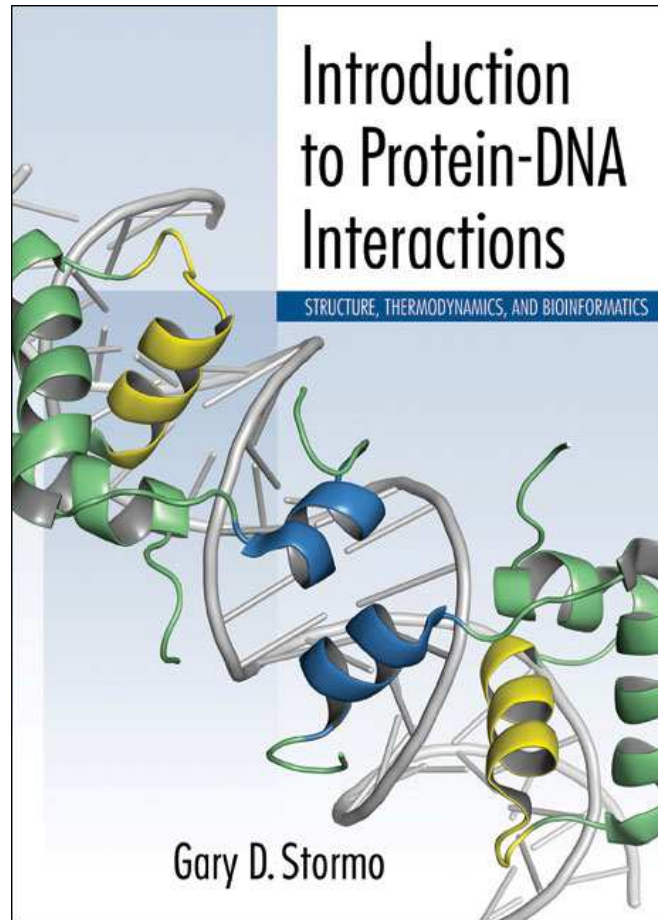


(DNA-)proteinové komplexy



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

DNA-proteinové komplexy

- Komplexy spojené s transkripcí (až 5% genomu)
- Komplexy spojené s duplikací genomu
- Komplexy podílející se na opravě genomu
- Chromatinové strukturní komplexy

G
E
N
O
M

The screenshot shows the NPIDB website. At the top, there is a navigation bar with 'Home', 'Browse', 'Download', 'Help', and 'About Us'. Below this is a search bar with a 'Search' button and radio buttons for 'PDB', 'Pfam', 'SCOP', 'GO terms', 'Fuzznuc', and 'BLAST'. The main content area is divided into two columns. The left column, titled 'NPIDB', contains a description of the database and a reference. The right column, titled 'List of complexes', contains information about the number of structures and details about individual complex pages.

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.

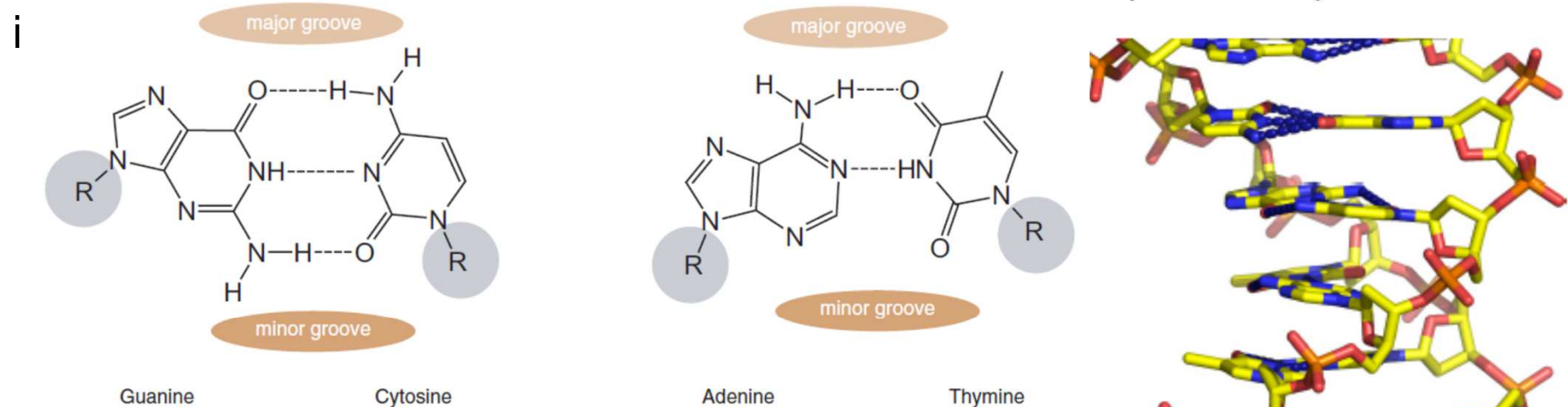
As of **27.11.2013** there are **4482** structures.

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)

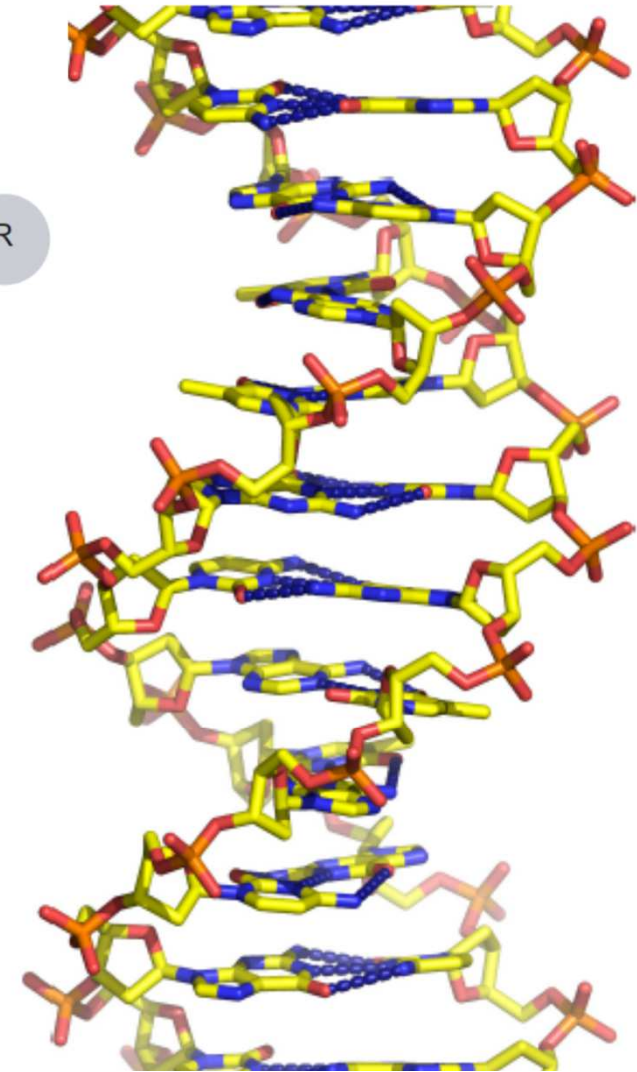
<http://npidb.belozersky.msu.ru/>

5000 struktur v PDB (v roce 2014)

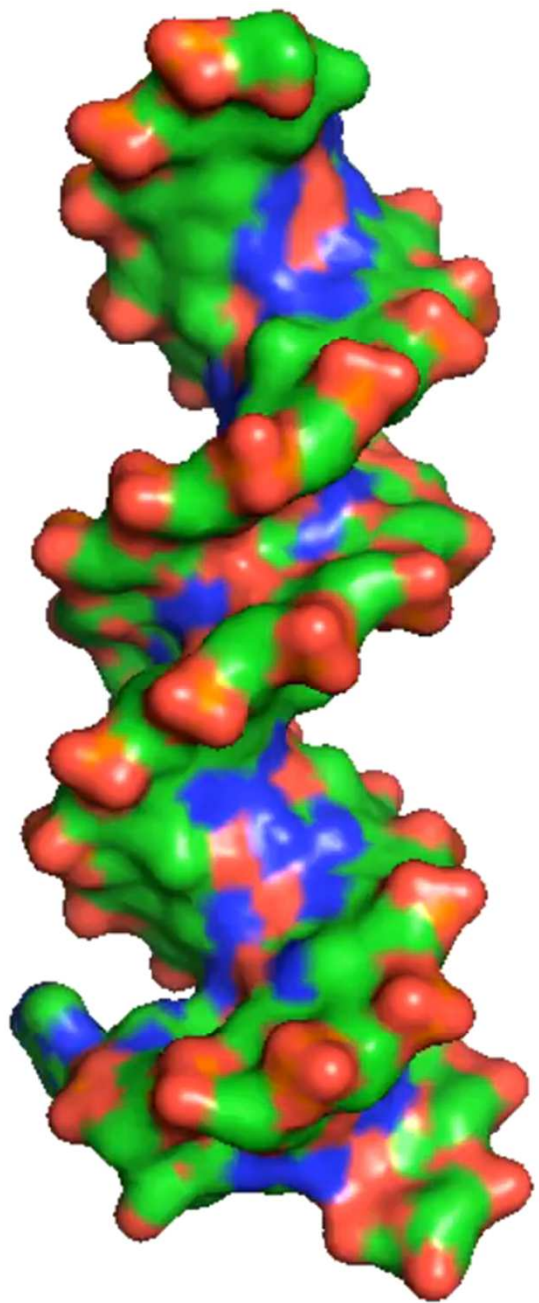
Tvarová a nábojová specifita DNA determinuje typy DNA-vazebných domén (oproti velké rozmanitosti protein-proteinových



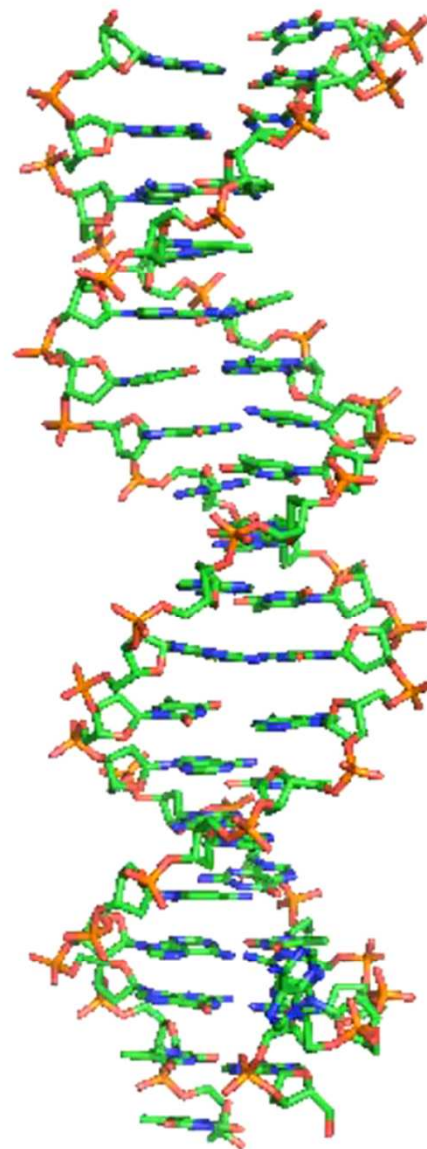
- proteiny interagují s cukrfofátovou kostrou (fosfát) nebo přes žlábký s bazemi (vod. vazba, tvar šroubovice)
- Interakce sekvenčně nespecifické (kostra – histony; strukturně specifické – HMG proteiny) nebo sekvenčně specifické (kostra+žlábký – kombinace: *Bgl*III (AGATCT) a *Bam*HI (GGATCC) kontaktují stejné báze a „čtou“ zakřivení okolní DNA ...)

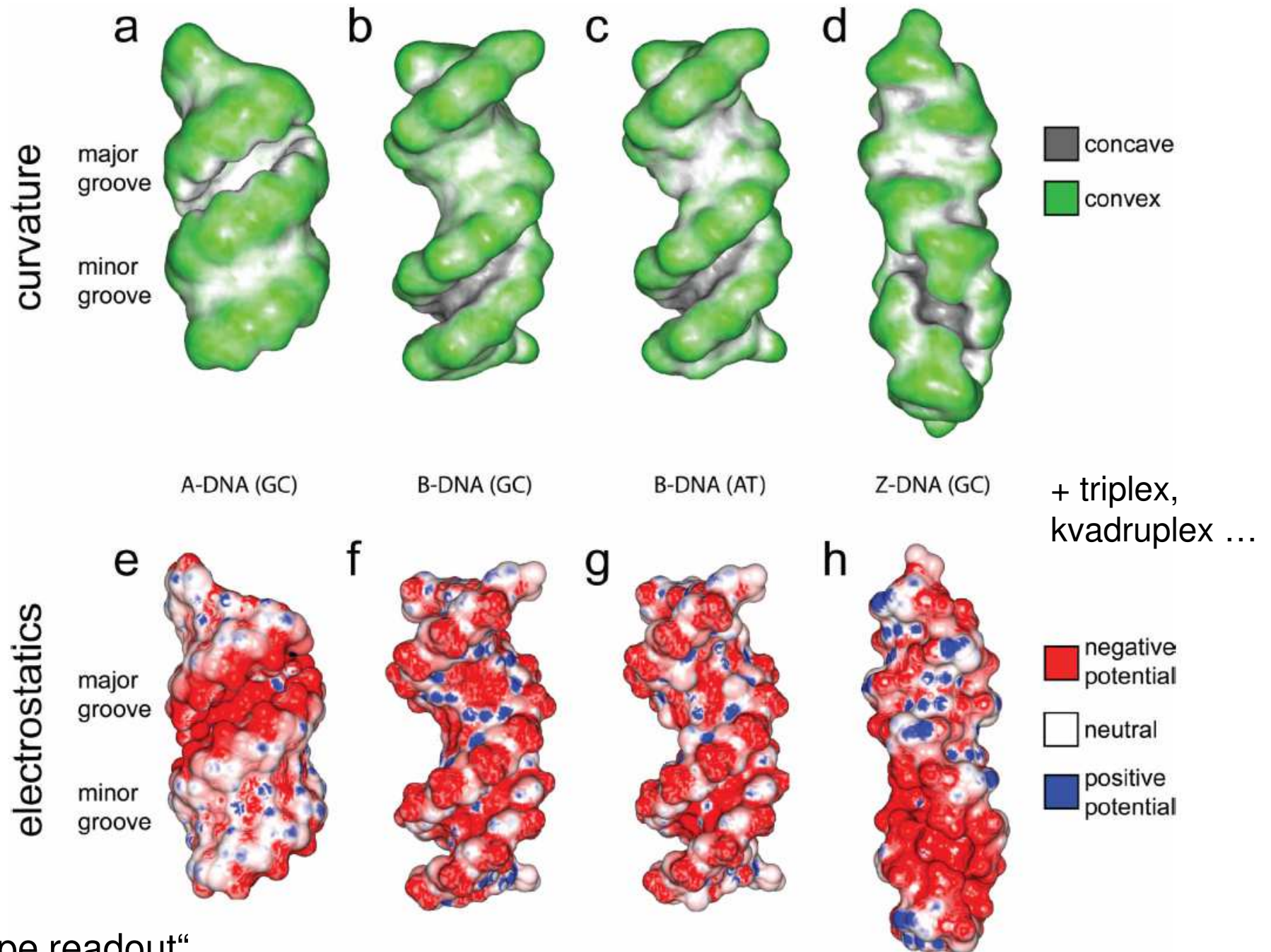


„shape readout“
zakřivení kostry souvisí se sekvencí



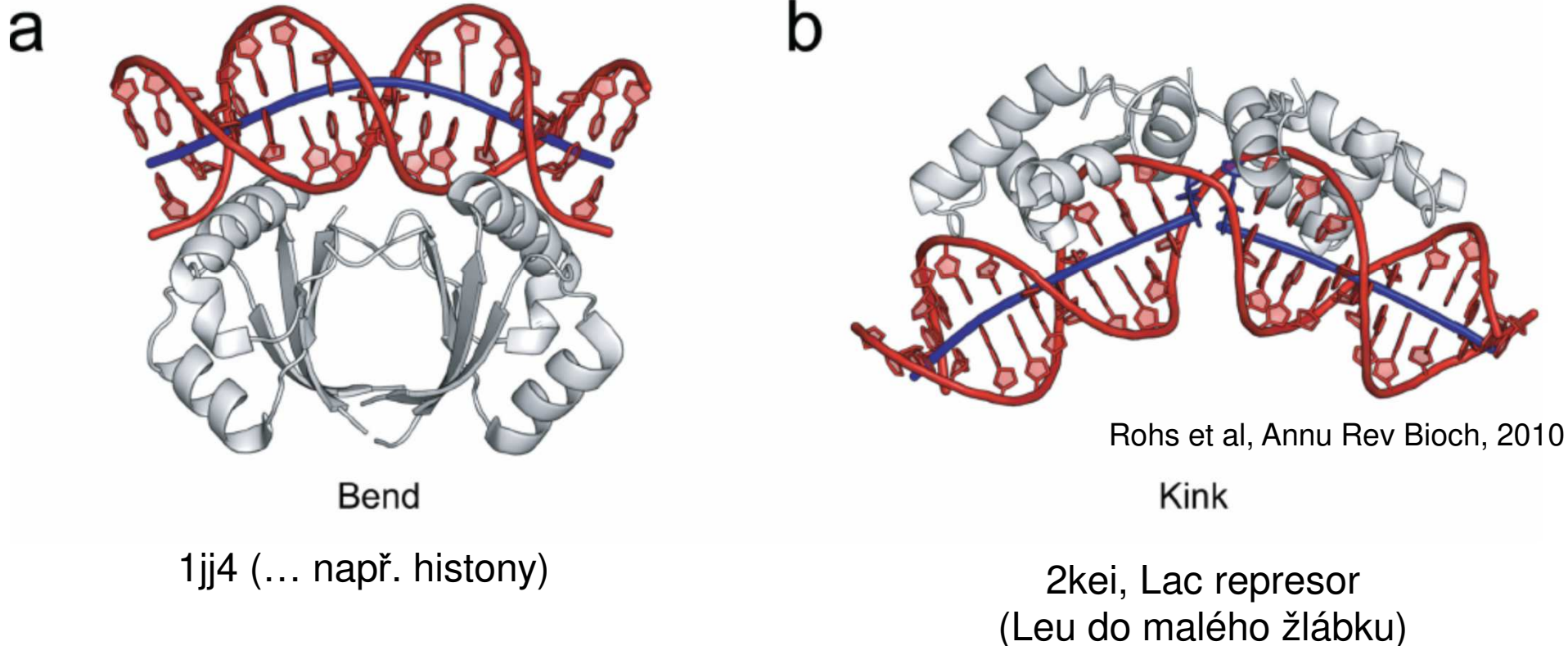
B-DNA





„shape readout“
zakřivení kostry - souvisí se sekvencí a prostředím

Vazba DNA-protein může indukovat změny



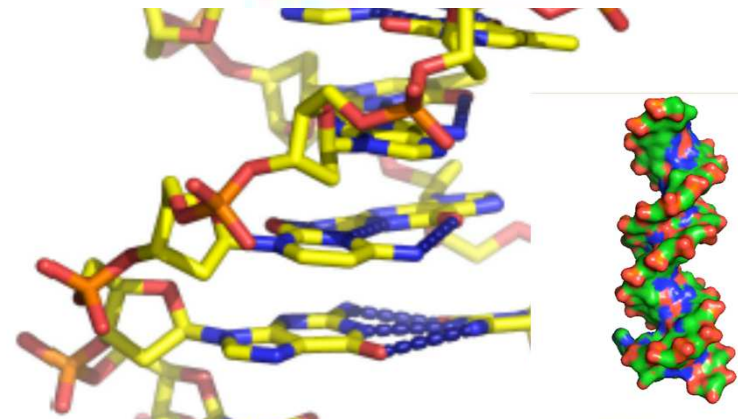
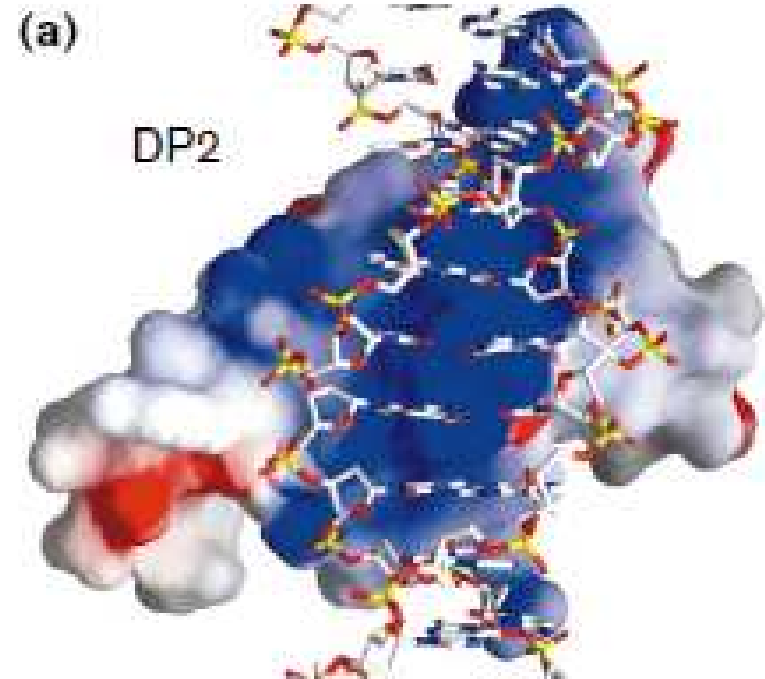
- vazba proteinu může indukovat změny ve struktuře DNA
- vazba DNA na protein často indukuje změny v jeho struktuře
- případně u nestrukturovaných proteinů strukturu indukuje (c-Jun/c-Fos = šroubovice až po navázání dimeru na DNA)

