

**PŘÍRODOVĚDECKÁ FAKULTA UNIVERZITY KARLOVY V PRAZE
KATEDRA GENETIKY A MIKROBIOLOGIE**

BAKALÁŘSKÁ PRÁCE

Databáze experimentálně doložených sekundárních struktur
v 5' nepřekládaných oblastech virových RNA

Markéta Pastuchová

Školitel: RNDr. Martin Pospíšek, Ph.D.
2006 / 2007

Abstrakt:

Tato práce se věnuje experimentálně doloženým sekundárním strukturám v 5' nepřekládaných oblastech +RNA virů a retrovirů. Zabývá se hlavními rysy struktury RNA a jejím srovnáním se strukturou DNA. Popisuje metody určování sekundární struktury se zaměřením na enzymatické a chemické mapování. Krátce shrnuje význam sekundárních struktur v 5' koncových oblastech. Hlavní část je věnována popisu sekundárních struktur v 5' koncových nepřekládaných oblastech konkrétních virů z čeledí: *Flaviviridae*, *Picornaviridae*, *Coronaviridae*, *Arteriviridae*, *Dicistroviridae*, *Potyviridae*, *Potexviridae* a *Retroviridae*.

Klíčová slova: databáze experimentálně doložených struktur, 5' UTR, sekundární struktury RNA, +RNA viry, mapování RNA

Abstract:

This work is devoted to experimentally documented secondary structures in the 5' untranslated region of +RNA viruses and retroviruses. It is focused on main characteristic attributes of the RNA structure and its comparison with the DNA structure. It describes methods of the RNA secondary structure prediction with specialization to enzymatic and chemical mapping. It briefly summarizes meaning of secondary structure elements in the 5' ending regions. Main part is aimed at the description of secondary structure elements in the 5' untranslated regions of specific viruses from families: *Flaviviridae*, *Picornaviridae*, *Coronaviridae*, *Arteriviridae*, *Dicistroviridae*, *Potyviridae*, *Potexviridae* and *Retroviridae*.

Keywords: 5' UTR, RNA secondary structure elements, +RNA viruses, RNA mapping

Seznam zkratek:

DNA	deoxyribonukleová kyselina
nt	nukleotid
A	adenin
C	cytosin
G	guanin
T	thymín
U	uracil
IRES	vnitřní vazebné místo pro ribozóm
IGR IRES	intergenové vnitřní vazebné místo pro ribozóm
ORF	otevřený čtecí rámec
PCBP	poly(C) vazebný proteinu
PTB	protein vázající se na pyrimidin bohaté oblasti
RNA	ribonukleová kyselina
gRNA	genomová RNA
sgRNA	subgenomová RNA
mRNA	mediátorová RNA
DI RNA	defektní interferující RNA
ssDNA/ssRNA	jednořetězcová DNA/jednořetězcová RNA
dsDNA/dsRNA	dvouřetězcová DNA/dvouřetězcová RNA
UTR	nepřekládaná oblast
BMH	konformace s rozvětvenými vlásenkami
DIS	dimerizační místo
LDI	konformace s interakcemi na dlouhé vzdálenosti
PBS	primer vazebné místo
PSI	balící signál ψ
SD	sestřihové místo
TAR	transkripci aktivující oblast
U5/U3	unikátní sekvence 5/unikátní sekvence 3 (retrovirů)
TRS	transkripci regulující sekvence
LTH	vedoucí TRS vlásenka

Obsah

1 Úvod	6
2 Struktura RNA	7
3 Metody určování sekundární struktury	9
3.1 Počítačové predikce	9
3.2 Enzymatické a chemické mapování sekundárních struktur	10
3.3 Gelová elektroforéza	12
3.4 Komparativní analýza	13
3.5 Krystalografie	13
3.6 NMR spektroskopie	14
4 +RNA viry	14
4.1 <i>Flaviviridae</i>	15
4.2 <i>Picornaviridae</i>	18
4.3 <i>Nidovirales</i>	20
4.3.1 <i>Coronaviridae</i>	20
4.3.2 <i>Arteriviridae</i>	20
4.4 <i>Dicistroviridae</i>	21
4.5 Rostlinné +RNA viry	23
5 <i>Retroviridae</i>	24
6 Význam 5' koncových sekundárních struktur +RNA virů	29
7 Diskuze	31
8 Závěr	32
9 Seznam citované literatury	33
10 Příloha	39

1 Úvod

Svou bakalářskou prací budu přispívat ke vzniku nové databáze sekundární struktury mediátorové RNA (mRNA), která bude obsahovat jen experimentálně ověřené struktury. Jedině kombinací počítačové predikce, enzymatickým a případně i chemickým mapováním a komparativní analýzou lze vytvořit opravdu spolehlivý model sekundární struktury. Proto studie, které jsem vybírala do své bakalářské práce, jsou založeny současně na všech třech těchto hlavních metodách.

Existuje již několik databází zabývajících se strukturami nukleových kyselin. Například UTRdb (Pesole *et al.*, 2000) je specializovaná databáze sekvencí v 5' a 3' nepřekládaných oblastech eukaryotické mRNA. UTRdb navíc obsahuje tři specializované sekce UTRef, UTRait a UTRexp, ta obsahuje struktury, u kterých byla experimentálně prokázána nějaká funkční role (Mignone *et al.*, 2005). Další databází je IRESdb, která zahrnuje primární strukturu eukaryotických i virových vnitřních ribozóm vazebných míst (Bonnal *et al.*, 2003).

Má práce je zaměřená na sekundární struktury nacházející se na 5' konci RNA. Ke zpracování těchto struktur jsem si vybrala +RNA viry, protože jsou unikátní v tom, že jejich RNA slouží zároveň jako genomová RNA, nese veškerou informaci pro reprodukci viru, a také jako mediátorová RNA kódující proteiny. +RNA je při vstupu do buňky sama o sobě infekční a virus si ve své virové partikuli nenese žádnou vlastní RNA polymerázu. Tím se liší od -RNA virů, které se také množí replikací RNA, ale virovou polymerázu si musí nést v infekční partikuli s sebou. Před translací proteinů kódovaných -RNA viry musí dojít k přepisu -RNA do komplementární + formy.

Pro srovnání s „obyčejnými +RNA viry“ jsem do své bakalářské práce zařadila také retroviry. Ve své virové partikuli si sice také nesou +RNA verzi svého genomu, ale jejich replikační strategie je naprosto odlišná. Jsou schopné reverzní transkripce do DNA a integrace do hostitelského genomu. Ve své virové partikuli si nesou kromě +RNA i virovou reverzní transkriptázu a specifickou tRNA, které slouží jako primer pro syntézu minus řetězce DNA.

Další nedílnou součástí mé bakalářské práce jsou kapitoly zabývající se strukturou RNA a jejím srovnáním se strukturou DNA, metodami mapování sekundární struktury a kapitola, v níž popisují význam sekundárních struktur nacházejících se na 5' konci mRNA.

Práce je doplněná přílohou, která obsahuje schematický přepis sekundárních struktur na 5' koncích + RNA virů a zápis všech experimentálních podmínek a činidel, které byly v jednotlivých studiích uvedeny.

2 Struktura RNA

Molekula RNA se skládá z polynukleotidového vlákna složeného ze čtyř typů nukleotidů (nt), ale na rozdíl od DNA jsou vlákna neboli řetězce RNA mnohem kratší. Primární struktura je dána pořadím neboli sekvencí nt. Nukleotidy RNA jsou tvořeny pětiuhlíkatým cukrem – ribosou a fosfátovou skupinou, dohromady tvoří cukr-fosfátovou páteř, a jednou ze čtyř bází: adeninem (A), cytosinem (C), guaninem (G) či uracilem (U), který zastupuje thymin (T) přítomný v DNA. V RNA mohou být zastoupeny také vzácné méně obvyklé báze jako jsou alkylované A, C, G, U nebo 7-methylguanosa, který tvoří čepičkovou (angl. „cap“) strukturu na 5' konci většiny mRNA, dihydrouridin, inosin a další (Saenger, 1984).

Tak jako u DNA i u RNA může docházet k párování bází (pár bází, pb) a RNA může nabývat dvouřetězcové formy (angl. „double stranded“, ds). Kromě klasického Watson-Crickovského párování bází u ssRNA nacházíme také kolísavé párování bází G-U (Varani a McClain, 2000) A-C a další nekanonické typy párování jako jsou třeba G-A nebo A-U páry (Nagaswamy *et al.*, 2000).

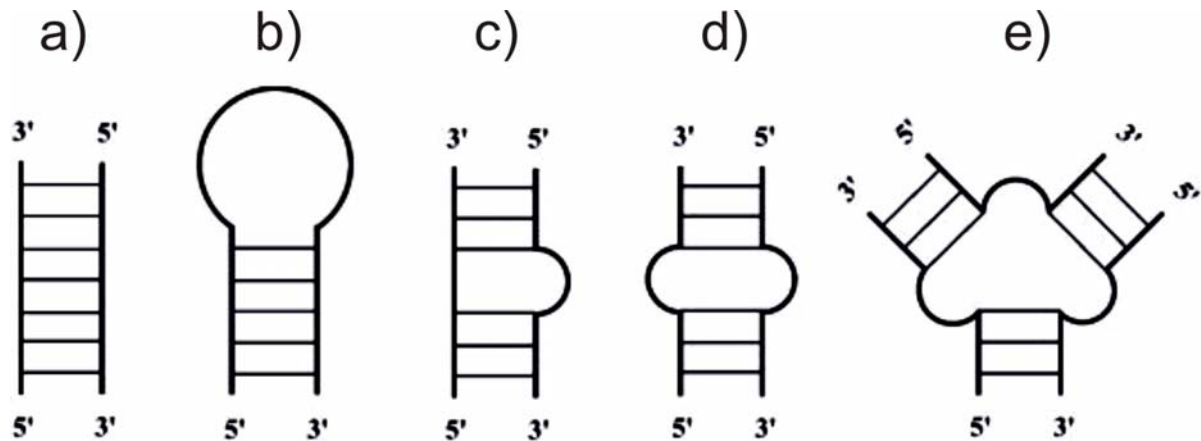
Molekula RNA může být tvořena dlouhými ds oblastmi, které jsou spojeny krátkými jednořetězcovými úseky (angl. „single strand“, ss), nebo naopak dlouhými ss oblastmi s krátkými ds úseky. Většina biologicky aktivních RNA jako je tRNA, rRNA a snRNA obsahuje množství ds oblastí. RNA může zajišťovat kromě funkce přenosné také funkci katalytickou, strukturní a regulační: RNA-RNA, RNA-DNA a RNA-protein interakce hrají roli v mnoha biologických pochodech (Saenger, 1984).

Narozdíl od DNA, jejíž dvouřetězcová struktura může nabývat mnoha forem, je dvouřetězcová struktura RNA díky přítomnosti ribosy místo deoxyribosy omezena jen na dvě blízké příbuzné pravotočivé formy dvoušroubovice A a A'.¹ Stejně jako u dsDNA jsou i u dsRNA vlákna antiparalelní. A forma obsahuje 11 pb na jeden závit a výška závitu je 2,85 nm. A' forma obsahuje 12 pb na závit, který je vysoký 3,6 nm. Obě A formy se vyznačují hlubokým velkým žlábkem (angl. „major groove“) a velmi plytkým malým žlábkem (angl. „minor groove“) (Saenger, 1984).

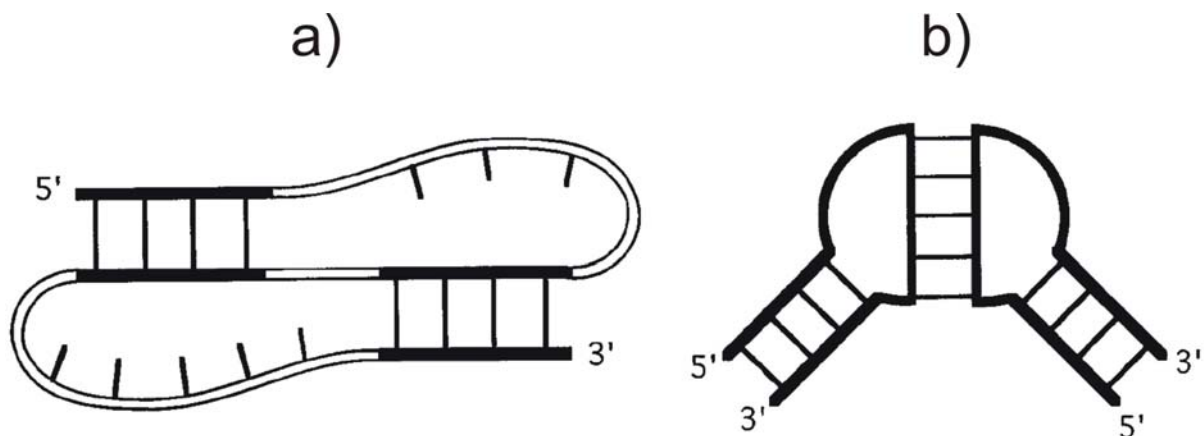
Sekundární struktura RNA je tvořena výše zmíněnými dvoušroubovicemi a množstvím dalších strukturních motivů jako jsou vlásenky (angl. „hairpin loop“) skládající

¹ Když nebereme v úvahu homopolymerní vlákna a nekanonické typy pb, které jsou schopny tvořit i jiné formy helixů, jako je (angl. „triple helix“)(Kankia, 2003).

se ze stonku (angl. „stem“) a smyčky (angl. „loop“), výdutě neboli boule (angl. „bulge“), smyčky vnitřní (angl. „interior loop“) a rozvětvené (angl. „multi – branched loop“) nebo pseudouzly (angl. „pseudoknot“)(Alberts *et al.*, 2004; Ding a Lawrence, 2003). Pseudouzly se ovšem stejně jako tzv. „kissing loop“ ve většině pramenů řadí již mezi prvky struktury terciální, kterou tvoří konečná trojrozměrná struktura molekuly (Weixlbaumer *et al.*, 2004; <http://wikipedia.org/wiki/Pseudoknot>).



Obrázek 1: Zjednodušené zobrazení elementů sekundární struktury: a) dvoušroubovice, b) vlásenka, c) výdutě neboli boule, d) vnitřní smyčka, e) rozvětvená smyčka
Převzato a upraveno z Ding a Lawrence, 2003



Obrázek 2: Schématické zobrazení elementů terciární struktury: a) pseudouzlel, b) „kissing loop“
Převzato a upraveno z Nowakowski a Tinoco, 1997

Struktura RNA je také ovlivněna složením iontového roztoku, ve kterém se nachází, například přítomnost hořečnatých iontů a polyaminů stabilizuje terciální strukturu molekuly (Knapp, 1989).

Pozitivně nabité ionty jako jsou kromě Mg^{2+} také Cu^{2+} , Ni^{2+} , Zn^{2+} , Co^{2+} , Ca^{2+} , Sr^{2+} , Ba^{2+} , Cd^{2+} , Fe^{3+} a další (Tan *et al.*, 2007; Duguid *et al.*, 1995) mají velký efekt na sekundární strukturu, nicméně pozitivní divalentní ionty mnohem více ovlivňují terciální strukturu. Navázání kovových iontů na dusíkaté báze (častěji ionty přechodných prvků) či na cukrfosfátovou kostru (častěji ionty kovů alkalických zemin) redukuje elektrostatickou repulzi fosfátových skupin sekundárních elementů, které se spolu dostávají do interakce v terciální struktuře. Bivalentní ionty můžeme rozdělit na dva druhy, oba se vážou na předem připravená místa v RNA jako je například G-U pár ve velkém žlábků helixu. První druh významně nemění strukturu těchto iont vazebných míst, druhý druh je součástí terciální struktury, bez kovového iontu by iont vazebné místo či terciální struktura jako jsou pseudoknoty či „kissing loop“ nemohly vzniknout. Dalším důležitým způsobem jakým ionty ovlivňují nukleové kyseliny je změna pH. (Tinoco a Bustamante, 1999; Duguid *et al.*, 1995).

I přes to, že nukleové kyseliny mají negativní náboj, se mohou anionty jako je SO_4^{2-} a Cl^- dostat do sousedství pozitivně nabitých skupin amino, imino nebo hydroxylových skupin adeninu, cytosinu a guaninu a tvořit s nimi přímé kontakty. Na thyminu nebo uracylu se anionty nevážou. Aniont vazebných míst je podstatně méně cca 13 než vazebných míst pro kladné ionty (Auffinger *et al.*, 2004).

