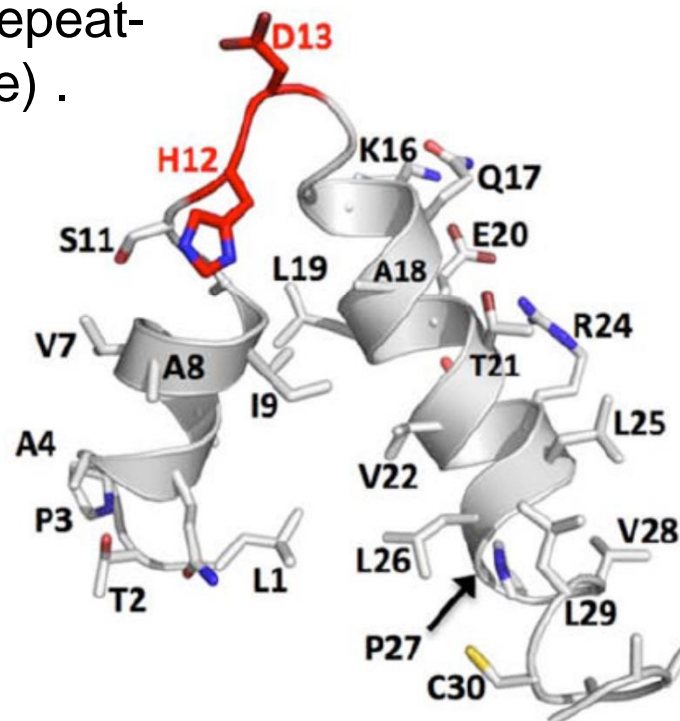


TALEN technologie

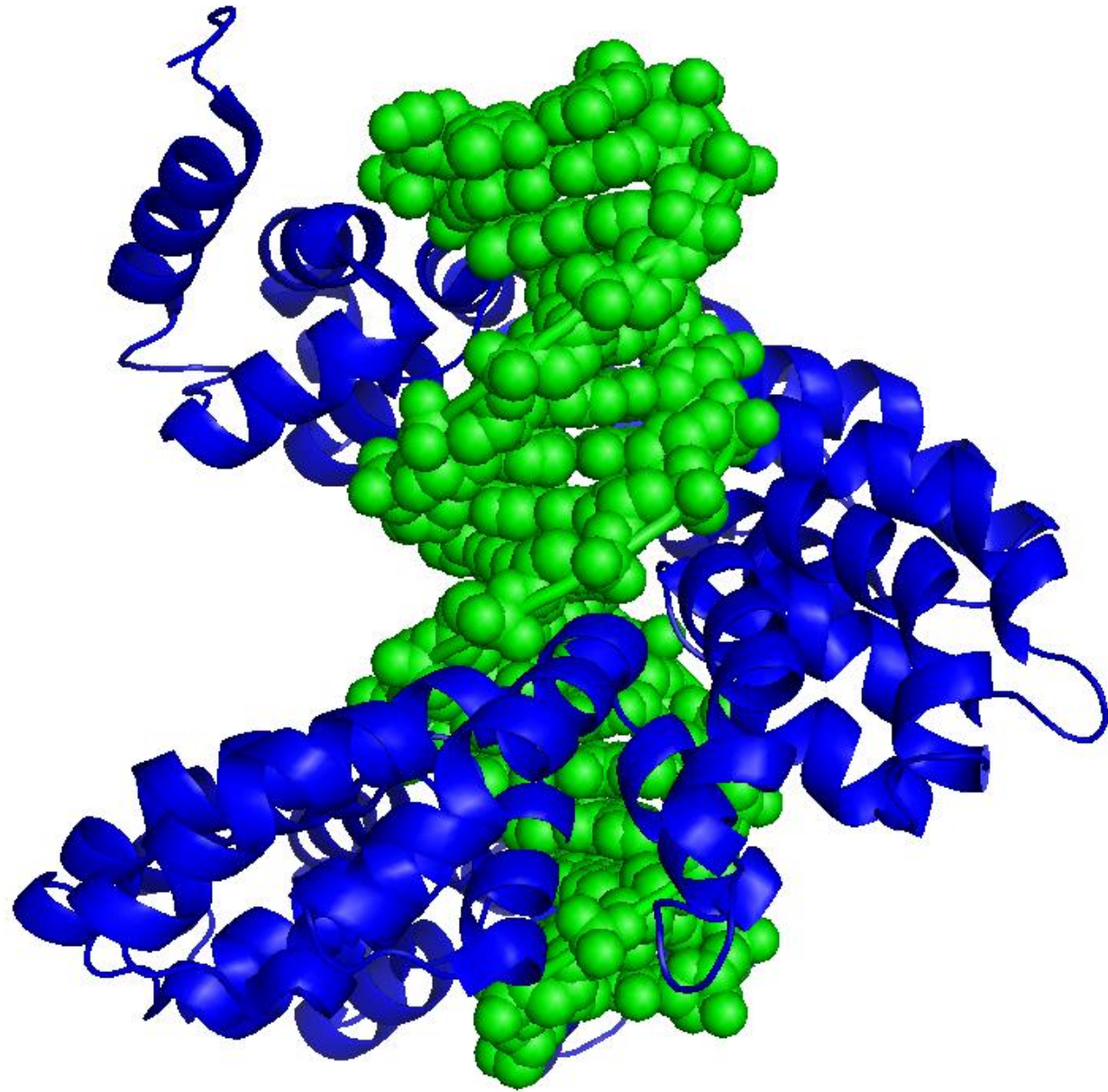
Mak et al, Science, 2012



Tandemové repetice (34)  
 AMK v pozicích 12 a 13  
 určují specifitu (repeat-  
 variable diresidue) .  
 hlavní:  
 HD, NG, NI, NN,  
 NS, HG, N\*



Interaguje otákač/turn  
 spíše než zroubovice



# Motivy DNA vazebných domén

## “ Zipper typ

- . Leucinový zip
- . Helix-loop-helix

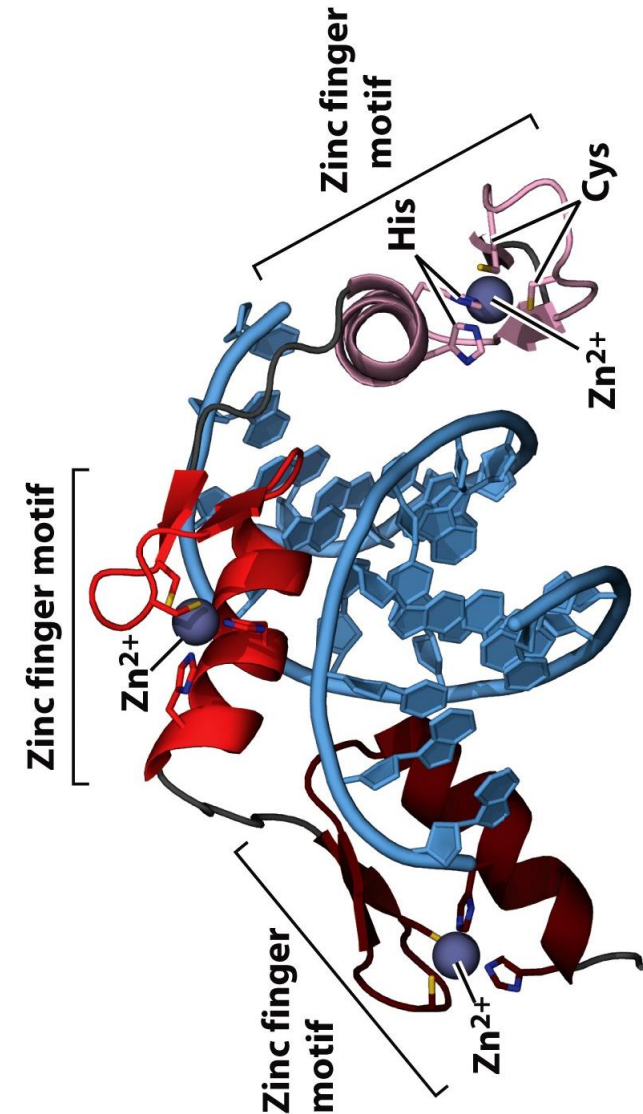
## “ Helix-otá ka-helix

- . HTH
- . Winged helix
- . TALE

## “ Zinkový prst

- .  $\beta\beta\alpha$  zinc-finger
- . Hormon-receptor
- . Loop-sheet-helix
- . Gal4

$\alpha$ -zroubovice



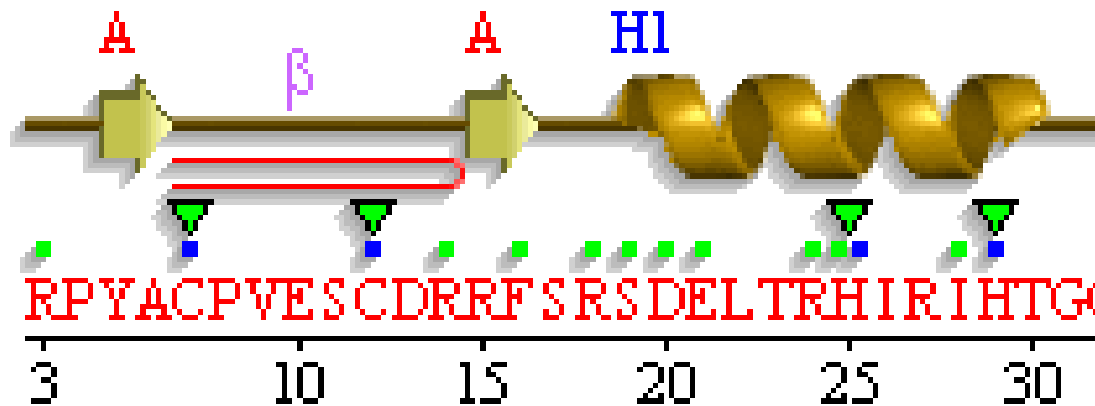
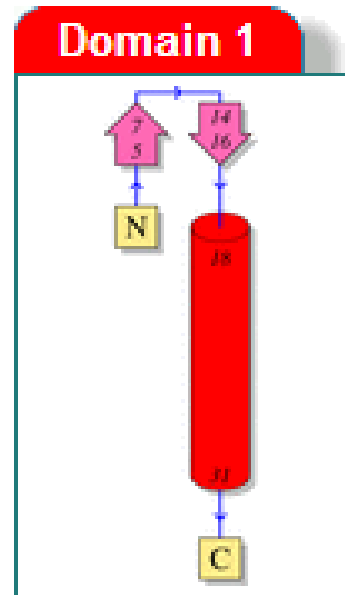
# Zinc-finger/Zinkový prst

- cca 30 AMK ve dvou krátkých antiparalelních  $\beta$ -listech a  $\alpha$ -zroubovici
- smyčka (shairpin) stabilizovaná (crosslinked)  $Zn^{2+}$  - koordinovaný 4xCys nebo 2xCys + 2xHis (tetraedrická struktura)

C2H2 motiv:

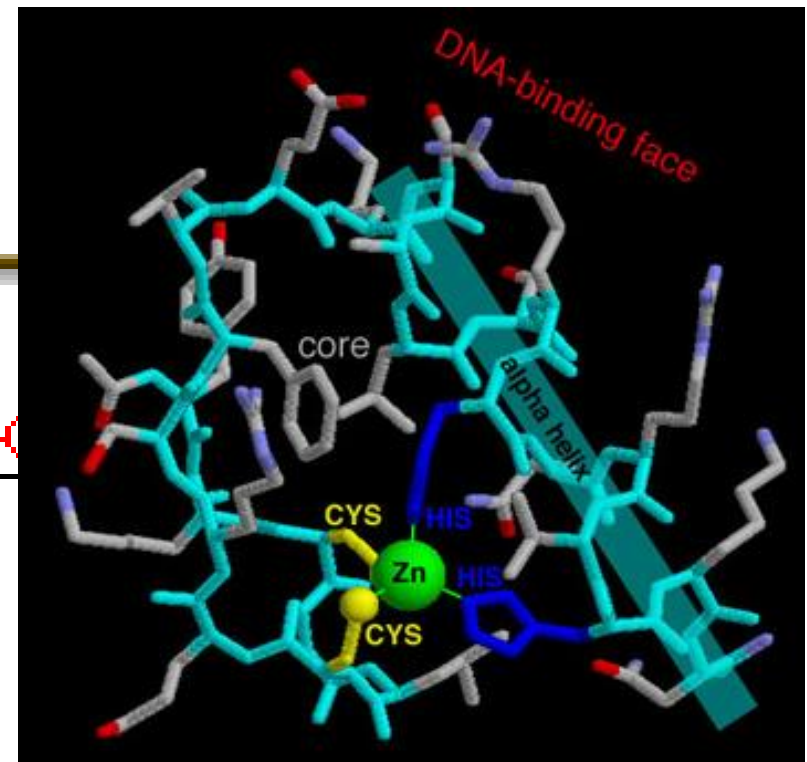
Cys-X<sub>2-4</sub>-Cys-X<sub>3</sub>-Phe-X<sub>5</sub>-Leu-X<sub>2</sub>-His-X<sub>3</sub>-His

