

Obsah

1	Úvod	1
2	Teoretická část	2
2.1	Metodologie	2
2.1.1	Molekulární dynamika	2
2.1.2	Silová pole RNA	4
2.1.3	Výpočty volné energie - termodynamická integrace	5
2.1.4	Výpočty pK_a	6
2.1.5	Molekulární dynamika za konstantního pH	9
2.2	Ribozymy	13
2.2.1	Vlásenkový ribozym	14
3	Praktická část	17
3.1	Cíle	17
3.2	Metody	18
3.3	Výsledky a diskuze	24
3.3.1	Parametrizace - termodynamická integrace	24
3.3.2	Titrace 9-methylguaninu	25
3.3.3	Aplikace na větší systémy	27
3.3.4	Vliv počtu iontů v boxu	31
3.3.5	Vliv velikosti boxu	33
3.4	Závěry	39
4	Literatura	41

1 Úvod

RNA je jedním ze dvou v současné době známých druhů nukleových kyselin. Na základě tzv. centrálního dogmatu molekulární biologie, které formuloval Francis Crick v roce 1958, bylo dlouhou dobu předpokládáno, že hlavní úlohou RNA v živých organismech je sloužit jako přenašeč genetické informace mezi sekvencí DNA a proteiny, přičemž DNA genetickou informaci uchovává a proteiny jsou interpretací genetické informace a jako jediné mají v organismu biologickou funkci. Za jedinou další možnou úlohou RNA v organismech pak bylo považováno uchovávání genetické informace v RNA virech.¹ Tento pohled na RNA byl však radikálně změněn v roce 1982, kdy Thomas Cech objevil ribozym intron skupiny I v bakterii *Tetrahymena thermophila*.²

Tento objev znamenal převratné zjištění. A sice, že některé sekvence DNA jsou přepisovány na RNA sekvence, a ačkoli u nich již dále neprobíhá proces translace, tj. nejsou dále přepisovány na proteiny, mají svoji biologickou úlohu. Takovéto sekvence RNA se začaly označovat jako nekódující RNA. Zjistilo se tedy, že nejen pouze proteiny, ale také molekuly RNA mohou mít biologickou funkci, tedy např. katalyzovat chemické reakce. Od té doby byla objevena celá řada nekódujících RNA sekvencí s biologickou funkcí. Jedním z představitelů této skupiny jsou ribozymy.³

Ribozymy neboli RNA enzymy jsou katalyticky aktivní molekuly či sekvence RNA, což znamená, že jsou schopny katalyzovat chemickou reakci, nejčastěji štěpení vlastní cukr-fosfátové páteře. Jejich objev znamenal průlom v molekulární biologii, protože do té doby byla schopnost enzymatické katalýzy považována za výhradní vlastnost proteinů. V mnoha organismech plní ribozymy životně důležitou funkci.⁴

Na mechanismu katalýzy ribozymů se významným způsobem nukleobáze v jejich nekanonickém protonačním stavu. Ty mají jednak přímý vliv na průběh katalýzy, a také mohou ovlivnit konformaci aktivního místa nebo konformaci celé molekuly RNA.⁵ Pro pochopení mechanismu katalýzy je proto vhodné studovat závislost strukturní dynamiku ribozymů na protonačních stavech nukleových bází. Výpočetní metodou, která umožňuje takovéto studium je molekulární dynamika za konstantního pH od Grubmüllera et al.⁶

Tato metoda je v současné době otestována pro proteiny a její aplikovatelnost na RNA není přímočará. Cílem této práce je otestovat funkčnost této metody pro molekuly RNA za použití silového pole AMBER ff99bsc0 χ_{OL3} , a to na vlásenkovém ribozymu.

2 Teoretická část

2.1 Metodologie

2.1.1 Molekulární dynamika

Molekulární dynamika (MD) je výpočetní metoda, která umožňuje počítat vývoj molekulového systému v čase. Využívá aproximace, kdy atomy jsou považovány za hmotné body a interakce mezi nimi jsou popsány pomocí empirických potenciálů. Systém pak lze popsat pomocí klasické mechaniky a vývoj systému v čase lze simulovat pomocí numerické integrace Newtonových pohybových rovnic po dostatečně malých časových krocích.⁷

Vzájemné interakce všech částic v systému lze popsat empirickým interakčním potenciálem. Ten je funkcí polohy všech částic, tedy jejich souřadnic. Sílu F působící na jednu konkrétní částici pak lze získat jako záporně vzatou derivaci potenciálu E podle souřadnice x této částice (rovn. 1), ze síly F a hmotnosti m částice lze podle druhého Newtonova zákona získat zrychlení a působící na tuto částici (rovn. 2) a konečně z aktuálního zrychlení $a(t)$ (a aktuální polohy $x(t)$ a rychlosti $v(t)$, respektive $v\left(t - \frac{1}{2}\Delta t\right)$) jsme schopni spočítat změnu rychlosti a polohy částice za určitý časový okamžik Δt , respektive hodnotu rychlosti v čase $t + \frac{1}{2}\Delta t$ $v\left(t + \frac{1}{2}\Delta t\right)$ (rovn. 3) a polohu částice v čase $t + \Delta t$ $x(t + \Delta t)$ (rovn. 4), kde Δt je hodnota integračního kroku.

$$F = -\frac{\partial E}{\partial x} \quad (1)$$

$$F = ma \quad (2)$$

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + a(t)\Delta t \quad (3)$$

$$x(t + \Delta t) = x(t) + v\left(t + \frac{1}{2}\Delta t\right)\Delta t \quad (4)$$

V klasické MD je poloha atomů je popsána pomocí Kartézských souřadnic. Výsledkem simulace spočítané pomocí MD je takzvaná trajektorie, což jsou souřadnice a rychlosti jednotlivých atomů v jednotlivých časových krocích.⁷⁻⁸

Empirický interakční potenciál E_{pot} je součtem potenciálů popisujících jednotlivé vazebné a nevazebné interakce v systému. Tvar těchto potenciálů je určen parametry, které jsou empiricky nastaveny tak, aby co nejpřesněji popsaly studovaný systém. Jejich parametrizace se provádí na malých molekulách pomocí fitu na přesná kvantově chemická data popř. experimentální data, např. vibrační frekvence vazeb, solvatační energie nebo elektrostatický potenciál okolo molekuly spočtený kvantově chemickými metodami. Nejčastější forma empirického potenciálu pro popis molekul je tvořena pěti termy (rovn. 5): energie spojená s vibrací kovalentních vazeb E_{bond} , energie vibrací vazebných úhlů E_{angle} , energie spojená s rotací okolo kovalentních vazeb tj. energie torzního úhlu E_{dihedral} , elektrostatické interakce $E_{\text{electrostatics}}$ a van der Waalsova interakce E_{vdw} , která je nejčastěji popsána Lenard-Jonesovým potenciálem.⁷

$$E_{\text{pot}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}} + E_{\text{electrostatics}} + E_{\text{vdw}} \quad (5)$$

$$E_{\text{bond}} = \sum \frac{1}{2} k_r (r - r_0)^2 \quad (6)$$

$$E_{\text{angle}} = \sum \frac{1}{2} k_\alpha (\alpha - \alpha_0)^2 \quad (7)$$

$$E_{\text{dihedral}} = \sum \frac{1}{2} k_\varphi (1 + \cos(n\varphi - \varphi_0)) \quad (8)$$

$$E_{\text{electrostatics}} = \sum \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{q_i q_j}{r_{ij}} \quad (9)$$

$$E_{\text{vdw}} = \sum \left[-2\epsilon_{ij} \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 + \epsilon_{ij} \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} \right] \quad (10)$$

Energie kovalentní vazby mezi atomy je nejčastěji popsána rovnicí harmonického oscilátoru (rovn. 6). Rovnovážná poloha vazby r_0 a konstanta tuhosti k_r jsou parametry v tomto termu a závisí na typu vazby, čili pro každou dvojici atomových typů je v empirickém poli definována rovnovážná poloha a konstanta tuhosti případné vazby, r je aktuální délka vazby daná aktuální polohou částic v simulaci. Obdobně je popsána i deformace vazebného úhlu (rovn. 7), kde α_0 je rovnovážná poloha úhlu, k_α konstanta tuhosti a α aktuální poloha úhlu. Energie torzního neboli dihedrálního úhlu je kvůli jeho

periodickému charakteru popsána goniometrickou funkcí (zpravidla funkcí kosinus) popř. součtem několika kosinů, které tvoří část Fourierovy řady (rovn. 8). Amplitudy k_φ , periody n a fáze φ_0 těchto kosinových termů jsou opět jako parametry empirického pole definovány pro každý torzní úhel, φ je aktuální poloha úhlu. Elektrostatické interakce jsou popsány jako interakce atomově centrovaných bodových parciálních nábojů pomocí Coulombova zákona (rovn. 9), kde q_i a q_j jsou hodnoty nábojů, r_{ij} je jejich vzdálenost, ϵ_0 je permitivita vakua a ϵ_r relativní permitivita. Van der Waalsovy interakce jsou popsány pomocí Lennard-Jonesova potenciálu skládajícího se z repulzní a disperzní části (rovn. 10). V Lennard-Jonesově potenciálu vystupují pro každou dvojici atomů dva empirické parametry, a to poloha minima σ_{ij} (součet van der Waalsových poloměrů obou atomů) a hloubka tohoto minima ϵ_{ij} . V praxi se tyto párové parametry mixují z atomových parametrů. Proměnnou v tomto termu je pak aktuální vzdálenost středů dvojice atomů r_{ij} .⁷

Rovnice popisující závislost potenciálu na poloze částic mají tedy pro jednotlivý typ interakce shodný tvar, ale liší se parametry. Parametry závisejí na druhu atomů, typu vazby a uskupení atomů v molekule. Aby nemusely být stanovovány parametry pro každou jednotlivou interakci a zvláště pro každou sloučeninu, předpokládá se určitá přenositelnost těchto parametrů mezi různými chemickými systémy tj. např. methylová skupina bude popsána stejnými Lennard-Jonesovými parametry uhlíku a vodíků a stejnými vazebnými parametry vazby C-H a úhlu H-C-H bez ohledu na to, v jaké konkrétní sloučenině se nachází. Tento popis vede na definici tzv. atomových typů, které rozlišují atomy nejen jako prvky, ale i podle typu jejich chemického okolí, např. alifatický uhlík, aromatický uhlík atd. Parametry se pak stanovují pro tyto atomové typy, u vazeb pro dvojice atomových typů atd. Soubor rovnic, kterými se počítají potenciály, a soubor parametrů pro jednotlivé atomové typy se nazývá silové pole.⁷

2.1.2 Silové pole RNA

V současnosti neexistuje žádné univerzálně použitelné silové pole, vhodnost daného pole závisí především na účelu použití.⁹ Pro MD simulace molekul RNA se nejčastěji používají¹⁰ silová pole AMBER (často nazývané i Cornellová et al.)¹¹ a CHARMM.¹²⁻¹³ Silová pole, která nebyla naparametrizována pro nukleové kyseliny, není vhodné používat, protože je vysoce nepravděpodobné, že by fungovala správně.¹⁴

Silové pole AMBER existuje v několika variantách, a to ff94, ff98, ff99, ff10, popřípadě verze ff02, ff03 nebo ff12, které nepředstavují novou verzi v popisu

nukleových kyselin. První verzi tzv. druhé generace silových polí AMBER byla verze ff94. Její reparametrizací, konkrétně změnou parametrů glykosidické χ torze a puckeru ribosy, která zlepšila stočení B-DNA do helixu, vzniklo pole ff98 a další zlepšení puckeru ribosy vedlo k poli ff99, které bylo po dlouhou dobu považováno za dostatečné pro popis nukleových kyselin. Při delších simulacích se však ukázalo, že toto silové pole způsobovalo při simulačních časech desítek ns nevratný přechod konformace páteře DNA, konkrétně změny torzí α (O3'-P-O5'-C5') a γ (O5'-C5'-C4'-C3') s typickou orientací úhlu γ v trans stavu. Tento přechod měl tendenci se kumulovat ve struktuře dvoušroubovice DNA a progresivně degradovat její tvar v průběhu simulace. Proto byla provedena reparametrizace torzních úhlů α a γ , tzv. reparametrizace parmbsc0.¹⁵⁻¹⁶ Podobná degradace kanonických struktur dvoušroubovic byla zaznamenána i v případě RNA, kdy na stovkách ns byla pozorována ztráta helikálního twistu A-RNA dvoušroubovic, které pak přešly do struktury podobné žebříku.¹⁷ Proto byla vyvinuta modifikace χ_{OL3} , která je navržena specificky pro simulace RNA.¹⁸ Tato modifikace je obsažena v poli ff10, což je standardní AMBER pole od roku 2010.¹⁵ Ukazuje se, že vylepšení silového pole pro RNA nemusí být přenositelná přímo do popisu DNA a obráceně¹⁵, proto byla vyvinuta také χ modifikace pro DNA nazvaná χ_{OL4} .¹⁹

Silové pole CHARMM má dvě varianty určené pro nukleové kyseliny, CHARMM27¹³ a novější CHARMM36.²⁰ Ani pro jednu z těchto variant CHARMM silového pole neexistuje dostatečně rozsáhlé množství testujících simulací pro popis RNA a z dostupných testů se zdá, že ani jedna varianta nepopisuje RNA správně a obě vedou mimo jiné k rozpadu struktury kanonické A-RNA dvoušroubovice.^{15, 21-22}

2.1.3 Výpočty volné energie - termodynamická integrace

Gibbsovu volnou energii lze počítat třemi základními metodami, metodou „free energy perturbation“, metodou distribuce pravděpodobnosti, nebo-li histogramovou metodou a metodou termodynamické integrace, která bude popsána podrobněji.²³

Termodynamická integrace je založena na postupném přeměňování systému z jednoho stavu do druhého. První stav, stav A, je popsán Hamiltoniánem H_A a druhý stav, stav B, je popsán H_B . Zvolíme párující parametr λ a sestrojíme Hamiltonián $H(\lambda)$, který je funkcí λ a pro který platí $H(\lambda=0) = H_A$ a $H(\lambda=1) = H_B$. Nejjednodušší vyjádření takového Hamiltoniánu má lineární tvar mixování:

$$H(\lambda) = (1 - \lambda)H_A + \lambda H_B \quad (11)$$

Rozdíl Gibbsovy volné energie ΔG_{BA} mezi stavy B (kterému přísluší G_B) a A (G_A) je pak dán vztahem:

$$\Delta G_{BA} = G_B - G_A = \int_0^1 \frac{dG(\lambda)}{d\lambda} d\lambda = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (12)$$

kde výraz $\langle \dots \rangle_{\lambda}$ znamená průměrování hodnoty uvnitř závorek, v tomto případě derivace energie podle λ , tj. záporně vzaté síly působící na parametr λ , přes soubor konformací získaný v simulaci při dané hodnotě λ . Systém je přeměňován ze stavu A do stavu B po diskretních krocích, tedy po určitých diskretních krocích se mění hodnota λ . Integrant z rovnice (12) je spočítán numericky z MD simulací provedených při dané hodnotě λ a integrál se počítá např. lichoběžníkovou metodou.²⁴

2.1.4 Výpočty pK_a

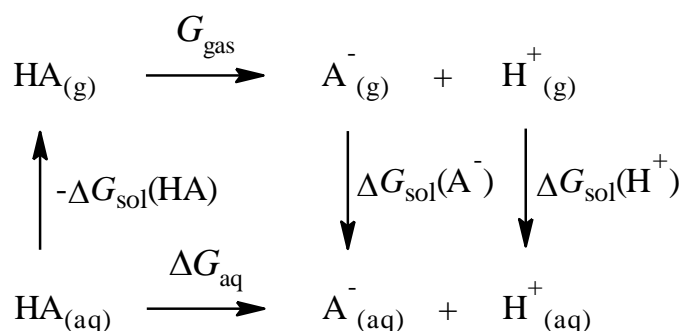
Disociační konstanty kyselin, které se častěji udávají jako záporně vzatý dekadický logaritmus této konstanty tj. pK_a , jsou nezbytné pro pochopení mnoha základních reakcí v chemii a biochemii. Většinou lze pK_a změřit velmi snadno experimentálně, avšak často nás také zajímají pK_a sloučenin, které dosud nebyly syntetizovány, a nebo experimentální měření není přímočaré. Například je velmi těžké určit hodnotu pK_a konkrétní aminokyseliny zabudované v polypeptidovém řetězci. Hodnoty pK_a volných aminokyselin a nukleobází jsou experimentálně změřené a tabelované, nicméně je známo, že tyto pK_a hodnoty mohou být ovlivněny strukturním kontextem proteinu²⁵ či RNA/DNA²⁶ a mohou být v biomolekulách posunuty až o několik jednotek. Proto je velmi důležitá schopnost spočítat hodnotu pK_a s vysokou přesností výpočetními metodami.²⁵

Základem pro výpočet pK_a je výpočet rozdílu Gibbsovy volné energie ΔG nedisociované a disociované formy dané molekuly v daném prostředí. Z tohoto rozdílu lze pak pK_a získat podle vztahu:

$$pK_a = -\log_{10} K_a = \frac{\Delta G}{RT \ln 10} \quad (13)$$

kde K_a je disociační konstanta, R je molární plynová konstanta a T termodynamická teplota.

