

Interakce DNA s proteiny

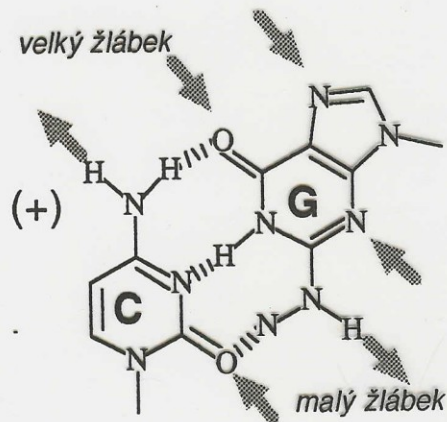
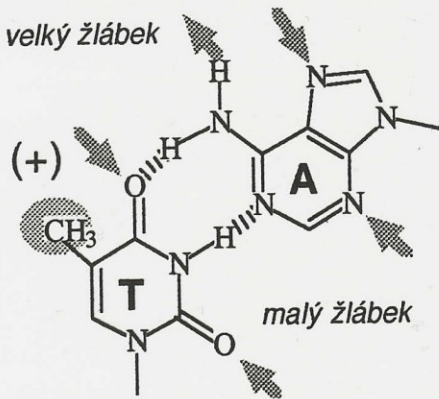
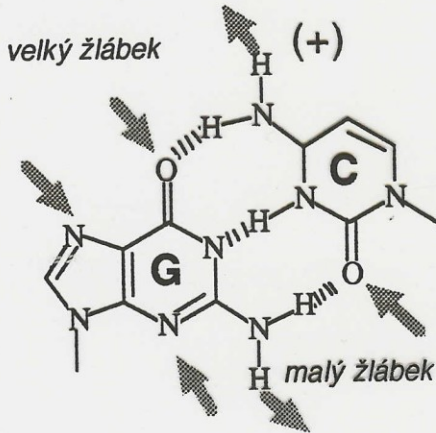
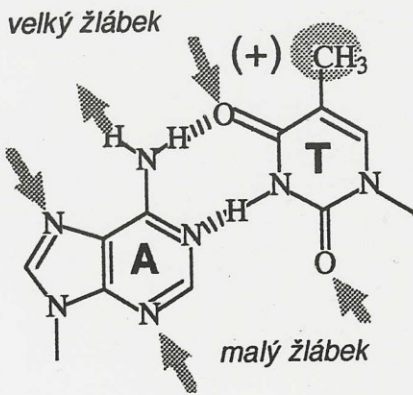
-veškeré funkce DNA: replikace, transkripce, rekombinace, regulace těchto procesů, opravy poškozené DNA

-výstavba chromatinu (a analogických struktur u prokaryot)

-interakce *sekvenčně specifické*: zejm. vodíkové vazby, solné můstky se zbytky bází v malém a velkém žlábkou dvoušroubovice

-interakce *sekvenčně nespecifické*: obecně interakce s cukrfofátovým řetězcem

FUNKČNÍ SKUPINY V PÁRECH BAZÍ, KTERÉ MOHOU VSTUPOVAT DO INTERAKCÍ S AMINOKYSELINAMI



← vodíková vazba

● van der Waalsova interakce

(+) elektrostatická interakce

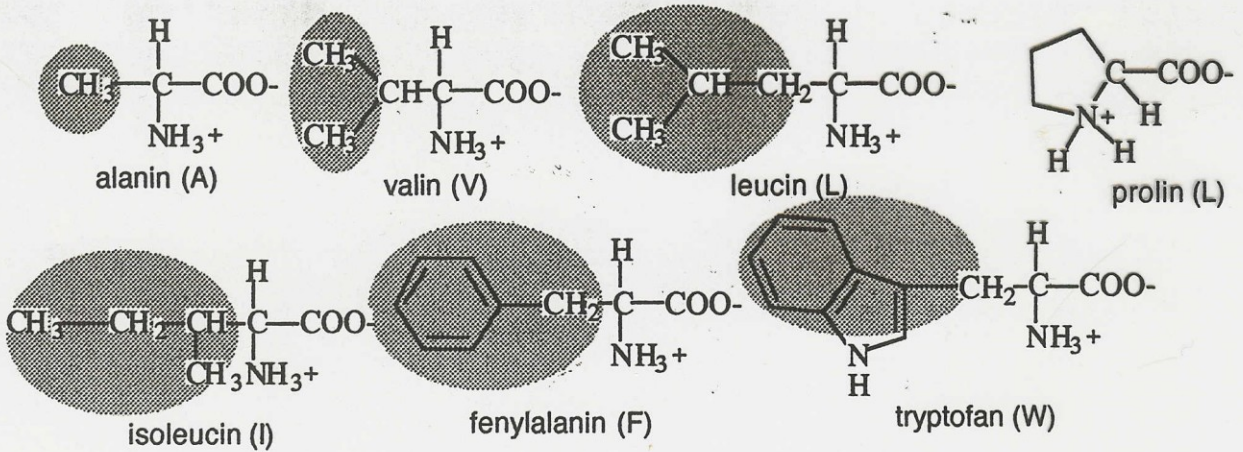
Velký žlábek: jeden donor a dva akceptory vodíkové vazby u všech bp; metylová skupina u T (a u 5-meC)

Malý žlábek: dva akceptory vodíkové vazby u všech bp; u C.G a G.C navíc je N2 (aminoskupina) G donorem vodíkové vazby

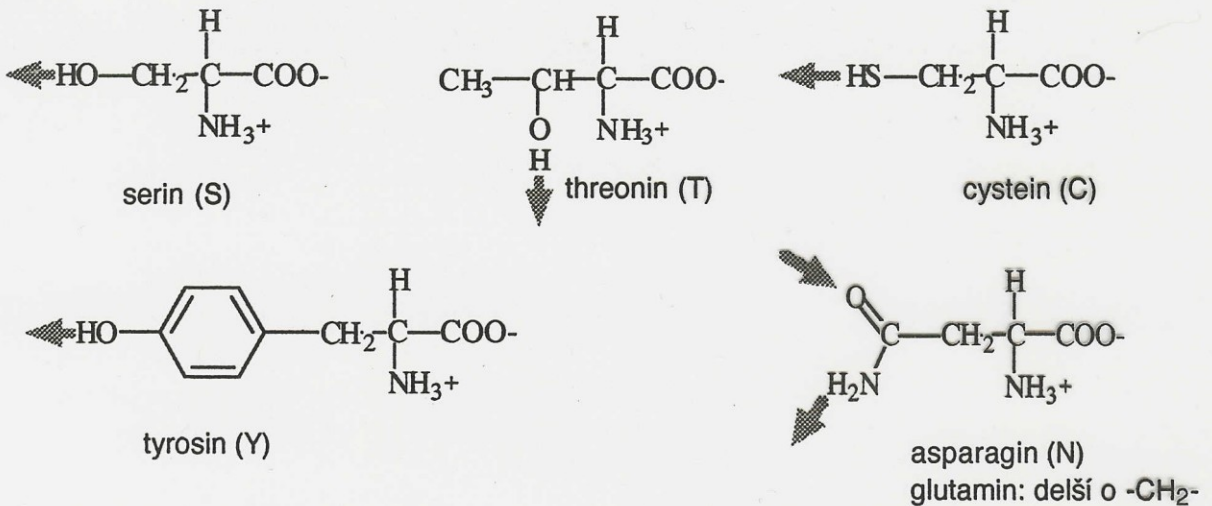
rozmístění interakcí je charakteristické pro každý bp