PDB grafika



Motifs:  $\beta$  beta turn beta hairpin

Residue contacts: to DNA/RNA to metal

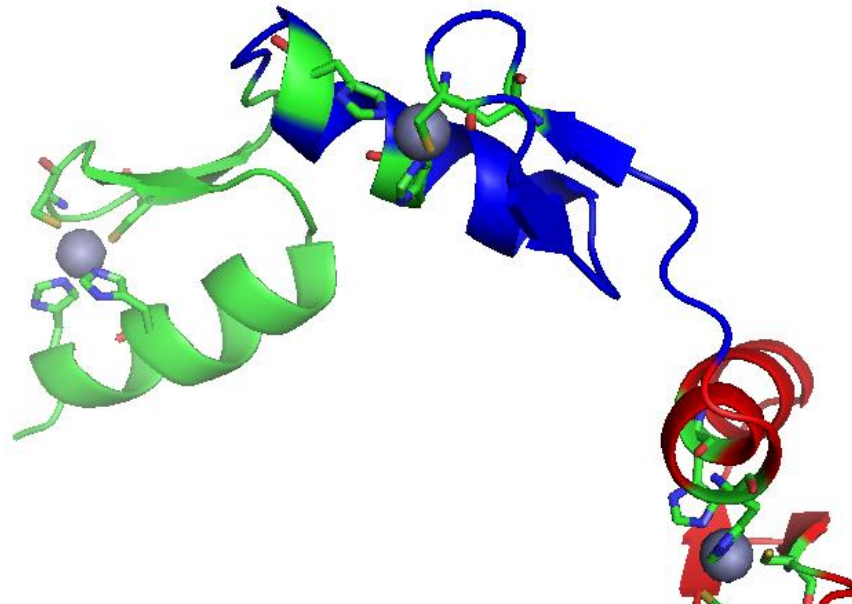
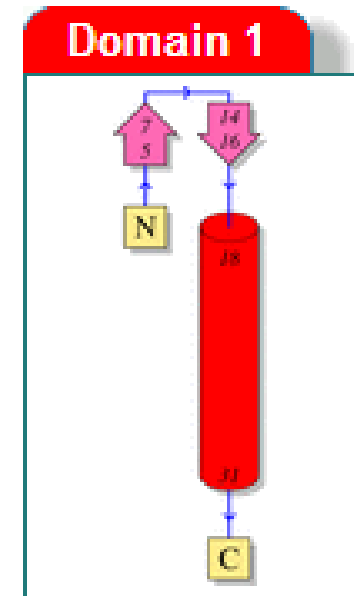


# Zinc-finger/Zinkový prst

- cca 30 AMK ve dvou krátkých antiparalelních  $\beta$ -listech a  $\alpha$ -zroubovici
- smyčka (shairpin) stabilizovaná (crosslinked)  $Zn^{2+}$  - koordinovaný 4xCys nebo 2xCys + 2xHis (tetraedrická struktura)

C2H2 motiv:

Cys-X<sub>2-4</sub>-Cys-X<sub>3</sub>-Phe-X<sub>5</sub>-Leu-X<sub>2</sub>-His-X<sub>3</sub>-His



- vazba na DNA je relativně slabá, takže se typicky vyskytuje tandemově několik domén za sebou (3x v Zif268, PDB=1zaa)
- $\alpha$ -zroubovice se váže do V<sub>ž</sub> . v tandemu obtáhne V<sub>ž</sub>

## Chain E (85 residues)

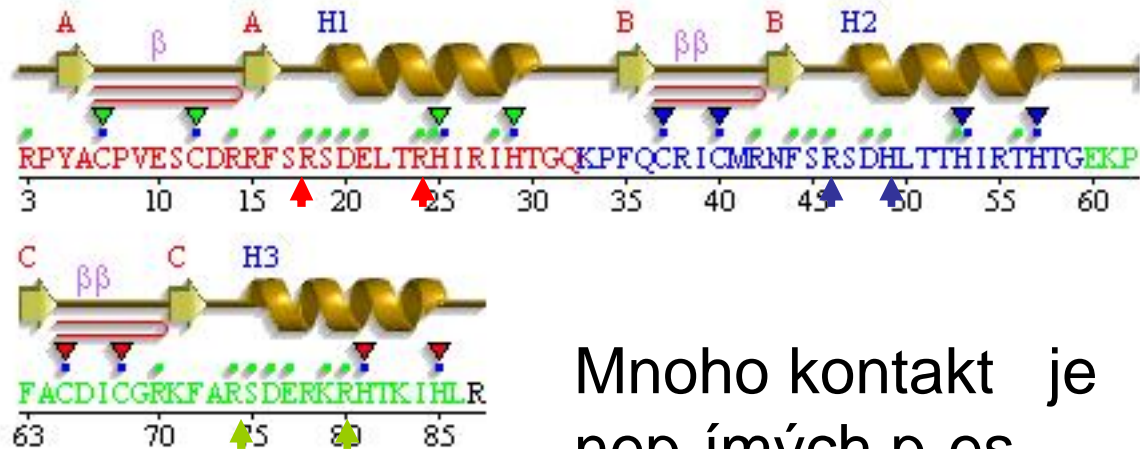
UniProt code: [P08046](#) (EGR1\_MOUSE) [Pfam]



Protein chain C highlighted  
(click to view)

**CATH** structural classification (3 domains):

Domain	Links	CATH no.	Class	Architecture
1	<a href="#">CATH</a>	3.30.160.60	= Alpha Beta 2-Layer Sandwich	
2	<a href="#">CATH</a>	3.30.160.60	= Alpha Beta 2-Layer Sandwich	
3	<a href="#">CATH</a>	3.30.160.60	= Alpha Beta 2-Layer Sandwich	

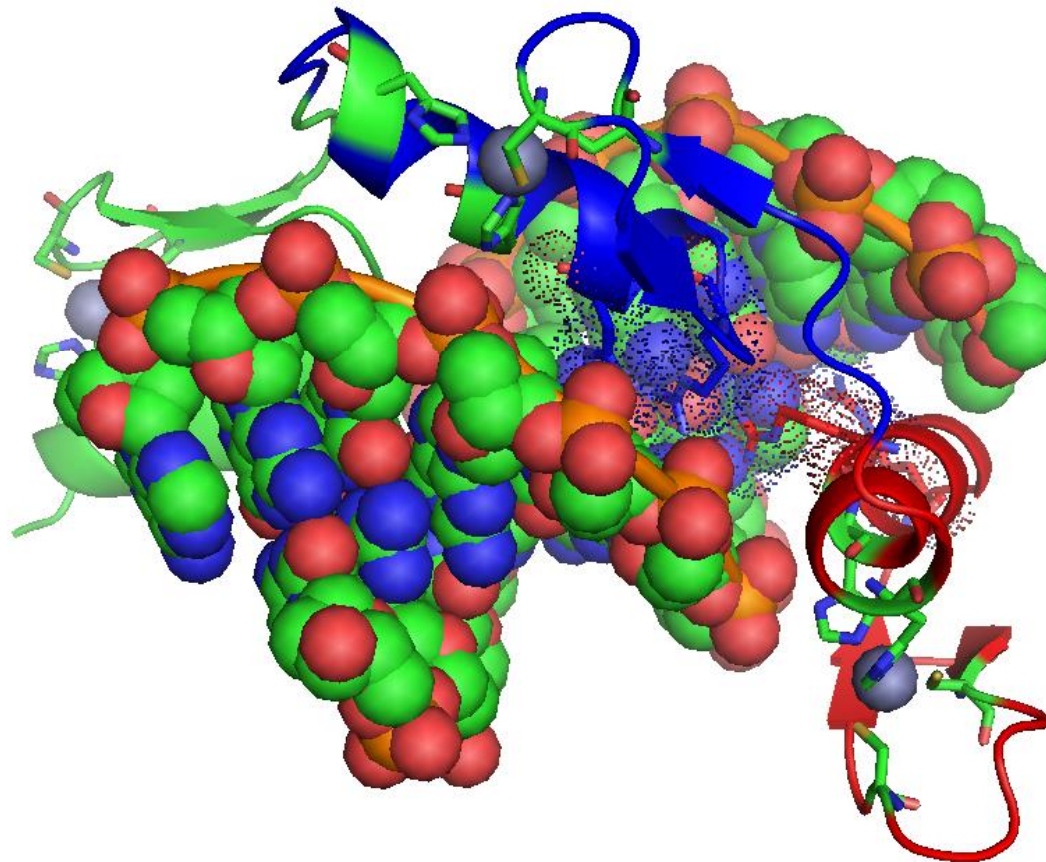
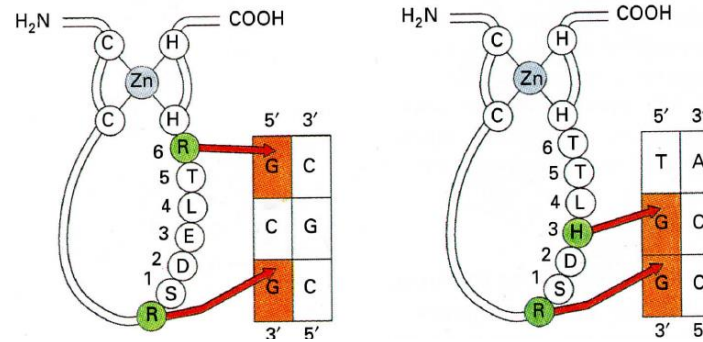


### Motifs

- Secondary structure
- Wiring diagram

Mnoho kontaktů je nepřímých přes molekuly vody

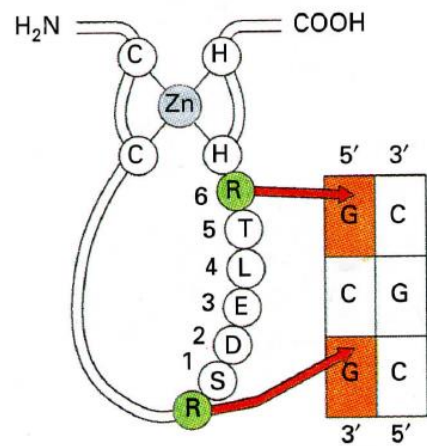
- 3x v Zif268, PDB=1zaa
- $\alpha$ -zroubovice se váže do VŽ . v tandemu obtáčí VŽ
- AMK na pozici 0 . 6; variancemi AMK => sekvenční specifita



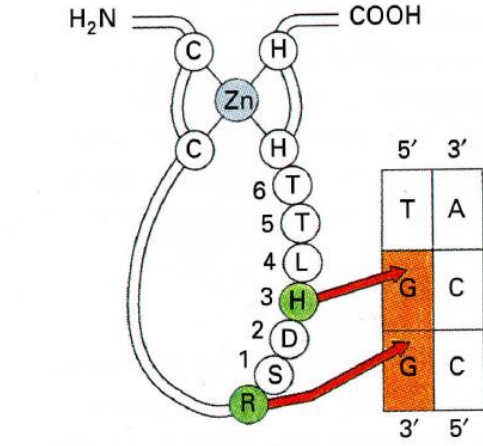


- AMK na pozici 0 . 6; variancemi AMK v t čhto pozicích lze dosáhnout r zné sekven ní specifity

-  $\alpha$ -zroubovice vá0e 2, 3 nebo 4 sousední páry bází



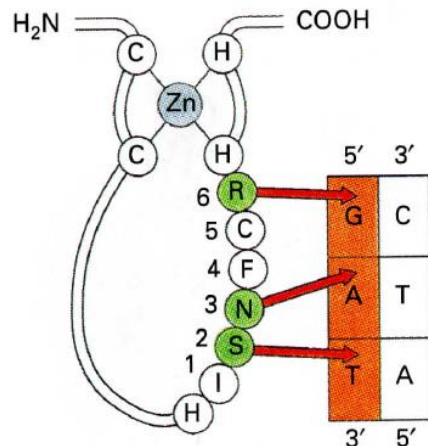
Zif Finger 1



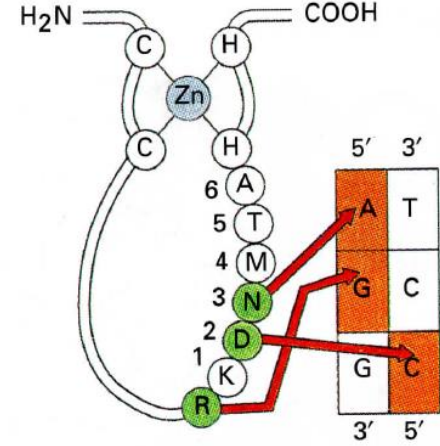
Zif268

Zif Finger 2

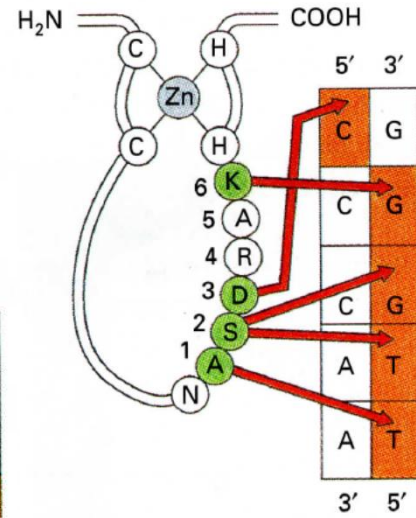
- nej ast jzí jsou kontakty Gua-Arg
- Gua se m 0e vázat i na His, Lys, Ser
- Ser se m 0e vázat na T i A



