

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DISERTAČNÍ PRÁCE



Mgr. Kamil Maláč

Počítačové modelování biomolekul – potenciálních chemoterapeutik

Fyzikální ústav Univerzity Karlovy

Vedoucí disertační práce: RNDr. Ivan Barvík Jr., PhD.

Studijní program: Fyzika

Studijní obor: Matematické a počítačové modelování

Praha 2013

Rád bych zde poděkoval všem svým blízkým okolo mě, kteří se mnou měli trpělivost při psaní této práce a podporovali mě. Dále bych chtěl poděkovat svému školiteli RNDr. Ivanu Barvíkovi Jr., PhD. za odborné vedení, množství cenných rad a ochotu a zájem, s nímž mou práci sledoval.

Prohlašuji, že jsem tuto disertační práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V dne.....

podpis

Název práce: Počítačové modelování biomolekul – potenciálních chemoterapeutik

Autor: Mgr. Kamil Maláč

Ústav: Fyzikální ústav Univerzity Karlovy

Vedoucí disertační práce: RNDr. Ivan Barvík Jr., PhD.

Abstrakt:

Hlavní metodou uplatněnou v této práci byly klasické molekulárně-dynamické simulace. Simulovanými systémy byly komplexy RNA-dependentní-RNA polymerázy, Ribonukleázy H, Argonautu a Ribonukleázy L s chemicky modifikovanými nukleovými kyselinami. Motivací bylo využití těchto chemicky modifikovaných nukleových kyselin jako potenciálních chemoterapeutik. Výkonné grafické karty, prostřednictvím nichž byly molekulárně-dynamické simulace provedeny, umožnily získat trajektorie o délce stovek nanosekund až jedné mikrosekundy, což umožnilo postihnout rozdíly ve vazbě různě modifikovaných nukleových kyselin k výše uvedeným enzymům. Zjištěné rozdíly přitom odpovídaly experimentálním výsledkům, což otevírá prostor pro racionální návrh struktury potenciálních chemoterapeutik na bázi chemicky modifikovaných nukleových kyselin.

Klíčová slova: molekulární dynamika, polymeráza, ribonukleáza

Title: Computer modelling of biomolecules – potential chemotherapeutics

Author: Mgr. Kamil Maláč

Department: Institute of Physics of Charles University

Supervisor: RNDr. Ivan Barvík Jr., PhD.

Abstract:

Classical molecular dynamics simulations were applied on complexes of RNA-dependent RNA-polymerase, Ribonuclease H, Argonaute and Ribonuclease L with chemically modified nucleic acids, which are studied as potential chemotherapeutic agents. Powerful graphics processing units, through which these molecular dynamics simulations were performed, enabled to acquire trajectory length from hundreds of nanoseconds to one microsecond. Molecular dynamics simulations allowed capture differences in binding of various modified nucleic acids to the above mentioned enzymes. These identified differences fitted well with experimental results. It opens the door for rational design of the structure of potential chemotherapeutic agents based on chemically modified nucleic acids.

Keywords: molecular dynamics, polymerase, ribonuclease

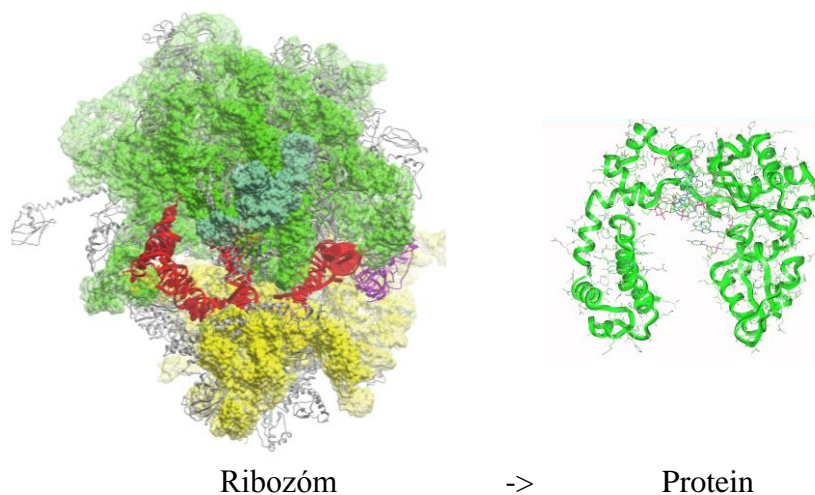
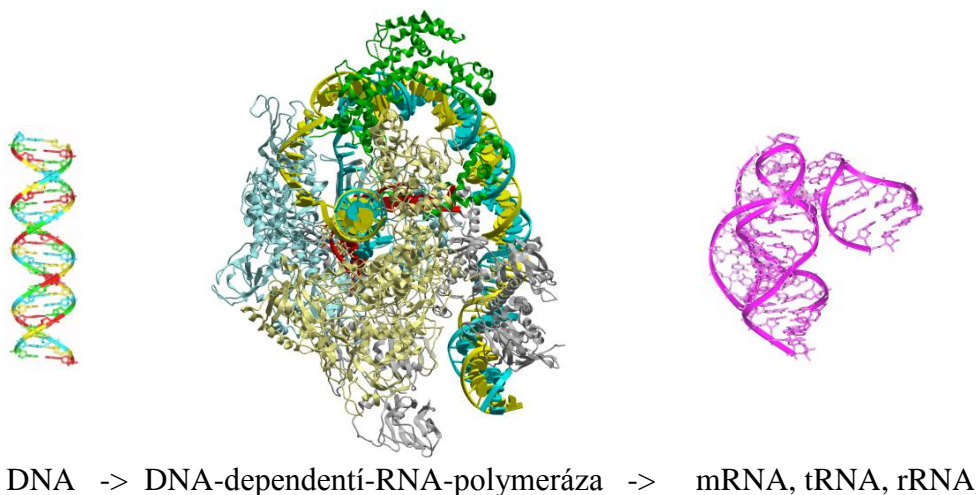
Obsah

1. Úvod	1
2. Biologická motivace	5
2.1. DNA dependentní RNA polymeráza	5
2.2. RNA dependentní RNA polymeráza	10
2.3. HIV reverzní transkriptáza	14
2.4. Ribonukleáza H	17
2.5. Argonaute	21
2.6. Ribonukleáza L	25
3. Metodika	28
3.1. Stacionární systémy	28
3.1.1. Popis stacionárních systémů	28
3.1.2. Silové pole AMBER	30
3.2. Dynamické systémy	33
3.2.1. Aproximace dynamických systémů	33
3.2.2. Numerické metody vhodné pro řešení pohybových rovnic	36
3.3. Použité modely	41
3.3.1. Termodynamické soubory – NVE, NVT, NPT	41
3.3.2. Periodické okrajové podmínky, Ewaldova sumace, Particle mesh metoda, Particle mesh Ewaldova metoda	44
3.4. Použitý software	49
4. Výsledky a diskuze	53
4.1. Molekulárně dynamické simulace komplexů RNA dependentní RNA polymerázy Noroviru s duplexem RNA:RNA a CTP či jeho analogy	53

4.2. Molekulárně dynamické simulace komplexů lidské Ribonukleázy H s fosfonátovými analogy oligonukleotidů	64
4.3. Molekulárně dynamické simulace komplexu enzymu Argonaute s duplexem DNA:RNA	79
4.4. Molekulárně dynamické simulace komplexů ankyrinové domény Ribonukleázy L s 2'-5' oligoadenyláty a jejich analogy	85
5. Závěr	99
Seznam použité literatury	103
Seznam tabulek	114
Seznam obrázků	115
Seznam použitých zkratk	120
Přílohy	122

1. Úvod

Prakticky všechny procesy probíhající v buňce jsou prováděny či modulovány proteiny popř. enzymy. Léky jsou obvykle malé molekuly, které se váží do tzv. aktivního místa proteinu a inhibují ho. Problémem při vývoji klasických léků je extrémně složitá prostorová struktura proteinů, která často ani nebývá známá na atomární úrovni. Jako potenciálně efektivnější se proto jeví terapeutická intervence již do procesu exprese genetické informace. Tedy selektivní zamezení vzniku nežádoucího proteinu.



Obrázek 1.1.: Schéma exprese genetické informace.

Informace o aminokyselinovém složení jednotlivých proteinů, nezbytné k jejich tvorbě, jsou uloženy jako jedinečná sekvence bází v genech v DNA. Vznikne-li v buňce potřeba syntetizovat určitý protein, je příslušná genetická informace přepsána z dvoušroubovice DNA do vlákna mRNA. Vlákno mRNA je syntetizováno tzv. DNA-dependentní-RNA polymerázou dle templátového DNA vlákna na principu komplementarity bází z jednotlivých nukleosidtrifosfátů. Tento krok exprese genetické informace je nazýván transkripce (Obrázek 1.1.).

mRNA následně přenáší genetickou informaci z jádra do cytoplasmy. Zde se mRNA váže s několika ribozómy – výsledný komplex se nazývá polyribosóm. Na polyribosómovém systému se uskutečňuje překlad (translace) genetické informace (Obrázek 1.1.). Pořadí aminokyselin v polypeptidovém řetězci určují vždy tři po sobě jdoucí báze v mRNA. Tyto tři báze dohromady tvoří kodon, který určuje buď konkrétní aminokyselinu nebo ukončení translace. Pro 20 proteinogenních aminokyselin máme 64 kodonů. Na kodon mRNA se vodíkovými vazbami váže antikodon tRNA (tRNA je specifická pro každou aminokyselinu). Trojice bází antikodonu tRNA musí být komplementární k nukleotidovému tripletu kodonu mRNA. Pohybem ribozómu po vláknu matrice mRNA (probíhá ve směru 5'→3') se postupně váží další molekuly tRNA přepravující aminokyseliny. Molekula tRNA, která se naváže na příslušný kodon mRNA, současně přebírá již vytvořenou část polypeptidového řetězce od předcházející molekuly tRNA a váže jej enzymaticky peptidickou vazbou s aminokyselinou, jejíž transport provedla. Objeví-li se na vláknu mRNA kodon, pro který neexistuje komplementární antikodon tRNA (tzv. stop kodon), ukončí se tvorba polypeptidického řetězce a protein se z ribozómu uvolní.