FUNKČNÍ SKUPINY AMINOKYSELIN

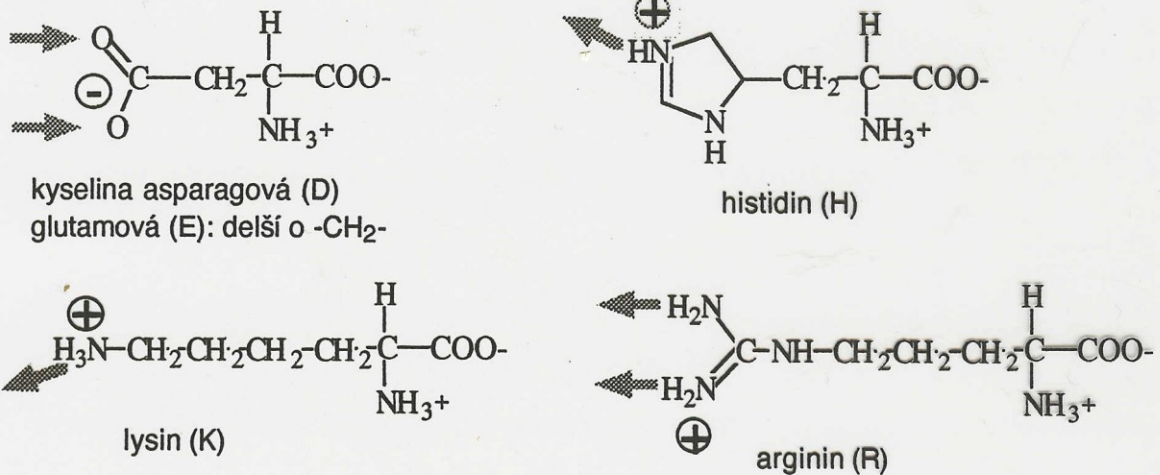
a/ nepolární boční řetězec



b/ polární

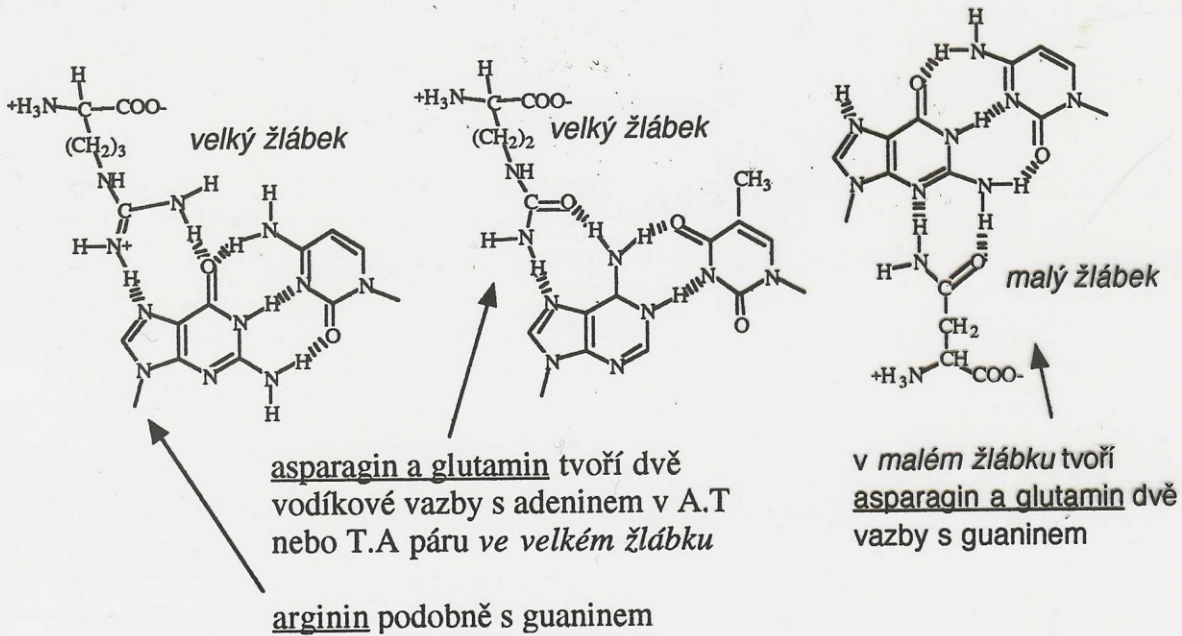


c/ nabíť



NEJVÝHODNĚJŠÍ INTERAKCE bp SE ZBYTKY AMINOKYSELIN

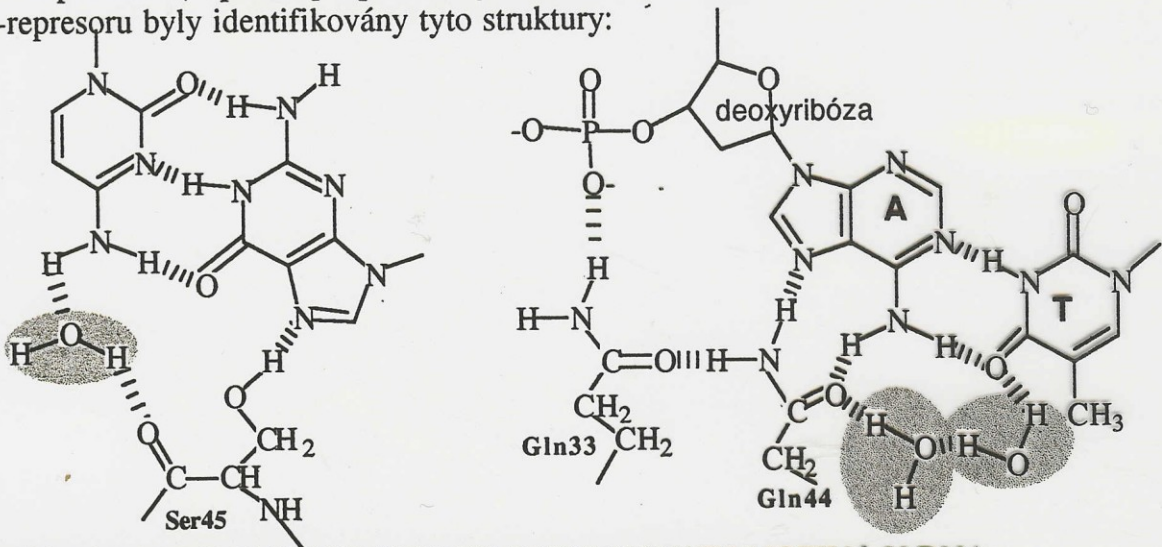
(těch, které tvoří více vazeb)



ÚLOHA MOLEKUL VODY

-hydratační obal DNA a proteinů musí být rozrušen

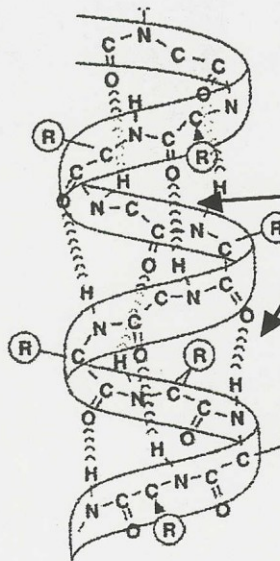
-jednotlivé molekuly vody často tvoří můstky zprostředkující vodíkové vazby mezi molekulami DNA a proteinů (např. v *trp* operátoru jsou všechny vodíkové vazby zprostředkovány vodou; v λ -repressoru byly identifikovány tyto struktury:



NEPŘÍMÉ ROZPOZNÁNÍ URČITÝCH SEKVENČNÍCH MOTIVŮ V DNA:

-vazba proteinů citlivá na *specifickou sekundární strukturu DNA, která je daná určitou sekvencí* (ta určuje lokální „twist“, „tilt“, „roll“ bází, ohyby dvoušroubovice a šířku žlábků a tím přesnou vzájemnou polohu donorových a akceptorových skupin vodíkových vazeb jak v bp, tak v cukr-fosfátové kostře - její geometrie často postačuje ke specifickému rozpoznání určité oblasti v DNA určitým proteinem)

ZÁKLADNÍ STRUKTURNÍ MOTIVY PROTEINŮ



α -helix

-pravotočivá šroubovice s 3.6 aminokyseliny na otáčku

-stabilizace vodíkovými vazbami mezi vodíkem amidové skupiny a karbonylovým kyslíkem o čtyři aminokyseliny dále

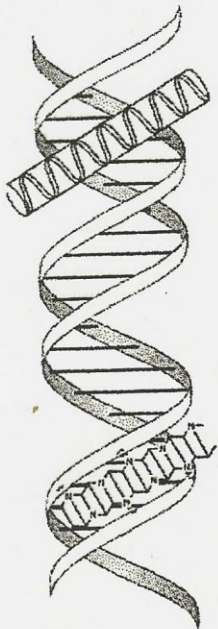
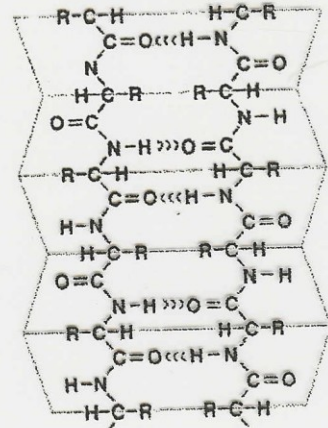
-postranní funkční skupiny aminokyselin jsou exponovány vně helixu a přístupné interakci s funkčními skupinami v DNA nebo s jinými segmenty téže molekuly proteinu (terciární struktura) či s dalším polypeptidovým řetězcem (kvarterní struktura)

β -(skládaný) list

-řetězce uloženy lineárně vedle sebe, antiparalelně nebo paralelně (mohou být ze stejné nebo různých molekul); relativně plochá struktura

-stabilizace vodíkovými vazbami mezi vodíkem amidoskupiny a karbonylovým kyslíkem

-postranní funkční skupiny aminokyselin směřují nad a pod rovinu β -listu

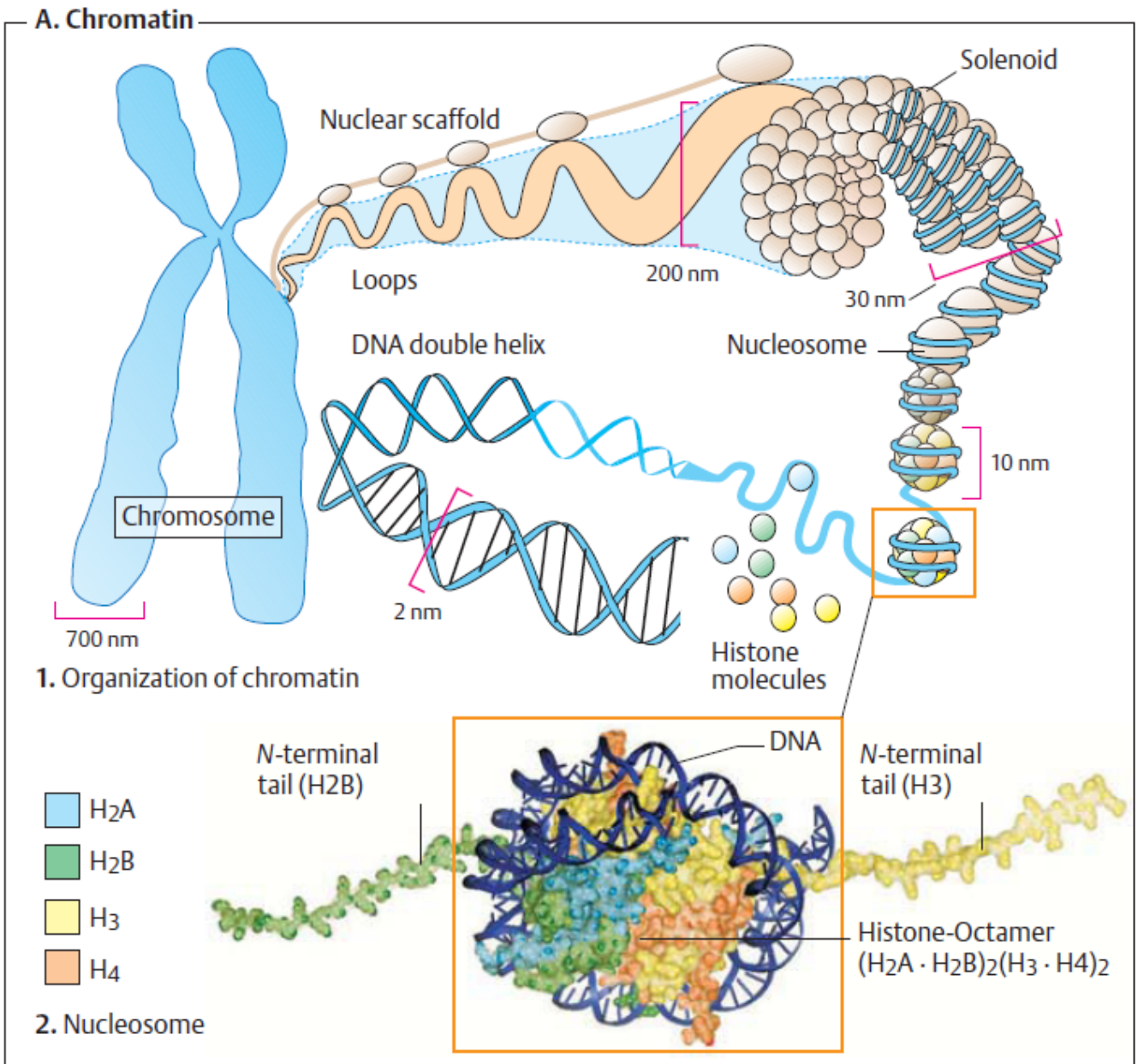


α -helix obvykle interaguje s DNA ve velkém žlábků (osa helixu ve směru dlouhé osy žlábků)

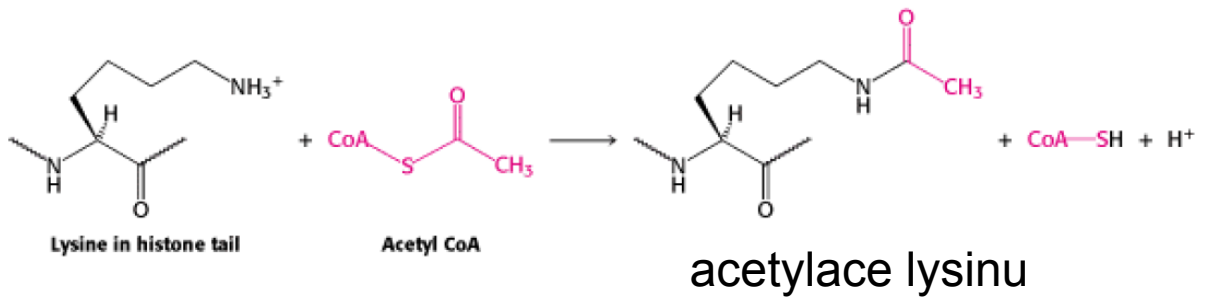
-postranní skupiny proteinu mohou interagovat s fosfátovými skupinami i pronikat žlábkem dovnitř dvoušroubovice a interagovat s páry bází (pravděpodobná tvorba specifické vazby)