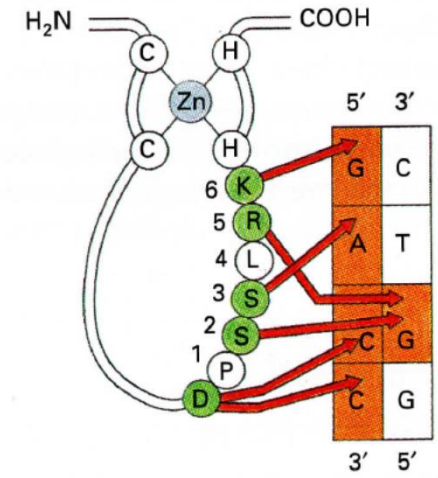
TTK Finger 1



TTK Finger 2



GLI Finger 4



GLI Finger 5

- Dobře charakterizované DNA-proteinové kontakty . je známá specifita ZFs pro všech 64 možných kombinací 3 sousedních bp
- Lze pro specifickou sekvenci DNA poskládat ZFs . nová technologie zinc nuclease%pro genové manipulace

The screenshot shows a web browser window with the URL <http://zf.princeton.edu/index.php>. The browser's address bar and tabs are visible at the top. The website's main heading is "Predicting DNA-binding Specificities for Cys<sub>2</sub>His<sub>2</sub> Zinc Finger Proteins", with a subtitle "A DNA binding site predictor for Cys<sub>2</sub>His<sub>2</sub> Zinc Finger Proteins".

On the left side, there is a sidebar with a "Přidat k oblíbeným položkám" button and a list of "Oblíbené p..." (Favorite pages) including "Encyclopedia of Life Sciences", "Ensembl release Homo sapiens Ex...", "Entrez-PubMed", "Essential Uncharacterized ORFs - Y...", "EST Profile - microarray", "EUROSCARF", "ExpASy - ProtParam tool", "ExpASy - ScanProsite", "ExpASy - Tools", "ExpASy Molecular Biology Server", "Forsburg lab S. pombe Technology", "Fralalyzer", "FreeFullPDF.com", "FUGUE Profile Lib Search", "GCG-instructions", "GCUA seqoverall optimal", "gene", "GeneCards", "GeneDB", "Genesilico Metaserver - meta2", "Genesilico", "GENESTIGATOR - shaping biolo...", "Gestation, Incubation, and Longev...", "gold Genomes OnLine Database H...", "GoPubMed", "HADDOCK docking", "HCPIN Database Site", "HMMER", "HotSpot wizard - Enzyme Enginee...", and "Human chromosome map".

The main content area has a "Home Page" button and a "Welcome to our new site!" message. Below this, there is a "Protein-DNA Form" button, a "Generate Sequence Logo" button, a "Downloads" button, an "Analyze Genomic Sequence" button, a "Help" button, and a "Contact Us" button. The text explains: "For a given C<sub>2</sub>H<sub>2</sub> zinc finger protein, we predict a position weight matrix representing its DNA binding specificity and display it as a sequence logo. This result can be used further to search genomic sequences for putative binding sites." A 3D ribbon diagram of a zinc finger protein structure is shown to the right.

At the bottom of the main content area, there is a footer with the text: "This research has been supported by NSF CCF-0542187 and NIH GM076275", a "Disclaimer: Works better with Firefox v2+, Safari v3+ or Opera v9+", and "This page has been visited 2369 times since 09-09-2010". The URL <http://zf.princeton.edu> is displayed prominently below the protein structure.

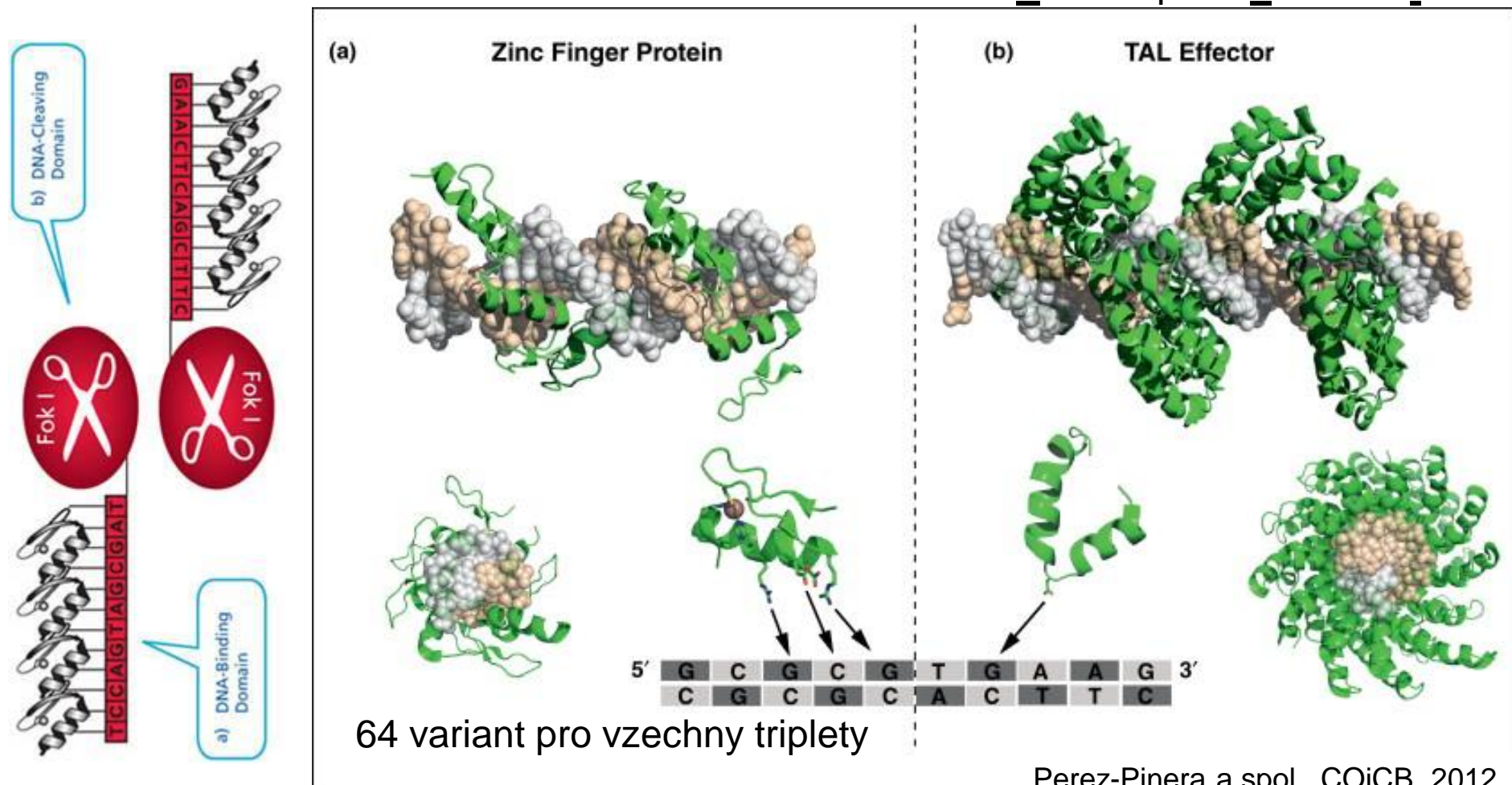
The browser's taskbar at the bottom shows various application icons, including Internet Explorer, Firefox, and several instances of Microsoft Word and PowerPoint. The system tray in the bottom right corner shows the time as 18:44 on 10.10.2014.

Persikov a Singh, NAR, 2014

- Dob e charakterizované DNA-proteinové kontakty . je známá specifita ZFs pro všech 64 možných kombinací 3 sousedních bp
- Lze pro specifickou sekvenci DNA poskládat ZFs . nová technologie šzinc nuclease%pro genové manipulace

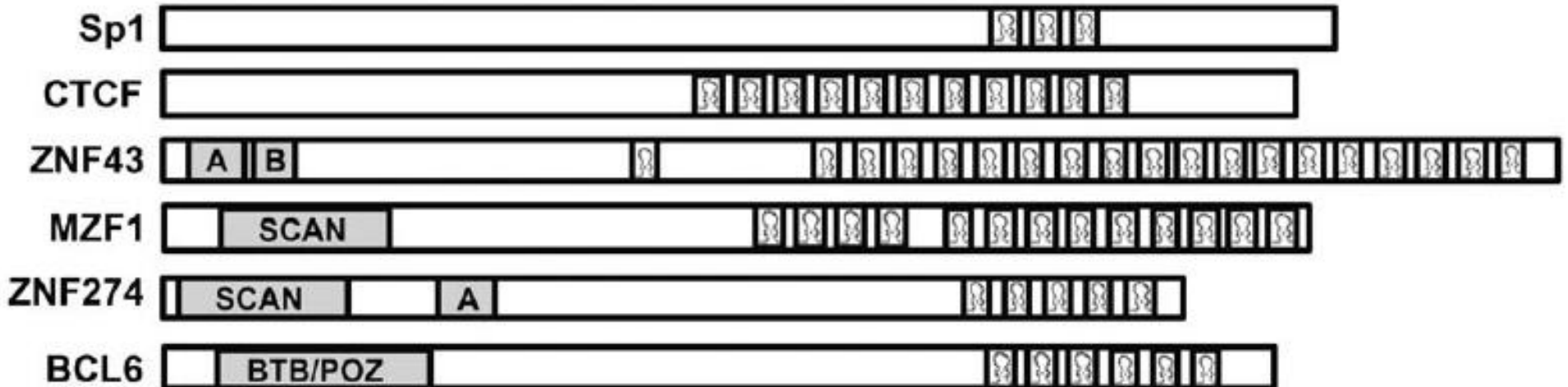
## Genome editing

### Transcription activator-like



- obsahuje 11 zinkových prstů . k vazb na DNA používá různé kombinace ZF

ZF1  
 Ts YQCEFCPYTNHKRYLLRHMKSHSEERP  
 Hs FQCELCSYTCPRRSNLDRHMKSHDTERP  
 Dm YSCPHCPYTASKKFLITRHSRSHDVEPS



ZF6  
 Ts YQCEVCNORFTQNSLKAHKLTHSG.SRPV  
 Hs YECYICHRFTQSGTMMKHILQHTENVAK.  
 Dm YQCDICKSRTQNSLKAHKLTHSVVDKPV

ZF7  
 Ts FQCKFCPSSCGRRKDLRIHVOKLHTA.SAP  
 Hs FHCYICDVIARKSDLVVHLRKHHSYIEQG  
 Dm FQCNICPTTCGRKADLRVHIKHMHTS.DVP

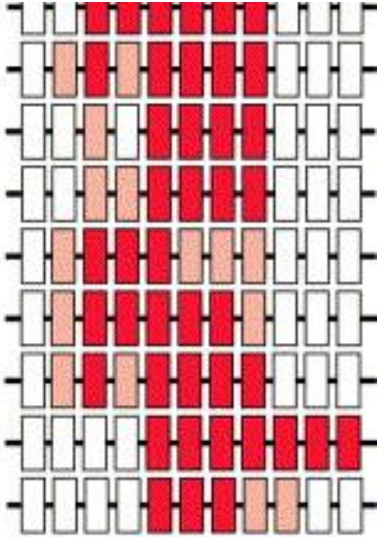
ZF8  
 Ts IKCKKCDRTFTDRYTFKLHCKEHDGERC  
 Hs KKCRYCDAVFHERYALIQHOKSHKNEKR  
 Dm MTCRRCGQLPDRYQYKLHVKSHEGEKC