Tzv. postranskripční inhibice exprese genetické informace je možno dosáhnout prostřednictvím krátkých „antisense“ vláken nukleových kyselin komplementárních k mRNA, která s ní vytvářejí dvoušroubovicovité komplexy. Pokud je oním krátkým vláknem fragment DNA, je enzym RNáza H schopen selektivně degradovat pouze vlákno mRNA z duplexu DNA:mRNA. Pokud je oním krátkým vláknem nukleové kyseliny fragment RNA (tzv. siRNA), degraduje mRNA z duplexu siRNA:mRNA enzym Argonaute (Obrázek 1.2.). K selektivní inhibici exprese genetické informace konkrétního proteinu je tedy potřeba znát pouze jeho primární strukturu - sekvenci aminokyselin - z níž je možno přímočaře určit sekvenci bází v příslušné mRNA jakož i komplementárním „antisense“ vláknem.

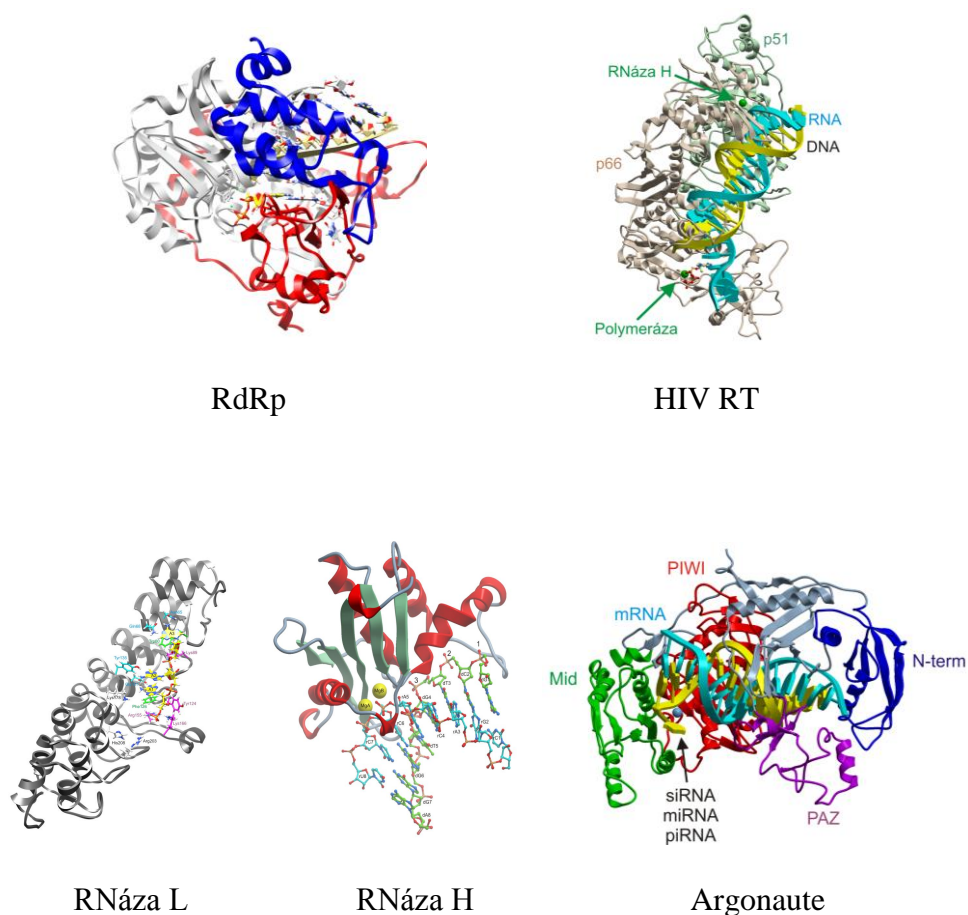
Replikaci virové RNA provádí enzym RNA-dependentní RNA polymeráza (RdRp). Přepis virové RNA do formy dvoušroubovice DNA provádí enzyme HIV reverzní transkriptáza (HIV RT) (Obrázek 1.2.). Virové polymerázy mají nesrovnatelně jednodušší strukturu než výše zmíněná mnohapoljednotková DNA-dependentní-RNA polymeráza, nicméně jejich aktivní místa jsou velmi podobná (a v tom ohledu, že substrát reakce je koordinován dvěma hořčíkovými ionty, se velmi podobají i aktivním místům výše zmíněných enzymů RNáza H a Argonaute (Obrázek 1.3.), jež nikoliv syntetizují, ale degradují nukleové kyseliny).

Zabránění exprese virové genetické informace je možno dosáhnout buď přímou inhibicí RdRp či HIV RT, nebo stimulací nespecifické degradace virové RNA enzymem RNáza L, který je aktivován krátkými fragmenty nukleových kyselin, které jsou pospojované atypickou 2'-5' fosfodiesterovou vazbou.

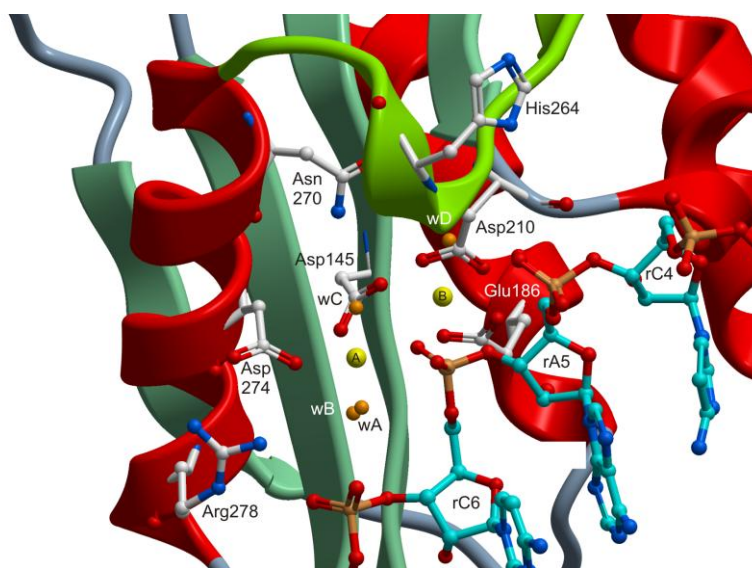
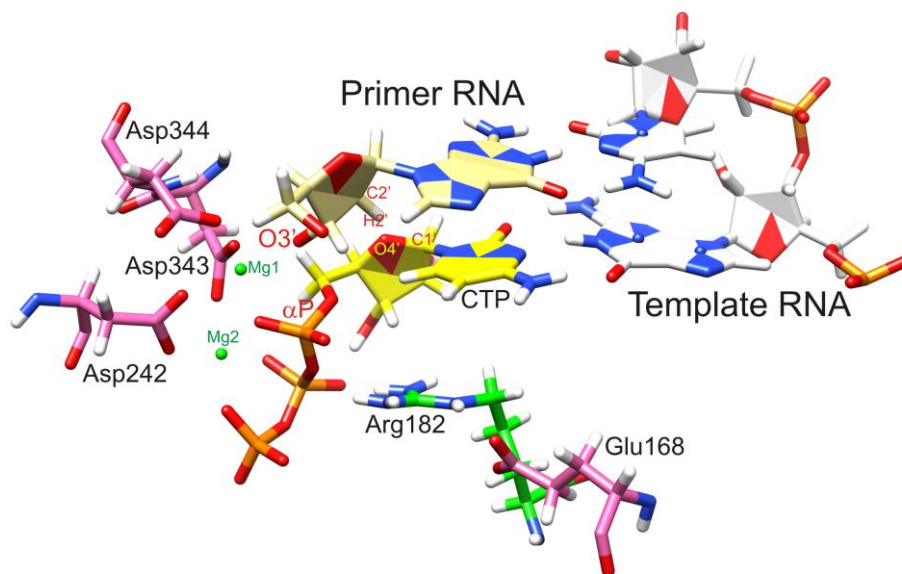
Předmětem zájmu této práce jsou enzymy syntetizující či degradující RNA (Obrázek 1.2.). Výsledkové kapitoly jsou věnovány RdRp, RNáze H, Argonautu a RNáze L. V teoretické části je navíc kapitola věnovaná mnohapoljednotkové DNA-dependentní-RNA polymeráze (jež má obdobnou funkci jako RdRp a jejíž produkt - mRNA - selektivně degradují RNáza H a Argonaute a neselektivně RNáza L). Teoretická část práce také obsahuje kapitolu věnovanou HIV reverzní transkriptáze,

neboť tento enzym obsahuje dvě aktivní místa, z nichž jedno je podobné aktivnímu místu RdRp a druhé aktivnímu místu RNázy H (Obrázek 1.2.).

Hlavní metodou použitou v této práci byly klasické molekulárně dynamické simulace. Simulovanými systémy byly komplexy RdRp, RNázy H, Argonautu a RNázy L s chemicky modifikovanými nukleovými kyselinami. Motivací je využití těchto chemicky modifikovaných nukleových kyselin jako potenciálních chemoterapeutik. Výkonné grafické karty, prostřednictvím nichž byly molekulárně dynamické simulace provedeny, umožnily získat trajektorie o délce stovek nanosekund až jedné mikrosekundy, což umožnilo postihnout rozdíly ve vazbě k výše uvedeným enzymům i u nepatrně odlišně modifikovaných nukleových kyselin. Zjištěné rozdíly přitom odpovídají experimentálním výsledkům, což otevírá prostor pro racionální návrh struktury potenciálních chemoterapeutik na bázi chemicky modifikovaných nukleových kyselin.



Obrázek 1.2.: Virové polymerázy (RdRp, HIV RT) a enzymy provádějící degradaci RNA (HIV RT - RNA v duplexech RNA:DNA; RNáza L - jednovláknová RNA; RNáza H - RNA v duplexech DNA:RNA; Argonaute - RNA v duplexech RNA:RNA). Tato práce se zabývá enzymy RdRp, RNázou L, RNázou H a Argonaute.