3 Metody určování sekundární struktury

3.1 Počítačové predikce

Počítačové predikce vycházejí z primární struktury nukleových kyselin. Počítačové algoritmy mohou přesně namodelovat, jaké uspořádání sekundární struktury bude mít nejmenší volnou energii, problém je v tom, že molekuly často termodynamicky nejvýhodnější strukturu nezaujímají. Navíc se pro zjednodušení programů většinou pracuje s předpoklady, které nejsou zcela přesné: třeba u patrových interakcí se předpokládá interakce s nejbližším sousedem, nepočítá se s možností tvorby pseudouzlů atd.² Tato zjednodušení, ačkoli jsou nutná pro minimalizaci ukládacího místa a času, které program vyžaduje, samozřejmě ovlivňují přesnost výsledného modelu.

² Samozřejmě existují i algoritmy, které dokáží predikovat i pseudouzly (Rivas a Eddy, 1999)

(http://www.comp.nus.edu.sg/~ksung/cs5238/project/final/rnafold_YeNan.pdf;
<http://cmgm.stanford.edu/biochem218/Projects%20Spring%202003/Silverman.pdf>).

Predikčním programem, se kterým se setkáme u drtivé většiny studií, jejichž součástí je mapování sekundární struktury, bude Mfold, s dalšími jako třeba s RNAfoldem či s Sfoldem se setkáme spíše zřídka.

Ve výsledku tedy počítačové predikce vytváří primární představu o sekundární struktuře, která se dále ověřuje a upřesňuje, bohužel zdaleka ne ve všech případech, mnozí autoři na počítačovou predikci zcela spoléhají (Rijnbrand *et al.*, 1995), experimentálně pomocí enzymatického a/nebo chemického mapování sekundárních struktur a komparativní analýzou, která je také prováděna pomocí počítačových programů.

3.2 Enzymatické a chemické mapování sekundárních struktur

Experimentální mapování sekundárních struktur se uskutečňuje pomocí ribonukleáz neboli RNáz či chemických činidel, při obou způsobech je nutné dodržovat pufovací podmínky (pH, iontová síla, teplota, koncentrace hořčikových iontů), které jsou pro každé činidlo specifické. Přesné složení reakčních a skladovacích pufrů včetně vhodné teploty a pH pro jednotlivá činidla se liší také podle výrobců³.

RNázy katalyzují hydrolýzu fosfodiesterové vazby v RNA. Existuje mnoho druhů RNáz, zmíním jen několik endoribonukleáz, se kterými jsem se v odborných publikacích setkala nejčastěji:

- **RNáza V1** z kobřího (*Naja naja*) jedu, někdy je nazývána také **RNáza CV** (angl. „Cobra Venom“, CV) štěpí nespécificky v ds oblastech, může štěpit i v ss oblastech, které přechodně zaujímají helikální strukturu nebo patrově interagují, generuje fragmenty s 5' P, optimální pH je mezi 4-9, potřebuje nejméně 1mM koncentraci Mg²⁺
- **RNáza T1** z *Aspergillus orizae* štěpí na nespárovaném G, generuje fragmenty s 3' P, má dvě pH optima 4,5 a 7,5, vyžaduje koncentraci solí optimální pro fyziologické procesy

³ Tyto informace nalezneme www stránkách mnoha výrobců, zde uvádím jen několik příkladů:

Fermentas - <http://www.fermentas.com>

Sigma-Aldrich - <http://www.sigmaaldrich.com>

Amersham - <http://www.amersham.com>

Ambion - <http://www.ambion.com>

Boehringer Ingelheim - <http://www.boehringer-ingelheim.com>

United States Biochemical - <http://www.usbweb.com>

- **RNáza T2** také původem z *Aspergillus o.* štěpí na nespárovaném A, generuje fragmenty s 3' P, pH optimum je 4,5, je však stabilní i při neutrálním pH, je inhibována ionty těžkých kovů zejména Cu^{2+}
- **RNáza A** štěpí na nespárovaném C a U, generuje fragmenty s 3' P vyžaduje fyziologickou koncentraci solí
- **nukleáza S1** původem z *Aspergillus o.* štěpí nukleotidově nespécificky v ss oblastech, generuje fragmenty s 5' P, optimální pH je 4,5 může být použita i v neutrálním pH, ale ve vyšší koncentraci, je stimulována ionty Zn^{2+} (Stern *et al.*, 1988; Ehresmann *et al.*, 1987)

Chemické mapování se provádí pomocí nespécifické alkalické hydrolyzy, vzniklé fragmenty mají Gaussovo rozdělení velikostí. Další chemická činidla jsou:

- dimethyl sulfát (**DMS**) modifikuje nespárovaný C, A a G, jenom modifikace C a A zastavuje reverzní transkripci, abychom detekovali modifikaci G je nutné ošetření sodium borhydridem a anilinem, které indukuje přestřížení řetězce
- diethylpyrokarbonát (**DEPC**), někdy zkracován jenom jako **DEP**, modifikuje nespárovaný A, nemá vliv na nt, které jsou zahrnuty v patrových interakcích
- 1-cyclohexyl-3-(2-morphopolinoethyl)-carbodiimid metho-p-toluen sulfonát (**CMCT**) modifikuje nespárovaný U a G, vyžaduje bazické pH a fyziologickou koncentraci solí
- 2-keto-3-ethoxybutyraldehyd, nazývaný zkráceně **kethoxal**, modifikuje nespárované G, také potřebuje bazické pH a obdobnou koncentraci solí jako CMCT (Stern *et al.*, 1988; Ehresmann *et al.*, 1987)
- hydroxymethylpsoralen zkráceně **psoralen** se vmezeřuje do ds oblastí a při ozařování ultrafialovým světlem o vlnové délce 366 nm tvoří kovalentní spojení s pyrimidiny hlavně s U (Harrison *et al.*, 1992)

název	RNáza V1	RNáza T1	RNáza T2	RNáza A	nukleáza S1	DMS	DEPC	CMCT	psoralen
místa	ds	nesp.	nesp.	nesp.	ss	nesp.	nesp.	nesp.	ds
působení	oblasti	G	A	C a U	oblasti	C a A	A	U a G	oblasti

Chemická modifikace i přerušení řetězce pomocí RNáz se oboje projevuje zastavením reverzní transkripce, tedy zastavením prodlužování primeru (HuQu *et al.*, 1983). Při chemické modifikaci se prodlužování primeru zastaví na nt, který se bude nacházet na 3' straně od modifikované báze. Při enzymatickém přestřížení se prodlužování zastaví na posledním nt před přestřížením. (Kolupaeva *et al.*, 1996)

Fragmenty – produkty prodlužování primeru připravené vhodnou směsí RNáz jsou následně rozdělovány gelovou elektroforézou (viz dále).

3.3 Gelová elektroforéza

Gelová elektroforéza se používá k dělení fragmentů RNA (i DNA) podle velikosti, princip je založen na rozdílné velikosti náboje fragmentů různých délek. Směs fragmentů se nanese do žlábků na jednom konci agarózového nebo polyakrylamidového gelu (agarózový gel je vhodný pro fragmenty o velikosti 200-6000 bp, polyakrylamidový gel se používá pro menší fragmenty o velikosti 100-1000 bp) a gel s fragmenty je umístěn do elektrického pole. RNA se díky svému zápornému náboji pohybuje směrem ke kladné elektrodě, přičemž dlouhé fragmenty se pohybují kvůli své velikosti pomaleji. Protože pohyblivost RNA v gelu je kromě velikosti fragmentů RNA ovlivněna také jejich konformací, provádí se elektroforéza v přítomnosti denaturantů jako je třeba glyoxal, formamid nebo formaldehyd (Farrell, 1993).

RNA je třeba v gelu nějak zviditelnit, jednou z metod detekce je smíchání RNA s ethidium bromidem - EtBr, který po navázání na RNA floreskuje v ultrafialovém světle. Dále se pro barvení RNA se používá akridinová oranž nebo toluidinová modř. Jinou detekční metodou je radioaktivní značení nukleotidů ^{32}P pomocí T4 polynukleotid kinázy, navazuje se gamma ^{32}P ATP na 5' konec RNA. Můžeme použít také radioaktivní značení pomocí RNA ligázy, v tomto případě se na 3' konec RNA navazuje alfa ^{32}P ATP. Takto označené fragmenty lze detekovat pomocí autoradiografie, protože částice emitované z ^{32}P způsobují změny na fotografickém filmu, po vyvolání je poloha všech RNA fragmentů viditelná (Farrell, 1993).

Podle rozložení fragmentů na gelu se určují ss a ds oblasti RNA a zpřesňuje se model vytvořený počítačovou predikcí.

3.4 Komparativní analýza

Srovnávací neboli komparativní analýza je založena na primární struktuře RNA a na srovnání se sekvencemi příbuznými zkoumané RNA. Její princip je založen na faktu, že sekundární struktura nekódující RNA, která se podílí na nějaké funkci, je v evoluci konzervována, například fylogeneticky příbuzné viry budou mít funkčně homologickou RNA stejné sekundární struktury, i když jejich primární struktura bude odlišná. Aby toto mohlo být splněno, muselo v evoluci docházet ke kompenzačním mutacím – záměnám bází, které se neprojeví na změně sekundární struktury, jinými slovy kompenzační záměny bází udržují komplementaritu (Parsch *et al.*, 2000).

Jeden z komparativních přístupů je kovarianční model založený na konsensu primární i sekundární sekvence. Je schopný rozpoznat i velmi mírně konzervované sekvence. K predikci sekundární struktury využívá primární strukturu RNA – kovariantní nukleotidové substituce (Eddy a Durbin, 1994).

Někdy se při vytváření modelu sekundární struktury vychází kromě postupů zmíněných výše také z modelu struktury terciární, proto jsem chtěla stručně zmínit aspoň principy dvou nejdůležitějších metod určování terciární struktury.

3.5 Krystalografie

Krystalografie se zabývá analýzou prostorové struktury malých molekul, největší možné rozlišení je jeden Angstrém. Je založena na difrakci (rozptylu) rentgenového záření. Krystal studované látky je ozařován monochromatickým rentgenovým zářením a působí jako mřížka, která ohýbá paprsky záření, které dopadají na stínítko. Z takto vzniklých interferenčních obrazců, lze pak vypočítat trojrozměrné mapy elektronové hustoty krystalů. Krystalografie je časově náročná metoda kvůli pomalé tvorbě krystalů (<http://en.wikipedia.org/wiki/Crystallography>).

3.6 NMR spektroskopie

Zkratka NMR se používá pro jadernou magnetickou rezonanci (angl. „nuclear magnetic resonance“). Narozdíl od krystalografie její využití nezávisí na schopnosti látek krystalizovat, protože se provádí v roztoku umístěném v silném magnetickém poli. Nukleární magnetickou rezonanci lze vyvolat jen u jader s magnetickým momentem. Mezi ně patří jádra, která mají lichý počet protonů nebo nukleonů, často se používá jádro izotopu vodíku H^1 (Alberts *et al.*, 2002; <http://en.wikipedia.org/wiki/NMR>).

Silné elektromagnetické pole způsobí přechod jádra do jiného spinového stavu. Když se excitované jádro vrací do původního (uspořádaného stavu), emituje měřitelné radiofrekvenční záření. Jádra ležící vedle sebe se navzájem ovlivňují, což se projevuje na povaze emitovaného záření, a právě tyto interakce dovolují určit vzdálenosti mezi různými částmi molekuly. Spolu se znalostí primární struktury umožňují určit strukturu terciální (Alberts *et al.*, 2002; <http://en.wikipedia.org/wiki/NMR>).

4 +RNA viry

+RNA viry obsahují „sense“ verzi svého genomu – genomová RNA (gRNA) slouží zároveň jako mediátorová RNA, která může být ihned po vstupu do hostitelské buňky překládána jejím translačním aparátem do proteinů. Po přeložení sekvence kódující virovou polymerázu, může začít prepis RNA do – formy, z níž je transkribována virová subgenomová RNA (sgRNA) a samozřejmě množství molekul +RNA, které slouží jako předloha pro další virové proteiny nebo jsou zabaleny do infekčních virových partikulí a slouží k dalšímu šíření viru (Závada, 1999).

Na obou koncích virových i většiny eukaryotických mRNA se nacházejí netranslatované oblasti (angl. „nontranslated region“ nebo „untranslated region“, NTR nebo UTR), jež nejsou překládány do proteinů nacházejí se před iniciačním kodónem a za kodónem terminačním, ale jsou vysoce strukturované a mají množství regulačních funkcí (Závada, 1999; Flint *et al.*, 2004; http://en.wikipedia.org/wiki/5%27_UTR).

5' UTR +RNA virů velmi často obsahuje IRES, byly nalezeny u flavivirů, pikornavirů, dicistrovirů, retrovirů a nacházejí se také u některých eukaryotických mRNA. IRES je sekvence s málo konzervovanou primární strukturou, ale s množstvím

konzervovaných sekundárních motivů, která řídí interní iniciaci translace navázáním 40S podjednotky ribozómu do 3' koncové části IRES. IRES je umístěno většinou před iniciačním AUG kodónem. IRES můžeme rozdělit do pěti druhů. Mechanismus funkce všech IRES není stejný např. u flavivirů probíhá iniciace translace bez přítomnosti některých eukaryotických iniciačních faktorů a je závislá na sekundární struktuře RNA kolem AUG iniciačního kodónu, pikornaviry zase navazují 40S nepřímo přes eIF4G (Schuler *et al.*, 2006, Fletcher a Jackson, 2002; Flint *et al.*, 2004).

4.1 čeleď: *Flaviviridae*

rody: *Pestivirus*, *Hepacivirus*, *Flavivirus*,

Hepaciviry a pestiviry narozdíl od rodu flavivirů nemají na svém 5' konci čepičkovou strukturu, tudíž mají na čepičce nezávislou iniciaci translace a ve své 5' UTR vždy obsahují přibližně 350 nt dlouhé IRES.

Hepaciviry: Virus žloutenky typu C (angl. „Hepatitis C virus“, HCV)

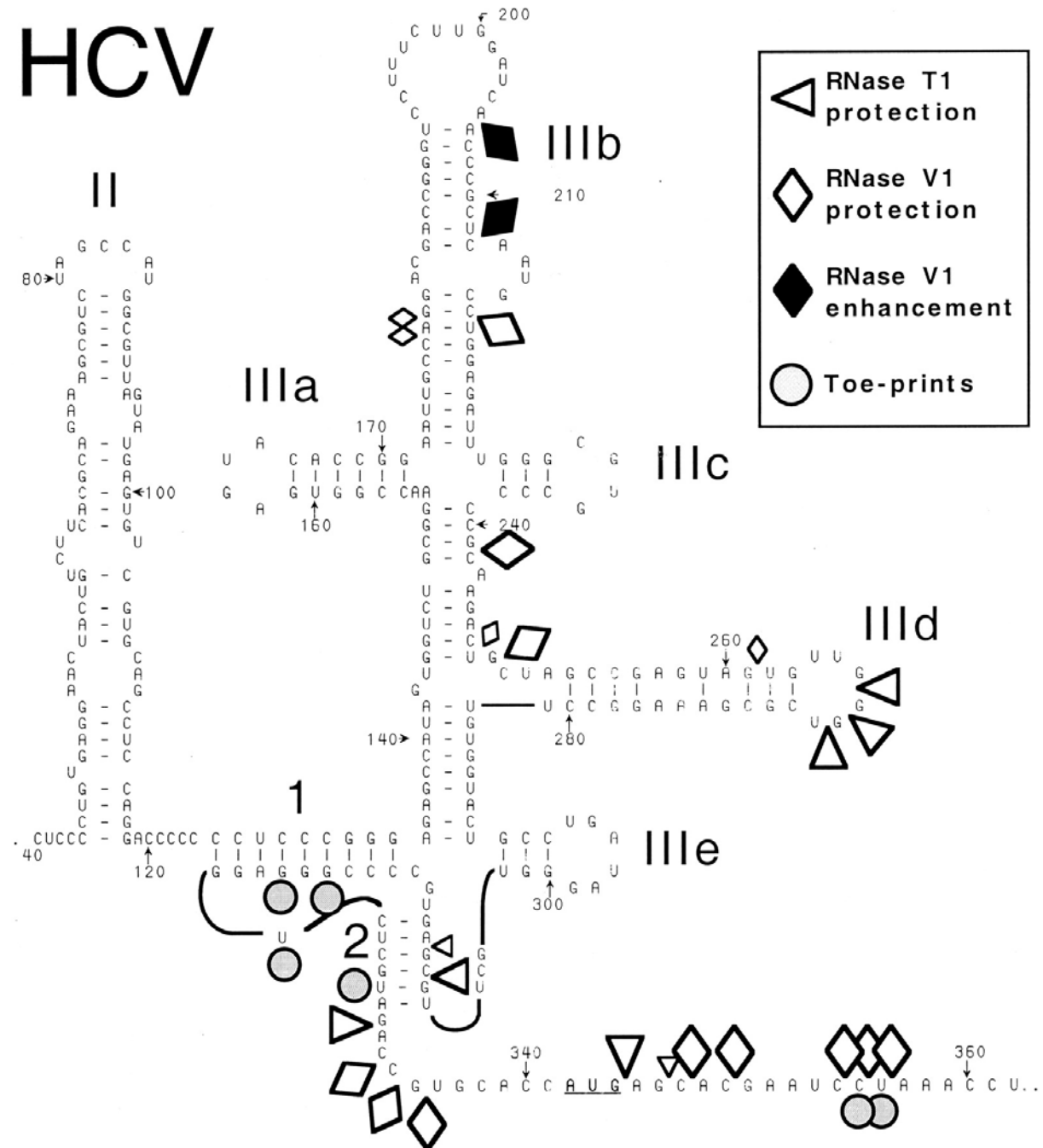
schéma: 1, 3, 4, 5⁴

I když mezi izoláty HCV z různých pacientů jsou velké sekvenční variace (tzv. kvazidruhy), sekundární struktura 5' UTR je velmi konzervovaná. IRES zabírá 330 nt z 341 nt dlouhé 5' UTR, ta se skládá ze čtyř hlavních domén I–IV. 5' UTR obsahuje rozsáhlé spárované úseky. Při iniciaci translace se ribozóm připojuje přímo na IRES bez pomoci iniciačních faktorů eIF4A, -4B, -4E a -4G a následná iniciace translace nevyžaduje ani ATP hydrolyzu jako u pikornavirové IRES. III. doména je nejčlenitější, z hlavní velké vlásenky, která má při základně pseudouzel, se odvětvují další vlásenky – subdomény IIIa – IIIe. Při iniciaci translace hrají roli hlavně domény II–III, ty se podílejí na vytváření četných interakcí s 40S podjednotkou ribozómu. I. doména se v řízení translace neuplatňuje a IV. působí dokonce jako negativní regulátor translace (Kolupaeva *et al.*, 2000).