Rozdíl Gibbsovy volné energie ΔG lze počítat různými metodami. Většina metod rozděluje proces disociace kyseliny, a tedy i výpočet ΔG na několik mezikroků. Rozdělení reakce na mezikroky lze popsat termodynamickým cyklem. Nejčastěji se používá takovéto rozdělení:



Obrázek 1 Jeden z termodynamických cyklů využívaných pro rozdělení Gibbsovy volné energie disociace kyseliny ve vodném prostředí ΔG_{aq} na složky G_{gas} (Gibbsova volná energie disociace v plynné fázi), $\Delta G_{\text{sol}}(\text{HA})$ (Gibbsova volná energie solvatace nedisociované kyseliny HA), $\Delta G_{\text{sol}}(\text{A}^{-})$ (Gibbsova volná energie solvatace deprotonované kyseliny, tj. její konjugované báze A^{-}) a $\Delta G_{\text{sol}}(\text{H}^{+})$ (Gibbsova volná energie solvatace protonu H^{+}).

Potom se počítá zvlášť Gibbsova volná energie reakce v plynném stavu a zvlášť Gibbsova volná energie solvatace jednotlivých částic uvedených v termodynamickém cyklu.

Všechny hodnoty Gibbsovy volné energie lze spočítat pomocí kvantové chemie, kromě $\Delta G_{\text{sol}}(\text{H}^{+})$, kterou lze spočítat jen velmi obtížně a k tomu s velkou nepřesností, a proto musí být určena experimentálně.

Výpočet Gibbsovy volné energie disociace v plynné fázi je velmi přesný za předpokladu, že je použita dostatečně přesná kvantově chemická metoda uvažující korelaci elektronů. Naproti tomu největším zdrojem chyb je výpočet Gibbsovy volné energie solvatace. Tu lze pomocí metod kvantové chemie spočítat za použití dvou různých modelů solventu, explicitního a implicitního.

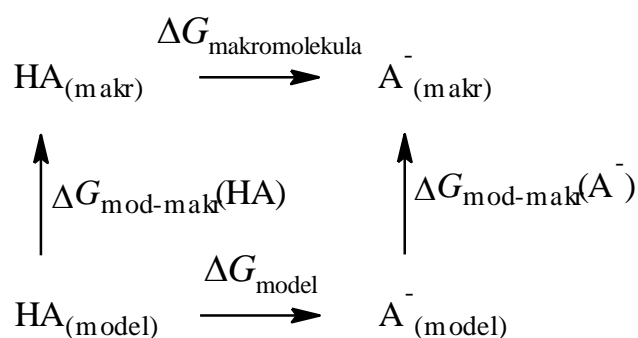
Explicitní model solventu popisuje solvent a jeho interakce v systému explicitně na úrovni atomů a jeho termodynamické vlastnosti jsou pak spočteny metodami statistické termodynamiky. Zahrnutí dostatečného počtu konfigurací solutu a solventu, které by byly schopny věrně reprezentovat jejich makroskopické vlastnosti, je největší obtíž tohoto přístupu. Naproti tomu implicitní model solventu popisuje solvent jako kontinuum a předpokládá, že požadované vlastnosti lze spočítat z rovnic popisujících jeho makroskopický stav. Většina metod používajících implicitní model solventu je založena na Poisson-Boltzmannově rovnici nebo na Bornově teorii solvatace. Ačkoliv

tyto metody zanedbávají explicitní popis interakce solventu se solutem, tak při vhodné parametrizaci mohou mít relativně dobré výsledky.^{25, 27}

Pokud nepotřebujeme spočítat absolutní hodnotu pK_a , ale relativní posunutí pK_a hodnoty vlivem strukturního kontextu například v biomolekule, lze využít molekulové mechaniky (MM) a spočítat pouze vliv okolního prostředí na titrovatelnou skupinu a případný posun její rovnováhy mezi disociovanou a nedisociovanou formou. Za využití MM se některou z metod uvedených v kapitole 1.1.3 spočte rozdíl Gibbsovy volné energie mezi protonovanou a deprotonovanou formou modelové molekuly $\Delta G^{\text{FF}}_{\text{model}}$, jejíž pK_a je známo, a rozdíl Gibbsovy volné energie mezi protonovanou a deprotonovanou formou analogické molekuly zabudované v makromolekule $\Delta G^{\text{FF}}_{\text{makr}}$. $\Delta G^{\text{FF}}_{\text{model}}$ ani $\Delta G^{\text{FF}}_{\text{makr}}$ nemají fyzikální význam, protože MM nezahrnuje všechny příspěvky Gibbsovy volné energie reálného systému. Předpokládá se ovšem, že nezahrnuté příspěvky Gibbsovy volné energie jsou v obou případech velmi podobné, takže by se v případě rozdílu $\Delta G^{\text{FF}}_{\text{makr}} - \Delta G^{\text{FF}}_{\text{model}}$ (označený jako $\Delta\Delta G^{\text{FF}}$) měly vyrušit a tento rozdíl by měl odpovídat skutečnému rozdílu $\Delta\Delta G^{\text{real}}$ Gibbsovy volné energie deprotonace modelové molekuly $\Delta G^{\text{FF}}_{\text{model}}$ a Gibbsovy volné energie deprotonace molekuly v makromolekule $\Delta G^{\text{FF}}_{\text{makr}}$:

$$\Delta\Delta G^{\text{FF}} = \Delta G^{\text{FF}}_{\text{makr}} - \Delta G^{\text{FF}}_{\text{model}} \approx \Delta G^{\text{real}}_{\text{makr}} - \Delta G^{\text{real}}_{\text{model}} = \Delta\Delta G^{\text{real}} \quad (14)$$

Za obdobných předpokladů lze $\Delta\Delta G^{\text{real}}$, respektive $\Delta\Delta G^{\text{FF}}$ spočítat tak, že se pomocí MM spočítá Gibbsova volná energie přeměny protonované formy modelové sloučeniny ve vodě na protonovanou formu molekuly v makromolekule $\Delta G_{\text{mod-makr}}(\text{HA})$ a Gibbsova volná energie přeměny deprotonované formy modelové sloučeniny ve vodě na deprotonovanou formu molekuly v makromolekule $\Delta G_{\text{mod-makr}}(\text{A}^-)$ a $\Delta\Delta G^{\text{real}}$, respektive $\Delta\Delta G^{\text{FF}}$ se pak spočítá pomocí termodynamického cyklu (Obrázek 2).^{6, 28}



Obrázek 2 Termodynamický cyklus používaný pro spočítání rozdílu $\Delta\Delta G$ Gibbsovy volné energie deprotonace modelové molekuly ΔG_{model} ve vodě a Gibbsovy volné energie deprotonace molekuly v makromolekule $\Delta G_{\text{makromolekula}}$, kde $\Delta G_{\text{mod-makr}}(\text{HA})$ je Gibbsova volná energie přeměny protonované formy modelové sloučeniny ve vodě na protonovanou formu molekuly v makromolekule a $\Delta G_{\text{mod-makr}}(\text{A}^{-})$ Gibbsova volná energie přeměny deprotonované formy modelové sloučeniny ve vodě na deprotonovanou formu molekuly v makromolekule.

Krom toho je možné posun $\text{p}K_{\text{a}}$ a hlavně sprážením posunu $\text{p}K_{\text{a}}$ se strukturní dynamikou biomolekuly vypočítat na MM úrovni v takzvané molekulární dynamice za konstantního pH,^{5-6, 29} která bude popsána dále.

2.1.5 Molekulární dynamika za konstantního pH

Hodnota pH je velmi důležitý parametr v systémech v kapalné fázi. Ovlivňuje protonační stavy titrovatelných skupin a tudíž i strukturu, dynamiku a funkci molekul v roztoku. Klasické MD simulace ovšem neumožňují zadat pH přímo jako jeden ze vstupních parametrů a už z podstaty charakteru empirických silových polí nepopisují změny protonačních stavů titrovatelných skupin. Pro simulace se používají struktury s fixními protonačními stavy, a to takovými, jaké jsou při daném pH nejpravděpodobnější. Tento přístup s sebou přináší dvě komplikace. Jednak ne vždy známe $\text{p}K_{\text{a}}$ konstantu dané titrovatelné skupiny v makromolekule tj. neznáme její posun vlivem strukturního kontextu této skupiny. Hodnotu této $\text{p}K_{\text{a}}$ konstanty lze získat jen poměrně náročnými experimenty, např. pomocí NMR spektroskopie. Navíc i takto naměřené hodnoty posunutých $\text{p}K_{\text{a}}$ konstanty představují pouze zprůměrovanou hodnotu a neumožňují tak popsat změnu protonačního stavu v závislosti na změně konformace makromolekuly. Druhá komplikace je spojená s fundamentálním omezením simulace za konstantního protonačního stavu, jelikož taková simulace nedovoluje správný popis rovnovážné fluktuační změny protonačního stavu. Proto byly vyvinuty metody, které umožňují

zadat hodnotu pH přímo jako jednu ze vstupních podmínek a v MD simulaci pak postihnout změny protonačních stavů. Tyto metody se označují jako molekulární dynamika za konstantního pH (cpH MD z anglického constant pH molecular dynamics).^{5-6, 29}

Metody cpH MD popisují protonační stav každé titrovatelné skupiny proměnnou λ . Tyto metody lze rozdělit do dvou skupin, podle toho, zda proměnná λ nabývá pouze diskrétních hodnot, nebo se může měnit kontinuálně. Do první skupiny patří zejména metody, které popisují protonační stav v MD simulaci pomocí metody Monte Carlo (MC). Změna protonačního stavu je popsána tak, že je v pravidelných časových intervalech během MD simulace proveden MC výpočet, který určí, jaký protonační stav bude mít daná skupina po následující interval. MC výpočet je založen na Gibbsově volné energii protonace dané skupiny, která se počítá při každém MC výpočtu. Přičemž Gibbsovou volnou energií se zde myslí hlavně Gibbsova volná energie solvatace protonované a deprotonované formy titrovatelné skupiny, kdežto rozdíl energie solutu se bere ve formě energie příslušející danému snímku simulace, v němž se provádí MC výpočet. Solvatační Gibbsovu volnou energii každého protonačního stavu lze získat pomocí Poisson-Boltzmannovy rovnice nebo zobecněných Bornových rovnic.⁶

Do druhé skupiny patří zejména metody, ve kterých je proměnná λ řešena jako další stupeň volnosti v systému, který se vyvíjí v čase podle svých pohybových rovnic. Potenciál, ve kterém se tento stupeň volnosti pohybuje, je funkcí pH a energie protonované a deprotonované formy. Síla působící na tento stupeň volnosti a tedy i jeho rychlost a nová hodnota λ se počítají ve stejných časových intervalech jako souřadnice a rychlosti pro všechny ostatní částice.⁶

Některé metody se používají v kombinaci s implicitním solventem, některé s explicitním. Pro simulace molekul RNA je však důležité, aby probíhaly v explicitním solventu, jelikož implicitní solvent popisuje solvataci polyanionického řetězce RNA velmi špatně a vede v simulacích k výrazným artefaktům.³⁰ Jednou z metod propagujících parametr λ v čase a používající explicitní solvent je cpH MD, kterou vyvinuli Grubmüller et al. a implementovali ji⁶ do výpočetního balíku GROMACS.³¹

Vliv pH a strukturní dynamiky systému na protonační stav titrovatelné skupiny popisuje tato metoda následovně. Protonační stav je popsán proměnnou λ . Je-li $\lambda = 0$, je skupina protonována, při $\lambda = 1$ je deprotonována. λ je při MD simulaci definována jako další stupeň volnosti systému s virtuální hmotností m_λ , souřadnicí λ a virtuální rychlostí v_λ . Hamiltonián systému $H(\lambda)$ je definován tak, že parametr λ slouží jako faktor

v lineárním mixování (viz kapitola 1.1.3) Hamiltoniánu protonované formy H_A s Hamiltoniánem deprotonované formy H_B titrovatelné skupiny:

$$H(\lambda) = (1 - \lambda)H_A + \lambda H_B \quad (15)$$

Síla F_λ působící na „částici“ λ je:

$$F_\lambda = -\frac{\partial V(\lambda)}{\partial \lambda} = V_A - V_B \quad (16)$$

$V(\lambda)$ je část Hamiltoniánu představující potenciální energii, V_A potenciální energie systému ve stavu A a V_B potenciální energie systému ve stavu B. Je-li znám potenciál $V(\lambda)$, je možno spočítat sílu F_λ působící na „částici“ λ . Účelem potenciálu $V(\lambda)$ je popsat změnu energie mezi protonovaným a deprotonovaným stavem a zároveň zajistit, aby četnost protonovaných a deprotonovaných stavů odpovídala disociačnímu stupni při daném pH. Protože pouze hodnoty $\lambda = 0$ a $\lambda = 1$ reprezentují stav systému s fyzikálním významem, potenciál $V(\lambda)$ by měl mít takový tvar, aby dynamika tohoto parametru splňovala, že bude dominantně populovat hodnoty okolo $\lambda = 0$ a $\lambda = 1$. Dále hodnota λ se musí pohybovat v intervalu od 0 do 1. Čas strávený v přechodném stavu mezi 0 a 1 by měl být tudíž krátký a čas strávený ve stavu s fyzikálním významem pak dostatečně dlouhý na to, aby mohly proběhnout konformační změny. Potenciál $V(\lambda)$ tedy musí mít mezi $\lambda = 0$ a $\lambda = 1$ bariéru a velikostí této bariéry pak lze řídit frekvence přechodů z jednoho stavu do druhého.

Splnění těchto podmínek je zajištěno následovně. Hodnota proměnné λ je vyjádřena pomocí úhlového parametru θ kružnice o poloměru r ($r=1$) podle rovnice:

$$\lambda = r \cos(\theta) + \frac{1}{2} \quad (17)$$

Dynamika tak probíhá v rámci proměnné θ a síla působící na θ F_θ je rovna:

$$F_\theta = -\frac{\partial V(\lambda(\theta))}{\partial \theta} = r \sin(\theta) \frac{\partial V(\lambda(\theta))}{\partial \lambda} \quad (18)$$

Tím je zajištěno, že λ bude v intervalu od 0 do 1. Dále je k potenciálu působícího na λ přidán takzvaný dodatečný potenciál $U^*(\lambda)$ ve tvaru:

$$U^*(\lambda) = 4h\lambda(1 - \lambda) \quad (19)$$

Ten zajistí, že větší čas bude strávený v hodnotách λ blízkých 0 nebo 1. Výškou potenciálu h lze regulovat frekvenci přechodů mezi stavy.

