

UNIVERZITA KARLOVA V PRAZE

Přírodovědecká fakulta

Studijní program: Speciální chemicko-biologické obory  
Studijní obor: Molekulární biologie a biochemie organismů



**Daniel Pinkas**

Modifikace 5' konce RNA bakterií

5' end modification of bacterial RNA

Bakalářská práce

Školitel: RNDr. Libor Krásný, Ph.D.

Praha, 2016

Děkuji zejména RNDr. Liboru Krásnému, Ph.D. a Natalii Panove, Ph.D. za pomoc při shromažďování zdrojů a kontrole textu. Také bych rád poděkoval všem ostatním, kteří mi jakýmkoliv způsobem pomohli.

**Prohlášení:**

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 12. 08. 2016

.....

## ***Abstrakt***

Regulace genové exprese je klíčovou vlastností všech organismů a může probíhat na několika úrovních počínaje iniciací transkripce až po degradaci proteinů. Jednou z možností je úprava stability mRNA pomocí nejrůznějších modifikací. Asi nejznámější modifikací je 7mG čepička eukaryot, která chrání RNA před RNázami. V nedávné době bylo díky kombinaci metod kapalinové chromatografie a hmotnostní spektrometrie objeveno několik nových modifikací také u prokaryot. Jednou z nich je nikotinamidadeninukleotid na 5' konci některých RNA, který je analogický s 7mG čepičkou. Tato práce popisuje tento fenomén v kontextu bakteriální transkripce.

***Klíčová slova: RNA, RNAP, NAD, 5' modifikace, regulace genové exprese***

## ***Abstract***

Regulation of gene expression is a key feature of all organisms and can occur at several levels ranging from transcription initiation to protein degradation. An important mechanism of this process is regulation of mRNA stability by various modifications. The best known modification is eukaryotic 7mG cap, which protects RNA from RNase degradation. Recently, several new prokaryotic modifications have been discovered thanks to advances in liquid chromatography and mass spectrometry methods. One such a modification is nicotinamide adenine dinucleotide at the 5' end of some RNA. Nicotinamide adenine dinucleotide is analogous to 7mG cap. This study describes this phenomenon in context of bacterial transcription.

***Key words: RNA, RNAP, NAD, 5' modification, regulation of gene expression***

## ***Seznam zkratek***

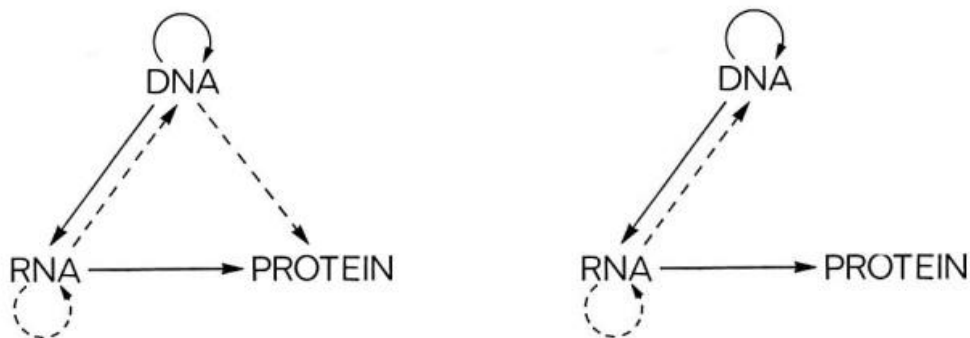
NAD <sup>+</sup>	nikotinamidadeninukleotid
LC/MS	kapalinová chromatografie – hmotnostní spektrometrie
m/z	podíl hmotnosti a náboje
tRNA	transferová RNA
mRNA	mediátorová RNA
rRNA	ribozomální RNA
RNAP	RNA polymeráza
ADPRC	adenosindifosfát-ribosylcykláza
RppH	RNA pyrofosfohydroláza
AMP	adenosinmonofosfát
NTP	nukleosidtrifosfát
NMP	nukleosidmonofosfát
dpCoA	defosfokoenzym A
polyA	polyadeninový
NTD	N-terminální doména
CTD	C-terminální doména
7mG	7metylguanosin
5mC	5metylcytosin

# Obsah

1. Úvod .....	1
2. Syntéza bakteriální RNA .....	3
2.1 Bakteriální RNA polymeráza.....	3
2.2 Promotor .....	4
2.3 Transkripce u bakterií .....	5
2.3.1 Iniclace.....	5
2.3.2 Elongace .....	5
2.3.3 Terminace .....	6
3. Modifikace bakteriální RNA .....	8
3.1 Metody pro identifikaci modifikací RNA.....	8
3.1.1 Kapalinová chromatografie – hmotnostní spektrometrie .....	9
3.1.2 RNA CaptureSeq .....	9
3.2 Nikotinamidadeninukleotid na 5' konci RNA .....	11
3.2.1 NAD captureSeq.....	12
3.2.2 RNAP, enzym zodpovědný za NAD <sup>+</sup> -RNA cap .....	13
3.2.3 Význam NAD <sup>+</sup> čepičky .....	14
3.3 Koenzym A na 5' konci RNA.....	16
4. Závěr.....	19
Seznam citované literatury .....	20

# 1. Úvod

Ribonukleová kyselina neboli RNA je známá jako intermediát mezi DNA a proteiny již velice dlouho. Jako první, kdo s tímto tvrzením přišel, byl roku 1956 Francis Crick, když ve svém článku „The Biological Replication of Macromolecules“ popsal tok genetické informace, dodnes uznávaný a známý jako centrální dogma molekulární biologie (Obr. 1). Kromě replikace DNA, přepisu DNA do RNA a překladau RNA do proteinové struktury také předpokládal, že by se RNA mohla sama replikovat a přepisovat do DNA a také, že se DNA může přepisovat přímo do proteinů. Kromě přepisu DNA přímo do proteinů jsou tyto varianty sice možné, avšak pro živé organismy netypické. Přepis z proteinů (ať už do DNA, RNA nebo replikace proteinu) nebyl pozorován, a proto se předpokládá, že k němu vůbec nedochází (Crick, 1970).



**Obr. 1. Tok genetické informace**

*Vlevo původní varianta, kterou předpokládal Crick, vpravo dnes uznávaná varianta bez přepisu DNA přímo do proteinů. Plné šipky znázorňují běžně uplatňovaný směr toku genetické informace, přerušované šipky znázorňují možný, avšak nepříliš využívaný směr toku (převzato a upraveno podle Crick, 1970).*

Jelikož RNA je nedílnou součástí výše zmiňovaného toku genetické informace, mohou její modifikace zásadním způsobem ovlivnit genovou expresi a přispívat tak k její regulaci. To se může dít různými mechanismy, například zvýšením stability RNA ochranou proti enzymům štěpícím RNA, čímž se prodlouží doba života dané RNA, nebo změnou frekvence překladau do proteinů.

Typickými modifikacemi RNA jsou například 7-metylguanositinová čepička na 5' konci, nebo polyadeninový úsek na 3' konci mediátorové RNA. O takovýchto modifikacích se dříve soudilo, že existují pouze u eukaryotických organismů. V případě polyadeninového 3' konce se zjistilo už relativně dávno, že ho obsahují i některé mRNA bakterií (Nakazato et al., 1975). Jeho funkce se však od funkce polyadeninového konce eukaryot poněkud liší. Zatímco polyadeninový konec eukaryot chrání RNA před degradací a reguluje míru translace (Laishram, 2014), u prokaryot naopak snižuje stabilitu některých mRNA (Mohanty a Kushner, 1999) tím,

že je „označí“ pro degradaci RNázou E (Mohanty a Kushner, 2011). Výjimkou u prokaryot je mRNA pro RNázu E, kde polyadenylace zvyšuje její stabilitu (Mohanty a Kushner, 2002). O modifikacích podobných 7mG čepičce u bakteriální mRNA se však dlouho vůbec nevědělo. Až v nedávné době bylo pomocí nových metod objeveno mnoho nových modifikací RNA, včetně nikotinamidadeninukleotidu (NAD<sup>+</sup>) či acetyl koenzymu A (AcCoA) na 5' konci některých malých RNA a mediátorových RNA u bakterií *Escherichia coli* a *Streptomyces venezuleae* (Chen et al., 2009).

Tato práce si klade za cíl seznámit čtenáře s některými nově objevenými modifikacemi bakteriální RNA. Blíže se pak bude věnovat NAD<sup>+</sup> – způsobu jeho objevení na 5' konci RNA, jeho inkorporaci do RNA a v neposlední řadě také jeho možné funkci jako regulátoru genové exprese.



## 2. Syntéza bakteriální RNA

RNA bývá v bakteriální buňce ve většině případů velice rychle degradována, a proto musí být i stejně rychle doplňována (Hambraeus et al., 2003). Důvodem této dynamiky je nutnost rychle reagovat na změny vnějších podmínek. K syntéze RNA dochází procesem zvaným transkripce, kde předlohou je obvykle DNA, i když v některých případech může templátem být i RNA (Gildehaus et al., 2007).

### 2.1 Bakteriální RNA polymeráza

Enzymem zodpovědným za transkripci je DNA dependentní RNA polymeráza (Hurwitz, 2005). Eukaryotické organismy využívají pro tyto účely několik různých RNA polymeráz, zatímco bakterie si vystačí s jedinou.