Vazba proteinů s DNA prostřednictvím solných můstků

- 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

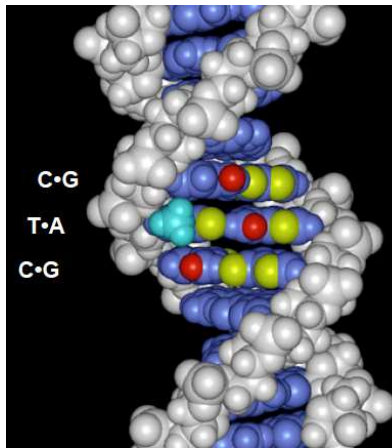
Table 2. The simplified relative charge set defined from CHARMM (10) used in the calculation of the electrostatic potential of atoms in the DNA-binding proteins

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

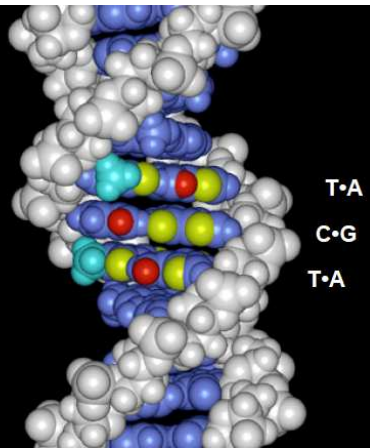


Gajiwala & Burley, COiSB, 2000

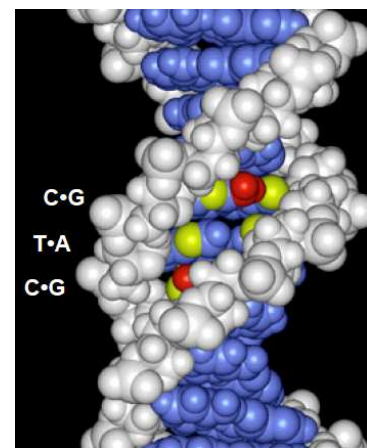
Major groove



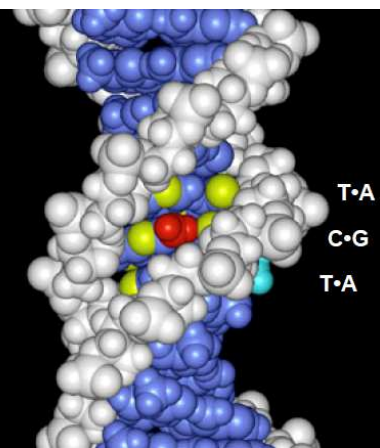
Major groove



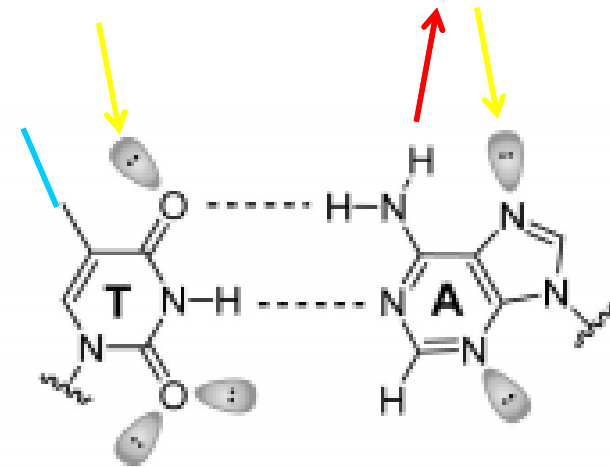
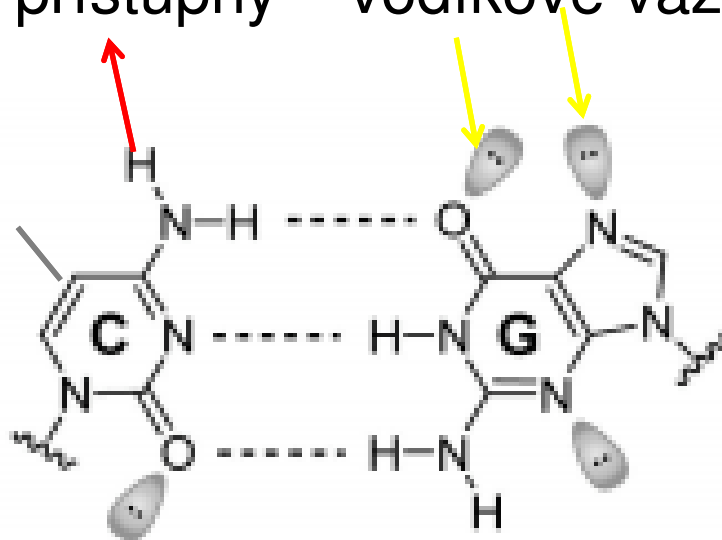
Minor groove

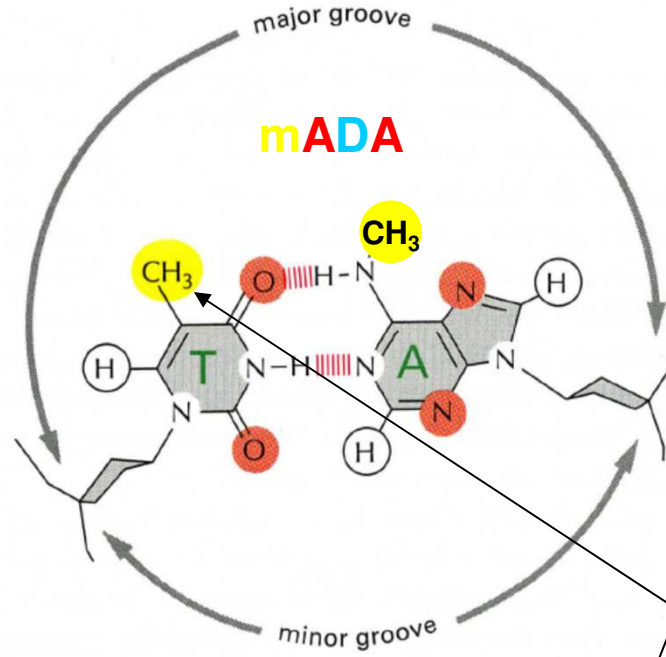
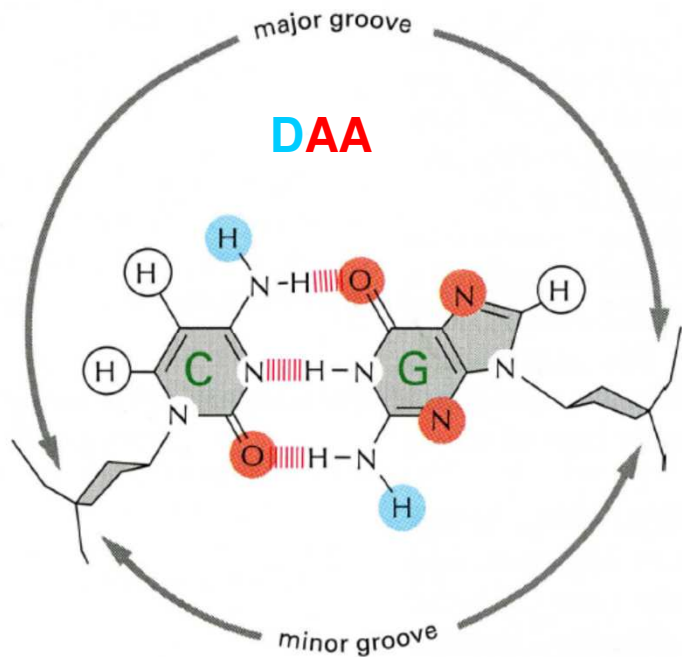


Minor groove



sekvenčně-specifický protein kontaktuje báze („direct“ readout) – skrze velký nebo malý žlábek – **velký žlábek** je lépe přístupný – vodíkové vazby (donor vs akceptor elektronu)

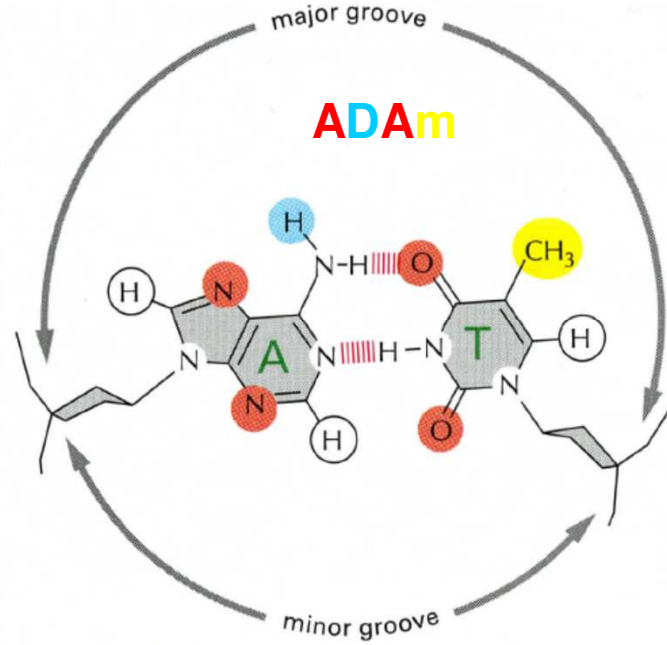
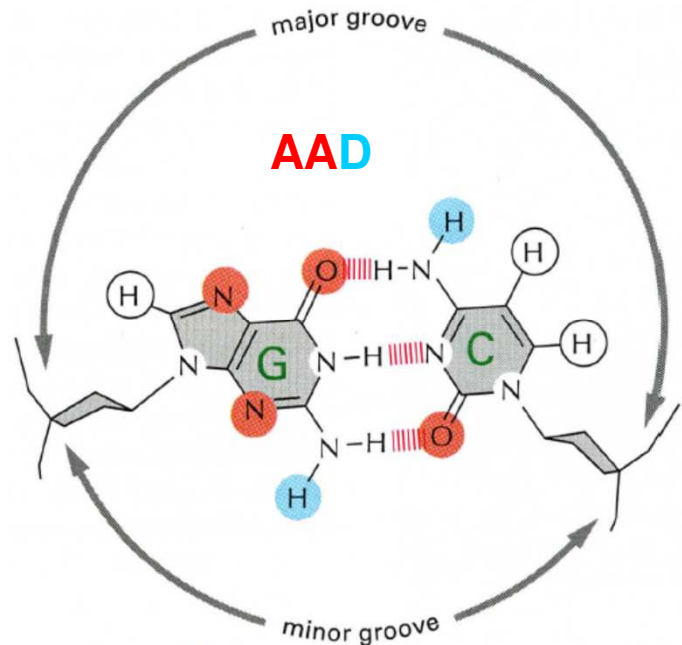




Jak odliší
protein různé
páry bazí?


“base readout”

Pozice donor vs
akceptor
+ metyl skupina



Methylace Ade
(C6NH₂) u bakterií
změna!

KEY:

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

Vazba proteinů s DNA prostřednictvím vodíkových vazeb

- **Velký žlábek** má velikost odpovídající rozměrům α -šroubovice a má exponované H-vazebné skupiny
- Ade zbytky C-6(NH₂) a N-7 mohou tvořit specifické vodíkové vazby s **Gln** a **Asn**
- Gua může tvořit specifické vodíkové vazby s **Arg**

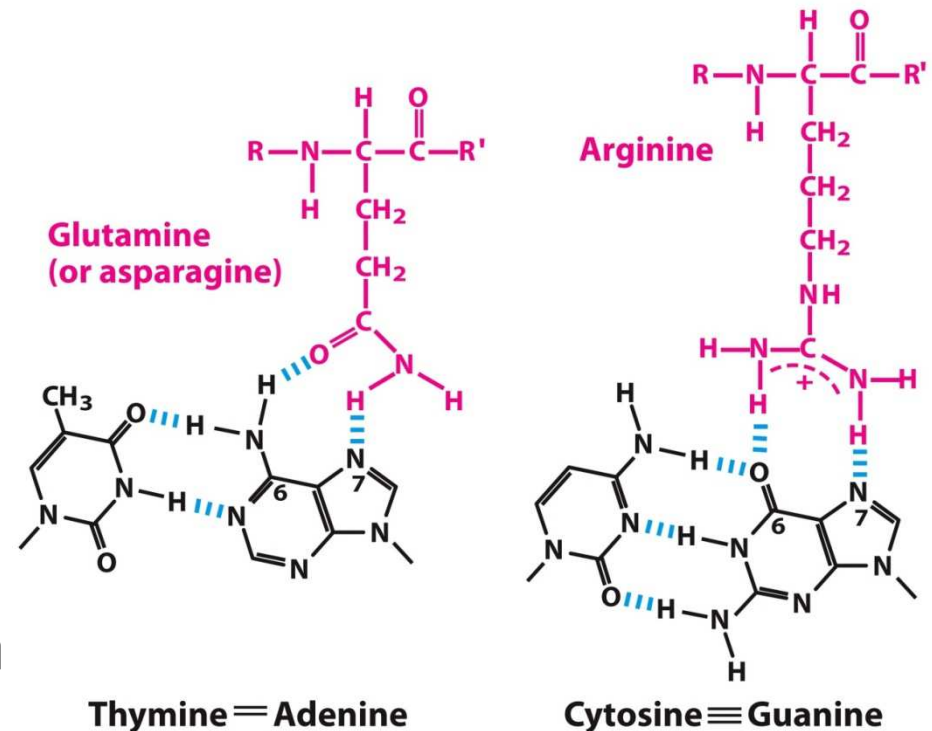
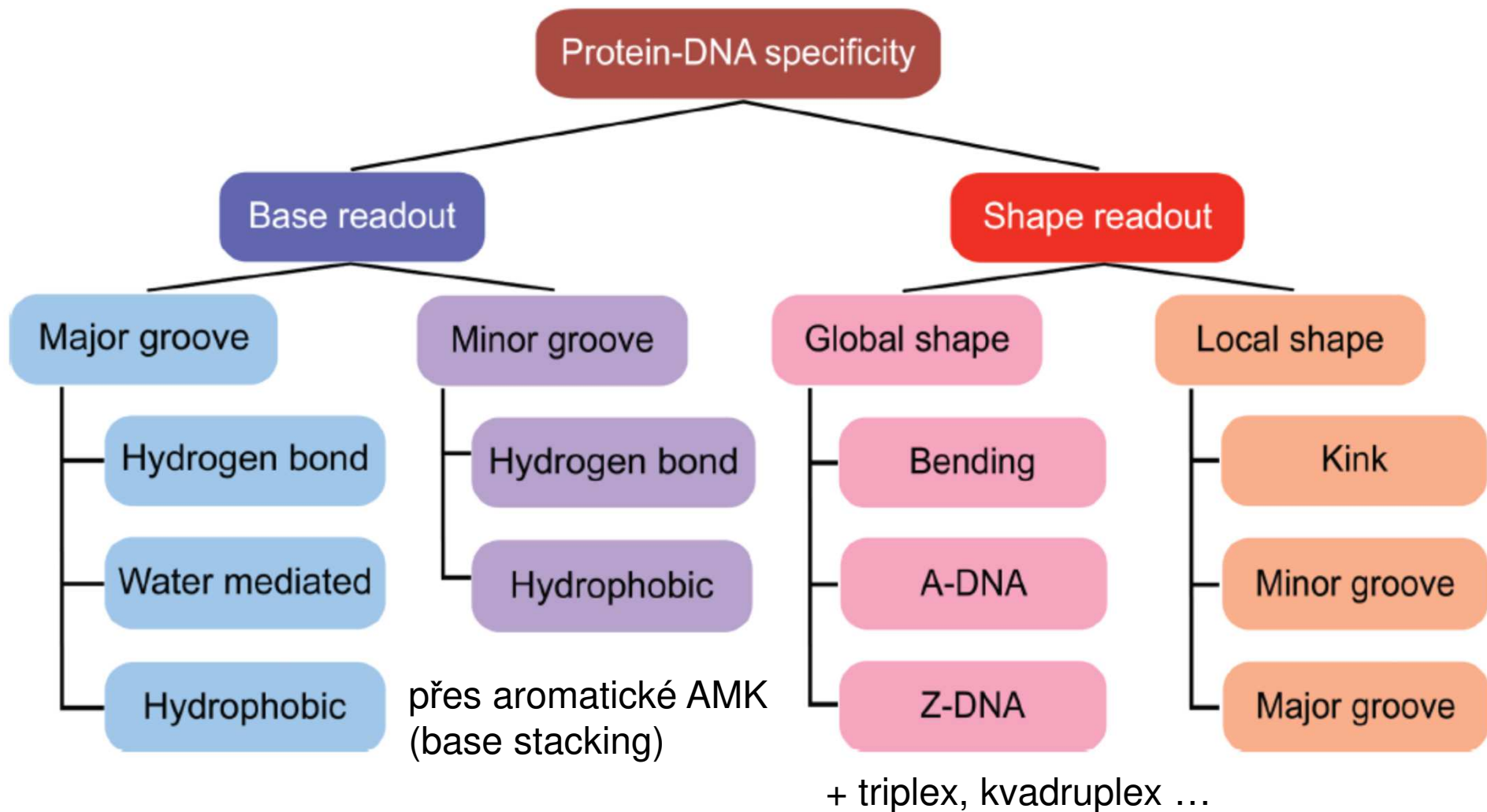


Figure 28-10
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Silná vazba, sekvenčně specifická - afinita nM – μ M
Slabá vazba, strukturně specifická - afinita μ M – mM



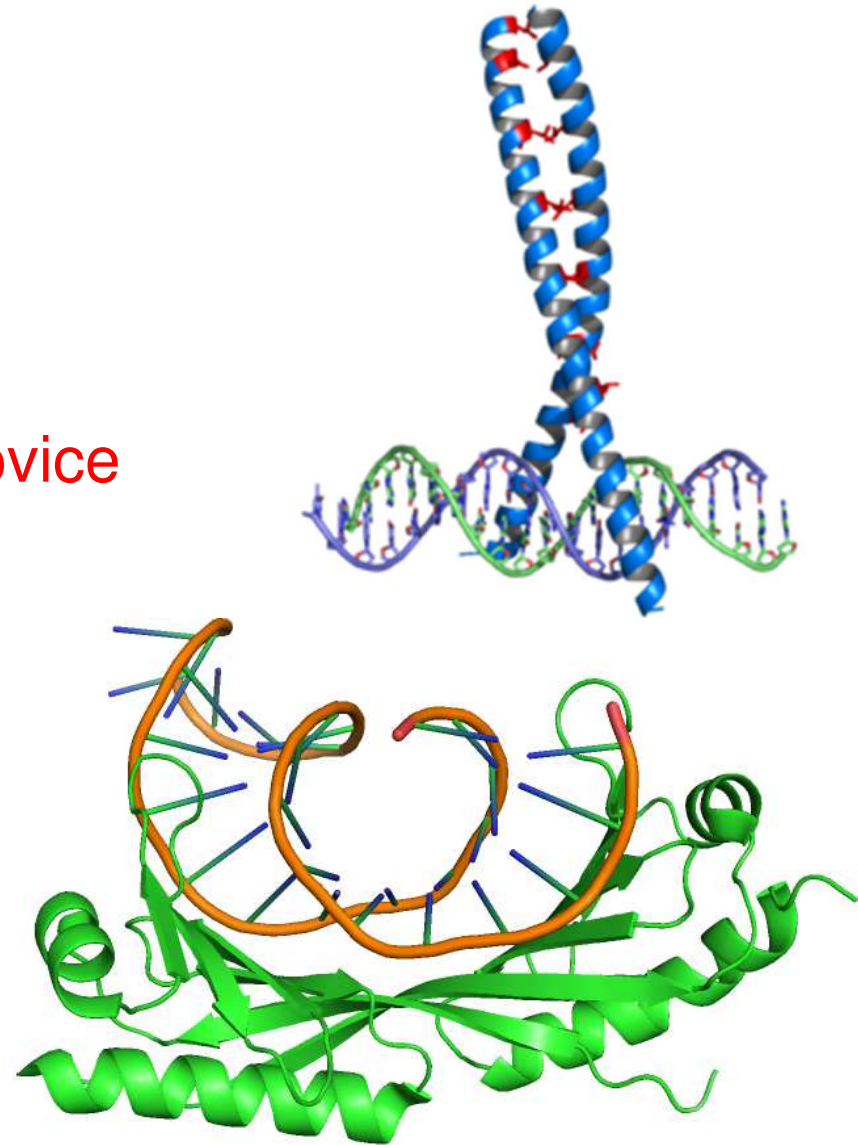
- více jak 70 SCOP superrodin (strukturních motivů)
- dle sekundárních struktur – α -šroubovice (17), β -listy (7), smíšené α/β motivy (48)