ZF9  
 Ts YQCHLCPYSAMAQRHLEAHTLHSDKPF  
 Hs FKDCDCYACRQERHMIMHKRTHTGKPF  
 Dm YSCKLCSYASVTQRHLASHMILHLEKPF

ZF10  
 Ts YKCVDCNLSEKQVSLKRVHVESTFAAANQ  
 Hs YACSHCDKTFRQKQLLDMHFVKRYDPNFV  
 Dm FHCDCQPAFRQRQLLRHMNLVNEEYQ

ZF11  
 Ts LNDNLASPSTSGVSVASASSSSPSSSTPNSS  
 Hs FA.....AFVCSKCGKTFRRNTMARHADNCA  
 Dm PEPREKLHKCPSCPREFTHKGNLMRHMETHD  
 insertion

human	GAGCGGAAGAGGAAGTCCACCGCGCTTTTGAACACGGA
8. PLK h/m	AGAGGAAGATTTAAGTAAAAGCTTCCTGGAGGAGGCGCAA TCTCCTTCTAAATTCATTTTCGAAGGACCTCCTCCGCGTT
9. PIM-1 human	CTTTTCCTTCCCGCCACGCTCGCGCGCGTAGAGACCATT GAAAAGGAAGGGCGGTCGACACCCGCCCATCTCTGSTAA
10. PIM-1 mouse	GGGGAGAGGGGTGTACCGCGAGGGGCGGAGCGGAGGG CCCTCTCCCCACATCGGCGCTCCCGCCCTCGCCTCC
11. p19ARF mouse	GCAGGGCCCGCCCGCTCCCTCCCGGCGCCTCTTGGGA CGTCCCGGGCGCGCGCGGAGGGGACCCCGGAGAACCT
12. DMD4 mouse	GGAAACGGAGCTACCGCGCGGCGCAGCATACTCCTATATA CCTTGCCTCGATGGCGGCCACCGCTCTATAGGATATAT
13. DMD7 mouse	CTAAATGGACAGACGATGCCCGCGCGGACATACAATAC GATTACCTGCTGCTACCGCGCACCACCGTCATGTTATG
14. "144" silen. rat	CCTGAGTGCATTTCCCTCATGATCCAAAAGAGGGATAAC GGACTCACGTAAAAGGAGTACTAGTPTTCTCCCTTATTG
15. APP human	TTCCCCGGCGGCGCGCGTACGGGCTCTCTCTCGGGTGGCA AAGGGCCCGCGCGCGGATCCCCAGGAGAGCCACGGCT

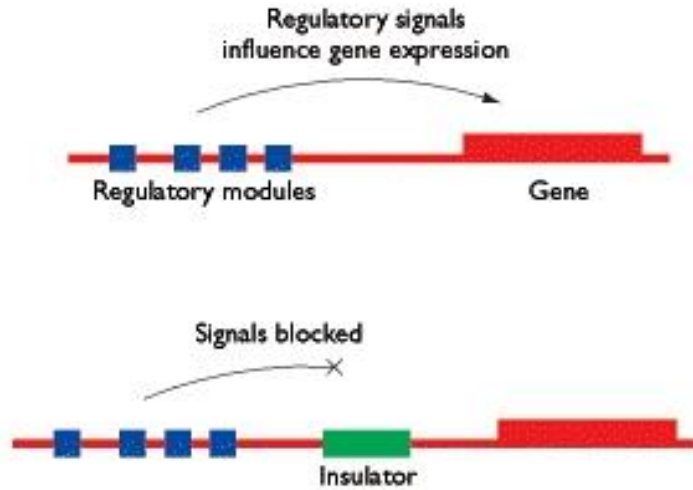


TRENDS in Genetics

Ohlsson a spol., TiG, 2001

# CTCF

A) Insulators block the regulatory signals that control gene expression



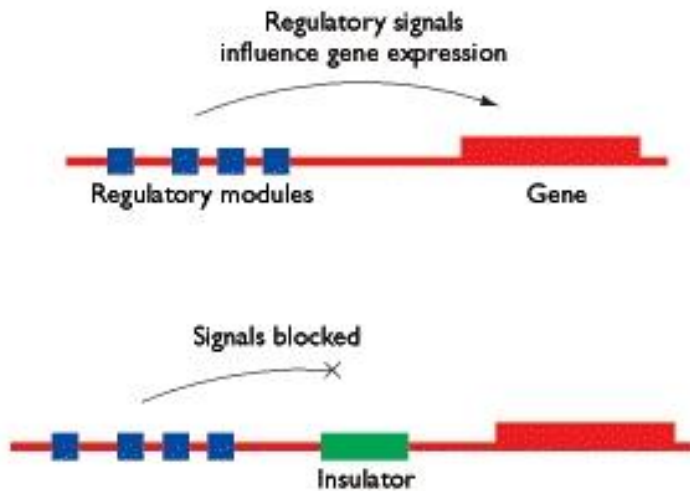
- CTCF (zkratka z CCCTC factor)
- izolátor/insulator brání transkripci
- vá0e se mezi transkrip ní aktivátory a obecné transkrip ní faktory

ZF1	<pre> Ts  YOC<sup>10</sup>EF<sup>10</sup>CPYTNH<sup>10</sup>KRYLLR<sup>20</sup>HMKSHSEERP Hs  FOCELCSYTC<sup>10</sup>PRRSNLD<sup>20</sup>RHMKSH<sup>10</sup>TDERP Dm  YSCPHCPYTAS<sup>10</sup>KKFLIT<sup>20</sup>RSRS<sup>10</sup>H<sup>10</sup>DVEPS           -1 23 6         </pre>
ZF2	<pre> Ts  FK<sup>30</sup>CTVCERC<sup>40</sup>FK<sup>40</sup>TNS<sup>40</sup>SLON<sup>50</sup>HINT<sup>50</sup>H<sup>50</sup>TGTRP Hs  HK<sup>30</sup>CHLCGRA<sup>40</sup>FR<sup>40</sup>TV<sup>40</sup>TL<sup>40</sup>LRN<sup>50</sup>HLN<sup>50</sup>H<sup>50</sup>TGTRP Dm  FK<sup>30</sup>CSI<sup>30</sup>CERS<sup>40</sup>F<sup>40</sup>RS<sup>40</sup>NVGL<sup>50</sup>Q<sup>50</sup>N<sup>50</sup>HINT<sup>50</sup>H<sup>50</sup>MGNKP           -1 23 6         </pre>
ZF3	<pre> Ts  HO<sup>60</sup>KG<sup>60</sup>CELAF<sup>70</sup>TS<sup>70</sup>GELI<sup>80</sup>FI<sup>80</sup>H<sup>80</sup>IRYK<sup>80</sup>H<sup>80</sup>TLEKP Hs  HK<sup>60</sup>CPD<sup>60</sup>CDMA<sup>70</sup>FV<sup>70</sup>TS<sup>70</sup>GELV<sup>80</sup>RR<sup>80</sup>RYK<sup>80</sup>H<sup>80</sup>THEKP Dm  HK<sup>60</sup>CKL<sup>60</sup>CESA<sup>70</sup>F<sup>70</sup>TS<sup>70</sup>GELV<sup>80</sup>RR<sup>80</sup>TRYK<sup>80</sup>H<sup>80</sup>TKEKP           -1 23 6         </pre>
ZF4	<pre> Ts  HK<sup>90</sup>CTE<sup>90</sup>C<sup>90</sup>SYASV<sup>100</sup>EL<sup>100</sup>SKL<sup>110</sup>R<sup>110</sup>H<sup>110</sup>IRSH<sup>110</sup>TGERP Hs  FK<sup>90</sup>CSM<sup>90</sup>CDYASV<sup>100</sup>EV<sup>100</sup>SKL<sup>110</sup>R<sup>110</sup>H<sup>110</sup>IRSH<sup>110</sup>TGERP Dm  HK<sup>90</sup>CTE<sup>90</sup>CTYASV<sup>100</sup>EL<sup>100</sup>TKL<sup>110</sup>R<sup>110</sup>H<sup>110</sup>MTC<sup>110</sup>H<sup>110</sup>TGERP           -1 23 6         </pre>
ZF5	<pre> Ts  YH<sup>120</sup>CPHC<sup>120</sup>CSYAS<sup>130</sup>PD<sup>130</sup>TYK<sup>140</sup>KL<sup>140</sup>R<sup>140</sup>H<sup>140</sup>LRV<sup>140</sup>H<sup>140</sup>TGEKP Hs  FO<sup>120</sup>CSL<sup>120</sup>CSYAS<sup>130</sup>RD<sup>130</sup>TYK<sup>140</sup>KL<sup>140</sup>R<sup>140</sup>H<sup>140</sup>MRT<sup>140</sup>H<sup>140</sup>S<sup>140</sup>G<sup>140</sup>EKP Dm  YO<sup>120</sup>CPHC<sup>120</sup>TYAS<sup>130</sup>Q<sup>130</sup>DM<sup>130</sup>EKL<sup>140</sup>R<sup>140</sup>H<sup>140</sup>MVI<sup>140</sup>H<sup>140</sup>TGEKK           -1 23 6         </pre>
ZF6	<pre> Ts  YO<sup>150</sup>CEV<sup>150</sup>CNQR<sup>160</sup>FT<sup>160</sup>OS<sup>160</sup>NSL<sup>170</sup>KA<sup>170</sup>H<sup>170</sup>KL<sup>170</sup>.I<sup>170</sup>H<sup>170</sup>SG<sup>170</sup>.SR<sup>170</sup>PV Hs  YE<sup>150</sup>CYI<sup>150</sup>CHAR<sup>160</sup>FT<sup>160</sup>OS<sup>160</sup>GM<sup>160</sup>M<sup>160</sup>HL<sup>160</sup>Q<sup>160</sup>K<sup>160</sup>H<sup>160</sup>TENV<sup>160</sup>AK<sup>160</sup>. Dm  YO<sup>150</sup>CDI<sup>150</sup>CKSR<sup>160</sup>FT<sup>160</sup>OS<sup>160</sup>NSL<sup>170</sup>KA<sup>170</sup>H<sup>170</sup>KL<sup>170</sup>.I<sup>170</sup>H<sup>170</sup>SV<sup>170</sup>V<sup>170</sup>DK<sup>170</sup>PK<sup>170</sup>V           -1 23 6         </pre>
ZF7	<pre> Ts  FO<sup>180</sup>CK<sup>180</sup>FC<sup>180</sup>PSS<sup>190</sup>CGR<sup>190</sup>KT<sup>190</sup>DL<sup>190</sup>LR<sup>190</sup>H<sup>190</sup>VQ<sup>190</sup>KL<sup>190</sup>H<sup>190</sup>TA<sup>190</sup>.SAP Hs  FH<sup>180</sup>CPH<sup>180</sup>CDT<sup>190</sup>VI<sup>190</sup>ARK<sup>190</sup>SD<sup>190</sup>LV<sup>190</sup>HL<sup>190</sup>RK<sup>190</sup>O<sup>190</sup>H<sup>190</sup>SY<sup>190</sup>IEQ Dm  FO<sup>180</sup>CNY<sup>180</sup>CP<sup>180</sup>T<sup>180</sup>CG<sup>190</sup>R<sup>190</sup>K<sup>190</sup>AD<sup>190</sup>LR<sup>190</sup>V<sup>190</sup>IK<sup>190</sup>H<sup>190</sup>M<sup>190</sup>TS<sup>190</sup>.DVP           -1 23 6         </pre>
ZF8	<pre> Ts  IK<sup>210</sup>CKK<sup>210</sup>CD<sup>210</sup>R<sup>210</sup>T<sup>210</sup>FT<sup>210</sup>DR<sup>220</sup>Y<sup>220</sup>TF<sup>220</sup>KL<sup>220</sup>H<sup>220</sup>CKE<sup>220</sup>H<sup>220</sup>DGERC Hs  KK<sup>210</sup>CRY<sup>210</sup>CD<sup>210</sup>AV<sup>210</sup>FE<sup>210</sup>HER<sup>210</sup>YAL<sup>210</sup>IQ<sup>210</sup>H<sup>210</sup>QKS<sup>210</sup>H<sup>210</sup>KNEKR Dm  MT<sup>210</sup>CR<sup>210</sup>RG<sup>210</sup>QQL<sup>210</sup>P<sup>210</sup>DR<sup>220</sup>Y<sup>220</sup>Q<sup>220</sup>KL<sup>220</sup>H<sup>220</sup>VKS<sup>220</sup>H<sup>220</sup>E<sup>220</sup>G<sup>220</sup>EKC           -1 23 6         </pre>
ZF9	<pre> Ts  YO<sup>230</sup>CH<sup>230</sup>LC<sup>230</sup>PYS<sup>240</sup>AMA<sup>240</sup>QR<sup>240</sup>H<sup>240</sup>LE<sup>240</sup>HA<sup>240</sup>HT<sup>240</sup>LL<sup>240</sup>H<sup>240</sup>SD<sup>240</sup>KP Hs  FK<sup>230</sup>CD<sup>230</sup>Q<sup>230</sup>CD<sup>230</sup>YAC<sup>240</sup>RO<sup>240</sup>ER<sup>240</sup>HM<sup>240</sup>IM<sup>240</sup>H<sup>240</sup>KRT<sup>240</sup>H<sup>240</sup>TG<sup>240</sup>EP Dm  YS<sup>230</sup>CK<sup>230</sup>LC<sup>230</sup>SYA<sup>240</sup>SV<sup>240</sup>T<sup>240</sup>QR<sup>240</sup>H<sup>240</sup>L<sup>240</sup>AS<sup>240</sup>H<sup>240</sup>ML<sup>240</sup>I<sup>240</sup>H<sup>240</sup>L<sup>240</sup>DE<sup>240</sup>KP           -1 23 6         </pre>
ZF10	<pre> Ts  YK<sup>260</sup>Q<sup>260</sup>VD<sup>260</sup>CN<sup>260</sup>LS<sup>260</sup>FK<sup>270</sup>Q<sup>270</sup>VS<sup>270</sup>LL<sup>270</sup>R<sup>270</sup>H<sup>270</sup>VEST<sup>270</sup>TA<sup>270</sup>AAANQ Hs  YA<sup>260</sup>CS<sup>260</sup>H<sup>260</sup>CD<sup>260</sup>K<sup>260</sup>T<sup>260</sup>FR<sup>260</sup>Q<sup>260</sup>QL<sup>260</sup>LD<sup>260</sup>M<sup>260</sup>FK<sup>260</sup>RY<sup>260</sup>DP<sup>260</sup>NFV Dm  FH<sup>260</sup>CD<sup>260</sup>Q<sup>260</sup>CP<sup>260</sup>QAF<sup>260</sup>R<sup>260</sup>Q<sup>260</sup>RL<sup>260</sup>LR<sup>260</sup>H<sup>260</sup>MN<sup>260</sup>LV<sup>260</sup>H<sup>260</sup>NEEYQ           -1 23 6         </pre>
ZF11	<pre> Ts  LNDNLASPSTSGVSVASASSSSSFSSSTSPNSS Hs  FA.....AFVGSKCGKTFTRRNTMARHADNCA Dm  PPEPREKLHKPSCPREFTHKGNLMRHMETED           -1 23 6           insertion         </pre>