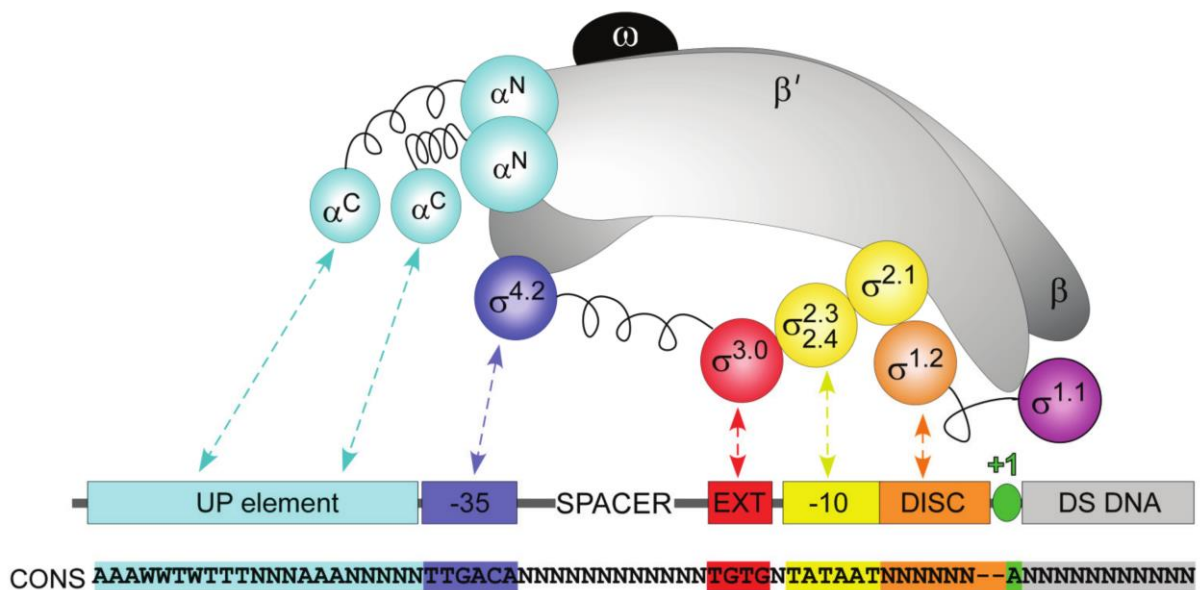
Bakteriální RNA polymeráza (RNAP) se skládá z pěti základních podjednotek ( $\alpha\alpha\beta\beta'\omega$ ) o souhrnné hmotnosti 370 kDa (Burgess, 1969), které tvoří proteinové jádro, jehož struktura je konzervovaná napříč celou doménou bakterií (Murakami, 2015). Alfa podjednotky tvoří homodimer, na který se váží podjednotky  $\beta$  a  $\beta'$ . Tvoří tedy jakési lešení, na kterém se formuje zbytek komplexu. Podjednotky  $\beta$  a  $\beta'$  tvoří typickou strukturu „krabího klepeta“, která váže DNA a je katalytickým centrem (Obr. 3) (Murakami, 2015). Podjednotka  $\omega$  se váže na  $\beta'$  a napomáhá sestavení celého komplexu (Ghosh et al., 2001). Některé gram pozitivní bakterie mají navíc ještě  $\delta$  podjednotku, která slouží k regulaci (Rabatinová et al., 2013), a  $\varepsilon$  podjednotku, jejíž funkce zatím není známá (Keller et al., 2014).

Kromě jádra je součástí RNAP ještě podjednotka  $\sigma$  (nebo také „faktor sigma“) (Richard et al., 1969), která dohromady s jádrem RNAP tvoří holoenzym. Její funkcí je rozpoznání promotorové oblasti na DNA a zprostředkování kontaktu mezi RNAP a DNA (Sorenson a Darst, 2006). Poté, co RNAP iniciuje transkripci,  $\sigma$  se může odpoutat od holoenzymu, navázat se na další RNAP a opět iniciovat transkripci (Travers a Burgess, 1969). V bakteriální buňce obvykle existuje více různých podjednotek  $\sigma$ . Množství závisí na organismu - může se lišit od jediného typu až po více než 100 různých  $\sigma$  podjednotek (Paget, 2015). Každá podjednotka  $\sigma$  propůjčuje RNAP afinitu vůči různým promotorům, což jsou regulační sekvence DNA, na kterých začíná transkripce. To má za následek transkripci rozdílných genů a tedy i jinou genovou expresi. Různé sigmy jsou bakteriemi produkovány v závislosti na podmínkách. Například  $\sigma^{70}$  je tzv. „House keeping“ faktor sigma bakterie *E. coli*, který je zodpovědný za transkripci většiny RNA v nestresových podmínkách (Feklistov et al., 2014). Dalšími faktory sigma jsou například  $\sigma^{32}$ , který je produkován v *E. coli* při teplotním šoku a řídí syntézu RNA

genů pro adaptaci na vysokou teplotu (Grossman et al., 1984), nebo SigE účastníci se transkripce při sporulaci *B. subtilis* (Haldenwang et al., 1981).

## 2.2 Promotor

Promotor je sekvence na DNA rozpoznávaná RNAP holoenzymem (Obr. 2), konkrétně jeho podjednotkou  $\sigma$ . Podjednotka  $\sigma$  se váže na pozici promotoru -10, tedy přibližně 10 párů bází proti směru transkripce od iniciačního +1 nukleotidu. Tato pozice -10, známá též jako Pribnowův nebo TATA box, je v evoluci zachovalý motiv, jehož konsensus sekvence je u mnoha bakteriálních druhů  $^{-12}\text{TATAAT}^{-7}$  (Harley a Reynolds, 1987). K tomuto motivu patří ještě takzvaná prodloužená -10 sekvence  $^{-15}\text{TGn}^{-13}$  (Keilty a Rosenberg, 1987). Tato sekvence je v kontaktu s faktorem sigma a její absence výrazně snižuje sílu promotoru (Burr et al., 2000). Další velice důležitou součástí promotoru, která je svou interakcí se  $\sigma$  podjednotkou zodpovědná za rozeznání promotoru a specifitu k danému promotoru (Hook-Barnard a Hinton, 2007), je -35 pozice s konsensus sekvencí  $^{-35}\text{TTGACA}^{-30}$  (Hawley a McClure, 1983).



**Obr. 2. Interakce RNAP s promotorovou oblastí**

Schéma znázorňuje interakci RNAP s promotorem. Různé domény podjednotky  $\sigma$  interagují s různými pozicemi na promotoru –  $\sigma^{2.4}$  s -10 pozicí,  $\sigma^{3.0}$  s -10 prodlouženou sekvencí (EXT) a  $\sigma^{4.2}$  s -35 pozicí. C-terminální domény (CTD)  $\alpha$  podjednotek mohou interagovat s některými sekvencemi promotoru, které jsou ještě dále proti směru transkripce než -35 pozice. Dále je na obrázku zobrazena promotorová konsensus sekvence (převzato a upraveno podle Ruff et al., 2015).

## 2.3 Transkripce u bakterií

Transkripce je enzymatický proces, který lze rozdělit na tři základní kroky: iniciace, elongace a terminace. Hlavním enzymem účastnícím se tohoto procesu je výše zmíněná RNAP, avšak i další faktory mohou významně ovlivnit průběh transkripce.

### 2.3.1 Iniciace

Prvním krokem v iniciaci transkripce je rozpoznání promotorové oblasti na DNA podjednotkou  $\sigma$  holoenzymu (samotná  $\sigma$  ani jádro RNAP toho nejsou schopny) a následné nasednutí RNAP na dvouvláknovou DNA za vzniku tzv. uzavřeného komplexu. Následné konformační změny indukují vznik otevřeného komplexu. Ten se vyznačuje přítomností 13 nukleotidů velké (nukleotidy -11 až +2) transkripční bubliny (rozvolněná dvoušroubovice DNA na jednotlivá vlákna) v aktivním místě mezi podjednotkami RNAP  $\beta$  a  $\beta'$  (Ruff et al., 2015).

Dalším krokem je párování iniciačního NTP s +1 pozicí a prodlužujícího NTP s +2 pozicí na templátovém vlákně, které probíhá standardním Watson-Crickovským párováním, čímž se stabilizuje otevřený komplex (Gaal et al., 1997). Iniciačním substrátem mohou být kromě jednotlivých nukleotidů také oligonukleotidy dlouhé 2 – 4 báze (Vvedenskaya et al., 2012), případně další molekuly (diskutováno v 3.2.2). Mezi těmito nukleotidy vzniká fosfodiesterová vazba, konkrétně mezi 3'OH skupinou ribózy iniciačního nukleotidu a  $\alpha$  fosfátem prodlužujícího nukleotidu, pro niž je energie dodána z hydrolýzy prodlužujícího NTP na NMP a pyrofosfát. Stejným způsobem se k nově vznikajícímu vlákně RNA připojí i třetí NTP. Avšak se vznikem druhé (a také každé další v průběhu iniciace) fosfodiesterové vazby RNAP vtahuje templátovou DNA dovnitř komplexu (a kódující vlákno ven z komplexu, přičemž sama RNAP zůstává uchycena k promotoru), tím posune +4 nukleotid do aktivního místa enzymu a zvětší transkripční bublinu (Kapanidis et al., 2006; Revyakin et al., 2006; Winkelman et al., 2015). Zde se pak může navázat další nukleotid a vytvořit vazbu s nukleotidem na předchozí pozici. Obdobně probíhá i připojování dalších nukleotidů v průběhu iniciace. V této fázi může docházet k abortivní iniciaci, při které se produkt dlouhý 2-15 nukleotidů oddělí od RNAP a tak musí začít transkripci nanovo (Goldman et al., 2009).