Victoria Kolupaeva (2000) zjišťovala přesnou polohu nukleotidů interagujících s 40S pomocí enzymového footprintingu před a po navázání ribozómové podjednotky, hledala které nt jsou nově chráněny před atakem RNáz. Nejvíce interakcí našla na vrcholu subdomény

⁴ Schémata - zjednodušený zápis sekundární struktury jsou spolu s experimentálními podmínkami, enzymy a chemickými činidly uvedeny v příloze.

IIIId, v helixu mezi subdoménami IIIc a IIIId a v pseudouzlu. Subdoména IIIc interaguje s eIF3, ale na interakci s 40S se nepodílí. Navázání ribozómu indukuje změny v IIIIb, pseudouzlu a v doméně IV.



Obrázek 3: Sekundární struktura 5' UTR HCV určena pomocí počítačové predikce, komparativní analýzy i enzymatického mapování. Sekundární struktura I. domény se v této studii neověřovala experimentálně, proto na obrázku chybí, doména IV (nt. 331 – 354) je pro zjednodušení znázorněna jako lineární sekvence. Na obrázku jsou trojúhelníčky a kosočtverci označeny místa se změněnou citlivostí k RNázám T1 a V1, v důsledku navázání ribozómové podjednotky 40S (tedy oblasti interagující s 40S). Kolečkem jsou označena místa, kde po navázání 40S docházelo k zastavení prodlužování primeru.
převzato z Kolupaeva *et al.*, 2000

Pseudouzlel (III. doména) je konzervován mezi všemi podtypy HCV i mezi členy pestivirové čeledi a jeho nepostradatelnost pro efektivní translaci odhalil už Changyu Wang v roce 1995.

Oblast IRES, a to jak primární sekvence, tak i sekundární struktura je velmi konzervovaná nejen v sérotypech HCV, ale i v příbuzném rodu pestivirů, primární struktura pestivirů a hepacivirů se shoduje z více než 70 %. Strukturní konzervovanost u flaviviridů, které infikují široké spektrum hostitelů a mají různorodé cílové tkáně, je důkazem, že flaviviridové struktury interagují s vysoce konzervovanými elementy hostitelského aparátu. I když jsou sekundární struktury 5' UTR hepacivirů a pikornavirů odlišné, sdílejí společný rys, který u pestivirů nenajdeme, krátké oligopyrimidinové úseky (Brown *et al.*, 1992).

Masao Honda (1999) navrhl na základě fylogenetického srovnání s se savčím hepatotropním virem GBV-B upravenou strukturu domény II.

Pestiviry: *Bovine Viral Diarrhea* (BVDV)

Classical swine fever virus (CSFV)

schéma: 2, 6

Stejně jako u HCV i BVDV většinu z cca 385 nukleotidů dlouhé 5' UTR tvoří IRES, i počet domén mají oba přízvěněné viry totožný (čtyři) (Brown *et al.*, 1992). Vlášenska nejbliže 5' konci shodně jako u HCV není potřebná pro IRES funkci (pro navázání ribozómu). Její stonek je tvořen nt (1-9 a 24-32), stonek lze ještě dále rozdělit na dva rozlišitelné elementy: nižší část stonku-6pb a vyšší část stonku-3pb, jejich primární struktura je konzervovaná, zatímco smyčka (10-23) je tvořena variabilními nukleotidy. Tato vlášenska působí při přepínání mezi virovou translací a replikací pomocí virového NS proteinu a pro efektivní translaci je nezbytná. Přerušení ve stonkové architektuře vlášensky (narušení párování bází pomocí mutace), vede k zastavení translace (v důsledku i k zastavení replikace viru). Variabilní smyčková oblast se podílí na translaci méně, respektive mutace v oblasti smyčky nemají na translaci takový vliv (Yu *et al.*, 2000).

5' UTR CSFV se skládá ze tří domén I-III. CSFV IRES začíná přibližně na 65 nt, obsahuje celou doménu II, ale ne sekvence před ní. Delece v doméně II včetně těch malého rozsahu (několik nt) redukovaly IRES aktivitu na 20%. Tato zbytková aktivita byla úplně odstraněna až mutací, která zasáhla pseudouzel nacházející se za doménou III. Bylo zjištěno, že mutace ve velmi konzervativní oblasti IIIa překvapivě translaci neovlivňují (Fletcher a Jackson, 2002).

4.2 čeleď: *Picornaviridae*

rody: *Enterovirus*, *Rhinovirus*, *Aphthovirus*, *Hepatovirus*, *Cardiovirus*

Enteroviry: *poliovirus*

schéma: 8, 10, 11, 14

5' UTR enterovirů obsahuje strukturu připomínající jetelový list složenou ze smyček A-C, která leží hned u 5' konce viru, je označována jako doména I, za ní následuje dalších 5 domén II-VI. 5' UTR polioviru je dlouhá 742 nt a je vysoce strukturovaná. Translace virových proteinů je ovlivňována vlásenkami II-VI, které mají IRES funkci, tyto vlásenky se nacházejí se mezi nt 135 – 585 (Gamarnik a Andino, 2000).

Andrea Gamarnik (2000) mapovala místa interakce buněčného poly(C) vazebného proteinu (PCBP) potřebného pro virovou replikaci s 5' UTR polioviru a prokázala interakce s doménou I a IV. Afinita domény IV k PCBP je šestkrát vyšší než k doméně I, nicméně tento poměr je modifikován interakcí virového proteinu 3 CD s doménou I. (ten dramaticky zvyšuje afinitu této domény k PCBP). Dokonce i různé další pikornaviry jako je coxsackievirus a virus encephalomyokarditidy a také buněčné proteiny mohou s poliovirem o buněčný PCBP kompetovat.

Evgeny Pilipenko (1989b) mapoval oblast 5' UTR polioviru mezi nt 236-628, našel tři domény A-C⁵, tyto domény odpovídají doménám IV-VI ve studii Andrey Gamarnik (2000). Doména A je tvořena velkou větvenou vlásenkou a má stejně jako doména I tvar trojlístku, nachází se mezi nt 236-442, doména B (nt 451-559) a doména C (nt 564-628) tvoří nevětvené vlásenky s výdutěmi.

Další model byl vytvořen Michaellem Skinnerem (1989), ve své studii se věnoval struktuře celé oblasti 5' UTR polioviru, avšak experimentálně podložil pouze sekundární strukturu oblasti mezi nt 240-540. Struktura dvou domén, které se v této oblasti nacházejí, je podobná struktuře odpovídajících domén ve výše zmíněných studiích.

Protein-RNA interakce studoval Lyle Najita (1990), zaměřil se na interakci oblasti mezi nt 178-224 s 50 kDa proteinem, ten obsahuje cysteinové oblasti, které jsou nezbytné pro navázání na RNA.

⁵ Evgeny Pilipenko nazval domény I-III, aby však nedošlo k záměně s doménami I-III z předchozí studie, jeho názvosloví jsem pro potřeby této práce upravila.

Aftoviry: virus slintavky a kulhavky (angl. „Foot and Mouth Disease Virus“, FMDV)

Kardioviry: virus encefalomyokarditidy (angl. „Encephalomyocarditis virus“, EMCV)

schéma: 9, 12

Ačkoli sekundární struktura 5' UTR aphtovirů a kardiovirů je hodně podobná, se strukturou 5' UTR hepacivirů nebo enterovirů má společného velmi málo. EMCV má 5' UTR tvořenou sérií vlásenek A–N (více než 65% nt je spárovaných), které tvoří pět strukturních domén I–V (Pilipenko *et al.*, 1989a).

IRES začíná na 400. nt a končí iniciačním kodónem AUG na 834. nt, zahrnuje vlásenky H–N, jeho ss úseky mají tendenci tvořit smyčky a výdutě, které nejspíš hrají přímou roli v interní iniciaci translace (Kolupaeva *et al.*, 1996).

Victoria Kolupaeva (1996) mapovala oblasti interagující s proteinem vázajícím se na pyrimidinové oblasti (angl. „polypyrimidin tract binding protein“, PTB) na IRES EMCV a spřízněného FMDV. Oba sice patří do různých rodů, ale sekundární struktura jejich 5' UTR je velmi podobná (5' UTR FMDV je tvořeno také pěti doménami s obdobnou strukturou), primární struktura vykazuje o něco nižší 40% identičnost. PTB, které svým navázáním stabilizuje aktivní konformaci IRES, se váže na šest míst v 5' UTR, z nichž čtyři tvoří oligopyrimidinové oblasti. Ovšem samotná sekvence těchto oblastí k navázání PTB nestačí, je nutná i přítomnost strukturních elementů v IRES, které tyto oblasti prezentují v konformaci, ve které je umožněno navázání PTB.

Hepaciviry: virus žloutenky typu A (angl. „Hepatitis A virus“, HAV)

schéma: 7, 13

Vysoce konzervovaná cca 730 nt dlouhá 5' UTR hepacivirů obsahuje šest hlavních strukturních domén I–VI a více než deset start kodónů AUG. IRES se nachází mezi nt 152–735 a nezahrnuje první dvě domény. Efekt IRES na translaci je mnohem nižší než u ostatních pikornavirů, v kompetici s EMCV IRES o translační faktory buňky je HAV IRES 25x méně aktivnější (Brown *et al.*, 1994).

5' UTR obsahuje dvě vysoce strukturované oblasti 1–94 (ve které se nachází domény I a II) a 155–735 (ve které se nachází domény III–VI). Obě oblasti jsou oddělené čistým polypyrimidinovým traktem 99–124, který obsahuje 5 tandemových repetice (U)UCC(C), u lidských hepatovirů se počet repetice pohybuje od tří do pěti. II. doména se skládá ze dvou menších vlásenek s tendencí k tvorbě pseudouzlů. Domény IV a V byly díky kovarianční analýze určeny nejpřesněji. IV. doména je velká větvená vlásenka s několika vnitřními

smyčkami, má zásadní význam pro translaci, o něco menší V. doména je tvořena také větvenou vlásenkou, která může tvořit pseudouzél (Brown *et al.*, 1991; Brown *et al.*, 1994).

HAV 5' UTR je odlišná od ostatních pikornavirů, má však několik strukturních motivů společných s aftoviry a cardioviry – vlásenkové struktury s tendencí k tvorbě pseudouzlů před IRES oblastí, polypyrimidinový trakt, doména IV HAV je velmi podobná doméně III EMCV, v obou případech je doména tvořena dlouhou vlásenkou s množstvím bublin zakončenou čtyřlístkovou strukturou. Menší podobnost sdílí 5' UTR HAV s rhinoviry a nejméně je podobná 5' UTR enterovirů (Shaffer *et al.*, 1994).

4.3 řád: *Nidovirales*

4.3.1 čeleď: *Coronaviridae*

Bovine Coronavirus (BCoV)

schéma: 15, 16

BCoV má 5' UTR dlouhou 210 nt, obsahuje čtyři vysoce uspořádané vlásenkové domény I–IV, které působí v replikaci, transkripci a translaci. Vlášenska I je tvořena z nt 11-42, vlásenka II z nt 51-84. Vlášenska III (nt 97-116) s nejjednodušší strukturou bez jakýchkoli výdutí či větvení se podílí na replikaci defektních interferujících (DI) RNA a je asociována s ORF, který začíná na 100. nt a potenciálně kóduje osm aminokyselin a možná funguje jako zesilovač replikace DI RNA. Iniciační kodón tohoto ORF udržuje stonkovou integritu vlásenky III. Vlášenska IV (nt 186-215) je také potřebná pro replikaci DI RNA. Vlášensky III a IV by mohly mít funkci v sestavování RNA replikačního komplexu. Vlášensky II–IV interagují s buněčnými proteiny při virové replikaci. Na vlásenku IV se může navázat šest buněčných proteinů (ještě nebylo zjištěno o jaké proteiny se jedná) o velikosti 25-58 kDa. Sekundární struktura 5' UTR je mezi koronaviry relativně různá (Raman *et al.*, 2003; Raman a Brian, 2005).

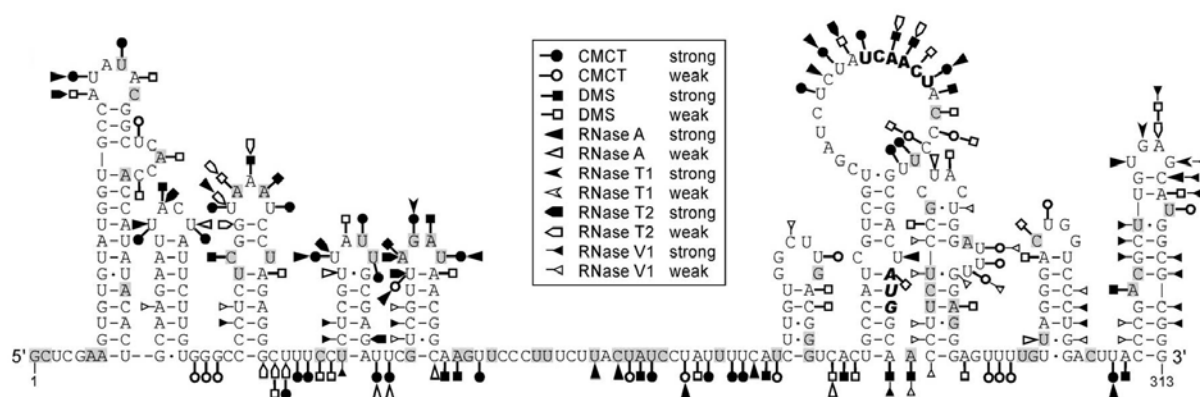
4.3.2 čeleď: *Arteroviridae*

Equine arteritis virus (EAV)

schéma: 17

Na 5' konci gRNA EAV se nachází čepičková struktura. 5' UTR EAV je dlouhá 244 nt a díky své vysoké strukturovanosti chrání RNA před degradací exonukleázami, umožňuje propojení 5' a 3' konců gRNA a má roli i při sgRNA syntéze. 5' UTR obsahuje IRES, které se uplatňuje v řízení translace. Erwin van de Born (2004) zjistil, že 313 koncových nt obsahuje deset vlásenek A–J. Vlášeny tvoří dvě konzervované domény (I. doména – vlášeny A–E, II. doména vlášeny F–J), domény jsou spojeny variabilní oblastí.