Motivy DNA-vazebných domén

- **Zipper typ**
 - Leucinový zip
 - Helix-loop-helix
- **Helix-otáčka-helix**
 - HTH
 - Winged helix
 - TALE
- **Zinkový prst**
 - $\beta\beta\alpha$ zinc-finger
 - Hormon-receptor
 - Loop-sheet-helix
 - Gal4
- **Histon, HMG-box**
- **β -sheet motivy**

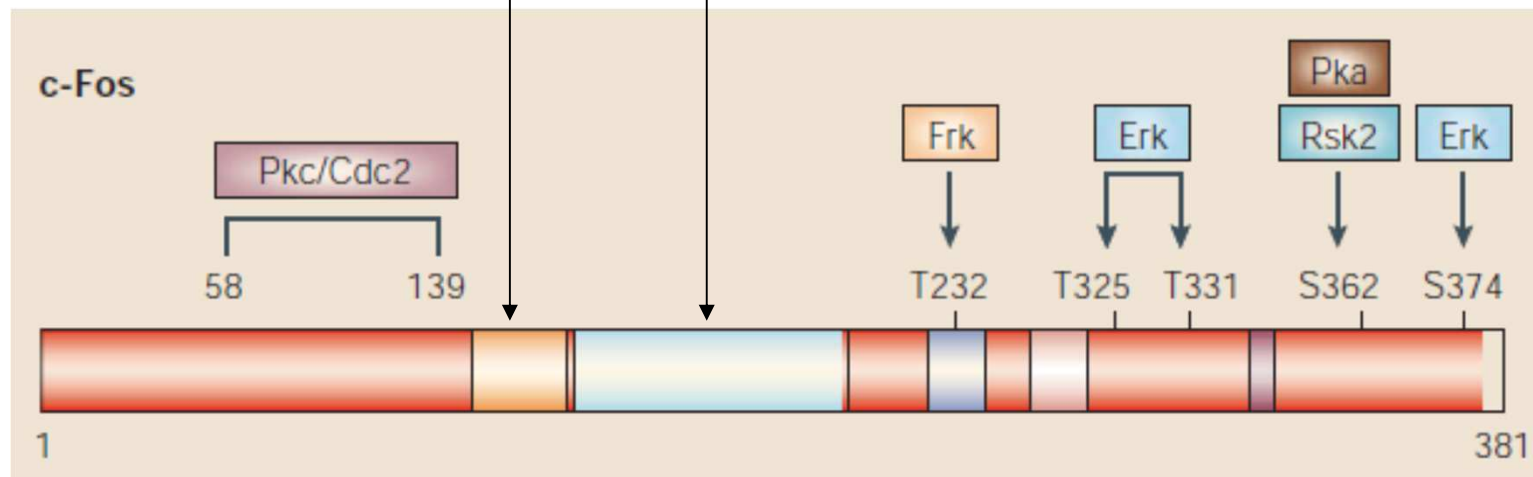
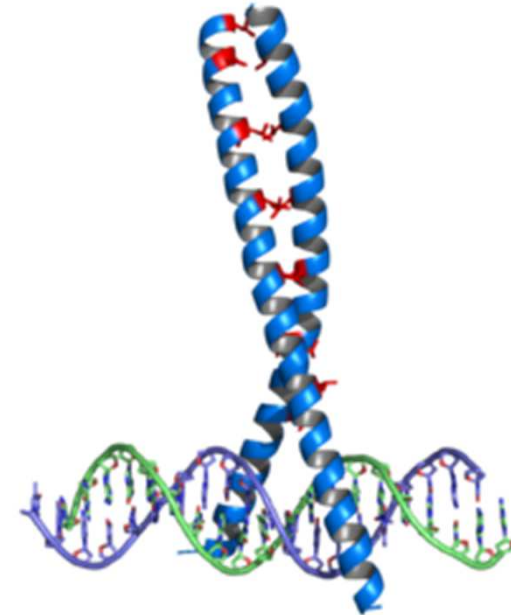
α -šroubovice

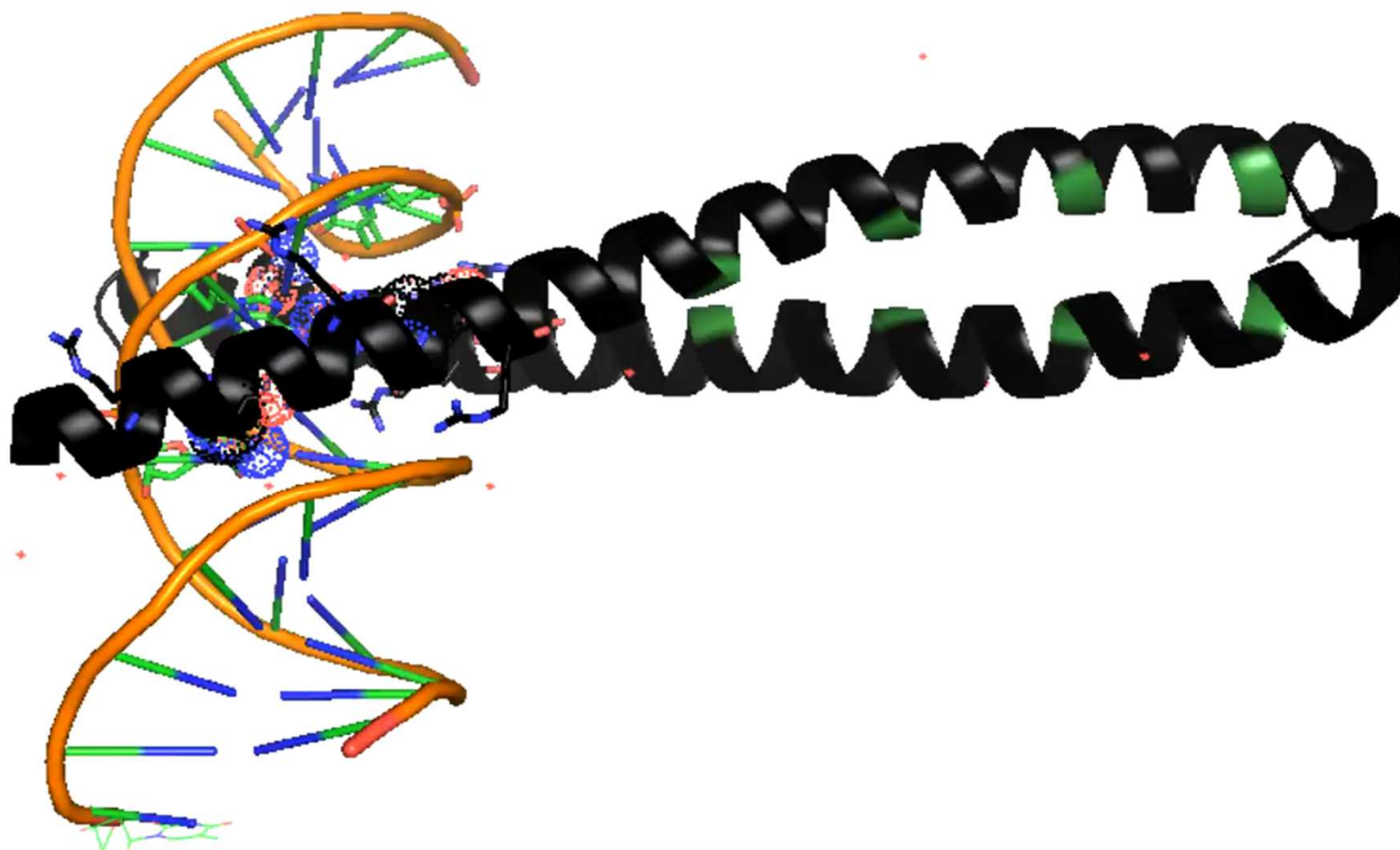
β -listy



Motivy DNA vazebných domén

- **Zipper typ** (dle způsobu dimerizace)
 - **Leucinový zip** (tzv. bZIP = basic)
(transkr. fakt. γ GCN4, c-Jun/c-Fos=AP-1)
 - 2 α -šroubovice
 - coiled-coil (>30AMK, Leu, C-term)
 - bazická část (N-terminus, navazuje na CC)
 - bazická šroubovice vázána do VŽ



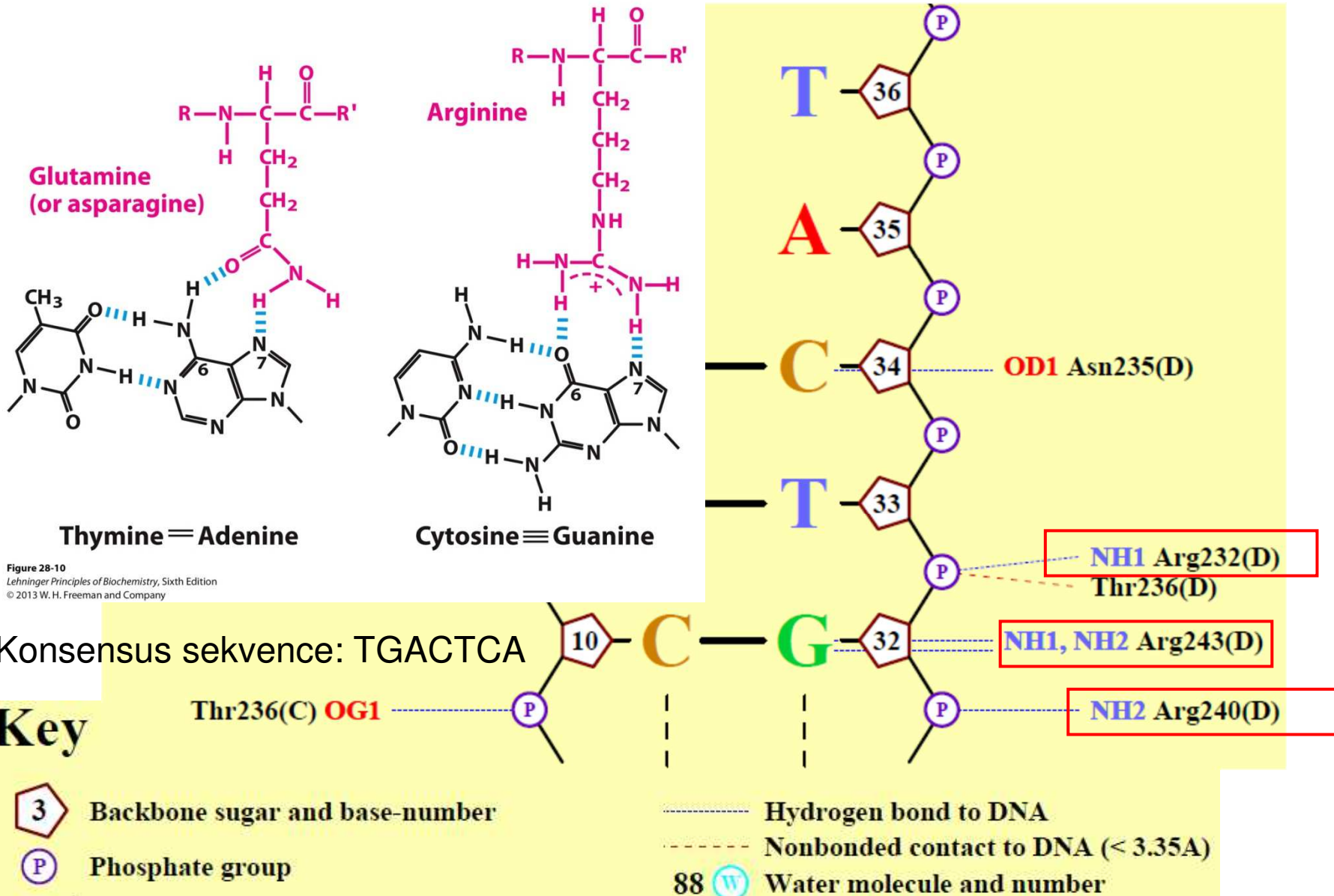


Interakce bazických AMK: Arg(232+240)=PO₄, Arg(243)=Gua

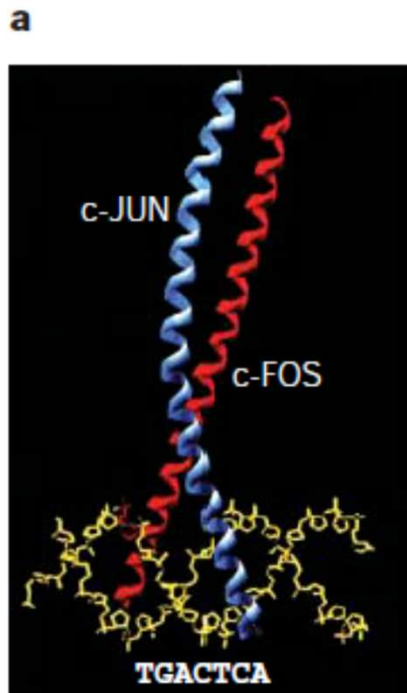
Konsensus sekvence: TGACTCA

GCN4 – regulace genů pro syntézu AMK

PDB: 1YSA



Jones a spol., NAR, 2003



AP-1

b

TRE: TGACTCA
ACTGAGT

CRE: TGACGTCA
ACTGCAGT

MARE I: TGCTGACTCAGCA
ACGACTGAGTCGT

MARE II: TGCTGACGTCAGCA
ACGACTGCAGTCGT

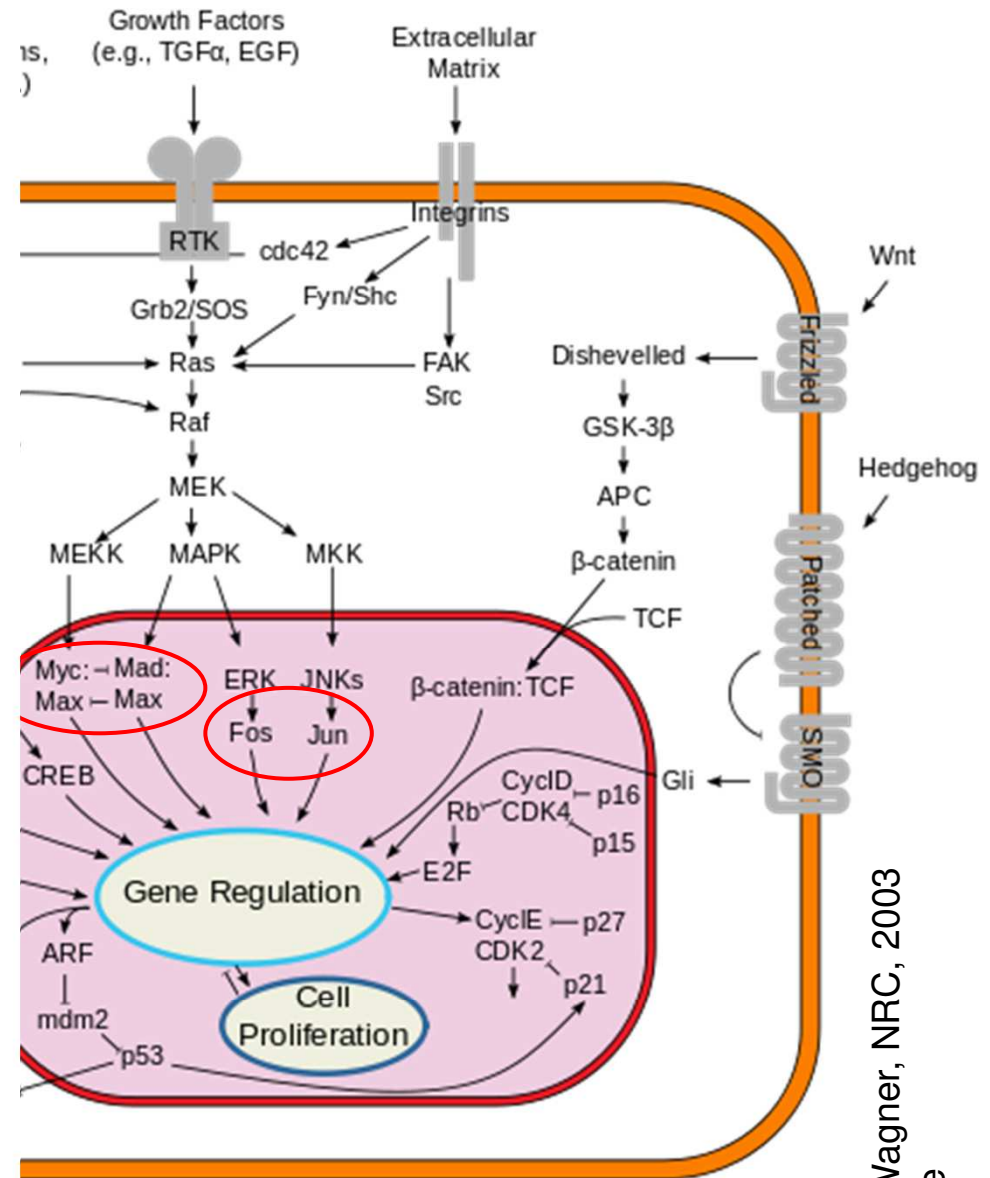
ARE: a/gTGACnnnGC
t/cACTGnnnCG

c homo/hetero

c-JUN	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

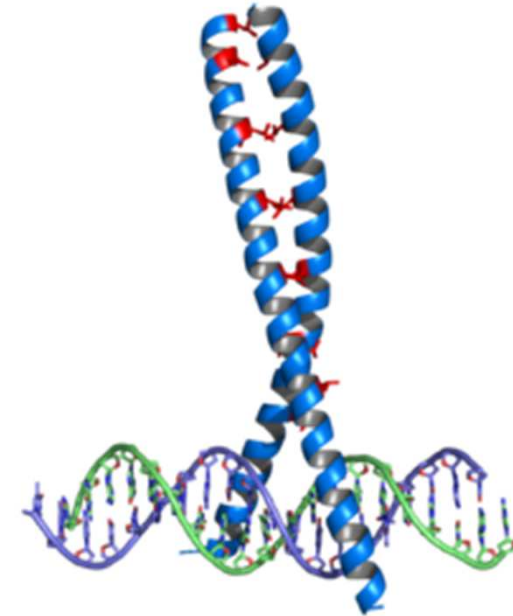
c-FOS	c-JUN	(TRE > CRE)
	JUNB	(TRE > CRE)
	JUND	(TRE > CRE)
	ATFa	(No binding)
	ATF2	(CRE > TRE)
	ATF4	(CRE)
	c-MAF	(MARE I/II)
	MAFA	(MARE I/II)
	MAFB	(MARE I/II)
	NRL	(TRE-related)
	MAFF/G/K	(MARE I/II)
	NRF2	(ARE)
	NFIL6	(TRE)



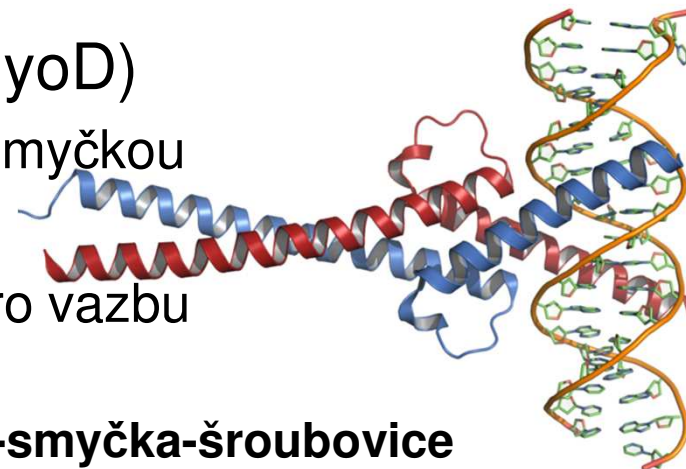
Kombinace – rŭzna specifita/afinita

Motivy DNA-vazebných domén

- **Zipper typ** (dle způsobu dimerizace)
 - **Leucinový zip** (tzv. bZIP = basic – transcr. fact. γ GCN4, c-Jun/c-Fos)
 - 2 α -helixy (2 x 60 AMK)
 - coiled-coil (30AMK, Leu, C-term)
 - bazická část (N-terminus, navazuje na CC)
 - bazická šroubovice vázána do VŽ



