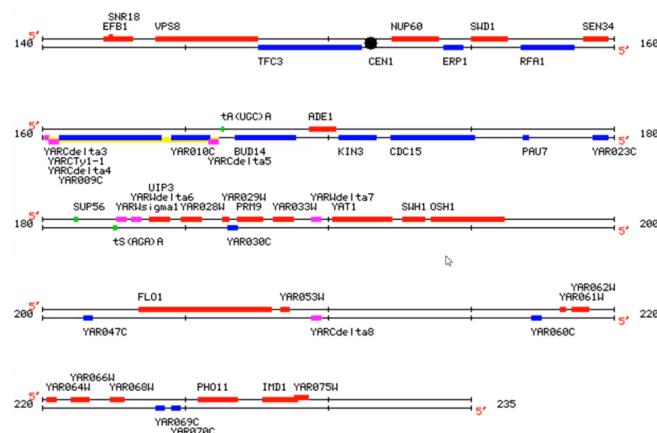


# Transkripce

- = přepis informace z DNA do RNA
- vlastnosti buňky (a její fenotyp) jsou dány vlastnostmi biopolymerů a biopolymerních struktur vyššího řádu (ve smyslu původu struktur pak především proteinů – největší podíl efektorových a strukturních molekul uvnitř buňky)
- čím jsou dány vlastnosti proteinů?
- kde se proteiny berou?
- jak jsou zajištěny správné koncentrační poměry proteinů, popřípadě jejich syntéza de novo v daném (správném) okamžiku či jako odpověď na daný podnět?
- regulace – viz konec přednáškového cyklu

## Gen

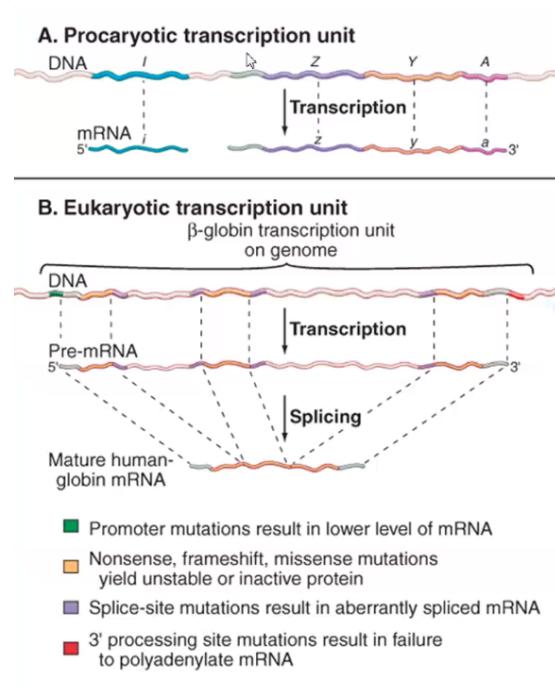
- definice genu – organizovaný úsek nukleové kyseliny projevující se a přenášející se jako základní jednotka dědičné (genetické) informace (kus NK, který se dědí a má nějaký projev)
  - (donedávna byla definice jednodušší, ale postupem času se zesložíuje)
  - dědí se
  - z hlediska transkripce nás zajímají nejvíce první 3 „druhy“ genů: gen kódující protein, RNA, regulační nebo strukturní oblast
  - gen kódující protein – kóduje polypeptid (Mendel, Morgan, Beadle, Tatum)
  - gen kódující funkční RNA – kóduje tRNA, rRNA, miRNA, snRNA a řadu dalších RNA
  - gen jako regulační nebo strukturní oblast – promotory, enhancery, centromery, telomery
  - gen jako dědičná variabilní oblast nukleové kyseliny – satelitní polymorfní DNA
- nejekonomičtější způsob regulace projevu (exprese) genů kódujících proteiny je na samém počátku = regulace syntézy mRNA, tedy regulace transkripce
- množství exprese na úrovni transkripce je dáno silou promotorů a genovou dózí
- fenotypy organismů v případě, že změníme regulační oblasti: viditelný vliv; pokud jsme schopni oblast lokalizovat a jsme schopni sledovat její dědičnost -> gen jako regulační nebo strukturní oblast
- alela = varianta genu; -> variabilní oblast NK může být považována za gen (nese s sebou určitý projev)
- geny kódující protein / RNA: jediné geny, které jsou přepisovány
- některé regulační oblasti jsou také přepisovány (v tu chvíli se ale chovají jako gen kódující funkční RNA / bílkovinu – „dvojatý“ gen)
- pouze geny kódující polypeptidy – využití přepsané RNA k translaci (jediný druh genu, který je takto využíván)



- SGD Chromosome I features between coordinates 141453 - 230203 bp Oct 20 2002
  - chromozom *Saccharomyces cerevisiae* s vyznačenými různými druhy genů
  - gen jako polymorfní oblast, která má nějakou funkci – transpozon TI11 (funkce se dá sledovat mezi jednotlivými kmeny a pokoleními kvasinky)

## Transkripcie

- je katalyzována DNA programovanou RNA polymerázou
- probíhá pouze ve směru 5' → 3'
- nesymetrická* – přepisováno pouze jedno vlákno DNA (minus vlákno)
- vlákna DNA na kódující DNA můžeme z hlediska kódující kapacity RNA (zejména se to týká RNA kódující bílkoviny) popsat jako + vlákno (*kódující, sense*) – **má shodnou sekvenci s přepisovanou RNA** (-> přepisuje se minus vlákno, *antisense, nekódující*)
- oblast začátku transkripce = *promotor*
- oblast ukončení transkripce = *terminátor* (ne vždy je přítomen!)
- oblast přepisovaná RNA polymerázou řízena jedním promotorem = *transkripční jednotka*
- start transkripce*
- číslování nukleotidů* – pokud mluvíme o transkripci, první nukleotid přepisované RNA odpovídá bodu startu na DNA (+1) (0 neexistuje – jen +1 a -1!)
  - sekvence před tímto bodem proti proudu transkripce = upstream (-)*
  - sekvence za tímto bodu po proudu transkripce = downstream (+)*
  - je důležité přemýšlet nad tím, od čeho upstream/downstream (je podstatné to zdůrazňovat, aby v tom nebyl bordel)
- čerstvě přepisovaná RNA (*primární transkript*) je obvykle v této formě velmi nestabilní:
  - řada dalších dějů navázaných na transkripci probíhá v průběhu transkripce
  - primární transkript = nascentní transkript*
  - bakterie, archea – velmi často rychle degradován – regulace genové exprese je zaměřena na iniciaci transkripce
  - eukaryota – častěji dochází k posttranskripčním modifikacím – mRNA je stabilnější
    - průměrný poločas života mRNA transkriptu eukaryot je cca 12 minut (-> potřeba další stupeň regulace přímo na transkriptu / regulace překládané bílkoviny -> regulace u eukaryot posunutá dál než u bakterií a archeí)
    - > regulace a váha regulací musí být u různých domén života různě zaměřena



© Elsevier B.V. All rights reserved. www.sciencedirect.com

- o bakterie, archea – *polycistronní transkripcie* – přepis dlouhého úseku (více genů), následně se překládají jednotlivé geny (pod jeden promotor spadá více produktů – např. Lac operon)
- o eukaryota – typická transkriční jednotka je *monocistronní* – čím výše v evoluci, tím delší jednotky, ve kterých menší procento zabírájí vlastní transkriční jednotky (přerušeny úseky, které nekódují vlastní protein) – až v procesu sestřihu (splicingu) vzniká monocistronní mRNA, která dává vznik jedné bílkovině

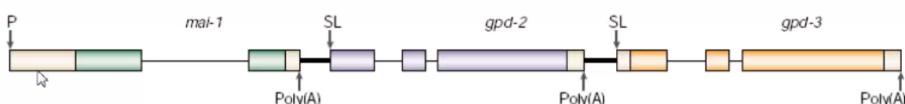
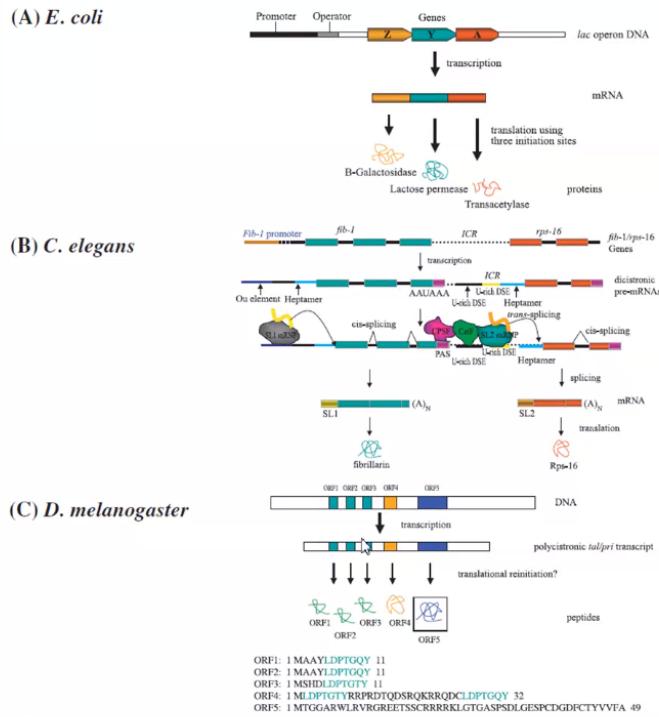
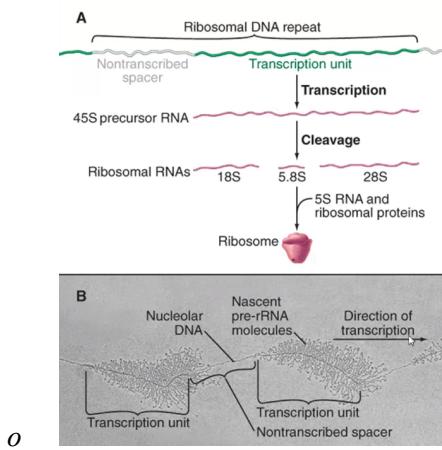


Figure 1 | The first operon identified in *Caenorhabditis elegans*. The 5'-most gene is mitochondrial ATPase inhibitor-1 (*mai-1*). This gene is not *TRANS-SPliced*. The two downstream genes, glyceraldehyde 3-phosphate dehydrogenase-2 (*gpd-2*) and glyceraldehyde 3-phosphate dehydrogenase-3 (*gpd-3*), encode isoforms of a glycolytic enzyme. The two intercistronic regions are ~100 bp long (bold). Exons and introns are drawn to scale. Coding regions are depicted by green, purple and orange boxes; non-coding regions are depicted by tan boxes. Arrows indicate the sites of 3'-end formation (poly(A) addition) and *trans*-splicing (addition of spliced leader (SL)). The single promoter is marked with a P.

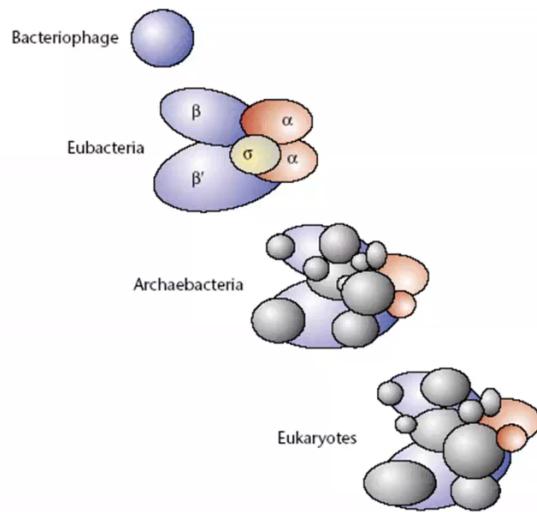
- o tricistronní transkriční jednotka *Caenorhabditis elegans* – následné translaci jsou nabídnuty monocistronní mRNA (jsou transkribovány společně a v rámci posttranskričních úprav jsou sestříhaný introny a transkript je rozdělen na vlastní části pro translaci)
- o -> nemůžeme tvrdit, že se u eukaryot polycistronní transkriční jednotky nevyskytují (Caenorhabditis elegans – mnohdy velmi dlouhé transkriční jednotky – až 8 genů v jedné jednotce – větší než u řady bakteriálních transkričních jednotek!)
- o polycistronní transkriční jednotky jsou typické pro tato eukaryota: Trypanosomy, Nematoda (zejména *Caenorhabditis*)



- *D. melanogaster* – receptory účastníci se chuti, rozpoznávání vůni – v polycistronních transkripcích jednotkách; jak je transkrypt upraven a co je následně přeloženo určuje poměr receptorů a jejich skládání do vyšších struktur -> zvyšuje možnost kombinace rozpoznání různých cukrů a chutí
- bicistronní jednotky najdeme prakticky u všech organismů včetně člověka!
  - o eukaryota – bud' se použijí oba cistrony, nebo jen jeden (tak trochu alternativní splicing)
- -> nemůžeme tvrdit, že eukaryota mají monocistronní transkripcní jednotky (ale můžeme tvrdit, že *eukaryota mají převážně monocistronní, bakterie+archea převážně polycistronní*)
- polycistronní jednotka je regulována jedním promotorem -> pro buňku je výhodné, aby v reakci na určitý signál spustila naráz transkripci celé sady genů, jejichž produkty jsou na tento signál potřebné (vyšší eukaryota – vyšší poločas existence transkriptů -> je lepší regulovat každý transkrypt zvlášť – je možné, aby se za některých podmínek některé transkripty přepisovaly spolu, jindy ne, jindy třeba s různou intenzitou)
- geny kódující rRNA – u všech organismů polycistronní



- o nepřepisovaný mezerník (*nontranscribed spacer*) – může oddělovat geny v rámci skupiny, která se přepisuje společně

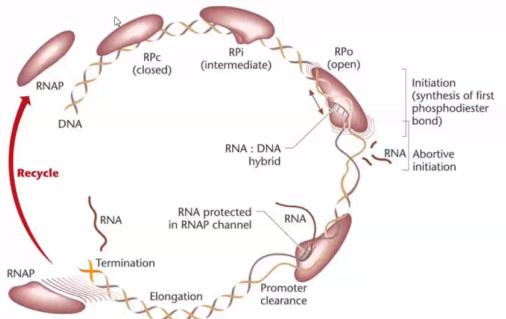


**Figure 1** Subunit complexity of RNA polymerases. The sizes of the subunits are approximate; the diagram is a generalization and is not intended to accurately display precise subunit interactions or enzyme structure.