Sedmá vlášenka G má nejdelší smyčku, ve které se nachází transkripční regulující sekvence (TRS), tato vlášenka se nazývá LTH (angl. „Leader TRS Hairpin“) a je nezbytná pro syntézu sgRNA, ze které se překládají všechny strukturální proteiny. Na stonku LTH se nachází iniciační kodón AUG. LTH vlášenka je konzervovaná i mezi ostatními arteroviry a zřejmě i koronaviry. Erwin Van de Born (2005) navrhl, že by se LTH a přilehlé oblasti mohly skládat do rozšířené konformace (eLTH), přilehlé oblasti by se spárovaly a tím rozšířily LTH stoněk. Toto konformační přepnutí by mohlo regulovat produkci sgRNA tím, že přilehlé oblasti v konformaci eLTH zvyšují efektivitu syntézy sgRNA



Obrázek 4: Model sekundární struktury 5' koncových 313 nt EAV. Tučným písmem je zvýrazněna konzervovaná sekvence nt v TRS ve vlásence LTH a iniciační kodón AUG. V šedých čtverečích se nacházejí nt, které se liší v různých izolátech EAV. Počítačem vytvořený model je podpořen výsledky fylogenetické analýzy a chemickým i enzymatickým mapováním sekundární struktury. Převzato z Van de Born, 2004

4.4 čeleď: *Dicistroviridae*

schéma: 18

V roce 2002 byla vytvořena na zasedání mezinárodní komise pro virovou taxonomii International Committee on Taxonomy of Viruses (ICTV) (http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/fs_dicis.htm) nová čeleď virů *Dicistroviridae*,

obsahující hmyzí viry. Do jediného rodu *Cripavirus* – pojmenován podle viru cvrččí paralýzy (angl. „Cricket paralysis virus“, CrPV), byly přidány viry Triatoma virus (TrV) and Black-queen cell virus (BQCV) a virus akutní včelí paralýzy (angl. „Acute bee paralysis virus“, ABPV) (Mayo 2002). Později byly do Cripavirů přidány ještě další viry: Rhopalosiphum padi virus (RhPV), Aphid lethal paralysis virus (ALPV), Drosophila C virus (DCV), Himetobi P virus (HiPV), Plautia stali intestine virus (PSIV) a Taura syndrome virus (TSV).

Původně byly dicistroviry řazeny do skupiny pikornavirů, liší se však v přítomnosti dvou otevřených čtecích rámců (angl. „open reading frame“, ORF). Jediný ORF picornavirů kóduje jeden polyprotein, z něhož jsou nastříhány všechny virové proteiny strukturní i nestrukturní. Na rozdíl od pikornavirů genom dicistrovirů obsahuje dva nepřekrývající se ORF a každému z nich předchází jedno IRES. Tato IRES jsou evolučně odlišná od IRES savčích a rostlinných virů. První ORF kóduje nestrukturní proteiny druhý proteiny strukturní. První IRES leží v 5' UTR oblasti (5' IRES) a druhé mezi oběma ORF v intergenovém regionu (IGR IRES). IGR IRES Dicistrovirů mají konzervovanou strukturu a fungují nejenom v hmyzích, ale i v rostlinných (extrakt ze zárodečných buněk pšenice) a savčích buňkách (lyzát z králičích retikulocytů), jsou nezávislé na iniciačních proteinech. IGR IRES jsou aktivnější než jejich 5' protějšky, což se projevuje ve zvětšení poměru vytvářených strukturních proteinů ku proteinům nestrukturním (Domier *et al.*, 2000; Wilson *et al.*, 2000).

5' IRES PSIV je funkční jenom v hmyzích buňkách, naproti tomu 5' IRES RhPV a TrV pracují nezávisle na translační specifitě hostitelského aparátu. 5' IRES CrPV fungují v savčích buňkách ale ne v rostlinných. Tyto rozdíly ukazují, že 5' IRES elementy mají zřejmě odlišné požadavky na iniciační proteiny a narozdíl od IGR IRES oblastí je mechanismus iniciace translace prekurzorů nestrukturních proteinů mezi různými dicistroviry odlišný. Ani sekundární struktura těchto oblastí není u dicistrovirů shodná, dokonce i iniciační místo translace není stejné, u DCV a TSV začíná na druhém AUG v ORF, jinde na prvním, v některých případech začíná translace i na nemethioninovém kodónu (Shibuya a Nakashima, 2006).

Plautia stali intestine virus (PSIV)

Martin Schuler (2006) pomocí vylepšené cryo – EM mikroskopie s rozlišením 0,73 nm zjistil, že IRES obsahuje tři domény, každá z nich obsahuje pseudoknot – doména I obsahuje pseudoknot II, doména II pseudoknot III a doména III pseudoknot I. Tyto tři domény jsou potřebné pro iniciaci translace na nemethioninovém kodónu. Každá z těchto domén má specifickou funkci: I. doména interaguje s 60S ribozómem v E a P místě, II. doména

interaguje s 40S ribozómem v E místě a III. doména umísťuje alaninový start kodón do A místa. IRES struktura je schopná podstupovať konformačné zmeny, III. doména se během dekodování alaninového start kodónu úplně vysouvá z A místa.

4.5 Rostlinné +RNA viry

+RNA viry rostlin můžeme rozdělit do dvou superrodin, jedna je příbuzná živočišným pikornavirům a druhá alfavirům.

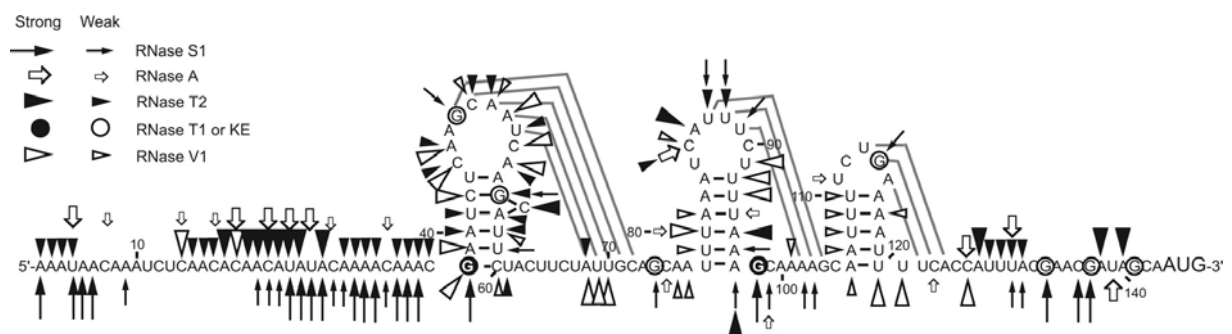
Pikornavirová superrodina

čeleď: *Potyviridae*

Virus leptané mozaiky tabáku (angl. „Tobacco Etch virus“, TEV)

schéma: 21

Na 5' konci gRNA obsahuje virový protein Vpg, který je však odstraněn před navázáním ribozómu. TEV obsahuje 143 nt dlouhou 5' UTR, na rozdíl od živočišných pikornavirů je kratší a méně strukturovaná a neobsahuje žádné AUG triplety před iniciačním AUG kodónem. 5' UTR obsahuje IRES, které svou velikostí patří mezi jedno z nejmenších. 5' UTR obsahuje domény I (nt 28-65) a II (nt 66-118), I obsahuje jeden pseudouzle, II dva. Pomocí delečních mutantů bylo zjištěno, že pro IRES funkci je kritická jenom oblast mezi nt 28-75 a že translace je spíše než jedním velkým regionem řízena několika menšími oblastmi (Zeenko a Gallie, 2005).



Obrázek 5: Sekundární struktura 5' UTR TEV. Počítačový model je podpořen výsledky enzymatického a chemického mapování. Šedými čarami jsou naznačeny pseudouzlové struktury. Zkratkou KE je označen kethoxal.

Převzato z Zeenko a Gallie, 2005

Alfavirová superrodina

čeleď: *Potexviridae*

Virus bramboru X (angl. „Potato virus X“, PVX)

schéma: 20

5' UTR PVX je velmi krátká, obsahuje jenom 84 nt. 5' UTR lze rozdělit na dvě oblasti α (nt 1-41) a β (nt 42-84), prvních 31 nt oblasti α je nespárovaných. 5' UTR (od nt 32) spolu s dalšími 22 nt gRNA tvoří vlásenku SL1, která kromě koncové čtyřnukleotidové smyčky obsahuje ještě další tři nespárované úseky (Miller *et al.*, 1998).

Konzervovaný oktanukleotid nacházející se před dvěmi hlavními sgRNA stejně jako elementy v 5' UTR ovlivňuje akumulaci sgRNA. Tyto dvě oblasti spolu párují, i když jsou odděleny dlouhou sekvencí nt (interakce na dlouhou vzdálenost). Mutace v 5' UTR, které snižují komplementaritu s oktanukleotidem, ovlivňovaly hladinu sgRNA i gRNA. Pro optimální akumulaci sgRNA je navíc kritická i primární struktura oktanukleotidu, mutanti v tomto úseku dokonce ztrácejí infekčnost. Interakce na dlouhé vzdálenosti mohou být ovlivňovány buněčnými či virovými proteiny, které mohou stabilizovat či destabilizovat RNA-RNA párování skrz RNA-proteinové interakce či protein-proteinové interakce (Kim a Hemenway, 1999).

čeleď: *Tombusviridae*

Virus keřičkovosti rajčat (angl. „Tomato Bushy Stunt Virus“, TBSV)

schéma: 19

5' UTR TBSV je dlouhá 166 nt. Baodong Wu (2001) se zaměřil na 5' koncových 78 nt, které se skládají do domény tvaru T. Tato doména je nezbytná pro efektivní replikaci viru, její funkce byla testována na replikaci DI, bez této domény akumulace DI poklesla až 20x. Doména tvaru T je uzavřena spárovaným úsekem mezi nt 1-10 a 69-78, tento úsek má na efektivitu replikace největší vliv.

5 čeleď: *Retroviridae*

Ačkoli Retroviry ve virové partikuli obsahují +RNA svého genomu, mají díky své odlišné replikační strategii (RNA je přepisována virovou reverzní transkriptázou do dsDNA, ta se pomocí virové integrázy začleňuje do genomu hostitele a integrovaný provirus je poté

přepisován jadernou RNA polymerázou II zpět do RNA) jiné složitější uspořádání konců genomu. 5' UTR obsahuje velké množství informací ve formě sekvencí a zároveň strukturních motivů, které jsou rozeznávány proteiny nebo jinými RNA a které regulují klíčové kroky virového cyklu jako je transkripce, translace, polyadenylace, sestřih, jaderný export a další. Další podstatný rozdíl od ostatních virů je diploidita retrovirů, infekční partikule obsahují dvě kopie virové RNA. Hostitelské buňky exprimují dva různé proviry a mohou produkovat i heterozygotní viriony. RNA genom na 5' konci začíná terminální repeticí, následuje unikátní sekvence (U5) a vazebné místo pro primer (angl. „primer binding site“, PBS), které je komplementární k buněčné transferové RNA (Závada, 1999).

Virus lidské imunodeficiencie (angl. „Human Immunodeficiency virus“, HIV)

Člověku jsou nebezpečné dva typy tohoto viru HIV-1 a HIV-2. HIV-1 sdílí největší homologii s opičím lentivirem (angl. „Simian Immunodeficiency viruses“, SIV) izolovaným z šimpanzů (angl. „chimpanzee“, SIV-CPZ) a primární struktura HIV-2 je nejvíce homologní se SIV izolovaným z mangabejů rudohlavých (angl. „Sooty Mangabey“, SIV-SMM) a se SIV izolovaným z makaků (angl. „Macaque“, SIV-MAC). Tyto homologie podporují představu o přenosu viru z opic do lidské populace a to dokonce dvakrát (Berkhout, 1992; Flint *et al.*, 2004)

S touto teorií evoluční příbuznosti HIV-SIV virů souhlasí také komparativní analýza transkripce aktivujícího regionu (TAR) HIV-1 s SIV-CPZ, která našla podobnosti v sekundární struktuře, u obou je TAR tvořeno stabilní jednoduchou vlásenkou. Další podobnost byla komparativní analýzou nalezena překvapivě mezi TAR HIV-2 a TAR opičím lentiviru mandril SIV-MND (angl. „mandrill“)⁶, i když ve fylogenetickém stromu vytvořeném na základě primární struktury, byly zařazeny do odlišných skupin. U obou je TAR tvořeno rozvětvenou vlásenkou (Berkhout, 1992).

Virus lidské imunodeficiencie 1 (angl. „Human Immunodeficiency virus 1“, HIV-1)

schéma: 22, 26, 27, 28, 30, 31, 33

5' UTR obsahující 335 nt je strukturně nejkonzervovanější část HIV-1 genomu a obsahuje regulační motivy, které se uplatňují ve všech krocích virového cyklu – TAR, polyadenylační (poly(A)) signální vlásenku, již zmíněné PBS, vlásenku obsahující dimerizační místo (DIS), sestřihové místo (angl. „splice donor“, SD), potřebné pro tvorbu

⁶ Mezi HIV-2 a SIV-MAC nebo SIV-SMM není téměř žádná sekvenční variace, a proto se pro ně předpokládá stejná sekundární struktura.

sgRNA a signál ψ , který zprostředkovává zabalení do virových partikulí (Abbink a Berkhout, 2003).

K enkapsidaci viru HIV do virové partikule je nezbytná interakce Gag prekurzorového proteinu (p55) a jeho derivátu nukleokapsidového proteinu (p12) s ψ signálem (Clever *et al.*, 1995; Damgaard *et al.*, 1998)

Takuma Hayashi (1993) mapoval oblast HIV-1 mezi nt 750-795, která je nezbytná pro enkapsidaci viru. Tuto oblast tvoří 46 nt dlouhá vlásenka nesoucí signál ψ . Svoji studii prováděl na genomové RNA, jež byla na svém 5' konci prodloužena o 455 nt (signál pro navázání čepičkového proteinu tato RNA nese až na 455. nt).

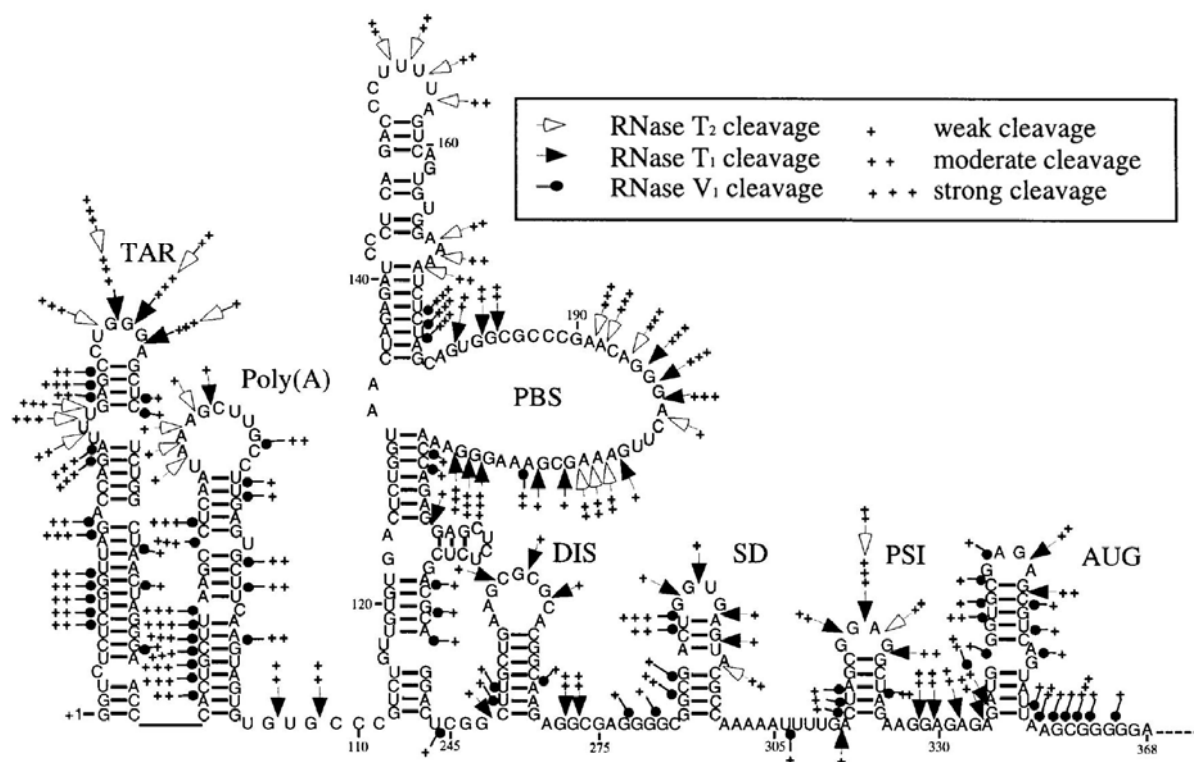
O rok dříve mapoval oblast ψ signálu HIV-1 Geoffrey Harrison (1992), svou studii prováděl taktéž na prodloužené RNA, ale s úplně odlišnými výsledky, našel rozvětvenou vlásenku tvaru čtyřlístku mezi nt 676-790, která zahrnuje také SD.