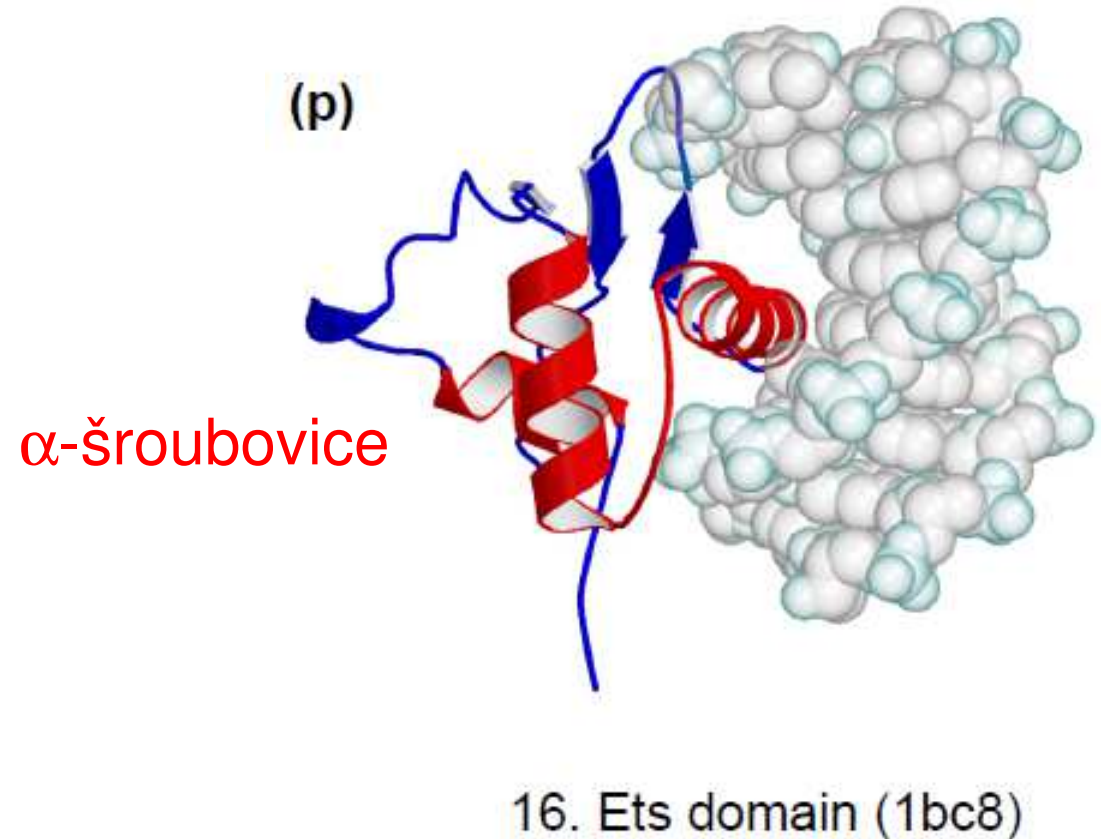
- **Helix-loop-helix** (c-Myc/Max, MyoD)
 - CC a bazické části jsou odděleny smyčkou
 - bazická šroubovice vázána do VŽ
 - smyčka poskytuje větší flexibilitu pro vazbu



šroubovice-smyčka-šroubovice

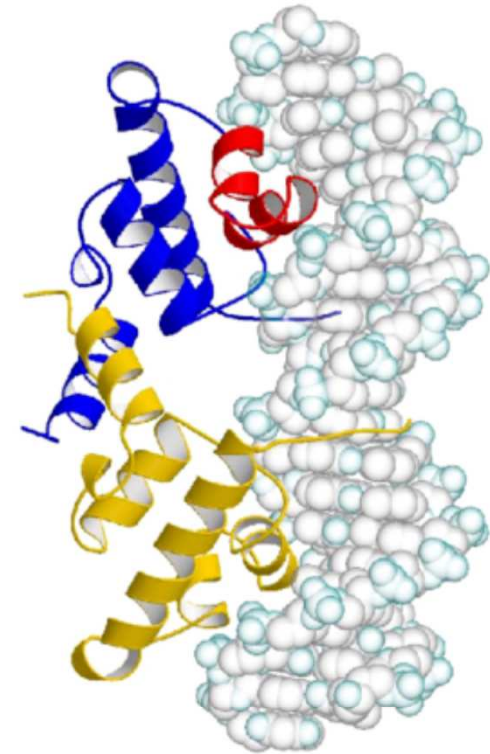
Motivy DNA vazebných domén

- **Zipper typ**
 - Leucinový zip
 - Helix-loop-helix
- **Helix-otáčka-helix**
 - HTH
 - Winged helix
 - TALE
- **Zinkový prst**
 - $\beta\beta\alpha$ zinc-finger
 - Hormon-receptor
 - Loop-sheet-helix
 - Gal4

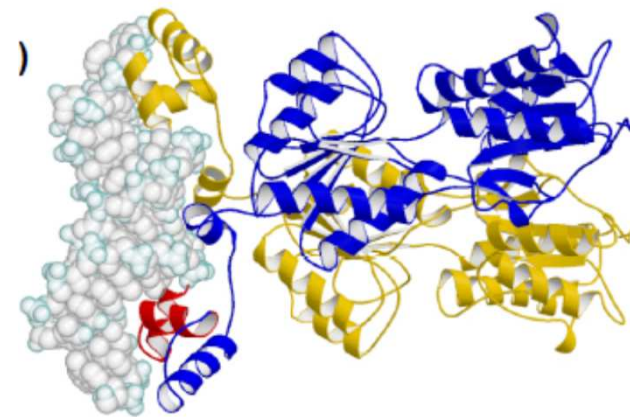


Helix-turn-helix motiv (HTH)

- Obsahuje ~ 20 AMK ve dvou šroubovicích vzájemně kolmých
 - α -helix pro vazbu na DNA („recognition“) - β -obrátká – druhá šroubovice
 - Sekvenčně-specifická vazba prostřednictvím „recognition“ šroubovice a velkého žlábků
 - nejčastější motiv u prokaryot - homodimery vážou palindrom. sekvence
 - HTH motiv se obvykle vyskytuje ve svazku 3-6 šroubovic (stabilizovaných hydrofobním jádrem)
 - motiv může být buď součástí hlavního proteinu (Cro) nebo z něj může pouze vybíhat (LacI)

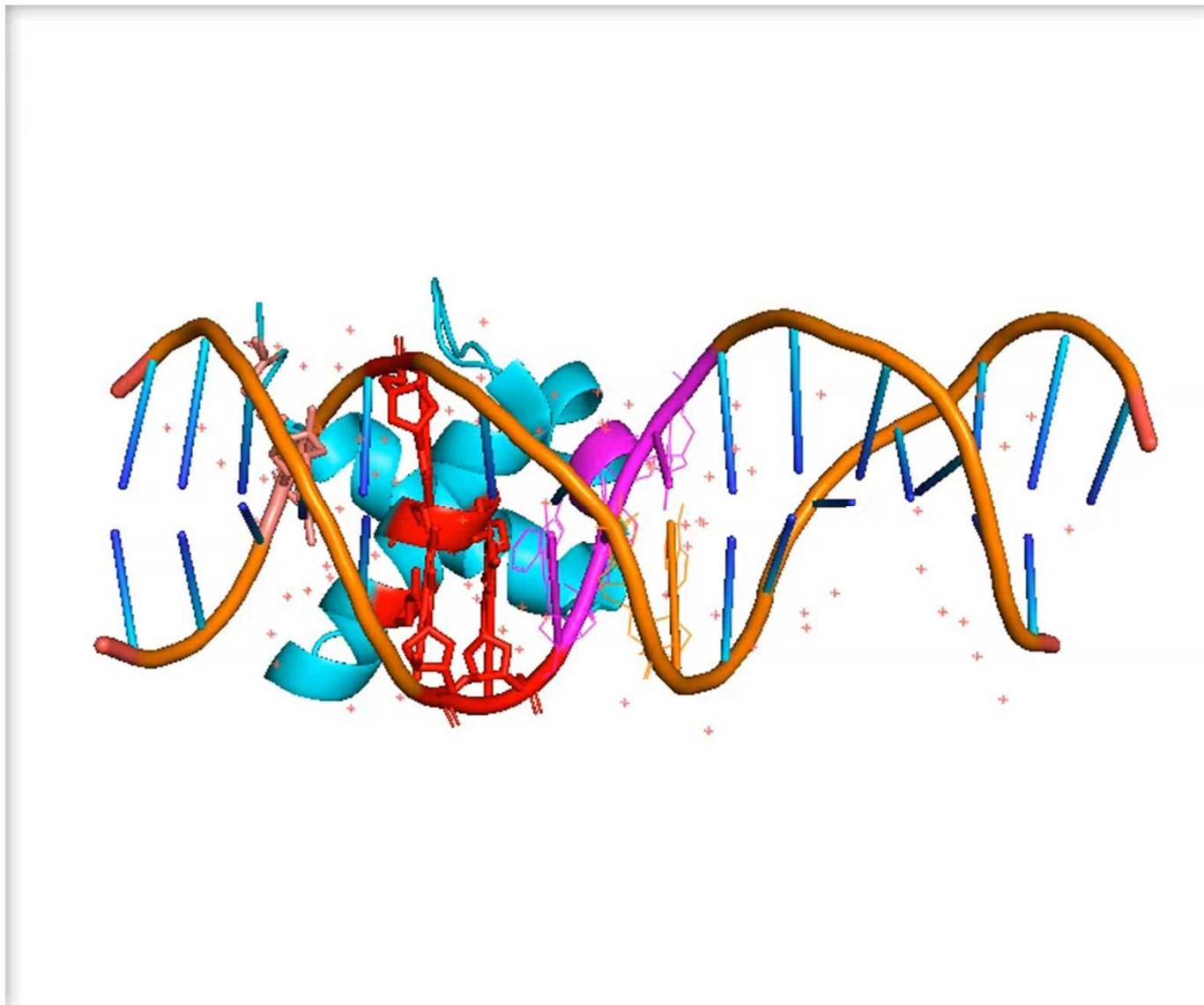


1. Cro and Repressor (1lmb)



Luscombe et al, Genome Biology, 2000

3. LacI repressor (1wet)



AMK motiv: **KDRWR**

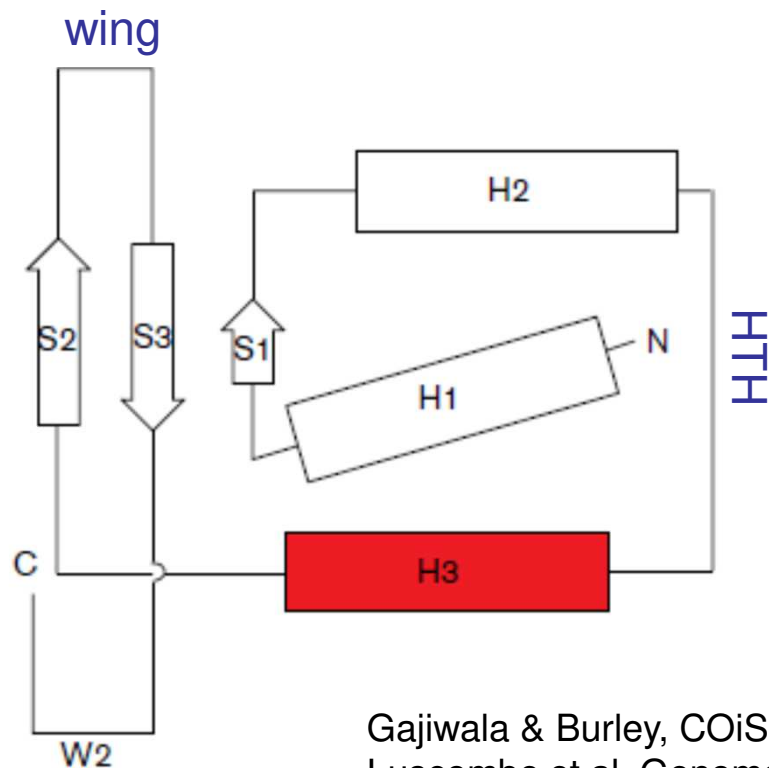
DNA sekvence: TTAG**GGG** (komplementární: CC**CTAA**)

TRF2 (myb) doména – vazba na telomery (Shelterin komplex)

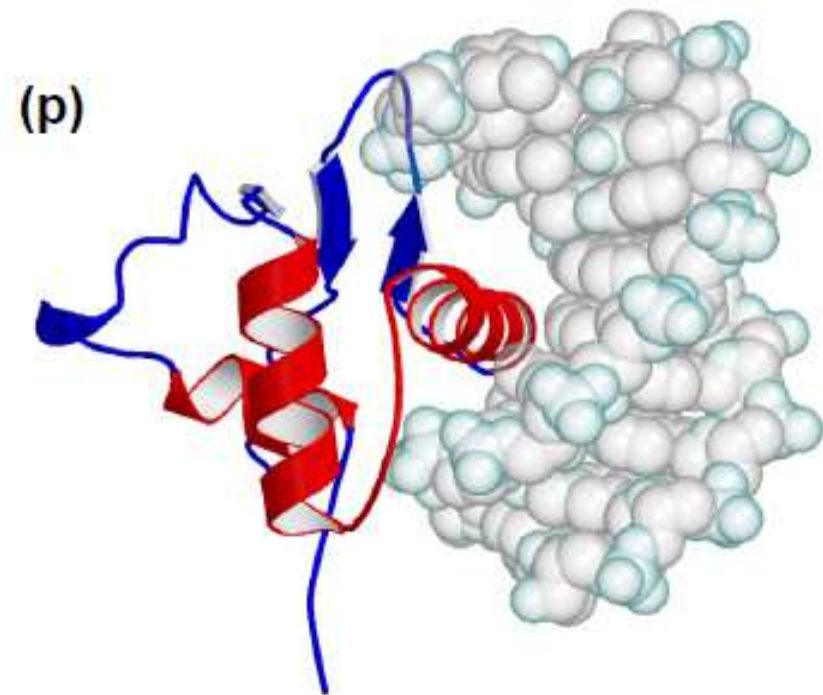
PDB: 1W0U

„Winged“ helix (okřídlená šroubovice)

- „winged“ HTH obsahuje „recognition“ šroubovici (**H3**) a β -listy, které poskytují další kontakty s DNA (smíšený α/β typ)



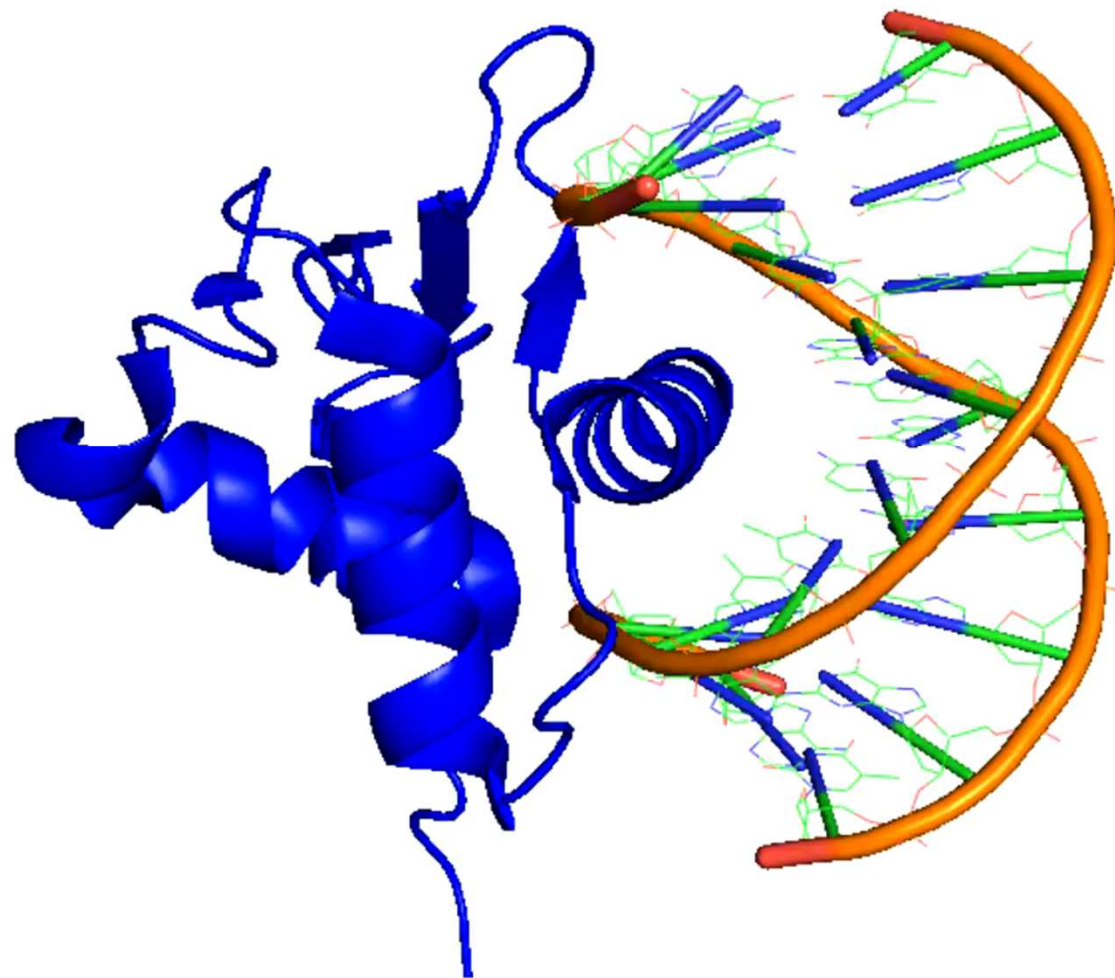
Gajiwala & Burley, COiSB, 2000
Luscombe et al, Genome Biology, 2000



16. Ets domain (1bc8)

Méně často křídlo ve VŽ a cukr-fosfátová kostra se šroubovicí (hRFX1)

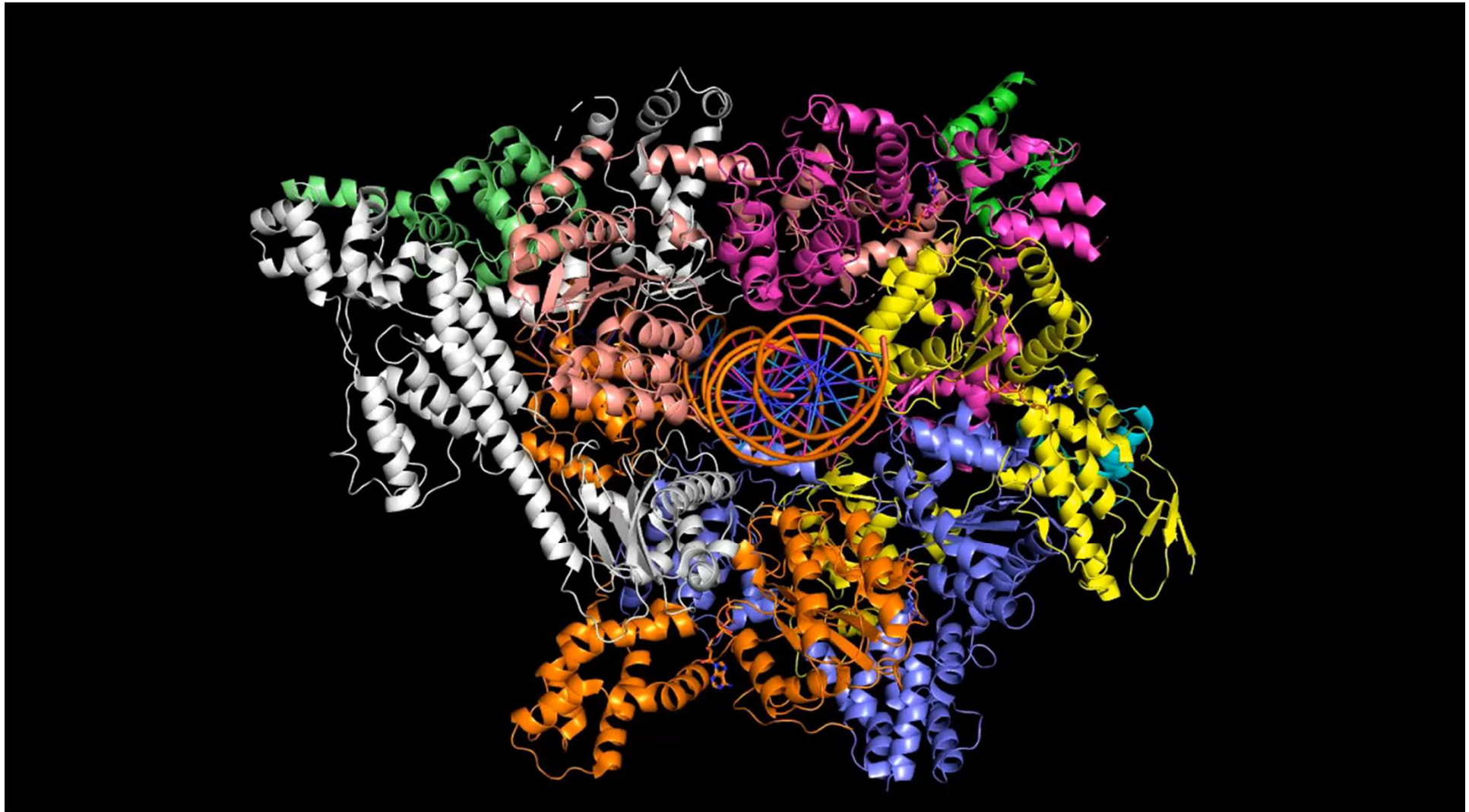
Interakce bazí se šroubovicí (H3) a křídla s cukr-fosfátovou kostrou