- - jak vypadají RNA polymerázy u různých organismů
  - viry: RNA polymeráza může být jednoúčelový rychlý enzym syntetizující příslušnou RNA podle virového / fágového genomu, jednopodjednotkové
  - bakterie – variabilní složení, které umožňuje rozeznávat konkrétní typy promotorů
  - eukaryotní RNA polymerázy jsou složitější, multipodjednotkové
  - archeální RNA polymerázy jsou více podobné eukaryotním než bakteriálním

## Bakteriální typ přepisu

- přepis není izolován od dalších buněčných dějů (translace)
- přepis je úzce spřažen s translací; tyto děje jsou u bakterií na sobě závislé, může docházet k regulacím na rozhraní přepisu a translace
- jediná RNA polymeráza
- *relativně nízká variabilita v promotorech* (v rámci jednoho druhu bakterií) a omezený počet dalších transkripčních faktorů
- mRNA obvykle není posttranskripčně modifikována – je ihned použita
- primární transkripty jsou obvykle polycistronní
- hned po zahájení syntézy se na nascentní konec transkriptu váže ribozom -> syntéza polypeptidu (-> není smysluplná posttranskripční modifikace)
- velké rozdíly v sekvencích promotorů a v AK sekvencích RNA polymeráz a transkripčních kofaktorů mezi jednotlivými druhy bakterií
- struktura RNA polymeráz – jsou si velmi podobné, podobné struktury katalytických jader RNA polymeráz



**Figure 1** The transcription cycle. RNA polymerases (RNAPs) have an intrinsic affinity for DNA. Promoter-specific DNA binding by RNAPs is facilitated by recruitment factors unique to each RNAP. Once RNAP binds promoter DNA it forms a stable complex with the closed duplex DNA (RPe). The formation of RPe promotes a conformational change between RNAP and the DNA that results in an intermediate complex (RPI) that rapidly converts to an open state (RPO). Elongation of the RNA chain results in a ternary complex that catalyzes RNA synthesis. However, RNAPs engage in non-productive cycles of abortive RNA synthesis. Transitions to a productive cycle occur when the RNA chain reaches a critical length. At this point, called promoter clearance, RNAP disengages from the promoter and elongates the RNA chain. During the elongation phase, RNAP is subject to DNA and protein influences that can effect the kinetics of RNA synthesis. Once RNAP reaches termination sequences in the DNA, the ternary complex dissociates liberating the newly formed RNA chain. The released RNAP can then recycle for a subsequent round of RNA synthesis.

- bakteriální RNA polymeráza má na rozdíl od eukaryotní RNA polymerázy *vnitřní afinitu k DNA* -> většinu času vázány někde na nukleoidu (když je DNA promotor volný, tak se na něj naváže RNA polymeráza)
- specifická asociace s promotorem – zajišťují specifické bílkoviny (sigma faktor – součást funkční RNA polymerázy, určuje, že na tomto konkrétním promotoru bude probíhat transkripce)
- RNA polymeráza má vlastní afinitu k jakékoli části DNA, může dokonce docházet k abortivní transkripcí -> *hlavní způsob regulace je represe* (navázání faktorů, které brání nasednutí RNA polymerázy – úplný opak regulace eukaryotních RNA polymeráz) (*pokud nemá probíhat transkripce, musí být promotor reprimován*)
  - o (x eukaryota – eukaryotní RNA Pol nemá takovou afinitu k DNA jako bakteriální RNA Pol – u eukaryot jsou v oblasti promotoru proteiny, které umožní nasednutí polymerázy – aktivace)

### Bakteriální RNA polymeráza

- Složení katalytického jádra enzymu

Podjednotka	funkce	gen	velikost (kD)
$\alpha$	vazba $\beta$ a $\beta'$ k sobě, vazba některých aktivátorů	rpo A	36
$\beta$	katalytická podjednotka	rpo B	150
$\beta'$	vazba k templátu	rpo C	160

- Katalytické jádro bakteriální RNA polymerázy má afinitu k DNA templátu, ale nasedá náhodně. Za specifitu k promotoru je zodpovědný faktor  $\sigma$ . Dojde-li k navázání  $\sigma$  faktoru a polymerázy na DNA, dojde k otevření DNA, usazení RNA polymerázy a zahájení transkripce. Po infekci T4 fágem dochází k ADP-ribosylaci  $\alpha$  podjednotky → snížená afinita k bakteriálním promotorům.
- *Regulace genové exprese v bakteriální buňce* – represory – brání nasednutí RNA polymerázy
  - o u eukaryot je tzv. aktivační typ – promotor se obsadí celou škálou faktorů, aby mohla začít transkripcí, eukaryotní RNA polymeráza nemá afinitu k DNA
- *Sekvence skládání RNA polymerázy*:
  - o 2 dimerizační podjednotky  $\alpha$ , katalytické jednotky  $\beta$ ,  $\beta'$ ,  $\omega$  a  $\sigma$  faktor
  - o  $\alpha + \alpha \rightarrow \alpha_2 + \beta \rightarrow \alpha_2\beta + \beta' \rightarrow \alpha_2\beta\beta' =$  katalytické jádro +  $\sigma = \alpha_2\beta\beta'\sigma =$  funkční iniciační komplex (holoenzym – cca 3000 molekul na buňku) +  $\omega$
- *bakteriální RNA polymeráza = katalytické podjednotky  $\beta$ ,  $\beta'$  + 2 dimerizační podjednotky  $\alpha + \sigma$  faktor*

- **$\sigma$  faktor** – určuje specifitu dané RNA polymerázy k danému typu promotoru
  - *bakteriální sigma faktory*
    - odpovědné za rozlišení konkrétního promotoru – základní je  $\sigma^{70}$
    - bakteriální buňka má ještě jiné  $\sigma$  faktory (pokud nejsou nepříznivé podmínky)
    - u sporulujících bakterií je sporulace řízená kromě jiného kaskádou syntézy různých  $\sigma$  faktorů
    - např.  $\sigma$  faktor 54 neumožňuje zahájení transkripce RNA polymerázy, ale umožní rozpozнат konkrétní promotor, pak v závislosti na interakci s proteiny, které umožní zahájení transkripce, spustí transkripci

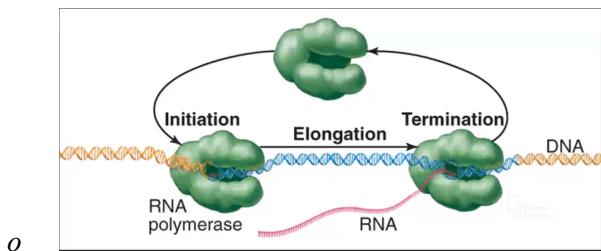
- **Příklady  $\sigma$  faktorů u *E. coli***

faktor	funkce	gen	velikost (kD)
$\sigma^{70}$	normální iniciace	<i>rpoD</i>	83
$\sigma^{32}$	teplotní šok	<i>rpoH</i>	32
$\sigma^{28(F)}$	geny pro chemotaxi a pohyb	<i>fliA</i>	28
$\sigma^{38(S)}$	stacionární fáze a stres	<i>rpoS</i>	38
$\sigma^{54}$	hladovění na dusík	<i>rpoN</i>	54

- promotory, které jsou neustále funkční (pro bílkoviny, které bakterie potřebuje po většinu svého života) – *faktor  $\sigma$ 70* (u *E. coli*, u ostatních bakterií homology tohoto faktoru; zodpovědný za iniciaci transkripce za normálních podmínek)
  - specifičejší regulace RNA polymerázy – další  $\sigma$  faktory; mohou buňku posunout do diferenciace
- specifické podmínky – hladovění na dusík, přechod do stacionární fáze, teplotní šok – pro tyto podmínky má bakterie další speciální  $\sigma$  faktory použité v těchto konkrétních podmírkách; po aktivaci takového  $\sigma$  faktoru může bakterie snadno a rychle přeprogramovat transkripcí a rychle se přizpůsobit novým podmínkám
- u některých bakterií přechod ke sporulaci – spuštění mRNA pro  $\sigma$  faktor → kaskáda spouštění, postupné přeprogramovávání
- některé bakterie mají velice málo  $\sigma$  faktorů a program transkripce spouštějí kaskádou postupného přeprogramovávání, jiné mají větší množství  $\sigma$  faktorů
- *neesenciální podjednotky* – nejsou esenciální, ale mají vliv na fitness
  - *omega podjednotka* – u všech bakterií ( $G^+$  i  $G^-$ ), má vliv na sestavování komplexu, zachycení a aktivaci sigma faktoru, stabilizace celého komplexu
  - *delta, epsilon podjednotka* – hlavně  $G^+$  bakterie
  - důležité pro stabilitu polymerázového komplexu
  - pokud tyto podjednotky nejsou přítomny, vede to ke špatnému rozpoznání  $\sigma$  faktorů a deregulaci genomu
  - změna fitness organisma po delecí některé z těchto podjednotek se dá měřit kompetitivním assay – kultury bakterií se kultivují v přesně stanovených podmírkách a sleduje se, které se namnoží více; více se namnoží wildtype

### Bakteriální přepis

- *fáze transkripce*
  - *rozpoznání templátu a iniciace*
  - *elongace*
  - *terminace*



## Iniciace transkripce

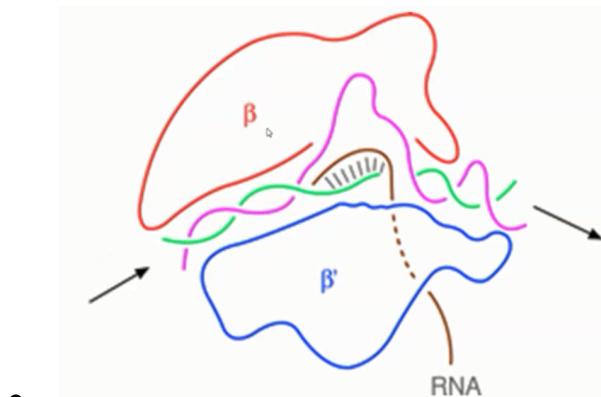
- sigma podjednotka RNA Pol rozpozná konsenzus sekvenci
- *E. coli* (a příbuzné bakterie) – u jiných vypadá podobně, ale popis je pro tento model
- **rozpoznání templátu + iniciace**
  - nalezení promotoru – vazba na DNA, syntéza prvního nukleotidu
  - Typický prokaryotický promotor obsahuje -10 konsensus (Pribnowovu) a 35 konsensus sekvence, kde -10 sekvence začíná 7 bází od +1 nukleotidu (obvykle purin) a -35 sekvence začíná 17 bází od -10 oblasti.
  - **-10 konsensus\***
    - $T_{80}A_{95}T_{45}A_{60}A_{50}T_{96}$
  - **-35 konsensus\***
    - $T_{82}T_{84}G_{78}A_{65}C_{54}A_{45}$
  - \* čísla u jednotlivých bází znamenají výskyt dané báze v procentech (u promotorů závislých na  $\sigma 70$ )
  - Pro různé příležitosti – různá síla promotorů.
  - Množství primárního transkriptu je určeno a) silou promotoru (někdy i dva v tandemě), b) genovou dozí (počet kopií genů, většina strukturních genů je unikátních)
  - RNA polymeráza v komplexu se  $\sigma$  faktorem nasedá na promotor do -35 oblasti a posunuje se k -10 oblasti. V této fázi chrání sekvenci v rozmezí 55 až +20. Vytvoření iniciační bubliny na DNA.
  - Negativně nadšroubovicová DNA = lepší templát pro zahájení transkripce, je uvolněnější
  - Kaskáda specifických  $\sigma$  faktorů využívána např. při sporulaci *B. subtilis* nebo některými bakteriofágy (např. SPO1 *B. subtilis*)
- všechny geny  $\sigma 70$  – můžeme definovat oblast, kam nasedá polymeráza (odkud transkripce začíná) – *promotor*
- bakteriální promotor – obsahuje *-10 konsensus* a *-35 konsensus* (posun proti proudu transkripce)
- *konsensus sekvence*
  - všechny promotory *E. coli* odpovídající  $\sigma 70$  – v každém promotoru najdeme -10 konsensus sekvenci, spočítáme, kolikrát se opakuje konkrétní nukleotid; nejčastější nukleotidy se vypíšeme  $\rightarrow$  sekvence určující nejčastěji se vyskytující nukleotid (+ dopíšeme počet opakování – viz výše)
  - konsensus sekvence se v organismu nemusí vůbec vyskytovat
  - konsensus sekvence nám umožňuje popsat danou oblast
  - konsensus sekvence není nejlepší! – nemusí být nejfunkčnější, nejaktivnější – není o funkčnosti nebo aktivitě sekvence, jen o tom, *co je na dané pozici nejčastější*
  - je to jen to, co se vyskytuje na jednotlivé pozici nejčastěji
- bakterie má různé  $\sigma$  faktory, různě silné promotory pro konkrétní  $\sigma$  faktory
- v promotorech a přilehlých oblastech *vazebná místa pro transkripční faktory* – aktivační x represní