Obrázek 1.3.: Porovnání aktivních míst Norovirové RNA-dependentní-RNA polymerázy [1] (*nahoře*) a lidské Rázy H [2] (*dole*). Ačkoliv obě aktivní místa mají opačnou funkci (syntéza vs. degradace nukleové kyseliny), jejich uspořádání je velmi podobné. Několik záporně nabitých acidických Asp/Glu reziduí poutá dva kladně nabitě hořčíkové ionty, které přímo interagují se záporně nabitou nukleovou kyselinou. Aktivní místa všech polymeráz (mnohapodjednotkové DNA-dependentní-RNA polymerázy, virových RNA-dependentních-RNA polymeráz, HIV reverzní transkriptázy) jsou si přitom velmi podobná. Stejně tak jsou si podobná aktivní místa enzymů provádějících degradaci RNA (RNázové domény HIV RT, lidské RNázy H a enzymu Argonaute)

2. Biologická motivace

2.1. DNA dependentní RNA polymeráza

Funkce RNAP

Úkolem DNA dependentní RNA polymerázy (RNAP) (Obrázek 2.1.) je přepis-transkripce genetické informace z DNA do mRNA. Tj. katalýza syntézy vláken RNA z nukleosidtrifosfátů ATP, CTP, GTP a UTP na základě komplementarity bází s templátovým vláknem DNA. Působením RNAP vznikají 3'-5' fosfodiesterové vazby mezi nukleosidtrifosfáty. V každém kroku je přidáván jeden nukleosid ve směru 5'→3' [3].

Na rozdíl od DNA polymerázy (provádějící replikaci DNA), RNAP nepotřebuje zvláštní enzym pro rozvinutí DNA a syntéza mRNA může začít *de novo*, bez použití primeru [3]. Také je třeba zmínit, že chyby v přepisu jsou u RNAP podstatně četnější než u DNA polymerázy [3].

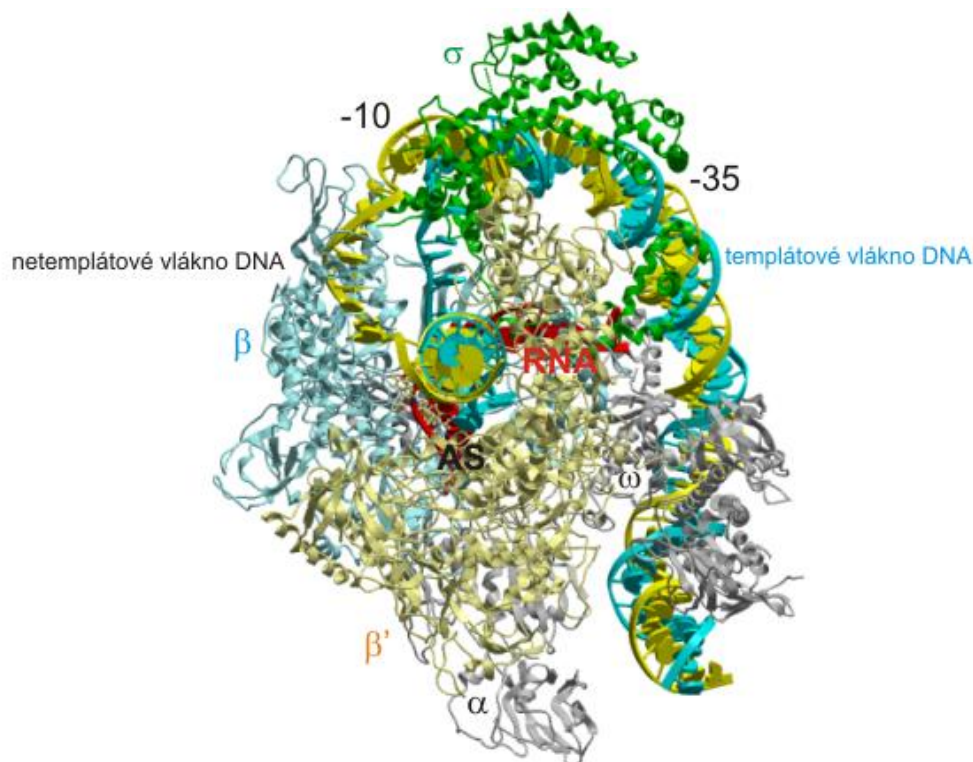
Struktura RNAP

Bakterie mají jen jednu RNAP. V buňkách vyšších organismů je hned několik specializovaných RNAP. Všechny mají velmi podobnou mnohopednostkovou strukturu. V následujících odstavcích je pro ilustraci popsána bakteriální varianta RNAP.

Tzv. holoenzym RNAP se skládá z jádra (core) RNAP, k němuž se váže iniciační σ -faktor [4]. Vlastní jádro RNAP se skládá z pěti podjednotek: β' , β , αI , αII a ω (Obrázek 2.1.). Evolučně jsou značně konzervované, ať už jde o sekvenci, strukturu či funkci, od bakterie až po člověka. Celková váha jádra RNAP je cca. 400kDa [3-6]. Největší podjednotka β' obsahuje část aktivního místa pro syntézu RNA. Druhá část aktivního místa se nachází ve druhé největší podjednotce β [4]. Tyto podjednotky jsou také klíčové pro sekvencně nespecifickou interakci s vlákny nukleových kyselin a regulačními transkripčními faktory jako je např. Gre [7-8]. Třetí největší je podjednotka α . V každé molekule RNAP se nacházejí dvě její identické kopie (αI a αII) [9]. Nejmenší podjednotkou je ω , která stabilizuje jádro RNAP a také usnadňuje jeho sestavení. Poslední část holoenzymu σ -faktor (Obrázek 2.1.) zvyšuje afinitu RNAP ke specifické promotorové oblasti na počátku genu a napomáhá tím zahájení transkripce na správném místě [3]. U σ -faktoru se rozlišují čtyři hlavní oblasti ($\sigma 1-4$) [4-5].

-10 a -35 element

Tzv. -10 a -35 elementy jsou dvě krátké sekvence nukleotidů nacházející se v promotorové oblasti DNA zhruba 10 a 35 nukleotidů před místem, v němž začíná transkripce [3, 5]. Konsenzuální sekvence pro element -35 je TTGACA a pro element -10 TATAAT [3]. V různých promotorech se obvykle zachovávají pouze tři nebo čtyři báze z šesti uvedených. Tyto sekvence jsou rozpoznávány bakteriálním σ -faktorem. Konkrétně element -10 je rozpoznáván regionem $\sigma 2.4$ a element -35 regionem $\sigma 4.2$ (Obrázek 2.2.) [4].



Obrázek 2.1.: Mnohapodjednotková DNA-dependentní-RNA polymeráza. Barevně odlišeny jsou největší podjednotky β (světlomodrá), β' (světležlutá) a σ (zelená). Dále je vyznačena poloha podjednotek α a ω . Dále jsou znázorněna vlákna nukleových kyselin vytvářející tzv. transkripční bublinu. Netemplátové vlákno DNA je obarveno žlutě, templátové vlákno DNA modře a vlákno syntetizované RNA červeně. Poloha aktivního místa je naznačena prostřednictvím symbolu AS.

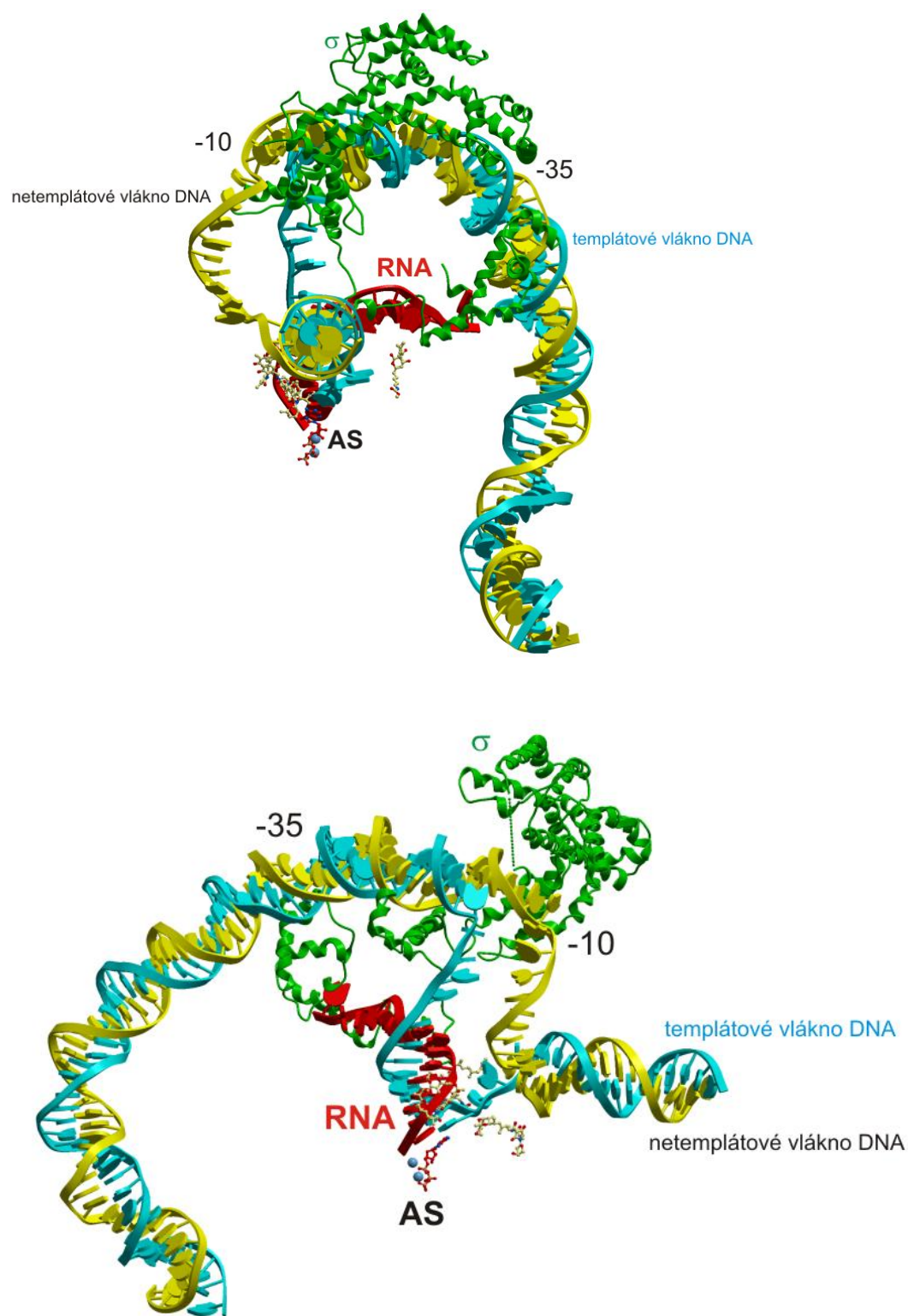
Aktivní místo RNAP

Ve štěrbině mezi podjednotkami β' a β leží aktivní místo, které je u všech bakteriálních RNAP velmi podobné. Tři Asp residua z konzervovaného motivu NADFDGD v β' váží dva hořčíkové ionty [3, 5-6], které stabilizují α -fosfát NTP a 3'OH konec syntetizovaného vlákna RNA, což v konečném důsledku umožní jejich propojení [3, 5-6].