PDB: 1BC8, SAP-1

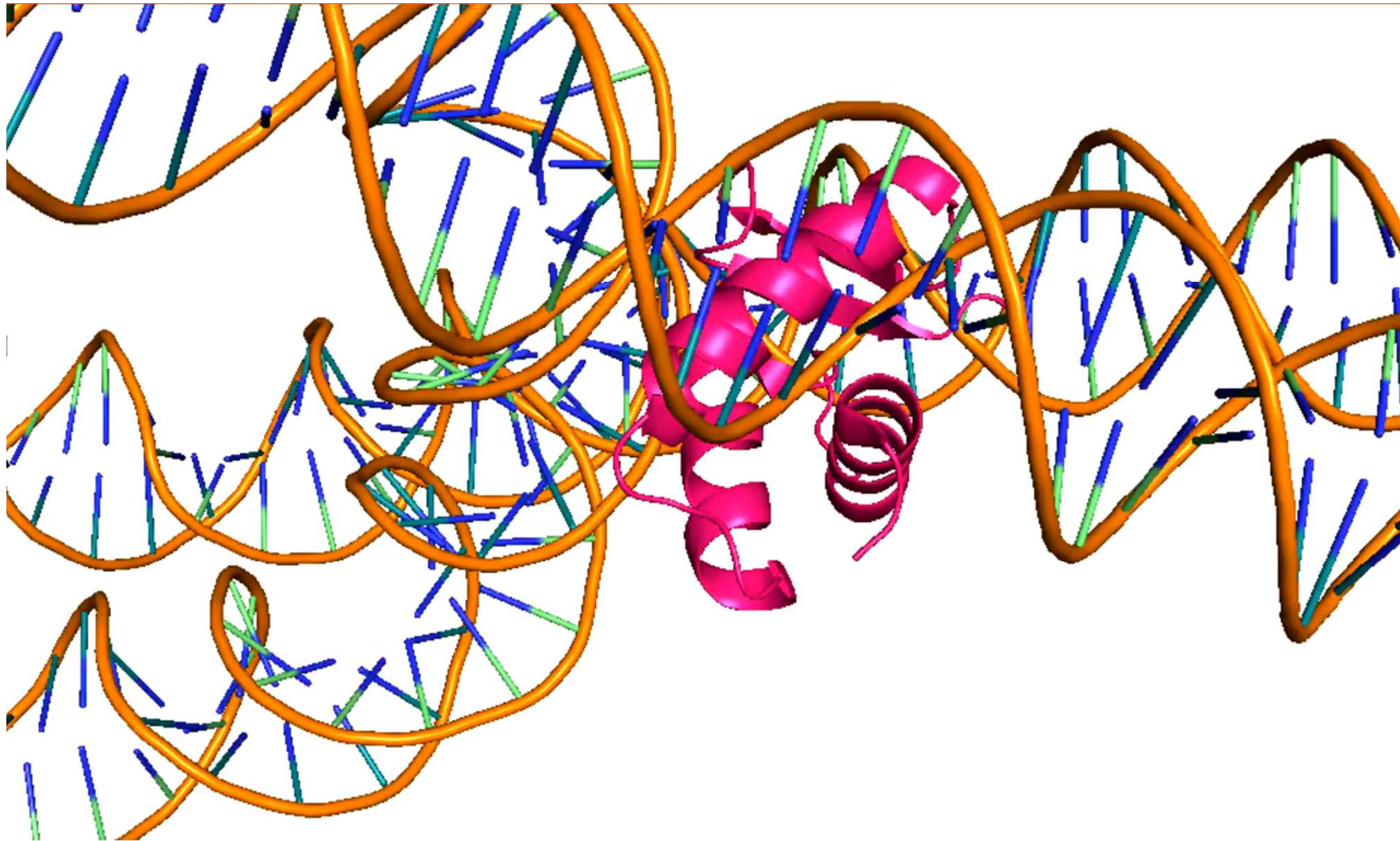
Méně často křídlo ve VŽ a cukr-fosfátová kostra se šroubovicí (hRFX1)

WH v ORC komplexu



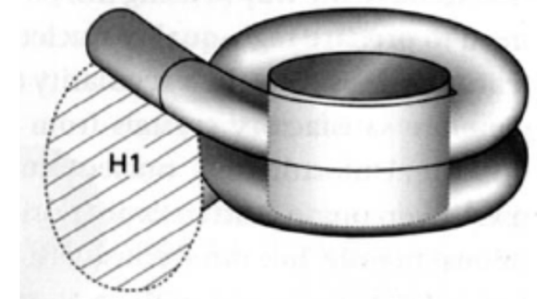
vazba WH přes křídla (špičky tvoří tunel; 6WGC)

Jaremko et al, eLife, 2020



PDB: 5NLO

Histon H1/H5 interaguje s DNA vybíhající z nukleosomů (kompaktnější struktura) – WH doména může vytvářet více kontaktů (intergaují: H2, H3 i křídlo s 3 DNA řetězci)



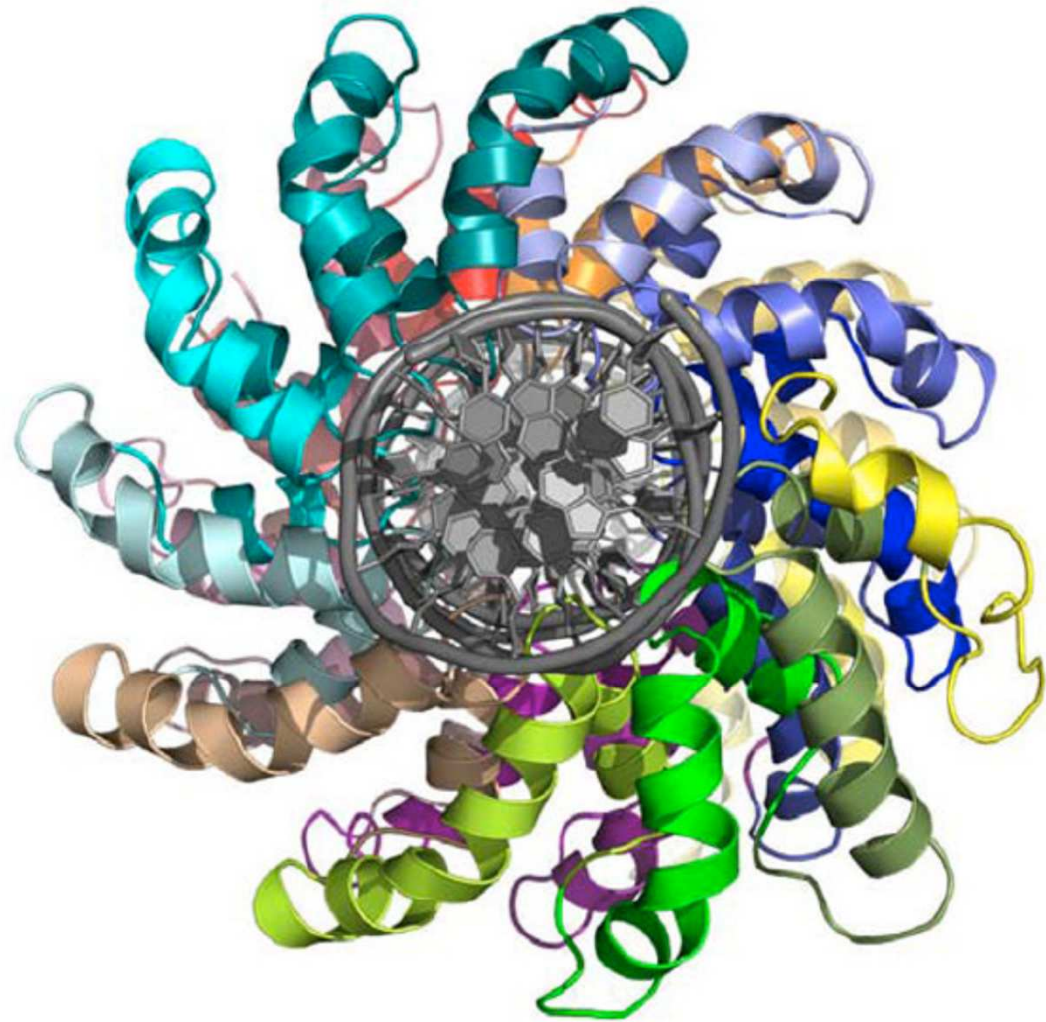
Transcription activator-like effectors (TALE)

Patogenní bakterie injikují do rostlinných buněk
ovlivňují transkripci hostitelských rostlinných promotorů



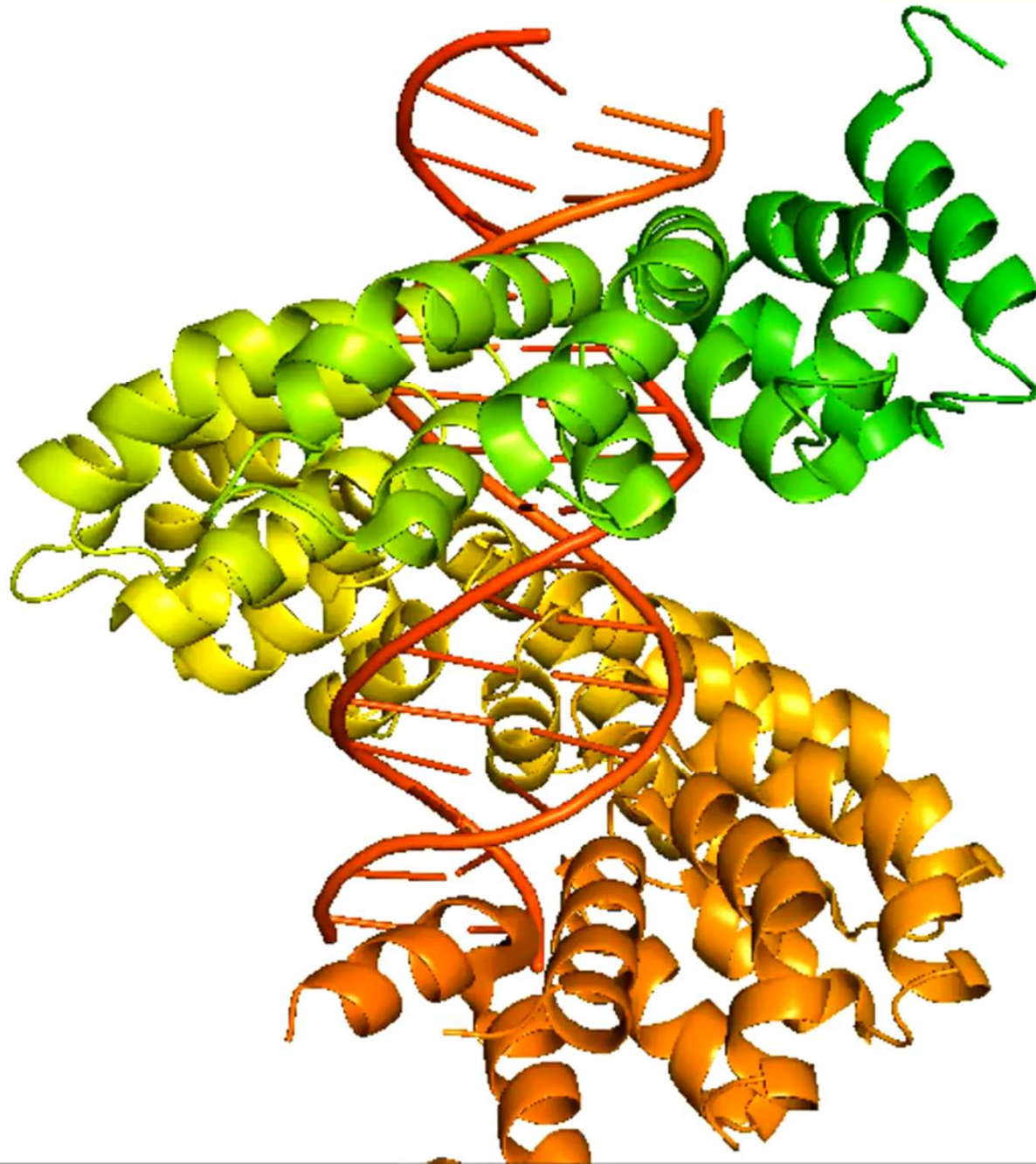
PthXo1

23 repetitivních obtáčků DNA ve VŽ

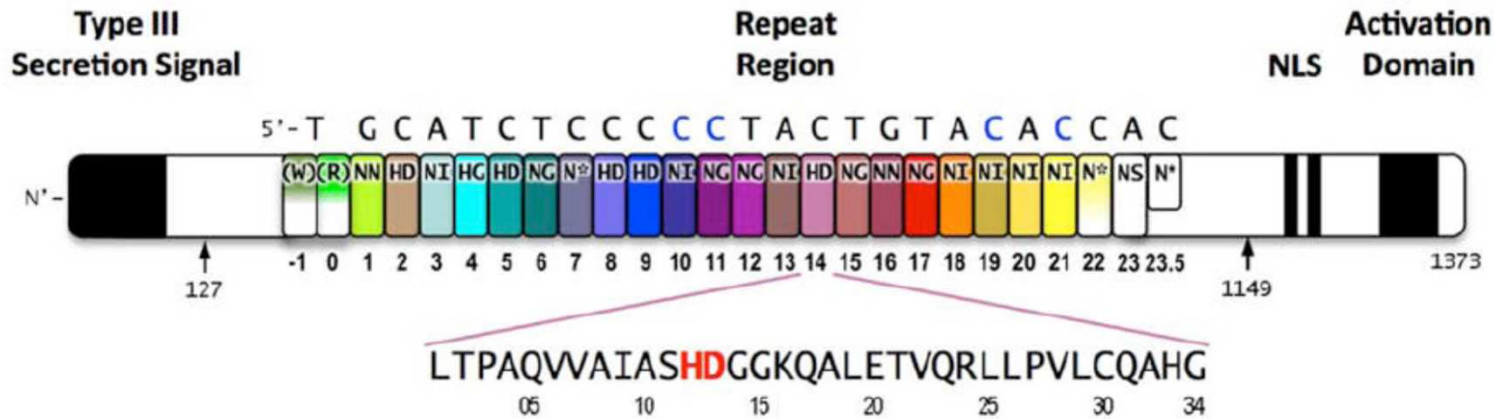


TALEN technologie

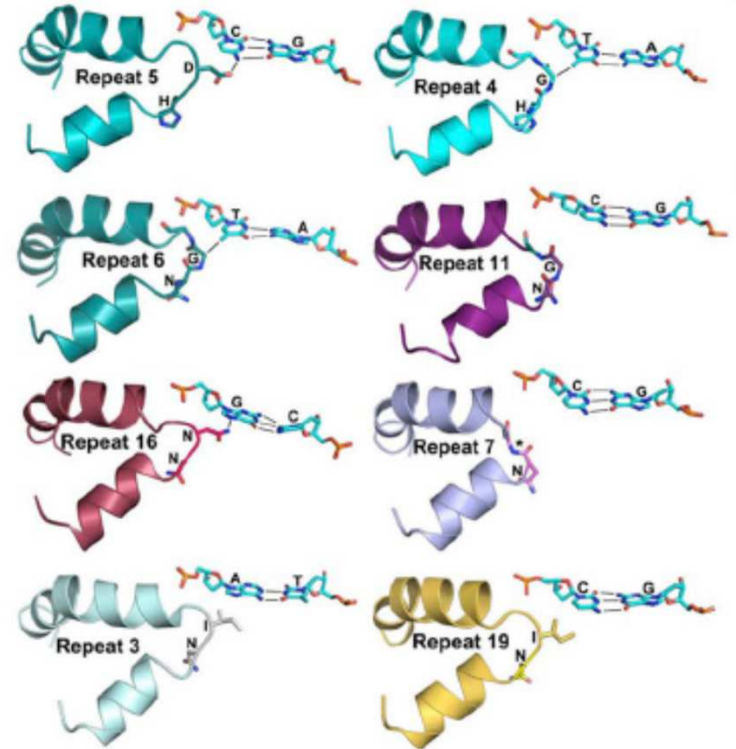
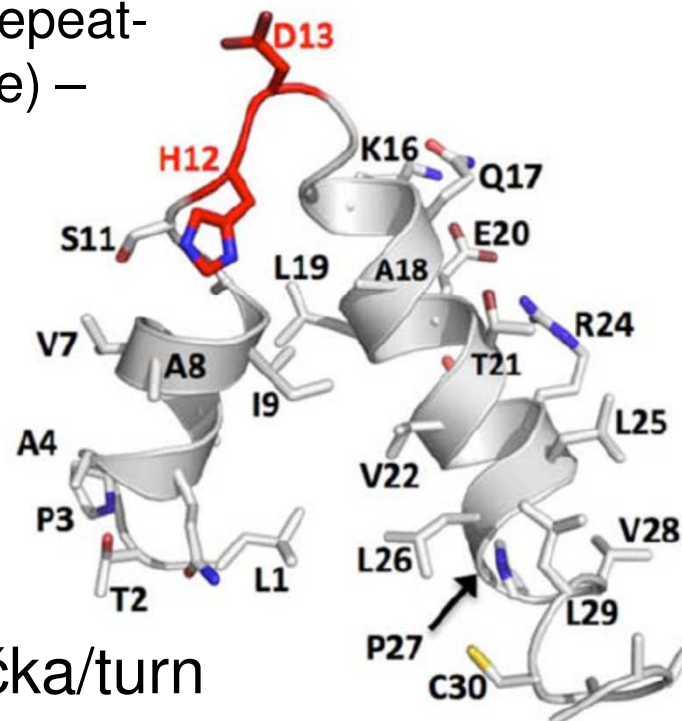
Mak et al, Science, 2012



Interaguje otáčka/turn spíše než šroubovice, PDB: 3V6T



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



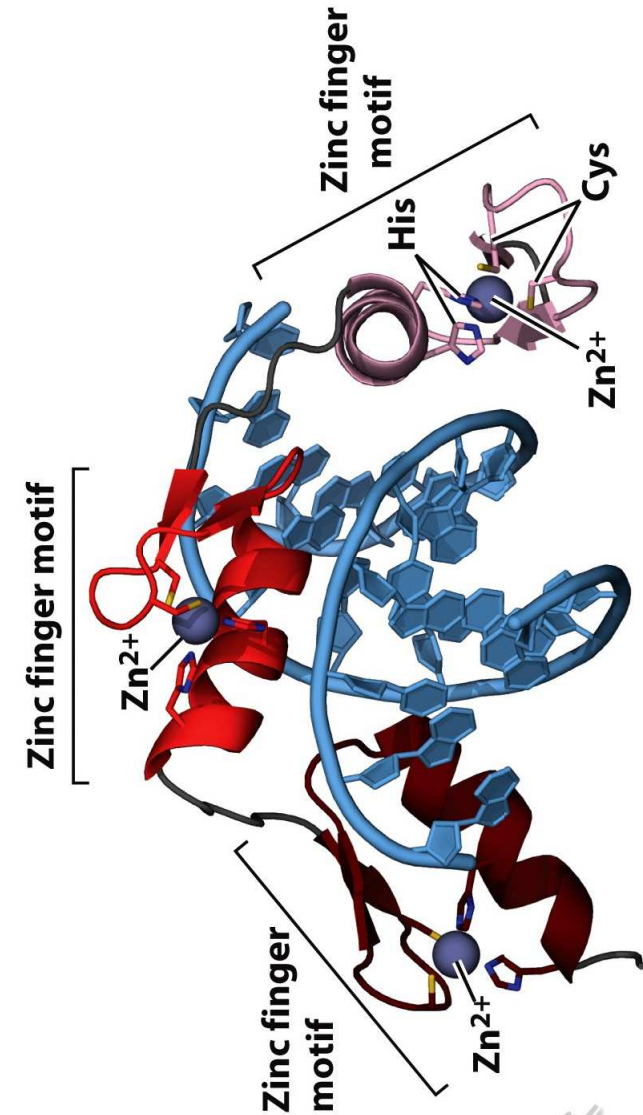
Interaguje otáčka/turn
 spíše než šroubovice

Mak et al, Science, 2012

Motivy DNA-vazebných domén

- **Zipper typ**
 - Leucinový zip
 - Helix-loop-helix
- **Helix-otáčka-helix**
 - HTH
 - Winged helix
 - TALE
- **Zinkový motiv**
 - $\beta\beta\alpha$ zinc-finger
 - Hormon-receptor
 - Loop-sheet-helix
 - Gal4