Ani sekundární struktura oblasti ψ signálu HIV-1 navržená Jardem Cleverem (1995) se s dvěma předchozími modely neshoduje (také prodloužená RNA). Clever mapoval oblasti interakce p55 a p12 s RNA HIV-1 a zjistil, že tyto proteiny se vážou na oblast dlouhou 206 nt (od 632. nt do 837. nt), obsahující 4 vlásenky SL1–SL4⁷. Každá z nich funguje jako nezávislé Gag vazebné místo a každá váže Gag protein s afinitou rovnající se čtvrtině afinity celé oblasti ψ signálu.

Christian Damgaard (1998) mapoval vazebná místa pro Gag protein a nukleokapsidový proteina na 5' koncových 401 nt gRNA HIV-1 (oblast 5' UTR + 66 nt z Gag ORF). Zjistil, že vlásenka nesoucí ψ , používá pro ni alternativní název PSI vlásenka (angl. „Packing Signal“, PSI), hraje hlavní roli v balení do virových partikulí. Ale ukázalo se, že i jiné oblasti jako je vlásenka s DIS, RNA zahrnující SD, balící signál a vlásenku s AUG mohou proces balení pozitivě ovlivňovat. Zjistil také, že proteiny HIV-1 interagovaly i s RNA HIV-2 a to především s vlásenkou PSI-3, která je homologní s vlásenkou PSI HIV-1.

5' koncových 352 nt HIV-1 včetně Gag iniciačního kodónu zaujímá za odlišných experimentálních podmínek dvě vzájemně se vylučující konformace – jednu s interakcemi na dlouhé vzdálenosti (angl. „long distance interaction“, LDI) a druhou s četnými rozvětvenými vlásenkami (angl. „branched multiple hairpin“, BMH). V obou konformacích tvoří nespárované nt 39 procent všech nt. Obě konformace se liší ve schopnosti tvořit diméry, LDI konformace je dimerizace neschopná, ale je termodynamicky stabilnější.

⁷ Odpovídají vlásenkám s DIS, s SD, s balícím signálem ψ a vlásence s AUG iniciačním kodónem (Damgaard *et al.*, 1998)



Obrázek 6: Sekundární struktura 5' UTR HIV-1. Model je podložen daty z enzymatického mapování RNázami T1, T2 a V1. Vlášeny jsou pojmenovány po regulačních motivech, které nesou. Převzato a upraveno z Damgaard *et al.*, 1998

Předpokládá se, že přepnutí konformace RNA z LDI formy do BMH formy iniciované přítomností nukleokapsidového proteinu, může koordinovat funkce spojené s UTR: způsobovat inhibici translace a aktivaci dimerizace a enkapsidace (Huthoff a Berkhout, 2001; Abbink a Berkhout, 2003).

V BMH konformaci existuje interakce na dlouhou vzdálenost – párování mezi iniciačním kodónem proteinu a unikátní oblastí (U5 – AUG duplex), toto párování se projevuje v regulaci translace a je jím aspoň částečně zvyšován také dimerizační efekt. U5 – AUG duplex je konzervovaný i mezi dalšími primáty lentiviry jako je SIV a HIV-2, i když v primární struktuře existují rozdíly. V LDI existuje párování mezi poly(A) a DIS, které způsobuje maskování DIS struktury a tím neschopnost dimerizace. V této konformaci má ψ vlášenska významně rozšířenou formu ψ E, obsahuje i SD signál a Gag start kodón (Huthoff a Berkhout, 2001; Abbink a Berkhout, 2003).

Jean Christophe Paillart (2004) provedl studie *ex vivo* v infikovaných buňkách a ve virionech a našel drobné rozdíly mezi svými výsledky a výsledky pokusů *in vitro*. Hlavní rozdíl je v nt 267-271 DIS, nejsou reaktivní *in vitro* ale *ex vivo* vykazují slabou reaktivitu,

DIS struktura je tedy destabilizována. *Ex vivo* studie ukazují, že naprosto majoritní část virové RNA se nachází v BMH formě, tudíž není moc termodynamicky stabilní, a naznačují také, že k dimerizaci genomové RNA by mohlo docházet již v cytoplazmě infikovaných buněk.

Virus lidské imunodeficiency 2 (angl. „Human Immunodeficiency virus 2“, HIV-2)
schéma: 24, 25

Primární struktura 5' UTR HIV-2 se rychle mění (velké odlišnosti v různých izolátech), ale některé sekundární struktury v této oblasti zůstávají stále stejné. Prvních 400 nt 5' UTR HIV je převážně spárovaných, skládá se z mnoha krátkých helixů, které jsou přerušovány různými strukturními nepravidelnostmi jako jsou výdutě a vnitřní smyčky. TAR oblast je tvořená dvouvětvnou vlásenkou a nachází se mezi nt 1-124 (sekvenčně i strukturně stejná rozvětvená vlásenka se nachází v TAR SIV-SMM a SIV-MAC, stejnou strukturu, ale jinou sekvenci obsahuje i TAR SIV-MND). Polyadenylační vlásenky se nacházejí mezi nt 125-184. Ben Berkhout (1993) předpokládá, že navázání primeru usnadňují i oblasti obklopující 18 nt PBS, proto PBS oblastí nazývá celou oblast mezi nt 197-379, která se skládá ze tří domén PBS-T, PBS-C, PBS-B (Berkhout, 1992; Berkhout a Schoneveld, 1993).

Moloney Murine Leukemia Virus (MoMLV)
schéma: 23, 32

5' UTR MoMLV je dlouhá 621 nk. Marylene Mougel (1993) mapovala sekundární strukturu oblastí nacházejících se před a za doménou s balícím signálem ψ . Tato nezávislá a zároveň vysoce strukturovaná oblast se nachází cca mezi nt 212-563 a skládá se z osmi vlásenek 4-11, i když pro zabalení virové RNA do partikulí nejsou pravděpodobně nutné ani vlásenky nacházející se za 400. nt (8-11) (Alford *et al.*, 1991).

Oblast U5-PBS domény obsahující poly(A), PBS (mezi nt 111-145) a SD může zaujímat dvě odlišné konformace, konformaci stabilní rozšířené helikální struktury (angl. „Extended helical structure“, E) a konformaci nepravidelné trojvětvné struktury (angl. „three Branched structure“, B). Vlásenka nejbliže 5' konci zabírající 26 nt z 68 nt dlouhé terminální repetice spolu s vlásenkou nacházející se mezi nt 113 a 143 svou konformaci nemění. Stejně jako u HIV je přepínání mezi oběma konformacemi řízeno zřejmě přítomností nukleokapsidového proteinu, který reguluje dimerizaci a enkapsidaci. gRNA ve formě monomeru můžeme najít v obou konformacích, gRNA diméry existují jenom v mírně

pozměněné konformaci B. AUG iniciační kodón je umístěn v ds oblasti vlásenky nacházející se mezi nt 590-630 (Mougel *et al.*, 1993).

Spleen necrosis virus (SNV)

schéma: 34

5' UTR SNV byl rozdělen do tří oblastí A nt 4-42, B nt 46-115 a C nt 115-165. A a C oblast tvoří dvě synergicky účinkující vlásenky, které jsou nezbytné k zachování translační aktivity. A vlásenka obsahuje jen jedinou výduť, která spolu se smyčkou vlásenky funguje dohromady jako vazebné místo pro protein či proteiny hostitelské buňky. Delece jenom jediného nt výduť významně snížila produkci proteinů. C vlásenka obsahuje dvě vnitřní smyčky a jednu výduť (Roberts a Boris-Lawrie, 2003).

Manson – Pfizer opičí virus (angl. „Manson – Pfizer monkey virus“, MPMV)

schéma: 29

MPMV způsobuje imunodeficienci u mláďat *macaca rhesus*. Geoffrey P. Harrison (1995) provedl mapování oblasti MPMV od 3' konce PBS do 130. nt nacházejícího se už v Gag kódující sekvenci – od cca 750 nt do 1020 nt. Objevil tzv. „stem with heptaloop“ - tedy velkou smyčku, ze které se větví sedm vlásenek, tato struktura je důležitá pro enkapsidaci. Mezi identifikovanými enkapsidačními signály retrovirů nebyla nalezena sekvenční konzervovanost, z toho vyplývá, že rozpoznání balící signálu ψ je nejspíš umožněno přítomností unikátních sekundárních či terciálních struktur. Stejně jako 5' koncová oblast HIV i 5' koncová oblast MPMV může *in vitro* zaujímat dvě rozdílné konformace. Vlásenka ležící mezi nt 838 až 863 je konzervována mezi mnoha příbuznými retroviry a i když se nachází před SD, ovlivňuje enkapsidaci vlásenky. I u fylogeneticky vzdálenějšího HIV a viru Rousova sarkomu byly nalezeny před SD sekundární struktury, které ovlivňují enkapsidaci.

6 Význam 5' koncových sekundárních struktur +RNA virů

Asi nejdůležitější funkce 5' koncových oblastí je regulace iniciace translace. Existují dva způsoby navázání ribozómu na RNA. První je na čepičce závislý, při kterém se iniciační komplex (podjednotka ribozómu 40S spolu s iniciačními proteiny) naváže na 5' koncovou čepičkovou strukturu a pohybuje se směrem k 3' konci, dokud nenarazí na iniciační kodón.

Druhý způsob je navázání iniciačního komplexu přímo na vnitřní ribozóm vazebné místo (angl. „internal ribosomal binding site“, IRES), které se nachází před iniciačním kodónem. Role sekundární struktury je u obou způsobů iniciace translace odlišná. Při iniciaci translace závislé na čepičkové struktuře je efektivita translace redukována přítomností stabilních elementů sekundární struktury a to zejména, když je vysoce strukturovaná oblast v blízkosti 5' konce. Důvodem je, že elementy sekundární struktury mohou blokovat navázání podjednotky ribozómu 40S či znesnadňovat nebo dokonce zcela zastavit pohyb ribozómové podjednotky směrem k iniciačnímu kodónu. Při iniciaci translace nezávislé na čepičkové struktuře bývá 5' koncová oblast vysoce strukturovaná a její strukturovanost je nezbytná pro navázání 40S ribozómové podjednotky (Flint *et al.*, 2004).

Další související uplatnění 5' koncových sekundárních struktur je v interakci, a to jak s dalšími strukturami RNA nacházejícími se mimo 5' koncovou oblast (Kim a Hemenway, 1999), tak s virovými a buněčnými proteiny (Kolupaeva *et al.*, 1996; Kolupaeva *et al.*, 2000; Raman a Brian, 2005) nebo celými nukleoproteinovými komplexy, jako je již dříve zmíněný iniciační komplex. Například u viru HIV je zabalení do virových partikulí zprostředkováno interakcí vládsenky se ψ signálem s virovými proteiny p55 a p12 (Harrison a Lever, 1992; Hayashi *et al.*, 1993; Damgaard *et al.*, 1998).

Strukturní elementy nacházející se v 5' UTR mohou být v některých případech nezbytné i pro virovou replikaci, přinejmenším pro efektivní virovou replikaci a to jak pro akumulaci + RNA (Miller *et al.*, 1998; Kim a Hemenaway, 1999) tak i pro vytváření - RNA formy (Yu *et al.*, 2006).

Další relativně nově objevenou schopností je přepínání mezi dvěma funkcemi spojenými s 5' UTR oblastí. Přepnutí je způsobeno konformační změnou, která je vyvolána interakcí s proteinem. U retrovirů je to přepnutí z translace viru na dimerizaci spojenou s enkapsidací pomocí nukleokapsidového proteinu, bylo objeveno jak u viru lidské imunodeficiency (Huthoff a Berkhout, 2001; Abbink a Berkhout, 2003) tak u *Moloney Murine Leukemia* viru (Mougel *et al.*, 1993). Podobné přepnutí bylo nalezeno i u arterovirů - *Equine arteritis virus*, zde je pomocí konformační změny regulována produkce sgRNA, ze které se překládají všechny strukturní proteiny (Van de Born *et al.*, 2005).

Analýza sekundární struktury RNA se díky konzervovanosti funkčních strukturních elementů může stát pomocným nástrojem pro určování fylogenetické příbuznosti (Berkhout, 1992)

7 Diskuze

Jednotlivé studie se i přes to, že pro mapování sekundární struktury použily kombinaci všech tří hlavních metodických postupů (počítačové predikce, enzymatické a chemické mapování, komparativní analýza), ve výsledných modelech sekundární struktury 5' UTR oblastí různých +RNA virů či retrovirů často liší.

Tyto odlišnosti mohou být způsobeny několika faktory, jedním z nich je používání různých sérotypů virů, dalším z faktorů je postupné zdokonalování počítačových programů, používání RNáz a chemických činidel od různých výrobců a za různých experimentálních podmínek nebo obtížně reprodukovatelné výsledky některých RNáz. Příkladem může být oblast HIV 1 obsahující ψ signál, sekundární strukturu namodelována různými studii se výrazně liší (Harrison a Lever, 1992; Hayashi *et al.*, 1993, Clever *et al.*, 1995) V některých studiích je přímo upozorněno na úseky, pro které nebyla udělána kovarianční analýza nebo kde existuje více možných uspořádání sekundární struktury (Skinner *et al.*, 1989).

Dalším problémem pro orientaci v sekundárních strukturách určených více studii se stává číslování domén, které nijak nezohledňuje přesnou polohu domén, ale jenom jejich pořadí ve zkoumané sekvenci. Takže ve studii, která provádí mapování 3' koncové části 5' UTR, je naprosto běžné a normální, že se domény začnou číslovat od jedničky (Pilipenko *et al.*, 1989b). Komplikace taktéž nastává, když různé studie určí odlišný počet domén, toto se však naštěstí neděje příliš často.

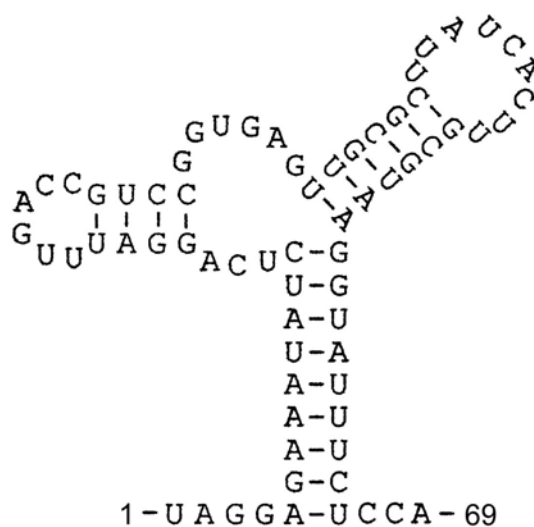
Proto mě při zpracovávání nebo spíše vyhledávání podkladů (odborných studií) pro mou bakalářskou práci nejvíce překvapilo, kolik studií se spoléhá jenom na jeden z metodických postupů (nejčastěji na počítačovou predikci). Nebo dokonce vyvozuje závěry týkající se funkcí úzce spjatých se sekundární strukturou (jako je virová replikace, translace, virulence, interakce s buněčnými a virovými proteiny...) na základě jediné studie sekundární struktury provedené před mnoha roky, aniž by autoři studie takovýto model jakkoli ověřili.⁸

Obávám se, že takovéto závěry nemohou být příliš důvěryhodné, a proto se vytvoření databáze sekundárních struktur spolu se stálým vylepšováním predikčních programů stává nezbytností pro spolehlivý výzkum funkcí souvisejících se sekundárními strukturami.

⁸ Typickým příkladem je používání modelů z Brownových studií 5' UTR HAV a HCV.

8 Závěr

Studium funkcí spojených se sekundárními strukturami RNA jako je translace či replikace a mnohé další je podmíněno znalostí těchto struktur. Zvláště u virů je znalost jejich sekundární struktury RNA velmi důležitá, oblasti s konzervovanými strukturálními elementy jako jsou IRES se mohou stát specifickým cílem léčiv. Kromě vylepšování počítačových predikčních programů bude jedním z prostředků, které by měly výzkum sekundárních struktur a s nimi spojených funkcí ulehčovat, také nově se vytvářející databáze experimentálně doložených sekundárních struktur v 5' UTR RNA, na jejímž vzniku se svou bakalářskou prací podílím. Tato databáze bude agregovat schematické přepisy sekundárních struktur RNA spolu s experimentálními podmínkami a použitými chemickými činidly a enzymy. Pomocí ní budou moci vědci vyhledávat a porovnávat informace o sekundární struktuře RNA virů či eukaryotické RNA, kterými se budou zabývat a podle shody dat se budou moci rozhodnout o možnosti zopakování určení sekundární struktury.



```

.....((((((((.....((((.....)))))).....((((.....)))))))))
UAGGAGAAUAUCUCAGGAUUUGACCGUCCGGUGAGUUGCGCUUAUCACUUGCGUAAGGUAU
))))...
UUCUCCA

```

Obrázek 7: Schématický zápis hypotetické sekundární struktury.