- Lac operon – CAP protein může mít za určitých okolností aktivační i represní funkci
- různá síla promotorů – síla může být laděna v závislosti na různých podmínkách podle toho, jak aktivní a v jakém množství k dispozici; k tomuto ladění dochází díky spoustě koaktivátorů
  - množství transkriptu, množství překládané bílkoviny – odpovídá síle promotoru v daném organismu v daných podmínkách
- genová dóze = kolikrát se vyskytuje gen, který poskytuje konkrétní transkript
- -> množství primárního transkriptu je dáno různou silou promotorů a genovou dózí
- -> multiplikace signálu (geny kódující bílkoviny) -> mRNA může být přeložena víckrát (-> tyto geny jsou u řady organismů unikátní)
- podobné geny kódující bílkoviny v genomu:
  - a) mají jinou funkci
  - b) produkt daného genu je naprosto esenciální
  - (zvláště jednobuněčné organismy s rychlou replikační strategií snažící se udržet genom malý -> většina genů unikátních)

## Elongace

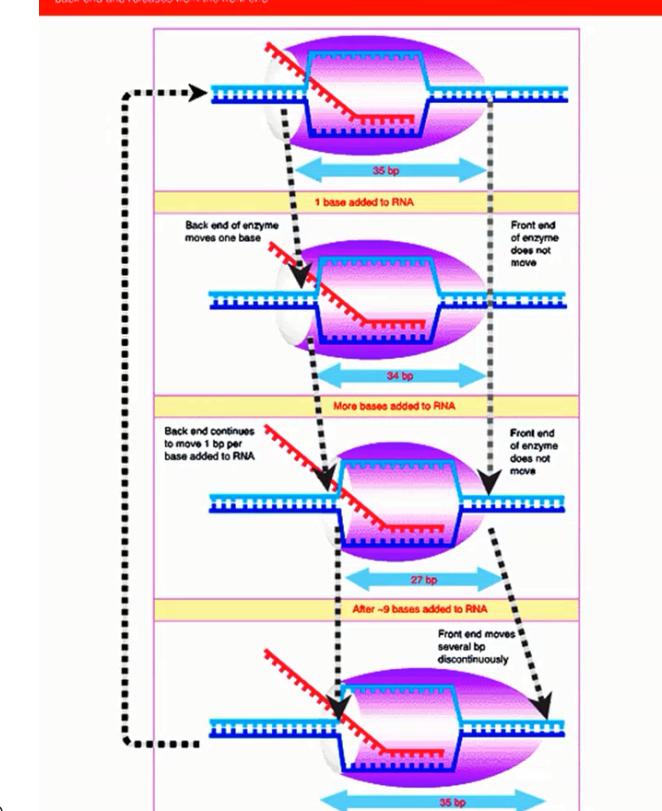
- vznik ternárního komplexu
- syntéza 40 až 60 bází / sekundu ( $48 \text{ b.s}^{-1}$  respektive až  $90 \text{ b.s}^{-1}$  pro rRNA)
- po syntéze cca 9 bází ( $5' \rightarrow 3'$ , NTPs, uvolnění pyrofosfátu) uvolnění σ faktoru (u *B. subtilis* zůstává asociován s RNA pol. po celou dobu) a případné navázání dalších faktorů účinkujících při elongaci a terminaci (NusA, NusB, NusG) a “proofreading” aktivitě? + uvolnění z místa předčasného zastavení (GreA, GreB); do uvolnění σ faktoru = tzv. abortivní iniciace (RNA pol. chrání cca 60 bp)
- Abortivní iniciace transkripce – všechny RNA polymerázy jsou schopny stálé syntézy malého množství nedokončené RNA, i když transkripce ještě neběží (jako když se zahřívá motor). U Gram<sup>-</sup> bakterií σ faktor po ukončení abortivní fáze opouští komplex, u Gram<sup>+</sup> se předpokládá, že tam zůstává.
- po syntéze cca 15 – 20 b další konformační změna RNA pol. (chrání cca 30 až 40 bp)
- RNA pol. má větší afinitu k sense řetězci DNA
- RNA polymeráza se pohybuje “píďalkovitě” v krocích 7 – 8 b nascentního transkriptu
- odvýjení a zavíjení DNA během elongace – šířka bubliny 12 – 20 bp
- rychlosť elongace – 40–60 nukleotidů / sekunda (u bakterií) – spřažená transkripce a translace -> znalostí tohoto čísla známe i rychlosť translace u bakterií (po tripletech -> 15–20 AMK/s)
- polycistronní transkript – za sebou geny A, B, C přepisované naráz, v genu B mutace na stop kodon -> A přeložen do proteinu, B zkrácený, C se nepřeloží vůbec (konec genu B může být odhalen – v oblasti mohou být terminátory; pokud se u bakteriální RNA odhalí delší úsek, který není příkryt ribozomy, mohou se objevit kryptické terminátory, které způsobí ukončení transkripce -> nejen že nebude překládán gen B, ale co je za ním, se vůbec nepřeloží (gen C) -> předčasná terminace) -> distribuce ribozomů má významný vliv z hlediska regulací
- rRNA je velmi strukturovaná, je syntetizovaná rychleji než mRNA -> některé RNA mohou být přepisovány rychleji než jiné -> 3. možnost regulace (kromě síly promotoru a kromě genové dóze 3. možnost regulace – některé RNA mohou být transkribovány rychleji)
- kromě faktorů účastnících se elongace a terminace (*antiterminační faktory NusA, NusB, NusC*) řada dalších podstatných faktorů – asociují s bakteriální RNA polymerázou
  - významné (homology u archeí a eukaryot – TF2S) – **GreA, GreB** – účastní se „proofreading“ aktivity (jiný typ než u DNA polymeráz!)
  - *GreA, GreB*

- mají slabou nukleázovou aktivitu
  - když RNA Pol zastaví, tak GreA nebo GreB odštěpí několik nukleotidů, vznikne čerstvé -OH očko a RNA Pol jede dál / celé se to disociuje
  - jeden z důvodů, proč RNA Pol zastaví, je špatné párování nukleotidů, dojde k odpojení 3' konce (nemá 3'->5' exonukleázovou aktivitu, takže tyto proteiny zastupují tuhle aktivitu)
- RNA polymeráza nemá 5'->3' exonukleázovou aktivitu; pokud se dostane do situace, kdy je RNA polymeráza zastavena, nemůže se pohybovat dál
- pokud má enzym navázat na 3'-OH skupinu další nukleotid, potřebuje očko „čerstvé“ – bud' hned syntetizováno, nebo dojde ke štěpení nukleové kyseliny (proč čerstvé? 3'-OH je reaktivní skupina -> snadno je modifikována, metylována apod. -> nemůže dále fungovat jako očko)
- proofreading aktivita RNA polymerázy: za pomoci accesory regulačních faktorů schopnost posunout se o několik nukleotidů zpět, faktory GreA nebo GreB aktivují vnitřní endonukleolytickou aktivitu -> vzniká 3'-OH očko -> RNA polymeráza může pokračovat v syntéze (nebo štěpení vede aspoň k rozpadu komplexu a uvolní místo na DNA -> může pokračovat jiný komplex)
- -> tato proofreading aktivita není o korekci chyb!
- zablokovaná RNA polymeráza na templátu se může uvolnit, nebo popojet dozadu, naštěpit mRNA a pokračovat v činnosti
- *elongační faktory*
  - **Nus faktory (A, B, G)** – u bakterií esenciální, zvyšují rychlosť polymerace u RNA polymerace, účastní se transkripce na genech pro ribozomální RNA
  - **GreA, GreB faktory** – malé proteiny, umožní RNA polymeráze při zastavení nebo zpomalení obnovit syntézu (když RNA polymeráza přeskočí – RNA polymerázy mají vyšší chybavost, vyšší než  $10^{-6}$ , než DNA polymerázy)
  - **Když RNA polymeráza udělá chybu, jsou 3 možnosti:**
    - a) RNA polymeráza tam chybu nechá a pokračuje
    - b) RNA polymeráza odpadne → vzniká nefunkční transkript
    - c) oprava – RNA polymeráza se zastaví a pak posune zpět, dojde o odštěpení 3' řetězce proteiny, které způsobily zastavení – vytvoří se nové 3'OH očko a zahájí se trankripce
- *abortivní iniciace* = iniciace, kdy vznikají spousty krátkých transkriptů (RNA polymeráza se uvolňuje) – σ faktor by měl být uvolněn (dříve se říkalo, že u  $G^+$  bakterií zůstává, teď se ukazuje, že zůstává u  $G^-$  bakterií -> nemusí být uvolněn)
- *topologie výhodná pro transkripci* – DNA v negativních nadobrátkách jsou přístupnější pro replikaci i pro transkripcí, DNA může být snadněji rozvolněná
- transkripce u bakteriální rRNA musí být rychlá – pokud se u bakterií přepisuje gen pro rRNA, tak RNA Pol spolu s antiterminačními faktory (NusA, NusB, NusG) zvýší rychlosť, geny pro rRNA jsou přepisovány rychleji (až 90 bází/sec.) – to se děje kvůli tomu, že bakteriální buňka transkribuje a translatuje najednou a potřebuje hodně ribozomů



- pohyb RNA polymerázy – „píďakovitý“

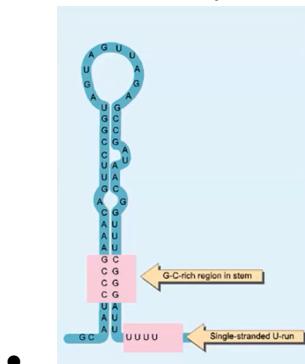
Figure 11.7 RNA polymerase moves like an inchworm during elongation, when it compresses from the back end and releases from the front end.



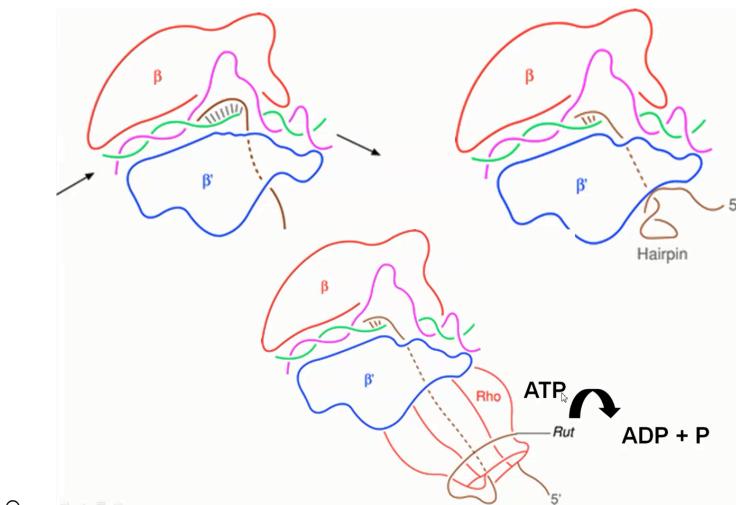
## Terminace

- ukončení transkripce a disociace ternárního komplexu
- terminace vždy závislá na struktuře a sekvenci nascentního transkriptu (záměna GTP za ITP)
- dva typy terminátorů, fungují na úrovni RNA (syntézu zastaví nově nasynthetizovaná RNA)
- a) **závislé na ρ** (50 – 90 b oblast primárního transkriptu bohatá na C, „upstream“ od konce)
  - faktor ρ funguje jako hexamer, má ATPázovou (helikázovou) aktivitu, interakce s NusA, NusG a ρ podjednotkou RNA pol.
  - pokud ρ faktor narazí na sekvenci, která je ρ faktorem rozpoznávána, tak je schopen se navázat na RNA polymerázu a pomocí helikázové aktivity „vytáhne“ RNA z RNA polymerázy
  - oblasti, které fungují jako ρ závislý terminátor jsou bohaté na cytosin
  - RNA polymeráza se dále zastaví v oblastech DNA, které jsou strukturně složité