α -šroubovice

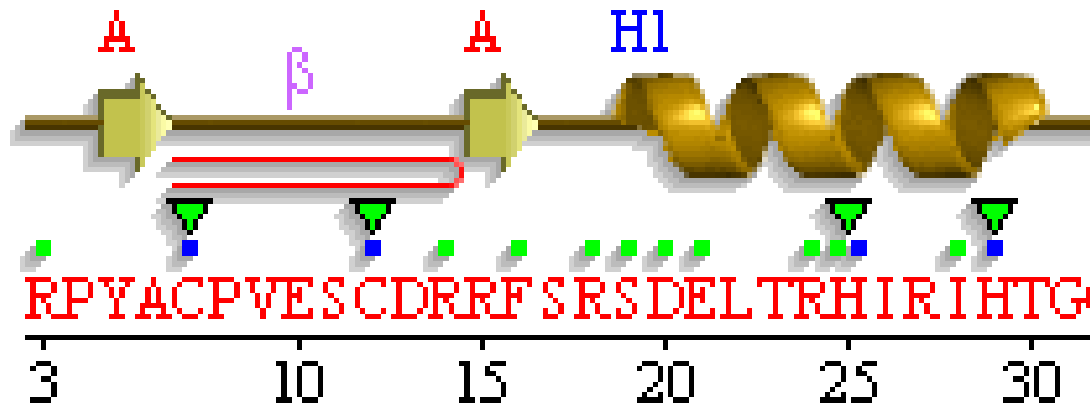
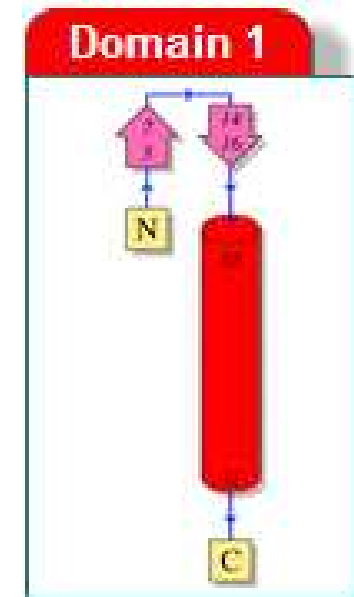


Zinc-finger/Zinkový prst

- cca 30 AMK ve dvou krátkých antiparalelních β -listech a α -šroubovici
- smyčka („hairpin“) stabilizovaná („crosslinked“) Zn^{2+} - **koordinovaný** 4xCys nebo 2xCys + 2xHis (tetraedrická struktura)

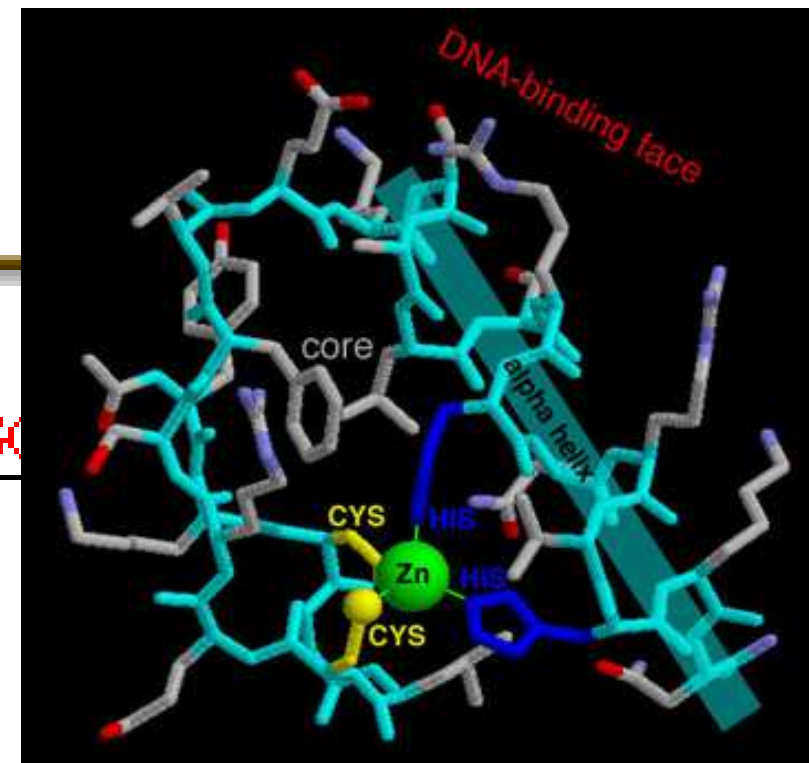
C2H2 motiv:

Cys-X₂₋₄-Cys-X₃-Phe-X₅-Leu-X₂-His-X₃-His



Motifs: β beta turn beta hairpin

Residue contacts: to DNA/RNA to metal

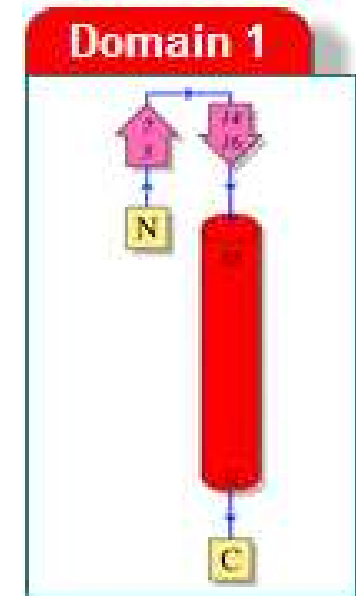
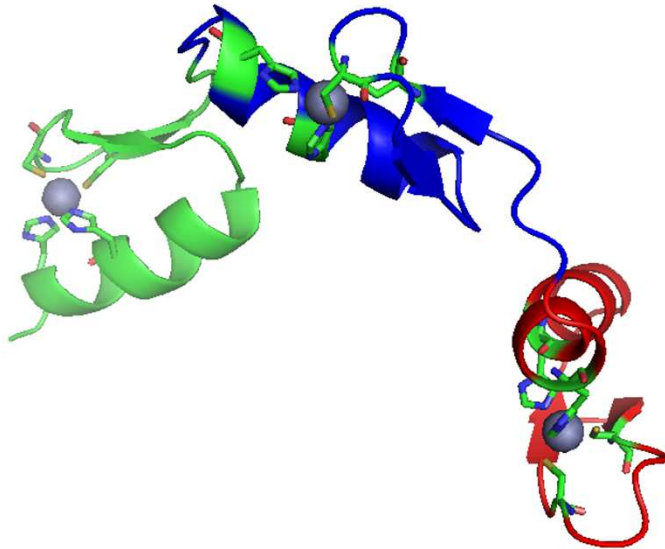


Zinc-finger/Zinkový prst

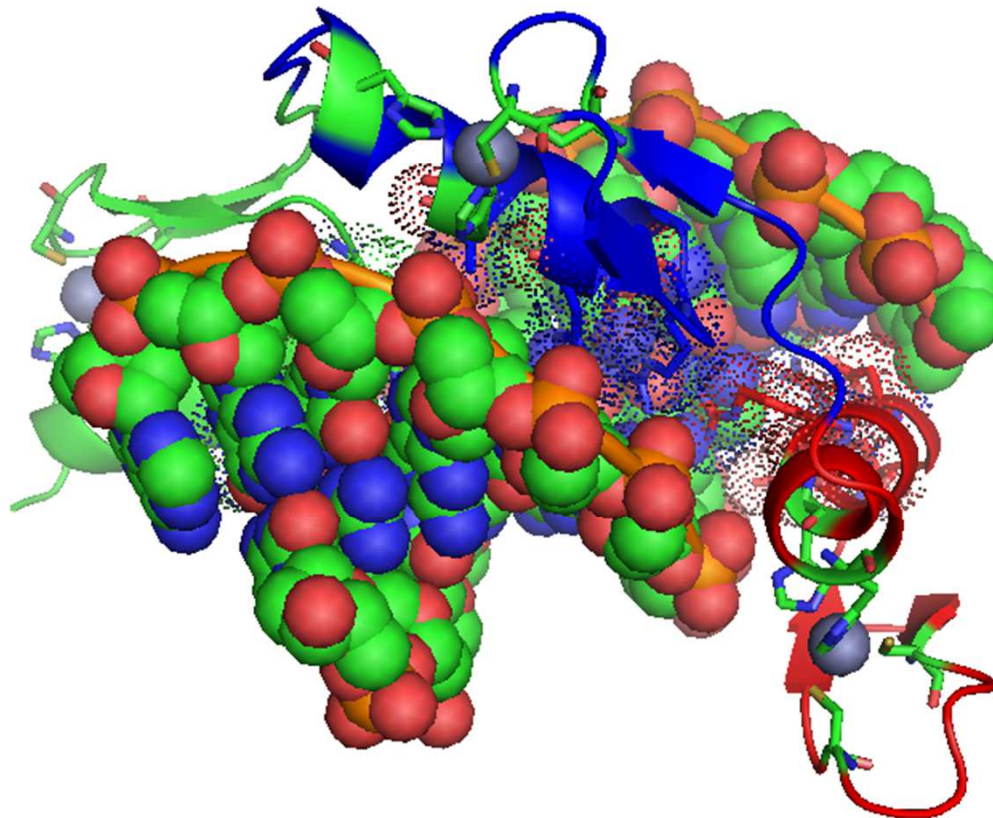
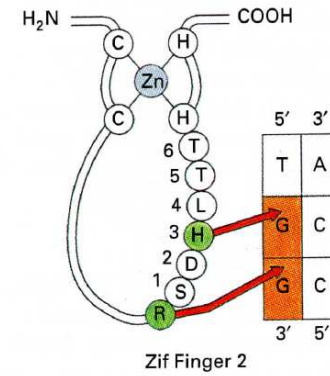
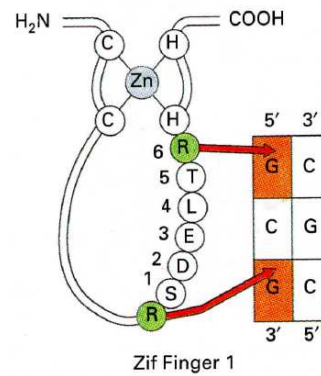
- cca 30 AMK ve dvou krátkých antiparalelních β -listech a α -šroubovici
- smyčka („hairpin“) stabilizovaná („crosslinked“) Zn^{2+} - **koordinovaný** 4xCys nebo 2xCys + 2xHis (tetraedrická struktura)

C2H2 motiv:

Cys- X_{2-4} -Cys- X_3 -Phe- X_5 -Leu- X_2 -His- X_3 -His

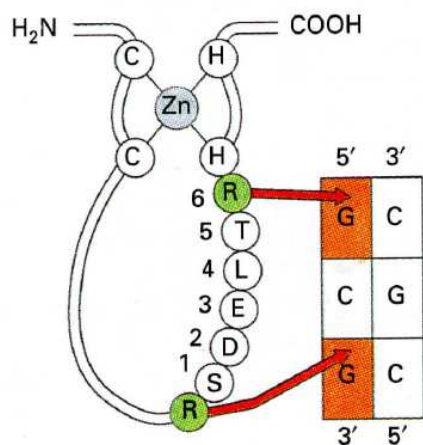


- 3x v Zif268, PDB=1zaa
- α -šroubovice se váže do VŽ – v tandemu obtáčí VŽ
- AMK na pozici 0 – 6; variancemi AMK => sekvenční specifita



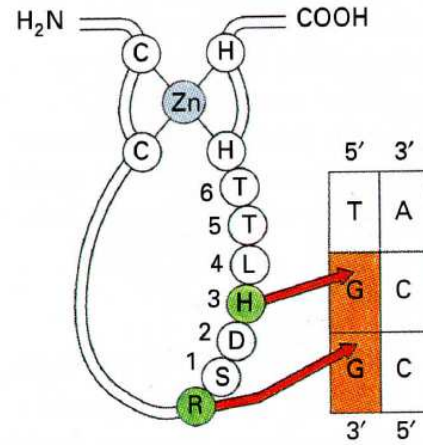
- AMK na pozici 0 – 6; variancemi AMK v těchto pozicích lze dosáhnout různé sekvenční specifity

- α -šroubovice váže 2, **3** nebo 4 sousední páry bazí



Zif Finger 1

Zif268



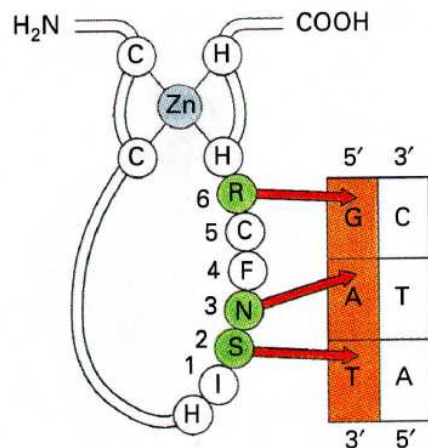
Zif Finger 2

- nejčastější jsou kontakty

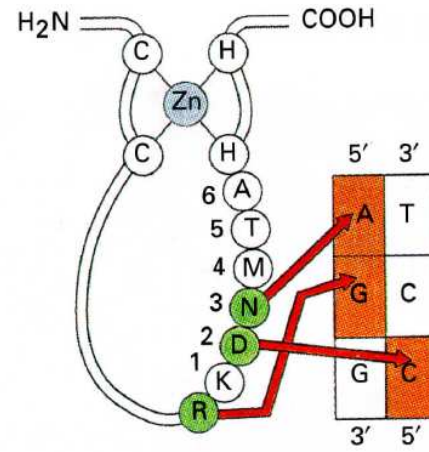
Gua-Arg

- Gua se může vázat i na His, Lys, Ser

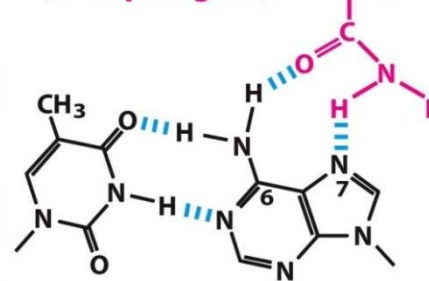
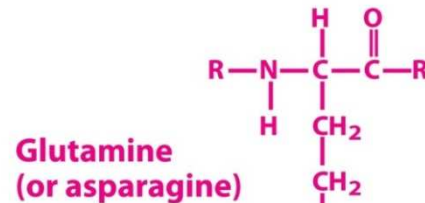
- Ser se může vázat na T či A



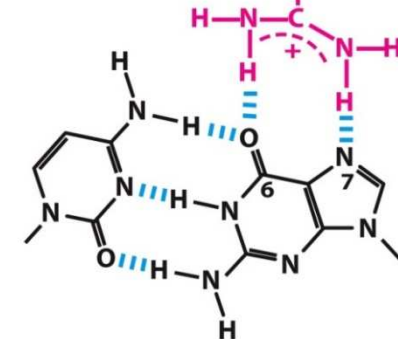
TTK Finger 1



TTK Finger 2



Thymine = Adenine



Cytosine ≡ Guanine

- 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 page title is "Predicting DNA-binding Specificities for Cys₂His₂ Zinc Finger Proteins" and the subtitle is "A DNA binding site predictor for Cys₂His₂ Zinc Finger Proteins".

On the left side, there is a sidebar with 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", "GENEVESTIGATOR - 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 features a navigation menu on the left with the following items: "Home Page", "Protein-DNA Form", "Generate Sequence Logo", "Downloads", "Analyze Genomic Sequence", "Help", and "Contact Us".

The main text area contains the following text:

Welcome to our new site!

For a given C₂H₂ 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.

This site serves as an interface between a user's input and a set of prediction algorithms that are able to create the mentioned logos. It consists of a protein input screen, followed by the selection of the desired set of fingers, and the algorithm to be used. The results are delivered through a simple HTML page that contains the generated sequence logo.

Below the text is a 3D ribbon diagram of a zinc finger protein structure, showing a blue ribbon structure with a green sphere representing a zinc ion coordinated by two cysteine residues.

At the bottom of the main content area, the URL <http://zf.princeton.edu> is displayed in large black text.

At the very bottom of the page, there is a footer with the following text:

This research has been supported by NSF CCF-0542187 and NIH GM076275

Disclaimer: Works better with Firefox v2+, Safari v3+ or Opera v9+

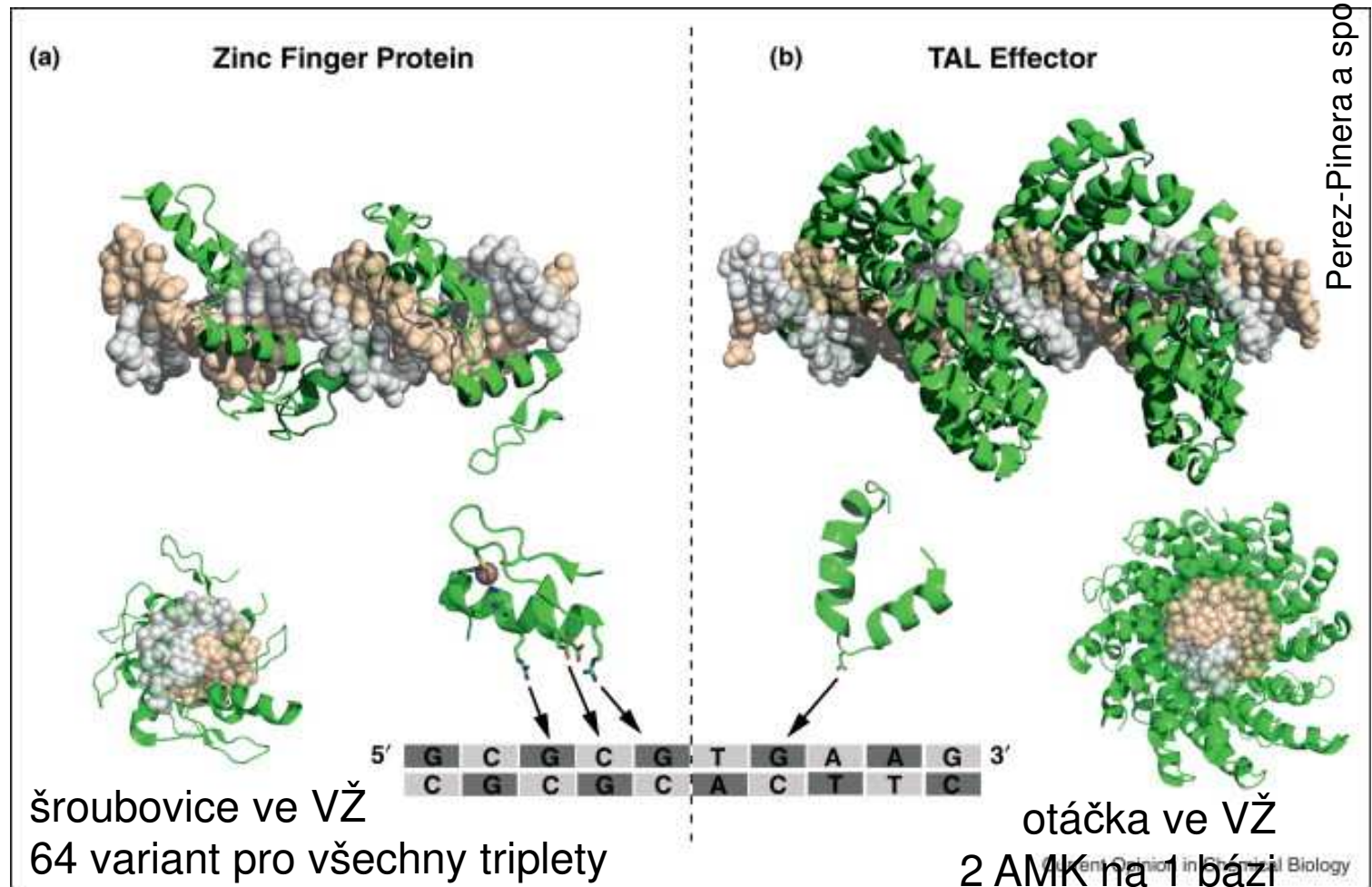
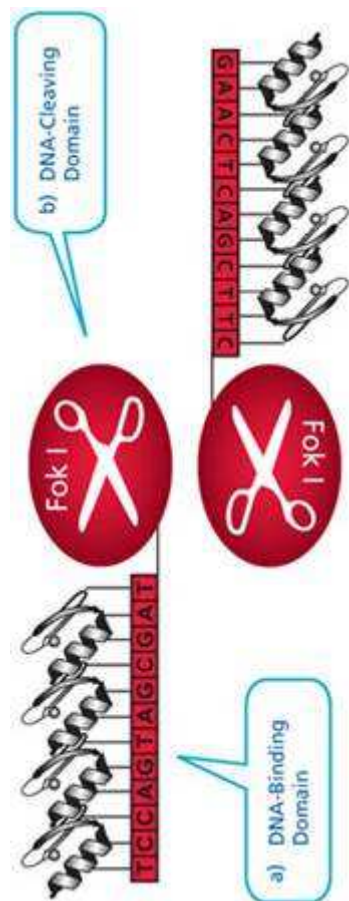
This page has been visited 2369 times since 09-09-2010

On the right side of the browser window, there is a vertical text label: "Persikov a Singh, NAR, 2014".