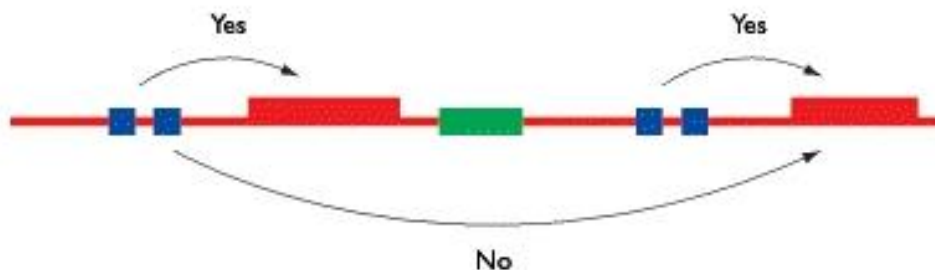
# CTCF

Izolátory chrání vzájemnou nezávislost sousedních domén, nedochází k vzájemnému rušení (B).

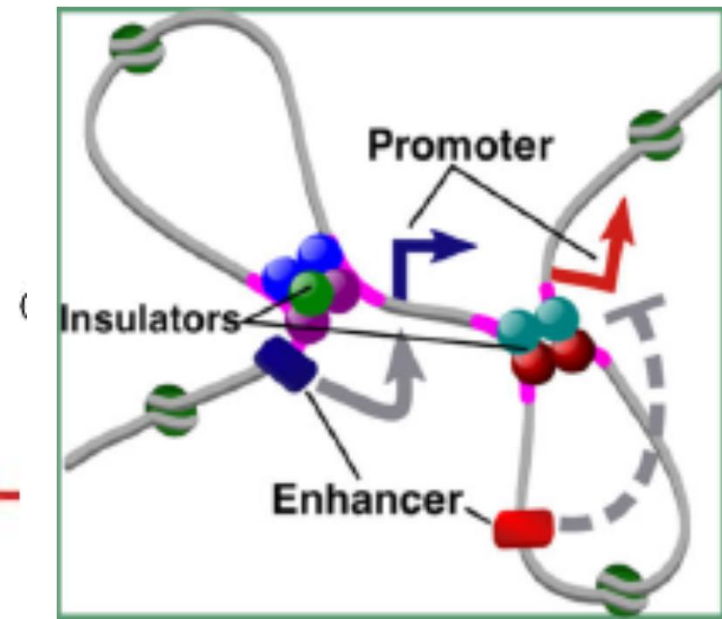
A) Insulators block the regulatory signals that control gene expression



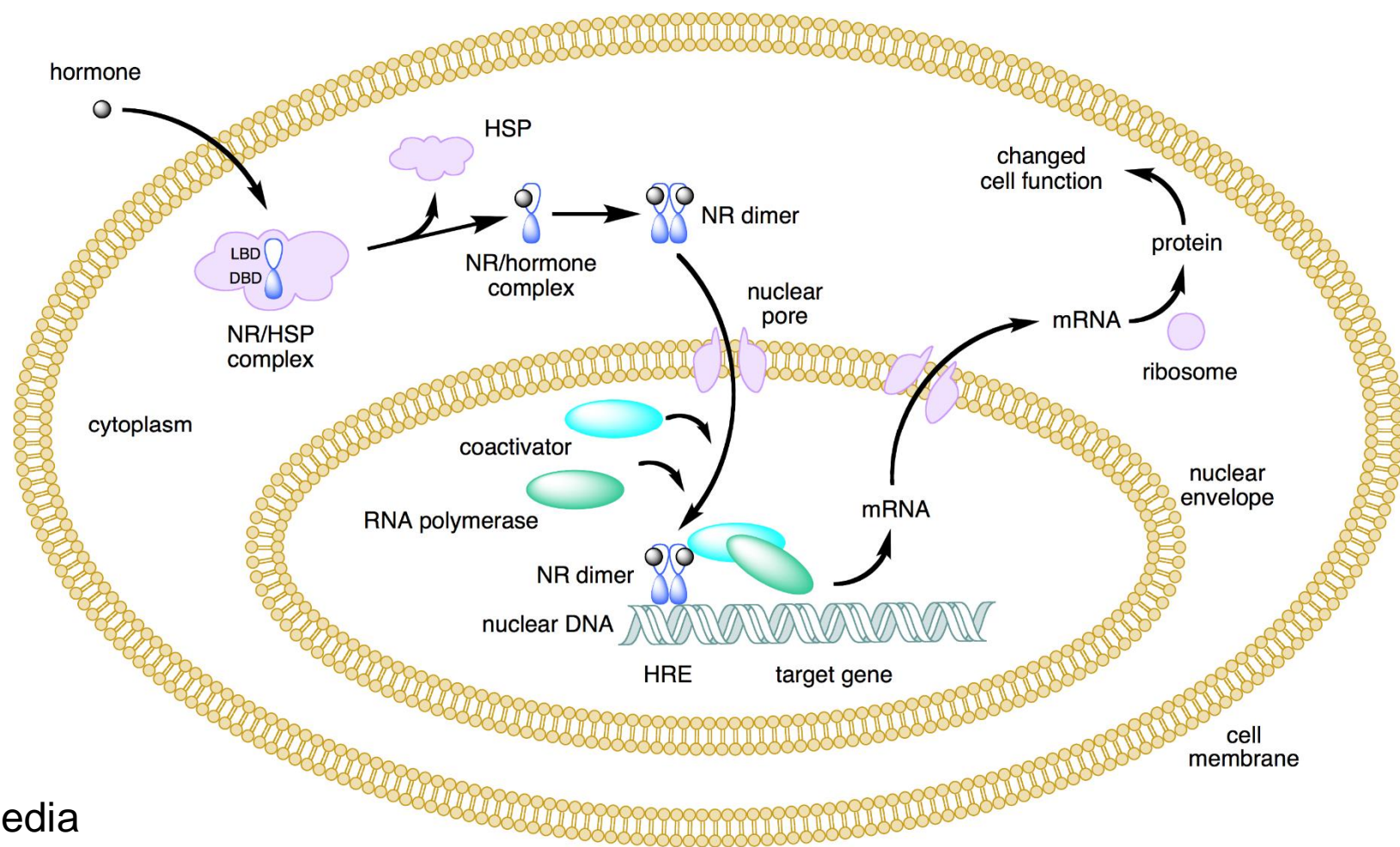
B) Insulators prevent cross-talk between functional domains



- CTCF funguje té0 jako kótva pro nukleosomy
- interaguje s kohesinem a podílí se na utvá ení vyzzích chromatinových struktur



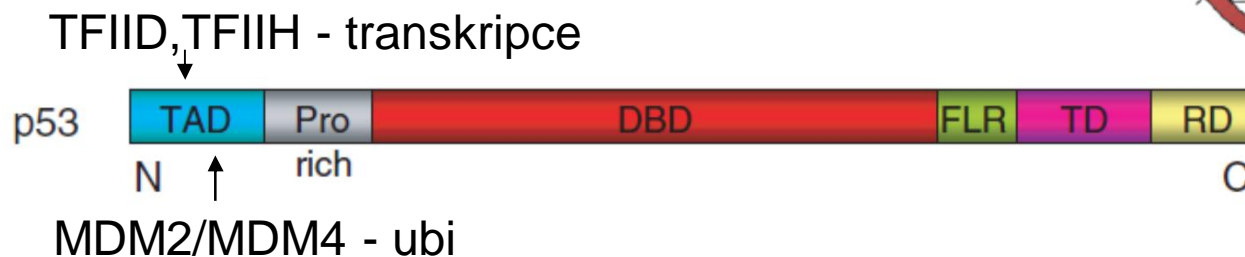
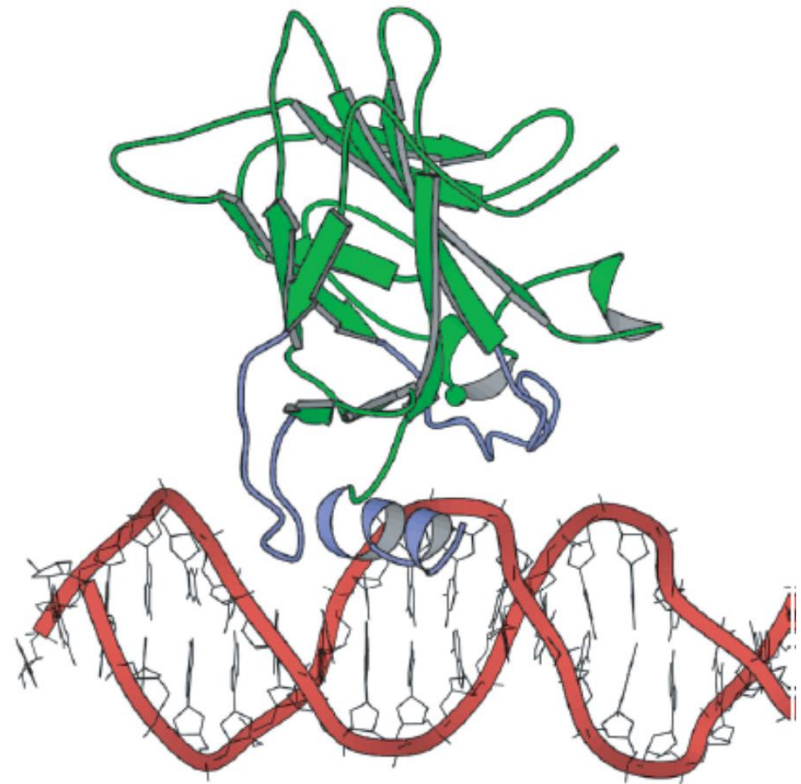
třída I (homodimery, cytoplasma) a třída II (heterodimery, jádro) . vazba ligandu také moduluje vazbu ko-aktivátor (dalších transkripčních faktorů nebo chromatinových remodelátorů)



# Loop-sheet-helix

- core/DNA-vazebná doména p53 . transkripční faktor dle0itý pro regulaci buněného cyklu, apoptozy a opravy poškozené DNA (nádorový supresor)

- smyčky vycházející mimo hlavní core doménu . vyčnívá  $\beta$ -list a  $\alpha$ -zroubovice
- 3 Cys a 1His koordinují Zn
- helix ve velkém lóábku a smyčka v malém lóábku
- Aktivace transkripce skrze kyselou TA doménu