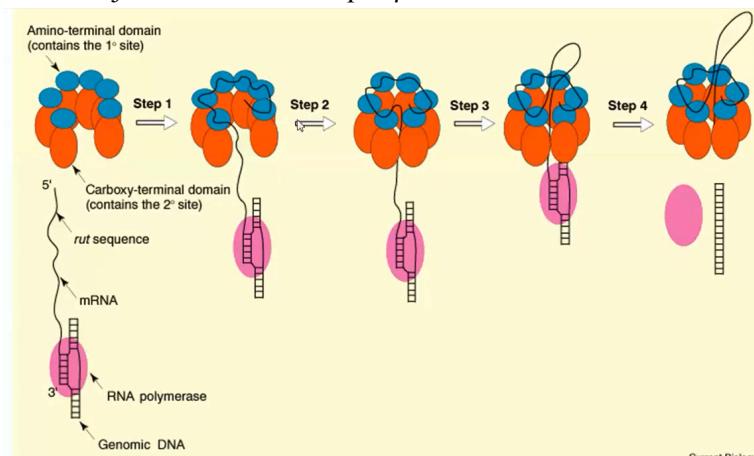
- pokud je RNA rychle syntetizována a obsazena ribozomy, tak se  $\rho$  faktor nemá kam navázat. Jakmile se v blízkosti RNA polymerázy objeví větší úsek, který je neobsazený ribozómy, tak se  $\rho$  faktor může navázat
- pokud se tam objeví potenciální terminátor závislý na  $\rho$  faktoru, může být využit a RNA polymeráza zastaví činnost
- nonsense mutace mají polární efekt na transkripci distálních genů =  $\rho$  zprostředkovaná terminace
- polární efekt mutací = mutace, která by se měla projevit na úrovni translace, může ovlivnit transkripci genů, které jsou uvnitř jedné transkripční jednotky
- antagonist  $\rho$  – antiterminační protein (bakteriofágy atd.)
- b) **nezávislé na  $\rho$**  = obligátní (vlásenka bohatá na GC, 4-8 U na konci)
  - bohaté na GC páry a jsou schopny tvořit stabilní vlásenku
  - ve směru 3' obsahují úsek, kdy do nascentní RNA je zabudováno několik zbytků kyseliny uridylové (uridin fosfát)
- sekvence pro terminaci se nachází v RNA
- závislá na terminátorech
  - a) mohou fungovat jako struktura, která vznikne na nascentní RNA a zabrání další transkripci – *nezávislé na faktoru  $\rho$*  (mají vnitřní strukturu)
  - b) terminátory rozpoznávané dalším bílkovinným faktorem, který umožní zastavit RNA polymerázu a na daném místě transkripcí ukončit – *faktor  $\rho \rightarrow$  terminátory závislé na faktoru  $\rho$*
- nesmyslné mutace (= mutace na stop kodon) mohou mít polární efekt (odhalí se místo, které umožní  $\rho$ -závislou terminaci)



- ○  $\rho$ -nezávislý terminátor – GC bohatá vlásenka (GC bohaté oblasti stabilizují vlásenku)
  - oblasti genomu, které jsou GC bohaté, vytvářejí vlásenky (vlásenka = dsRNA úsek)
  - dojde k vytvoření vlásenky uvnitř aktivního centra RNA Pol, za touhle GC bohatou vlásenkou je hodně UUUU, AU oblasti jsou méně stabilní, je tady celkem velká šance, že se z RNA Pol uvolní
- 3' konec – zbytek kys. uridylové (U bohatá oblast – málo stabilní) – proti všem U adenosin  $\rightarrow$  nestabilní nukleotidové páry  $\rightarrow$  pokud dojde k tvorbě složité struktury v aktivním místě, může dojít k destabilizaci a uvolnění celého komplexu
- v sekvenci se může vyskytnout oblast, na kterou se preferenčně váže  $\rho$  faktor – složité tím, že takových oblastí může být v genomu celá řada; bohaté na C; funguje na úrovni nascentní RNA



- rozpoznání  $\rho$  faktorem – hexamer;  $\rho$  faktor za spotřeby ATP „vyvlekne“ RNA z aktivního centra
- $\rho$  terminátor putuje po DNA, musí „dohonit“ RNA polymerázu na místě, které zároveň rozpoznává (je schopen se na něj navázat, zastavit RNA polymerázu a zastavit transkripci)
- sekvence uvnitř RNA, bohaté na cytidin =  $\rho$  utilization sequence, na ni se váže faktor  $\rho$  (homohexamer, ATPáza, helikáza), ta rozpoznává, jako ATPáza umožní vytáhnout nascentní RNA z RNA Pol za spotřeby ATP
- ve chvíli, kde je RNA plně obsazena ribozomy, tak si  $\rho$  moc neškrtné
- ve chvíli, kdy se naskytne na nascentním transkriptu oblast, která nekóduje bílkoviny, tak je dost dobré možné, že se jedná o  $\rho$  závislý terminátor
- polární efekt na transkripci distálních genů: mutace na antisense vlákně může ovlivnit vedlejší geny – např. posunutí čtecího rámce, který způsobí, že se najednou odkryje  $\rho$  závislý terminátor a ovlivní takto i okolní geny
- RNA molekula obsazena ribozomu – dokud tomu tak je (geny, které něco kódují),  $\rho$  faktor „nemá přílišnou šanci“
- přepsání posledního genu v polycistronní jednotce – downstream nejspíše zafunguje nějaká sekvence jako vazebné místo pro  $\rho$  faktor



**Figure 1. Rho factor: transcription termination in four steps.**  
In step 1, the amino-terminal domain of Rho factor (blue), which contains the 1' mRNA binding site, binds to the *rut/rut'* sequence of the mRNA transcript. Rho factor is shown as an open ring, but it may be in equilibrium between open and closed states at this step. In step 2, the carboxy-terminal domain of Rho (orange), which contains the 2' mRNA binding site, binds the mRNA downstream from *rut*, and the Rho hexameric ring closes. In step 3, the carboxy-terminal domain of Rho cyclically binds and hydrolyzes ATP to propel itself along mRNA in the 5' to 3' direction. The amino-terminal domain may disengage from *rut* during this step. In step 4, Rho factor acts as a helicase to couple ATP binding and hydrolysis to nucleic acid unwinding. This results in disengagement of the mRNA from the genomic DNA and RNA polymerase (pink).

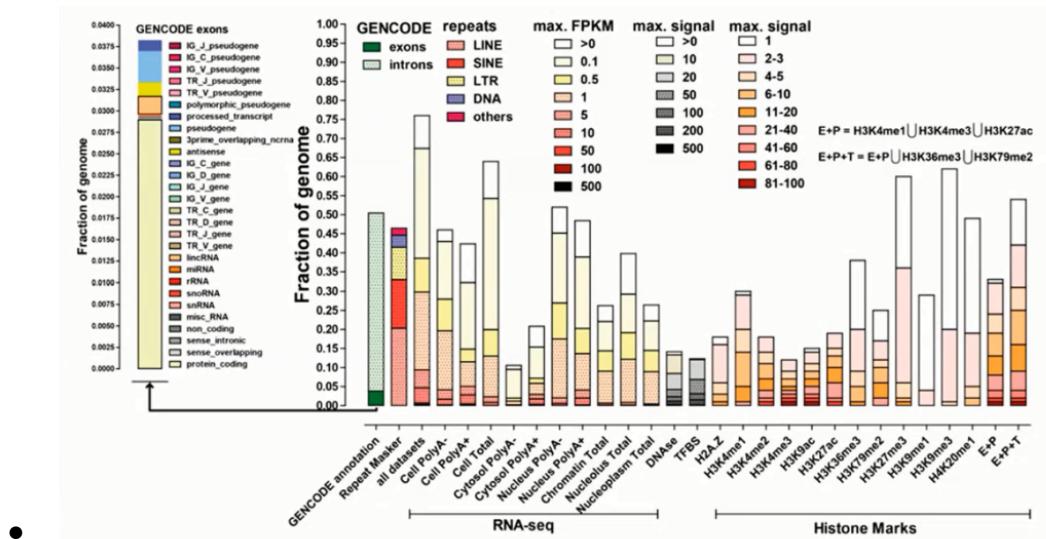


## Obrat RNA v buňce („turnover“)

- rRNA ~ 70 – 80 % celkového obsahu RNA
- tRNA ~ 15 %
- mRNA ~ 3 – 4 %
- rRNA a tRNA – stabilizovány modifikacemi, terciární strukturou, asociací s proteiny
- mRNA – nestabilní; průměrný T<sub>1/2</sub> = 40 sekund
- degradace mRNA většinou nastartována endonukleotickým štěpením od 5' konce a následným 3' → 5' štěpením vzniklých fragmentů
- účast enzymů aktivních v posttranskripčních úpravách
- iniciace transkripce je děj, který maximálně reguluje většinu genové exprese u bakterií (stoupá stabilita transkribovaných RNA a jejich délka – posun v regulacích dál)
- *degradace mRNA u bakterií*
  - většinou nastartována *endonukleolytickým štěpením* od 5' konce -> „nasekání“ na více malých kousků; 3' konce označeny oligo-A řetízkem (-> značka pro exonukleázu) -> štěpení exonukleázami (3'->5')

## Eukaryotní typ přepisu

- není spřažena transkripce s translací – kompartmentace
- export RNA z jádra – mRNA překládána
- více RNA polymeráz (v základu RNA polymeráza I, RNA Pol II, RNA Pol III, u rostlin navíc RNA Pol IV a RNA Pol V)
- převážně monocistronní transkripty
- eukaryotní geny často přerušeny introny
- mRNA obvykle posttranskripčně modifikována
- navíc transkripce v organelách
- poločas rozpadu RNA delší
- častější regulace i na úrovni translace
- kombinatorický a kooperativní způsob regulace transkripce
- *typy buněčných RNA*
  - mRNA – messenger RNA
  - rRNA – ribozomální RNA = 5,8S, 18S, 28S, 5S, rRNA (savci)
  - hnRNA – heterogenní jaderná RNA = prekurzorová mRNA
  - tRNA – transferová RNA
  - snRNA, snoRNA, scRNA, miRNA – malé buněčné RNA
  - lncRNA – dlouhé nekódující RNA
  - (ENCODE – Encyclopedia of DNA elements – regulační oblasti, chromatinové modifikace, DNA elementy, modifikace DNA, souvislosti s transkripcí, vazba bílkovin na různé oblasti genomu a souvislost s transkripcí atd.)

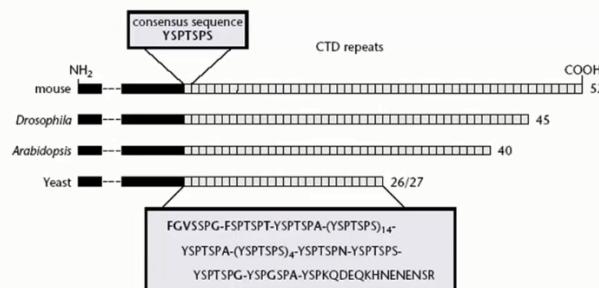


- pohled na transkripci genomu – ještě nedávno nebylo možné
- podjednotky ribozomu
  - prokaryotní ribozom:
    - 70S: 30S a 50S
    - 30S podjednotka: 16S RNA + 21 proteinů
    - 50S podjednotka: 5S + 23S + 31 proteinů
  - eukaryotní ribozom:
    - 80S: 40S a 60S
    - 40S podjednotka: 18S RNA + 33 proteinů
    - 60S podjednotka: 5S RNA + 28S RNA + 5,8S RNA + 46 proteinů
- buněčné RNA polymerázy

	syntetizuje	citlivost k (-amanitinu)	potřeba ATP pro iniciaci
RNA polymeráza I	ribozomální RNA v jadérku	ne	ne
RNA polymeráza II	hnRNA (strukturní geny) a řada malých RNA	ano	ano
RNA polymeráza III	tRNA, 5S rRNA a řada malých RNA	druhově specifické	ne

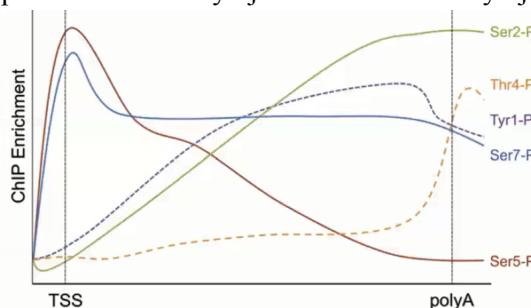
Podobné s bakteriálními polymerázami = struktura RNA polymeráz je evolučně velmi stará na rozdíl od DNA polymerázy

- Typicky velké proteiny o 8 – 14 podjednotkách (*S. cerevisiae* RNA pol. II 10 – 11); největší podjednotka **RNA pol. II nese CTD** ("carboxy-terminal domain")
- mnohonásobná repetice motivu 7 AMK – YSPTSPS (cca 26x u kvasinek, 34x u *Caenorhabditis*, 43x u octomilky a 55x u savců)
- 2 proliny udělají zlom, fosforylovatelné aminokyseliny dávají výrazný záporný náboj
- unikátní pro eukaryontní RNA polymerázu II, katalytické jádro je homologní k  $\alpha_2\beta\beta'$  *E. coli*
- esenciální, její nepřítomnost u všech eukaryot je neslučitelná s životem buňky
- kvasince, když odstraníte 2/3 CTD, tak kvasinka zahyne
- v průběhu evoluce se počet těchto jednotlivých repetic velmi zvětšuje
  - vícepodjednotkové enzymy
  - liší se tím, že mají různé podjednotky, citlivostí k inhibitorům (tradičně se popisuje citlivost k  $\alpha$ -amanitinu) a tím, co syntetizují (viz výše)
  - většina RNA je syntetizována RNA Pol II
  - RNA syntetizované RNA Pol II se ve většině případů vyznačuje strukturou na 5' konci – čepičkou



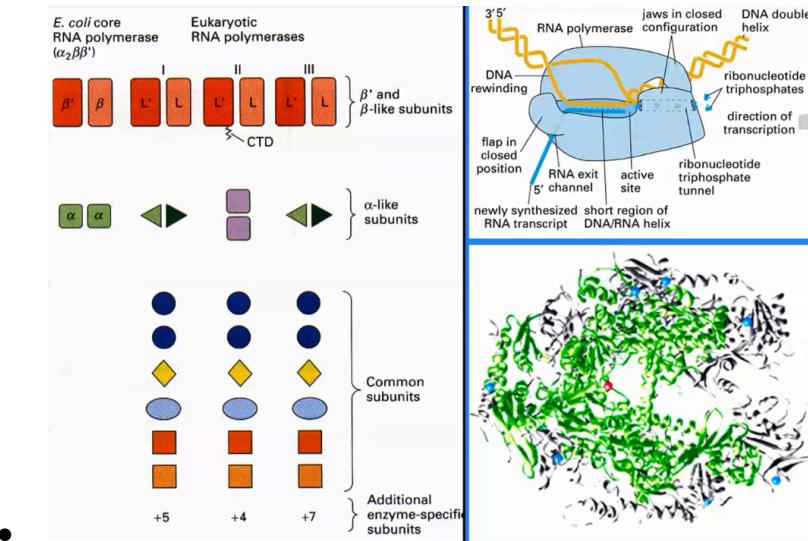
**Figure 5** C-terminal domain (CTD) length and sequence in RNA polymerase II. Each open box represents one CTD heptapeptide, the black rectangles (not drawn to scale) represent RNA polymerase II large subunit amino acids in front of the CTD. The sequence of the *Saccharomyces cerevisiae* CTD (bold-faced residues display deviation from the canonical consensus sequence) from a commonly used laboratory strain is shown. The yeast CTD shows strain variability, comprising either 26 or 27 repeats.