Skutečná hodnota změny Gibbsovy volné energie ΔG mezi protonovaným a neprotonovaným stavem je rovna:

$$\Delta G = (\ln 10)RT (\text{p}K_a - \text{pH}) \quad (20)$$

kde pH je zvolené pH systému, pK_a je záporný dekadický logaritmus disociační konstanty titrovatelné skupiny, R je molární plynová konstanta a T termodynamická teplota. Proto by pro molekulu, jejíž pK_a známe, měl potenciál $V(\lambda)$ mít tvar:

$$V(\lambda) = [(\ln 10)RT (pK_a - pH)]\lambda + U^*(\lambda) \quad (21)$$

Dá se však očekávat, že potenciál vzniklý mixováním empirických potenciálů protonované a deprotonované molekuly (rovn. 15) nebude mít takový tvar, proto je nutné závislost takto zmixovaného potenciálu na parametru λ spočítat a v cpH MD jej odečíst a místo něj vložit výše zmíněný potenciál (rovn. 21), jak bude popsáno dále.

Změnu Gibbsovy volné energie mezi dvěma stavy či dvěma molekulami lze spočítat na MM úrovni pomocí termodynamické integrace tak, jak byla popsána v kapitole 1.1.3, a to podle rovnice (12). Takto lze spočítat také změnu Gibbsovy volné energie mezi protonovanou a deprotonovanou formou ΔG^{FF} . Jak bylo řečeno, takto spočtená změna Gibbsovy volné energie ovšem nemá fyzikální význam a není rovna skutečné změně Gibbsovy volné energie. V tomto případě však termodynamické integrace využijeme pro spočtení závislosti energie (v popisu empirickým polem) na parametru λ u referenčního systému (typicky volné aminokyseliny nebo nukleobáze) a tuto závislost využijeme pro parametrizaci cpH MD a tyto parametry následně přeneseme do strukturního kontextu biomolekuly. Celkový potenciál má tedy tvar:

$$V(\lambda) = (1 - \lambda)V_A + \lambda V_B - \int_0^\lambda \left\langle \frac{\partial E}{\partial \lambda} \right\rangle_\lambda d\lambda + [(\ln 10)RT (pK_a - pH)]\lambda + U^*(\lambda) \quad (22)$$

Kde první dva členy jsou zmixované energie protonované a deprotonované formy libovolného systému s titrovatelnou skupinou (tj. např. biomolekuly), třetí člen je odečtení závislosti energie na parametru λ spočtené na referenčním systému (např. volná aminokyselina nebo nukleobáze). Čtvrtý člen zajišťuje správnou bilanci mezi protonovanou a deprotonovanou formou, přičemž parametr pK_a je záporný dekadický logaritmus disociační konstanty referenčního systému. A konečně pátý poslední člen je bariéra mezi hodnotami λ 0 a 1 zajišťující dominantní populaci fyzikálně smysluplných hodnot tohoto parametru. Je potřeba zmínit, že pro referenční systém se z principů parametrizace této metody první tři členy odečtou na nulu, kdežto v případě aplikace na jiný systém, např. biomolekulu, bude jejich součet nenulový v důsledku vlivu strukturního okolí na titrovanou skupinu, který vede k posunu jejího pK_a .⁶

Do balíku GROMACS je tato metoda implementována následovně. Molekula s titrovatelnou skupinou má takzvanou dvojakou topologii, což znamená že jsou uvedeny parametry pro molekulu v protonovaném stavu (stav A) a deprotonovaném

stavu (stav B), respektive pro stav B stačí v topologii uvést ty parametry, které se liší od stavu A.³¹ Hamiltonián systému pak závisí na λ podle vztahu (15). Protože během MD simulace nelze měnit počet atomů v systému, je deprotonovaný stav v topologii realizován tak, že atom vodíku je ve stavu B nahrazen atomem s nulovými Lennard-Jonesovými parametry a nulovým nábojem. Parametry ostatních atomů jsou takové, jako by atom vodíku nebyl přítomen.

Výraz $\langle \frac{\partial E}{\partial \lambda} \rangle_{\lambda}$, respektive závislost $\frac{\partial E}{\partial \lambda}$ na λ se získá postupem termodynamické integrace. Spustí se tedy MD s několika fixními hodnotami λ (zpravidla od 0 do 1 po intervalech 0,1). Samotná integrace se ale neprovádí, pouze se graf závislosti $\frac{\partial E}{\partial \lambda}$ na λ proloží kubickou rovnicí.

Jako vstupní parametry pro MD simulaci se pak kromě vstupních hodnot pro klasickou MD zadávají požadované pH, pK_a referenční molekuly a koeficienty kubické rovnice pro závislost $\frac{\partial E}{\partial \lambda}$ na λ spočtené na této referenční molekule.⁶

Metoda tak, jak ji vyvinuli Grubmüller et al., je určena a otestována pro proteiny za využití silového pole GROMOS.⁶ Tato metoda by ale mohla být teoreticky použitelná i pro dynamiku nukleových kyselin za použití silového pole AMBER. To je třeba otestovat. Vhodným modelovým systémem pro tento účel je vlásenkový ribozym.

2.2 Ribozymy

Vlásenkový ribozym patří mezi ribozymy. Ribozymy, nebo-li RNA enzymy jsou molekuly RNA, které mají schopnost katalyzovat chemickou reakci, nejčastěji štěpení své vlastní cukr-fosfátové páteře.⁴ Ribozymy se podílejí na regulaci genové exprese a mají klíčový význam pro životaschopnost některých organismů.³² Podle jedné z teorií existovaly molekuly RNA a ribozymy mnohem dříve než DNA a proteiny, tato hypotéza se nazývá teorie RNA světa.³³ V současné době se mimo jiné zkoumá možné využití ribozymů jako potenciálních cílů léčiv inhibujících jejich aktivitu.³⁴

Podle struktury a reakčního mechanismu lze ribozymy rozdělit obecně do dvou skupin, a to na velké a malé ribozymy. Mezi velké ribozymy patří self-splicing introny skupiny I a skupiny II a ribonukleasa P. Skládají se z několika set až 3000 nukleotidů a jimi katalyzované reakce vedou k produktům s volnou 3'-hydroxylovou a volnou 5'-fosfátovou skupinou.^{4, 32}

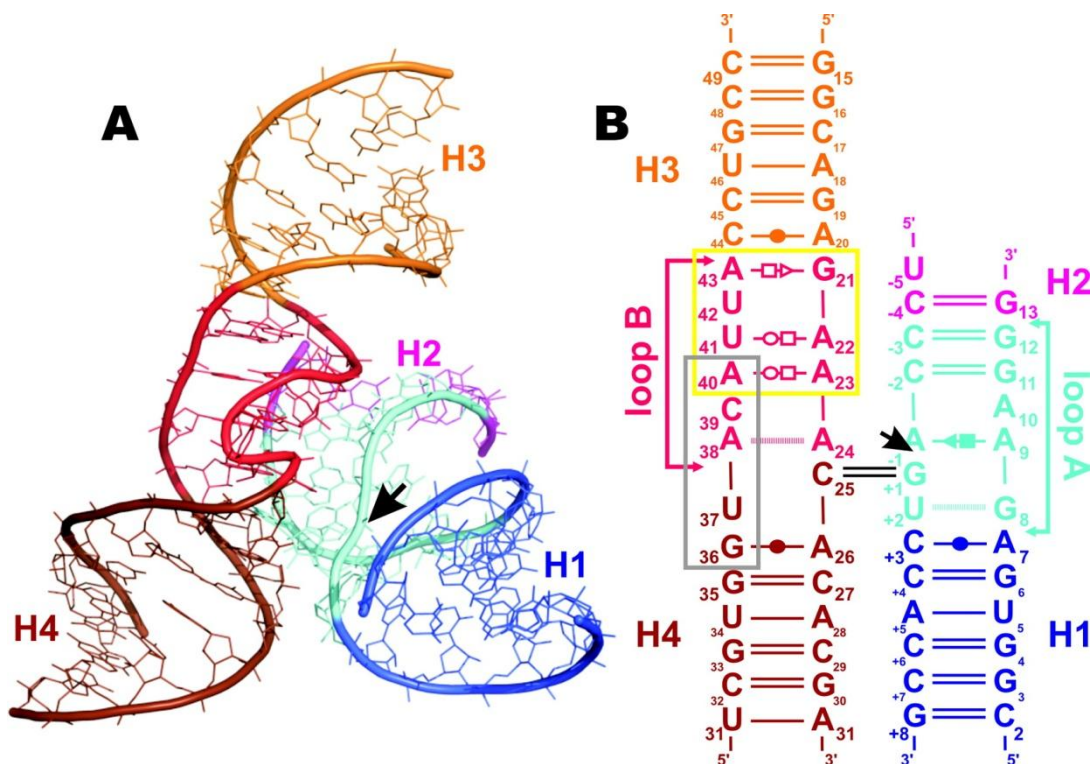
Mezi malé ribozymy patří např. ribozymy vlásenkový, hammerhead, HDV (hepatitis delta virus) a VS (varkud satellite). Mají délku od 30 do 150 nukleotidů a jimi

katalyzované reakce vedou k produktům s volnou 2'-3'-cyklickou fosfátovou a volnou 5'-hydroxylovou skupinou.^{4, 32, 35}

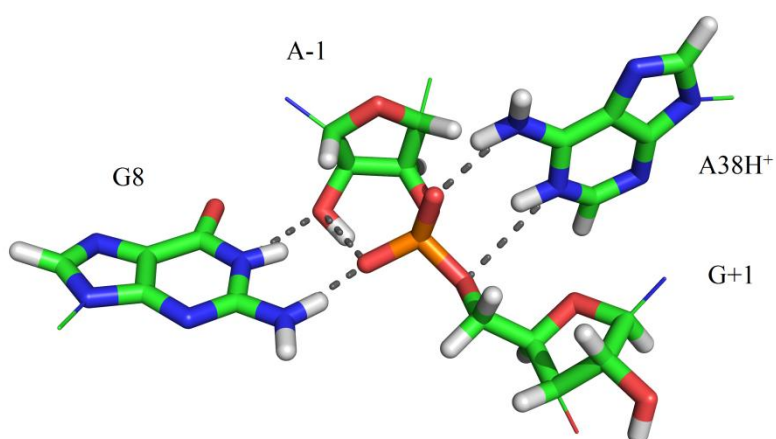
2.2.1 Vlášekový ribozym

Vlášekový ribozym (z anglického hairpin ribozyme) patří do skupiny malých ribozymů. Skládá se ze dvou domén, označovaných A a B. Každá z domén se skládá ze dvou šroubovic, doména A ze šroubovic H1 a H2 a doména B z H3 a H4. Nukleotidy v doméně B jsou číslovány od 1 do 50. Nukleotidy v doméně A od -5 do +9 tak, aby fosfodiesterová vazba, která může být ribozymem rozštěpena, byla mezi nukleotidy -1 a +1 (Obrázek 3). Ke své funkci nepotřebuje žádný dvojmocný kovový kation a je plně aktivní i bez přítomnosti těchto iontů. Nachází se v tzv. minusovém vlákně satelitní RNA přidružené k viru tabákové mozaiky. Má schopnost katalyzovat samoštěpení, tedy odstřížení sebe sama a následné navázání sebe sama, což je ve viru využíváno při replikaci jeho genetické informace tzv. double-circle rolling mechanismem. Odstřížení je zajištěno rozštěpením fosfodiesterové vazby na specifickém místě, vlášekový ribozym tedy katalyzuje tuto reakci. Katalýza je specifická na určité místo své vlastní molekuly RNA. Pro experimentální studium tohoto RNA enzymu je však důležité, že je možné upravit vlášekový ribozym tak, že je rozdělen na dvě RNA vlákna, jedno katalytické a druhé štípané, které vystupuje v roli substrátu tohoto enzymu. Vlášekový ribozym může tedy v této úpravě také katalyzovat štěpení fosfodiesterové vazby na jiné molekule RNA, opět specificky na určitém místě.^{17, 36-37}

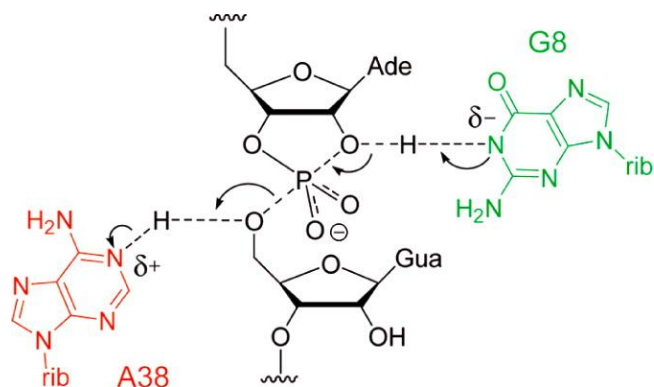
Reakce probíhá v aktivním místě, které tvoří guanin 8, adenin 38, adenin -1 a guanin +1 (Obrázek 4). Toto aktivní místo vzniká při sbalení ribozymu do jeho nativní konformace, přičemž bezprostředně poté dochází k rozštěpení jeho cukr-fosfátové páteře mezi adeninem -1 a guaninem +1. Aktivní místo vlášekového ribozymu je narozdíl od většiny ostatních ribozymů uzavřené, což znamená, že k němu nemají přístup molekuly solventu a iontů z okolního prostředí. Štěpena je fosfodiesterová vazba mezi adeninem -1 a guaninem +1, výsledkem je volná 2'-3'-cyklická fosfátová skupina na adeninu -1 a volná 5'-hydroxylová skupina na guaninu +1.^{17, 37}



Obrázek 3 Struktura vlásenkového ribozymu v úpravě na dvě RNA vlákna, jedno katalytické a druhé štípané. Dvoušroubovice H1-H4 a smyčky jsou barevně rozlišeny, místo štěpení je vyznačeno černou šipkou. (A) 3D struktura. (B) Sekundární struktura s číslováním nukleových bází. Obrázek byl převzat z článku ¹⁷.



Obrázek 4 Aktivní místo vlásenkového ribozymu, nukleové báze G8, A38H⁺ a ribosy, na kterých jsou navázané nukleové báze A-1 a G+1.



Obrázek 5 Znáznornění reakce probíhající v aktivním místě vlásenkového ribozymu, popsány jsou nukleové báze A38 a G8 a zkratky Ade a Gua reprezentují nukleové báze A-1 a G+1. Obrázek byl převzat z článku³⁸.