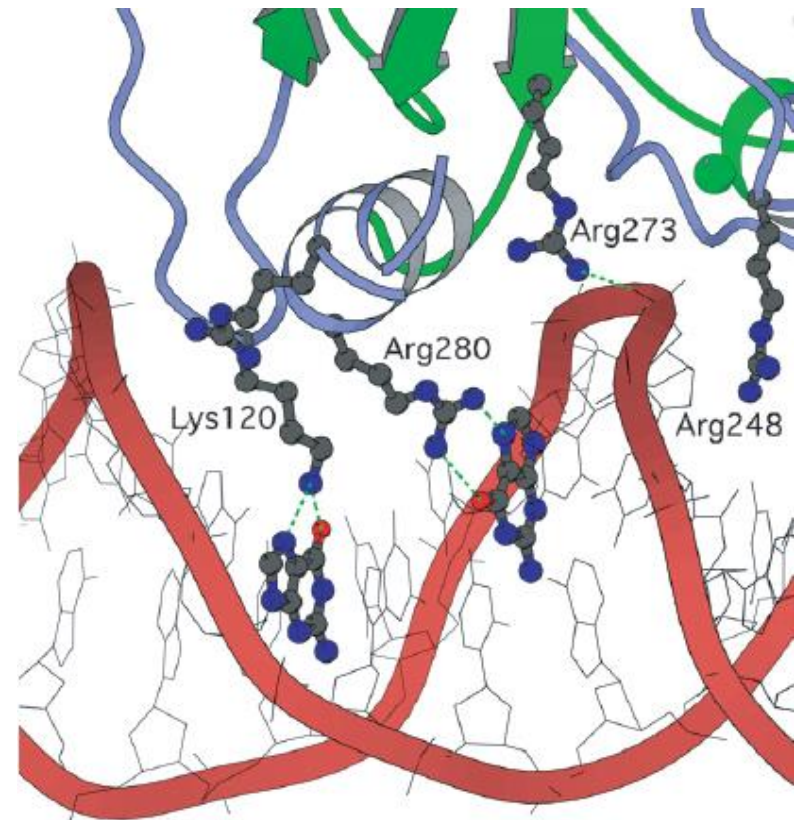


# Loop-sheet-helix

- core/DNA-vazebná doména p53 . transkripční faktor dle0itý pro regulaci buněného cyklu, apoptozy a opravy poškozené DNA (náborový supresor)

- Konsensus sekvence  
PuPuPuC(A/T)(T/A)GPyPyPy  
(v promotorech p21, PUMA)
- 95% nádorových mutací je v core doméně (R273H)
- Regulace/aktivace modifikací C-koncové domény

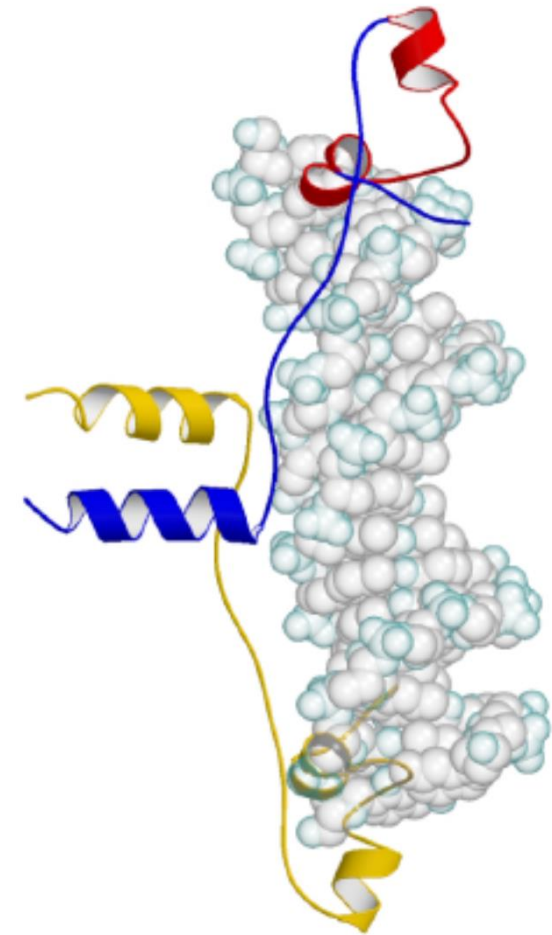
Protein se váže jako tetramer (C-koncová doména)



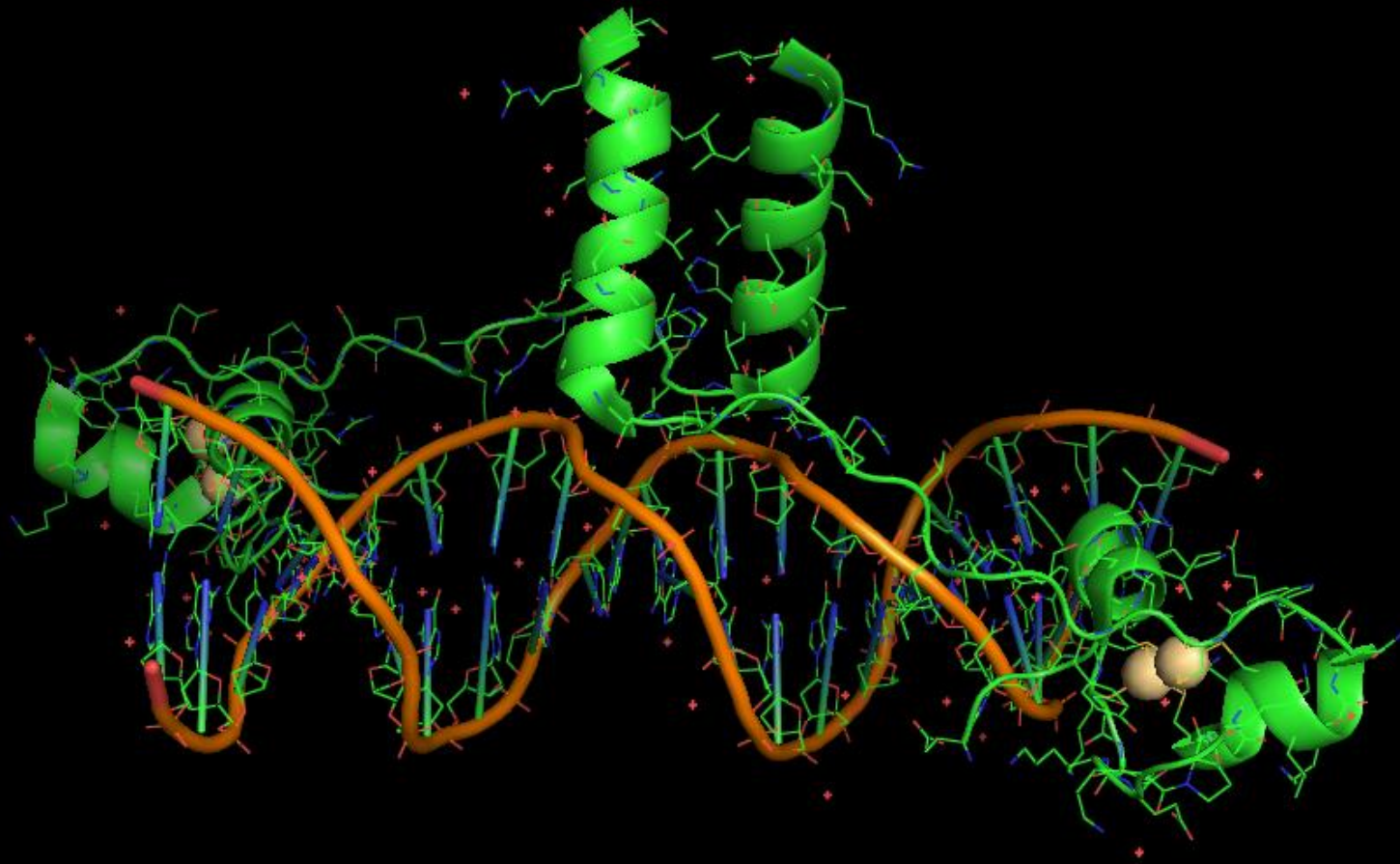
# Gal4

- transkripční faktor reguluje v kvasinkách metabolismus galaktosy (kvasinkový dvou-hybridní systém)

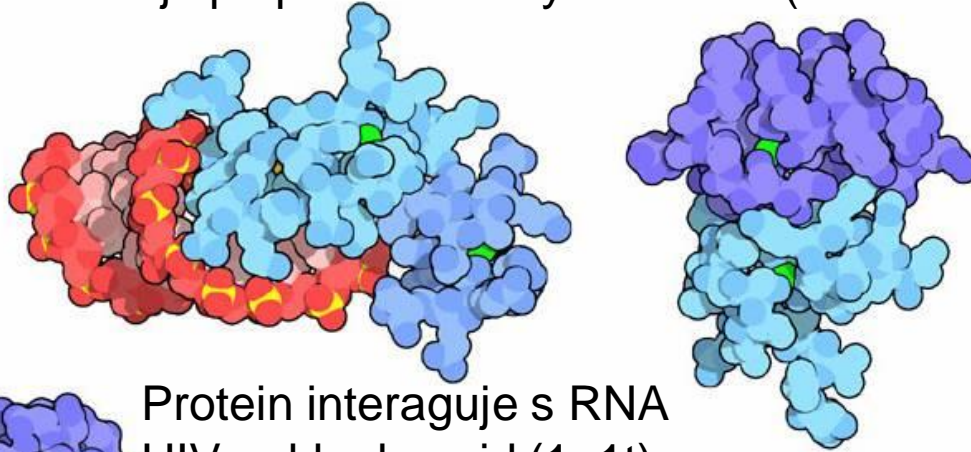
- 2  $\alpha$ -zroubovice
- 6 Cys koordinuje 2 Zn (2 Cys sdílené 2 Zn)
- 1. helix ve velkém lóbulu a 2. kontakt s cukr-fosfátovou kostrou
- Dimerizuje přes krátký CC segment



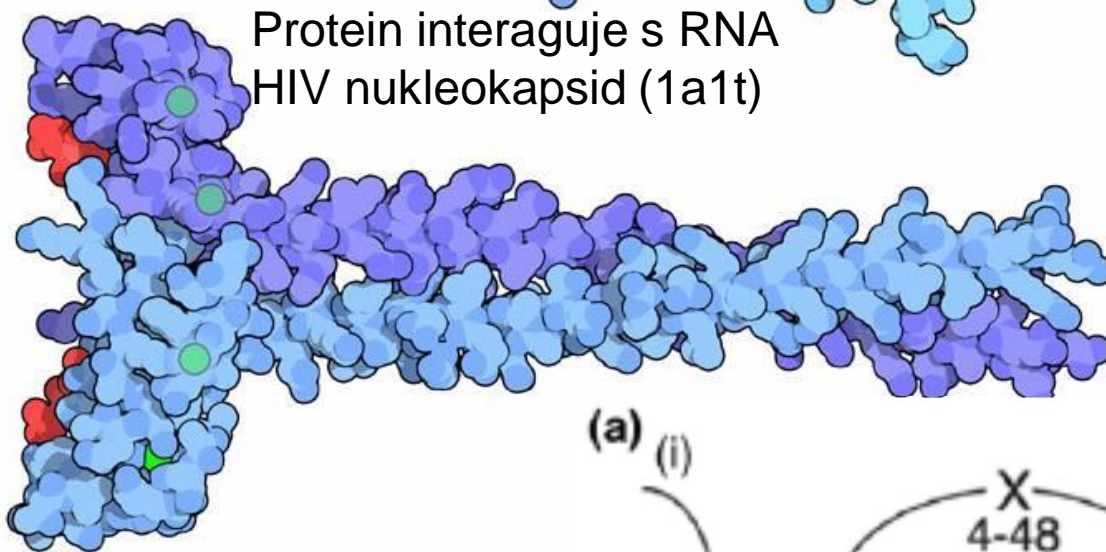
# Gal4



- jeden z nejlépe prostudovaných motiv (váže DNA, RNA i v jiných typech protein )

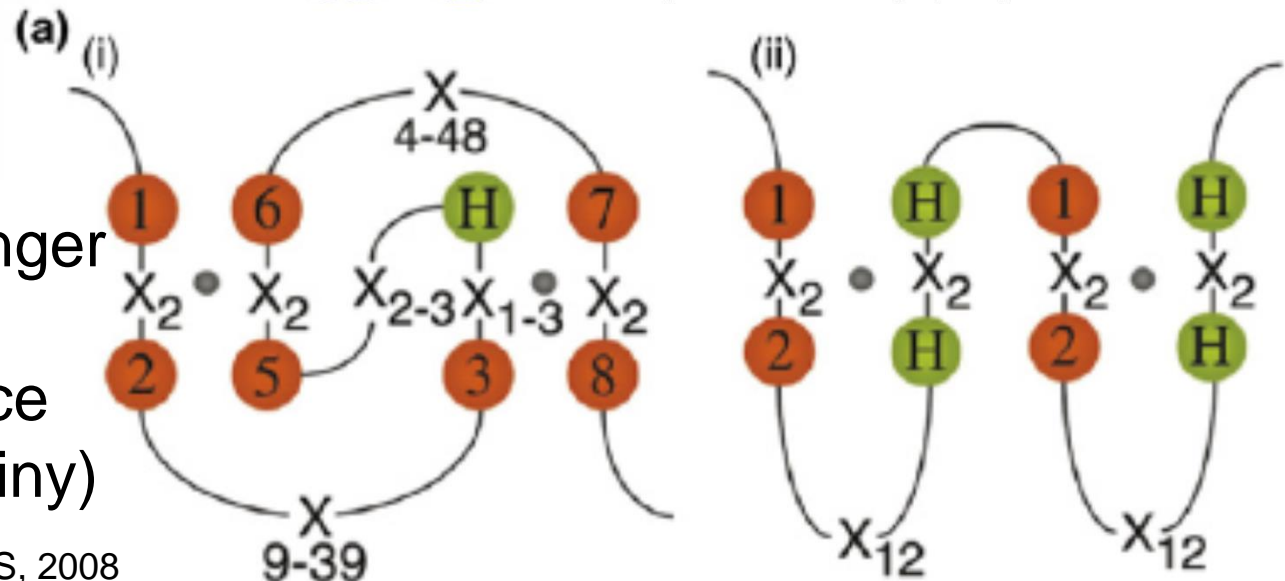


Vazba 2 zinkových prstů spojuje 2 transkripční faktory (další ZnF váže DNA): 1y0j