- *C terminální doména RNA polymerázy* sestává z repetic heptapeptidů (7 AK), repetitive je spousta
- konsensus repetitive – **tyrosin** – serin – prolin – threonin – serin – prolin – serin
- x ne v každé C terminální doméně konsensus repetitive vypadá takto!
- AK konsensus sekvence mají –OH skupinu, mohou být fosforylovány
- opakující se prolin – nestrukturovaný dlouhý úsek přístupný k různým modifikacím, zároveň může být modifikován násobně -> doména má regulační funkci (důležitou pro funkci RNA polymerázy -> pro transkripci a život buňky)
  - ověření: připravíme eukaryotní linii, původně budeme krátit část, která kóduje C-terminální doménu RNA polymerázy; jakmile zkrátíme repetitive pod 10 opakování, neslučitelné se životem
  - buněčné linie savčích buněk – hranice pro život 23 repetitive
  - expanze a regulace C-terminální domény hráje obrovskou roli v regulaci genové exprese
- C-terminální doména v průběhu evoluce expandovala, jednotlivé repetitive se měnily
- *tam, kde probíhá masivní transkripce, je C-terminální doména RNA Pol II masivně fosforylována*
- *konsensus sekvence RNA Pol II* – u člověka se vyskytuje jen pákrát (4–5, 7–21, 28–3, ostatní pozměněné)
- v blízkosti transkripčního počátku je obvykle málo fosforylován serin 2 na C-terminální doméně RNA Pol II, ale více Ser-5, v průběhu transkripce se naopak spíše serin 2 fosforyluje a serin 5 defosforyluje (viz graf níže)

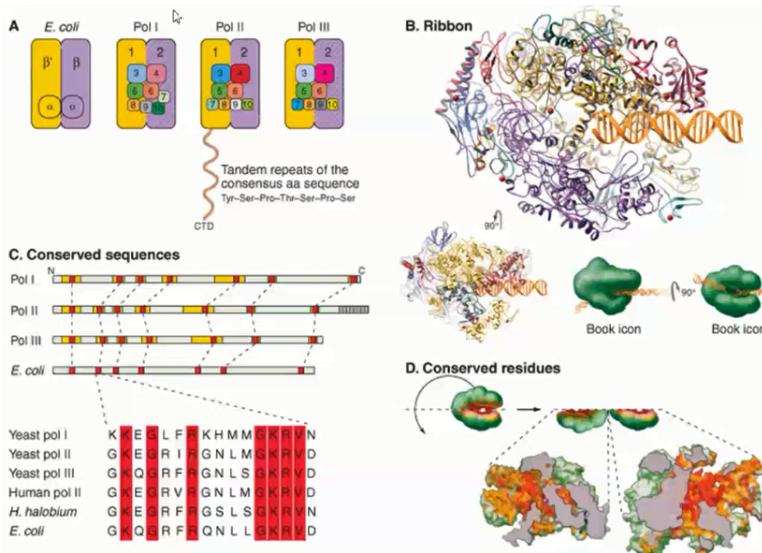


**Fig. 2.** Average profile of CTD phosphorylation marks in genes revealed by chromatin immunoprecipitation (ChIP) experiments. RNAPII attached to the cellular DNA is purified and correlated to define gene sections using monoclonal antibodies which target specific CTD modifications. Schematic representation of genome-wide occupancy profiles for all CTD phosphorylation marks. Tyr1-P, Ser2-P, Thr4-P, Ser5-P and Ser7-P show similar distribution of the modified RNAPII along certain gene regions. While Ser5-P and Ser7-P peak at the transcription start site (TSS) of genes, Tyr1-P, Ser2-P and Thr4-P signals increase toward the 3' end and polyadenylation site.

M. Heideemann et al. / *Biochimica et Biophysica Acta* 182



- porovnání RNA polymeráz



© Elsevier Pollard et al: Cell Biology 2e - www.studentconsult.com

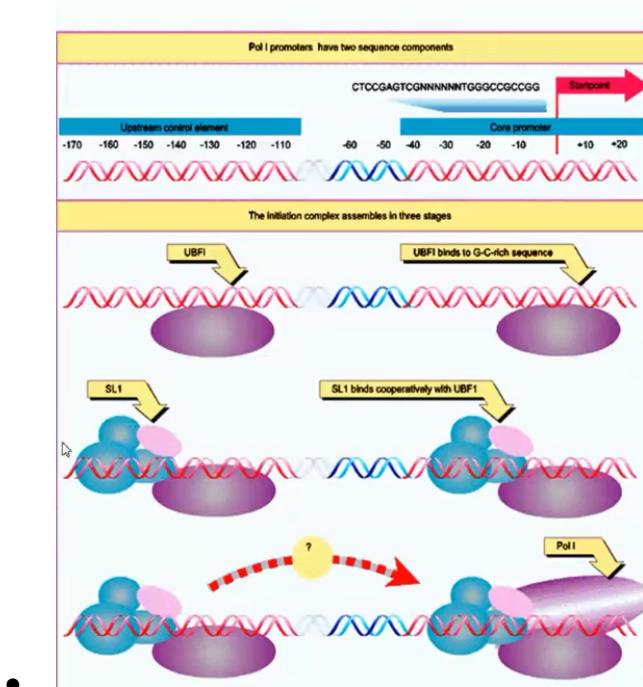
## Transkripční faktory

- hrají podstatnou roli v iniciaci transkripcie
- *obecné transkripční faktory* = transkripční faktory rozpoznávající promotor (označený modifikacemi chromatinu nebo transkripčními aktivátory/resresory), umožňují sestavení pre-iniciačního a iniciačního komplexu RNA polymerázy
  - provozní = housekeeping geny
  - spolu s RNA polymerázou tvoří základní transkripční aparát
- *speciální transkripční faktory*
  - *konstitutivní* – jen v určité tkáni (u mnohobuněčných) – každá tkáň má jiný bílkovinný obsah -> sada transkripčních faktorů způsobujících transkripci konkrétních bílkovin typických pro danou tkáň
  - *indukovatelné* – objevují se jen jako reakce na nějaký vnější nebo vnitřní podnět
- iniciace transkripcie u eukaryot vyžaduje řadu TF fungujících v definovaném sledu a s definovanými a přesně určenými vazebnými specifitami

## Trankripce RNA Pol I

- rRNA přepisována jako velká prekursorová RNA (45S u savců) a poté rozstříhána na jednotlivé 18S, 5,8S a 28S rRNA

- Geny pro 45S pre-rRNA a jednotlivé rRNA uvnitř této jednotky jsou odděleny mezerníky.
- Promotor má dvě hlavní konsensus oblasti
  - **-45 až +20 = základní oblast**; dostačující pro zahájení transkripce
  - **-180 až -107 = regulační oblast (UCE - "upstream control element")**
  - Obě oblasti z cca 85% identické a velice GC bohaté. Sekvence obou oblastí rozeznává UBF1 faktor a váže se na ně. V přítomnosti UBF1 se váže SL1 (4 polypeptidy = 3 + TBP)



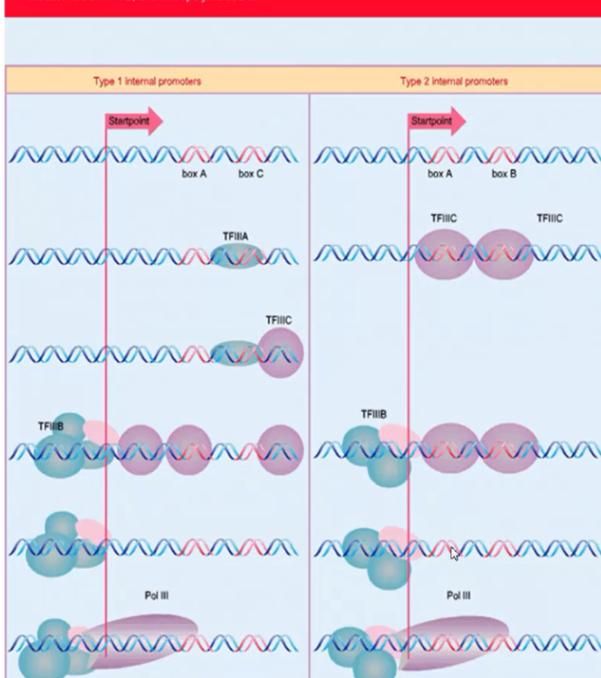
- přepisuje rDNA (= DNA kódující velkou prekurzorovou RNA pro ribozomy)
- v buňce musí být masivní produkce rRNA, pool ribozomů musí být neustále doplňován
  - → silné promotory
  - → velká genová dóza
- lze lokalizovat *jadérko* = oblast, kde jsou zmnoženy geny pro rDNA, masivní přepis
- jak je zajištěna síla promotoru?
  - faktor *UBF1* – rozpoznává promotor pro RNA Pol I – váže se na něj komplex faktorů SL1
    - upstream (proti proudu transkripce) od tohoto promotoru jsou oblasti, kde se může také navázat UBF1 a SL1 – slouží k navýšení efektivní koncentrace transkripčních faktorů
    - čím více vazebných míst, tím větší šance, že dojde k transkripci
  - aby byl promotor silnější, tak ve směru upstream je ještě jedno promotorové místo – vniká vyšší koncentrace faktorů a vyšší koncentrace RNA Pol
- UBF1 je dále rozpoznáván SL1 (TBP – TATA binding protein) – součást všech 3 rozpoznávacích komplexů polymeráz

## Transkripce RNA Pol III

- Dva základní druhy promotorů:
  - pro 5S RNA a tRNA – interní (1 a 2 typ)
  - pro snRNA
- Hlavní faktory účastnící se iniciace transkripce RNA pol. III na interních typech promotorů
  - **TFIIIA**, rozeznává Box C
  - **TFIIB** (TBP + 2) jednou navázaný nepotřebuje navázané faktory TFIIIA a TFIIIC pro iniciaci transkripce

- **TFIIC** ( $> 500$  kD a  $> 5$  polypeptidů), rozeznává Box B
- vždy TFIIC umožňuje nasednutí TFIIIB
- promotor pro tRNA vyžaduje TFIIC a TFIIIB, pro 5S rRNA navíc ještě TFIIA
- Promotory pro přepis snRNA jsou strukturou podobné promotorům rozeznávaným RNA pol. II.
- --- OCT --- PSE --- TATA --- + 1 ---

**Figure 28.7** Initiation via the internal pol III promoters involves the assembly factors TFIIA and TFIIC, the initiation factor TFIIIB, and RNA polymerase III.