9 Seznam citované literatury

Abbink T. E. M., Berkhout B. (2003) A Novel Long Distance Base-pairing Interaction in Human Immunodeficiency Virus Type 1 RNA Occludes the Gag Start Codon. *The Journal of Biological Chemistry* 278:11601-1161

Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P. *Molecular biology of the cell*. 4th ed. New York: Garland Science, 2002. p. 299-520

Alford R. L., Honda S., Lawrence C. B., Belmont J. W. (1991) RNA Secondary Structure Analysis of the Packaging Signal for Moloney Murine Leukemia Virus. *Virology* 183:611-619

Auffinger P., Bielecki L., Westhof E. (2004) Anion Binding to Nucleic Acids. *Structure* 12:379-388

Berkhout B. (1992) Structural features in TAR RNA of human and simian immunodeficiency viruses: a phylogenetic analysis. *Nucleic Acids Research* 20:27-31

Berkhout B., Schoneveld (1993) Secondary structure of the HIV-2 leader RNA comprising the tRNA-primer binding site. *Nucleic Acids Research* 21:1171-1178

Bonnal S., Boutonnet Ch., Prado-Lourenco L., Vagner S. (2003) IRESdb: the Internal Ribosome Entry Site database. *Nucleic Acids Research* 31:1427-1428

Brown E. A., Day S. P., Jansen R. W., Lemon S. M. (1991) The 5' Nontranslated Region of Hepatitis A Virus RNA: Secondary Structure and Elements Required for Translation In Vitro. *Journal of Virology* 65:5828-5838

Brown E. A., Zhang H., Ping L., Lemon S. M. (1992) Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Research* 20:5041-5045

Brown E. A., Zajac A. J., Lemon S. M. (1994) In Vitro Characterization of an Internal Ribosomal Entry Site (IRES) Present within the 5' Nontranslated Region of Hepatitis A Virus RNA: Comparison with the IRES of Encephalomyocarditis Virus. *Journal of Virology* 68:1066-1074

Clever J., Sasseti Ch., Parslow T. G. (1995) RNA Secondary Structure and Binding Sites for *gag* Gene Products in the 5' Packaging Signal of Human Immunodeficiency Virus Type 1. *Journal of Virology* 69:2101-2109

Damgaard Ch. K., Dyhr-Mikkelsen H., Kjems J. (1998) Mapping the RNA binding sites for human immunodeficiency virus type-1 Gag and NC proteins within the complete HIV-1 and -2 untranslated leader region. *Nucleic Acids Research* 26:3667-3676

Ding Y., Lawrence Ch. E. (2003) A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Research* 31:7280-7301

- Domier I.L., McCoppin N.K., D'Arcy C.J. (2000) Sequence Requirements for Translation Initiation of Rhopalosiphum padi Virus ORF2. *Virology* 268:264-71
- Duguid J. G., Bloomfield V. A., Benevides J. M., Thomas G. J. (1995) Raman Spectroscopy of DNA-Metal Complexes. II. The Thermal Denaturation of DNA in the Presence of Sr^{2+} , Ba^{2+} , Mg^{2+} , Ca^{2+} , Mn^{2+} , Co^{2+} , Ni^{2+} , and Cd^{2+} . *Biophysical Journal* 69:2623-2641
- Eddy S. R., Durbin R. (1994) RNA sequence analysis using covariance models. *Nucleic Acids Research* 22:2079-2088
- Ehresmann Ch., Baudin F., Mougél M., Romby P., Ebel J., Ehresmann B (1987) Probing the structure of RNAs in solution. *Nucleic Acids Research* 15:9109-9128
- Farrell R. E. RNA Methodologies: A Laboratory Guide for Isolation and Characterization. 1st ed. London: Academic Press 1993. p.112-155
- Fletcher S. P., Jackson R. J. (2002) Pestivirus Internal Ribosome Entry Site (IRES) Structure and Function: Elements in the 5' Untranslated Region Important for IRES Function. *Journal of Virology* 76:5024-5033
- Flint S. J., Enquist L. W., Racaniello V. R., Skalka A. M. Principles of Virology: Molecular Biology, Pathogenesis, and Control of Animal Viruses. 2nd ed. Washington DC: ASM Press, 2004. p. 182-214, 378-410, 623-627
- Gamarnik A. V., Andino R. (2000) Interactions of Viral Protein 3CD and Poly(rC) Binding Protein with the 5' Untranslated Region of the Poliovirus Genome. *Journal of Virology* 74:2219-2226
- Harrison G. P., Lever A. M. L. (1992) The Human Immunodeficiency Virus Type 1 Packaging Signal and Major Splice Donor Region Have a Conserved Stable Secondary Structure. *Journal of Virology* 66:4144-4153
- Harrison G. P., Hunter E., Lever A. M. L. (1995) Secondary Structure Model of the Mason-Pfizer Monkey Virus 5' Leader Sequence: Identification of a Structural Motif Common to a Variety of Retroviruses. *Journal of Virology* 69:2175-2186
- Hayashi T., Ueno Y., Okamoto T. (1993) Elucidation of a conserved RNA stem-loop structure in the packaging signal of human immunodeficiency virus type 1. *FEBS Letters* 327:213-218
- Honda M., Beard M. R., Ping L., Lemon S. M. (1999) A Phylogenetically Conserved Stem-Loop Structure at the 5' Border of the Internal Ribosome Entry Site of Hepatitis C Virus Is Required for Cap-Independent Viral Translation. *Journal of Virology* 73:1165-1174
- Hu Qu L., Michot B., Bachelleri J. (1983) Improved methods for structure probing in large RNAs: a rapid 'heterologous' sequencing approach is coupled to the direct mapping of nuclease accessible sites. Application to the 5' terminal domain of eukaryotic 28S rRNA. *Nucleic Acids Research* 11:5903-5920

- Huthoff H., Berkhout B. (2001) Two alternating structures of the HIV-1 leader RNA. *RNA* 7:143-157
- Kankia R. (2003) Mg²⁺-induced triplex formation of an equimolar mixture of poly(rA) and poly(rU). *Nucleic Acids Research* 31:5101-5107
- Kim K. H., Hemenway C. L. (1999) Long-distance RNA-RNA interactions and conserved sequence elements affect potato virus X plus-strand RNA accumulation. *RNA* 5:636-645
- Knapp G. (1989): Enzymatic approaches to probing of RNA secondary and tertiary structure. (review) *Methods in Enzymology*: 180:192-212
- Kolupaeva V. G., Hellen C. U., Shatsky I. N. (1996) Structural analysis of the interaction of the pyrimidine tract-binding protein with the internal ribosomal entry site of encephalomyocarditis virus and foot-and-mouth disease virus RNAs. *RNA* 2:1199-1212
- Kolupaeva V. G., Pestova T. V., Hellen Ch. U. T. (2000) An Enzymatic Footprinting Analysis of the Interaction of 40S Ribosomal Subunits with the Internal Ribosomal Entry Site of Hepatitis C Virus. *Journal of Virology*: 74:6242-6250
- Mayo M. A. (2002) Virus Taxonomy – Houston 2002. *Archives of Virology* 147/5
- Mignone F., Grillo G., Licciulli F., Iacono M., Liuni S., Kersey P. J., Duarte J., Saccone C., Pesole G. (2005) UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research* 33:141-146
- Miller E. D., Plante C. A., Kim K., Brown J. W., Hemenway C. (1998) Stem-loop Structure in the 5' Region of Potato Virus X Genome Required for Plus-strand RNA Accumulation. *Journal of Molecular Biology* 284: 591-608
- Mougel M., Tounekti N., Darlix J., Paoletti J., Ehresmann B., Ehresmann Ch. (1993) Conformational analysis of the 5' leader and the gag initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. *Nucleic Acids Research* 21:4677-4684
- Nagaswamy U., Voss N., Zhang Z., Fox G. E. (2000) Database of non-canonical base pairs found in known RNA structures. *Nucleic Acids Research* 28:375-376
- Najita L., Sarnow P. (1990) Oxidation-reduction sensitive interaction of a cellular 50-kDa protein with an RNA hairpin in the 5' noncoding region of the poliovirus genome. *Proceedings of the National Academy of Science of the USA* 87:5846-5850
- Nowakowski J., Tinoco I. (1997) RNA Structure and Stability. *Seminars in Virology* 8:153-165
- Paillart J., Dettenhofer M., Yu X., Ehresmann Ch., Ehresmann B., Marquet R. (2004) First Snapshots of the HIV-1 RNA Structure in Infected Cells and in Virions. *The Journal of Biological Chemistry* 279:48397-48403
- Parsch J., Braverman J. M., Stephan W. (2000) Comparative Sequence Analysis and Patterns of Covariation in RNA Secondary Structures. *Genetics* 154:909-921

- Pesole G., Liuni S., Grillo G., Licciulli F., Larizza A., Maklowski W., Saccone C. (2000) UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research* 28:193-196
- Pilipenko E. V., Blinov V. M., Chernov B. K., Dmitrieva T. M., Agol V. I. (1989) Conservation of the secondary structure elements of the 5'-untranslated region of cardio- and aphthovirus RNAs. *Nucleic Acids Research* 17:5701-5711
- Pilipenko E.V., Blinov V. M., Romanova L. I., Sinyakov A. N., Maslova S. V., Agol V. I. (1989) Conserved Structural Domains in the 5'-Untranslated Region of Picornaviral Genomes: An Analysis of the Segment Controlling Translation and Neurovirulence. *Virology* 168:201-209
- Raman S., Bouma P., Williams G. D., Brian D. A. (2003) Stem-Loop III in the 5' Untranslated Region Is a *cis*-Acting Element in Bovine Coronavirus Defective Interfering RNA Replication. *Journal of Virology* 77:6720-6730
- Raman S., Brian D.A. (2005) Stem-Loop IV in the 5' Untranslated Region Is a *cis*-Acting Element in Bovine Coronavirus Defective Interfering RNA Replication. *Journal of Virology* 79:12434-12446
- Rijnbrand R., Bredenbeek P., van der Straten T., Whetter L., Inchauspe G., Lemon S., Spaan W. (1995) Almost the entire 5' non-translated region of hepatitis C virus is required for cap-independent translation. *FEBS Letters* 365:115-119
- Rivas E., Eddy S. R. (1999) A Dynamic Programming Algorithm for RNA Structure Prediction Including Pseudoknots. *Journal of Molecular Biology* 285:2053-2068
- Roberts T. M., Boris-Lawrie K. (2003) Primary Sequence and Secondary Structure Motifs in Spleen Necrosis Virus RU5 Confer Translational Utilization of Unspliced Human Immunodeficiency Virus Type 1 Reporter RNA. *Journal of Virology* 77:11973-11984
- Saenger W. Principles of nucleic acid structure. 1st ed. New York: Springer, 1984.
- Shaffer D. R., Brown E. A., Lemon S. M. (1994) Large Deletion Mutations Involving the First Pyrimidine-Rich Tract of the 5' Nontranslated RNA of Human Hepatitis A Virus Define Two Adjacent Domains Associated with Distinct Replication Phenotypes. *Journal of Virology* 68:5568-5578
- Shibuya N., Nakashima N. (2006) Characterization of the 5' internal ribosome entry site of *Plautia stali* intestine virus. *Journal of General Virology* 87:3679-3686
- Schuler M., Connell S. R., Lescoute A., Giesebrecht J., Dabrowski M., Schroeder B., Mielke T., Penczek P. A., Westhof E. & Spahn Ch. M. T. (2006) Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nature structural & Molecular biology*, 13:1092-1096
- Skinner A. M., Racaniello V. R., Dunn G., Cooper J., Minor P. D., Almond J. W. (1989) New Model for the Secondary Structure of the 5' Non-coding RNA of Poliovirus is Supported

- by Biochemical and Genetic Data That Also Show That RNA Secondary Structure is Important in Neurovirulence. *Journal of Molecular Biology* 207:379-392
- Stern S., Moazed D., Noller. (1988) Structural Analysis of RNA Using Chemical and Enzymatic Probing Monitored by Primer extension. *Methods in Enzymology* 164:481-489
- Tan L., Lai W., Lee Ch. T., Kim D. S., Choe W. (2007) Differential interactions of plasmid DNA, RNA and endotoxin with immobilised and free metal ions. *Journal of Chromatography* 1141:226-234
- Tinoco I. J., Bustamante C. (1999) How RNA folds. *Journal of Molecular Biology* 293:271-281
- Van den Born E., Gulyaev A. P., Snijder E. J. (2004) Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. *RNA* 10:424-437
- Van den Born E., Posthuma C. C., Gulyaev A. P., Snijder E. J. (2005) Discontinuous Subgenomic RNA Synthesis in Arteriviruses Is Guided by an RNA Hairpin Structure Located in the Genomic Leader Region. *Journal of Virology* 79:6312-6324
- Varani G., McClain W. (2000) The G-U wobble base pair: A fundamental building block of RNA structure to RNA function in diverse biological systems (review) *EMBO reports* 1:18-23
- Wang Ch., Le S., Ali N., Siddiqui A. (1995) An RNA pseudoknot is an essential structural element of the internal ribosome entry site located within the hepatitis C virus 5' noncoding region. *RNA* 1:526-537
- Weixlbaumer A., Werner A., Flamm Ch., Westhof E., Schroeder R. (2004) Determination of thermodynamic parameters for HIV DIS type loop-loop kissing complexes. *Nucleic Acids Research* 32: 5126-5133
- Wilson J.E., Powell M.J., Hoover S.E. and Sarnow, P. (2000) Naturally Occurring Dicistronic Cricket Paralysis Virus RNA is Regulated by Two Internal Ribosome Entry Sites. *Molecular Cellular Biology* 20: 4990-4999
- Wu B., Vanti W. B., White A. K. (2001) An RNA Domain within the 5' Untranslated Region of the Tomato Bushy Stunt Virus Genome Modulates Viral RNA Replication. *Journal of Molecular Biology*. 305: 741-756
- Yu H., Isken O., Grassmann C. W., Behrens S. (2000) A Stem-Loop Motif Formed by the Immediate 5' Terminus of the Bovine Viral Diarrhea Virus Genome Modulates Translation as well as Replication of the Viral RNA. *Journal of Virology* 74:5825-5835
- Závada V. *Molekulární virologie I.: RNA viry, viroidy a viry replikující se reverzní transkripcí*. 1. vyd. Praha: Peres, 1999.
- Zeenko V., Gallie D. R. (2005) Cap-independent Translation of Tobacco Etch Virus Is Conferred by an RNA Pseudoknot in the 5' Leader. *The Journal of Biological Chemistry* 280:26813-26824

<http://cmgm.stanford.edu/biochem218/Projects%20Spring%202003/Silverman.pdf>
http://www.comp.nus.edu.sg/~ksung/cs5238/project/final/rnafold_YeNan.pdf
http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/fs_dicis.htm
http://en.wikipedia.org/wiki/5%27_UTR
<http://en.wikipedia.org/wiki/Crystallography>
<http://en.wikipedia.org/wiki/NMR>
<http://en.wikipedia.org/wiki/Pseudoknot>

PŘÍLOHA

FLAVIVIRIDAE

1 HCV (Brown et al., 1992)

```
....((((((.....))))).((((.....((((((.....))
GCCAGCCCCUGAUGGGGCGACACUCCACCAUGAAUCACUCCCCUGUGAGGAACUACUGUC
))))).(((.....))))).(((.....))))).(((.....))))).(((.....))))).
UUCACGCAGAAAGCGUCUAGCCAUGGCGUUAGUAUGAGUGUCGUGCAGCCUCCAGGACCCCC
.(((((((.....((((((.....))))).(((.....))))).(((.....))))).
CCCUCCCGGAGAGCCAUAUGUGGUCUGCGGAACCGGUGAGUACACCGGAAUUGCCAGGCGAC
((((.....))))).(((.....))))).(((.....))))).(((.....))))).
CGGGUCCUUUCUUGGAUAAACCCGCUCAAUGCCUGGGAUUUGGGCGUGCCCCCGCAAGACUG
..((((.....))))).(((.....))))).(((.....))))).(((.....))))).
CUAGCCGAGUAGUGUUGGGUCGCGAAAGGCUUGUGGUACUGCCUGAUAGGGUGCUUGCGAGU
((((.....))))).(((.....))))).(((.....))))).(((.....))))).
GCCCCGGAGGUCUCGUAGACCGUGCAUCAUG
```

od 0,2 do 0,5 U RNáza T1 (Pharmacia), RNáza V1 (Pharmacia),
RNáza T2 (Bethesda Research Laboratories)
pufr: 10 mM Tris (pH 7,5), 10 mM MgCl₂, 50 mM KCl
50 U S1 nukleáza (Pharmacia)
pufr: 10 mM Tris (pH 7,5), 10 mM MgCl₂, 50 mM KCl, 1 mM ZnSO₄
pokojová teplota

2 CSFV (Fletcher a Jackson, 2002)

```
.....(((((((.....((((((.....))))).(((.....))))).
CUAGGGCACCCUCCAGCGACGGCCGAAAUGGGCUAGCCAUGCCCAUAGUAGGACUGCAAAC
))))).(((((((.....((((((.....))))).(((.....))))).
GGAGGGACUAGCCGUAGUGGCGAGCUCUCCUGGGUGGUCUAAGUCCUGAGUACAGGACAGUCG
((.....))))).(((.....))))).(((.....))))).(((.....))))).
UCAGUAGUUCGACGUGAGCADUAGCCCACCUCGAGAUGCACGUGGACGAGGGCAUGCCCAA
))))).(((.....))))).(((.....))))).(((.....))))).(((.....))))).
GACACACCUUAACCCUGGCGGGGUCGCUAGGGUGAAAUCACAUUAUGUGAUGGGGGUACGA
((((.....))))).(((.....))))).(((.....))))).(((.....))))).
CCUGAUAGGGUGCUGCAGAGGCCACUAGCAGGCUAGUAUAAAAAUCUCUGCUGUACAUGGC
.....
ACAUGGAGUUGAAUCAUUUGA
```

RNáza A, RNáza T1, kethoxal

přítomna pseudouzlová struktura

3 HCV (Honda et al., 1999)

```
.....(((((((.....((((((.....))))).(((.....))))).
AUCACUCCCCUGUGAGGAACUACUGUCUUCACGCAGAAAGCGUCUAGCCAUGGCGUUAGUAU
))))).(((.....))))).(((.....))))).(((.....))))).
GAGUGUCGUGCAGCCUCCAGGACCCCC
```

RNáza V1 (Pharmacia)
pufr: 10mM MgCl₂-10 mM Tris (pH 7,6)
RNáza S1 (Pharmacia)
pufr: 10mM MgCl₂-10 mM Tris (pH 7,6), 1 mM ZnSO₄
pokojová teplota
čas: 10 min.