Přestože byla katalýza vlásenkového ribozymu intenzivně studována experimentálně i teoreticky a její mechanismus se dá považovat za objasněný, stále zůstávají některé jeho aspekty ne úplně objasněny. Biochemická a strukturní experimentální data ukázala, že dvě nukleobáze v aktivním místě jsou klíčové pro katalýzu, a to guanin 8 a adenin 38. Na základě naměřené závislosti rychlostní konstanty štěpení na pH, strukturních a mutačních dat a teoretických MD simulací a QM/MM výpočtů se jeví jako nejpravděpodobnější mechanismus, kde reakci začíná nekanonický deprotonovaný guanin 8, který funguje jako akceptor protonu od 2'-OH skupiny adeninu -1, a tedy jako obecná báze. 2'-OH skupina v reakci funguje jako nukleofil a je tudíž deprotonovaným guaninem 8 aktivována pro nukleofilní ataku štípaného fosfátu. Při nukleofilní substituci O5' kyslíku za atakující O2' kyslík na štípaném fosfátu je odstupující O5' skupina protonována nekanonickým protonovaným adeninem 38, který funguje jako obecná kyselina (Obrázek 5). Z dostupných teoretických simulací a experimentů se však zdá, že zatímco pK_a adeninu 38 je posunuto strukturním kontextem aktivního místa k neutralitě a jeho přítomnost v protonované formě je důležitá pro stabilitu aktivního místa, guanin 8 nemá významně posunutou pK_a a jeho přítomnost v deprotonované formě spíše není aktivním místem tolerována.^{17,37}

3 Praktická část

3.1 Cíle

Protonační stavy nukleobází v aktivních místech katalyticky aktivních molekul RNA mají jednak přímý vliv na samotnou katalýzu, ale také i vliv na strukturní dynamiku aktivních míst těchto RNA molekul, která ovlivňuje katalytickou aktivitu nepřímo.³⁹ Velmi účinným nástrojem pro studium strukturní dynamiky těchto molekul jsou MD simulace, které umožňují detailní náhled na dynamické chování aktivních míst RNA enzymů a studium strukturní podstaty některých jevů.¹⁰ Proto by bylo velmi užitečné, pokud by v MD simulacích bylo možno popsat také dynamické změny protonačních stavů. Tato možnost by mohla pomoci objasnit například mechanismus katalýzy nebo závislost katalytické aktivity na pH.³⁹

Metod, které umožňují v MD simulaci v explicitním solventu popsat dynamickou změnu protonačních stavů, existuje v současnosti několik, ovšem extenzivně byly testovány pouze pro klasické enzymy tj. pro proteiny.⁶ Metoda, která by umožnila spolehlivě popsat dynamickou změnu protonačních stavů v molekulách RNA ve standardně dlouhých (tj. několik μ s) MD simulacích, však dosud nebyla dostatečně otestována. (Metodu cpH MD v explicitním solventu pro molekuly RNA popisují Brooks et al.³⁹, ti ovšem využívají silového pole CHARMM36, jehož funkčnost pro RNA je zatím značně diskutabilní, přičemž některé práce poukazují na významné nedostatky tohoto silového pole.²²) Některé z metod vyvinutých pro proteiny jsou navrženy tak, aby byly použitelné univerzálně. Proto se jeví jako vhodný postup otestovat funkčnost některé z těchto metod původně navržených a testovaných pro proteiny i na molekulách RNA.

Cílem této práce bylo zjistit, zda metoda cpH MD vyvinutá Grubmüllerem et al. a otestovaná na proteinech za použití silového pole GROMOS⁶, je použitelná také pro RNA za použití silového pole AMBER ff99bsc0 χ_{OL3} ^{11, 18, 40}. Silové pole AMBER bylo zvoleno proto, že jiná silová pole zatím pro RNA nefungují spolehlivě.²² V případě pole ff99bsc0 χ_{OL3} je tak nejpravděpodobnější, že bude nejen schopno správně popsat strukturní dynamiku RNA, ale i dostatečně přesně popsat ne vazebné interakce mezi titrovanou skupinou a zbytkem RNA molekuly, které determinují případný posun pK_a titrované skupiny v rámci strukturního kontextu RNA. Jako struktura pro otestování této metody byl zvolen vlásenkový ribozym, který je vhodný proto, že tato metoda by mohla podstatným způsobem přispět k objasnění mechanismu jím katalyzované reakce.

V průběhu řešení této práce se však ukázalo, že v případě vlásenkového ribozymu použitá metoda cpH MD významně přeceňovala posun pKa, což může signalizovat dosud neidentifikovanou nepřesnost v poli ff99bsc0 χ OL3. Abychom lépe prostudovali podstatu selhání metody cpH MD v případě vlásenkového ribozymu, provedli jsme analogické výpočty i u menších systémů, a to dinukleotidu GpG (guanylyl(3'→5')guanosin) a guanosinu. Tyto simulace pak byly použity pro systematické hledání, kde dochází v popisu dynamiky protonačních stavů k případným odchýlkám od očekávaného chování.

3.2 Metody

Pro cpH MD simulace byl použit balík GROMACS 3.3.3⁴¹ s implementovanou metodou cpH MD od Grubmüllera et al.⁶ verze 1.3.

Pro startovní strukturu vlásenkového ribozymu byla použita krystalová struktura s označením 2OUE z databáze Protein Data Bank (PDB), která byla získána Röntgenovou strukturní analýzou monokrystalu s rozlišením 2,05 Å.⁴² Pomocí programu tLeaP z balíku AMBER⁴³ byly doplněny atomy vodíku a struktura byla umístěna do boxu s explicitním modelem vody TIP3P. Vzdálenost stěn boxu od solutu byla nastavena na 10 Å a minimální vzdálenost solventu od solutu na 0,7 Å. Výsledná velikost boxu činila 6,92 x 8,87 x 8,59 nm³. Náboj struktury byl vykompenzován přidáním příslušného počtu K⁺ iontů ($R = 1,705$ Å a $\epsilon = 0,1936829$ kcal/mol⁴⁴) do boxu. A za použití silového pole AMBER ff99bsc0 χ OL3 byly programem tLeaP vytvořeny soubor s koordináty a soubor s topologií (obojí ve formátu pro AMBER). Programem sander z balíku AMBER pak byla provedena minimalizace a ekvilibrace struktury včetně pomalého vyhřátí na teplotu 298,15 K. A pomocí programu acpype byly koordináty s ekvilibrovanou geometrií a topologie převedeny do formátu pro GROMACS.

Obdobným způsobem byla vytvořena topologie pro vlásenkový ribozym s nekanonickým, tedy deprotonovaným guaninem G8 na dusíku N1. Parametry tohoto guaninu byly převzaty z článku¹⁷. Z obou topologií vlásenkového ribozymu s kanonickým i deprotonovaným guaninem 8 ve formátu GROMACSu pak byla vytvořena tzv. dvojaká topologie pro cpH MD. Tato dvojaká topologie obsahovala jak parametry pro vlásenkový ribozym s kanonickým guaninem 8, tak informace o odlišnostech topologie vlásenkového ribozymu s deprotonovaným guaninem 8, hlavně

pak odlišné náboje a van der Waalsovy parametry guaninu 8 v obou jeho protonačních formách.

Jak již bylo uvedeno, během MD simulace nesmí být měněn počet atomů v systému, proto je deprotonovaný stav realizován tak, že je na místě G8(H1) vodíku atom s nulovými Lennard-Jonesovými parametry a s nulovým nábojem. Navíc celkový náboj systému s deprotonovaným G8 by byl o jeden elementární náboj menší než v případě kanonického guaninu. Nicméně během MD simulace se rovněž nesmí měnit celkový náboj systému, který navíc musí zůstat kvůli konvergenci Ewaldovy sumace používané pro výpočet dalekodosahové elektrostatiky nulový. Proto je náboj systému s deprotonovaným G8 kompenzován kontinuální nábojovou hustotou rovnoměrně rozmístěnou po celém boxu. Vzhledem k velikosti boxu a faktu, že parametrizace metody cpH MD byla prováděna na 9-methylguaninu umístěném v přibližně stejně velkém boxu, by tato nábojová hustota měla mít na chování systému minimální vliv.⁶

Pro explicitní otestování, zda je tento vliv zanedbatelný či nikoli, byl použit také druhý způsob kompenzace záporného náboje systému s nekanonickým G8, a to pomocí Cl⁻ iontu spárovaného s protonačním stavem G8. Pokud je G8 v protonovaném stavu, tento Cl⁻ je v systému přítomen, a pokud je G8 v deprotonovaném stavu, tento Cl⁻ není přítomen. To je realizováno tak, že Cl⁻ má obdobně jako další atomy, jejichž parametry závisejí na protonačním stavu G8, uveden v dvojité topologii parametry pro oba protonační stavy a mění se tedy spolu s parametry titrovaného guaninu. Je-li G8 protonovaný, Cl⁻ má být v systému přítomen, v tomto stavu (stavu A) jsou tedy v topologii uvedeny standardní parametry pro Cl⁻ ion (R = 2,513 Å a $\epsilon = 0,0355910$ kcal/mol⁴⁴), je-li G8 deprotonovaný, Cl⁻ nemá být v systému přítomen, což je realizováno tak, že v tomto stavu (stavu B) jsou namísto Cl⁻ iontu parametry atomu s nulovými Lennard-Jonesovými parametry a nulovým nábojem. Celkový náboj systému je tedy konstantní a nezávislý na protonačním stavu G8.

Topologie a koordináty tohoto systému byly vytvořeny s obdobnými parametry jako pro systém bez Cl⁻ a obdobným způsobem pomocí programu tLEaP. V tomto případě ale byla struktura neutralizována Na⁺ ionty (R = 1,369 Å a $\epsilon = 0,0874393$ kcal/mol⁴⁴), a navíc abychom otestovali i další možnou příčinu selhání metody cpH MD, byla použita modifikace $\epsilon_{\zeta_{OL1}}$ silového pole AMBER.⁴⁵ Výsledná velikost boxu činila 8,35 x 9,21 x 9,27 nm³.

Protože parametry atomů Cl⁻ i nukleobáze G8 závisejí na proměnné λ spojující oba protonační stavy, stav A se stavem B (rovn. 15), a zároveň na sebe Cl⁻ a G8 působí

elektrostatickými silami a silami popsanými Leonard-Jonesovým potenciálem, není žádoucí, aby se k sobě mohly přiblížit. Proto byl pro všechny MD simulace tohoto systému nastaven dodatečný potenciál mezi atomem Cl⁻ a N1 dusíkem guaninu 8, který penalizuje struktury s menší vzdáleností těchto atomů než 10 Å (v případě větší vzdálenosti Cl⁻ iontu od G8(N1) dusíku se potenciál neuplatňuje, resp. je nulový).

Startovní struktury 9-methylguaninu, guanosinu a guanylyl(3'→5')guanosinu (GpG) byly odvozeny z krystalové struktury vlásenkového ribozymu. Krystalová struktura vlásenkového ribozymu s označením 2OUE ve formátu PDB byla vizualizována v programu PyMOL⁴⁶ a pomocí tohoto programu byly zkopírovány požadované části struktury a uloženy ve formátu PDB. V případě 9-methylguaninu a guanosinu byl využit přímo guanin 8, v případě GpG byla využita jiná vhodná část struktury, která obsahuje dva po sobě jdoucí guaninové nukleotidy a její cukr-fosfátová páteř má konformaci A-RNA.

Obdobně jako u struktury ribozymu pak byly pomocí programu tLEaP u každé z těchto struktur doplněny atomy vodíku a struktura byla umístěna do boxu s explicitním modelem vody TIP3P. 9-Methylguanin, který byl použit jako referenční systém pro parametrizaci celé metody, byl umístěn do boxu o stejné velikosti jako vlásenkový ribozym, aby byly zajištěny stejné podmínky a chování systému. Velikost boxu činila 8,77 x 8,63 x 8,22 nm³. Pro guanosin byla vzdálenost stěny boxu od solutu nastavena na 18 Å a minimální vzdálenost solutu od solventu na 0,7 Å a pro GpG na 20 Å a 0,7 Å. Velikost boxu pro guanosin činila 5,10 x 4,51 x 5,01 nm³ a pro GpG 5,76 x 5,30 x 5,60 nm³. Dále bylo do boxu ke každé struktuře vždy přidáno tolik K⁺ a Cl⁻ iontů, aby jejich koncentrace v boxu činila 150 mmol/l. Od každé struktury pak byly programem tLEaP za použití silového pole ff99bsc0χ_{OL3} vytvořeny koordináty a topologie ve formátu pro AMBER a programem acpype převedeny do formátu pro GROMACS.

Obdobně jako u struktury ribozymu byla pro každou strukturu získána topologie s deprotonovaným N1 dusíkem na guaninu a vytvořena dvojaká topologie pro cpH MD. Protože ve struktuře GpG jsou dva guaniny, byly v tomto případě vytvořeny dvě dvojaké topologie pro dvojí různé simulace. V jedné se mění protonační stav prvního guaninu (pořadí guaninů je uvažováno od 5' konce řetězce RNA, tato topologie je dále v textu značena jako GpG1) a v druhé druhého guaninu (GpG2).

Pro referenční systém, 9-methylguanin, byly stejně jako v případě ribozymu použity dva způsoby vykompenzování změny náboje systému při změně protonačního

stavu guaninu, kontinuální nábojová hustota a spárovaný Cl^- ion, neboť změna způsobu kompenzování náboje mění závislost celkové energie na parametru λ , a je tudíž potřeba parametrizovat metodu pro oba tyto způsoby kompenzování náboje zvlášť. Pro struktury guanosinu a GpG bylo použito pouze způsobu spárovaného Cl^- iontu.

Z topologie a koordinátů 9-methylguaninu se spárovaným Cl^- iontem byly dále odvozeny koordináty a topologie s menším počtem iontů ve stejně velkém boxu a koordináty a topologie s menší velikostí boxu (a menším počtem iontů tak, aby zůstala jejich koncentrace v boxu přibližně zachována). Menšího počtu iontů bylo dosaženo editací souborů s topologií a koordináty v textovém editoru. Původní systém s 9-methylguaninem obsahoval 48 Cl^- a 48 K^+ iontů, úpravami byly vytvořeny systém s 24 Cl^- a 24 K^+ ionty a systém s 8 Cl^- a 8 K^+ ionty při stejné velikosti boxu. Menší velikosti boxu bylo dosaženo úpravami souborů s topologií a koordináty pomocí programu napsaného v jazyce C, který existující box nejprve vycentroval tak, aby byl 9-methylguanin ve středu boxu, a následně ořezal o požadovanou velikost z každé strany. V textovém editoru se pak počty iontů upravily tak, aby koncentrace iontů v boxu činila 150 mmol/l. Takto byl vytvořen systém s 9-methylguaninem v boxech o rozměrech 7,45 x 7,30 x 6,89 nm³, 6,55 x 6,40 x 5,99 nm³, 5,55 x 5,40 x 4,99 nm³ a 4,55 x 4,40 x 3,99 nm³.

Rozměry a velikosti boxů všech systémů udává Tabulka 1. U každé MD simulace byla nejprve provedena ekvilibrace, tj. energetická relaxace systému a jeho pozvolné vyhřátí na 300K, a poté vlastní MD simulace. Ekvilibrace byla pro všechny systémy 0,1 ns dlouhá, délky vlastních MD simulací se lišily (Tabulka 2). Všechny simulace byly spuštěny za izobaricko-izotermických podmínek při teplotě 300 K a tlaku 1 bar. U všech ekvilibrací i vlastních MD simulací byl použit termostat, který je součástí SD integrátoru, u všech vlastních MD simulací byl použit barostat Parrinello-Rahman a v případě ekvilibrací byl použit pro různé systémy různý barostat (Tabulka 2).

V případě cpH MD simulací byla jako vstupní údaj zadána hodnota pK_a 9-methylguaninu 9,56⁴⁷ a rovněž byly jako vstup zadány požadované hodnoty pH zkoumaných systémů (Tabulka 2).