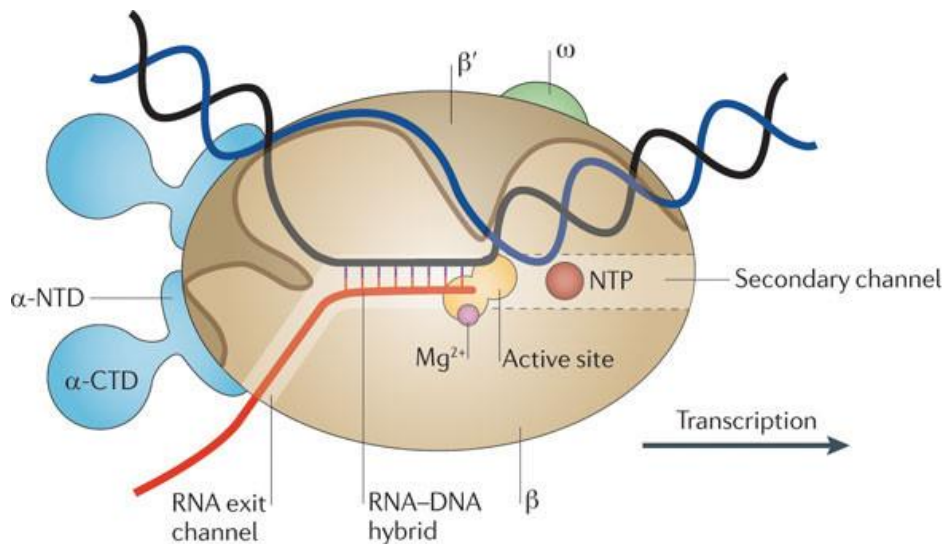
Iniciace transkripce je zřejmě nejvíce regulovaným procesem v buňce (Yang a Lewis, 2010) a její frekvence se v závislosti na sekvenci promotoru může lišit až o 4 řády (McClure, 1985).

### 2.3.2 Elongace

Pokud se RNAP podaří syntetizovat úsek dlouhý alespoň 12 nukleotidů, naruší tím vazbu mezi sigma podjednotkou a zbytkem enzymu. Faktor sigma disociuje od RNAP a tím jí umožní

odpoutat se od promotorové oblasti na DNA, což je považováno za počátek elongace (Nickels et al., 2005). Podobně jako při iniciaci, i při elongaci se do aktivního místa RNAP váží NTP na základě Watson-Crickovského párování bází s templátovým vláknem DNA, přičemž je katalyzován vznik fosfodiesterové vazby a dochází k prodloužení transkriptu. Rozdílem je však to, že posun aktivního místa je způsoben pohybem RNAP po vlákně, nikoliv vtahováním vlákna do RNAP (Vassilyev, 2009).

V této fázi jsou rekrutovány další faktory regulující průběh transkripce. Proteiny GreA a GreB pomáhají vyprostit zastavenou RNAP, která sklouzla zpět například kvůli chybnému párování bází, stimulaci její nukleolytické aktivity (Borukhov et al., 2001). Tímto se zasluhují o hladký průběh elongace, která je tak při dostatečné koncentraci substrátu relativně rychlým procesem.



### Obr. 3. Elongační komplex

*Schéma elongačního komplexu RNAP, DNA a RNA znázorňuje základní podjednotky RNAP, aktivní místo, interakci mezi jednotlivými molekulami a směr transkripce (převzato a upraveno podle Santangelo a Artsimovitch, 2011).*

### 2.3.3 Terminace

Transkripce končí oddělením vlákna nově vznikající RNA od komplexu RNAP s DNA. K tomu může dojít buď pouze na základě sekvence DNA, takzvaná vnitřní terminace, nebo s přispěním vnějších faktorů, jako jsou proteiny Rho nebo DksA (Washburn a Gottesman, 2015).

Vnitřní terminace, někdy také označovaná jako Rho nezávislá, je umožněna GC bohatou invertovanou repeticí následovanou A-bohatou sekvencí na templátovém vlákně přepisované DNA (Carafa et al., 1990). Invertovaná repetice vytváří vlásenku, která za podpory oslabení

duplexu RNA/DNA AU bohatou oblastí (AU párování je slabší než GC) vytrhne RNA z elongačního komplexu (Komissarova et al., 2002). Účinnost takovéto terminace se může lišit v závislosti na velikosti vlásenky (Washburn a Gottesman, 2015).

Rho dependentní terminace je alternativní možností, kterou využívá většina známých bakterií. *E. coli* tuto možnost využívá přibližně ve 20 % případů terminace (Peters et al., 2009) a je pro ni tudíž nezbytná. Rho je hexamer s RNA/DNA helikázovou aktivitou (Brennan et al., 1987) složený ze šesti shodných podjednotek (Skordalakes a Berger, 2003). Rho skenuje nově vznikající RNA, a jakmile narazí na tzv. *rut* sekvenci, která kóduje Rho dependentní terminaci, naváže se na něj a stimuluje tak svoji helikázovou aktivitu, čímž uvolní RNA z elongačního komplexu (Koslover et al., 2012).

### **3. Modifikace bakteriální RNA**

Většina donedávna známých modifikací RNA, kterých bylo do roku 2013 známo 144, se týká hlavně transferové RNA a ribozomální RNA (Machnicka et al., 2013). Tyto zahrnují například izomery, z nichž nejznámějším je asi pseudouridin ( $\Psi$ ) (Charette a Gray 2000), metylaci na 2' hydroxylové skupině ribózy nebo adice různých struktur přímo na nukleotidy (Machnicka et al., 2013). Nicméně pomocí technik vyvinutých v nedávné době začaly být ve větší míře objevovány i modifikace mediátorové RNA a některých nekódujících RNA (Schwartz et al., 2014), a to nejen u eukaryotických organismů, kde byly takové modifikace známé již dříve, ale také u organismů prokaryotního typu, konkrétněji u některých bakterií (Chen et al., 2009). To je poměrně důležitý objev, protože výzkum funkce a samotné biologické podstaty podobných modifikací je u bakterií nepoměrně jednodušší než u eukaryotických organismů.

Nejčastěji bývají modifikace RNA prováděny post transkripčně přidáváním různých zbytků (například metyl) na řadu specifických míst. To může mít za následek například změnu konformace vlákna RNA, ale i jiné párování bází. Také různé proteiny mohou specificky rozpoznávat pouze místa, která jsou modifikována určitým způsobem. Toto jsou jen některé příklady funkcí, které mohou modifikace na RNA zastupovat (Yue et al., 2015).

#### **3.1 Metody pro identifikaci modifikací RNA**

Bakterie byly jako modelový organismus využívány pro studium modifikací rRNA a tRNA, kterých bylo do roku 2011 známo přes 60 (Cantara et al., 2010), díky jejich nenáročnosti a rychlosti reprodukce. Pro mRNA toto však až do nedávna neplatilo, protože podobné modifikace nebyly příliš známé. To se změnilo s dostupností nových technik pro detekci a identifikaci „abnormálních“ částic v molekule RNA. Starší techniky detekce, jako například chromatografie na tenké vrstvě (TLC) nebo UV spektrofotometrie (Lane, 1998), byly nahrazeny kombinovanou technikou kapalinové chromatografie a hmotnostní spektrometrie (LC/MS) (Kowtoniuk et al., 2009). Novější metody zvládnou určit a kvantifikovat molekulu RNA, ve které se daná modifikace nachází. Tyto metody sestávají ze selekce modifikované RNA a následné sekvenace (tzv. RNA CaptureSeq). To umožňuje výzkum funkce a biologického významu těchto modifikací, který se proto dnes začíná rozmáhat (Marbaniang a Vogel, 2016).

### 3.1.1 Kapalinová chromatografie – hmotnostní spektrometrie

Kombinace metod vysokoúčinné kapalinové chromatografie (HPLC) a hmotnostní spektroskopie (MS), dále již jen LC/MS, byla pro detekci modifikovaných nukleotidů poprvé použita před několika lety, ačkoliv metoda samotná je známá již delší dobu. Použitím LC/MS metody jsou nejprve změřeny retenční faktory vzorku, který je pak za vysokého tlaku z HPLC zaveden do hmotnostního spektrometru, kde je převeden do plynné fáze a ionizován. Jednotlivé částice jsou pak separovány na základě poměru své hmotnosti a svého náboje. Výsledkem je soubor molekulárních hmotností (a retenčních faktorů) jednotlivých molekul (Pomerantz a McCloskey, 1990).

Pokud je cílem detekce neznámých modifikací RNA, je třeba před LC/MS metodou vystavit vzorek izolované RNA nukleázám, aby byly měřeny hmotnosti pouze jednotlivých nukleotidů. Výsledný soubor molekulových hmotností je porovnán s molekulovými hmotnostmi kanonických nukleotidů a modifikovaných nukleotidů, které jsou již známy. Pokud molekulová hmotnost neodpovídá žádné z dosud známých molekulových hmotností, může se tak s určitou pravděpodobností jednat o zatím neznámou modifikaci RNA (Kowtoniuk et al., 2009).

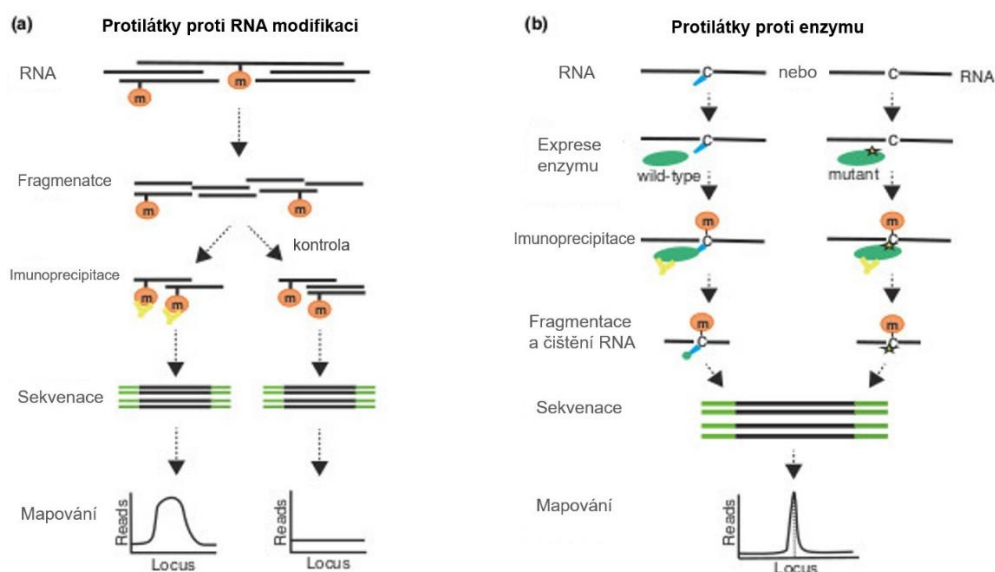
Některé molekuly lze díky jejich hmotnosti identifikovat přímo, jindy je třeba znát alespoň sumární vzorec. Kultivací bakterií v médiu, kde jsou jediným zdrojem uhlíku sloučeniny s uhlíkem  $^{13}\text{C}$ , podrobení RNA z tohoto vzorku LC/MS a porovnáním molekulových hmotností se vzorkem, který je z bakterií rostoucích na uhlíku  $^{12}\text{C}$ , lze určit přesný počet uhlíků v daném vzorku. Stejný postup je možno aplikovat i na ostatní prvky, jako je dusík nebo kyslík, dokud nebude zastoupení jednotlivých prvků v neznámé molekule zcela zřejmé. Ze znalosti sumárního vzorce již lze predikovat strukturu a porovnáním MS/MS analýzy se známým vzorkem stejné molekuly je možné tuto predikovanou strukturu potvrdit (Chen et al., 2009).