- 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

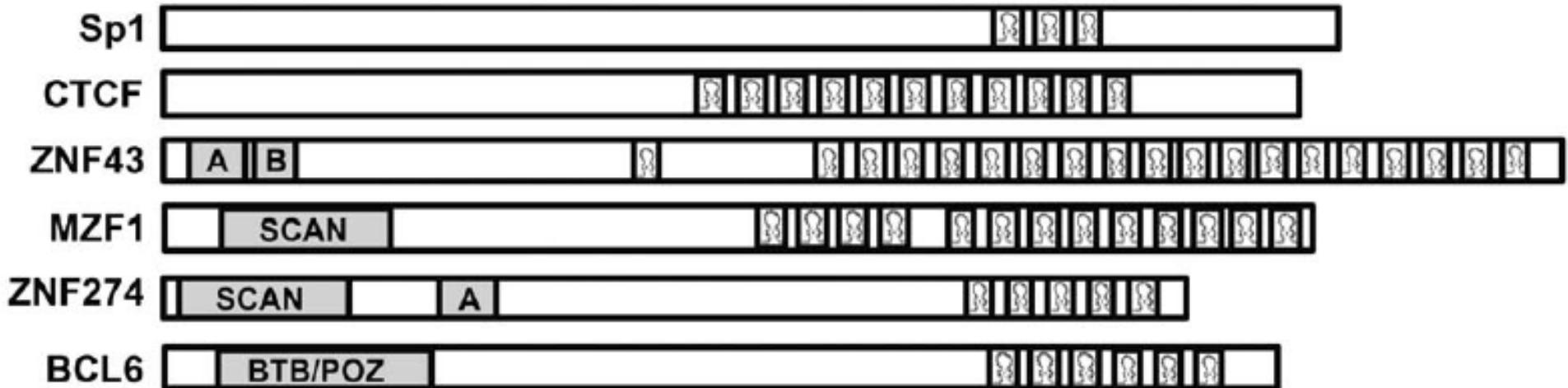
„genome editing“

Transcription activator-like



- CTCF obsahuje 11 zinkových prstů – k vazbě na DNA používá v různých org. různé kombinace ZF

ZF1
 Ts YQCEFCPYTNHRRYLLRHMKSSEERP
 Hs FQCELCSYTCPRRSNLDLHMKSHTDERP
 Dm YSCPHCPYTASKKFLITRHSRSDVEPS



ZF6
 Ts YQCEVCNQRFTQNSLKAHKLTHSG.SRPV
 Hs YECYICHARFTQSGMKNHILQHTENVAK.
 Dm YQCDICKSRFTQNSLKAHKLTHSVVDKPV

ZF7
 Ts FQCKFCPSSCGRRKDLRIHVQKLDHTA.SAP
 Hs FHCYICDVIARKSDLVHLRKHHSYIEQG
 Dm FQCNYPCTTCGRKALRVHIKHMHTS.DVE

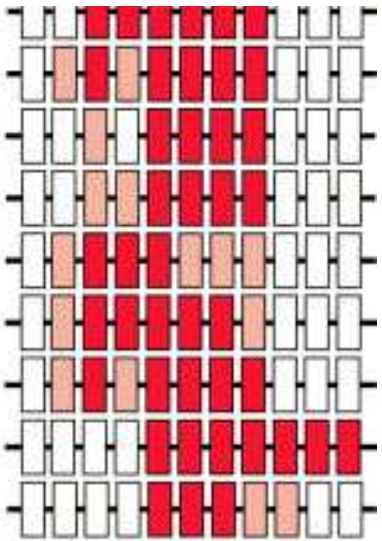
ZF8
 Ts IKKKKCDRTFTDRYTFKLHCKEHDGERC
 Hs KKCRYCDAVFHERYALIQHOKSHKNEKR
 Dm MTCRRCGQLPDRYQYKLVHVKSEGEKC

ZF9
 Ts YQCHLCPYSAMAQRHLEAHTLHSDKPF
 Hs FKCDQCDYACRQERHMHKRTHTGKEKP
 Dm YSCKLCSYASVTQRHLASHEMLHLDKEKP

ZF10
 Ts YKQVDNLSFKQVSLKRVHVESTHAANQ
 Hs YACSHCDKTFROKQLLDMHFKRYHDPNFV
 Dm FHCQCPQAFRQRLRRHMNLVHNEEYQ

ZF11
 Ts LNDNLASPSSTSGVSVASASSSSSPSSSTPNS
 Hs FA.....AFVCSKCGKTFRRNTMARHADNCA
 Dm PPEPREKLHKCPSPREFTHKGNLMRHMETHD

human	GAGCGGAAGAGGAAGTCCACCGCGTTTTTGAACACGGGA
8. PLK h/m	AGAGGAAGATTTAAGTAAAAGCTTCCTGGAGGAGGCGCAA TCTCCTTCTAAATTCATTTTCGAAGGACCTCCTCCGCGTT
9. PIM-1 human	CTTTTCCTTCCCGCCACGCTCGGGCGCGTAGAGACCATT GAAAAGGAAGGGCGGTCACACCCGCCGCATCTCTGGTAA
10. PIM-1 mouse	GGGGAGAGGGGTGTACCCCGAGGGGGCGAGCGGAGGG CCCTCTCCCCACATCGGCGCTCCCGCCCTCGCCTCC
11. p19ARF mouse	GCAGGGCCCGCCCGCCCTCCCGCTGGGGCCCTCTGGGA CGTCCCGGGCGCGCGCGAGGGGACCCCGGAGAACCCT
12. DMD4 mouse	GGAAACGGAGCTACCGCCCGGCGCAGCATACTCCTATATA CCTTGCCCTCGATGGCGCGCCACCCTCTATAGGATATAT
13. DMD7 mouse	CTAAATGGACAGACGATGCCCGGGGCGCAATACAATAC GATTACCTGCTGCTACGGCGCACCACCGTCATGTTATG
14. "144" silen. rat	CCTGAGTGCATTTCCCTCATGATCCAAAAGAGGCAATAAC GGACTCACGTAAGGGGAGTACTAGGTTTCTCCCTTATTG
15. APP human	TTCCCGCGCGCGCGCGTACCCCGGCTCTCTCTCGGGTCCGA AAGGCCCCGCGCGCGCGTACCCCGGCTCTCTCTCGGGTCCGA

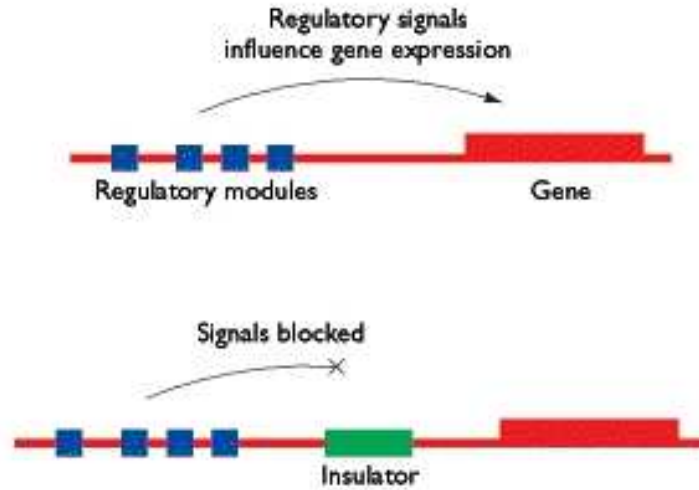


TRENDS in Genetics

Ohlsson a spol., TiG, 2001

CTCF

A) Insulators block the regulatory signals that control gene expression



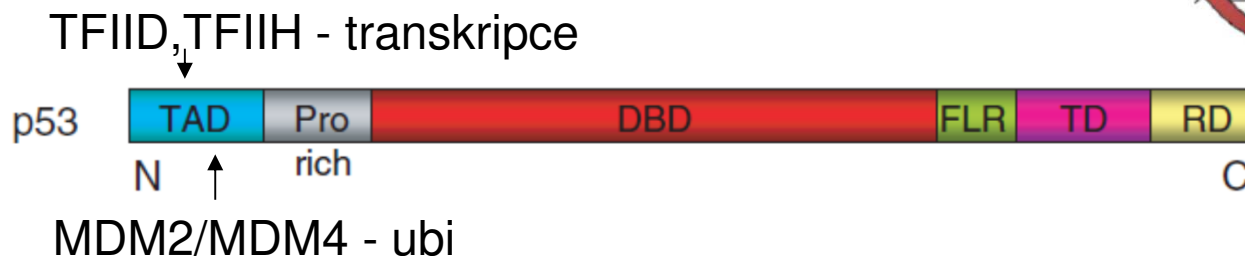
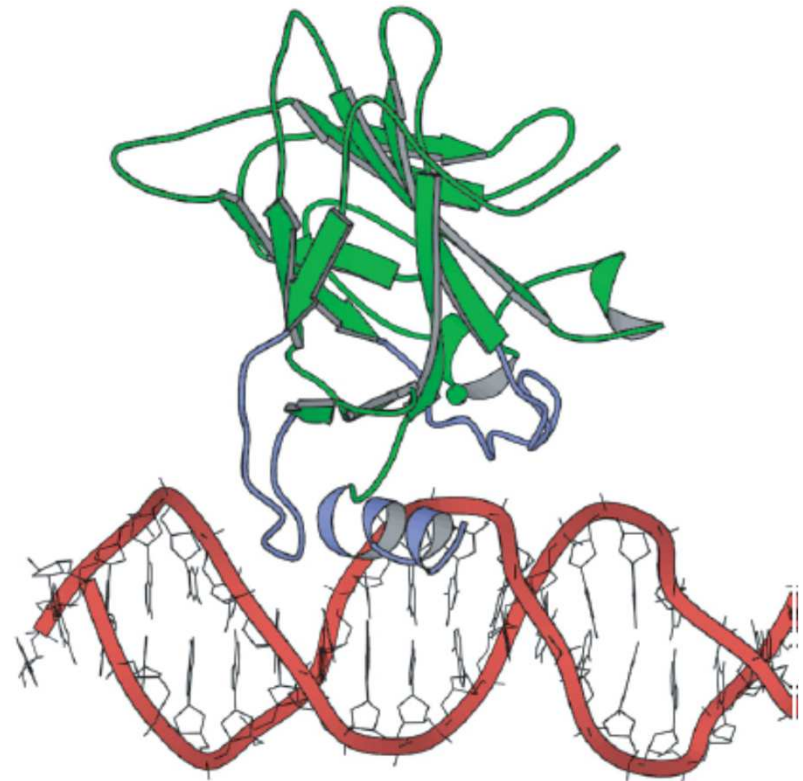
- CTCF (zkratka z CCCTC factor)
- izolátor/insulator brání transkripci
- váže se mezi transkripční aktivátory a obecné transkripční faktory

ZF1	Ts YOC ¹⁰ EF ²⁰ CPYTNH ³⁰ KRYLLR ⁴⁰ HMKS ⁵⁰ HSEERP Hs FOC ¹⁰ EL ²⁰ CSYTCP ³⁰ RRSNLDR ⁴⁰ HMKS ⁵⁰ HTDERP Dm YSC ¹⁰ PH ²⁰ CPYTAS ³⁰ KKFLIT ⁴⁰ HSRS ⁵⁰ H ⁶⁰ DVEPS
ZF2	Ts FK ³⁰ CTV ⁴⁰ CERC ⁵⁰ FK ⁶⁰ TNS ⁷⁰ SLON ⁸⁰ HINT ⁹⁰ HTGTRP Hs HK ³⁰ CHL ⁴⁰ CGRAF ⁵⁰ RTV ⁶⁰ TLLRN ⁷⁰ HLNTH ⁸⁰ HTGTRP Dm FK ³⁰ CSI ⁴⁰ CERS ⁵⁰ F ⁶⁰ R ⁷⁰ SNVGLQ ⁸⁰ NHINT ⁹⁰ H ¹⁰⁰ MGNKP
ZF3	Ts HO ⁶⁰ CKG ⁷⁰ CELAF ⁸⁰ TTS ⁹⁰ GELI ¹⁰⁰ HHIRYK ¹¹⁰ HTLEKP Hs HK ⁶⁰ CPD ⁷⁰ CDMAF ⁸⁰ V ⁹⁰ TSGELV ¹⁰⁰ R ¹¹⁰ RRYK ¹²⁰ HTHEKP Dm HK ⁶⁰ CKL ⁷⁰ CESA ⁸⁰ F ⁹⁰ TSGELV ¹⁰⁰ R ¹¹⁰ TRYK ¹²⁰ HTKEKP
ZF4	Ts HK ⁹⁰ CTE ¹⁰⁰ CSYASV ¹¹⁰ ELSKL ¹²⁰ R ¹³⁰ HIRSH ¹⁴⁰ TGERP Hs FK ⁹⁰ SM ¹⁰⁰ CDYASV ¹¹⁰ EVSKL ¹²⁰ R ¹³⁰ HIRSH ¹⁴⁰ TGERP Dm HK ⁹⁰ CTE ¹⁰⁰ CTYASV ¹¹⁰ ELTKL ¹²⁰ R ¹³⁰ HMTCH ¹⁴⁰ TGERP
ZF5	Ts YH ¹²⁰ CPHC ¹³⁰ SYASP ¹⁴⁰ DTYK ¹⁵⁰ LKR ¹⁶⁰ HLRV ¹⁷⁰ HTGEKP Hs FOC ¹²⁰ SL ¹³⁰ CSYASR ¹⁴⁰ DTYK ¹⁵⁰ LKR ¹⁶⁰ HMRT ¹⁷⁰ HTS ¹⁸⁰ GEEKP Dm YOC ¹²⁰ PH ¹³⁰ CTYASQ ¹⁴⁰ DMFKL ¹⁵⁰ R ¹⁶⁰ H ¹⁷⁰ MVI ¹⁸⁰ HTGEKK
ZF6	Ts YO ¹⁵⁰ CEV ¹⁶⁰ CNQR ¹⁷⁰ FTOS ¹⁸⁰ NSLKA ¹⁹⁰ HKL ²⁰⁰ .I ²¹⁰ HSG ²²⁰ .SRP Hs YBC ¹⁵⁰ YI ¹⁶⁰ CHAR ¹⁷⁰ FTOS ¹⁸⁰ GM ¹⁹⁰ MK ²⁰⁰ HL ²¹⁰ QK ²²⁰ HTENVAK. Dm YOC ¹⁵⁰ DI ¹⁶⁰ CKSR ¹⁷⁰ FTOS ¹⁸⁰ NSLKA ¹⁹⁰ HKL ²⁰⁰ .I ²¹⁰ HS ²²⁰ VVDKPV
ZF7	Ts FO ¹⁸⁰ CK ¹⁹⁰ CFPSS ²⁰⁰ CGR ²¹⁰ KTDL ²²⁰ LR ²³⁰ HVQKL ²⁴⁰ HTA ²⁵⁰ .SAP Hs FH ¹⁸⁰ CPH ¹⁹⁰ CDT ²⁰⁰ VIAR ²¹⁰ KSD ²²⁰ LV ²³⁰ HLR ²⁴⁰ KO ²⁵⁰ HSYIEQG Dm FO ¹⁸⁰ CNY ¹⁹⁰ CP ²⁰⁰ T ²¹⁰ CGR ²²⁰ KAD ²³⁰ LRV ²⁴⁰ HIK ²⁵⁰ HM ²⁶⁰ TS ²⁷⁰ .DVP
ZF8	Ts IK ²¹⁰ CKK ²²⁰ CDRT ²³⁰ FTDR ²⁴⁰ Y ²⁵⁰ TFKL ²⁶⁰ HCKE ²⁷⁰ H ²⁸⁰ DGERC Hs KK ²¹⁰ CRY ²²⁰ CDAV ²³⁰ FHER ²⁴⁰ YALIQ ²⁵⁰ HQKS ²⁶⁰ H ²⁷⁰ KNEKR Dm MT ²¹⁰ CR ²²⁰ CGQ ²³⁰ QLP ²⁴⁰ DRY ²⁵⁰ OYKL ²⁶⁰ HVKS ²⁷⁰ H ²⁸⁰ EGEKC
ZF9	Ts YO ²³⁰ CHL ²⁴⁰ CPYSAMA ²⁵⁰ QR ²⁶⁰ HLE ²⁷⁰ FA ²⁸⁰ HTLL ²⁹⁰ HSDK ³⁰⁰ P Hs FK ²³⁰ DD ²⁴⁰ CDYAC ²⁵⁰ ROER ²⁶⁰ MM ²⁷⁰ HK ²⁸⁰ RT ²⁹⁰ HTGEEKP Dm YSC ²³⁰ KL ²⁴⁰ CSYASV ²⁵⁰ TQR ²⁶⁰ H ²⁷⁰ LAS ²⁸⁰ HMLI ²⁹⁰ H ³⁰⁰ LDKEP
ZF10	Ts YK ²⁶⁰ QVD ²⁷⁰ CNLS ²⁸⁰ E ²⁹⁰ FK ³⁰⁰ QVSL ³¹⁰ LK ³²⁰ R ³³⁰ HVEST ³⁴⁰ HA ³⁵⁰ AAANQ Hs YAC ²⁶⁰ SH ²⁷⁰ CDK ²⁸⁰ TFR ²⁹⁰ QK ³⁰⁰ QLDM ³¹⁰ HF ³²⁰ KRY ³³⁰ DPNFV Dm FH ²⁶⁰ QD ²⁷⁰ CPQAF ²⁸⁰ R ²⁹⁰ QRL ³⁰⁰ LR ³¹⁰ H ³²⁰ MNLV ³³⁰ H ³⁴⁰ NEEYQ
ZF11	Ts LNDNLASPSTSGVSVASASSSSSSFSSTSPNSS Hs PA...AFV ²⁹⁰ SK ³⁰⁰ CGK ³¹⁰ TFR ³²⁰ TRNT ³³⁰ MA ³⁴⁰ RHADNCA Dm PPEPREKLHK ²⁹⁰ CP ³⁰⁰ CPRE ³¹⁰ FT ³²⁰ HKG ³³⁰ NLM ³⁴⁰ RHMET ³⁵⁰ HD insertion

Loop-sheet-helix

- core/DNA-vazebná doména p53 – transkripční faktor důležitý pro regulaci buněčného cyklu, apoptozy a opravy poškozené DNA (nádorový supresor)

- smyčky vycházející mimo hlavní core doménu – vyčnívá β -list a α -šroubovice
- 3 Cys a 1His koordinují Zn
- helix ve velkém žlábkku a smyčka v malém žlábkku
- Aktivace transkripce skrze kyselou TA doménu

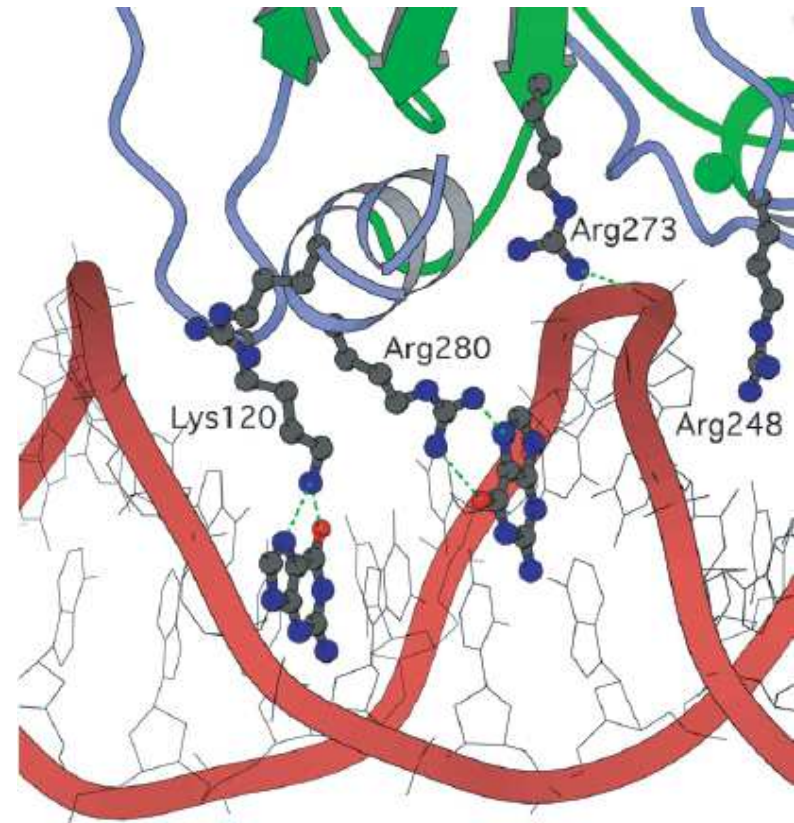


Loop-sheet-helix

- core/DNA-vazebná doména p53 – transkripční faktor důležitý pro regulaci buněčného cyklu, apoptozy a opravy poškozené DNA (nádorový 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)