Tabulka 1 Rozměry a velikosti boxů. (Použité zkratky: **HrRz** = vlásenkový ribozym, **9-MeGua** = 9-methylguanin, **TI** = termodynamická integrace, **Cl** = kompenzace změny náboje při deprotonaci spárovaným Cl iontem, u systémů uvedených v tabulce bez přídomku Cl je použita kompenzace kontinuální nábojovou hustotou, **GpG1**, resp. **GpG2** = guanylyl(3'→5')guanosin, u kterého se mění protonační stav N1 dusíku na prvním, resp. druhém guaninu, pořadí guaninů je uvažováno od 5' konce řetězce RNA, **48 iontů** = 24 K⁺ a 24 Cl⁻ iontů, **16 iontů** = 8 K⁺ a 8 Cl⁻ iontů, **ořez 0,5** = ořez boxu způsobem popsáným v kapitole Metody o 0,5 nm z každé strany, analogicky **ořez 1,0; 1,5 a 2,0**)

system	rozměry boxu [nm ³]	velikost boxu [nm ³]
9-MeGua TI	8,77 x 8,63 x 8,22	622,35
HrRz	6,92 x 8,87 x 8,59	527,37
9-MeGua	8,77 x 8,63 x 8,22	622,35
9-MeGua Cl TI	8,77 x 8,63 x 8,22	622,35
HrRz Cl	8,35 x 9,21 x 9,27	712,50
9-MeGua Cl	8,77 x 8,63 x 8,22	622,35
Guanosin Cl	5,10 x 4,51 x 5,01	115,33
GpG1 Cl	5,76 x 5,30 x 5,60	171,08
GpG2 Cl	5,76 x 5,30 x 5,60	171,08
9-MeGua Cl 48 iontů	8,77 x 8,63 x 8,22	622,35
9-MeGua Cl 16 iontů	8,77 x 8,63 x 8,22	622,35
9-MeGua Cl ořez 0,5	7,45 x 7,30 x 6,89	374,85
9-MeGua Cl ořez 1,0	6,55 x 6,40 x 5,99	251,21
9-MeGua Cl ořez 1,5	5,55 x 5,40 x 4,99	149,62
9-MeGua Cl ořez 2,0	4,55 x 4,40 x 3,99	79,93

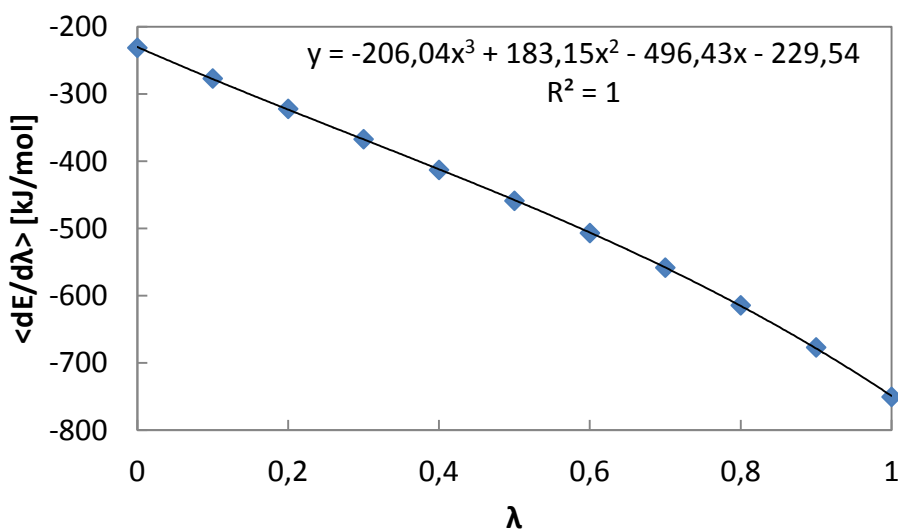
Tabulka 2 Rozdílné parametry MD simulací - barostat použitý při ekvilibraci, délka vlastní MD simulace a hodnoty pH nastavené při cpH MD. (Použité zkratky jsou definovány u Tabulky 1).

system	ekvilibrace - barostat	MD - délka (ns)	MD - hodnoty pH
9-MeGua TI	1,0 bar Parrinello-Rahman	5	
HrRz	1,0 bar Parrinello-Rahman	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 13,36, 14,36, 16,36, 17,36, 18,36
9-MeGua	1,0 bar Parrinello-Rahman	5	4,01; 5,01; 6,01 7,01; 12,01; 13,01; 14,01
		20	8,01; 9,01; 9,51; 10,01; 11,01
9-MeGua CI	1,0 bar Parrinello-Rahman	5	
TI			
HrRz CI	1,0 bar Parrinello-Rahman	20	6,36, 8,36, 10,36, 12,36, 14,36, 16,36, 18,36
9-MeGua CI	1,0 bar Parrinello-Rahman	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
guanosin CI	1,0 bar Parrinello-Rahman	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
GpG1 CI	1,0 bar Parrinello-Rahman	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
GpG2 CI	1,0 bar Parrinello-Rahman	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
9-MeGua CI	1,0 bar Parrinello-Rahman	20	77,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
48 iontů			
9-MeGua CI	1,0 bar Parrinello-Rahman	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
16 iontů			
9-MeGua CI	1,0 bar Berendsen	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
ořez 0,5			
9-MeGua CI	1,0 bar Berendsen	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
ořez 1,0			
9-MeGua CI	1,0 bar Berendsen	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
ořez 1,5			
9-MeGua CI	1,0 bar Berendsen	20	7,36, 8,36, 9,36, 9,86, 10,36, 11,36, 12,36
ořez 2,0			

3.3 Výsledky a diskuze

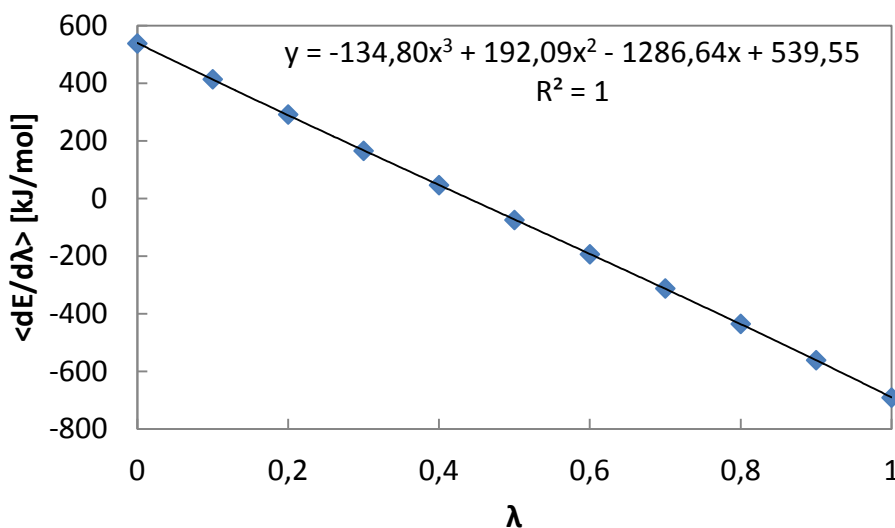
3.3.1 Parametrizace - termodynamická integrace

Nejprve byla provedena parametrizace metody na 9-methylguanin, jehož pK_a je známo. 9-Methylguanin byl zasolvatován do stejně velkého boxu jako vlásenkový ribozym a bylo spuštěno 11 MD simulací, každá 5 ns dlouhá, s fixními hodnotami proměnné λ o hodnotách 0 až 1, postupně se zvětšující o 0,1. Celkem byly provedeny dvě parametrizace celé metody pro dva různé způsoby kompenzace náboje guaninu v deprotonované formě. V prvním případě byla změna náboje při deprotonaci kompenzována kontinuální nábojovou hustotou. Pro každé nastavení parametru λ byla získána v každém snímku simulace hodnota $dE/d\lambda$, podobně jako v metodě termodynamické integrace. Po dokončení MD simulací byl spočítán aritmetický průměr hodnoty $dE/d\lambda$ pro každou hodnotu λ . Tím jsme získali závislost $\langle dE/d\lambda \rangle$ na proměnné λ . Nicméně na rozdíl od metody termodynamické integrace jsme tuto závislost neintegrovali až k celkové ΔG přechodu kanonické a deprotonované formy guaninu, ale tuto závislost jsme metodou nejmenších čtverců aproximovali kubickou rovnicí (Graf 1) a parametry této kubické rovnice byly použity jako vstupní parametry pro cpH MD.



Graf 1 Závislost $\langle dE/d\lambda \rangle$ na proměnné λ pro systém 9-methylguanin - deprotonovaný 9-methylguanin, kde náboj deprotonované formy byl kompenzován kontinuální nábojovou hustotou. Data byla nafitována kubickou funkcí a parametry této funkce byly použity jako vstupní parametry pro cpH MD.

Analogickým způsobem byla provedena také parametrizace na 9-methylguanin spárovaný s Cl⁻ iontem (Graf 2), která byla využita pro cpH MD simulace systémů rovněž obsahující spárovaný Cl⁻.



Graf 2 Závislost $\langle dE/d\lambda \rangle$ na proměnné λ pro systém 9-methylguanin - deprotonovaný 9-methylguanin s kompenzací náboje pomocí spárovaného Cl⁻ iontu. Rovnice regrese udává parametry kubické rovnice pro tuto parametrizaci.

3.3.2 Titrace 9-methylguaninu

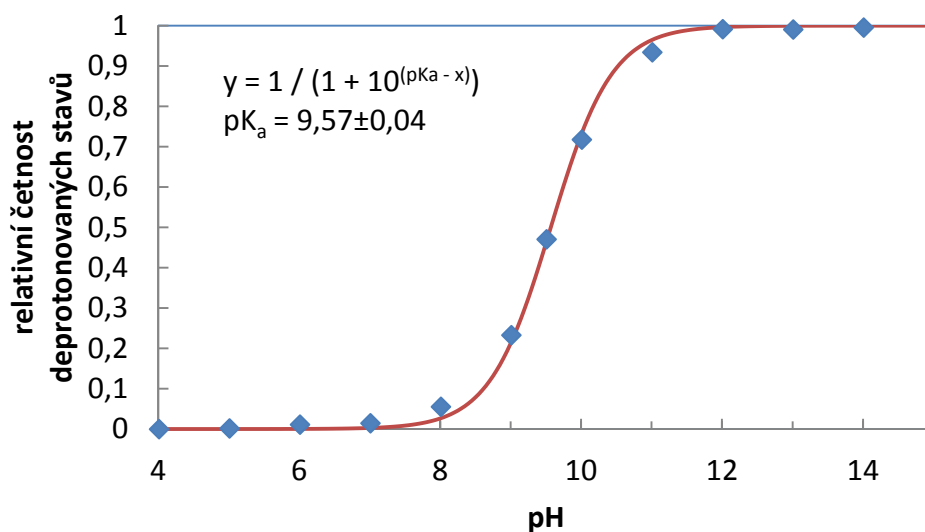
Abychom ověřili, že parametrizace byla provedena správně, byla provedena tzv. titrace 9-methylguaninu. Titrací v případě cpH MD rozumíme zjištění závislosti četnosti deprotonovaných stavů ($\lambda > 0,5$) a kanonických (protonovaných) stavů ($\lambda < 0,5$) v simulaci na pH a titrační křivkou graf této závislosti.⁶ Bylo spuštěno 12 cpH MD simulací, každá 5 ns dlouhá, každá s jinou nastavenou hodnotou pH. Změna náboje systému při deprotonaci byla kompenzována jak kontinuální nábojovou hustotou (Graf 3), tak také pomocí spárovaného iontu (Graf 4), a to vždy za použití příslušné parametrizace. Nastavené hodnoty pH byly v rozsahu od 4,01 do 14,01 (viz též Tabulka 2). Sledovala se závislost relativní četnosti deprotonovaných stavů ($\lambda > 0,5$) po dobu simulace na pH.

V reálném systému by na základě Henderson-Hasselbalchovy rovnice měla relativní četnost deprotonovaných stavů p_{A^-} na pH záviset podle vztahu:

$$p_{A^-} \cong \frac{[A^-]}{[HA] + [A^-]} = \frac{1}{1 + 10^{pK_a - pH}} \quad (23)$$

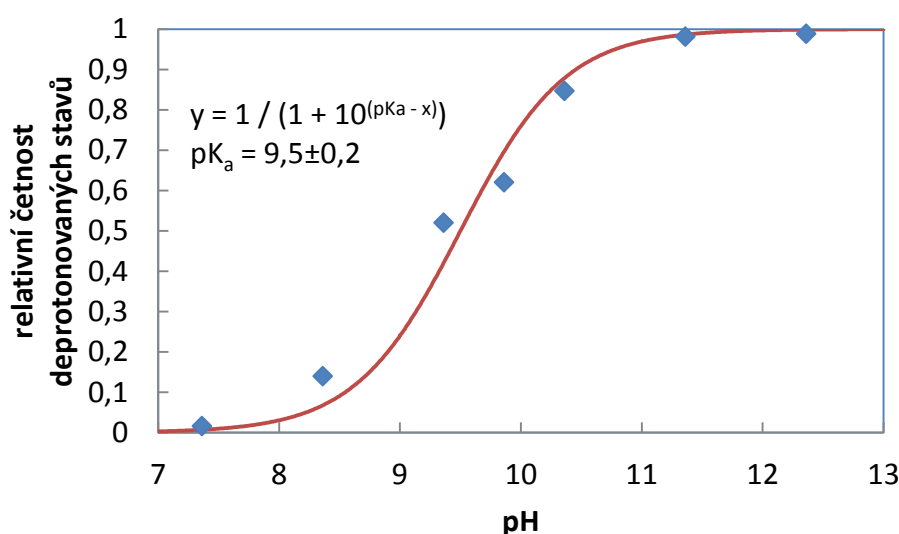
kde $[A^-]$ je rovnovážná koncentrace deprotonované formy, $[HA]$ rovnovážná koncentrace protonované formy a pK_a záporný dekadický logaritmus příslušné

titrovatelné skupiny. Zjištěná závislost četnosti deprotonovaných stavů na pH byla naitována rovnicí v tomto tvaru a tím bylo zjištěno pK_a , které pro danou titrovatelnou skupinu v daném systému spočetla metoda cpH MD (Graf 3).



Graf 3 Závislost relativní četnosti deprotonovaných stavů 9-methylguaninu v průběhu simulace (změna náboje při deprotonaci je kompenzována kontinuální nábojovou hustotou) na nastavené hodnotě pH, rovnice regrese a intervalový odhad pK_a .

Obdobně byla provedena titrace, ve které změnu náboje při deprotonaci kompenzoval spárovaný Cl^- ion. V tomto případě bylo spuštěno 7 cpH MD simulací o nastavených hodnotách pH v rozmezí od pH 7,36 do pH 12,36 (viz též Tabulka 2). Zjištěnou závislost četnosti deprotonovaných stavů na pH s naitovanou rovnicí a intervalovým odhadem pK_a znázorňuje Graf 4.



Graf 4 Závislost relativní četnosti deprotonovaných stavů systému 9-methylguanin se spárovaným Cl^- iontem v průběhu simulace na nastavené hodnotě pH, rovnice regrese a intervalový odhad pK_a .