### 3.1.2 RNA CaptureSeq

Příchod RNA CaptureSeq metod umožnil určení identity (a také kvantity) RNA, která je modifikována určitým způsobem (Cahová et al., 2015). Díky tomu je možné začít zkoumat nejen význam modifikací pro RNA, ale také relevanci pro celý organismus. Přístupy k identifikaci RNA se mohou lišit v závislosti na charakteru modifikace, obecně se ale dají rozdělit podle způsobu provedení na imunoprecipitační (Obr. 4) a na chemickou modifikaci RNA (Obr. 5), která je následovaná reverzní transkripcí.

Imunoprecipitační metody využívají, jak již název napovídá, protilátky specifické k dané modifikované bázi (Dominissini et al., 2012). Alternativou je využití protilátek k enzymu, který je zodpovědný za modifikaci báze. Pro využití druhého přístupu je třeba nejprve upravit danou

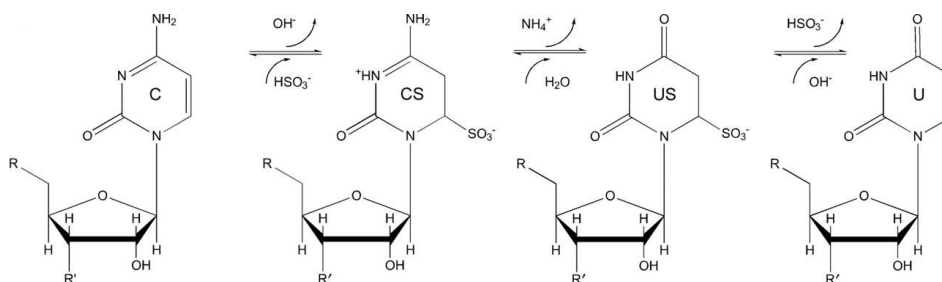
bázi (Khoddami a Cairns, 2013) nebo enzym (Hussain et al., 2013) tak, aby enzym zůstal kovalentně připojen k RNA. RNA značenou protilátkami, tedy tu, která obsahuje dané modifikace, lze izolovat, purifikovat, převést do cDNA knihoven a sekvenovat.



**Obr. 4. RNA CaptureSeq metody založené na imunoprecipitaci**

Vazba protilátek (a) přímo na modifikovanou bázi nebo (b) na enzym kovalentně vázaný na RNA v místě jeho působení, následovaná sekvenací a mapováním (převzato a upraveno podle Marbaniang a Vogel, 2016).

Druhý přístup zahrnuje různé chemické modifikace RNA, díky kterým je možné po reverzní transkripci a sekvenaci identifikovat modifikovanou RNA. Příkladem je využití hydrogensířičitanu pro identifikaci RNA obsahující 5mC, cytosin metylovaný na pozici 5 (Schaefer et al., 2009). Při správně nastavených podmínkách je hydrogensířičitan schopen deaminovat veškeré cytosiny v molekule RNA (Obr. 5) s výjimkou 5mC. Při provedení reverzní transkripce se deaminované cytosiny, které jsou shodné s uracilem, párují s adeninem, zatímco metylovaný cytosin se klasicky páruje guanosem. Ve vlákně nově syntetizované DNA se tedy vyskytují guanosiny pouze na místech, kde templátová RNA obsahovala 5mC.



**Obr. 5. Konverze cytosinu na uracil**

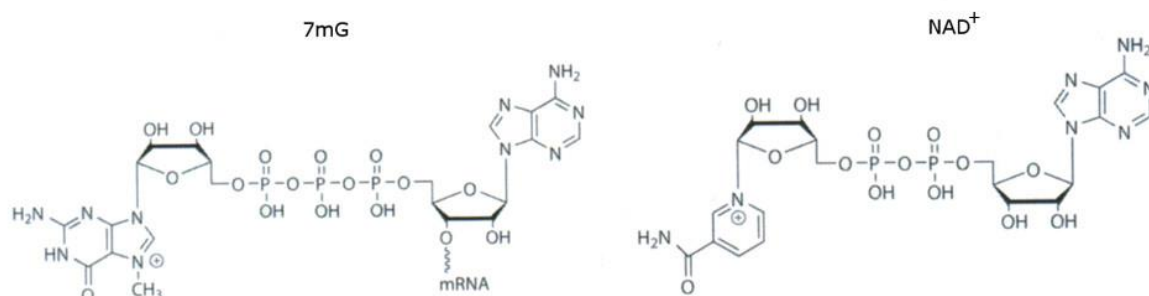
Deaminací cytosinu v molekule RNA katalyzovanou hydrogensířičitanem vzniká uracil (převzato a upraveno podle Schaefer et al., 2009).



## 3.2 Nikotinamidadenindinukleotid na 5' konci RNA

Jedna z molekul, které byly nově objeveny ve vzorcích celkové RNA z bakterií *Escherichia coli* a *Streptomyces venezuelae* díky LC/MS, měla poměr hmotnosti k náboji  $m/z = 540,0533$ . Po kultivaci bakterií na médiu, kde byla jediným zdrojem uhlíku glukóza obsahující pouze izotop  $^{13}\text{C}$ , se molekulová hmotnost zvětšila o 15 daltonů. Obdobně se molekulová hmotnost zvedla o 5 daltonů po kultivaci na médiu s jediným zdrojem dusíku  $^{15}\text{N}$ . Tyto výsledky umožnily odvodit sumární vzorec neznámé molekuly:  $\text{C}_{15}\text{H}_{20}\text{N}_5\text{O}_{13}\text{P}_2$ . Tandemová hmotnostní spektrometrie identifikovala část této molekuly jako adenosindifosfát. Jelikož zbytek molekuly ( $\text{C}_5\text{H}_7\text{O}_3$ ) by mohl být nesaturovaný zbytek ribózy, vzniklý rozpadem nějaké větší molekuly, byla LC/MS analýza nastavena na mírnější ionizační podmínky. Po provedení této upravené LC/MS se objevila nová neznámá molekula s poměrem hmotnosti k náboji 662,1032, přičemž molekula s  $m/z = 540,0533$  nebyla detekována.  $M/z = 662,1032$  odpovídá hmotnosti koenzymu nikotinamidadenindinukleotidu, takže nově modifikovaná báze byla identifikována jako  $\text{NAD}^+$  (Chen et al., 2009).

Struktura  $\text{NAD}^+$ , která je analogická s 7mG čepičkou prokaryot (Obr. 6), napovídá, že by se mohlo jednat o dinukleotid na úplném 5' konci RNA. Tato hypotéza byla testována a potvrzena pomocí P1 nukleázy. P1 nukleáza štěpí molekulu RNA od 5' konce a vytváří nukleotidmonofosfáty, které nově obsahují OH skupiny z vody využitě pro nukleofilní atak. Jedinou bází, která tuto hydroxylovou skupinu neobsahuje, je ta, která byla původně na 5' konci. Pokud je v reakční směsi voda, jež obsahuje izotop kyslíku  $^{18}\text{O}$ , dojde k hmotnostnímu posunu oněch nukleosidmonofosfátů, který je možné změřit. U  $\text{NAD}^+$  k hmotnostnímu posunu nedošlo, což naznačuje jeho pozici na samotném 5' konci a potvrzuje výše zmíněnou hypotézu (Chen et al., 2009).



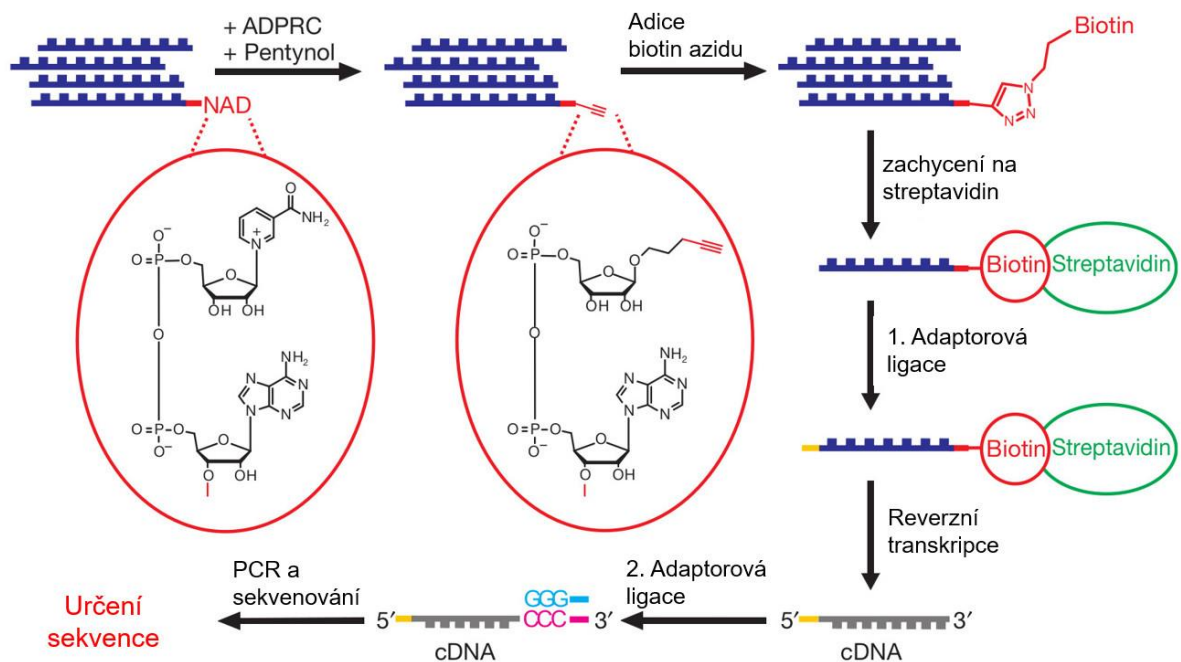
**Obr. 6. 7 metylguanosenová čepička a nikotinamidadenindinukleotid**

Porovnání struktury 7 metylguanosenové čepičky na 5' konci molekuly RNA (vlevo) a struktury  $\text{NAD}^+$  (vpravo) (převzato z Panova a Krásný., nepublikovaná data).