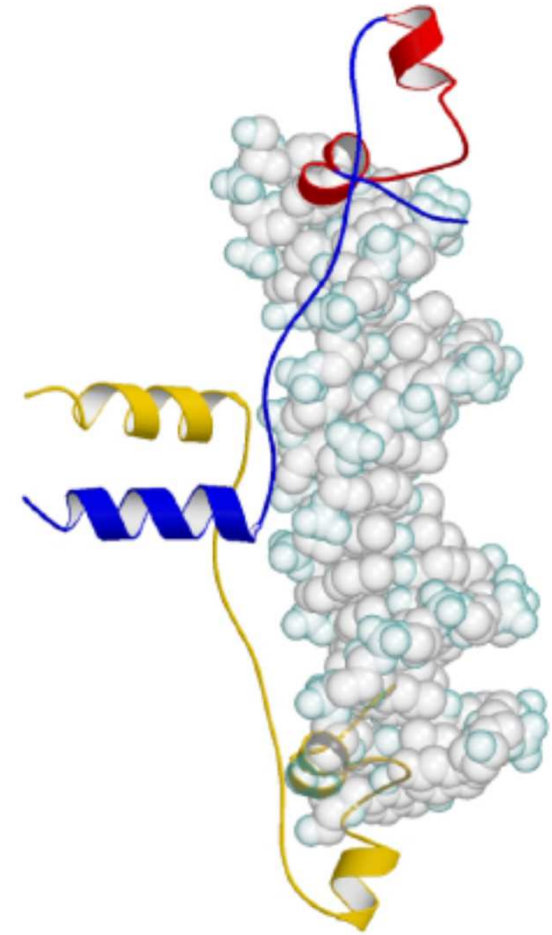
p53



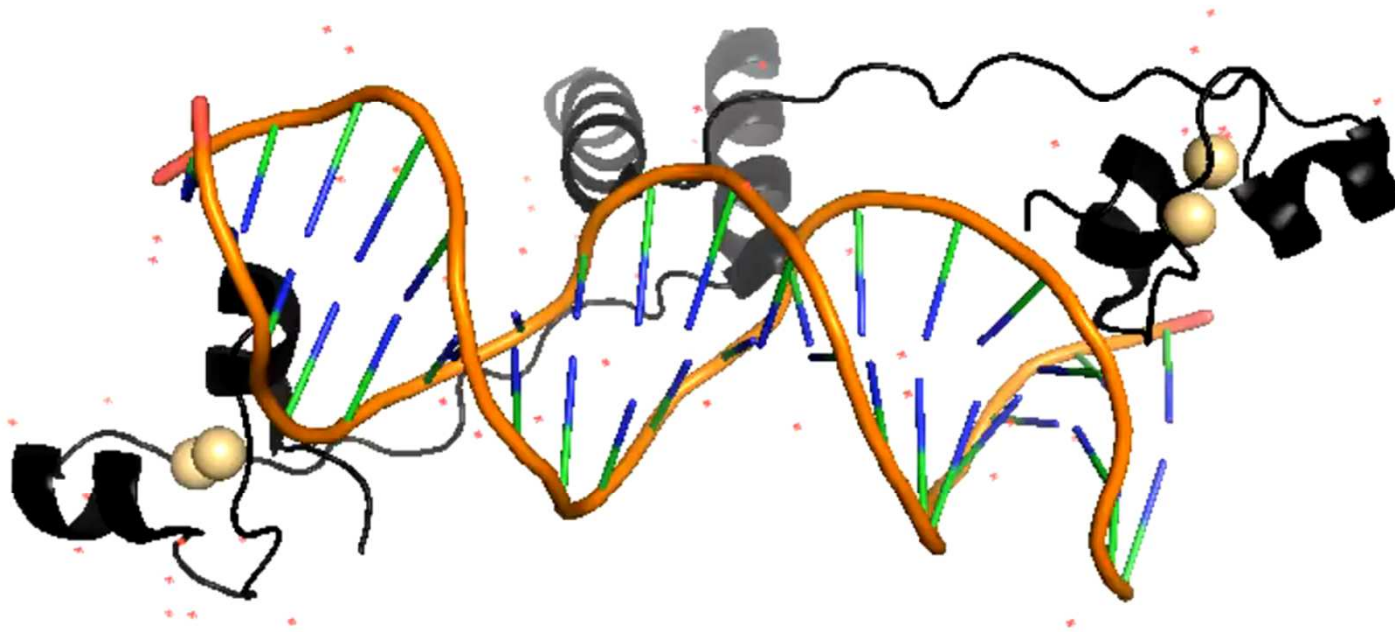
Gal4

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

- 2 α -šroubovice
- 6 Cys koordinuje 2 Zn (2 Cys sdílené 2 Zn)
- 1. šroubovice ve velkém žlábků a smyčka k 2. šroubovici kontaktuje cukr-fosfátovou kostrou
- Dimerizuje přes krátký CC segment



Gal4



PDB: 1D66

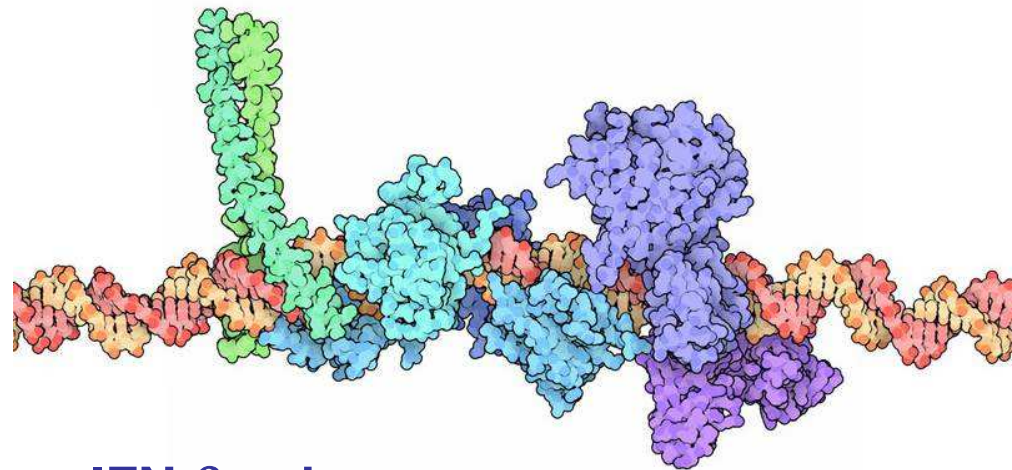
Motivy DNA vazebných domén

- **Zipper typ**
 - Leucinový zip
 - Helix-loop-helix
- **Helix-otáčka-helix**
 - HTH
 - Winged helix
 - TALE
- **Zinkový prst**
 - $\beta\beta\alpha$ zinc-finger
 - Hormon-receptor
 - Loop-sheet-helix
 - Gal4

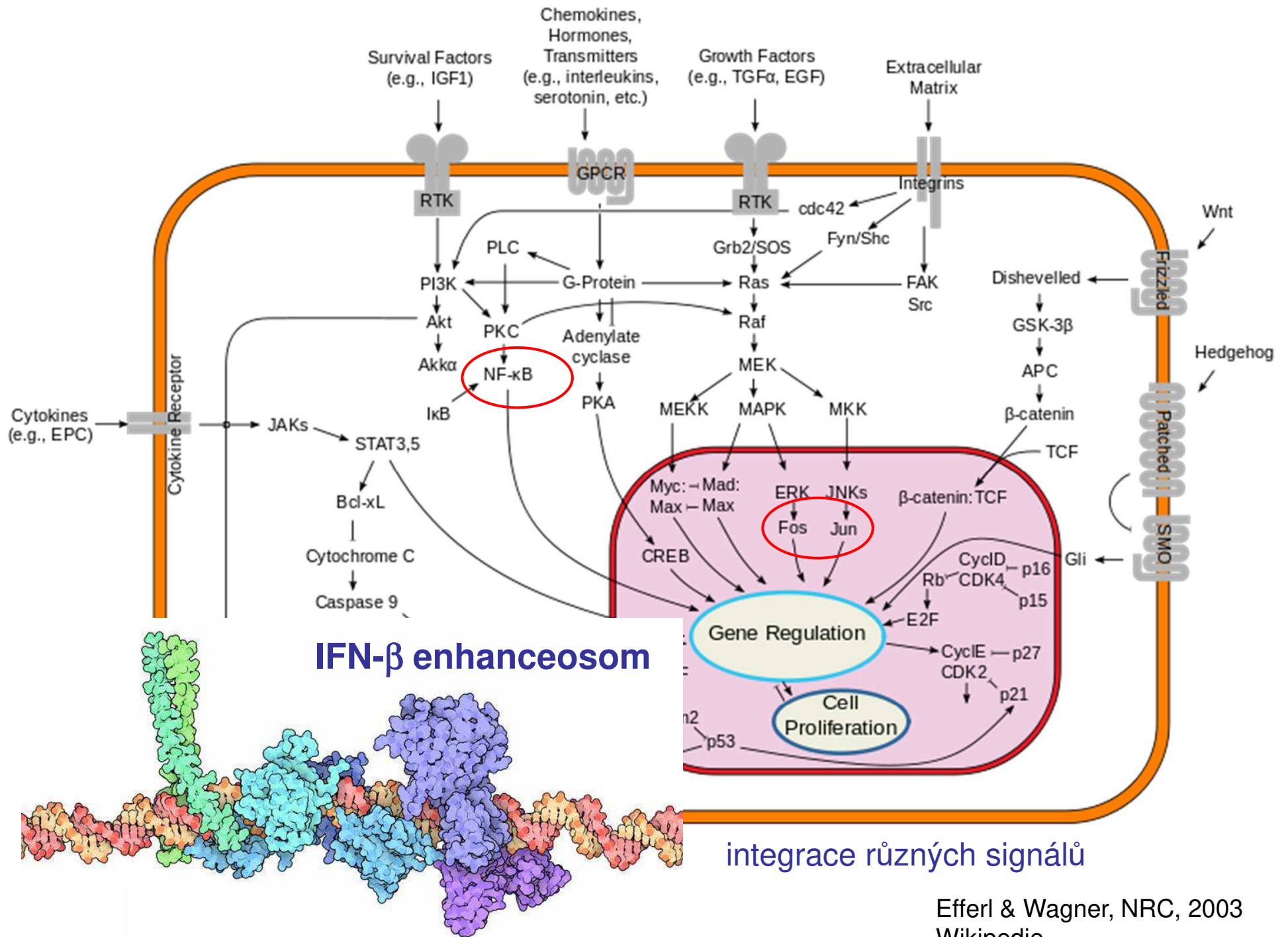
Transkripce ...

Kombinace více proteinů ...

Vliv chromatinu ...



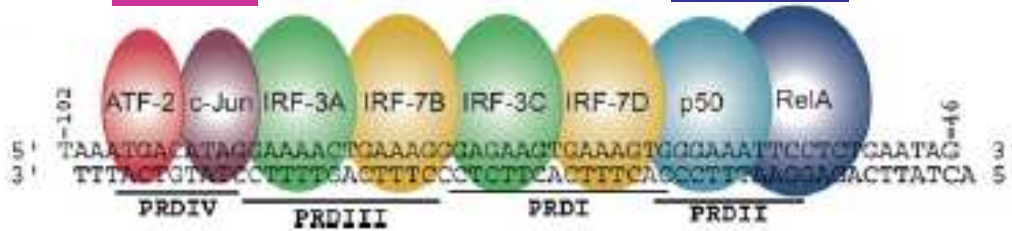
IFN- β enhanceosom



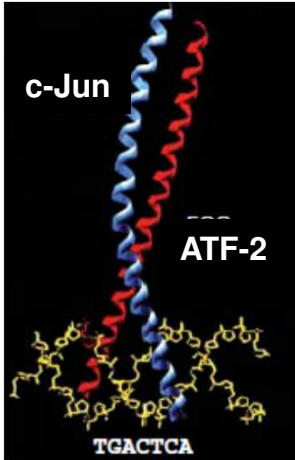
AP-1

NF-κB

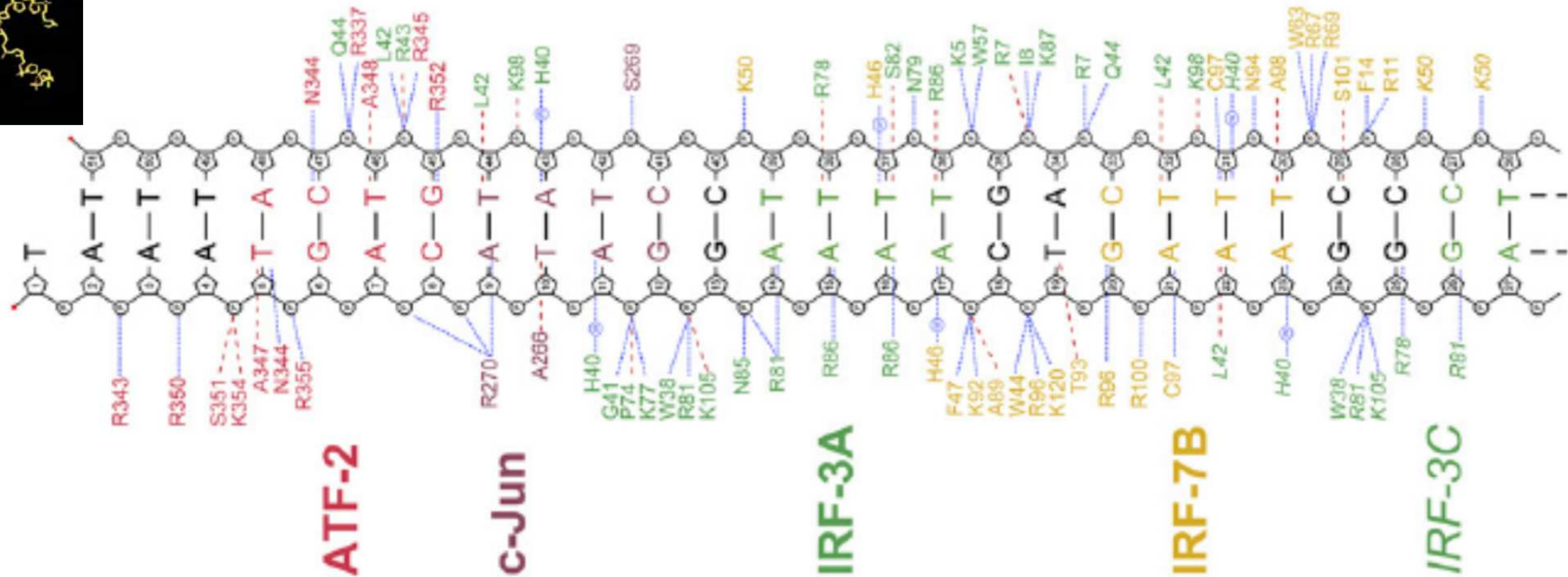
IFN-β enhanceosom



transkripce



- jeden z nejlépe popsaných enhancerů u vyšších eukaryot – induk. viry
- sekvence -102 až -47 básí upstream od počátku transkripce
- TF pokrývají 72% povrchu DNA (těsné sbalení DB-domén) – málo PPI
- nicméně vazba 8 proteinů 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)

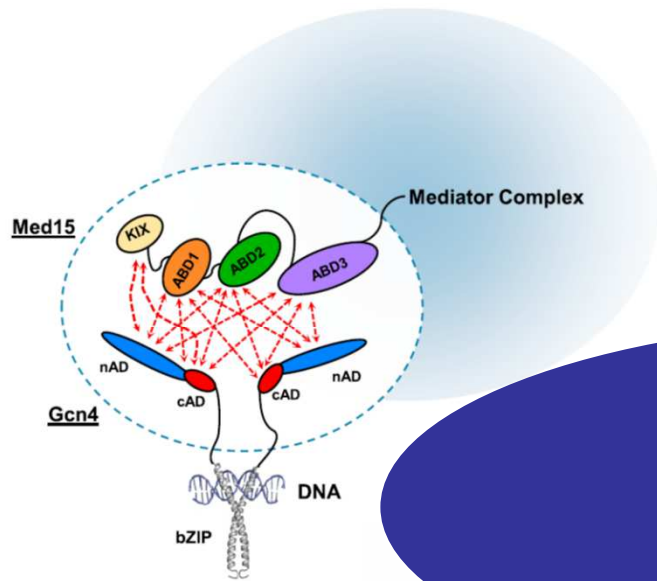
IFN- β enhanceosom



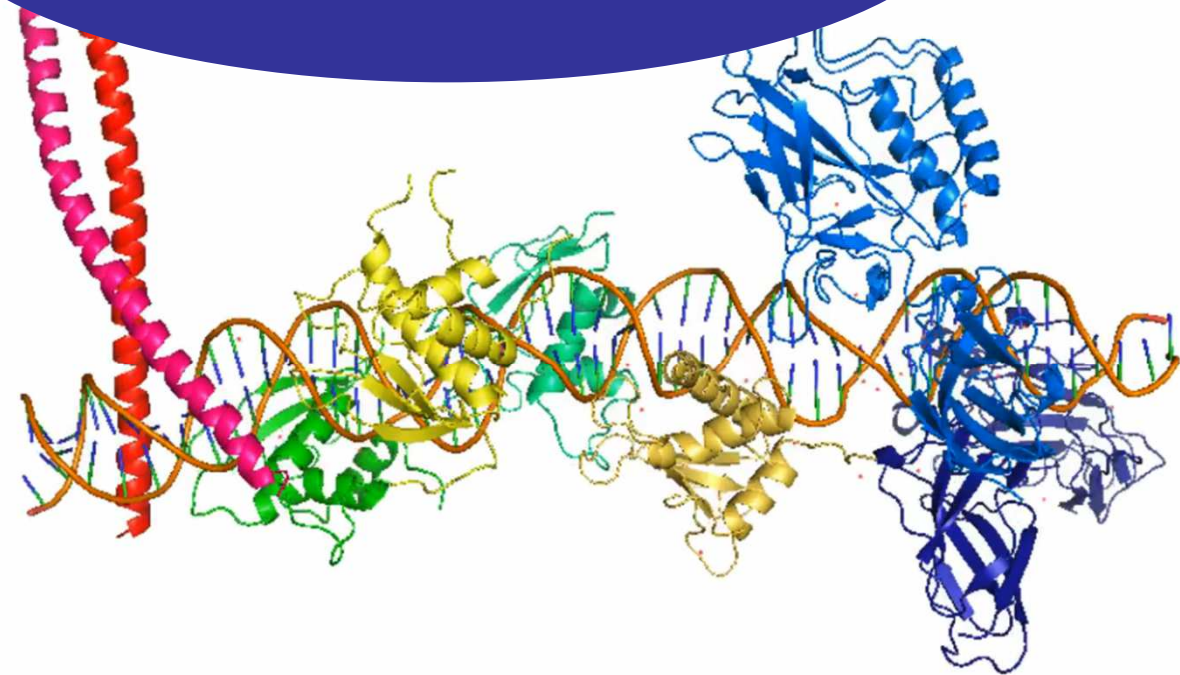
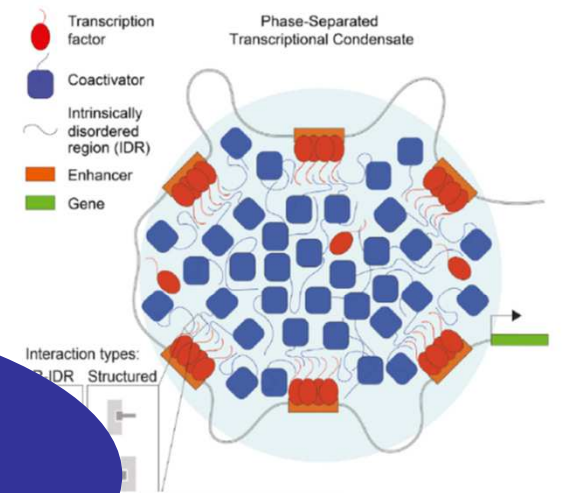
AP-1 leucin zipper, IRF – podobný WHD (směs), NF κ B – komplexní motiv

červené tečky – molekuly vody

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



mediator



Boija et al, Cell, 2018
 Tuttle et al, Cell Rep, 2018

- TF obsahují **aktivační doménu** – na AD se váže mediator komplex – integruje/propojí TF – zprostředkuje vazbu s RNA polymerasou - iniciaci transkripce

Souhrn:

- vazba většiny TF pomocí šroubovice ve velkém žlábkku (leucinový zip, HTH, zink-finger ...)
- transkripční komplexy (enhanceosom ...)

Příště:

- Histon, HMG-box
- β -sheet motivy
- enhanceosom ... a počátek transkripce