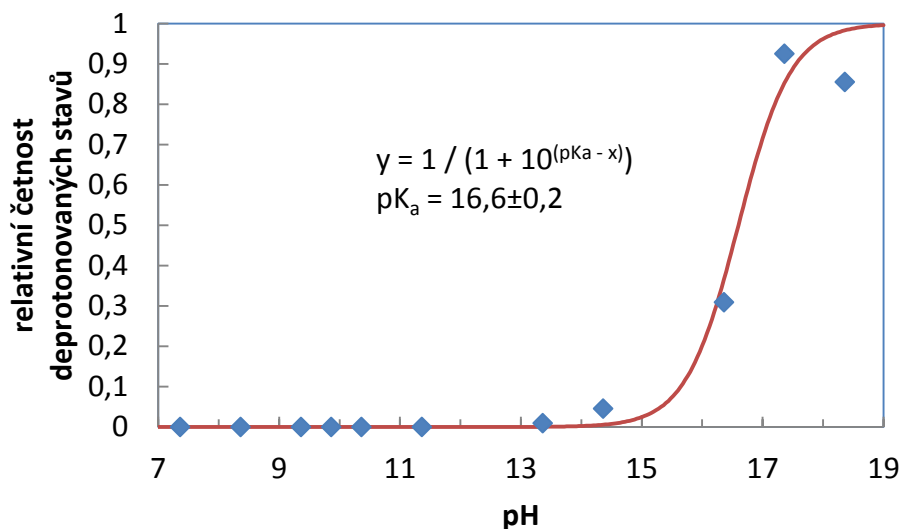
Je třeba poznamenat, že pK_a 9-methylguaninu spočtené tímto způsobem nemá význam jako takové, protože hodnota pK_a 9-methylguaninu byla zadána jako jeden ze vstupních parametrů cpH MD simulací a metoda byla naparametrizována tak, aby reprodukovala tuto hodnotu pro 9-methylguanin. Tato hodnota pK_a a stejně tak průběh titrační křivky ale slouží jako ověření, zda byla metoda naparametrizována správně, tj. zda byly arteficiální síly silového pole působící na parametr λ dostatečně přesně spočtené metodou termodynamické integrace (resp. její modifikací) a metoda cpH MD dokáže přesně reprodukovat nastavenou hodnotu pK_a pro referenční systém, tj. 9-methylguanin.

V obou případech kompenzace náboje odpovídá vypočtené pK_a hodnotě zadaného pK_a 9-methylguaninu, a tedy v obou případech je metoda naparametrizována správně.

3.3.3 Aplikace na větší systémy

Po ověření správné parametrizace titrací 9-methylguaninu byla provedena titrace větších systémů, aby bylo ověřeno, zda metoda funguje správně a lze ji dále využít k popisu dynamiky daných systémů. Nejprve byla provedena titrace vlásenkového ribozymu za využití silového pole ff99bsc0 χ_{OL3} s kompenzací změny náboje při deprotonaci kontinuální nábojovou hustotou. V tomto případě bylo spuštěno 14 cpH

MD simulací o nastavených hodnotách pH v rozmezí (viz též Tabulka 2) od pH 6,36 do pH 18,36 (Graf 5).

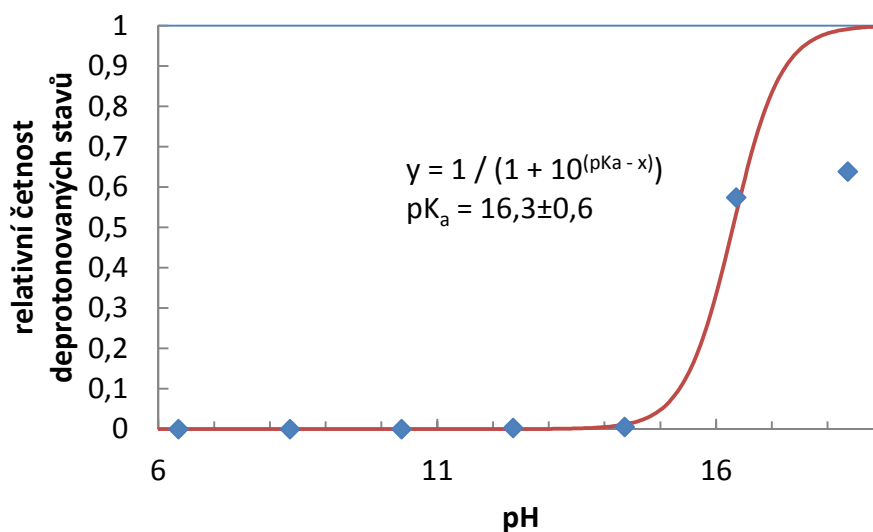


Graf 5 Závislost relativní četnosti deprotonovaných stavů vlásenkového ribozymu (změna náboje při deprotonaci je kompenzována kontinuální nábojovou hustotou) v průběhu simulace na nastavené hodnotě pH, rovnice regrese a intervalový odhad pK_a .

Spočtené pK_a ($16,6 \pm 0,2$) se významně liší od experimentálně naměřené hodnoty ($9,55^{48}$), dokonce až o 7,05 jednotek. Takto výrazný rozdíl naznačuje, že může existovat výrazná nerovnováha v molekulárně mechanickém popisu mezimolekulárních interakcí různých protonačních forem nukleobází s vodou a se zbytkem RNA. Tato nerovnováha popisu mezimolekulárních interakcí by se pak nejspíš primárně týkala popisu nekanonických protonačních forem nukleobází. Proto bylo dále zkoumáno, které parametry simulací mohou mít vliv na posun pK_a od jeho skutečné experimentálně zjištěné hodnoty.

Na systému vlásenkového ribozymu se zjišťovalo, zda posun pK_a může být zlepšen přesnějším způsobem kompenzace změny náboje při změně protonačního stavu, tedy pomocí spárovaného Cl^- iontu, nebo zda může být zlepšen použitím korekcí ϵ_{OL1} silového pole $ff99bsc0\chi_{OL3}$, u kterých bylo prokázáno mírné zlepšení dynamiky RNA, konkrétně zvýšená populace reaktivní formy aktivního místa vlásenkového ribozymu.⁴⁹ Byla provedena titrace vlásenkového ribozymu s oběma těmito změnami parametry, tedy se spárovaným Cl^- iontem a za použití silového pole $ff99bsc0\chi_{OL3}$ s korekcí ϵ_{OL1} . (Je velmi nepravděpodobné, že by změna způsobu kompenzace náboje a použití korekcí pro silové pole měly na posun pK_a protichůdný vliv, proto byly tyto dva kroky

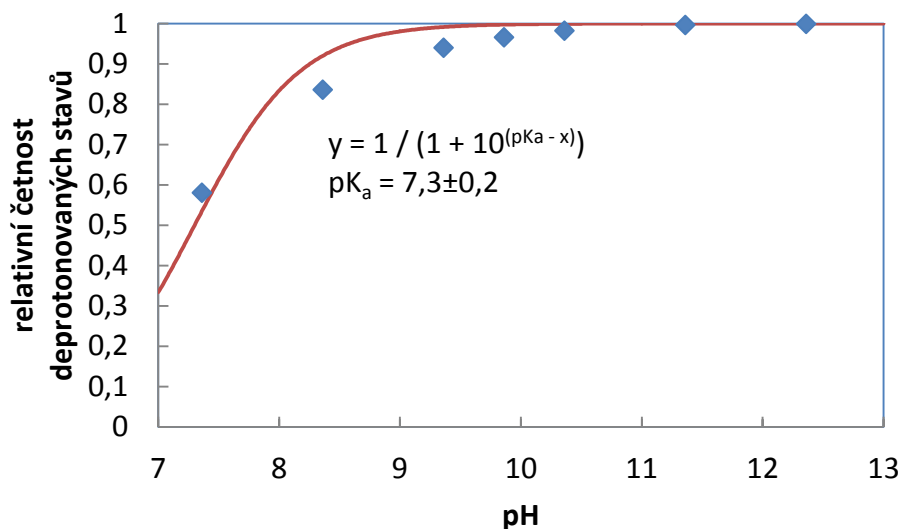
provedeny zároveň a až v případě úspěchu by byl efekt dekomponován.) Bylo spuštěno 7 cpH MD simulací s nastavenými hodnotami pH (viz též Tabulka 2) od pH 6 do pH 18 (Graf 6).



Graf 6 Závislost relativní četnosti deprotonovaných stavů systému vlásenkový ribozym se spárovaným Cl⁻ iontem v průběhu simulace na nastavené hodnotě pH, rovnice regrese a intervalový odhad pK_a.

Spočtené pK_a (16,3) se opět velmi výrazně liší od experimentálně určeného (9,55), v tomto případě o 6,75 jednotek. Tento výsledek ukazuje, že enormně velký arteficiální posun pK_a v kontextu vlásenkového ribozymu je spíše zapříčiněn hlubší nerovnováhou v popisu mezimolekulárních interakcí nukleobází v různých protonačních stavech a není ovlivněn změnou způsobu kompenzace náboje deprotonované formy G8 nebo zahrnutím εζ_{OLI} korekce. Abychom alespoň částečně odkryli, kde v molekulárně mechanickém popisu protonačních stavů nukleobází je domnělá nerovnováha, byly provedeny další cpH MD simulace na jednodušších systémech. Pro tento účel byl studován posun pK_a predikovaný metodou cpH MD ve dvou dostatečně jednoduchých systémech, u nichž je experimentálně tento posun pK_a stanoven. Konkrétně jsme studovali posun pK_a v molekule guanosinu, který místo methylu má na guaninu navázanu celou ribosu, a dinukleotidu GpG. Druhá uvedená struktura obsahuje dva guaniny, a proto byl zkoumán predikovaný posun pK_a u obou těchto guaninů za využití topologií GpG1 a GpG2, jejichž příprava byla popsána v kapitole Metody. Simulace všech těchto struktur byly provedeny za použití silového pole ff99bsc0χ_{OL3}. Ke kompenzaci změny náboje při deprotonaci byl použit spárovaný Cl⁻ ion.

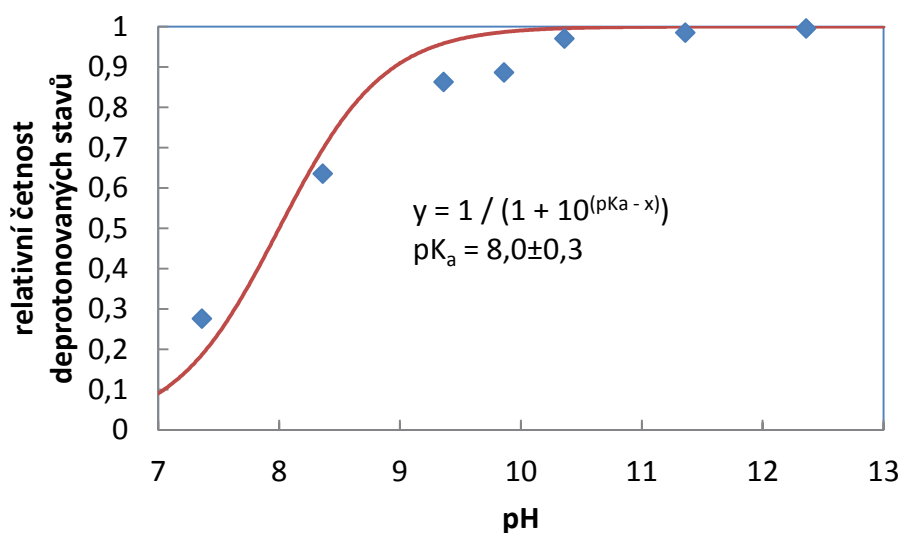
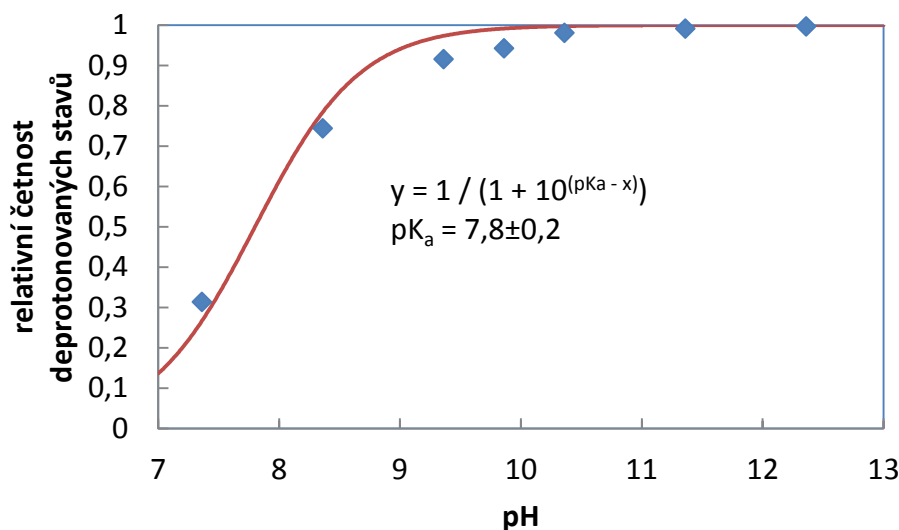
Pro guanosin bylo spuštěno 7 cpH MD simulací s nastavenými hodnotami pH v rozmezí (viz též Tabulka 2) od pH 7,36 do pH 12,36 (Graf 7).



Graf 7 Závislost relativní četnosti deprotonovaných stavů systému guanosin se spárovaným Cl⁻ iontem v průběhu simulace na nastavené hodnotě pH, rovnice regrese a intervalový odhad predikované pK_a.

Spočtené pK_a guanosinu (7,3±0,2) se opět významně liší od experimentálně určeného (9,22±0,02⁵⁰⁻⁵¹), ačkoliv chyba při odhadu pK_a je v tomto systému mnohem menší, než tomu bylo v případě vlásenkového ribozymu.

Pro každou z topologií GpG1 a GpG2 bylo opět spuštěno 7 cpH MD simulací s týmiž hodnotami pH jako u referenčního systému, tj. 9-methylguaninu (Grafy 8 a 9).



Graf 8 a 9 Závislost relativní četnosti deprotonovaných stavů systémů GpG1 (horní graf) a GpG2 (dolní graf), oba systémy se spárovaným Cl⁻ iontem, v průběhu simulace na nastavené hodnotě pH, rovnice regrese a intervalový odhad pK_a.

Spočtená pK_a systémů 5'-guaninu (systém GpG1) a 3'-guaninu (systém GpG2) (7,8±0,2 a 8,0±0,3) jeví podobný posun od experimentálně určených (9,34±0,07 a 10,38±0,10⁵²) jako samotný guanosin.