β -listu se předpokládá nescifická vazba do malého žlábků, kde vytváří vodíkové vazby amidoskupin peptidové vazby a fosfáty

Interakce DNA s bazickými (= pozitivně nabitými) proteiny určuje strukturu chromatinu (obecně umožňuje kondenzaci DNA do malého prostoru)



Posttranslační modifikace histonů ovlivňuje jejich bazicitu a tudíž interakce s DNA, strukturu chromatinu a genovou expresi



B. Histone

	A	P	A	A	A	A	A	M																		
	Acetylation		Phosphorylation				Methylation																			
Animals	S	G	R	G	K	G	G	K	G	L	G	K	G	G	A	K	R	H	R	K	V	L	R	D	N	I
Plants	S	G	R	G	K	G	G	K	G	L	G	K	G	G	A	K	R	H	R	K	V	L	R	D	N	I
Yeast	S	G	R	G	K	G	G	K	G	L	G	K	G	G	A	K	R	H	R	K	I	L	R	D	N	I
Animals	Q	G	I	T	K	P	A	I	R	R	L	A	R	R	G	G	V	K	R	I	S	G	L	I	Y	E
Plants	Q	G	I	T	K	P	A	I	R	R	L	A	R	R	G	G	V	K	R	I	S	G	L	I	Y	E
Yeast	Q	G	I	T	K	P	A	I	R	R	L	A	R	R	G	G	V	K	R	I	S	G	L	I	Y	E
Animals	E	T	R	G	V	L	K	V	F	L	E	N	V	I	R	D	A	V	T	Y	T	E	H	A	K	R
Plants	E	T	R	G	V	L	K	I	F	L	E	N	V	I	R	D	A	V	T	Y	T	E	H	A	R	R
Yeast	E	V	R	A	V	L	K	S	F	L	E	S	V	I	R	D	S	V	T	Y	T	E	H	A	K	R
Animals	K	T	V	T	A	M	D	V	V	Y	A	L	K	R	Q	G	R	T	L	Y	G	F	G	G		
Plants	K	T	V	T	A	M	D	V	V	Y	A	L	K	R	Q	G	R	T	L	Y	G	F	G	G		
Yeast	K	T	V	T	S	L	D	V	V	Y	A	L	K	R	Q	G	R	T	L	Y	G	F	G	G		

Amino acid sequence of histone H4

INTERAKCE MOTIVU HELIX-OTÁČKA-HELIX S DNA

- dvacetiaminokyselinový motiv nalezený u řady proteinů interagujících s DNA
- 8 AK v helixu, pak pravotočivá otáčka ze tří AK a další helix z 9 AK
- přičemž jsou konzervovány: alanin v pozici 5, glycin 9 (první AK v otáčce) isoleucin nebo valin v 9 (tyto AK zřejmě udržují strukturu motivu)
- v ostatních pozicích vysoká heterogenita (zachována je hydrofilita těch částí helixů, které směřují „ven z proteinu“ do rozotku nebo k DNA; „vnitřní“ povrchy helixů jsou hydrofobní)

- interaguje s DNA ve *velkém žlábkku*, jemuž odpovídá rozměry (typický α -helix má průměr 1.2 nm, velký žlábek je v B-DNA 1.2 nm široký a 0.6 - 0.8 nm hluboký)
- pokud je α -helix paralelně se žlábkem, může interagovat s **4-6 bp**, než žlábek „zahne“; na této délce dochází k rozpoznání specifické sekvence určitými AK

λ -repressor

-236 AK protein, dimerní
 -dvě funkční domény: N-koncová (AK 1- 92) se váže na DNA, C-koncová má dimerizační funkci

-N-koncová doména sestává 5 α -helixů; z nich 2. a 3. a pět AK mezi nimi vytváří *helix-otáčka-helix* motiv

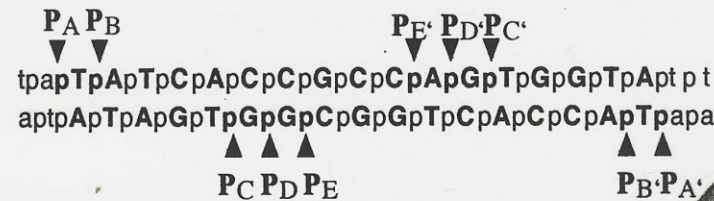
-řada specifických vodíkových vazeb na bp (Asn, Gln, Ser, Lys)

-Lys 4 tvoří *solný můstek* G v šestém bp

-*van der Waalsovy kontakty* (mezi hydrofobními skupinami proteinu a metylskupinami T)

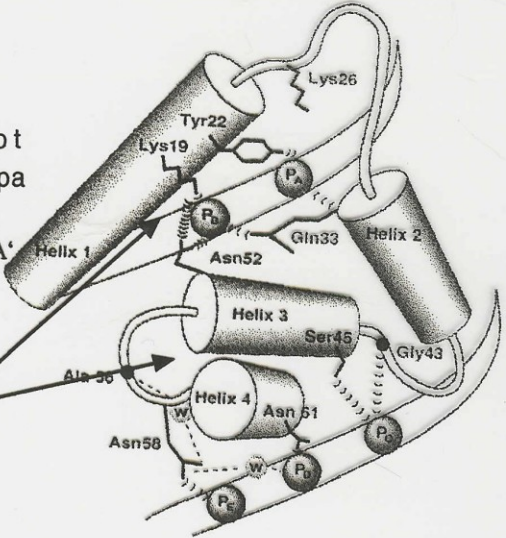
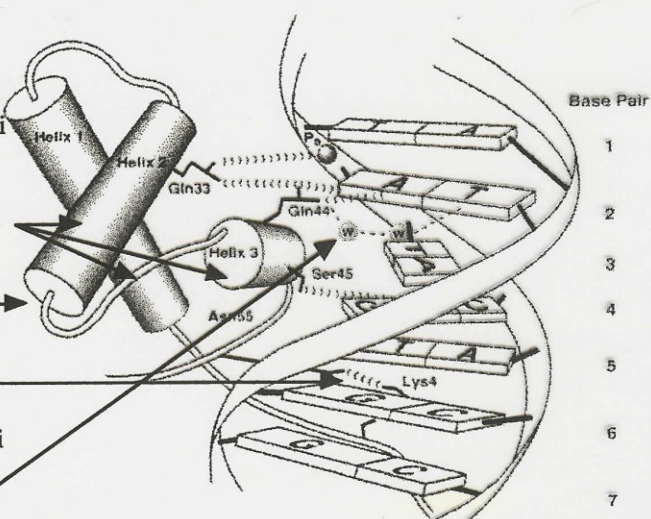
-vazby zprostředkované *molekulami vody*

-interakce helixu 3 s *fosfátovými skupinami*::



vzhledem k otáčce dvoušroubovice jsou vyznačené fosfáty A-E a A'-E' vždy ve stejném místě velkého žlábků (značení odpovídá převrácené symetrii operátoru)