### 3.2.1 NAD captureSeq

Týmu vědců z německé univerzity v Heidelbergu se podařilo vyvinout metodu na izolaci RNA s NAD<sup>+</sup> na svém konci, kterou pojmenovali NAD captureSeq (Obr. 7). Tato metoda specificky upravuje NAD<sup>+</sup> tak, že je možné ho kvantitativně ze vzorku vychytat a tím i všechny molekuly k němu připojené (hlavně tedy RNA).

Enzym adenosindifosfát-ribosylcykláza (ADPRC) z mořského měkkýše *Aplysia californica* dokáže velice specificky provést transglykosylaci alkynylalkoholem (jako nejlepší se ukázal 4-pentyn-1-ol) na dusíku v nikotinamidovém cyklu (Preugschat et al., 2008). Na ten je možné adicí katalyzovanou mědí přidat azid spojený s biotinem (Tornøe et al., 2002). Biotinilovanou molekulu lze vychytat na streptavidinové kuličce. Výsledný vzorek je pak dále možné amplifikovat, sekvenovat a určit jeho identitu (Cahová et al., 2015).



**Obr. 7. NAD CaptureSeq protokol**

Schéma popisující jednotlivé kroky NAD CaptureSeq metody. Jako alkynylalkohol pro transglykosylaci byl použit 4-pentyl-1-ol (převzato a upraveno podle Cahová et al., 2015).

Této metodě byla podrobena celková RNA z exponenciální fáze bakterie *E. coli* (konkrétně kmeny JM109 a K-12). Mezi identifikovanými molekulami bylo velké množství malých regulačních RNA, například RNA I, jejíž funkce je kontrola replikace plazmidů typu ColE1 (Lacatena a Cesareni, 1981). Kromě toho byly identifikovány i fragmenty mRNA genů pro některé enzymy důležité v metabolismu a ve stresové odpovědi (Cahová et al., 2015). Zatím není zcela jasné, proč se NAD<sup>+</sup>-mRNA vyskytuje pouze jako fragmenty, ale s největší

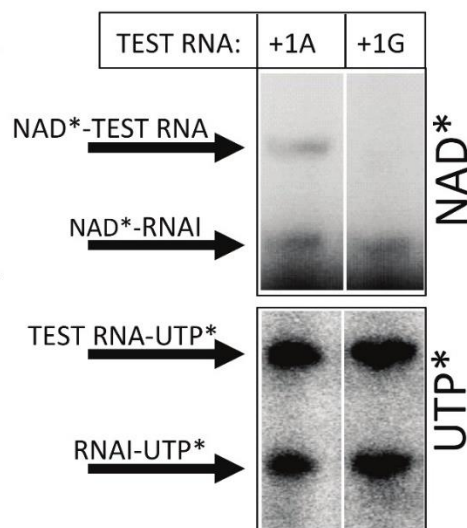
pravděpodobností se nejedná o artefakt použité metody. V úvahu tedy připadá předčasná terminace elongace, nebo nějaká neznámá endonukleáza specificky rozpoznávající RNA s NAD<sup>+</sup> čepičkou.

### 3.2.2 RNAP, enzym zodpovědný za NAD<sup>+</sup>-RNA cap

Další přirozenou otázkou je, jaký enzym a v jaké fázi „života“ RNA je zodpovědný za přítomnost nikotinamidové „čepičky“ na konci RNA. Na tuto otázku se názory a důkazy rozcházejí. Jedna skupina vědců se domnívá, že, podobně jako 7mG čepička eukaryotních organismů (Topisirovic et al., 2011), je NAD<sup>+</sup> připojován k RNA až posttranskripčně (Chen et al., 2009; Luciano a Belasco, 2015). Novější studie však přicházejí s důkazy, že tomu tak není (diskutováno níže).

Bakteriální RNA polymeráza je známá tím, že může za určitých podmínek iniciovat transkripci nejen jednotlivými nukleosidtrifosfáty, ale také takzvanými nanoRNA, tedy oligonukleotidy dlouhými povětšinou 2 – 4 báze (Vvedenskaya et al., 2012). Nebylo by proto až tak překvapující, kdyby jako jeden ze substrátů pro iniciaci replikace mohla RNAP využívat také NAD<sup>+</sup>. Tato hypotéza byla v současnosti potvrzena na modelu RNAP z *E. coli*. Jako modelová RNA byla vybrána RNA I, u které již dříve byla zjištěna nikotinamidová čepička. Po přidání radioaktivně značeného NAD<sup>+</sup> do transkripční směsi byl na screenu polyakrylamidového gelu patrný proužek (pouze radioaktivně značené jsou na screenu detekovatelné) RNA o velikosti odpovídající RNA I (Bird et al., 2016), a viz obrázek 8. Díky tomu je možné RNAP definitivně označit jako enzym zodpovídající za přítomnost NAD<sup>+</sup> na 5' konci RNA.

To, na čem závisí, zda si RNAP pro svou iniciaci vybere jako substrát NAD<sup>+</sup>, je v současné době ve fázi intenzivního výzkumu. Kromě dostupnosti (koncentrace) jednotlivých potenciálních iniciačních molekul je, zdá se, hlavním faktorem sekvence promotoru, převážně pak její oblasti v okolí +1 nukleotidu. Zásadní podmínkou je přítomnost adeninu na pozici +1 (Obr. 8), protože se, stejně jako adenin, může NAD<sup>+</sup> párovat pouze s tyminem na templátovém vlákně přepisované DNA (Bird et al., 2016).



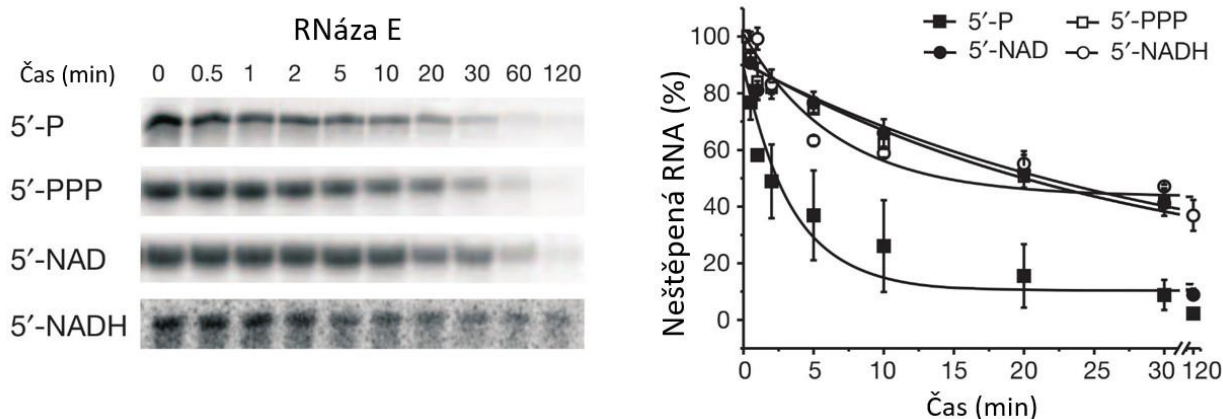
**Obr. 8. Transkripce s radioaktivně značeným  $NAD^+$**

*In vitro* transkripce s radioaktivně značeným  $NAD^+$  (nebo  $UTP^*$  jako kontrola funkčnosti promotoru) ze dvou stejných promotorů v jiné oblasti plazmidu ( $NAD^*$ -TEST RNA a  $NAD^*$ -RNA I). Experiment byl proveden dvakrát, kdy poprvé měla TEST RNA na pozici +1 adenin a v druhém případě guanosin (RNA I měla v obou případech na pozici +1 adenin). Výsledky dokazují schopnost RNAP začít transkripci  $NAD^+$  a také nutnost výskytu adeninu na pozici +1 pro takovouto iniciaci (převzato z Panova a Krásný, nepublikovaná data).

### 3.2.3 Význam $NAD^+$ čepičky

O roli nikotinamidové čepičky se v současné chvíli spíše diskutuje. Existuje sice již několik hypotéz a jejich důkazů, její skutečný význam zatím však není zcela objasněn. Přihlédneme-li k analogii s 7mG čepičkou mRNA prokaryot, nabízí se otázka, zda by jejich funkce nemohla být analogická. Jednou z mnoha funkcí 7mG čepičky je ochrana RNA před degradací nukleázami (Shimotohno et al., 1977).