3.3.4 Vliv počtu iontů v boxu

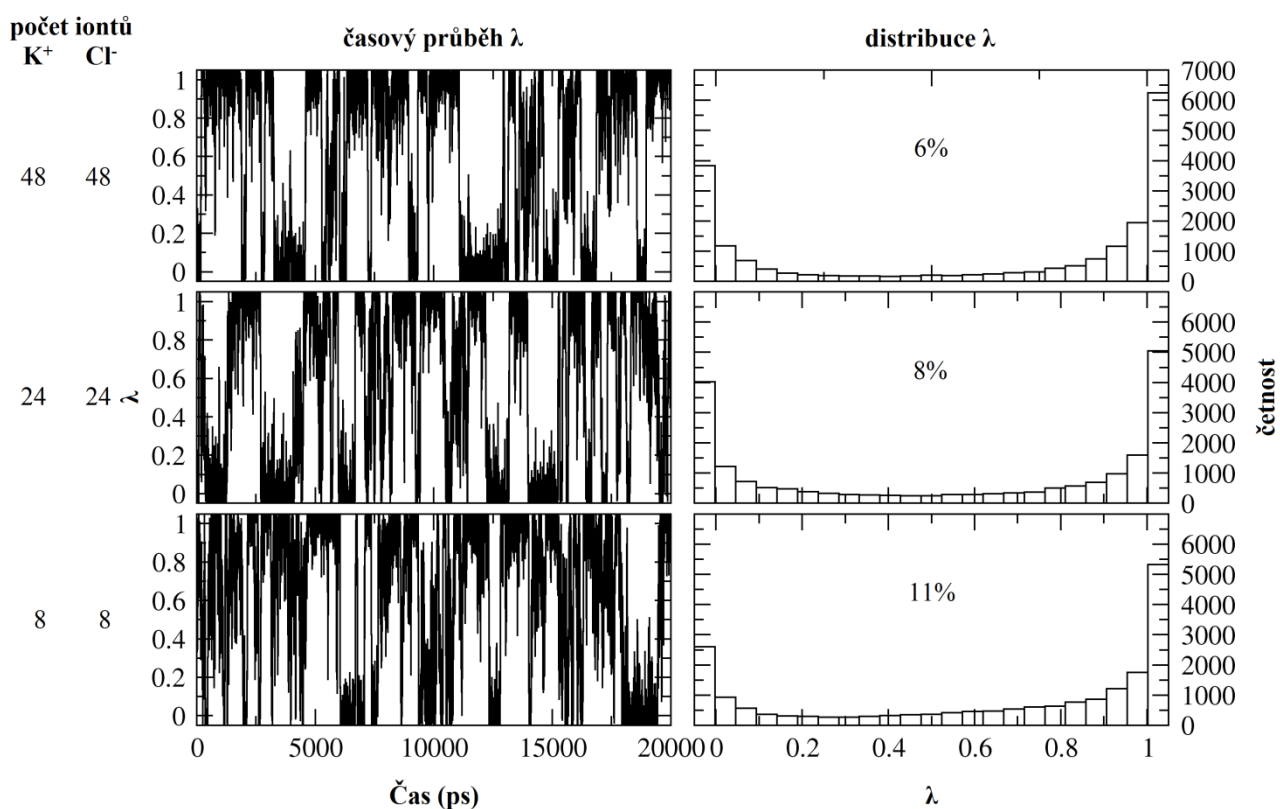
Další z parametrů cpH MD simulací, který jsme testovali jako potenciální zdroj, který by mohl mít vliv na posun pK_a, je počet iontů v boxu, tj. iontová síla v simulaci. Tento vliv byl zkoumán na samotném referenčním systému tj. na 9-methylguaninu. Byla provedena titrace 9-methylguaninu v boxu s třemi různými hodnotami iontové síly

tj. různými počty iontů ve stejně velkém simulačním boxu (viz Tabulka 3). Pro každý z boxů bylo, obdobně jako v předchozích případech, spuštěno 7 cpH MD simulací při týchž sedmi hodnotách pH jako v případě titrace referenčního systému, tj. 9-methylguaninu s počtem iontů, který byl použit při parametrizaci pomocí termodynamické integrace. Spočtené hodnoty pK_a byly porovnány (Tabulka 3) s hodnotou pK_a referenčního 9-methylguaninu, tj. spočtenou v boxu s 48 Cl^- a 48 K^+ ionty, viz kapitola 2.3.2 Titrace 9-methylguaninu.

Tabulka 3 Spočtená hodnota pK_a 9-methylguaninu při různém počtu iontů v boxu téže velikosti.

Počet K^+	Počet Cl^-	pK_a
48	48	9,5±0,2
24	24	9,6±0,1
8	8	9,6±0,1

Bylo zjištěno, že na spočtenou hodnotu pK_a nemá iontové prostředí významný vliv. Krom odhadu pK_a jsme sledovali případný vliv různého iontového prostředí i pomocí distribuce hodnot proměnné λ v závislosti na počtu iontů (Graf 10). Ani v tomto parametru jsme však nezaznamenali vliv iontového prostředí na chování cpH MD.



Graf 10 Grafy v levém sloupci ukazují časový průběh hodnoty λ a histogramy v pravém sloupci distribuci hodnot λ do 23 intervalů v průběhu simulace (při nastavené hodnotě pH 9,86) pro trojí různý počet iontů ve stejně velkém boxu ($622,35 \text{ nm}^3$). Číslo v procentech v pravém grafu udává relativní četnost hodnot proměnné λ v nefyzikálním intervalu od 0,25 do 0,75.

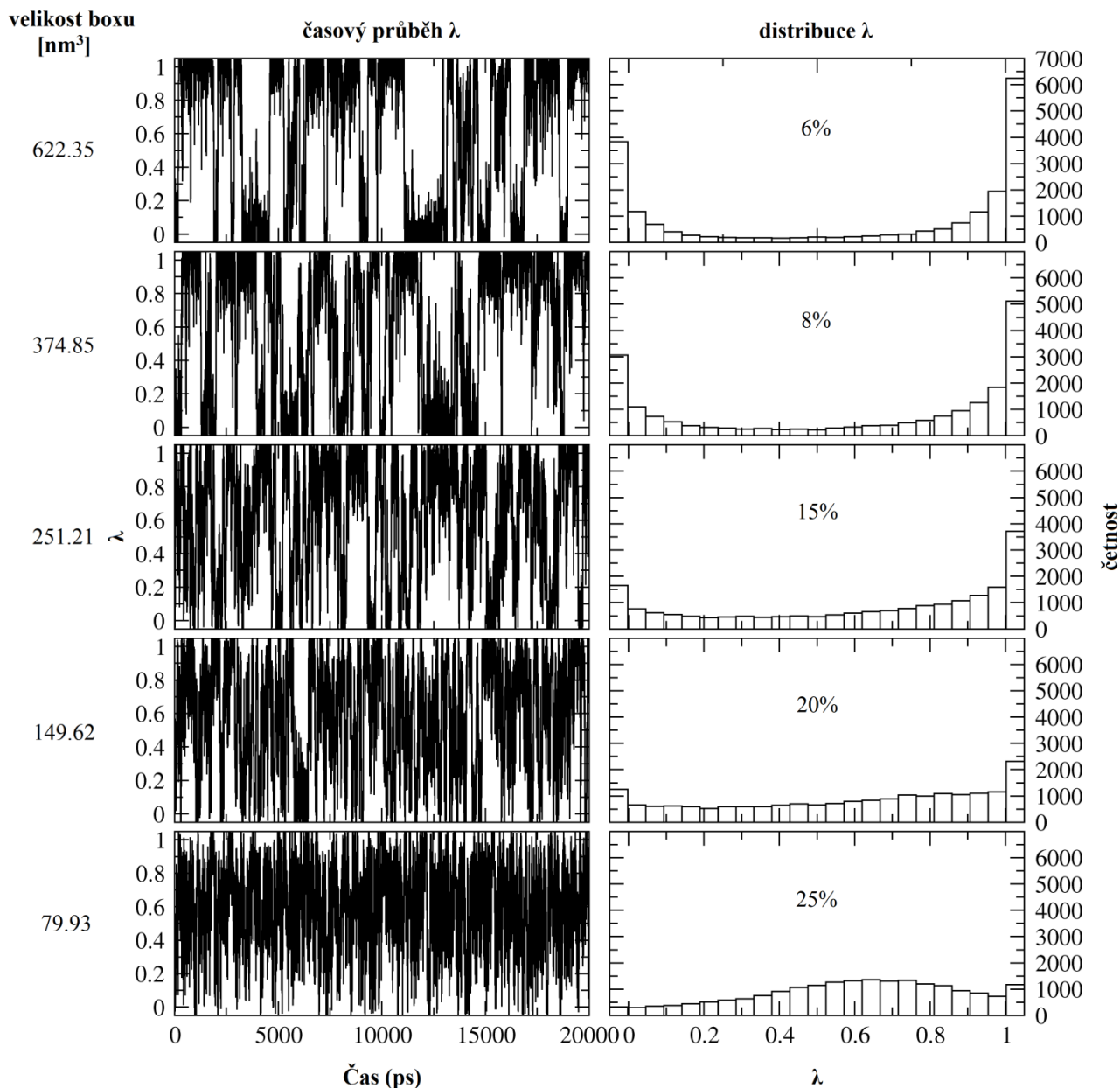
3.3.5 Vliv velikosti boxu

Dalším z parametrů, který by mohl mít na posun pK_a vliv je velikost boxu. Tento vliv byl rovněž zkoumán na systému 9-methylguaninu. Byla provedena titrace v boxech s čtyřmi různými velikostmi uvedenými v Tabulce 1 s označením „ořez“ při konstantní koncentraci iontů v simulačním boxu. Pro každý z boxů bylo, obdobně jako v předchozích případech, spuštěno 7 cpH MD simulací při týchž sedmi hodnotách pH (viz též Tabulka 2) jako v případě titrace referenčního systému, tj. 9-methylguaninu v boxu, který byl použit při parametrizaci pomocí termodynamické integrace. Spočtené hodnoty pK_a byly porovnány (Tabulka 4) s hodnotou pK_a referenčního 9-methylguaninu, tj. spočtenou v boxu o velikosti $622,35 \text{ nm}^3$, viz kapitola 2.3.2 Titrace 9-methylguaninu.

Tabulka 4 Spočtená hodnota pK_a 9-methylguaninu při různé velikosti boxu.

Velikost boxu [nm ³]	pK_a
622,35	9,5±0,2
374,85	9,6±0,1
251,21	9,58±0,06
149,62	9,6±0,2
79,93	9,6±0,3

Bylo zjištěno, že na spočtenou hodnotu pK_a nemá velikost boxu významný vliv, nicméně chování cpH MD simulací se přesto lišilo v závislosti na velikosti boxu. Tyto rozdíly v chování simulací v závislosti na velikosti boxu lze demonstrovat pomocí distribuce hodnot proměnné λ v závislosti na velikosti boxu (Graf 11).



Graf 11 Grafy v levém sloupci ukazují časový průběh hodnoty λ a histogramy v pravém sloupci distribuci hodnot λ do 23 intervalů v průběhu simulace (při nastavené hodnotě pH 9,86) pro pět různě velkých boxů. Číslo v procentech v pravém grafu udává relativní četnost hodnot proměnné λ v nefyzikálním intervalu od 0,25 do 0,75.

Naměřená data ukazují, že se zmenšující se velikostí boxu dochází k nárůstu populace proměnné λ okolo hodnot $\lambda=0,5$, které nemají fyzikální význam. Ačkoliv tento efekt, jak bylo ukázáno, nemá vliv na odhad pK_a , je možné, že může ovlivnit přesnost tohoto odhadu právě tím, že parametr λ častěji populuje nefyzikální hodnoty okolo

$\lambda=0,5$. Tento jev však lze kompenzovat velikostí bariéry pro hodnoty λ okolo hodnot $\lambda=0,5$, která je vstupním parametrem cpH MD simulací (viz rovnice 19).

Tento jev lze vysvětlit na základě elektrostatických interakcí titrovaného guaninu G8, spárovaného Cl^- iontu s jejich periodickými kopiemi. Lze ho dobře znázornit i v jednorozměrném prostoru na atomech A a B, mezi nimiž se se změnou parametru λ přenáší záporný náboj (pro vysvětlení tohoto jevu postačí, když budeme uvažovat interakci atomů A a B s jejich jednou periodickou kopií A' a B').

Je-li $\lambda=0$, mají A a A' nulový náboj a B a B' náboj -1. Jediná elektrostatická interakce, která se v tomto případě uplatní, je tedy mezi B a B'. Je-li $\lambda=0,5$, mají A, A', B a B' náboj -0,5. Uplatní se tedy elektrostatické interakce mezi všemi těmito čtyřmi částicemi. A konečně je-li $\lambda=1$, mají A a A' náboj -1 a B a B' nulový náboj. Jediná elektrostatická interakce, která se v tomto případě uplatní, je tedy mezi A a A'.

Je-li velikost boxu dostatečná (např. stonásobná vzhledem k průměrné vzdálenosti mezi A a B), jsou vzdálenost mezi A a A' a vzdálenost mezi B a B' relativně velké a elektrostatické interakce mezi A a A' a mezi B a B' jsou tedy velmi malé. V případech $\lambda=0$ a $\lambda=1$ tedy bude celková hodnota elektrostatické složky interakčního potenciálu v systému velmi malá (Tabulka 5). Oproti tomu v případě $\lambda=0,5$ budou mezi A a B nacházejícími se v minimální vzdálenosti působit relativně silné elektrostatické interakce (Tabulka 5). Proto bude tento stav energeticky znevýhodněn. Pokud je paramterizace cpH MD metody prováděna a takto velkém boxu, pak je toto energetické znevýhodnění popsáno v termodynamické integraci, ze závislosti energie na parametru λ je odečteno a nahrazeno zvolenou bariérou, viz rovnice (19).

Tabulka 5 Hodnota elektrostatické složky interakčního potenciálu mezi atomy A a B a jejich periodickými kopiemi A' a B' pro hodnoty $\lambda=0$; 0,5 a 1 při vzdálenosti mezi A a B = 1 Å a velikosti jednorozměrného boxu = 100 Å. (Použité značky a zkratky: r = vzdálenost mezi A a B, a = velikost boxu, q(A) = náboj atomu A, q(B) = náboj atomu B, $E_{e(AB)}$ = hodnota elektrostatické složky interakčního potenciálu mezi atomy A a B, analogicky $E_{e(AA')}$, $E_{e(AB')}$, $E_{e(BA')}$ a $E_{e(BB')}$, $E_{e\ celk}$ = celková hodnota elektrostatické složky interakčního potenciálu mezi atomy A a B a jejich periodickými kopiemi A' a B' v rámci jedné periodicity, poměr = relativní poměr hodnot $E_{e\ celk}$ pro různé hodnoty λ)

λ	r	a	q(A)	q(B)	$E_{e(AB)}$	$E_{e(AA')}$	$E_{e(AB')}$	$E_{e(BA')}$	$E_{e(BB')}$	$E_{e\ celk}$	poměr
0	1	100	0	-1	0,000	0,000	0,000	0,000	0,010	0,01	1
0,5	1	100	-0,5	-0,5	0,250	0,003	0,002	0,003	0,003	0,26	26,1
1	1	100	-1	0	0,000	0,010	0,000	0,000	0,000	0,01	1

Je-li velikost boxu malá (např. pouze dvojnásobná vzhledem k průměrné vzdálenosti mezi A a B), je vzdálenost mezi A a A' a B a B' natolik malá, že elektrostatické interakce působící mezi A a A' a mezi B a B' jsou už relativně významné. V případech $\lambda=0$ a $\lambda=1$ tedy bude celková hodnota elektrostatické složky interakčního potenciálu vzhledem k systému s 100 Å velkým boxem mnohonásobně větší (Tabulka 6). Stejně tak v případě $\lambda=0,5$ budou elektrostatické interakce mezi atomy A, B a jejich periodickými kopiemi signifikantně větší, ale toto zvětšení repulzní interakce je mnohem menší než v případech $\lambda=0$ a $\lambda=1$. (Tabulka 6). Proto bude v takto malém boxu stav $\lambda=0,5$ energeticky mnohem méně znevýhodněn než v systému ve velkém boxu. Pokud byla tedy cpH MD metoda parametrizována s větším boxem bude tento stav s λ okolo hodnoty 0,5 v malém boxu více populovaný než ve velkém boxu.

Tabulka 6 Hodnota elektrostatických odpuzivých sil pro hodnoty $\lambda=0; 0,5$ a 1 při vzdálenosti A a $B = 1 \text{ \AA}$ a velikosti jednorozměrného boxu $= 2 \text{ \AA}$. (Použité značky a zkratky jsou definovány u Tabulky 5.)

λ	r	a	$q(A)$	$q(B)$	$E_{e(AB)}$	$E_{e(AA')}$	$E_{e(AB')}$	$E_{e(BA')}$	$E_{e(BB')}$	$E_{e \text{ celk}}$	poměr
0	1	2	0	-1	0,000	0,000	0,000	0,000	0,500	0,50	1
0,5	1	2	-0,5	-0,5	0,250	0,125	0,083	0,250	0,125	0,83	1,7
1	1	2	-1	0	0,000	0,500	0,000	0,000	0,000	0,50	1

3.4 Závěry

Tato práce byla zaměřena na aplikaci metody cpH MD od Grubmüllera et al.⁶ na studium nekanonických protonačních stavů molekul RNA a na případnou analýzu problémů v aplikaci této metody (původně vyvinuté pro proteiny) na RNA. Bylo zkoumáno, jak spolehlivě je tato metoda schopna popsat posun pK_a nukleové báze zabudované v makromolekule a v menších strukturách (nukleosid a dinukleotid), použijeme-li silové pole AMBER ff99bsc0 χ_{OL3} a ff99bsc0 χ_{OL3} s korekcí $\epsilon\zeta_{OL1}$. Funkčnost této metody byla otestována na struktuře vlásenkového ribozymu a jeho nukleové bázi G8.