4 HCV (Kolupaeva *et al.*, 2000)
.....(((((((.....(((.....((((.....((((.....)))))).....))))))
CUCCCCUGUGAGGAACUACUGUCUUCACGCAGAAAGCGUCUAGCCAUGGCGUUAGUAUGAGU
).....)))))).....(((((((.....((((.....((((.....)))))).....))))))
GUCGUGCAGCCUCCAGGACCCCCCUCGCGGAGAGCCAUAGUGGUCUGCGGAACCGGUCAG
.....)))))).....(((((((.....((((.....((((.....)))))).....))))))
UACACCGGAAUUGCCAGGACGACCGGUCCUUCUUGGAUCAACCCGCUCAUUGCCUGGAGA
).....)))))).....(((((((.....((((.....((((.....)))))).....))))))
UUUGGGCGUGCCCCCGCAAGACUGCUAGCCGAGUAGUGUUGGUCGCGAAAGGCCUUGUGGU
).....)))))).....(((((((.....((((.....((((.....)))))).....))))))
ACUGCCUGAUAGGGUGCUUGCGAGUGCCCCGGAGGUCUCGUAGACCGUGCACCAUGAGCAC
.....
GAAUCCUAAACCU

0,0007, 0,00105 U/ml RNáza V1 (Amersham Pharmacia Biotech),
0,015, 0,025 U/ml RNáza T1 (Amersham Pharmacia Biotech)
pufr: 20 mM Tris-HCl (pH 7,5), 2,5 mM octan hořečnatý, 100 mM
KCl, 2 mM dithiothreitol
teplota: 37°C
čas: 10 min.

přítomna pseudouzlová struktura

5 HCV (Wang *et al.*, 1995)
.....)))))).....)))))).....)))))).....)))))).....)))))).....))))))
GCUUGCGAGUGCCCCGGAGGUCUCGUAGACCGUGCAUCAUG

40-100 U RNáza PhyM (USB), 0,1-0,4 U RNáza V1 (USB), 0,4-1 U
RNáza T1 (USB), 100-400 RNáza A (USB)
pufr: 10 mM Tris (pH 7,5), 10 mM MgCl₂, 50 mM KCl
pokojová teplota
čas: 15 min. (pro RNázu PhyM 30 min.)
1 µl DMS (Aldrich)
pufr: 100 mM Hepes (pH 7,5), 10 mM MgCl₂, 50 mM KCl
pokojová teplota
čas: 5 min.

přítomna pseudouzlová struktura

6 BVDV (Yu *et al.*, 2000)
((((((((((.....)))))))))).....
GUAUACGAGAAUUAGAAAAGGCACUCGUAUACAUUAUG

0,5, 1 ml DMS (Fluka)
pufr: 20 mM HEPES-KOH (pH 7,9), 60 mM KCl, 12 mM MgCl₂
teplota: 30°C
čas: 5 min.
210, 420 mg CMCT (Sigma, Deisenhofen, Germany)
pufr: 80 mM borát sodný (pH 8,1), 60 mM KCl, 12 mM MgCl₂
teplota: 30°C
čas: 10 min.

PICORNAVIRIDAE

7 HAV (Brown *et al.*, 1991)
((((((((((((((.....)))))))))))))).....((((((((.....
UUCAAGAGGGGUCUCCGGGAAUUCCGGAGUCCUCUUGGAAGUCCAUGGUGAGGGGACUUG
.....)))))).....((((.....)))).
AUACCUCACCGCCGUUUGCCUAGGGUAUAGGCUAAAUUUUCCCUUCCCUUUUCCCUUCCU
.....((((.....)))).
AUUCCCUUUGUUUGCUUGUAAAUAUUAUUCUGCAGGUUCAGGGUUCUAAAUCUGUUUC
(.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
UCUAUAAGAACACUCAUUUUUCACGCUUUCUGUCUUCUUCUUCUCCAGGGCUCCCCUUGCCCU
(.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
AGGCUCUGGCCAUUGCGCCCGGGGUC AACUCCAUGAUUAGCAUGGAGCUGUAGGAGUCU
.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
AAAUUGGGGACACAGAUCUUUGGAACGUCACCUUGCAGUGUUAACUUGGCUUUCAUGAAUCU
(.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
CUUUGAUCUUCACAAAGGGGUAGGCUACGGGUGAAACCUCUAGGCUAUACUUCUAUGAAG
.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
AGAUGCCUUGGAUAGGGUAACAGCAACAGAUUAUUGGUGAGUUGUUAAGACAAAACCAUUCA
(.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
ACGCCGGAGGACUGACUCUCAUCCAGUGGAUGCAUUGAGUGGAUUGACUGUCAGGGCUGUCU
.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
UUAGGCUAAAUCAGACCUCUCUGUGCUUAGGGCAAACAUCAUUUGGCCUAAAUGGGAUU
(.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
CUGUGAGAGGGGAUCCCUCCAUGACAGCUGGACUGUUCUUGGGGCCUUAUGUGGUGUUUG
(.....)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).)))).
CCUCUGAGGUACUCAGGGGCAUUUACCUUUUCCUCAUUCUAAAUAUAUAUG

od 0,2 do 0,5 U RNáza T1 (Pharmacia), RNáza V1 (Pharmacia),
RNáza T2 (Bethesda Research Laboratories)
pufr: 10 mM Tris (pH 7,5), 10 mM MgCl₂, 50 mM KCl
50 U S1 nukleáza (Pharmacia)
pufr: 10 mM Tris (pH 7,5), 10 mM MgCl₂, 50 mM KCl, 1 mM ZnSO₄
pokojová teplota

výskyt pseudouzlových struktur

RNáza V1, RNáza ONE, DMS, CMCT

10 poliovirus (Najita a Sarnow, 1990)

.....(((((((.....((((((.....)))))).....)))))).....
CCCGGUGAUGUCGUAUAGACUGCUUGCGUGGUUGAAAGCGACGGAUC

0,2 µl α-sarcin

pufr: 5 mM Hepes (pH 7,6), 25 mM KCl, 2 mM MgCl₂, 0,1 mM EDTA,
3.8% (vol/vol) glycerol, 2 mM dithiothreitol
čas: 30 min.

0,0016 U/µl RNáza T1

pufr: 5 mM Hepes (pH 7,6), 25 mM KCl, 2 mM MgCl₂, 0,1 mM EDTA,
3.8% (vol/vol) glycerol, 2 mM dithiothreitol

teplota: 30°C

čas: 5 min.

DMS

11 poliovirus (Pilipenko *et al.*, 1989)

((((((((((.....((((((.....((((((((.....((((.....((((
CUCAUGAUCUUCGAGAAGCCUAGUAUCGCUCUGGAAUCUUCGACGCGUUGCGCUCAGCACUC
.....)))))))).))))((((((((.....((((((((.....)))))))))
AACCCCGGAGUGUAGCUUGGGCCGAUGAGUCUGGACAGUCCCCACUGGCGACAGUGGUCCAG
)).....))))))((((((((.....))))))))).....)))))).....
GCUGCGCUGGGCGGCCACCUGUGGGCCCAAAGCCACGGGACGCUAGUUGUGAACAGGGUGUGA
.....)))))).....)))))).....((((.....((((.....((((.....((((
AGAGCCUAUUGAGCUACAUGAGAGUCCUCCGGCCCCUGAAUGCGGCUAUAUCCUAACCAUGGA
.....)))))).....)))))).....)))))).....)))))).....
GCAGGCAGCUGCAACCCAGCAGCCAGCCUGUCGUAACGCGCAAGUCCGUGGCGGAACCGACU
.....)))))).....)))).....((((.....)))))).....)))))).....
ACUUUGGGUGUCCGUGUUUCCUUUAUUCUUGAAUGGCUGCUUAUGGUGACAAUCAUAGAUAU
)))))))).)))).....))))
GUUAUCAUAAAGCGAGUUGGA

25, 50 ng RNáza CV1 (the Academy of Sciences of the Estonian
SSR, Tallin)

pufr: 200 mM NH₄Cl, 10 MgCl₂, 20 mM Tris-HCl (pH 7,5)

teplota: 37°C

čas: 10 min.

2 U nukleáza S1 (Pharmacia)

pufr: 50 mM NaCl, 5 mM MgCl₂, 2 mM ZnCl₂ . 3 NH₄Cl, 40 mM
acetát sodný (pH 4,7)

teplota: 37°C

čas: 10 min.

0-0,25% DMS

pufr: 100 mM kakodylát sodný (pH 7,0), 20 mM MgCl₂

teplota: 37°C

čas: 20 min.

12 EMCV (Pilipenko *et al.*, 1989)

(((((.....))))).))....(((.....((((((((.....)))))).....
GGAAGGGUAUGUGUUGCCCUUCCUUCUUGGAGAACGUGCGCGGCGGUCUUCCGUCUCUCG
.....)))))))).))..((((.....(((.....(((.....(((.....(((.....(((
ACAAGCGCGUGCAACAUACAGAGUAACGCGAAGAAAGCAGUUCUCGGUCUAGCUCUAGUG
.....((((.....(((.....))))((((.....(((.....))))((((.....))))
CCCACAAGAAAACAGCUGUAGCGACCACACAAAGGCAGCGGAACCCCCCUCCUGGUAACAGG
))))).)))((.....))..((((.....)))))))).))..((((.....))))
AGCCUCUGCGGCCAAAAGCCACGUGGAUAAGAUCACCUUUGUGUGCGGUGCAACCCCAGCA
))))).))))))....))))))....))))..))))..))))..))))..))))))((
CCUGGUUUCUUGGUGACACUCUAGUGAACCCUGAAUGGCAAUCUCAAGCGCCUCUGUAGG
((((.....(((.....(((.....(((.....)))))))).))))..((((.....(((
GAAGCCAAGAAUUCAGGAGGUACCCCUUCUCUCGGAAGGGAUCUGACCUGGAGACACAUC
((((.....))))..))))..))))..))))..))))..))))..))))..))))..))))
ACACGUGCUUACACCGUGCUUGUGUUUAAAAUUGUUACAGCUUCCCCGAACCAAGUGGU
))))))
CUUGGUU

25, 50 ng RNáza CV1 (the Academy of Sciences of the Estonian
SSR, Tallin)

pufr: 200 mM NH₄Cl, 10 MgCl₂, 20 mM Tris-HCl (pH 7,5)

teplota: 37°C

čas: 10 min.

2 U nukleáza S1 (Pharmacia)

pufr: 50 mM NaCl, 5 mM MgCl₂, 2 mM ZnCl₂ . 3 NH₄Cl, 40 mM
acetát sodný (pH 4,7)

teplota: 37°C

čas: 10 min.

0-0,25% DMS

pufr: 100 mM kakodylát sodný (pH 7,0), 20 mM MgCl₂

teplota: 37°C

čas: 20 min.

13 HAV (Shaffer *et al.*, 1994)

((((((((((.....)))))))))).....((((.....
UUCAAGAGGGGUCUCCGGGAAUUCCGGAGUCCUCUUGGAAGUCCAUGGUGAGGGGACUUG
.....)))))).....((((.....)))).....
AUACCUCACCGCCGUUUGCCUAGGCUAUAGGCUAAAUUUUCCCUUCCCUUUUCCCUUCCU
.....((((.....(((.....))))((((.....(((.....))))((((.....))))
AUUCCCUUUGUUUUGCUUGUAAAUUAAUUCUGCAGGUUCAGGGUUCUAAAUCUGUUUC
.....))))..))))..))))..))))..))))..))))..))))..))))..))))
UCUAUAAGAACACUCAUUUUUCACGCUUUCUGUCUUCUUCUUCUCCAGGGCUCUCCCCUUGCC
((((.....(((.....)))))))).))))..))))..))))..))))..))))..))))..))))
CUAGGCUCUGGCCGUUGCGCCCCGGCGGGGUC AACUCCAUGAUUAGCAUGGAGCUGUAGGAGU
((((.....))))
CUAAAUUGGGGAC

RNáza T1 (Pharmacia), RNáza T2 (BRL), RNáza V1 (Pharmacia)

pufr: 10 mM MgCl₂-10 mM Tris (pH 7,6)

RNáza S1 (Pharmacia)

pufr: 10 mM MgCl₂-10 mM Tris (pH 7,6), 1 mM ZnSO₄
pokojová teplota

přítomny pseudouzlové struktury

14 poliovirus (Skinner *et al.*, 1989)

.....((((((((.....))))))))).....((((.....((((
UUAAAACAGCUCUGGGGUUGUUCCCACCCCAGAGGCCACGUGGCGCUAGUACACUGGUAU
.....)))))).....
CACGGUACCUUUGUACGCCUGUUUAUACUCCUCCCGCAACUUAGAAGCAUACAAUUCA
.....((((.....((((((((.....))))))))).....
AGCUCAAUAGGAGGGGUGCAAGCCACCGCCUCCGUGGGCAAGCACUACUGUUUCCCGGUG
.....((((.....((((((((.....))))))))).....
AGGCCGCAUAGACUGUCCCACGGUUGAAAGUGGCCGAUCCGUUAUCCGCUCAUFUACUUCG
.....((((.....((((((((.....))))))))).....
AGAACCCUAGUAUCGGUCUGGAAUCUUCGACGCGUUGCGCCUCAGCACUCAACCCCGGAGUG
.....((((.....((((((((.....))))))))).....
UAGCUUGGCCCGAUGAGUCUGGACAGUCCCCACUGGGGACAGUGGUCCAGGUUUGCGUGGCG
.....((((.....((((((((.....))))))))).....
GCCACCUGUGGCCCAAAGCCACGGGACGCUAGUUGUGAACAGGGUGUGAAGAGCCUAUUGA
.....((((.....((((((((.....))))))))).....
GCUACAUGAGAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAACCAUGGAAGCAGGCAGCUC
.....)))))).....
CACCCAGCAGCCAGCCUGUCGUAACGCGCAAGUCCGUGGCGGAACCGACUACUUUGGUGUC
.....((((.....((((((((.....))))))))).....
CGUUUCCUUUUAUUCUUGAAUGGCUGCUUAUGGUGACAAUCAUAGAUUGUUAUCAUAAAG
)
C

0,5-50 U/ml RNáza T1 (Pharmacia, England), 5-500 U/ml RNáza T2
(Pharmacia, England), 8-800 U/ml RNáza CV1 (Pharmacia,
England)

pufr: 30 mM Tris HCl (pH 7,8), 20 mM MgCl₂, 300 mM KCl
čas: 30 min.

0-3% DMS (BDH, England), 0,1-0,5 mg/ml kethoxal (United States
Biochemical Corporation, U.S.A.)

pufr: 70 mM Hepes (pH 7,4), 10 mM MgCl₂, 270 mM KCl
teplota: 30°C
čas: 5 min.