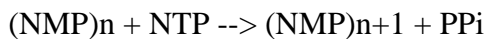
Transkripční bublina

RNA polymeráza se navazuje k promotorové oblasti DNA, která je rozpoznávána pomocí σ -faktoru [3]. Po navázání vytvoří RNAP s DNA tzv. transkripční bublinu (Obrázek 2.2.), což je cca. 17 párů bází dlouhý úsek, v němž jsou jednotlivá vlákna dvoušroubovice DNA oddělena [3, 5]. Tzv. templátové vlákno DNA a syntetizované vlákno RNA pak vytváří krátký (8-9 párů bází) heteroduplex DNA-RNA nad aktivním místem [3, 5-6].

Vzhledem k silné vazbě RNAP na promotor nezapočne obvykle její pohyb podél DNA na první pokus. Výsledkem je syntéza krátkých úseků RNA [4-6]. Kompletní přepis genu proběhne, až když jsou přerušeny kontakty se σ -faktorem [3-4]. Přitom jsou ribonukleotidy přidávány k 3'-konci RNA vlákna dle reakční rovnice:



Obrázek 2.2.: Interakce σ podjednotky (zelená) DNA-dependentní-RNA polymerázy s tzv. transkripční bublinou. Netemplátové vlákno DNA je obarveno žlutě, templátové vlákno DNA modře a vlákno syntetizované RNA červeně. Poloha aktivního místa je naznačena prostřednictvím symbolu AS. Jsou v něm patrné dva hořčikové ionty (modré kuličky) a přistupující nukleosidtrifosfát (NTP).



rychlostí cca. 50 nukleotidů za sekundu [3, 5, 10].

Prodlužování RNA prostřednictvím RNAP zahrnuje i opravný mechanismus, v rámci kterého mohou být nahrazeny nesprávně včleněné báze.

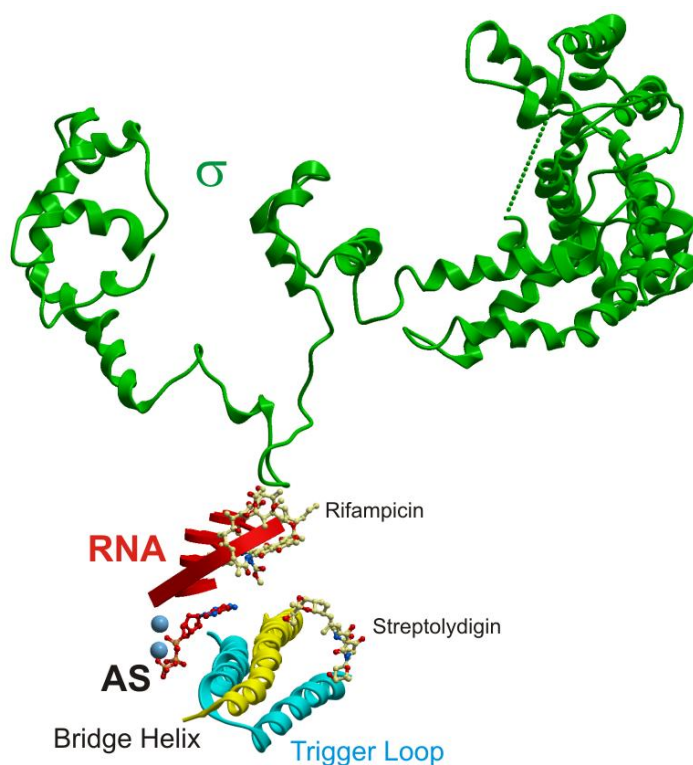
Ukončení transkripce provádí speciální bakteriální protein RHO patřící do třídy hexamerických helikáz závislých na ATP. RHO se váže na tzv. RHO utilizační místo, což je oblast syntetizovaného vlákna RNA s nízkým obsahem guaninu a vysokým obsahem cytosinu [3]. Následně se RHO faktor pohybuje po tvořeném RNA vlákne směrem k jeho 3'-konci. Dohání RNAP a destabilizuje interakce mezi DNA templátem a tvořeným vláknem RNA.

Při tzv. intrinsické (bez RHO kofaktoru) terminaci transkripce se využívá palindromická sekvence obsažená v tvořeném vlákně RNA. Protein nusA, který je navázán na RNAP, se naváže i na onu palindromickou smyčku a pozastaví pohyb RNAP [11].

Někdy dochází k tomu, že více RNAP používá ve stejnou chvíli jeden DNA templát k paralelní transkripci. Tímto způsobem může být z jednoho genu rychle vytvořeno velké množství vláken mRNA.

Inhibitory RNAP

V klinické praxi se již více než 40 let používá antibiotikum Rifampicin (Obrázek 2.3.), které se váže nad aktivní místo RNAP a stericky brání tvorbě vláken RNA delších než 2-3 nukleotidy [3, 5]. Dalším významným antibiotikem je Actinomycin D, který se váže na dimer GC vlákna DNA a blokuje tím transkripci [3]. Chemicky modifikovaná analoga NTP (s nimiž se potkáme v následujících kapitolách) se jako antibiotika neuplatňují, neboť aktivní místa lidské a bakteriální RNAP jsou téměř identická. Proto pokud modifikované NTP vykazuje významnou antibakteriální aktivitu, tak bývá obvykle i značně toxické.



Obrázek 2.3.: Interakce iniciační σ podjednotky (zelená) s RNA (červená) v aktivním místě DNA-dependentní-RNA polymerázy. Poloha aktivního místa je naznačena prostřednictvím symbolu AS. Jsou v něm patrné dva hořčíkové ionty (modré kuličky) a přístupující nukleosidtrifosfát (NTP). Dále jsou schematicky vyznačeny helixy tzv. Bridge Helixu (BH) a Trigger Loopu (TL), jejichž pohyb otevírá a uzavírá aktivní místo RNAP a přispívá k extrémní přesnosti přepisu genetické informace. Antibakteriální látka Streptolydigin inhibuje RNAP sterickou blokádou BH a TL. Antibiotikum Rifampicin narušuje interakci mezi σ podjednotkou RNAP a nukleovými kyselinami vytvářejícími transkripční bublinu. V konečném důsledku tak zabraňuje prodloužení RNA na délku větší než 4-5 nt.

2.2. RNA dependentní RNA polymeráza

Funkce RdRp

RNA dependentní RNA polymeráza (RdRp) replikuje genetickou informaci RNA virů [12-13] jako jsou např. virus hepatitidy typu C (HCV) způsobující cirhózu jater a virus Norwalk (NV), který u lidí způsobuje epidemickou nebakteriální akutní gastroenteritidu [12-13].

Struktura RdRp

Délka HCV genomu je cca. 9600 nukleotidů. HCV RdRp je zakódována v genu NS5B [14-15]. Celý protein je složený z 591 reziduí (66 kDa) [16-18]. HCV RdRp bývá ukotvená k buněčné membráně C-terminální peptidovou kotvou dlouhou 21 aminokyselinových zbytků [10-11, 19]. Katalytická část HCV RdRp sestává z 531 aminokyselinových zbytků, které tvoří 21 α -helixů a 18 β -strandů [14, 18, 20-21]. 3D struktura RdRp je podobná tvaru pravé ruky. Jednotlivé domény připomínají dlaň, prsty a palec. Prst a palec obklopují aktivní místo v dlani a vytvářejí žlábek pro navázání templátového vlákna RNA a kanálek, kterým pronikají NTP do aktivního místa. Zatímco struktury palců a prstů se u různých RdRp velmi liší, dlaň s katalytickým centrem je pozoruhodně konzervovaná [20]. To svědčí o společném původu RdRp [12].

Na rozdíl od ostatních polymeráz jsou však strukturní domény prstů a palce u RdRp propojeny a aktivní místo je jimi obklopeno. V důsledku propojení prstů a palce nejsou tyto domény schopny provádět nezávislé konformační změny [21], které by např. umožnily přemístění předpřipraveného vznikajícího páru bazí z přípravného do katalytického aktivního místa. To by totiž vyžadovalo značné pohyby podjednotky prstů. Proto RdRp pro zachycení komplementárních NTP používají templátovou bázi plně umístěnou v aktivním místě [22].