- vodíkové vazby s fosfáty tvoří *dvě třídy*:
- a/* vazby ve velkém žlábkku tvořené motivem *helix-otáčka-helix* a helixem 4
- b/* mimo velký žlábek od aminokyselin z helix 1 a AK mezi helixy 1 a 2



INTERAKCE Zn-VAZEBNÝCH DOMÉN S DNA

-velký počet proteinů specificky interagujících s DNA váže zinečnaté ionty (často regulační proteiny)

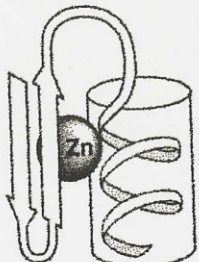
-tři typy Zn-vazebných domén:

Zinkový „prst“ - Zn(Cys₂-His₂)

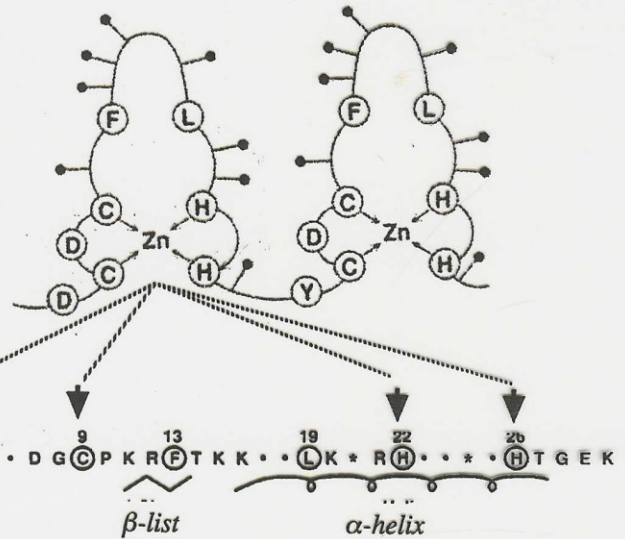
transkripční faktor IIIA:(1985)

třicetiaminokyselinový repetitivní motiv, v němž jsou konzervovány tyto aminokyseliny:

-cysteiny a histidiny koordinují zinečnatý ion



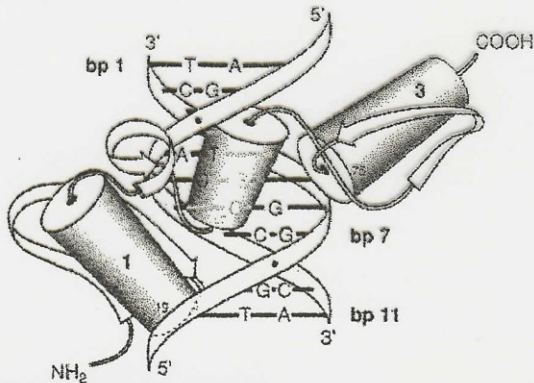
β -list α -helix



zbytky 2-14 a 3-17 tvoří *antiparalelní* β -list, z jehož jedné strany váží cysteiny 4 a 9 Zn

zbytky 16-27 tvoří α -helix, který je „přidržován“ podél β -listu koordinací histidinů 22 a 26 se Zn

V TFIIIA je devět těchto motivů, vazebné místo zahrnuje 27 bp -na obou koncích *tři* vnější „obtačejí“ dvoušroubovici a sledují směr osy velkého žlábků (interagují vždy s 10 bp - trojice) -*tři* centrální interagují napříč malým žlábkem z jedné strany dvoušroubovice:



Protein Zif268 : kokystalová struktura

-tři Zn-prsty váží 11 bp sekvenci A GCG TGG GCG T, každý k jednomu podtrženému trinukleotidu

-protein „obtač“ DNA podél *velkých žlábků* a váže se na *guaniny* v rekogniční sekvenci

-N konec každého α -helixu směřuje do velkého žlábků

-před helixem každého prstu je *arginin* (polohy 18, 46, 84), který tvoří *vodíkovou vazbu* s *posledním* guaninem každého tripletu

-helix prvního a třetího prstu má uprostřed další *arginin* (24 a 80), vázající se na *první* guaniny příslušných tripletů; v prostředním helixu je *histidin* (49), vázající *centrální* guanin

-množství vazeb na fosfáty

-struktury zcela odlišná od helix-otáčka-helix!

Zn.,twist“ - [Zn(Cys4)]₂ doména

receptory pro steroidní hormony (glukokortikoidy, estrogeny):(1990)

-šedesátiaminokyselinová vazebná doména, obsahuje dvě oblasti -

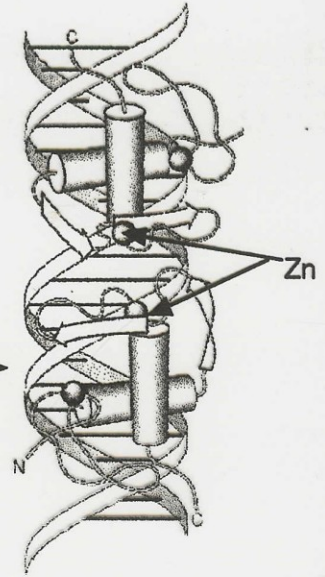
každá s dvěma cysteiny; vysoce konzervativní

-vazebná doména je uspořádána do globulární struktury tvořené dvěma na sebe kolmými α -helixy

-Zn je vázán poblíž N-konce α -helixů, které obsahují množství bazických zbytků soustředěných na jedné straně helixů

-tyto zbytky tvoří množství specifických vodíkových vazeb ve velkém žlábků DNA

- vazebné místo DNA je dimerní, zcela symetrické



Zn.,klastř“ - Zn₂(Cys₆) doména

kvasničný transkripční faktor GAL4

-třicetiaminokyselinová vazebná doména, obsahuje šest cysteinových zbytků

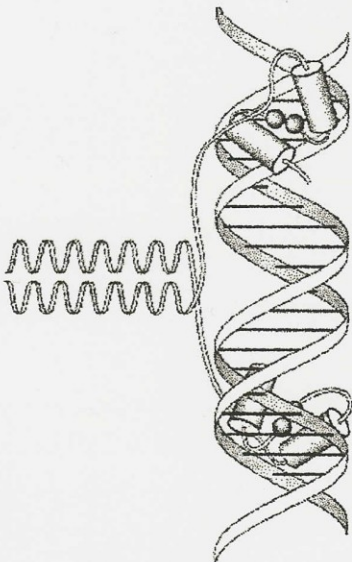
-rozmístění cysteinů je vysoce konzervativní u řady druhů kvasinek a hub

-vždy dvojice cysteinů váže zinek, dva atomy Zn jsou spojeny můstkem z další dvojice: Cys₂-Zn-Cys₂-Zn-Cys₂

-cysteiny jsou organizovány okolo dvoujaderného zinkového klastru do dvou α -helixů, spojených smyčkou na druhé straně atomů zinku (po dvou cysteinech v helixech, dva ve smyčce)

- jeden z helixů interaguje s DNA ve velkém žlábků s tripletem CCG

-vazebné místo DNA je dimerní (celkem 17 bp), podjednotky spojené „svinutým klubkem“ (coiled coil) hydrofobních zbytků interagují z opačných stran dvoušroubovice (CCG triplety jsou separovány 1.5 otáčky dvoušroubovice)