- faktor TFIIIB zůstává déle navázán na promotor -> více cyklů RNA Pol III ( $\rightarrow$  šetří se čas na skládání celého komplexu)
- rozpad nebo regulace -> komplex se poskládá znovu (promotory pro 5S ribozomální RNA nebo tRNA – promotory rozpoznány faktory – TFIIC pro tRNA – kooperativně se váže do oblasti startu transkripce +1 -> sestavení TFIIIB -> start transkripce RNA Pol III)
- transkripční faktory jsou číslovány – TF-číslo-písmeno
- *interní promotor* – regulační oblast je uvnitř přepisované oblasti -> takový gen je kompaktnější (např. tam, kde v rRNA dochází k repeticím – repetice jsou více za sebou)
  - *box B & box A* – název promotoru typického pro tRNA
  - *box C & A* – typický pro 5S
- *jak se na to přišlo?*
  - jak člověk odhalí promotor? – vezmeme plazmid (vektor, se kterým budeme pracovat) -> vložení reportérového genu (umožňuje sledovat funkci – např. GFP fluorescenční protein – pokud se v buňce produkuje, došlo k transkripci; pokud před něj nedáme promotor, k transkripci nedochází); neznámá oblast genomu se vloží před reportérový gen -> vložení do buňky – dochází k produkci reportérového genu? produkce -> vložili jsme promotor
  - kde přesně se promotor vyskytuje a jeho bližší specifikace? – vložený úsek zkracujeme, snažíme se vymezit nejkratší část, která ještě funguje jako promotor; pak můžeme mutovat a sledovat, co mutace dělají, můžeme sledovat konsensus sekvence

## Transkripce RNA Pol II

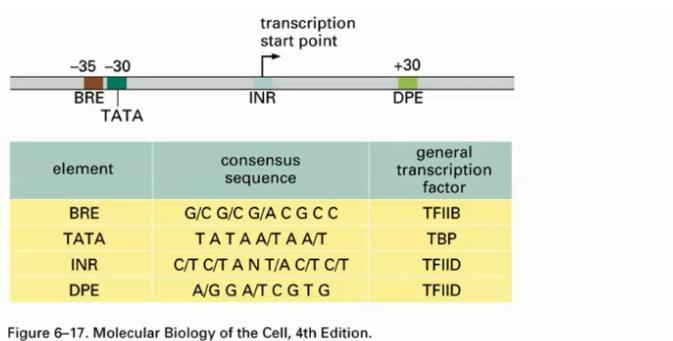
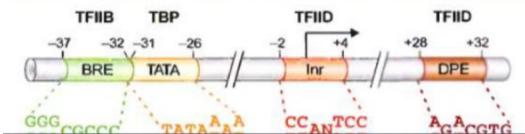
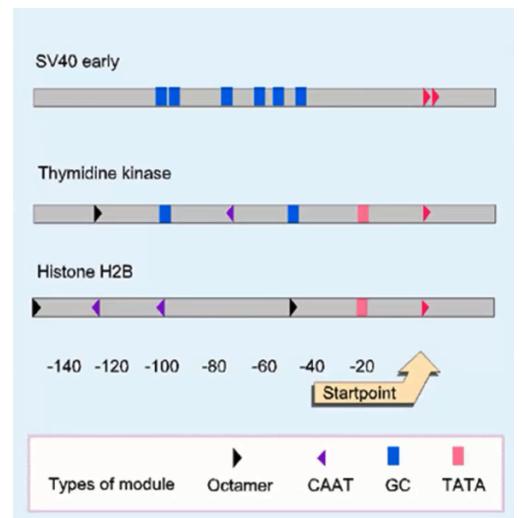


Figure 6-17. Molecular Biology of the Cell, 4th Edition.



- konsensus v oblasti startu = Inr – “initiator” – (Py2CAPy5); +1 obvykle A; geny bez Inr – často alternativní 5' konec mRNA (101 až 102 bp)
- cca -25 oblast = **TATA box**; často ohraničený GC bohatou oblastí; oblast vazby **TFIID** (obsahuje cca 30 kD veliký **TBP** = “TATA binding protein” a řadu **TAF** = “**TBP associated factors**”); **TFIID** = cca 800 kD (TBP + cca 9 TAF); mutace nezabrání iniciaci, ale mají vliv na pozici startu
- cca -75 oblast = **CAAT box**; funkční v obou orientacích, mutace mají vliv na sílu promotoru
- cca -90 oblast = **GC box**; funkční v obou orientacích, často ve více kopiích, vliv na sílu promotoru, vazba faktoru SP1; GpC = “marker” promotorové oblasti savců
- cca -120 oblast = **promotorové proximální oblasti** (oktamery, kB, ATF...)
- promotorové distální oblasti – **enhancery**
- kombinování různých kontrolních elementů
- mnoho genů pro rRNA, mnoho různých promotorů pro tRNA
- geny kódující mRNA – 20 000 promotorů
- mnoho promotorů -> regulace, vazba obecných a speciálních transkripčních faktorů – mnohem složitější
- promotory mnohem širší a náročnější (vyšší počet sekvencí)
- „typický eukaryotní promotor“ – nejde přesně určit, v každé učebnici něco jiného
- kdybychom analyzovali promotory, uvidíme, že řada promotorů některé z typických oblastí nemá / má zmnoženy / otočeny
- -> eukaryotní promotor pro RNA Pol II je velmi komplexní a nelze ho definovat konkrétními sekvencemi

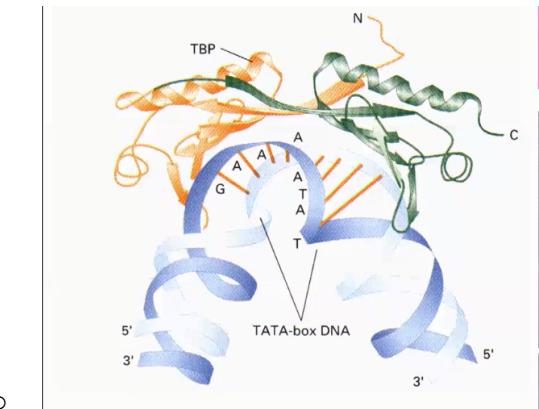


- TATA vazebný protein – popsán vazbou na promotory RNA Pol II, na TATA box (= místo typické sekvenčí TATA), kam se protein váže
  - u člověka něco podobného TATA boxu u pouhé čtvrtiny genů kódujících bílkoviny
  - konsensuální TATA box jen u 10 % genů transkribovaných RNA Pol II
- *oblast INR* = oblast iniciačního regionu – zásadně významná pro zahájení transkripce a konsensus v ní; jen u necelé poloviny genů přepisovaných RNA Pol II (u člověka)
- promotory pro RNA Pol II jsou „skládačka“ celé řady modulů

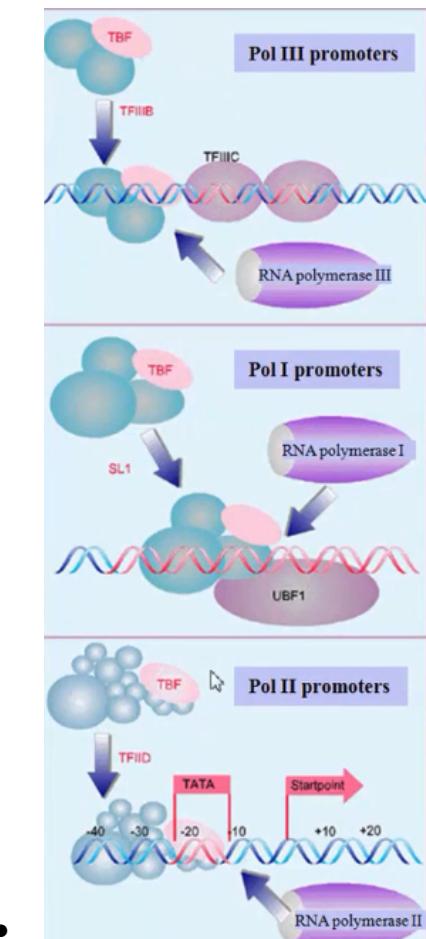
SUMMARY OF EUKARYOTIC RNA POLYMERASE II GENERAL TRANSCRIPTION FACTORS			
Factor	Number of Subunits	Subunit M (KD)	Functions
TFIIA	3	12, 19, 35	Stabilizes binding of TBP and TFIIB
TFIIB	1	25	Binds TBP, selects start site, and recruits polymerase II
TFIID	12	15–250	Interacts with regulatory factors
(TBP)	1	38	Subunit of TFIID; specifically recognizes the TATA box
TFIIE	2	34, 57	Recruits TFIIF
TFIIF	2	30, 74	Binds polymerase II and TFIIB
TFIIH	9	35–98	Unwinds promoter DNA; phosphorylates CTD (C-terminal domain of RNA polymerase II)
Polymerase II	12	10–220	Catalyzes RNA synthesis
<b>TOTALS</b>	<b>42</b>	<b>~1000</b>	

- TFIID – obsahuje TATA vazebný protein; interaguje s TFIIB (faktor, který můžeme najít i u archeí – jeden z historicky nejstarších faktorů)
  - TFIIF – váže RNA Pol (katalytický komplex) a přináší ho do oblasti promotoru
  - regulační faktory – TFIIH (oprava mutací NER; účastní se fosforylace C-terminální domény RNA Pol II – na této fosforylacii se podílí konkrétně její součást Cdk 7), TFIIE (zprostředkovává nasednutí a aktivitu TFIIH)
  - dohromady obrovský komplex – 1 MDalton
- interakce s dalšími transkripčními faktory a enhancery – *mediátorový komplex* (také 1 MDalton)
- *iniciace transkripce RNA Pol II – posloupnost kroků*

- 1) *vazba TFIID* (sestává z TBP + 11 TAFs; cca 750 kDa) na TATA box, chrání cca -45 až 10
  - 2) *vazba TFIIA*, aktivace TBP
  - 3) *vazba TFIIB*, váže se na komplex více po proudu transkripce, komplex chrání cca -10 až +10, rozeznáván RNA pol. II
  - 4) *vazba RNA pol. II* zprostředkovaná faktorem TFIIF (dvě podjednotky – jedna s ATázovou / helikázovou aktivitou, druhá – homologní k bakteriálnímu  $\sigma$ -faktoru s afinitou ke katalytickému jádru RNA pol. II), komplex chrání cca -45 až +20
  - 5) *vytvoření první fosfodiesterové vazby* = vznik *ternárního komplexu*
  - 6) *vazba TFIIIE* – posun komplexu až k +30, předpoklad pro vazbu TFIIH
  - 7) *vazba TFIIH* (ATPasa, helikasa, kinasa – fosforylace CTD konce RNA pol. II, sestává z více podjednotek a vyskytuje se v různých formách) a TFIIJ
  - 8) *fosforylace CTD RNA pol. II* – zahájení syntézy RNA (fosforylace je signálem pro celou řadu dalších proteinových a riboproteinových komplexů, které se pak účastní dalších dějů, je to taková značka v místech kde probíhá transkripce)
- *TFIIH*
  - může hrát roli i při elongaci
  - hydrolýza ATP a funkce TFIIH a TFIIE jsou nezbytné pro pohyb RNA pol. II na lineárním templátu. Na DNA v nadšroubovicové formě tomu tak není.
  - vliv na opravy DNA asociované s transkripcí (mutace v TFIIH = zvýšená citlivost k UV a častější výskyt rakoviny)
  - na promotory bez TATA boxu (některé provozní geny) se váží některé TF do oblasti Inr = podobný model iniciace jako u promotorů RNA pol I a interních promotorů RNA pol. III.
- -> iniciační komplex transkripce je obrovský – více než 2 MDaltony (transkripční faktory, mediátorový komplex)
- *TBP protein* – součást transkripčních faktorů u všech třech eukaryotních RNA polymeráz
  - váže se do malého žlábku
  - = TATA binding protein
  - způsobuje ohnutí DNA -> (zvláště u DNA v negativních nadobrátkách) usnadnění tání dvoušroubovice na specifickém místě -> vazba RNA polymerázy -> určení +1 místa, odkud bude zahájena transkripce
  - vazba do menšího žlábku DNA, způsobuje ohyb a rozšíření menšího žlábku, funguje jako umísťovací faktor funkční při iniciaci přepisu u většiny promotorů všech eukaryotických RNA polymeráz, chrání cca 1 obrátku DNA (typicky -37 až -25), celý TFIID pak -45 až -10
  - má 2 domény, jedna rozpoznává TATA box a druhá udělá ohnutí DNA
  - význam: nalezení přesné pozice RNA polymerázy, vytvoření pnutí a vytvoření bubliny
  - pokud je v komplexu s proteiny, které ho přímo navedou na pozici, TATA boxu není zapotřebí
  - za určitých okolností může být součástí obecných transkripčních faktorů všech třech základních polymeráz



- *promotory pro RNA Pol II*
  - promotory jsou hodně specifický pro konkrétní gen
    - cca 40 nt dlouhý
    - BRE element: vazba TFIIB
    - TATA box
    - inr(initiator) CATTC sekvence
    - DPE (downstream promoter element)
  - faktor TFIID (největší z faktorů) obsahuje TBP protein, který rozpoznává oblast aktivátorů transkripce
  - váže se do malého žlábků DNA a způsobuje ohyb – ohyb znamená, tání dvoujšroubovice, zároveň je to značka, podle které může být pozičně určen zahajovací komplex
  - Tento komplex (DNA + TFIID + TBP) je stabilizován TFIIA a následně TFIIB, poté se naváže RNA Pol II, která tam přichází spolu s faktorem TFIIF, který umožňuje navázání TFIIH, TFIIH umožňuje přeměnu pre-iniciačního komplexu na otevřený komplex – funguje jako helikáza a spotřebovává při tom ATP. Faktor TFIIH umožňuje nasednutí TFIIE a TFIJ – tyhle dva faktory mají kinázovou aktivitu, jsou přítomny s RNA polymerázou i v průběhu elongace a modifikují C terminální doménu (fosforylují atd. – tam, kde je fosforylovaná C terminální doména, tam probíhá masivní transkripcie)
  - Modifikace C term. domény má vliv na RNA Pol (kdy bude ukončenej transkript atd...).