Protein interaguje s RNA HIV nukleokapsid (1a1t)

EEA1 protein se váže na specifický lipid obsažený v endosomech. esenciální pro transport molekul do buněčných kompartment (1ioc)



Rozdíl mezi Zinc-finger a RING-finger doménami (interakce mezi E2 a E3 proteiny)

# Motivy DNA vazebných domén

## “ Zipper typ

- . Leucinový zip
- . Helix-loop-helix

## “ Helix-otá ka-helix

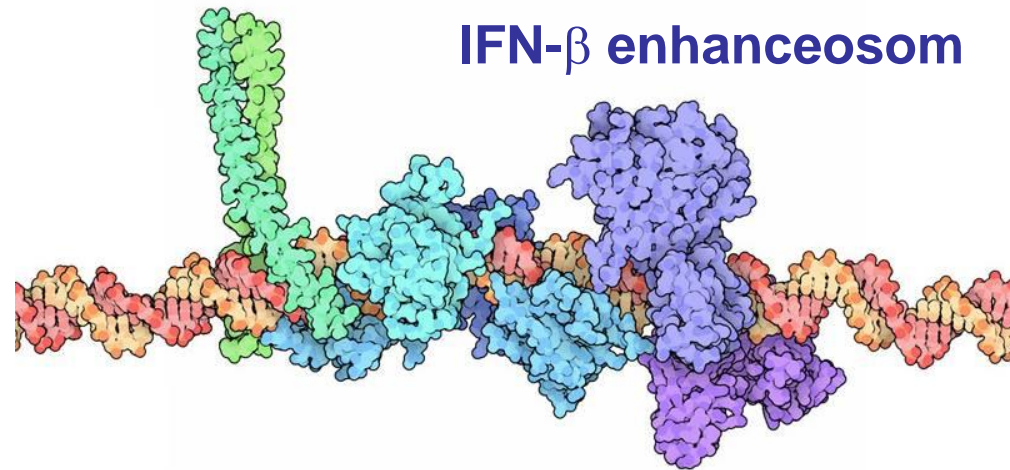
- . HTH
- . Winged helix
- . TALE

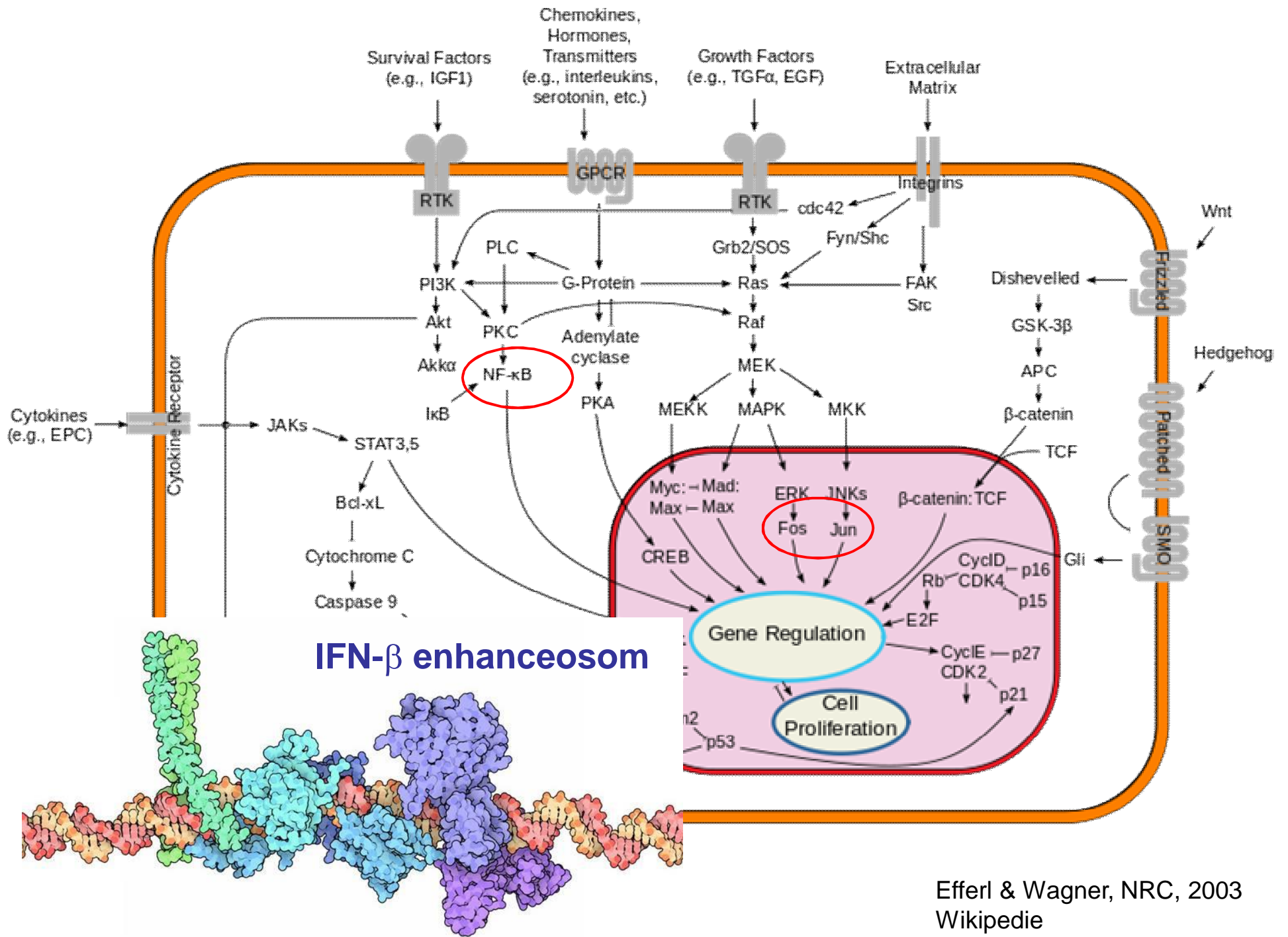
## “ Zinkový prst

- .  $\beta\beta\alpha$  zinc-finger
- . Hormon-receptor
- . Loop-sheet-helix
- . Gal4

**Kombinace motiv (ýroubovice, Zn Å )  
Å nej ast ji Vp a ýroubovice**

**Kombinace více protein Å**





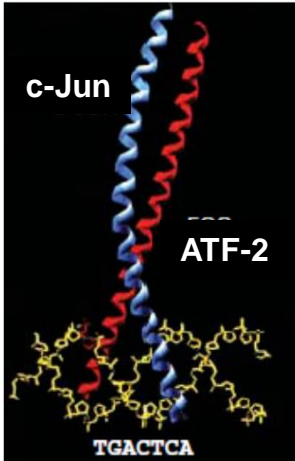
AP-1

NF-κB

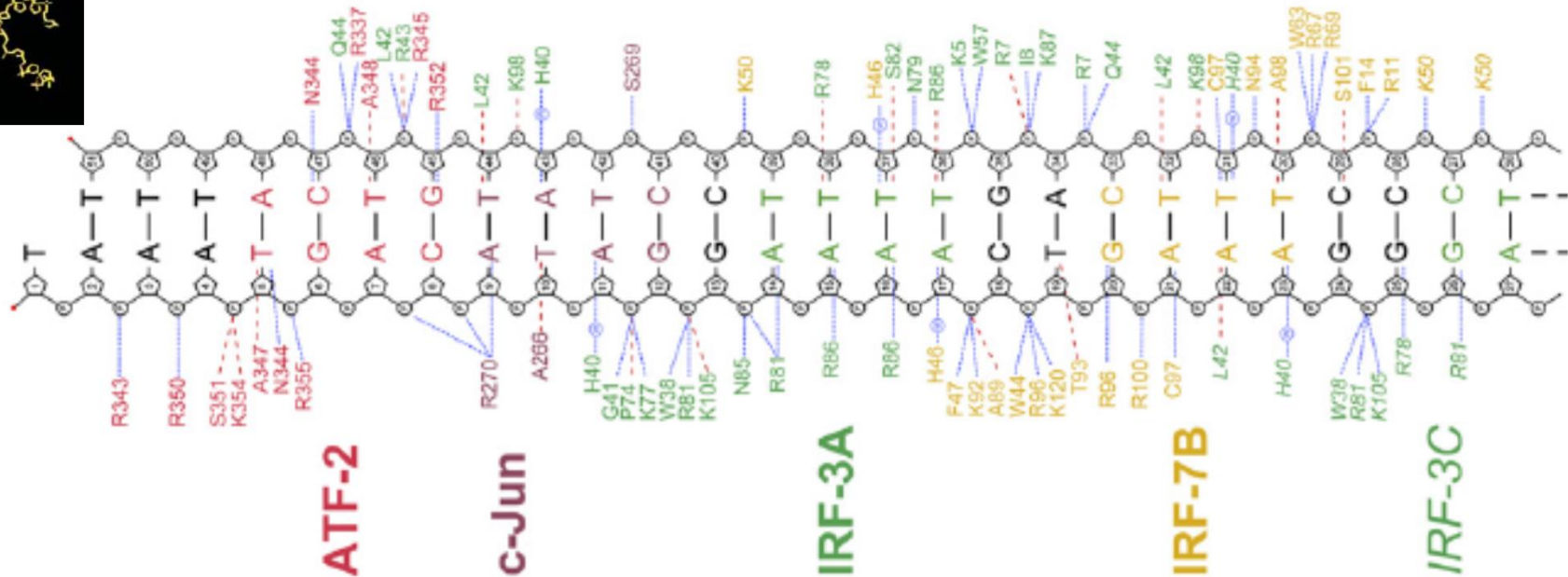
# IFN-β enhanceosom



transkripce



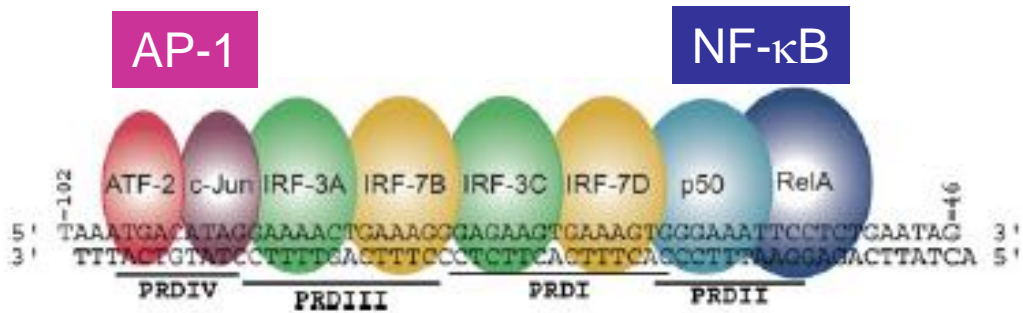
- jeden z nejlépe popsaných enhancer u vyzích eukaryot . induk. viry
- sekvence -102 a0 -47 básí upstream od po átku transkripce
- TF pokrývají 72% povrchu DNA (t sné sbalení DB-domén) . málo PPI
- nicmén vazba 8 protein je koordinovaná (AP-1, IRF-3, IRF-7, NF-κB)