LEUCINOVÝ „ZIP“

Trankripční faktory rodiny bZip (kvasničné faktory GCN4, onkoproteiny fos a jun)

-obsahují dimerizační doménu a bazický motiv interagující s DNA

-**dimerizační doména** - tzv. **leucinový zip** - je tvořena

amfipatickým α -helixem, ve kterém je každá sedmá AK leucin:

protože otáčka α -helixu je 3.6 AK, vychází všechny leuciny *na jednu stranu helixu* (ob otáčku)

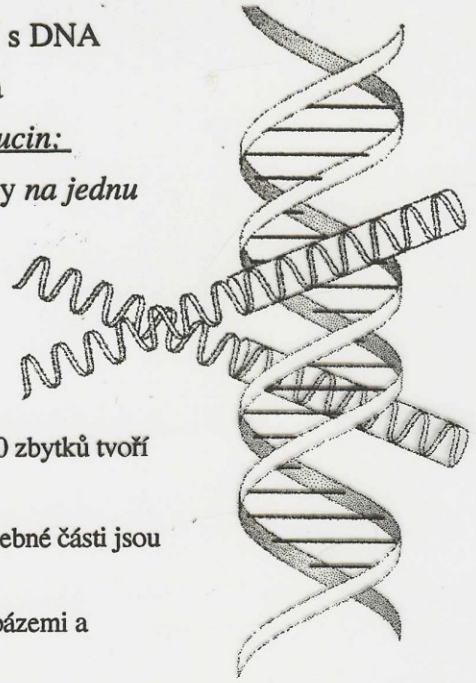
-dimerizační domény obou podjednotek spolu interagují hydrofobními vazbami leucinů, tvoří „svinuté klubko“

-**vazebná doména** obsahuje bazické AK zbytky

-monomer GCN4 obsahuje 56 AK a tvoří jeden kontinuální α -helix; 30 zbytků tvoří leucinový zip, 25 zbytků bazickou doménu

-“zip“ je v komplexu s DNA orientován kolmo k dvoušroubovici a vazebné části jsou „nastaveny“ do velkého žlábků na opačných stranách

-ve velkém žlábků tvoří vodíkové vazby zbytky Asp, Ser, Arg, Lys s bázemi a fosfáty



TATA-box -vazebné proteiny

-TBP: DNA-vazebná podjednotka proteinového komplexu, který rozpoznává TATA-box (8 bp)

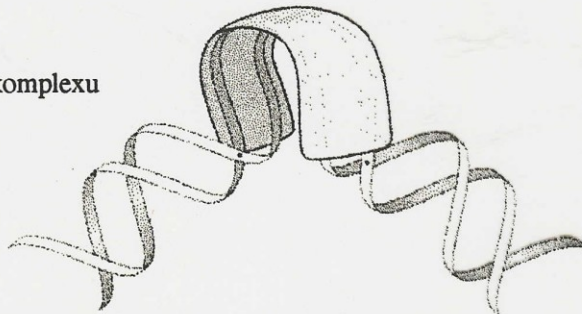
-TBP je tvořen *sedlovitě zakřiveným β -listem z 10 antiparalelních řetězců* ; na vrcholu této struktury jsou 4 α -helixy

-při tvorbě komplexu se DNA zakříví podle tvaru proteinu (o 100° ve směru malého žlábků); to způsobí odvinutí TATA boxu o 110°

-malý žlábek je otevřen, bp jsou exponovány uvnitř „sedla“

- na koncích TATA-elementu jsou mezi bp (tj bp 1,2 a 7,8) vsunuty dvě dvojice zbytků fenylalaninu

- DNA v těsném sousedství komplexu je normální B-forma!



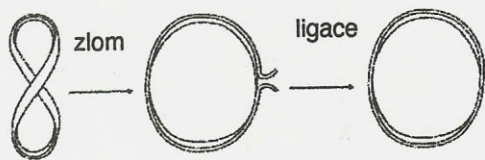
INTERAKCE β -LISTU S DNA

β -list obvykle leží v **malém žlábk**u a tvoří *vodíkové vazby s fosfáty* DNA

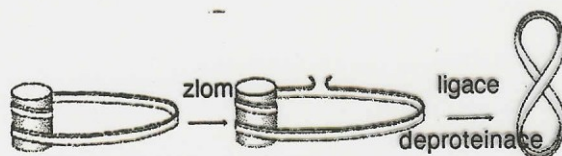
-dvě možné orientace: výhodnější interakce, je-li *N-konec β -listu ve směru 3' \rightarrow 5'* toho DNA řetězce, se kterým tvoří vodíkové vazby

HU-proteiny: rodina malých bazických bakteriálních proteinů

-váží se na DNA nespecificky, indukují ohyby, „konzervují“ negativní superhelicitu in vivo (jejich vazbou na negativně sc DNA se relaxuje zbytek molekuly, takže molekula není relaxována topoisomerázou)



volná molekula scDNA zavedením jednořetězcového zlomu zrelaxuje, superhelicita se ztratí



vazbou na HU-protein se spotřebuje část superhelicity (tj. ΔW_r , zakřivení DNA okolo proteinu); volná část molekuly je relaxovaná, takže její topologický stav se zavedením zlomu nezmění: superhelicita je „uvězněna“ (restrained) ve vazbě na HU protein

v HU-proteinech je β -list ze tří řetězců;

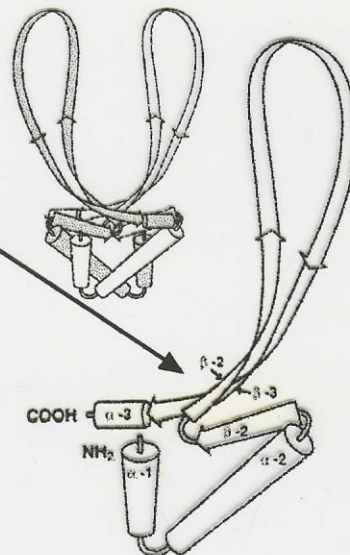
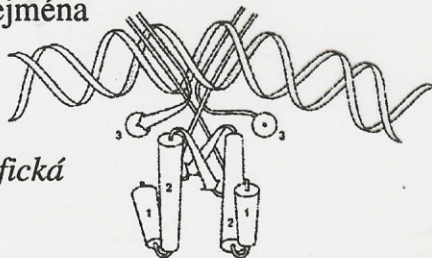
dva (antiparalelní) z nich pokračují v dlouhé rameno tři α -helixy vytvářejí strukturu, která interaguje s druhou podjednotkou za tvorby dimeru

v dimeru tvoří ramena „račí klepato“, které obejmě dvoušroubovici DNA;

β -listy interagují v *malém žlábk*u (ramena jsou zahnutá tak, aby byla interakce co nejvýhodnější, a obsahují vysoce konzervované zbytky AK, zejména

bazické: Arg, Lys, které interagují s *fosfáty*)

- vazba je *sekvenčně nespecifická*



sekvenčně specifická interakce β -listu:

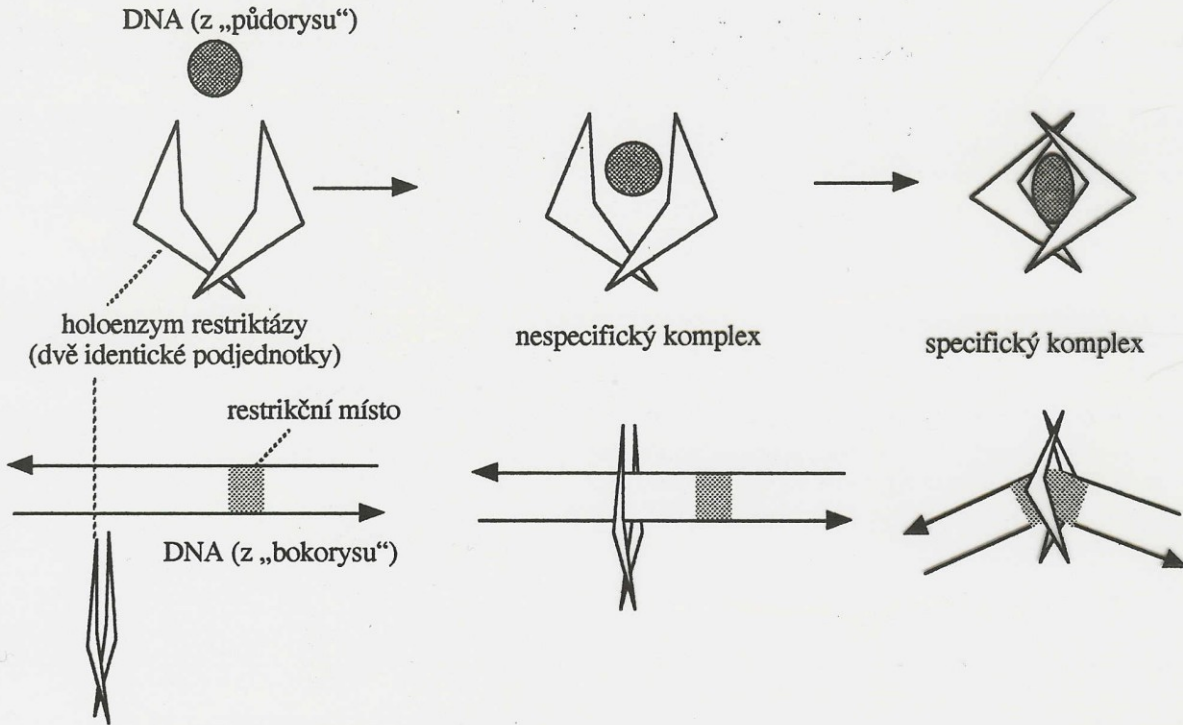
např, *MetJ-represor* z *E.coli*: β -list tvoří ve *velkém žlábk*u speciifický kontakt s palindromickou sekvencí AGACGTCG; ta je 5x tandemově opakována

vodíkové vazby tvoří zbytky Lys a Thr z β -listu s druhým a třetím nukleotidem (G a A)

Interakce restriktáz typu II s DNA

- velká podobnost jednotlivých enzymů
- společné reakční schéma:

nespecifická interakce kdekoli na molekule -> *lineární difuze* k restriktčnímu místu -> tvorba *specifického komplexu* spojená s *konformační změnou* proteinu i DNA -> *štěpení DNA*



Nespecifická vazba

- relativně *slabá* za podmínek optimálních pro štěpení (tj. mM Mg^{2+})
- *silnější* v nepřítomnosti hořčíku
- *značně závisí na koncentraci solí*: to ukazuje na značný elektrostatický příspěvek (tj. iontové interakce s fosfáty)

Eco RV: nejlépe prostudovaná:

- komplex je držen pohromadě *pěti kontakty aminokyselin* z každé podjednotky s *cukrfosfátovou kostrou*

- katalytické centrum *není v nespecifickém komplexu v aktivní konformaci* a ty části molekuly, které interagují s restriktčním místem, nejsou „správně“ uspořádány

- v těchto částech je volně vázáno velké množství molekul vody (~100 na holoenzym) a iontů, které jsou při tvorbě specifického komplexu uvolněny

(interference s vazbou/uvolněním molekul vody, např. osmotický tlak, je zodpovědná za tzv. „hvězdičkovou aktivitu“, tj. méně stringentní štěpení)

Lineární difuze

-“nasednutí“ enzymu na DNA v kterémkoli místě a hledání restrikčního místa podél řetězce (= redukce na jednorozměrný problém) zvyšuje pravděpodobnost a rychlost jeho nalezení např. *EcoR I* postupuje rychlostí 7×10^{-6} bp/s !!

- enzym „klouže“ podél molekuly DNA a opisuje helikální zkrut; děj je podmíněn citlivou rovnováhou přitažlivých a odpuzivých elektrostatických sil

(alternativní mechanismus - „skákání“, tj. mikroskopické disociace/asociace- není pravděpodobný: enzym „nepřehlédne“ žádné restrikční místo a může být zablokovan neobvyklou strukturou DNA, např. triplexem; navíc se „odrážá“ od konců lineární DNA)

-“hvězdičková“ místa enzym „zdrží“ až na 20 s (i za optimálních podmínek, kdy nejsou štěpena; enzym „ověruje“, zda není na „své“ sekvenci, a poté pokračuje. Čas, po který se „zdrží“, koreluje s podobností místa ke kanonické sekvenci a tedy s relativní afinitou enzymu k danému místu)

Specifická vazba a rozpoznání restrikčního místa

-nejlépe prostudován je enzym *EcoRV* (GATATC)

- ve specifickém komplexu je DNA ohnuta o 55° , čímž je místo „odvinuto“, u centrálních bp porušena stacking interakce

- komplex je **těsný**, enzym těsně „obejme“ DNA kolem dokola

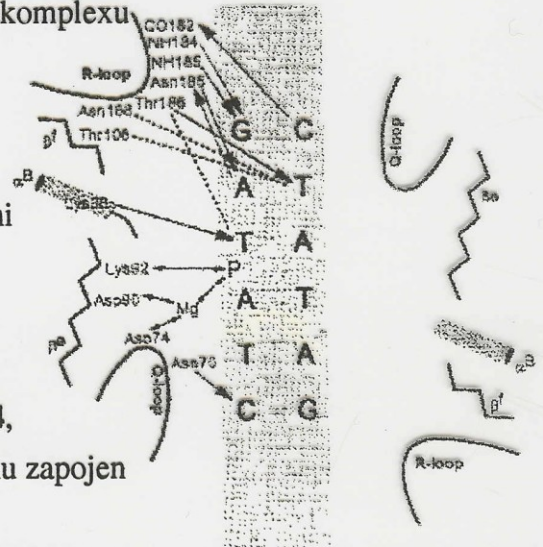
- v molekule enzymu tři smyčky, které v nespecifickém komplexu jsou neuspořádané, vytvoří specifický kontakt s DNA ve velkém a malém žlábků

-tzv. R-smyčka vytvoří 12 přímých vodíkových vazeb s bázemi, 12 vazeb zprostředkovaných molekulami vody s fosfáty a dva van der Waalsové kontakty s vnějšími thyminy

-Q-smyčka („bohatá na glutamin“) tvoří čtyři vodíkové vazby s hranami bází v malém žlábků a obsahuje **Asp74**, který má katalytickou funkci; tam je rovněž do komplexu zapojen **hořčnatý ion**

-centrální bp nevstupuje do žádné přímé interakce (tam je v důsledku ohybu hluboký a úzký malý žlábků)

-dalších 24 aminokyselin (mimo R a Q smyčky) tvoří vodíkové nebo iontové můstky s fosfáty



interakce je **vysoce specifická**: k téměř úplné *inaktivaci* komplexu vede *řídzená substituce*

aminokyselin, které tvoří vodíkové vazby s bázemi, *mutace restrikčního místa* (včetně centrálních bp, které netvoří přímé vodíkové vazby), a ve většině případů i *chemické změny fosfátových skupin* (nahrazení kyslíků) a *substituce aminokyselin s nimi interagujících* (např. za alanin)