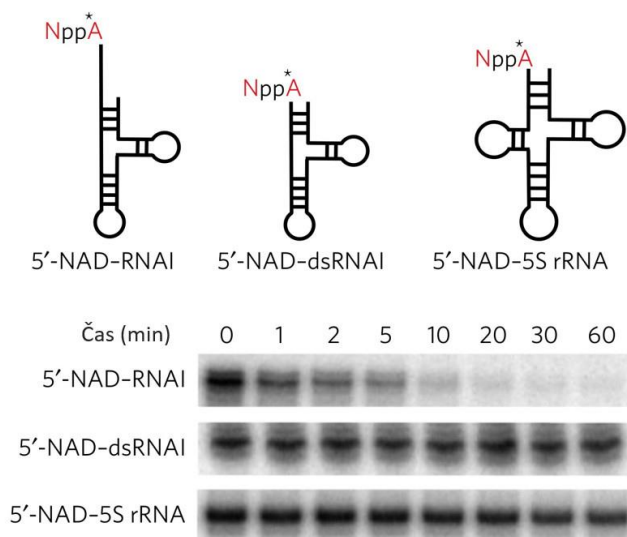
RNáza E je enzym vyskytující se v multiproteinových komplexech degradosomech, a je zodpovědný za degradaci mnoha mRNA v bakterii (nejen) *E. coli* (Mackie, 1998). Jeho účinnost závisí především na charakteru 5' konce RNA (Obr. 9) – substrátem s nejlepší efektivitou degradace je nukleosidmonofosfát na 5' konci. Je-li na 5' konci RNA nukleosidtrifosfát (ten bývá nejčastěji použit jako iniciační substrát pro transkripci), pak je efektivita degradace poměrně nízká (Mackie, 1998). Stejně nízká, nebo dokonce ještě nižší efektivita degradace RNázou E byla potvrzena pro  $NAD^+$  (Cahová et al., 2015). V buňce však existují mechanismy, které spouštějí degradaci mRNA RNázou E. RNA pyrofosfohydroláza (RppH) převádí trifosfát na 5' konci RNA na monofosfát, a tím v podstatě spouští její degradaci RNázou E (Deana et al., 2008). Pokud má však RNA na svém konci  $NAD^+$ , stává se „rezistentní“ vůči RppH, a tím i RNáze E (Cahová et al., 2015).



**Obr. 9. Degradace RNA RNázou E v čase**

Na obrázku vlevo jsou vidět radioaktivně značené  $\alpha$ -fosfáty v celém vlákně RNA. Graf sestavený z těchto dat znázorňuje rychlost degradace RNA RNázou E v čase v závislosti na charakteru 5' konce RNA (převzato a upraveno podle Cahová et al., 2015).

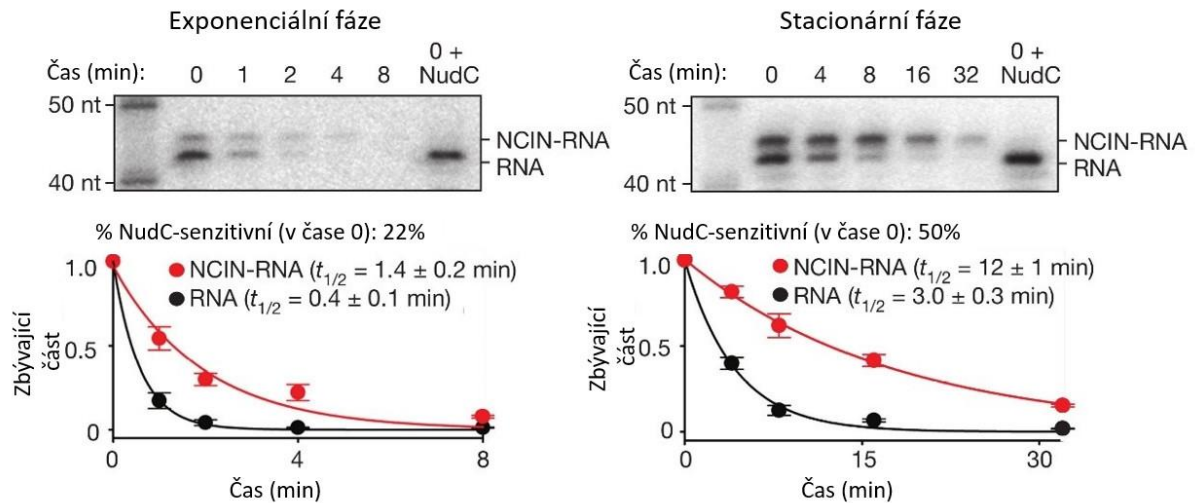
Mohlo by se tedy zdát, že  $\text{NAD}^+$  čepička zásadním způsobem zvyšuje stabilitu RNA. V *E. coli* však fungují i další enzymy spouštějící degradační dráhu. NudC je fosfohydroláza patřící do rodiny nudix proteinů (stejně jako RppH - Song et al., 2013), která hydrolyzuje  $\text{NAD}^+$  na nikotinamidmononukleotid a AMP (Frick a Bessman, 1995). Tento enzym přednostně štěpí  $\text{NAD}^+$ , které je v komplexu s jednovláknovou RNA (Höfer et al., 2016; Obr. 10) a spouští tak degradaci takové RNA RNázou E nezávisle na RppH (Cahová et al., 2015).



**Obr. 10. Jednovláknová NAD-RNA jako substrát pro NudC**

Data ukazující efektivitu NudC fosfohydrolázy pro různé substráty. Jednovláknová RNA I s  $\text{NAD}^+$  na 5' konci (5'-NAD-RNAI) je štěpena velice účinně, zatímco dvouvláknová RNA I (5'-NAD-dsRNAI) a kontrolní dvouvláknová 5S rRNA (5'-NAD-5S rRNA) s  $\text{NAD}^+$  na 5' konci nejsou štěpeny vůbec, nebo jen velice pomalu (upraveno podle Höfer et al., 2016).

Experimenty s kmenem *E. coli* bez funkčního genu pro NudC byla potvrzena schopnost  $\text{NAD}^+$  čepičky prodloužit životnost RNA *in vivo*. RNA s  $\text{NAD}^+$  na svém 5' setrvala v buňce až čtyřikrát déle (Obr. 11) než RNA bez čepičky. Dalším významným zjištěním při těchto experimentech bylo, že se ve stacionární fázi růstové křivky bakteriální kultury oproti exponenciální fázi výskyt  $\text{NAD}^+$  čepičky u některých RNA více než zdvojnásobil (Bird et al., 2016).



**Obr. 11. Efekt 5' čepičky na stabilitu RNA I *in vivo***

Doba života RNA I s čepičkou na 5' konci (NCIN-RNA) a bez ní (RNA) v exponenciální fázi (vlevo) a ve stacionární fázi (vpravo) růstové křivky bakteriální kultury (převzato a upraveno podle Bird et al., 2016).

Existují i další hypotézy, které však zatím postrádají jakýkoliv důkaz. Vzhledem k tomu, že  $\text{NAD}^+$  je koenzymem mnoha enzymů působících v oxidačně-redukčních dějích, mohlo by zde zastávat podobnou úlohu. RNA by mohla zastupovat funkci transportéru, který by na základě své sekvence dopravil  $\text{NAD}^+$  na místo potřeby, případně by rekrutovala různé proteiny. Také by naopak onu funkci transportéru/rekrutace proteinu mohlo plnit  $\text{NAD}^+$ , protože některé proteiny mají vysokou afinitu pro  $\text{NAD}^+$ . Navázaná RNA by pak mohla mít regulační funkci – například by snižovala/zvyšovala aktivitu daného proteinu, nebo by mohla fungovat opět jako transportér (Jäschke et al., 2016).

### 3.3 Koenzym A na 5' konci RNA

Podobně jako  $\text{NAD}^+$ , koenzym A (CoA) byl objeven díky využití kombinace metod vysokotlaké kapalinové chromatografie a hmotnostní spektrometrie. Nicméně cesta od objevení neznámého RNA konjugátu po zjištění, že se jedná o CoA, byla daleko komplikovanější

a vlastně předcházela metodu, pomocí níž byl identifikován  $\text{NAD}^+$  a další molekuly asociované s RNA.

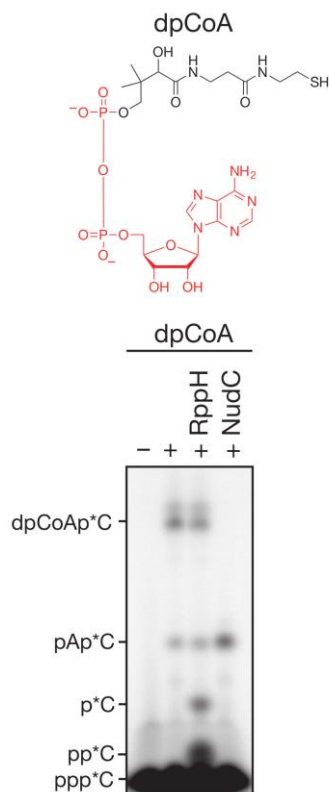
Na rozdíl od metody použité při identifikaci  $\text{NAD}^+$  (popsané výše), vzorek celkové RNA, pocházející též z bakterií *E. coli*, případně *S. venezuelae*, nebyl nejprve vystaven působení nukleáz. Místo toho byl vystaven takovým podmínkám, při kterých byly od hlavního vlákna RNA odštěpeny malé molekuly citlivé k těmto podmínkám. Konkrétně se jednalo buď o slabou bázi (pH 8.0), nebo n-butylamin v acetonitrilu, který jako nukleofil napadá a štěpí některé vazby. Frakce větších molekul (tedy hlavně vlákna RNA) byla poté odstraněna za pomoci gelové chromatografie a frakce menších molekul byla podrobena LC/MS analýze (Kowtoniuk et al., 2009).