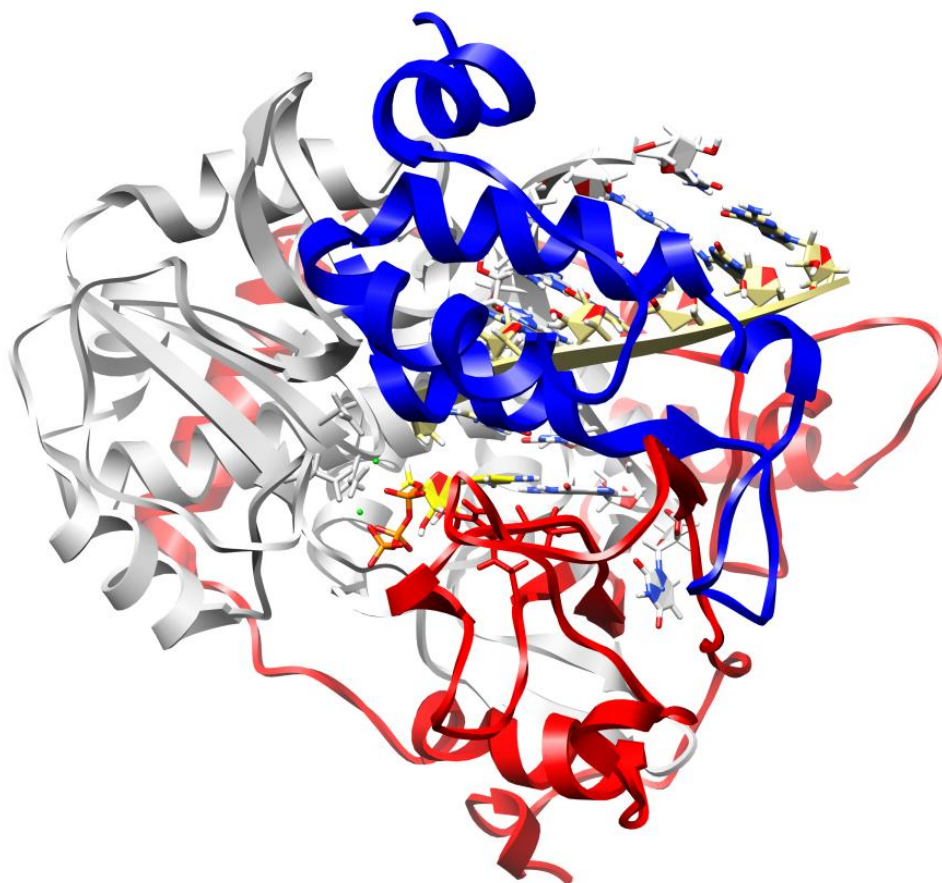
Obecně se cyklus katalýzy využívaný polymerázami skládá z pěti kroků. Nejprve se nukleotidový substrát váže do komplexu polymeráza-primer-templát. V dalším kroku tento komplex izomerizuje do katalyticky aktivní formy. Následuje krok transferu fosforylu a po něm druhá konformační změna. Poslední krok je závěrečné uvolnění pyrofosfátu [22].

Vzhledem k ohraničenosti aktivního místa používají RdRp pro katalýzu pouze nevelké strukturní přeskupení v jednotce dlaně. Nejdůležitějším krokem při uzavírání aktivního místa je u NV/3Dpol RdRp posunutí tzv. motivu A v podjednotce dlaně, které umísťuje do aktivního místa rezidua Asp233 a Asp242, která jsou důležitá pro vázání kovových iontů a následnou katalýzu [22-23].

HCV RdRp má cca. 30Å od aktivního místa vazebné místo s nízkou afinitou pro GTP. Předpokládá se, že se jedná o alosterický regulátor interakcí mezi doménami palce a prstů [14-15, 20-21].

NV RdRp (Obrázek 2.4.) má N-terminální oblast spojující podjednotky prstů a palce [13, 24] a C-terminální konec vážící se v nepřítomnosti substrátu do štěrbin aktivního místa [23]. Navázání RNA templátu odsune C-terminální konec ze štěrbin aktivního místa a umožní tak centrálnímu helixu palce rotovat (o cca. 22 stupňů) a interagovat se syntetizovaným vláknem RNA resp. s minor žlábkem RNA:RNA duplexu (Obrázek 2.4.) [23]. Přestože jsou strukturně velmi rozdílné,

připomíná umístění fragmentu C-terminálního konce NV RdRp ve spáře aktivního místa C-terminální konec a β -loop palce HCV RdRp, které napomáhají iniciaci syntézy RNA [24].

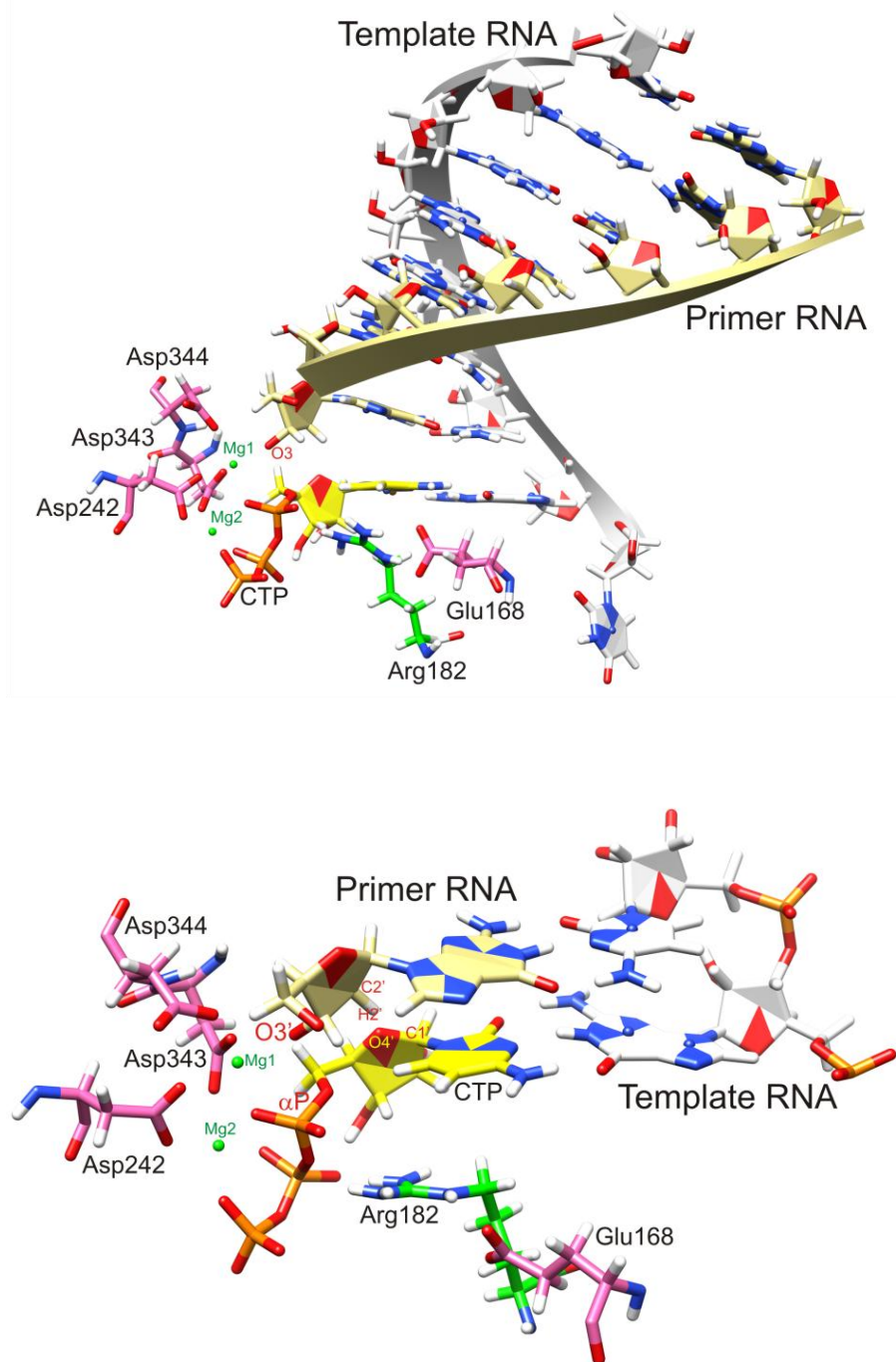


Obrázek 2.4.: Norovirová RNA-dependentní-RNA polymeráza (NV RdRp) v komplexu s nukleovou kyselinou (PDB id: 3BSO [23]). NV RdRp má typickou strukturu polymeráz připomínající pravou ruku (dlaň - šedá, palec - modrá, prsty - červená). V RdRp bývají palec a prsty spojené a tak nemohou měnit své konformace nezávisle [1].

Aktivní místo RdRp

V aktivním místě NV RdRp vytváří NTP vodíkové vazby s komplementární bází v templátovém vlákne RNA. Vklíní se přitom mezi 3'-konec syntetizovaného vlákna a postranní řetězec Arg182 (Obrázek 2.5.). Argininové residuum přitom zaujímá v blízkosti α -fosfátu NTP podobnou pozici ve všech známých strukturách virových RdRp [23]. 2'-OH skupina NTP vytváří vodíkové vazby s Asn309 a Ser300, který se váže k Asp247. Síť vodíkových vazeb mezi těmito konzervovanými residui umožňuje RdRp rozlišovat NTP od dNTP [23].

Dva hořčíkové ionty v aktivním místě NV RdRp interagují se třemi konzervovanými residui Asp a s trifosfátovou částí NTP tak, aby katalyzovaly reakci pomocí mechanismu „dvou kovových iontů“. Ion A je zkoordinován s 3'-OH skupinou syntetizovaného vlákna RNA, α -fosfátem NTP, Asp242, Asp343, Asp344 a také



Obrázek 2.5.: CTP v aktivním místě NV RdRP v kontextu primerového a templátového vlákna RNA. Mg^{2+} ionty se váží na konzervovaná rezidua Asp344, Asp343 a Asp242 v doméně dlaně NV RdRp [1].

s molekulou vody (Obrázek 2.5.). Interakcí s fosfodiesterovou skupinou syntetizovaného vlákna RNA a α -fosfátovou skupinou NTP se tato molekula vody aktivuje, což jí umožňuje oddělit proton z 3'-OH skupiny syntetizovaného vlákna RNA, která pak může atakovat α -fosfátovou skupinou NTP.

Iont B je zkoordinován s residui Asp242, Asp343, Tyr243 a také s jedním atomem kyslíku z α , β a γ fosfátových skupin NTP (Obrázek 2.5.) [23].