DIMERNÍ VAZEBNÁ MÍSTA

- častý jev zejména u regulačních proteinů (dimerní rekogniční místo interaguje s dimerní nebo tetramerní molekulou proteinu)
- tím se zvyšuje afinita vazby (více vazebných míst neboli větší vazebné místo znamená více interakcí) a její specificita (protein se musí navázat na obě místa současně)

organizace dimerních vazebných míst: v zásadě je možná jako

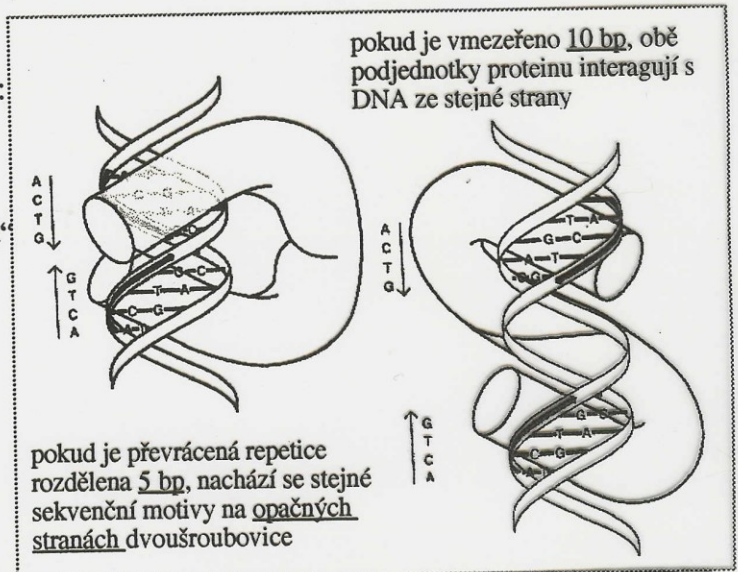


-nejobvyklejší je *převrácená* repeticie (*lac operátor*, λ -*operátor*; *restriktázy*):

nejvíce symetrie

-*přímé* (*tandemové*) repeticie jsou typické pro proteiny s tzv. „*Zn-prstec*“

-interakce dimerního proteinu se *zrcadlovou* repeticí je málo pravděpodobná z důvodu nedostatku symetrie takové sekvence v dvoušroubovici (nicméně případy jsou známy)



SEKVENČNĚ NESPECIFICKÉ INTERAKCE

-využívají zejména cukrfosfátové kostry, která má *relativně uniformní tvar* (s výjimkou lokálních změn konformace, které jsou závislé na určité sekvenci) a *negativní elektrický náboj*

-DNA-polymeráza I: (musí „projít“ celou molekulou a kontaktovat každý bp), interaguje s fosfátovými skupinami

-nespecifická vazba restriktáz typu II: silná závislost na iontové síle ukazuje na značný podíl elektrostatické interakce; u *EcoRV* bylo identifikováno 5 vodíkových vazeb na fosfátové skupiny (žádná na báze)

-bakteriální HU protein: účast β -listu zanořeného do malého žlábků, který interaguje vodíkovými vazbami mezi argininovými zbytky a fosfátovými skupinami

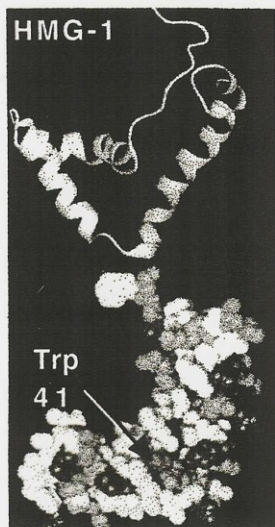
-DNáza I: vodíkové vazby mezi 10 aminokyselinami a fosfátovými zbytky; kromě toho stacking interakce mezi tyrosinovým kruhem a pyrimidinovým zbytkem a 3 vodíkové vazby mezi argininovým zbytkem a O-atomy v pyrimidinovém zbytku

Proteiny s motivem HMG-boxu (High Mobility Group)

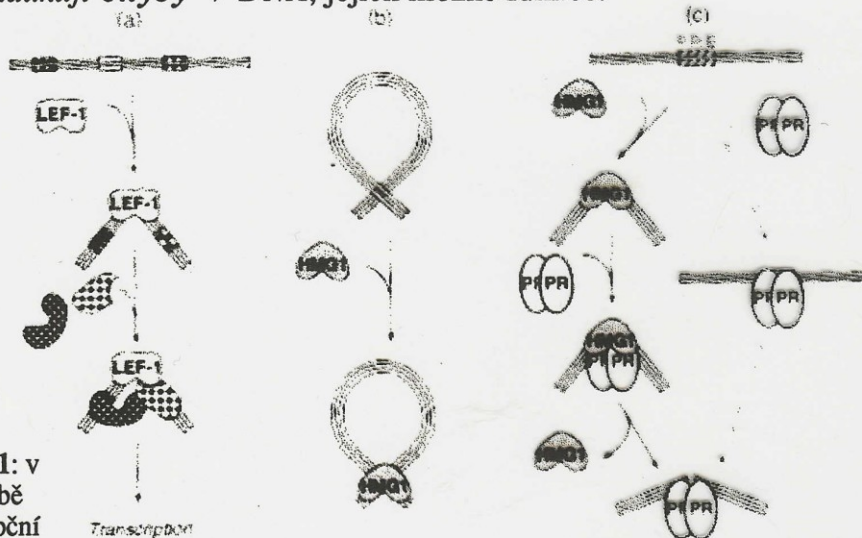
- DNA vazebná doména společná pro řadu proteinů *chromatinu*, obecně *transkripčních faktorů* a transkripčních regulátorů, které řídí *diferenciaci a determinaci pohlaví*
- rozpoznávají zejména **konformační rysy DNA** (některé se váží *sekvenčně specificky*)
- HMG-box tvoří asi 70 *aminokyselin*; vysoký podíl *prolinů, aromatických a bazických aminokyselin*
- konzervované zbytky: aromatický v poloze 11, Trp 41, aromat. 52; několik dalších méně konzervativních
- mutace v HMG-boxu dramaticky snižují afinitu k DNA
- celkově je HMG box spíše hydrofilní, s hydrofobním úsekem na N-konci

-3 skupiny:

- 1-savčí HMG-1 a podobné chromatinové proteiny (ne zcela jasná funkce, snad mají úlohu při transkripci tím, že vážou histony a napomáhají tvorbě iniciačního komplexu)
- 2-transkripční faktory RNA polymerázy I a mitochondriální proteiny (včetně TBP)
- 3-poměrně heterogenní skupina zahrnující mj. savčí faktor určující samčí pohlaví (SRY)



indukují ohyby v DNA; jejich možné funkce:



transkripční faktor LEF-1: v důsledku ohybu se k sobě přiblíží dva další transkripční faktory

HMG-1: možná stabilizace smyček DNA v důsledku „svázání“ dvou vazebných míst (tato struktura připomíná čtyřcestné spojení, které HMG-1 in vitro rozpoznává)

progesteronový receptor (PR) s navázaným hormonem se váže efektivně na příslušný DNA element v přítomnosti HMG-1 (ale ne v jeho nepřítomnosti); ten mu může DNA „vytvarovat“