NIDOVIRALES

15 BCoV (Raman *et al.*, 2003)

.....((((((((.....))))))))).....
CACUCCCUGUAUUCUAUGCUUGUGGGCGUAGAUUUUUCAUAG

0,05, 0,5, 1, 2, 4 U RNáza CV1 (Pharmacia), 0,05, 0,5, 1, 5,
10 RNáza T1 (Gibco), 0,05, 1, 2, 3, 4 U RNáza T2 (Gibco)

pufrr: 30 mM Tris HCl (pH 7,5), 20 mM MgCl₂, 300 mM KCl
teplota: 25°C
čas: 15 min.

16 BCoV (Raman a Brian, 2005)
...((((((((((.....))))))..)))).....
CUAGGCAGUGGCCGACCCAUAAGGUCACAAUGUCGAAGAUAACA

0,05, 0,5, 1, 2, 4 U RNáza CV1 (Ambion), 0,05, 0,5, 1, 5, 10
URNáza T1 (Gibco), 0,05, 1, 2, 3, 4 U RNáza T2 (Gibco)
pufrr: 30 mM Tris HCl (pH 7,5), 20 mM MgCl₂, 300 mM KCl
teplota: 25°C
čas: 15 min.

17 EAV (van den Born et al., 2004)
.....((((((((((((((((((((.....))))))))))))))))))))((((((((.....)))
GCUCGAAGUGUGUAUGGUGCCAUAUACGGCUCACCACCAUAUACACUGCAAGAAUUACUAUU
))))))....((((((.....)))..))))..((((.....)))..((((.....)))..((((.....)))
CUUGUGGGCCCCUCUCGGUAAAUCCUAGAGGGCUUUCUCUCGUUAUUGCGAGAUUCGUCGU
.....)))))).....((((.....)))..((((.....)))..((((.....)))..((((.....)))
UAGUAACGGCAAGUCCCUUUCUAUCCUAUUUUAUCUUGUGGCUUGACGGGUCACU
((((.....)))..))))))..))))))..))))))..))))))..))))))..))))))..))))))
GCCAUCGUCGUCGAUCUCUAUCAACUACCCUUGCGACUAUGGCAACCUUCUCCGCUACUGGA
.....)))))).....((((.....)))..))))))..))))))..))))))..))))))..))))))
UUUGGAGGGAGUUUUGUUAGGGACUGGUCUCCUGGACUUAACCCGACGCUUGUGAGCAUGGCGC
))
GGG

0,00005, 0,0001, 0,0005 U RNáza A (Amersham Pharmacia Biotech)
0,1, 0,5 U RNáza T1 (Invitrogen), 0,01, 0,05, 0,1 RNáza T2
(Invitrogen), 0,0005, 0,001, 0,005 RNáza V1 (Kemetex Bio Ltd.)
pufrr: 30 mM Tris-HCl (pH 7,5), 20 mM MgCl₂, 300 mM KCl
teplota: 37°C
čas: 30 min.
0,25, 0,5, 0,75, 1% DMS (Fluka)
pufrr: 80 mM kakodylát sodný (pH 7,2), 5 mM MgCl₂, 100 mM KCl
teplota 37°C
čas: 10 min.
4,2, 8,4, 12,6, 16,8 mg/ml CMCT (Aldrich)
pufrr: 50 mM borát sodný (pH 8,0), 2,5 mM MgCl₂, 12,5 mM KCl
teplota 37°C
čas: 10 min.

DISTROVIRIDAE

18 PSIV (Shibuya a Nakashima, 2006)

.....((((.....)))...((((.....))))).((((.....)))...
GAGUUGGAUCUAGAAGAUCUAUGCGCUGGCAUAGUAAUCCAGUGUUUCCUGUUUUAGGAUGA
..((((.....)))...((((.....)))...((((.....)))
CCUCUGAAAGUAGAUGACCGUGGAAAGUCACGUAGUGCCCCAAUAAGCACGUUUGGGCAGCG
)...((((.....)))...((((.....)))...((((.....)))...
UGCGCUAUCACAAGGCUUGAUCUCCGAGGAGCCCCUUGUUUAGCUGGCUGGAAGCCAAUGA
.....((((.....)))...((((.....)))...((((.....)))
UCUUAAGUAGUAAAGUGCUGUUGCUUGUAGUUAACAGAAAGCUUUGAGUACGUCUUUCUUG
..))))))...((((.....)))...((((.....)))...((((.....)))
CGAGAAAGAACAUGCAUUCUUAUGCUCUCAAUUCUAUUUUUUUUUUUGGGCGAAAGGA
..((((.....)))...((((.....)))...((((.....)))
AAGCUCUCACGCGAGUACGAAUAGCCAACCCUUUAUAUGAUGUUUUCUCUUAUUCUUUAAA
.....
UAGCCGGACUGAUUUUACA

5 U RNáza T1 (Ambion)

pufr: 10 mM Tris-HCl (pH 7,5), 1 mM MgCl₂, 50 mM KCl

teplota: 37°C

čas: 6 min.

40 mM DMS

pufr: 50 mM kakodylát draselný (pH 7,2), 2 mM MgCl₂, 100 mM KCl

teplota: 20°C

čas: 30 min.

20 mM CMCT

pufr: 50 mM borát sodný (pH 7,4), 2 mM MgCl₂, 100 mM KCl

teplota: 20°C

čas: 40 min.

přítomna pseudouzlová struktura

ROSTLINNÉ VIRY

20 PVX (Miller et al., 1998)

.....((((.....)))...((((.....))))).((((.....)))...
GAAAACUAAACCAUACACCACCAACACAACCACCACCCACGCCAAUUGUUACACACCCG
((((.....)))...((((.....)))...((((.....)))...
CUUGAAAAGCAAGUCUGACAAAUGGCCAAAGUGCGGAGUUUACCAAUCCUUUACAGACU
.....((((.....)))...((((.....)))...((((.....)))
CCACCACAAAACUCUCAUCCAAGAUGAGGCUUAUAGAAAUAUUCGUCCCAUCAUGGAAAAA
.....((((.....)))...((((.....)))...((((.....)))
CAUAAACUAGCUAACCCGUACGCUCAACGGUUGAAGCGGCUAA

0,04 U/ml RNáza T1, 0,0014 U/ml RNáza V1

pufr: 100 mM Tris-HCl (pH 7,6), 100 mM NaCl, 20 mM KCl

teplota: 25°C
čas: 15 min., 10 min.
0-5 mM DMS
pufr: 80 mM draselný kakodylát (pH 7,2), 300 mM KCl, 20 mM
octan hořečnatý
teplota: 40°C, 37°C, 90°C
čas: 60 min, 20 min., 2 min.

21 TEV (Zeenko a Gallie, 2005)

```
.....((((((.....))))))..  
AAAUAAcAAAUUCUCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUA  
.....((((((.....)))))).....((((((.....))))))....  
CUUCUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAGCAAUUUUCUGAAAAUUUUCA  
.....  
CCAUUUACGAACGAUAGCAAUG
```

RNáza T1, RNáza T2, RNáza A, RNáza S1, RNáza V1

přítomny pseudouzlové struktury

19 TBSV (Wu et al., 2001)

```
((((((((((..((((.....)))))).....((((.....  
GGAAUUCUCCAGGAUUUCUCGACCUAGUUCGUUUAUCUGGUGACUUGCGCUACCGUUGC  
))))))))))))))  
GCGUAGAGAAUUUCU
```

0,0125 U/μl RNáza T1

pufr: Tris-HCl (pH 7,5), 5 mM MgCl₂, 1 mM EDTA, 200 mM NaCl

teplota: 25°C

čas: 1 min.

0,0175 U/μl RNáza V1

pufr: Tris-HCl (pH 7,5), 5 mM MgCl₂, 1 mM EDTA, 200 mM NaCl

teplota: 25°C

čas: 1 min.

2 μl (50 mg/ml) CMCT

pufr: Tris-HCl (pH 8), 5 mM MgCl₂, 1 mM EDTA, 200 mM NaCl

teplota: 25°C

čas: 15 min.

10% DEPC

pufr: Tris-HCl (pH 7,5), 5 mM MgCl₂, 1 mM EDTA, 200 mM NaCl

teplota: 25°C

čas: 20 min.

4 µl kethoxal
pufr: TNE buffer
15 µl CMCT
pokojová teplota
čas: 5 min.

24 HIV 2 (Berkhout, 1992)

(((((((((.....))))))))).(((((((.GGUCGCUCUGCGGAGAGGCUGGCAGAUUGAGCCCUGGGAGGUUCUCUCCAGCACUAGCAGGU.((((.....))))))))).(((((((.....))))))))).AGAGCCUGGGUGUCCCUGCUAGACUCUCACCAGCACUUGGCCGGUGCUGGGCAGACGGCCC
.....
CACGCU

0, 0,1, 0,3, 0,9 U RNáza S1, 0, 0,001, 0,003, 0,01 U RNáza CV,
0, 0,001, 0,003, 0,01, 0,033, 0,1 U RNáza T1
0, 0,5, 1, 2% DEP

25 HIV 2 (Berkhout a Schoneveld, 1993)

(((((((((.....))))))))).(((((((.GGUCGCUCUGCGGAGAGGCUGGCAGAUUGAGCCCUGGGAGGUUCUCUCCAGCACUAGCAGGU.((((.....))))))))).(((((((.....))))))))).AGAGCCUGGGUGUCCCUGCUAGACUCUCACCAGCACUUGGCCGGUGCUGGGUAGACGGCCC
(((((((.....))))))))).(((((((.....))))))))).CACGCUUGCUUGCUUAAAACCUCUUAUAAGCUGCCAGUUAGAAGCAAGUUAAGUGUGUG
.....(((((((.....))))))))).(((((((.....))))))))).CUCCCAUCUCUCCUAGUCGCCGCCUGGUCAUUCGGUGUUCACCUGAGUAACAAGACCCUGGU
(((((((.....))))))))).(((((((.....))))))))).CUGUUAGGACCCUUCUUGCUUUGGAAACCGAGGCAGGAAAUCCCUAGCAGGUUGGCGCCU
.....(((((((.....))))))))).(((((((.....))))))))).GAACAGGGACUUGAAGAAGACUGAGAAGUCUUGGAACACGGCUGAGUGAAGGCAGUAAGGGC
)))))).....
GGCAGGAACAAACCACGA

0, 0,001, 0,005, 0,0025 U RNáza T1, 0, 0,001, 0,003, 0,0010 U
RNáza CV, 0, 0,1, 0,3, 0,9 U nukleáza S1
teplota: 37°C
čas: 5 min.
0, 1-2, 4 µl DEP, 0, 4 µl kethoxal, 0, 0,6, 1,2, 1,8 µl DMS

26 HIV 1 (Clever et al., 1995)

.....(((((((.....))))))))).(((((((.UCGACGCCAGGACUCGCUUGCUGAAGCGCGCACGGCAAGAGGGCGGCGCACUUGGUGA
)).))))).(((((((.....))))))))).GUACGCCAAAAUUUGACUAGCGGAGGCUAGAAGGAGAGAGAUGGGUGCGAGAGCGUCAGU
.....
AUUAAGCGGGGGAG

od 0,02 do 0,1 U RNáza T1 (Pharmacia)
teplota: 0°C
čas: 20 min.
od 0,15 do 0,7 U Rnáza V1 (Pharmacia)
teplota: 0°C
čas: 20 min.
od 0,5 do 1% DMS (Aldrich)
teplota: 0°C
čas: 20 min.
od 1,5 do 3% karbodiimid (Sigma)
teplota: 25°C
čas: 30 min.
od 4 do 8 mg kethoxal (U.S. Biochemical)
teplota: 0°C
čas: 60 min.
pufr: 30 mM HEPES (pH 7,5), 50 mM KCl, 10 mM ZnCl₂, 2 mM
dithiothreitol

27 HIV 1 (Damgaard *et al.*, 1998)

(((((.....)))))).....(((.....))....
GGUCUCUCUGGUUAGACCAGAUUUGAGCCUGGGAGCUCUCUGGCUAACUAGGGAACCCACUG
.....(((.....))....
CUUAAGCCUCAAUAAAGCUUGCCUUGAGUGCUUCAAGUAGUGUGGCCGUCUGUUGUGUGA
.....(((.....))....
CUCUGGUAACUAGAGAUCCCUCAGACCCUUUUAGUCAGUGUGGAAAUCUCUAGCAGUGGCG
.....(((.....))....
CCCGAACAGGGACUUGAAAGCGAAAGGGAAACCAGAGGAGCUGUCUCGACGCAGGACUCGGC
.....(((.....))....
UUGCUGAAGCGCGCACGGCAAGAGGGCGAGGGGCGGCGACUGGUGAGUACGCCAAAAUUUUG
.....(((.....))....
ACUAGCGGAGGCUAGAAGGAGAGAGAUGGGUGCGAGAGCGUCAGUAUUAAGCGGGGA

12 U/ml RNase T1, 200 U/ml RNase T2, 350 U/ml RNase V1
teplota: 0°C
čas: 20 min.

28 HIV 1 (Harrison a Lever, 1992)

.....(((.....))....
GCAGUGGCGCCCGAACAGGGACCUGAAAGCGAAAGGGAAACCAGAGCUCUCUCGACGCAGGA
.....(((.....))....
CUCGGCUUGCUGAAGCGCGCACGGCAAGAGGGCGAGGGGCGGCGGUCUGGUGAGUACGCCAAAA
.....(((.....))....
AUUUUGACUAGCGGAGGCUAGAAGGAGAGAGAUGGGUGCGAGAGCGUCAGUAUUAAGCGGGG
.....
GAGAAUUGGAUCGAUGG

od 0 do 1 U RNáza V1 (Pharmacia)
teplota: 0°C

čas: 30 min.
od 0 do 50 U RNáza T1 (Boehringer, Lewes, England)
teplota: 0°C
čas: 30 min.
od 0,05 do 0.3% DMS (BDH, Poole, England)
teplota: 20°C
čas: 10 min.
kethoxal (United States Biochemical) od 10 do 35 mg/ml ve 20%
ethanol
teplota: 20°C
čas: 10 min
psoralen - UV 370 nm
pokojová teplota
čas: 10 min.
pufr: 30 mM Tris HCl (pH 7,8), 20 mM MgCl₂, 300 mM KCl

29 MPMC (Harrison *et al.*, 1995)

.....((((.....))).....((((.....
UGGCGCCAACGUGGGGUGGAUACGAGGGAAUUUCGUGAGGAAGACGACGCGUUCGCCGGCC
))))((.....))((.....((((.....))))))((.....))
GGCGAUUAAAAGUGAAAGUAAACUCUCUUGGCCGCCGCGGGAACCGCGUUGGACCUGA
.).....((((.....)))).....((((.....)).....).
AAGUAAGUGUUGCGCUCGGAUAUGGGGCAAGAAUUAAGCCAGCAUGAACGUUAUGUAGAACA
.....)))).).....((((((((((((((((.....)))))))))).....)))))
AUUGAAGCAGGCUUAAAAGACACGGGGAGUAAAGGUUAAAUAUGCUGAUCUUUUGAAUUUU
)).....((((.....))))
UUGAUUUUGUGAAGGAUACUUGUCCUU

od 0 do 1 U RNáza V1 (Pharmacia)
teplota: 0°C
čas: 30 min.
od 0 do 50 U RNáza T1 (Boehringer Mannheim)
teplota: 0°C
čas: 30 min.
od 0,05 do 0.3% DMS (Fluka)
teplota: 20°C
čas: 10 min.
kethoxal (United States Biochemical) od 10 do 35 mg/ml ve 20%
ethanolu
teplota: 20°C
čas: 10 min.
psoralen (Sigma)
UV 366 nm
pokojová teplota
čas: 10 min
pufr: 30 mM Tris HCl (pH 7,8), 20 mM MgCl₂, 300 mM KCl

CCCGAACAGGGACUUGAAAGCGAAAGGGAAACCAGAGGAGCUCUCUCGACGCAGGACUCGGC
((((((.....))))))...(((.....))).....
UUGCUGAAGCGCGCACGGCAAGAGGGCGAGGGGCGGCGACUGGUGAGUACGCCAAAAUUUUG
.....
ACUAGCGGAGCCUAGAAGGAGAGACAUGGGUG

3 microl Me₂SO₄ (Acros Organics)
teplota: 37°C
čas: 4, 8 min.

34 SNV (Roberts a Lawrie, 2003)

vlásenka A

((((((((.....)))))))))
GGUCGCCGUCCUACACAUGUUGUUGUGACGUGCGGCC

vlásenka C

((((.....))))))
GAUUUUGUUCGUGGUGUUGGCUCGCCUACUGGGUGGGCGCAGGGAUCCGGACUGAAUCC

RNáza T1 (Ambion), RNáza A (Ambion), RNáza V1 (Ambion)
pufr: 100 mM Tris (pH 7), 7 M KCl, 100 mM MgCl₂
pokojová teplota
čas: 15 min.