Takto bylo objeveno celkem 5 neznámých molekul, které byly potenciálními modifikacemi RNA. Kontrola provedená se synteticky vytvořenou RNA potvrdila, že se skutečně jedná o modifikující molekuly, nikoliv o artefakty způsobené vystavením kanonických bází daným podmínkám. Nejmenší z takto nově objevených molekul měla poměr hmotnosti k náboji  $m/z = 101,0232$ . Na základě této hmotnosti, které odpovídá sumárním vzorcem  $\text{C}_4\text{H}_4\text{O}_3$ , a dalších provedených experimentů byla molekula identifikována jako anhydrid kyseliny jantarové. Takováto molekula však může být pouze fragmentem nějaké větší molekuly, která byla rozštěpena působením použitých podmínek. Bylo tedy potřeba vyvinout novou „neinvazivní“ metodu, která by nevytvářela podobné artefakty. Tomu vyhovuje výše popsání metoda štěpení RNA nukleázami na jednotlivé nukleotidy, které jsou dále podrobené analýze pomocí LS/MS. Nově objevených molekul bylo po aplikaci tohoto nového přístupu ještě daleko více, a podle očekávání mezi nimi nebyla žádná s  $m/z = 101,0232$ . Aby bylo možné identifikovat molekulu, které patří fragment s  $m/z = 101,0232$ , byla hledána taková molekula, která by po aplikaci obou metod současně měla poměr hmotnosti a náboje o toto číslo nižší než po vystavení vzorku pouze nukleázám. Taková molekula se ve vzorku skutečně vyskytovala a podle hmotnosti  $m/z = 787,166$  ( $m/z = 686,1209$  v případě vystavení oběma podmínkám) byla identifikována (obdobným postupem jako v případě  $\text{NAD}^+$ ) jako 3'-defosfosuccinyl-CoA, případně 3'-defosfo-CoA (Kowtoniuk et al., 2009).

Pro ověření, že může RNAP iniciovat transkripci dpCoA, podobně jako nikotinamidadenin nukleotidem, byla provedena *in vitro* transkripce s RNAP z *E. coli*, promotorem pro RNA I (podobně jako u  $\text{NAD}^+$  je základem CoA ADP, takže +1 nukleotid promotoru musí být adenin), dpCoA jako iniciačním nukleotidem a cytidintrifosfátem radioaktivně značeným na  $\alpha$  uhlíku. Z výsledků je patrné (Obr. 12), že RNAP je zodpovědná i za dpCoA na 5' konci RNA (Bird et al., 2016).



Výsledný transkript byl vystaven účinkům RppH a NudC. Výsledky (Obr. 12) byly obdobné jako v případě  $\text{NAD}^+$  a vzhledem k tomu, že CoA zastává v metabolismu podobnou funkci (je taktéž významným koenzymem v energetickém metabolismu), je možné předpokládat, že bude mít i podobný význam jako čepička na 5' konci RNA.



**Obr. 12. Transkripce in vitro s dpCoA a [ $\alpha^{32}\text{P}$ ]-CTP**

Transkripce provedená s dpCoA jako iniciačním nukleotidem a cytidintrifosfátem radioaktivně značeným na  $\alpha$  uhlíku jako prodlužujícím nukleotidem. (-)kontrola bez dpCoA, (+) transkripce proběhla s využitím dpCoA jako iniciačního nukleotidu, (+RppH) přidání RNA pyrofosfohydrolázy nemělo na fragment RNA s dpCoA na začátku žádný vliv, (+NudC) přidání enzymu NudC vedlo k rozštěpení čepičky tak, že na 5' konci zůstal adenosinmonofosfát (převzato a upraveno podle Bird et al., 2016).



## 4. Závěr

*Escherichia coli* je intenzivně zkoumaným živým organismem a její kompletně osekvenovaný genom je k dispozici již téměř 20 let (Blattner et al., 1997). Stále nové objevy nám přesto dokazují, že o fungování ani tohoto relativně jednoduchého organismu nevíme zdaleka všechno. Regulace genové exprese je klíčová pro přežití *E. coli* (a vlastně všech organismů) v různých podmínkách. Její podrobnější pochopení by mohlo pomoci porozumění dalším molekulárním mechanismům a tím nás o krůček přiblížit úplnému pochopení fungování této bakterie.

*In vitro* a následně i *in vivo* experimenty prokázaly, že NAD<sup>+</sup> (případně dpCoA) čepička je schopná ovlivnit stabilitu RNA a tím případně regulovat genovou expresi. Nyní je třeba pomocí *in vivo* experimentů prokázat, že bakterie tuto možnost regulace skutečně využívají a že se nejedná pouze o „omyl“ při iniciaci transkripce způsobený nespecifitou bakteriální RNAP, případně že NAD<sup>+</sup> na 5' konci RNA nezastává úplně jinou funkci (diskutováno výše).

Budoucí výzkum se zaměří na pochopení promotorových sekvenčních pravidel pro inkorporaci NAD<sup>+</sup>, studium biologické role 5' modifikací RNA a na průzkum dalších organismů (včetně patogenů) z hlediska těchto modifikací.

## ***Seznam citované literatury***

Bird, J. G., Y. Zhang, Y. Tian, N. Panova, I. Barvík, L. Greene, M. Liu, B. Buckley, L. Krásný and J. K. Lee (2016). "The mechanism of RNA 5' capping with NAD<sup>+</sup>, NADH and desphospho-CoA." Nature.

Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode and G. F. Mayhew (1997). "The complete genome sequence of *Escherichia coli* K-12." science **277**(5331): 1453-1462.

Borukhov, S., O. Laptenko and J. Lee (2001). "[7]-*Escherichia coli* Transcript Cleavage Factors GreA and GreB: Functions and Mechanisms of Action." Methods in enzymology **342**: 64-76.

Brennan, C. A., A. J. Dombroski and T. Platt (1987). "Transcription termination factor rho is an RNA-DNA helicase." Cell **48**(6): 945-952.

Burgess, R. R. (1969). "Separation and characterization of the subunits of ribonucleic acid polymerase." Journal of Biological Chemistry **244**(22): 6168-6176.

Burr, T., J. Mitchell, A. Kolb, S. Minchin and S. Busby (2000). "DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study." Nucleic acids research **28**(9): 1864-1870.

Cahová, H., M.-L. Winz, K. Höfer, G. Nübel and A. Jäschke (2015). "NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs." Nature.

Cantara, W. A., P. F. Crain, J. Rozenski, J. A. McCloskey, K. A. Harris, X. Zhang, F. A. Vendeix, D. Fabris and P. F. Agris (2010). "The RNA modification database, RNAMDB: 2011 update." Nucleic acids research: gkq1028.

Carafa, Y. d. A., E. Brody and C. Thermes (1990). "Prediction of rho-independent *Escherichia coli* transcription terminators: A statistical analysis of their RNA stem-loop structures." Journal of molecular biology **216**(4): 835-858.

Crick, F. (1970). "Central dogma of molecular biology." Nature **227**(5258): 561-563.

Deana, A., H. Celesnik and J. G. Belasco (2008). "The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal." Nature **451**(7176): 355-358.

Dominissini, D., S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, S. Osenberg, K. Cesarkas, J. Jacob-Hirsch, N. Amariglio and M. Kupiec (2012). "Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq." Nature **485**(7397): 201-206.

Feklistov, A., B. D. Sharon, S. A. Darst and C. A. Gross (2014). "Bacterial sigma factors: a historical, structural, and genomic perspective." Annual review of microbiology **68**: 357-376.

Frick, D. N. and M. J. Bessman (1995). "Cloning, Purification, and Properties of a Novel NADH Pyrophosphatase." Journal of Biological Chemistry **270**(4): 1529-1534.

Gaal, T., M. Bartlett, W. Ross, C. Turnbough Jr and R. Gourse (1997). "NTP concentration as a regulator of transcription initiation: control of rRNA synthesis in bacteria." science **278**: 2092-2097.

Ghosh, P., A. Ishihama and D. Chatterji (2001). "*Escherichia coli* RNA polymerase subunit  $\omega$  and its N-terminal domain bind full-length  $\beta'$  to facilitate incorporation into the  $\alpha 2\beta$  subassembly." European Journal of Biochemistry **268**(17): 4621-4627.

Gildehaus, N., T. Neußer, R. Wurm and R. Wagner (2007). "Studies on the function of the riboregulator 6S RNA from *E. coli*: RNA polymerase binding, inhibition of *in vitro* transcription and synthesis of RNA-directed *de novo* transcripts." Nucleic acids research **35**(6): 1885-1896.

Goldman, S. R., R. H. Ebright and B. E. Nickels (2009). "Direct detection of abortive RNA transcripts *in vivo*." science **324**(5929): 927-928.

Grossman, A. D., J. W. Erickson and C. A. Gross (1984). "The *htpR* gene product of *E. coli* is a sigma factor for heat-shock promoters." Cell **38**(2): 383-390.

Haldenwang, W. G., N. Lang and R. Losick (1981). "A sporulation-induced sigma-like regulatory protein from *B. subtilis*." Cell **23**(2): 615-624.

Hambraeus, G., C. von Wachenfeldt and L. Hederstedt (2003). "Genome-wide survey of mRNA half-lives in *Bacillus subtilis* identifies extremely stable mRNAs." Molecular genetics and genomics **269**(5): 706-714.

Harley, C. B. and R. P. Reynolds (1987). "Analysis of *E. coli* promoter sequences." Nucleic acids research **15**(5): 2343-2361.

Hawley, D. K. and W. R. McClure (1983). "Compilation and analysis of *Escherichia coli* promoter DNA sequences." Nucleic acids research **11**(8): 2237-2255.

Höfer, K., S. Li, F. Abele, J. Frindert, J. Schlotthauer, J. Grawenhoff, J. Du, D. J. Patel and A. Jäschke (2016). "Structure and function of the bacterial decapping enzyme NudC." Nature chemical biology.

Hook-Barnard, I. G. and D. M. Hinton (2007). "Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters." Gene regulation and systems biology **1**: 275.

Hurwitz, J. (2005). "The discovery of RNA polymerase." Journal of Biological Chemistry **280**(52): 42477-42485.