Bylo zjištěno, že hodnota pK_a guaninu G8 ((de)protonovaného na dusíku N1) zabudovaného ve struktuře vlásenkového ribozymu, kterou spočetla cpH MD za použití silového pole ff99bsc0 χ_{OL3} a při kompenzaci změny náboje při změně protonačního stavu pomocí kontinuální nábojové hustoty, $16,6\pm 0,2$, se od skutečné hodnoty, $9,5^{48}$, podstatně liší. Popis pK_a se nezlepší ani při použití kompenzace změny náboje při změně protonačního stavu pomocí spárovaného Cl^- iontu a použití silového pole ff99bsc0 χ_{OL3} s korekcí $\epsilon\zeta_{OL1}$ (spočtená hodnota pK_a $16,3\pm 0,6$). Pro vlásenkový ribozym tedy metoda cpH MD za použití silového pole AMBER ff99bsc0 χ_{OL3} ani ff99bsc0 χ_{OL3} s korekcí $\epsilon\zeta_{OL1}$ není schopna spolehlivě určit hodnotu pK_a nukleové báze zabudované v makromolekule. Je tedy zřejmé, že metoda cpH MD není přímo přenositelná na studium protonačních rovnováh v RNA. Jedna z možných hypotéz, která by vysvětlila selhání metody cpH MD v RNA systémech je existence dosud ne zcela popsané nepřesnosti v popisu mezimolekulárních interakcí nukleobází v jejich nekanonickém protonačním stavu se zbytkem RNA. Bylo ukázáno, že tato nerovnováha není zapříčiněna způsobem kompenzace změny náboje při změně protonačního stavu, ani nelze dohledat přímou souvislost s konformační variabilitu aktivního místa vlásenkového ribozymu, kde změna protonačního stavu probíhá.

Abychom lépe porozuměli příčinám selhání metody cpH MD v RNA systémech, provedli jsme testy na jednodušších systémech, konkrétně volném nukleosidu a dinekleotidu. Hodnota pK_a guaninu ((de)protonovaného na dusíku N1) v guanosinu, kterou spočetla cpH MD za použití silového pole ff99bsc0 χ_{OL3} a při kompenzaci změny náboje při změně protonačního stavu pomocí Cl^- iontu, činí $7,3\pm 0,2$. Tato hodnota také neodpovídá experimentální hodnotě pK_a $9,22\pm 0,02^{50-51}$. Hodnoty pK_a systémů GpG1 a GpG2, tj. dinukleotidu GpG deprotonovaného na N1 dusíku 5'- a 3'-guaninu, spočetla metoda cpH MD na $7,8\pm 0,2$ a $8,0\pm 0,3$. Tyto hodnoty také neodpovídají

experimentálními hodnotám $9,34 \pm 0,07$ a $10,38 \pm 0,10^{52}$. Jinými slovy, ačkoli absolutní hodnota chyby v predikci posunu pK_a je u obou těchto systémů menší, projevuje se nesprávný popis pK_a již na těchto jednoduchých elementech RNA. Je proto možné, že nesprávný popis pK_a G8 ve vlásenkovém ribozymu, ale i v guanosinu nebo GpG dinukleotidu může být pravděpodobně zapříčiněn nepřesným nastavením některých parametrů silového pole. Chování vlásenkového ribozymu, popř. dalších RNA systémů by tedy mohla zlepšit další reparametrizace silového pole.

Bylo také otestováno, jaký vliv na spočtenou hodnotu pK_a a distribuci proměnné λ má velikost boxu a počet iontů v boxu. Výsledky ukazují, že velikost boxu má zanedbatelný vliv na spočtenou hodnotu pK_a , ale má vliv na distribuci proměnné λ . Se zmenšující se velikostí boxu roste četnost hodnot, které nemají fyzikální význam. Tento efekt nemá však vliv na odhad pK_a a může být korigován nastavením velikosti bariéry mezi oběma protonačními stavy. Počet iontů v boxu má zanedbatelný vliv jak na spočtenou hodnotu pK_a , tak na distribuci proměnné λ .

Nalezení příčiny, proč metoda cpH MD od Grubmüllera et al. za využití silového pole AMBER ff99bsc0 χ_{OL3} ani ff99bsc0 χ_{OL3} s korekcí $\epsilon\zeta_{OL1}$ není schopna spočítat správně pK_a , bude pravděpodobně vyžadovat další testy na dalších strukturách, popř. dodatečnou analýzu cpH MD simulací guanosinu, abychom odhalili, odkud se bere posun pK_a v tomto jednoduchém systému. Výsledky této práce však znamenají jistý pokrok v pochopení příčin selhání přímé přenositelnosti metody cpH MD původně vyvinuté a testované pro proteiny na RNA, a tedy tato práce může přispět ke zprovoznění této metody pro molekuly RNA obecně. Stejně tak úplné objasnění mechanismu katalýzy vlásenkového ribozymu bude rovněž vyžadovat další studie a doladění parametrů cpH MD bude zřejmě v těchto studiích hrát svou nezastupitelnou roli.

4 Literatura

1. Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561-&.
2. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **1982**, *31*, 147-57.
3. Mattick, J. S.; Makunin, I. V. Non-coding RNA. *Hum Mol Genet* **2006**, *15*, R17-R29.
4. Ward, W. L.; Plakos, K.; DeRose, V. J. Nucleic Acid Catalysis: Metals, Nucleobases, and Other Cofactors. *Chemical Reviews* **2014**, *114*, 4318-4342.
5. Goh, G. B.; Knight, J. L.; Brooks, C. L. Toward Accurate Prediction of the Protonation Equilibrium of Nucleic Acids. *J Phys Chem Lett* **2013**, *4*, 760-766.
6. Donnini, S.; Tegeler, F.; Groenhof, G.; Grubmuller, H. Constant pH Molecular Dynamics in Explicit Solvent with lambda-Dynamics. *J Chem Theory Comput* **2011**, *7*, 1962-1978.
7. Vangunsteren, W. F.; Berendsen, H. J. C. Computer-Simulation of Molecular-Dynamics - Methodology, Applications, and Perspectives in Chemistry. *Angew Chem Int Edit* **1990**, *29*, 992-1023.
8. Nezbeda, I.; Kolafa, J.; Kotrla, M., *Úvod do počítačových simulací Metody Monte Carlo a molekulární dynamiky*. Karolinum: Praha, **1998**.
9. Schlick, T., *Molecular modeling and simulation*. Springer: Berlin Heidelberg New York, **2002**.
10. Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. Molecular Dynamics and Quantum Mechanics of RNA: Conformational and Chemical Change We Can Believe In. *Accounts Chem Res* **2010**, *43*, 40-47.
11. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J Am Chem Soc* **1995**, *117*, 5179-5197.
12. Foloppe, N.; MacKerell, A. D. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* **2000**, *21*, 86-104.
13. MacKerell, A. D.; Banavali, N. K. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* **2000**, *21*, 105-120.
14. Sponer, J.; Spackova, N. Molecular dynamics simulations and their application to four-stranded DNA. *Methods* **2007**, *43*, 278-290.
15. Sponer, J.; Cang, X. H.; Cheatham, T. E. Molecular dynamics simulations of G-DNA and perspectives on the simulation of nucleic acid structures. *Methods* **2012**, *57*, 25-39.
16. Cheatham, T. E.; Case, D. A. Twenty-Five Years of Nucleic Acid Simulations. *Biopolymers* **2013**, *99*, 969-977.
17. Mlynsky, V.; Banas, P.; Hollas, D.; Reblova, K.; Walter, N. G.; Sponer, J.; Otyepka, M. Extensive Molecular Dynamics Simulations Showing That Canonical G8 and Protonated A38H(+) Forms Are Most Consistent with Crystal Structures of Hairpin Ribozyme. *J Phys Chem B* **2010**, *114*, 6642-6652.
18. Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E.; Jurecka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J Chem Theory Comput* **2011**, *7*, 2886-2902.
19. Krepl, M.; Zgarbova, M.; Stadlbauer, P.; Otyepka, M.; Banas, P.; Koca, J.; Cheatham, T. E.; Jurecka, P.; Sponer, J. Reference Simulations of Noncanonical Nucleic Acids with Different chi Variants of the AMBER Force Field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *J Chem Theory Comput* **2012**, *8*, 2506-2520.

20. Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D. Impact of 2'-Hydroxyl Sampling on the Conformational Properties of RNA: Update of the CHARMM All-Atom Additive Force Field for RNA. *J Comput Chem* **2011**, *32*, 1929-1943.
21. Orozco, M.; Noy, A.; Perez, A. Recent advances in the study of nucleic acid flexibility by molecular dynamics. *Curr Opin Struct Biol* **2008**, *18*, 185-193.
22. Besseova, I.; Banas, P.; Kuhrova, P.; Kosinova, P.; Otyepka, M.; Sponer, J. Simulations of A-RNA Duplexes. The Effect of Sequence, Solute Force Field, Water Model, and Salt Concentration. *J Phys Chem B* **2012**, *116*, 9899-9916.
23. Chipot, C.; Pohorille, A., *Free Energy Calculations*. Springer: Berlin Heidelberg New York, **2007**.
24. Dolenc, J.; Oostenbrink, C.; Koller, J.; van Gunsteren, W. F. Molecular dynamics simulations and free energy calculations of netropsin and distamycin binding to an AAAAA DNA binding site. *Nucleic Acids Res* **2005**, *33*, 725-733.
25. Alongi, K. S.; Shields, G. C. Theoretical Calculations of Acid Dissociation Constants: A Review Article. *Ann Rep Comp Chem* **2010**, *6*, 113-138.
26. Tang, C. L.; Alexov, E.; Pyle, A. M.; Honig, B. Calculation of pK(a)s in RNA: On the structural origins and functional roles of protonated nucleotides. *Journal of Molecular Biology* **2007**, *366*, 1475-1496.
27. Ho, J.; Coote, M. L. pK(a) Calculation of Some Biologically Important Carbon Acids - An Assessment of Contemporary Theoretical Procedures. *J Chem Theory Comput* **2009**, *5*, 295-306.
28. Mongan, J.; Case, D. A. Biomolecular simulations at constant pH. *Curr Opin Struct Biol* **2005**, *15*, 157-163.
29. Mongan, J.; Case, D. A.; McCammon, J. A. Constant pH molecular dynamics in generalized born implicit solvent. *J Comput Chem* **2004**, *25*, 2038-2048.
30. Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. Protonation States of the Key Active Site Residues and Structural Dynamics of the glmS Riboswitch As Revealed by Molecular Dynamics. *J Phys Chem B* **2010**, *114*, 8701-8712.
31. van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. Gromacs User Manual version 3.3. *www.gromacs.org* **2005**.
32. Serganov, A.; Patel, D. J. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet* **2007**, *8*, 776-790.
33. Tsukiji, S.; Ramaswamy, K.; Suga, H. Ribozymes that use redox cofactors. *Pure Appl Chem* **2004**, *76*, 1525-1536.
34. Abera, G.; Berhanu, G.; Tekewe, A. RIBOZYMES: NUCLEIC ACID ENZYMES WITH POTENTIAL PHARMACEUTICAL APPLICATIONS - A REVIEW. *Pharmacophore (An International Research Journal)* **2012**, *3*, 164-178.
35. Butcher, S. E. Structure and function of the small ribozymes. *Curr Opin Struct Biol* **2001**, *11*, 315-320.
36. Walter, N. G.; Burke, J. M. The hairpin ribozyme: structure, assembly and catalysis. *Curr Opin Chem Biol* **1998**, *2*, 24-30.
37. Mlynsky, V.; Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. QM/MM Studies of Hairpin Ribozyme Self-Cleavage Suggest the Feasibility of Multiple Competing Reaction Mechanisms. *J Phys Chem B* **2011**, *115*, 13911-13924.
38. Kath-Schorr, S.; Wilson, T. J.; Li, N. S.; Lu, J.; Piccirilli, J. A.; Lilley, D. M. J. General Acid-Base Catalysis Mediated by Nucleobases in the Hairpin Ribozyme. *J Am Chem Soc* **2012**, *134*, 16717-16724.
39. Goh, G. B.; Knight, J. L.; Brooks, C. L. pH-Dependent Dynamics of Complex RNA Macromolecules. *J Chem Theory Comput* **2013**, *9*, 935-943.
40. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers. *Biophys J* **2007**, *92*, 3817-3829.

41. Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J Comput Chem* **2005**, *26*, 1701-1718.
42. Salter, J.; Krucinska, J.; Alam, S.; Grum-Tokars, V.; Wedekind, J. E. Water in the active site of an all-RNA hairpin ribozyme and effects of Gua8 base variants on the geometry of phosphoryl transfer. *Biochemistry* **2006**, *45*, 686-700.
43. Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G. R.; D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*, University of California: San Francisco: **2012**.
44. Joung, I. S.; Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* **2008**, *112*, 9020-9041.
45. Zgarbova, M.; Luque, F. J.; Sponer, J.; Cheatham, T. E.; Otyepka, M.; Jurecka, P. Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J Chem Theory Comput* **2013**, *9*, 2339-2354.
46. DeLano, W. L. *The PyMOL Molecular Graphics System*, **2002**.
47. Song, B.; Zhao, J.; Griesser, R.; Meiser, C.; Sigel, H.; Lippert, B. Effects of (N7)-coordinated nickel(II), copper(II), or platinum(II) on the acid-base properties of guanine derivatives and other related purines. *Chem-Eur J* **1999**, *5*, 2374-2387.
48. Liu, L.; Cottrell, J. W.; Scott, L. G.; Fedor, M. J. Direct measurement of the ionization state of an essential guanine in the hairpin ribozyme. *Nat Chem Biol* **2009**, *5*, 351-357.
49. Mlynsky, V.; Kuhrova, P.; Zgarbova, M.; Jurecka, P.; Otyepka, M.; Sponer, J.; Banas, P. Reactive Conformation of the Active Site in the Hairpin ribozyme Achieved by Molecular Dynamics Simulations with ϵ/ζ Reparametrizations. (*článek v přípravě*).
50. Sigel, H.; Massoud, S. S.; Corfu, N. A. COMPARISON OF THE EXTENT OF MACROCHELATE FORMATION IN COMPLEXES OF DIVALENT METAL-IONS WITH GUANOSINE (GMP(2-)), INOSINE (IMP(2-)), AND ADENOSINE 5'-MONOPHOSPHATE (AMP(2-)) - THE CRUCIAL ROLE OF N-7 BASICITY IN METAL ION-NUCLEIC BASE RECOGNITION. *J Am Chem Soc* **1994**, *116*, 2958-2971.
51. Da Costa, C. P.; Sigel, H. Lead(II)-binding properties of the 5'-monophosphates of adenosine (AMP(2-)), inosine (IMP(2-)), and guanosine (GMP(2-)) in aqueous solution. Evidence for nucleobase-lead(II) interactions. *Inorganic Chemistry* **2000**, *39*, 5985-5993.
52. Da Costa, C. P.; Sigel, H. Acid-base and metal ion binding properties of guanylyl(3' \rightarrow 5')guanosine (GpG(-)) and 2'-deoxyguanylyl(3' \rightarrow 5')-2'-deoxyguanosine d(GpG)(-) in aqueous solution. *Inorganic Chemistry* **2003**, *42*, 3475-3482.