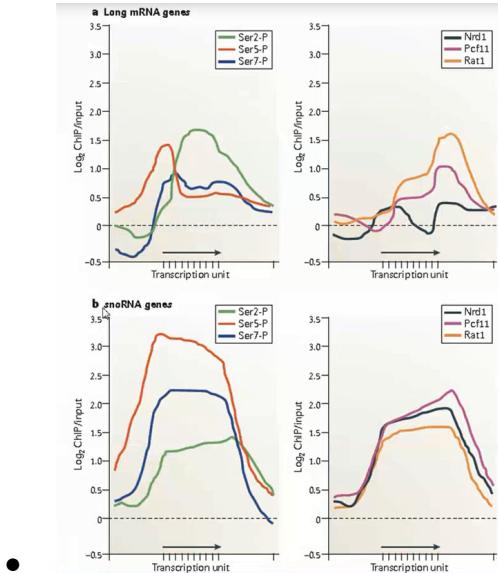


- - demonstrace toho, že TBF protein je součástí transkripčních faktorů všech tří RNA Pol

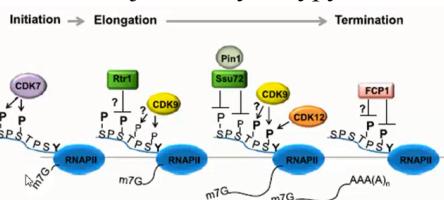
## Regulace

- regulace na úrovni sestavení pre-iniciačního a iniciačního komplexu RNA Pol II
- Metazoa – RNA Pol II rozběhne syntézu a zůstává stát po určitou dobu asociovaná s templátem ve vzdálenosti cca 50 nukleotidů (a více – peak pravděpodobnosti jejího výskytu je kolem 50.–60. nukleotidu) – zde polymeráza zůstává – *pauza RNA polymerázy* -> další regulace: výměna faktorů, navázání dalších faktorů (zodpovědné za část další regulace v průběhu elongace – dostávají se sem elongační faktory, jejich akce na RNA polymeráze je výrazně regulována)
  - RNA Pol II se po nasynthetizování cca 100 nukleotidů zastaví – můžou za to negativní regulátory – DSIF a NELF. RNA Pol II čeká. Uvolnění spočívá ve fosforylací negativních regulátorů faktorem PFB (osahuje kinázu Cdk 9). Faktor NELF se po fosforylaci uvolní, faktor DSIF je fosforylován a zůstává přítomen. RNA Pol II se rozjede -> další úroveň regulace transkripcie
- RNA Pol zde může čekat maximálně sekundy (na transkripci dlouhý čas)
- přesnost regulace konkrétní skupiny genů – neznáme
- faktory uvolňující RNA polymerázu z pauzy – u eukaryot TFIIS (funkční homolog bakteriálních faktorů Gre-A, Gre-B; homology i u archeí)
- regulace probíhá dalším faktorem – PTFB (pozitivní transkripční elongační faktor) – jeho součástí je CDK kináza 9 (ta, která fosforyluje C-terminální doménu RNA Pol II na serinech – na těch, které se mění v průběhu elongace)
- -> fosforylace C-terminální domény vede k uvolnění RNA Pol II
- došlo k iniciaci transkripcie a pauze -> RNA Pol II může pokračovat

- polyténní chromozomy – různé stavy RNA Pol II z hlediska její fosforylace nebo nefosforylace C-terminální domény
- další stupeň regulace – v každé chvíli transkripce je C-terminální doména a její modifikace jiná (vypadá jinak vzhledem k dalším interagujícím molekulám) – „silnice, která se mění z pěšiny lesem na dálnici“

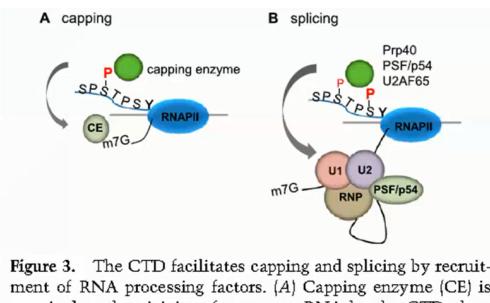


- jak vypadají dlouhé transkripty přepisující pre-mRNA z genů kódujících bílkoviny a krátké transkripty přepisující snoRNA (krátké jadérkové RNA)
- pre-iniciační komplex + iniciační komplex v přítomnosti dalších transkripčních faktorů -> určuje specifické kombinace faktorů, modifikace C-terminální domény -> určuje, které faktory se preferenčně váží
- na obrázku *Nrd1* faktor – faktor zejména zodpovědný za ukončování transkriptů, které nejsou polyadenylovány (kódují geny různých krátkých RNA)
- *Pcf11* – faktor účastnící se polyadenylace (syntéza na 3' konci většiny eukaryotních mRNA – 3' polyA řetízek)
- *Rat1* – 5'→3' exonukleáza, která hraje roli při ukončování transkripce, zároveň interahuje s dalšími faktory (zejména *Pcf11*) – u tohoto typu transkriptů navázání faktorů reflektuje rozdíly ve fosforylací C-terminální domény, reflektuje rozdíly mezi jednotlivými typy mRNA



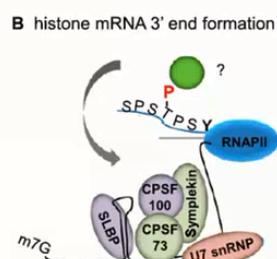
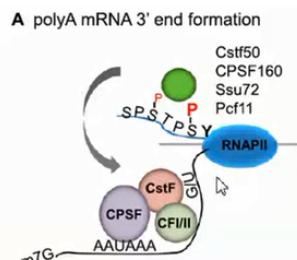
**Figure 2.** Dynamic modification of the CTD during the transcription cycle. At transcription initiation, CDK7 phosphorylates Ser5 and Ser7 residues. Later, during elongation, CDK9 phosphorylates Ser2 and perhaps Thr4, while the phosphate groups on Ser5 and Ser7 are gradually removed by phosphatases. For example, Rtr1, likely indirectly with another phosphatase, and Ssu72, with the aid of the prolyl isomerase Pin1, dephosphorylate Ser5-P early and late during elongation, respectively. Ssu72 also dephosphorylates Ser7-P. CDK12 likely also contributes to Ser2 phosphorylation during elongation of at least some genes. As RNAP II nears termination, Fcp1 dephosphorylates Ser2-P, regenerating unphosphorylated RNAP II that can be recycled for another round of transcription.

- transkripční faktor TFIIH, dochází k fosforylaci časně fosforylovaných serinů (5 a 7), v průběhu transkripce jsou tyto seriny spíše defosforylovány
- další AK (seriny, threoniny, tyrosiny) fosforylovány v průběhu elongace, následně blízko ke konci transkripce dochází k finální defosforylací C-terminální domény

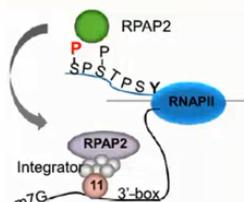


**Figure 3.** The CTD facilitates capping and splicing by recruitment of RNA processing factors. (A) Capping enzyme (CE) is recruited to the vicinity of nascent mRNA by the CTD phosphorylated on Ser5. (B) During transcription, the CTD is phosphorylated on Ser2, while the Ser5-P is dephosphorylated and is involved in recruiting the indicated splicing factors, which defines splice sites and facilitates assembly of the spliceosome. In this and subsequent figures, green spheres above the CTD represent relevant CTD-binding proteins, while assembled functional complexes are indicated below.

- interakční schopnost C-terminální domény se mění v průběhu transkripce
- na počátku vysoká pravděpodobnost interakce s komplexem enzymů, které syntetizují 5' čepičku
- C-terminální doména se mění -> větší šance, že bude interagovat se sestřihovým aparátem, který vyštěpuje introny z pre-mRNA
- -> větší šance, že se objeví bílkoviny, které jsou zodpovědné za odstřížení nascentní mRNA a nastavení 3'-konce polyA řetízkem (polyadenylace)
- není polyA závislá terminace transkripce – může být Nrd1 závislá terminace transkripce (pokud je místo pro vazbu Nrd1 komplexu)



#### C snRNA 3' end formation



- terminace transkripce
- ne všechny RNA u eukaryot kódují mRNA – mRNA ukončena polyA řetízkem
- histonové mRNA – roli hraje U7 snRNP (ribonukleová partikule – ribonukleoprotein – RNA asociované bílkoviny)

**Figure 4.** The CTD functions in 3' processing of both polyadenylated and nonpolyadenylated RNAs. (A) At 3' ends of polyadenylated mRNA, Ser2-P serves to recruit Pef1, a component of CFII (human nomenclature is used for all factors). Other 3' end factors, such as CstF50 and AAUAAA-binding factor CPSF-160 (Yhh1 in yeast), also bind the CTD, whereas Ssu72, with the aid of Pin1 (Ess1), must dephosphorylate Ser5-P. The AAUAAA element and G/U elements are bound by CPSF and CstF, respectively. Loading of some factors, including CPSF and CstF, may occur upstream, perhaps at the promoter [see the text]. (B) The 3' end of histone pre-mRNA contains a stem-loop motif bound by SLBP and a downstream element recognized by U7 snRNP. A complex containing CPSF73, CPSF100, and Symplekin is recruited for 3' cleavage. Thr4-P facilitates this process, likely through a yet-to-be-identified factor. (C) The 3' end of snRNA genes contains a 3' box that interacts with the Integrator complex. RPAP2 binds to Ser7-P on the CTD and to recruit the Integrator, and Int 11, an Integrator subunit, cleaves the RNA.

- roli hrají též CPSF faktory účastnící se štěpení a rozpoznání polyadenylačního signálu na mRNA
- existují modifikace chromatinu a histonů (viz genetika)
- distribuce histonových modifikací v rámci transkripční jednotky, která je přepisována – vysledovatelné gradienty (např. acetylace, metylace) – různé u různých transkripčních jednotek
- transkripční jednotku a transkripční aparát si nelze představovat jako vláček s lokomotivou, která někam dojede a tam zase skončí – proces je dynamický, stochastický, proměnlivý, mění se v každý konkrétní okamžik; každý konkrétní okamžik dochází k syntéze všech přítomných vlivů, výsledkem je, jak běží elongace a jak běží posttranskripční úpravy
- s C-terminální doménou interaguje řada faktorů a komplexů, které se účastní dalších úprav mRNA (u vyšších eukaryot velmi komplexní a regulovaná záležitost, která se může měnit na základě velkého množství vnějších vlivů; mohou se odrážet nejen v regulaci konkrétních faktorů, ale i v podobě chromatinu, interakci s C-terminální doménou atd.)

## Zesilovače (enhancers)

- často repetice nebo tandemové uspořádání
- i velmi daleko od vlastního promotoru
- základní role enhanceru spočívá ve zvýšení koncentrace TF v oblasti promotoru
- specifický příklad enhanceru = kvasinkové UAS (“upstream activating sequence”)
- *eukaryotické TF mají často několik nezávislých domén*
  - *DNA vazebnou*
  - *aktivační* – velmi nespecifické, obsahují bazické oblasti bílkovin, mohou reagovat stejně
  - *regulační* (interakce s dalšími proteiny)
  - “*upstream TF*” nejčastěji interagují s TF, které účinkují v časné iniciaci transkripce (TFIID, TFIIB, TFIIA) a to buď přímo nebo přes kofaktory
  - některé TF mají “*kyselou*” *aktivační doménu* = tzv. kyselé TF (Gal4, Gcn4, VP16 Herpes simplex viru atd.); zvyšují koncentraci v oblasti promotorů, vazba TFIIB a dalších
- enhancer může být na sekvenci výrazně vzdálen od vlastní promotorové oblasti
- může být rozpoznáván dalšími transkripčními faktory -> mění (zvyšuje) efektivní koncentraci dalších faktorů v blízkosti pre-iniciačního komplexu (promotoru), zvyšuje interakci komplexu mediátoru a jeho umístění do pre-iniciačního komplexu
- reaguje s komplexem mediátoru (1 MDaltonový komplex)
  - pro transkripci *in vivo* nestačí pouze obecné transkripční faktory, jsou potřebné transkripční aktivátory, mediátorový komplex, remodelační proteiny, histon acetylasy
  - pre-iniciační komplex se při iniciaci transkripce skládá s mediátorovým komplexem, proteiny regulujícími transkripci a remodelačními proteiny

- mediátorový komplex zprostředkovává interakci RNA pol II s aktivátory transkripce
- komplex 30 různých polypeptidů
- pomáhá iniciaci transkripce tím, že reguluje CTD kinázu
- enhancery mohou být prostorově blízko, ale sekvenčně daleko
- řada eukaryotních transkripčních faktorů má několik nezávislých domén
- existují nástroje umožňující např. vzít DNA vazebnou doménu, fúzovat ji s něčím jiným a sledovat, co se stane, když bude nějaká struktura umístěna poblíž DNA, zda má aktivační/inhibiční účinek na transkripci apod.
- aktivační domény transkripčních faktorů – bývají poměrně nespecifické a kyselé – nespecificky zvyšují pravděpodobnost sestavení pre-iniciačního komplexu – některé domény jen zvyšují efektivní koncentraci jiných transkripčních faktorů

## Terminace transkripce

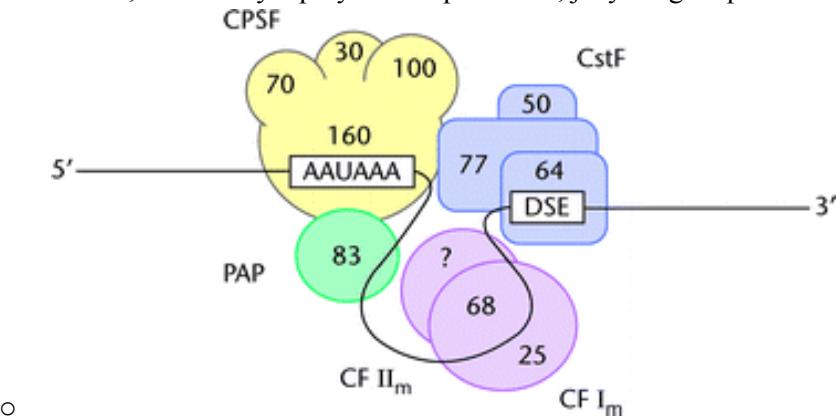
- *RNA Pol I* – končí na terminátorové sekvenci (cca 18 bp) ve vzdálenosti > 1 kbp od reálného 3' konce pre-rRNA; vyžaduje pomocný faktor, trochu funkčně připomíná ρ faktor u bakterií (ale na rozdíl od prokaryotického faktoru se váže na DNA) -> terminace probíhá jako rozpoznání sekvence, vazby faktoru, zastavení RNA polymerázového komplexu
- *RNA Pol III* – připomíná ρ nezávislou terminaci prokaryot – lze rozeznat vlásenkovité struktury na konci transkripčních jednotek; složitější struktury na 3' konci nascentní RNA ukončeny několik zbytky kys. uridylové (2–4 U, které jsou přepsány z GC bohaté oblasti)
- archaea – RNA Pol je vícepodjednotková, více připomíná eukaryotní RNA Pol
  - z hlediska bazálních transkripčních faktorů (obecné – zásadní pro každou transkripci archeální RNA Pol II) hráje zásadní roli faktor TBP – homolog eukaryotního TBP faktoru (součást pre-iniciačního komplexu u všech tří RNA polymeráz)
  - faktor TFB – historický obecný transkripční eukaryotní faktor zásadní pro rozeznání promotoru a pro navázání RNA polymerázy do oblasti promotoru; analog TFIIB (zásadní pro přivedení RNA Pol II do oblasti promotoru)