Hussain, S., A. A. Sajini, S. Blanco, S. Dietmann, P. Lombard, Y. Sugimoto, M. Paramor, J. G. Gleeson, D. T. Odom and J. Ule (2013). "NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs." Cell reports **4**(2): 255-261.

Charette, M. and M. W. Gray (2000). "Pseudouridine in RNA: what, where, how, and why." IUBMB life **49**(5): 341-351.

- Chen, Y. G., W. E. Kowtoniuk, I. Agarwal, Y. Shen and D. R. Liu (2009). "LC/MS analysis of cellular RNA reveals NAD-linked RNA." Nature chemical biology **5**(12): 879-881.
- Jäschke, A., K. Höfer, G. Nübel and J. Frindert (2016). "Cap-like structures in bacterial RNA and epitranscriptomic modification." Current opinion in microbiology **30**: 44-49.
- Kapanidis, A. N., E. Margeat, S. O. Ho, E. Kortkhonjia, S. Weiss and R. H. Ebright (2006). "Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism." science **314**(5802): 1144-1147.
- Keilty, S. and M. Rosenberg (1987). "Constitutive function of a positively regulated promoter reveals new sequences essential for activity." Journal of Biological Chemistry **262**(13): 6389-6395.
- Keller, A. N., X. Yang, J. Wiedermannová, O. Delumeau, L. Krásný and P. J. Lewis (2014). "ε, a new subunit of RNA polymerase found in Gram-positive bacteria." Journal of bacteriology **196**(20): 3622-3632.
- Khoddami, V. and B. R. Cairns (2013). "Identification of direct targets and modified bases of RNA cytosine methyltransferases." Nature biotechnology **31**(5): 458-464.
- Komissarova, N., J. Becker, S. Solter, M. Kireeva and M. Kashlev (2002). "Shortening of RNA: DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination." Molecular cell **10**(5): 1151-1162.
- Koslover, D. J., F. M. Fazal, R. A. Mooney, R. Landick and S. M. Block (2012). "Binding and translocation of termination factor rho studied at the single-molecule level." Journal of molecular biology **423**(5): 664-676.
- Kowtoniuk, W. E., Y. Shen, J. M. Heemstra, I. Agarwal and D. R. Liu (2009). "A chemical screen for biological small molecule–RNA conjugates reveals CoA-linked RNA." Proceedings of the National Academy of Sciences **106**(19): 7768-7773.
- Lacatena, R. and G. Cesareni (1981). "Base pairing of RNA I with its complementary sequence in the primer precursor inhibits ColE1 replication." Nature **294**: 623-626.
- Laishram, R. S. (2014). "Poly (A) polymerase (PAP) diversity in gene expression–Star-PAP vs canonical PAP." FEBS letters **588**(14): 2185-2197.
- Lane, B. G. (1998). "Historical perspectives on RNA nucleoside modifications." Modification and editing of RNA. ASM Press, Washington, DC: 1-20.
- Luciano, D. J. and J. G. Belasco (2015). "NAD in RNA: unconventional headgear." Trends in biochemical sciences **40**(5): 245-247.
- Mackie, G. A. (1998). "Ribonuclease E is a 5'-end-dependent endonuclease." Nature **395**(6703): 720-724.

- Machnicka, M. A., K. Milanowska, O. O. Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz and K. M. Rother (2013). "MODOMICS: a database of RNA modification pathways—2013 update." Nucleic acids research **41**(D1): D262-D267.
- Marbaniang, C. N. and J. Vogel (2016). "Emerging roles of RNA modifications in bacteria." Current opinion in microbiology **30**: 50-57.
- McClure, W. R. (1985). "Mechanism and control of transcription initiation in prokaryotes." Annual review of biochemistry **54**(1): 171-204.
- Mohanty, B. K. and S. R. Kushner (1999). "Analysis of the function of *Escherichia coli* poly (A) polymerase I in RNA metabolism." Molecular microbiology **34**(5): 1094-1108.
- Mohanty, B. K. and S. R. Kushner (2002). "Polyadenylation of *Escherichia coli* transcripts plays an integral role in regulating intracellular levels of polynucleotide phosphorylase and RNase E." Molecular microbiology **45**(5): 1315-1324.
- Mohanty, B. K. and S. R. Kushner (2011). "Bacterial/archaeal/organelle polyadenylation." Wiley Interdisciplinary Reviews: RNA **2**(2): 256-276.
- Murakami, K. S. (2015). "Structural biology of bacterial RNA polymerase." Biomolecules **5**(2): 848-864.
- Nakazato, H., S. Venkatesan and M. Edmonds (1975). "Polyadenylic acid sequences in *E. coli* messenger RNA." Nature **256**(5513): 144-146.
- Nickels, B. E., S. J. Garrity, V. Mekler, L. Minakhin, K. Severinov, R. H. Ebright and A. Hochschild (2005). "The interaction between  $\sigma 70$  and the  $\beta$ -flap of *Escherichia coli* RNA polymerase inhibits extension of nascent RNA during early elongation." Proceedings of the National Academy of Sciences of the United States of America **102**(12): 4488-4493.
- Paget, M. S. (2015). "Bacterial sigma factors and anti-sigma factors: structure, function and distribution." Biomolecules **5**(3): 1245-1265.
- Peters, J. M., R. A. Mooney, P. F. Kuan, J. L. Rowland, S. Keleş and R. Landick (2009). "Rho directs widespread termination of intragenic and stable RNA transcription." Proceedings of the National Academy of Sciences **106**(36): 15406-15411.
- Pomerantz, S. C. and J. A. McCloskey (1990). "[44] Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry." Methods in enzymology **193**: 796-824.
- Preugschat, F., G. H. Tomberlin and D. J. Porter (2008). "The base exchange reaction of NAD<sup>+</sup> glycohydrolase: Identification of novel heterocyclic alternative substrates." Archives of biochemistry and biophysics **479**(2): 114-120.
- Rabatinová, A., H. Šanderová, J. J. Matějčková, J. Korelusová, L. Sojka, I. Barvík, V. Papoušková, V. Sklenář, L. Židek and L. Krásný (2013). "The  $\delta$  subunit of RNA polymerase is required for rapid changes in gene expression and competitive fitness of the cell." Journal of bacteriology **195**(11): 2603-2611.

Revyakin, A., C. Liu, R. H. Ebright and T. R. Strick (2006). "Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching." science **314**(5802): 1139-1143.

Richard, R., A. Andrew, J. J. Dunn and E. K. Bautz (1969). "Factor stimulating transcription by RNA polymerase." Nature **221**: 43.

Ruff, E. F., M. T. Record and I. Artsimovitch (2015). "Initial events in bacterial transcription initiation." Biomolecules **5**(2): 1035-1062.

Santangelo, T. J. and I. Artsimovitch (2011). "Termination and antitermination: RNA polymerase runs a stop sign." Nature Reviews Microbiology **9**(5): 319-329.

Shimotohno, K., Y. Kodama, J. Hashimoto and K.-I. Miura (1977). "Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis." Proceedings of the National Academy of Sciences **74**(7): 2734-2738.

Schaefer, M., T. Pollex, K. Hanna and F. Lyko (2009). "RNA cytosine methylation analysis by bisulfite sequencing." Nucleic acids research **37**(2): e12-e12.

Schwartz, S., D. A. Bernstein, M. R. Mumbach, M. Jovanovic, R. H. Herbst, B. X. León-Ricardo, J. M. Engreitz, M. Guttman, R. Satija and E. S. Lander (2014). "Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA." Cell **159**(1): 148-162.

Skordalakes, E. and J. M. Berger (2003). "Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading." Cell **114**(1): 135-146.

Song, M.-G., S. Bail and M. Kiledjian (2013). "Multiple Nudix family proteins possess mRNA decapping activity." Rna **19**(3): 390-399.

Sorenson, M. K. and S. A. Darst (2006). "Disulfide cross-linking indicates that FlgM-bound and free  $\sigma$ 28 adopt similar conformations." Proceedings of the National Academy of Sciences **103**(45): 16722-16727.

Topisirovic, I., Y. V. Svitkin, N. Sonenberg and A. J. Shatkin (2011). "Cap and cap-binding proteins in the control of gene expression." Wiley Interdisciplinary Reviews: RNA **2**(2): 277-298.

Tornøe, C. W., C. Christensen and M. Meldal (2002). "Peptidotriazoles on solid phase:[1, 2, 3]-triazoles by regiospecific copper (I)-catalyzed 1, 3-dipolar cycloadditions of terminal alkynes to azides." The Journal of organic chemistry **67**(9): 3057-3064.

Travers, A. A. and R. R. Burgess (1969). "Cyclic re-use of the RNA polymerase sigma factor." Nature **222**: 537-540.

Vassilyev, D. G. (2009). "Elongation by RNA polymerase: a race through roadblocks." Current opinion in structural biology **19**(6): 691-700.

Vvedenskaya, I. O., J. S. Sharp, S. R. Goldman, P. N. Kanabar, J. Livny, S. L. Dove and B. E. Nickels (2012). "Growth phase-dependent control of transcription start site selection and gene expression by nanoRNAs." Genes & development **26**(13): 1498-1507.

Washburn, R. S. and M. E. Gottesman (2015). "Regulation of transcription elongation and termination." Biomolecules **5**(2): 1063-1078.

Winkelman, J. T., B. T. Winkelman, J. Boyce, M. F. Maloney, A. Y. Chen, W. Ross and R. L. Gourse (2015). "Crosslink mapping at amino acid-base resolution reveals the path of scrunched DNA in initial transcribing complexes." Molecular cell **59**(5): 768-780.

Yang, X. and P. J. Lewis (2010). "The interaction between bacterial transcription factors and RNA polymerase during the transition from initiation to elongation." Transcription **1**(2): 66-69.

Yue, Y., J. Liu and C. He (2015). "RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation." Genes & development **29**(13): 1343-1355.