Iniciace

Replikace RNA začíná v okolí 3'-konce templátového vlákna RNA jedním z několika možných způsobů. Nejčastěji se tak děje *de novo* (bez použití primeru) spojením dvou NTP [13, 20].

U picornavirů (včetně NV [25]) bývá jako primer použit protein VPG [20, 26-27]. Aby protein mohl fungovat jako primer, musí poskytovat volnou OH skupinu. Toho bývá dosaženo jeho „uridylicí“ na třetí pozici, kde se nachází tyrosinové residuum [13, 28]. Výsledkem je VPGpUpUOH [26]. Calicivirus, je v nepřítomnosti VPG neinfekční, zatímco poliovirus je infekční i v nepřítomnosti VPG.

Další možností iniciace přepisu genetické informace je "cap-snatching" mechanismus popsán např. u viru chřipky A. Z mRNA hostitelské buňky jsou pomocí virové endonukleázy odštěpnuty úseky o délce 10-20 nukleotidů, které jsou následně použity jako primery pro syntézu virové RNA [23, 29-30].

Inhibitory RdRp

RdRp je jedním z klíčových cílů při vývoji nových léků proti RNA virům [21]. V buňkách savců za normálních okolností není RdRp přítomna a mnohapanadotková DNA dependentní RNA polymeráza je velmi odlišná. To dává jistou naději, že inhibitory RdRp nebudou toxické.

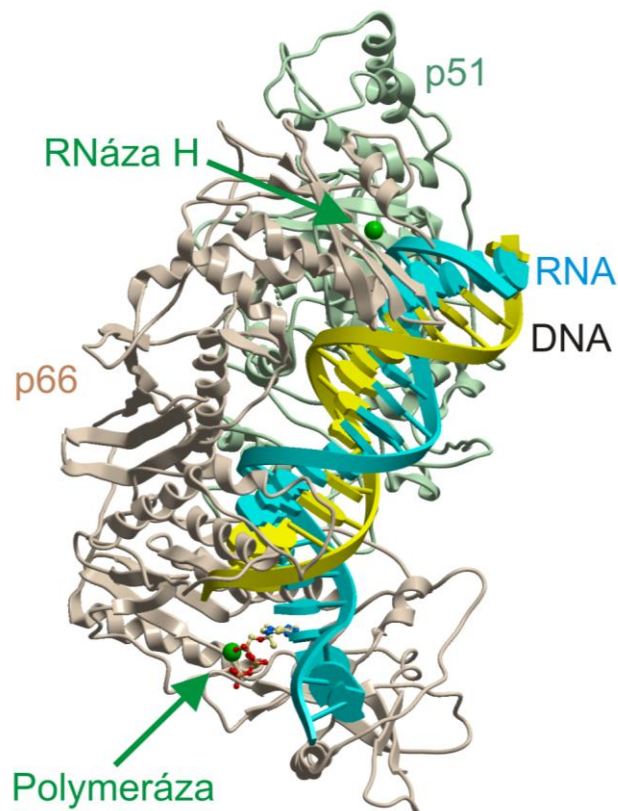
V současné době jsou známy tři skupiny inhibitorů HCV RdRp, které procházejí klinickými zkouškami: nukleosidové (NI), non-nukleosidové (NNI) a analoga pyrofosfátu (PPi) [31].

NI se vážou do aktivního místa RdRp, kde konkurují přirozeným NTP. Pokud jsou zapracovány do tvořeného vlákna RNA, způsobí to ukončení replikace virového genomu [16], ačkoliv NI většinou mají 3'-hydroxylovou skupinu, která v principu umožňuje připojení dalšího nukleotidu. Tomu však brání chemická modifikace ve struktuře NI, která stericky zabraňuje navázání dalšího NTP do aktivního místa RdRp [31]. Příkladem potentního NI je C2'-methyl-adenosin [32].

NNI se vážou buď do alosterických kavit na povrchu RdRp nebo přímo do jejího aktivního místa [16]. U HCV RdRp byla objevena celkem čtyři alosterická vazebná místa. Zatímco místa I a II jsou lokalizována v palci, místa III a IV se nacházejí v dlani poblíž aktivního místa [31].

PPi inhibitory napodobují zbytkový pyrofosfát uvolněný v průběhu reakce z NTP. Skládají se ze dvou částí: kotevní - která interaguje se dvěma katalytickými ionty v aktivním místě a specifické - která se váže na jiné části RdRp [31].

2.3. HIV reverzní transkriptáza



Obrázek 2.6.: HIV reverzní transkriptáza sestávající ze dvou podjednotek (p51 a p66). Templátové vlákno virové RNA je obarveno modře. Syntetizované vlákno DNA je obarveno žlutě. V dolní části je polymerázové aktivní místo (syntetizující DNA) s přístupujícím dNTP. V horní části je aktivní místo domény RNázy H degradující vlákno virové RNA.

Funkce HIV RT

Schéma přenosu genetické informace dle centrálního dogmatu molekulární biologie (zformulovaného Crickem) je následující:

DNA → transkripce → mRNA → translace → PROTEIN

Během tzv. transkripce se tedy genetická informace přenáší z DNA do RNA. Cílem procesu, který provádí reverzní transkriptáza (RT), je naopak přenést genetickou informaci viru uloženou v molekule RNA do molekuly DNA [33-34]. Začlenění vysyntetizované DNA s genomem viru do DNA v hostitelské buňce se následně děje pomocí retrovirální integrázy [35-36]. Nejznámějším virem, který využívá tento mechanismus je virus HIV. Genetická informace viru HIV sestává pouze z devíti genů, v nichž je uložena veškerá informace nutná pro produkci proteinů tvořících virový obal a všech enzymů - reverzní transkriptázy, integrázy, ribonukleázy a proteázy [35]. HIV RT má dvě aktivní místa (Obrázek 2.6.) a vykonává tři biochemické aktivity. Postupně funguje jako RNA dependentní DNA polymeráza, RNáza H a DNA dependentní DNA polymeráza [35-37]. U virů postrádajících DNA dependentní DNA polymerázovou aktivitu je druhé vlákno duplexu virové DNA

vytvořeno pomocí hostitelské DNA polymerázy δ . HIV RT patří mezi DNA polymerázy. Pro zahájení přepisu genetické informace potřebuje primer s hydroxylovou skupinou na 3' konci. V blízkosti 3' konce RNA genomu viru HIV se nachází 18 nukleotidů, které tvoří vazebné místo primeru (PBS). To je komplementární k osmnácti nukleotidům u 3' konce hostitelské tRNA^{Lys,3}, která slouží jako primer [34, 36].

Struktura HIV RT

HIV1 RT je heterodimer, který se skládá ze dvou podjednotek (p66 a p51), které připomínají pravou ruku včetně dělení na dlaň, prsty a palec [33, 36-37]. Podjednotka p66 je větší (66-kDa, celkem 560 aminokyselinových zbytků) než p51 (51-kDa, 440 aminokyselinových zbytků). Obě podjednotky jsou přitom kódovány jedním genem. p51 vznikne z p66 po syntéze proteolytickým odstraněním C-terminální RNázové domény virovou proteázou [34]. Navzdory stejné sekvenci aminokyselinových zbytků je prostorové uspořádání navzájem si odpovídajících oblastí podjednotek p66 a p51 rozdílné [34, 37].

Aktivní místo polymerázy

V dlani podjednotky p66 se nachází aktivní místo polymerázy HIV RT (Obrázek 2.6.). V tomto aktivním místě se nachází tři katalytická residua Asp110, Asp185 a Asp186, která vážou dva divalentní ionty nezbytné pro katalýzu. *In vivo* se s největší pravděpodobností jedná o hořčikové ionty Mg²⁺ [34-36]. Tyto ionty koordinují atomy kyslíku ze všech tří fosfátových skupin příchozího dNTP a stabilizují náboje reakčních meziproductů. Residua Asp110 a Asp186 váží iont Mg²⁺, který je navázán také na β a γ fosfátové skupiny příchozího dNTP. Druhý iont Mg²⁺ koordinuje kromě Asp110 a Asp186 ještě Asp185 a α -fosfátovou skupinu dNTP [34]. Ionťová koordinace usnadňuje atak aktivovaného 3'-OH konce syntetizovaného vlákna DNA na α -fosfátovou skupinu příchozího dNTP.

Residua Asp185 a Asp186 jsou součástí motivu Tyr-X-Asp-Asp, který je společný mnoha retrovirálním RT [36]. Tento motiv utváří β -turn a zaujímá konformaci II', která je typická pro β -hairpiny. U HIV1 RT je v proměnné pozici X residuum Met184. U ostatních retrovirů se v ekvivalentní pozici vyskytují residua Val, Leu, nebo Ala [34]. Met184 je stabilizován vodíkovou vazbou na Gln182 [37]. Vazebné místo pro dNTP spoluvytváří u HIV1 RT i další residua, jež mají své protějšky i u jiných reverzních transkriptáz. Jsou to Tyr115, Arg72, Lys65 a Gln151. Tyr115 tvoří sterickou bránu, která od sebe rozlišuje dNTP a NTP. Arg72 a Lys65 vytvářejí solné můstky s β - a γ -fosfátovou skupinou příchozího dNTP. Gln151 interaguje s 3'-OH skupinou příchozího dNTP a stabilizuje postranní řetězec residua Arg72 [34, 36].

Aktivní místo RNázy H

Aktivní místo RNázy H (Obrázek 2.6.) se nachází na opačném konci HIV RT než aktivní místo polymerázy [36]. RNáza H štěpí virovou RNA z hybridu RNA:DNA při tvorbě prvního vlákna DNA, které pak slouží jako templát při tvorbě druhého vlákna DNA. Dále pak RNáza H štěpí tRNA primer a PPT [33-34, 36].

Také HIV1 RNáza H používá mechanismus "dvou kovových iontů", které interagují s residui Asp443, Glu478, Asp498 a Asp549 [34]. Z biochemických studií je zřejmé,

že aktivita HIV1 RNázy H je vyšší se dvěma ionty Mn^{2+} než se dvěma ionty Mg^{2+} [38].