**Table 2**  
Differences and commonalities on the gene regulation in archaea, bacteria and eukarya.

Structural element	Archaea	Bacteria	Eukarya
DNA-binding TFs	The helix-turn-helix is the most abundant structure	Typically bacterial-like type. The helix-turn-helix is the most abundant structure	Different structures have been identified as Zn-fingers, homeodomains and helix-loop-helix. These structures have not been identified in archaea
DNA topology	Relaxed or positively supercoiled (in hyperthermophilic archaea that contain reverse gyrase)	Negatively supercoiled	Negatively supercoiled
Promoter structure	TATA box and/or initiator element	The most common promoter consists to two conserved boxes at -10 (TATAAT) and -35 (TTGACA) relative to the transcription start site	TATA box and/or initiator element
RNA polymerase	One type; complex subunit structure (from 8 to 13 subunits similar to eukaryal RNA polymerase II); it requires basal transcription factors for promoter recognition	One type. Relatively simple subunit composition (two subunits alpha, beta, beta', omega and sigma factor that can be interchanged). It binds directly to promoter	Three types; complex subunit compositions; require basal transcription factors for promoter recognition/binding
Basal transcription factors	TBP, TFIIB, and TFIIS homologs of eukaryal RNA polymerase II-associated factors described thus far	No	TBP, TAFs, TFIIA, TFIIB, TFIIE, TFIIF, TFIH required for RNA polymerase II initiation; P-TEFb, TFIIS, TFIIF, elongin, and ELL required for elongation
Operonic organization	Yes	Yes	No

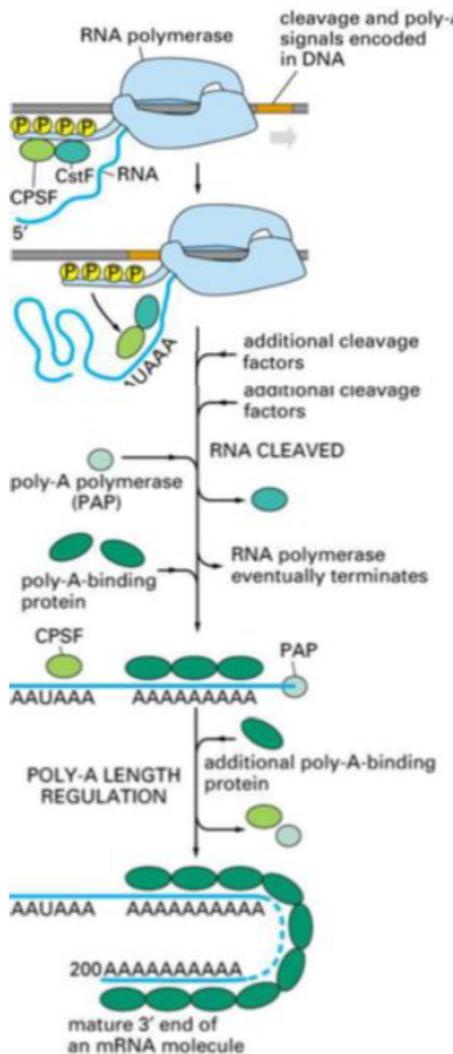
- archeální transkripce – velmi často polycistronní transkripční jednotky (pro jednobuněčné organismy výhodné)
  - z hlediska komplexity genomů je do určité míry výhodné regulovat RNA Pol spíše inhibicí, regulační transkripční faktory jsou spíše bakteriálního typu
- *RNA Pol II*
  - ukončení je třeba rozdělit na dvě nezávislé oblasti: *ukončení vlastního transkriptu a vlastní ukončení transkripce*
  - *ukončení vlastního transkriptu* (v přesně daném místě) – rozpozná polyadenylační signál včetně elementů po proudu transkripce (včetně downstream elementů) –

- sestaven polyadenylační komplex, který vede k přestřížení transkriptu a jeho nastavení na polyA konci řetízkem; transkript dále používán
- *vlastní ukončení transkripce* – RNA Pol II pokračuje i po ukončení transkriptu
    - ustříhnutí transkriptu -> postupná signalizace na C-terminální doméně -> postupná defosforylace C-terminální domény, postupné zastavování RNA polymerázy, produkce nascentního transkriptu (s původně přepisovaným genem nemá mnoho společného); volný 5' konec – exonukleáza degraduje od 5' konce nascentní transkript až k RNA polymeráze, tak zároveň zpomaluje (je defosforylována C-terminální domén) -> zastavení RNA polymerázy po degradaci nascentního transkriptu -> rozpad komplexu
    - 5'->3' exonukleáza – RAD1, XRN2
  - dlouhé terminační oblasti
  - *polyadenylace*
    - 3' konec je generován štěpením syntetizované hnRNA (faktory CFI, CGII – „cleavage factor“ a CStF – „cleavage stimulatory factor“) cca 11 až 30 b za **polyadenylačním signálem AAUAAA + GU nebo U bohatá oblast po proudu transkripce** (= konsensus sekvence + downstream element) (rozeznáván CPSF – „cleavage and polyadenylation specificity factor“) a následnou polyadenylací (*PAP* = *polyA polymeráza*)
    - polyadenylace (polyA) souvisí s ukončením transkripce
    - na RNA polymeráze II jsou celá řada bílkovinových faktorů, které se účastní rozpoznání klíčových signálních sekvencí; většina těchto faktorů je závislá na interakci s fosforylovanou C-terminální doménou
    - na nově vzniklé transkriptu, kde je RNA polymeráza, která stále syntetizuje, dojde k odštěpení vzniklé RNA a k uvolnění celého komplexu RNA polymerázy II
    - syntéza polyA řetízku je kooperativní a závisí na polyA vazebním proteinu
    - délka toho polyA řetízku určuje, jak dobře bude příslušná RNA překládána, tzn. jak moc se z ní budou tvořit proteiny, s jakou frekvencí a jak bude stabilní
    - při každém průběhu translace se tento řetízek o něco zkracuje → je to jakási značka, kolikrát byla polymeráza přeložena, jakýsi signál pro účinnost translace



- - Model štěpícího komplexu („cleavage complex“). Uspořádání poly(A) polymerázy (PAP), štěpících faktorů I a II (CF I a CF II) na RNA není známo. Jsou uvedeny pouze dvě podjednotky CF II. DSE = downstream element.
  - faktory CFI, CFII, CStF, CPSF – rozpoznávají místo polyadenylace, štěpí v konkrétním místě nascentní transkript, umožňují PAP syntetizovat polyA řetízky (bez templátu syntéza cca 200 zbytků kys. adenylové)
  - *polyadenylace probíhá ve dvou krocích*
    - I) závislá na rozpoznávacím faktoru CPSF a tedy i AAUAAA sekvenci – prvních cca 10 b

- 2) nezávislá na CPSF; vyžaduje další faktory rozpoznávající oligo(A) konec (PABP = “poly(A)-binding protein” a další) – dalších cca 200 b
- mRNA pro replikativní histony obratlovců nejsou polyadenylovány (-> *ne všechny geny mají polyA řetízek*). Tvoří vlásenku na 3' konci a reasociuje s U7 snRNA v místě konsensus sekvence 5'-AAGAAAGA-3' (na mRNA)
  - histony, které se dosyntetizovavají ve chvíli, kdy neběží replikace, tzn. neúčastní se replikace, tak obvykle polyA řetízek mají
  - Konec histonů v mRNA tvoří kovalentní vazbu s malou jadernou RNA (u7RNA), ta tam přinese bílkoviny, které jsou zodpovědné za štěpení vlastního transkriptu. Transkript není polyadenylován, protože součástí komplexu není poly A polymeráza.
  - SnRNA, snoRNA – ukončení NRD komplexem, který rozezná NRD sekvenci na RNA, umožní vytržení RNA z transkripčního komplexu. Komplex EXOSOME/Tramp – účastní se degradace RNA od 3' konce, dokud nenarazí na RNA protective element.
  - RNA může mít sekvenci pro NRD, ale pokud není chráněná speciálními proteiny, tak je rychle degradovaná EXOSOME/Tramp (například CUT - cryptic unstable transcript)
- neexistuje „nahá“ RNA – spousta dalších bílkovin, které ji degradují, štěpí apod., spousta RNA v buňce je reálně obalena bílkovinami / obaleny bílkovinami
- polyA řetízek je rozpoznáván polyA vazebními proteiny (zároveň po navázání na první A zbytky zvýší procesivitu a aktivitu polyA polymerázy -> syntéza polyA řetízku běží ve dvou stupních: pomalejší počáteční syntéza a rychlá syntéza po navázání polyA proteinu)
- vliv na celou řadu dalších dějů (co se s RNA bude dít): maturace pre-mRNA, její transport, stabilitu, jak bude translatována (iniciace translace)
- za různých podmínek vznikají různé pre-mRNA s různě dlouhými 3' konci (než jsou ukončeny polyA signálem); v každém buněčném typu za určitých podmínek může být určité spektrum RNA s různě dlouhými 3' konci ukončenými polyA -> i kdyby kódovaly stejnou bílkovinu, 3' nepřekládaný konec může mít různé regulační možnosti, může být buňkou využit pro různé regulace; zastoupení konkrétních forem může regulovat intenzitu překládání určitých proteinů
- řada sekvencí může fungovat jako polyadenylační signál; je podobný u všech eukaryot, některé organismy mohou používat signál odlišný (např. samá A)



#### Vysvětli mechanismus polyadenylace:

- ukončení transkripce: polyadenylační komplex rozezná polyadenylační signál, naváže se na nascentní transkript (RNA Pol II stále syntetizuje) a přeštipne tu nascentní RNA a začne tam syntetizovat PolyA řetízek. Nascentní RNA se uvolní. Pro RNA Pol II je to signál, aby došlo k defosforylací C term. domény (odpadnutí elongačních faktorů). Zároveň se na tom přeštiplym místě objeví nechráněný 5' konec RNA, na kterej se naváže komplex RNAs, který štěpí RNA a jedou hned za RNA Pol II a zastaví ji (torpedo model)
- nascentní transkript je rozpoznán faktory CF1, CF2 (cleavage factor), CStF (cleavage stimulatory factor), poly A polymeráza (syntetizuje RNA bez předlohy) nasyntetizuje pář Áček a pak dojde k rozpadnutí tohoto komplexu. Nově syntetizovaný poly A řetízek je během syntézy poly A polymerázou obsazen poly A binding proteinem (PABP) (hraje roli v transportu RNA z jádra a v iniciaci translace). Ve chvíli, kdy je na RNA navázáno několik PABP, tak to zvýší procesivitu poly A polymerasy a ta nasyntetizuje delší řetízek. PABP taky neobjasněným způsobem určuje délku poly A konce.

## Transkripce v organelách

- Polymerázy i regulace transkripce připomínají transkripci u prokaryot
- V mitochondriích je RNA Pol malá, připomíná tu bakteriální. (kvasinky 145 + 43 kDa)
- V chloroplastech jsou 2 typy polymeráz: jedna připomíná bakteriální (má i sigma faktor); druhá je jednosložková a připomíná virovou Pol.
- Archaeální RNA Pol připomíná eukaryotní Pol. Regulace iniciace transkripce u Archaea – TFB (homologní TFIIB u eukaryot) se váže na T response element, TBP

## Transkripce u Archaea

- jedna RNA polymeráza, ale 13 až 14 podjednotek, některé homologie s podjednotkami eukaryotických RNA pol.
- TF homologní k eukaryotickým TFIIB a TBP se váží na A/T bohatou oblast promotoru
- transkripční jednotky obvykle polycistronní, tzn. způsob regulace transkripce, i ve smyslu, že se přemísťují celé dráhy, je spíše podobný bakteriálnímu typu
- u archebakterií najdeme transkripční aktivátory a transkripční represory
- dále celá řada helix proteinů, které rozpoznávají konkrétní elementy v oblasti promotorů, které se vyskytují u všech typů organismů, jsou typické represory u bakterií
- Iniciace transkripce u archeí: Prvním krokem rozpoznání promotoru je vazba TBP na archaeální TATA box. Tento komplex je stabilizován asociací s TFB. Vazba TFB interaguje s purin-bohatou BRE sekvencí 5' TATA boxu. Tento komplex připojuje RNA polymerázu,

která se váže na oblast DNA po proudu („downstream“) TATA boxu a překrývá transkripční start stranu a oblast po proudu DNA do pozice +18