Panne et al, Cell, 2008  
Panne, CO in SB, 2008

AP-1

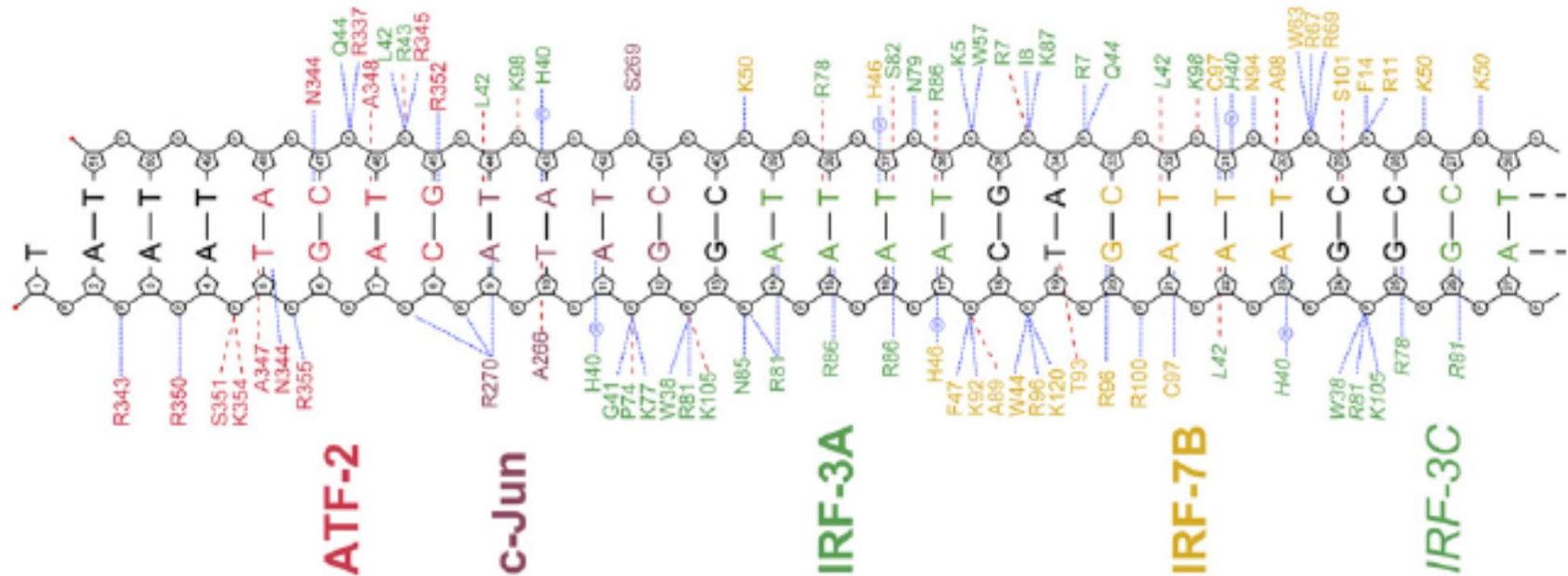
Activator Protein = b-ZIP (basic leucine zipper)



# IFN-β enhanceosom

transkripce

- koordinovaná vazba 8 protein (AP-1, IRF-3, IRF-7, NF-κB)
- AP-1 slab interaguje s IRF proteinem, ale IRF proteiny mezi sebou nemají žádný kontakt
- zroubovice IRF-3 ve Vž ohýbá DNA, což **stimuluje** vazbu dalšího IRF
- ohyby se po 1/2 otoce kompenzují, takže DNA je v tomto úseku **ROVNÁ**



Panne et al, Cell, 2008  
 Panne, CO in SB, 2008

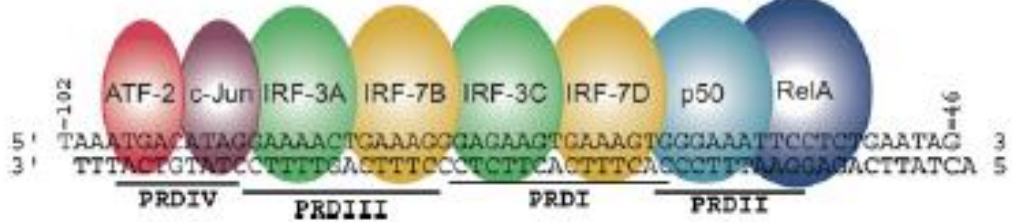
IRF . interferon regulation factor = zroubovice ve velkém Olábkú a smy ka (His) kontaktuje base v malém Olábkú (var WHD)



AP-1

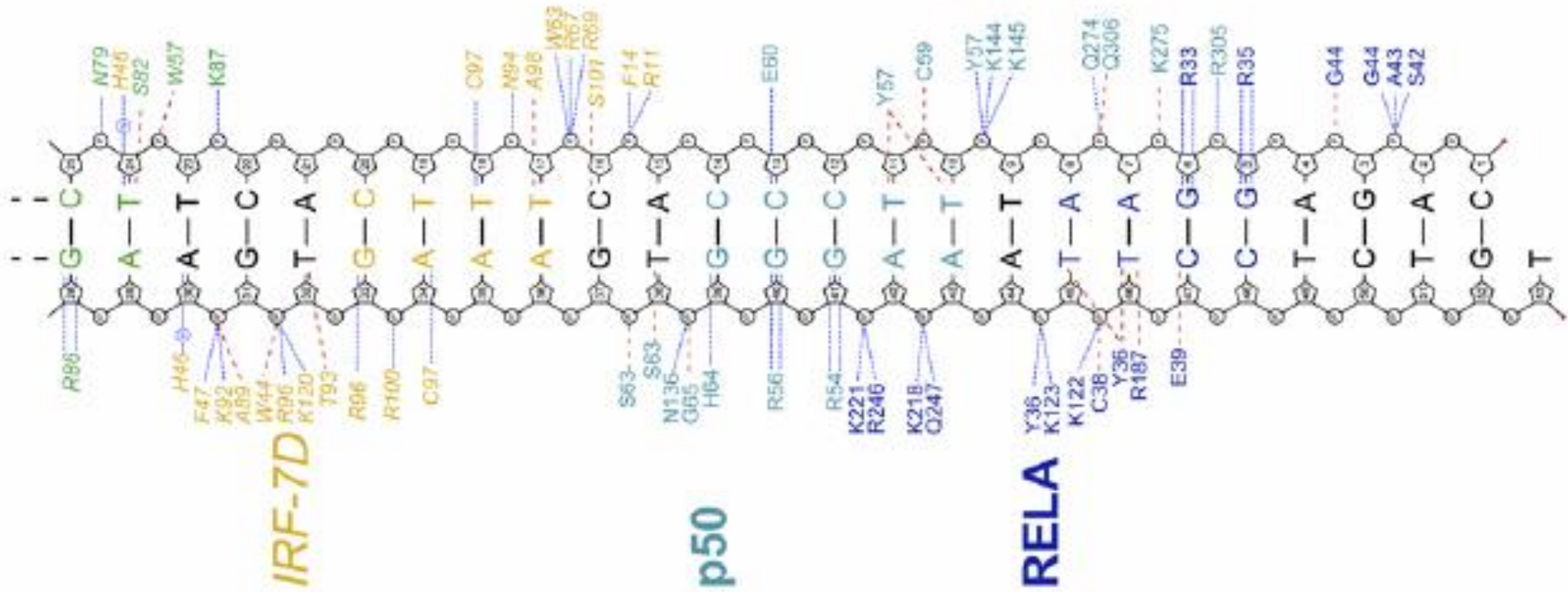
NF-κB

# IFN-β enhanceosom



transkripce

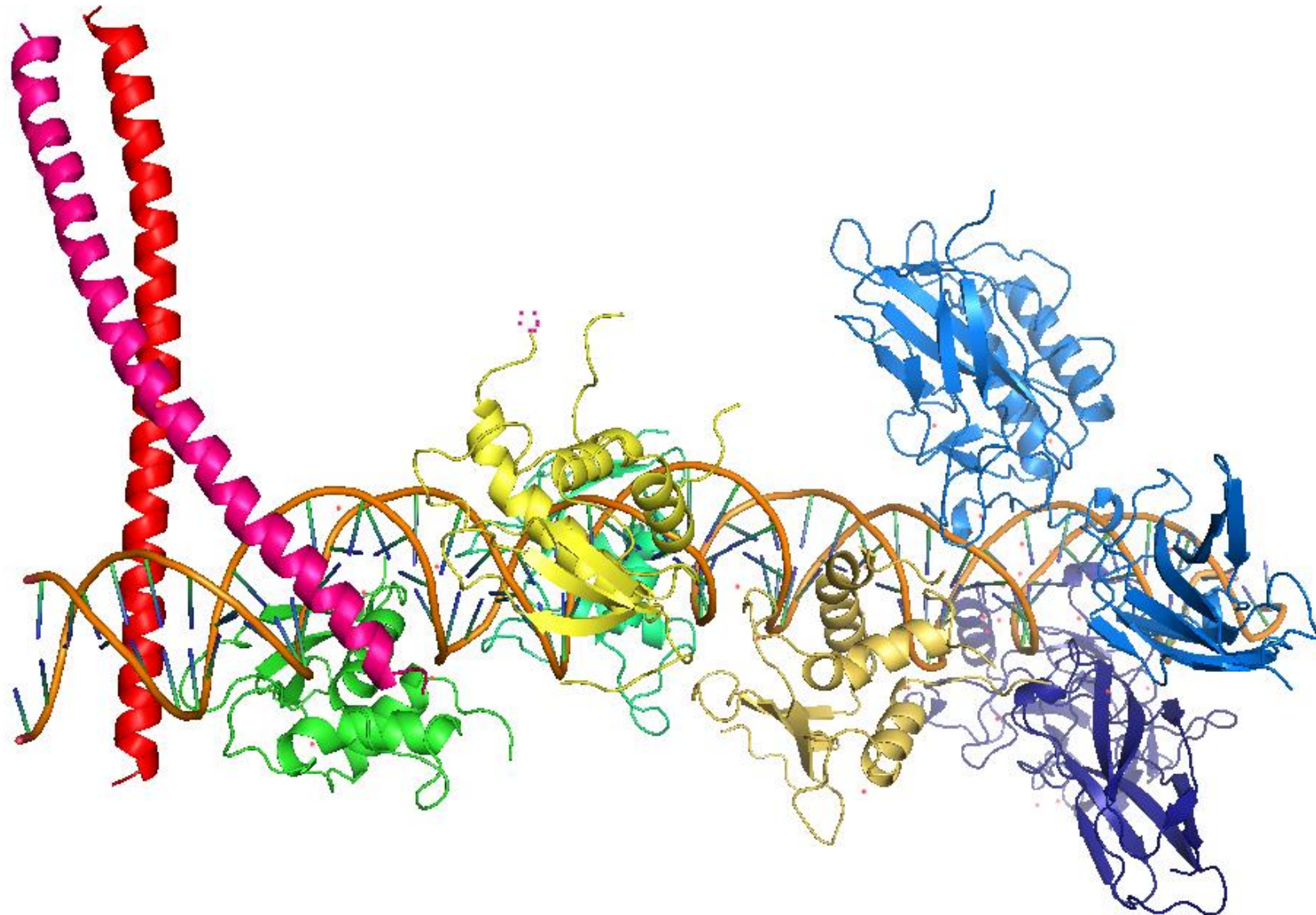
- koordinovaná vazba 8 protein (AP-1, IRF-3, IRF-7, NF-κB)
- p50/REL-A dimerizují (β-listy) - p50 slab interaguje s IRF-7
- vazba do Vž ō



NF-κB

Panne et al, Cell, 2008  
Panne, CO in SB, 2008

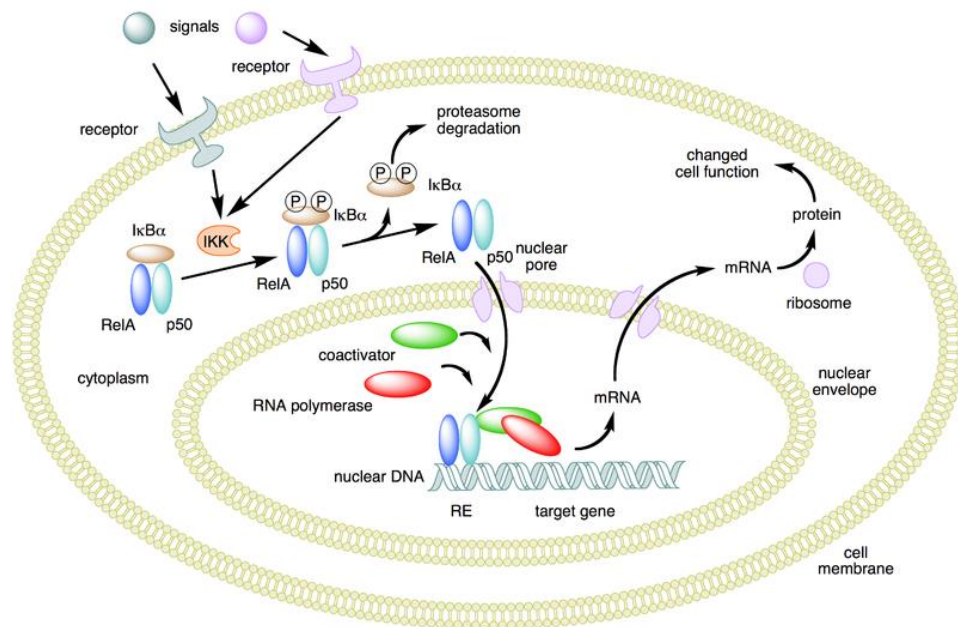
# IFN- $\beta$ enhanceosom



ervené te kv . molekulv vodv

<http://www.rcsb.org/pdb/101/motm.do?momID=122>

# CBP/p300



- enhanceosom interaguje (sp itáhne%) koaktivátory (CBP/p300 histon acetylasy), modifikuje chromatinovou strukturu ò umo0ní vazbu RNA polymerasy a iniciaci transkripce

## Souhrn:

- vazba v týiny TF pomocí ýroubovice ve velkém ýlábku (leucinový zip, HTH, zink-finger Å )
- transkrip ní komplexy (enhanceosom ...)

## P íýt :

- Histon, HMG-box
- $\beta$ -sheet motivy
- enhanceosom ... a po átek transkripce