Vazba nukleové kyseliny k HIV RT

Vazebná štěrbina pro nukleovou kyselinu na povrchu HIV RT je utvořena tak, že duplex DNA:RNA je současně v kontaktu s aktivním místem polymerázy i RNázy H. Vzdálenost mezi těmito aktivními místy je cca. 60 Å - tedy 17 párů bází u DNA:DNA duplexu nebo 18 párů bází u RNA:DNA duplexu [34, 36]. Při syntéze duplexu DNA:DNA dojde k rotaci palce podjednotky p66 o cca. 32°. Děje se tak především díky konformačním změnám residuí Pro243 a Val317 [37]. Geometrie HIV1 RT se díky tomu nastaví do polohy optimální pro sevření vznikající nukleové kyseliny. Dochází přitom k rozsáhlému kontaktu mezi α -helixy palce podjednotky p66 a páteřemi vláken primeru a templátu [36-37]. Vlásačka β 12- β 13 v podjednotce p66 pomáhá stabilizovat 3'-konec syntetizovaného vlákna DNA v aktivním místě polymerázy [36-37].

Inhibitory HIV RT

HIV RT dělá při přepisu virové genetické informace relativně mnoho chyb. Na rozdíl od ostatních DNA polymeráz přitom není schopna tyto chyby opravit [33-34]. To často vede ke ztrátě účinnosti inhibitorů, které jsou cílené vůči virovým proteinům. Obvykle se totiž relativně rychle vyprofilují rezistentní kmeny viru s bodovými mutacemi ve vazebných místech inhibitorů [33, 36, 39]. Zejména se to týká tzv. alosterických inhibitorů, které se vážou mimo aktivní místo HIV RT, jež ovlivňují na dálku. V alosterických vazebných místech totiž může docházet k drobným variacím ve struktuře HIV RT, aniž by byla ohrožena její funkčnost.

Alternativou jsou inhibitory na bázi chemicky modifikovaných deoxynukleosidů (např. AZT), které postrádají terminální 3'OH skupinu. Ty se navážou do vznikajícího vlákna DNA, které se tak nemůže dále prodlužovat [33].

Výhodou chemicky modifikovaných deoxynukleosidů je to, že relativně snadno pronikají skrz buněčné membrány. Nevýhodou je pak to, že se v buňce musí třikrát fosforylovat, aby vzniklo NTP - substrát pro HIV RT. Zejména první fosforylace buněčnými enzymy je díky chemickým modifikacím obtížná.

Východiskem proto jsou chemicky modifikované nukleotidy (jako například PMPA/Tenofovir vytvořený A. Holým a I. Rosenbergem v UOCHB AV ČR), které již jednu fosfátovou skupinu obsahují a v buňce se tak musí fosforylovat pouze dvakrát. Díky přítomnosti fosfátové skupiny však špatně procházejí skrz buněčné membrány. Proto se v klinické praxi užívají jejich konjugáty s dalšími skupinami atomů, které jsou po průchodu skrz buněčnou membránu odštěpnuty buněčnými enzymy.

Ve světě je rozšířen zejména virus HIV1, který je velmi virulentní a infekční. Na území západní Afriky však převládá virus HIV2 [40]. Léčiva jsou vyvíjena zejména proti kmenu B viru HIV1 a to i přesto, že tímto kmenem je postiženo pouze 12% z celkového počtu celosvětově infikovaných jedinců. Je to způsobeno tím, že subtyp B je predominantní v západním světě, zatímco celosvětově dominují subtypy A a C [33, 39].

2.4. Ribonukleáza H

Funkce RNázy H

Ribonukleáza H (RNáza H) je enzym ze skupiny nukleáz (EC 3.1.26.4), který štěpí vlákna RNA hydrolyzací fosfodiesterových vazeb [41]. Jedná se o nespecifickou endoribonukleázu, která degraduje RNA v hybridních duplexech DNA/RNA. Neštěpí samotná vlákna RNA ani vlákno DNA z hybridního duplexu DNA/RNA. Rychlost degradace RNA podle typu báze je následující: rA > rU > rC > rG. Rychlost štěpení rA je přitom čtyřikrát větší než rG.

RNáza H pracuje prostřednictvím mechanismu “dvou kovových iontů” stejně jako funkčně podobné transportázy, retrovirální integrázy, resolvázy Hollidayových spojů a nukleáza Argonaute v komplexu RISC [42].

Výsledkem štěpení P-O3' vazby jsou produkty ukončené 3'-hydroxylovými a 5'-fosfátovými skupinami. Například ribonukleáza A zanechává naopak produkty s 3'-fosfátovými a 5'-hydroxylovými skupinami. Nukleové kyseliny zakončené 3'-hydroxylovými skupinami mohou sloužit jako substráty pro DNA polymerázu či DNA ligázu [43].

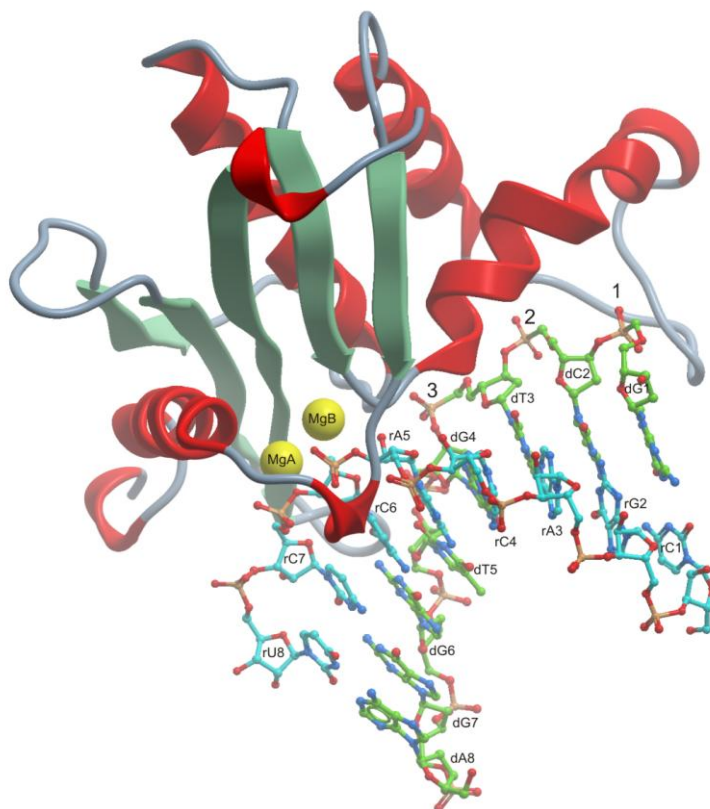
Při replikaci DNA odstraňuje RNáza H primer RNA a tím napomáhá kompletaci syntetizované DNA. RNáza H je nezbytná pro množení retrovirů, neboť je součástí virové reverzní transkriptázy [44-45], která odstraňuje vlákno RNA s genomem viru, štěpí tRNA primer a PPT [42-43, 45].

RNáza HI vs. RNáza HII

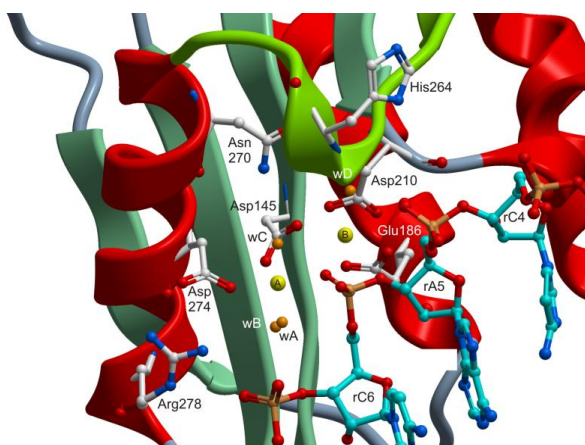
Existují dva hlavní typy RNázy H. Alespoň jeden z nich se nachází ve většině živých organismů [44, 46-47]. Savčí RNáza HI je větší (68-90 kDa) než RNáza HII (cca. 40 kDa). V baktérii *Escherichia Coli* se nachází RNáza HI i RNáza HII, jejíž úloha není doposud známa. Má totiž velmi slabou aktivitu a nemá žádnou sekvenční homologii s RNázou HI. Ve většině kompletně určených genomů archebakterií se nachází geny pouze pro tvorbu RNázy HII. Doposud byly objeveny pouze tři archebakterie nespňující toto pravidlo: *Halobacterium NRC-1*, *Sulfolobus tokodaii* a *Pyrobaculum aerophilum*. Genom těchto archebakterií obsahuje i geny pro RNázu HI, která je svými enzymatickými vlastnostmi blízká RNáze H retrovirů [47-48].

Struktura RNázy H

Prokaryotické RNázy HI mají pouze ribonukleázovou doménu. Eukaryotické RNázy HI obsahují navíc i N-terminální oblast pro navázání hybridů (HBD), která jim propůjčuje afinitu k DNA/RNA substrátu [44, 46]. Například lidská RNáza HI se tak skládá z 286 aminokyselinových zbytků (Obrázek 2.7.), kdežto RNáza HI bakterie *Escherichia Coli* ze 155 aminokyselinových zbytků. Sekvenční identita ribonukleázových domén lidské a *Escherichia Coli* RNázy HI je pouze 34%. Přesto mají překvapivě podobnou třidimenzionální strukturu [44, 46]. V retrovirálních reverzních transkriptázách je oblast RNázy H kovalentně navázána na N-terminální oblast polymerázy a postrádá díky tomu C-terminální α -helix s pozitivním nábojem, který zvyšuje vazebnou kapacitu enzymu [42-45].



Obrázek 2.7.: Pět internukleotidových spojek z 3' konce DNA se podílí na stabilizujících interakcích s povrchem HS RNázy H [2].



Obrázek 2.8.: Aktivní místo HS RNázy H [2].

Aktivní místo RNázy H

Prostorová struktura katalytického elementu RNázy HI se skládá z 5-vláknového β -sheetu obklopeného α -helixy [42]. Čtyři α -helixy (A, B, C a D) jsou na jedné straně β -sheetu a pátý α -helix (E) na straně druhé.

Residua v aktivním místě vytváří motiv Asp-Glu-Asp-Asp zachovávající se u RNáz HI i HII. U lidské RNázy HI jsou to residua Asp145, Glu186, Asp210 a Asp274 (Obrázek 2.8.) [44]. U RNázy H viru HIV1 mají analogickou funkci residua Asp443, Glu478, Asp498 a Asp549 [44-45]. Biochemické studie prováděné na mutantních proteinech potvrdily esencialitu těchto aminokyselin pro enzymatickou aktivitu. Residuum Asp145 je situované uprostřed vlákna 1 v β -sheetu, Glu186 v helixu A, Asp210 na konci vlákna 4 v β -sheetu a Asp274 v helixu E [44].

Jako kofaktor hydrolýzy probíhající mechanismem dvou kovových iontů působí ionty hořčíku [42-43]. Divalentní kovový iont A je koordinován residui Asp145, Asp274 a Sp kyslíkem ze štípané fosfátové skupiny RNA. Divalentní kovový iont B je koordinován pěti ligandy - residui aktivního místa Asp145, Glu186, Asp210 a kyslíky O3' a Sp ze štípané fosfátové skupiny (Obrázek 2.8.) [44].

Vazba nukleové kyseliny k RNáze H

V krystalové struktuře interaguje lidská RNáza HI s 11 z 18 párů bází hybridního duplexu DNA:RNA. Většina interakcí enzym-substrát se při tom realizuje podél tzv. minor žlábků duplexu [46]. Povrch lidské RNázy HI obsahuje dva žlábků oddělené residui Asn151, Asn182 a Gln183 [44, 48]. Páteře vláken RNA a DNA (Obrázek 2.7.) dosedají do těchto žlábků. Minor žlábek hybridního duplexu DNA:RNA interaguje s residui Asn151, Asn182 a Gln183.

Vlákno RNA je v konformaci A [44] a prostřednictvím čtyř 2'-OH skupin interaguje s enzymem v oblasti aktivního místa. Interakce mezi enzymem a vláknem RNA, zprostředkovávají rezidua Glu186, Cys148, Ser150, Asn151 a Met212 [46, 48].

Lidská RNáza HI interaguje také se sedmi nukleotidy vlákna DNA. Residua Arg179, Thr181 a Asn240 interagují s fosfátem nukleotidu, který se nachází dva páry bází před štípanou fosfodiesterovou vazbou [44, 48]. Žlábek skládající se z residui Trp221, Trp225 a Ser233 tvoří další důležité vazebné místo vlákna DNA. Jeho páteř zde interaguje prostřednictvím vodíkových vazeb a van der Waalsových interakcí [44, 48].

Antisense oligonukleotidy

Biologická role domény RNázy H u HIV RT je dobře prostudována a byla diskutována výše. Vzhledem k podobnostem ve struktuře a funkci s RNázou HI napadených buněk je třeba dávat pozor, aby případná léčiva zaměřená proti retrovirální replikaci neinterferovala s buněčným enzymem [46].

K terapeutickému zamezení exprese specifických genů lze použít tzv. antisense deoxyoligonukleotidy [42, 49], které stimulují RNázu H k degradaci cílové mRNA. Teoreticky by mělo stačit zvolit cílovou sekvenci nukleotidů a vnesením krátkých syntetických antisense deoxyoligonukleotidů s komplementární sekvencí bází do buněk získat pozitivní výsledky [49].

Přirozené antisense deoxynukleotidy jsou však rychle štěpeny buněčnými nukleázami a proto je nezbytné je chemicky modifikovat [50]. V minulosti byla testována celá řada chemických modifikací antisense oligonukleotidů. Pouze několik málo z nich však splňuje kritéria kladená na potenciální chemoterapeutika. Jedná se zejména o vyvolání dostatečné biologické aktivity lidské RNázy H, netoxičnost, odolnost proti nukleázám a schopnost Watson–Crickovského párování bází s přirozenou nukleovou kyselinou [50].

Fosforothioátové deoxyoligonukleotidy jsou schopné odolat nukleázám, zformovat heteroduplex a vyvolat degradaci mRNA prostřednictvím lidské RNázy H [49]. Jejich efektivita je však silně ovlivněna typem buňky, cílovou sekvencí a délkou oligonukleotidu [49-50]. Nespecificky bohužel interagují s buněčnými proteiny, což vyvolává četné vedlejší efekty.

Oligonukleotidy s fosfonátovou internukleotidovou spojkou s vazbou P-C jsou odolné vůči nukleázám. Konformační adaptibilita takto modifikovaných vláken DNA napomáhá tvorbě stabilních hybridních duplexů DNA:RNA [51-52]. Pokud fosfonátové spojky alternují s přirozenými spojkami, jsou příslušné deoxyoligonukleotidy schopné vyvolat aktivitu RNázy H.

2.5. Argonaute

Funkce AGO

Proteiny Argonaute (dělicí se na PIWI a AGO) jsou evolučně konzervované. Poprvé byly objeveny u rostlin a jsou pojmenovány po fenotypu Argonaute mutantů Huseníčku rolního (*Arabidopsis thaliana*) [53]. Proteiny AGO jsou esenciální pro vývoj u většiny rostlin a zvířat a také ochraňují buňky proti virové infekci. PIWI proteiny jsou unikátní u zvířat, kde potlačují expresi transposonních genů a zajišťují úspěšnou produkci spermatu a vajíček [53-55]. U eukaryotů byly zaznamenány vysoké koncentrace proteinů AGO v cytoplasmě v oblastech známých jako "P-bodies", kde dochází k rozkladu mRNA [54]. Proteiny z třídy Argonaute jsou rozšířeny i mezi archebakteriemi a některými bakteriemi [56]. Na základě studií prováděných pomocí komparativní genomiky se předpokládá, že se třída Argonaute, vyvinula z komponentů systému iniciace translace [57]. Je zde totiž významná sekvenční shoda s eIF2 (eukaryotickým Iniciačním Faktorem 2).

Proteiny Argonaute se podílejí na selektivní inhibici exprese genetické informace prostřednictvím tzv. RNA interference (RNAi) [54-55, 58, 59-60]. Vážou krátká vlákna RNA o délce 21-35 nukleotidů, pomocí nichž identifikují mRNA s komplementární sekvencí bází, kterou následně degradují anebo alespoň zabraňují translaci genetické informace v mRNA obsažené. Proteiny Argonaute a krátká vlákna RNA tvoří dohromady tzv. RNA-induced silencing complex (RISC), ve kterém jsou proteiny Argonaute katalytickou komponentou [60-61], jejíž úlohou je endonukleolytické štěpení specifické fosfodiesterové vazby cílové mRNA. Struktura proteinů Argonaute, zajišťuje, že štěpená vazba leží mezi desátým a jedenáctým nukleotidem v mRNA [53, 55-56, 58, 62]. Komplexy RISC mohou obsahovat i další proteiny, které rozšiřují funkci Argonaute proteinů nebo směřují RISC do specifických buněčných podoblastí [54]. Krátká vlákna RNA se dělí na endogenní miRNA (micro RNA) dlouhé 21-23 nukleotidů, exogenní (generované např. z virové dsRNA) siRNA (small interfering RNA) o délce 21 nukleotidů a piRNA (Piwi-interacting RNA) dlouhé 23-30 nukleotidů (Obrázek 2.9.) [53-54, 56, 58]. Krátké siRNA a miRNA duplexy jsou vyštěpnuté z dlouhých dvouvláknových RNA prekurzorů enzymy RNáza III jako je tzv. Dicer [53, 56, 58, 62-63].

Mnoho živých organismů má více než jeden protein Argonaute [53-54, 61-62]. Geny proteinů Ago1, Ago3, Ago4 jsou v lidském genomu shromážděny v prvním chromozomu, zatímco gen proteinu Ago2 se nachází v chromozomu osmém [54]. Endonukleázová (Slicer) aktivita vůči vláknům mRNA byla zaznamenána pouze pro lidské Ago2 a to i přesto, že katalytická rezidua jsou více méně konzervovaná i u ostatních variant lidského AGO [53-54, 58]. V současné době není známo, zda se funkce Ago1, Ago3 a Ago4 liší. Je možné, že se váží na různé regulační proteiny, které jim propůjčují specifitu [53-54].

Struktura AGO

Všechny proteiny Argonaute a Piwi se skládají ze čtyř klíčových domén - PAZ, MID, PIWI a N-terminální (Obrázek 2.10.), které jsou propojeny linkery L1 (pojí N-terminální a PAZ doménu) a L2 (spojuje PAZ a MID) [53-55, 58-59].

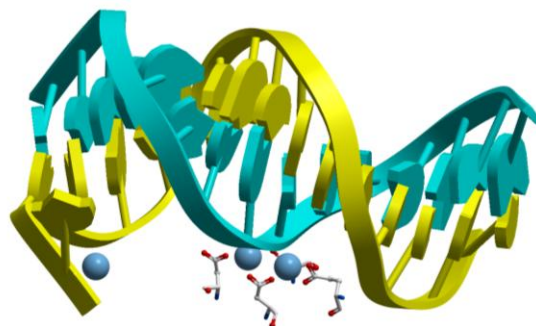
PIWI doména sestává z cca. 300 aminokyselinových zbytků a připomíná RNázu H [53-55, 58-59, 61]. V C-terminální oblasti domény PIWI se nachází endonukleázové

aktivní místo, které využívá Mg^{2+} ionty pro štěpení mRNA na produkty s 3'-hydroxylovou a 5'-fosfátovou skupinou. Podobně jako RNáza H i doména PIWI obsahuje čtyři evolučně konzervovaná a negativně nabitá rezidua (Asp-Asp-Asp-Glu), která tvoří katalytickou triádu vážící dva Mg^{2+} ionty (Obrázek 2.9.). Na rozdíl od RNázy H a antisense oligonukleotidů zůstávají krátké siRNA v komplexu s proteiny Argonaute po mnoho cyklů, během nichž se štěpí mnoho kopií cílové mRNA [55, 62].

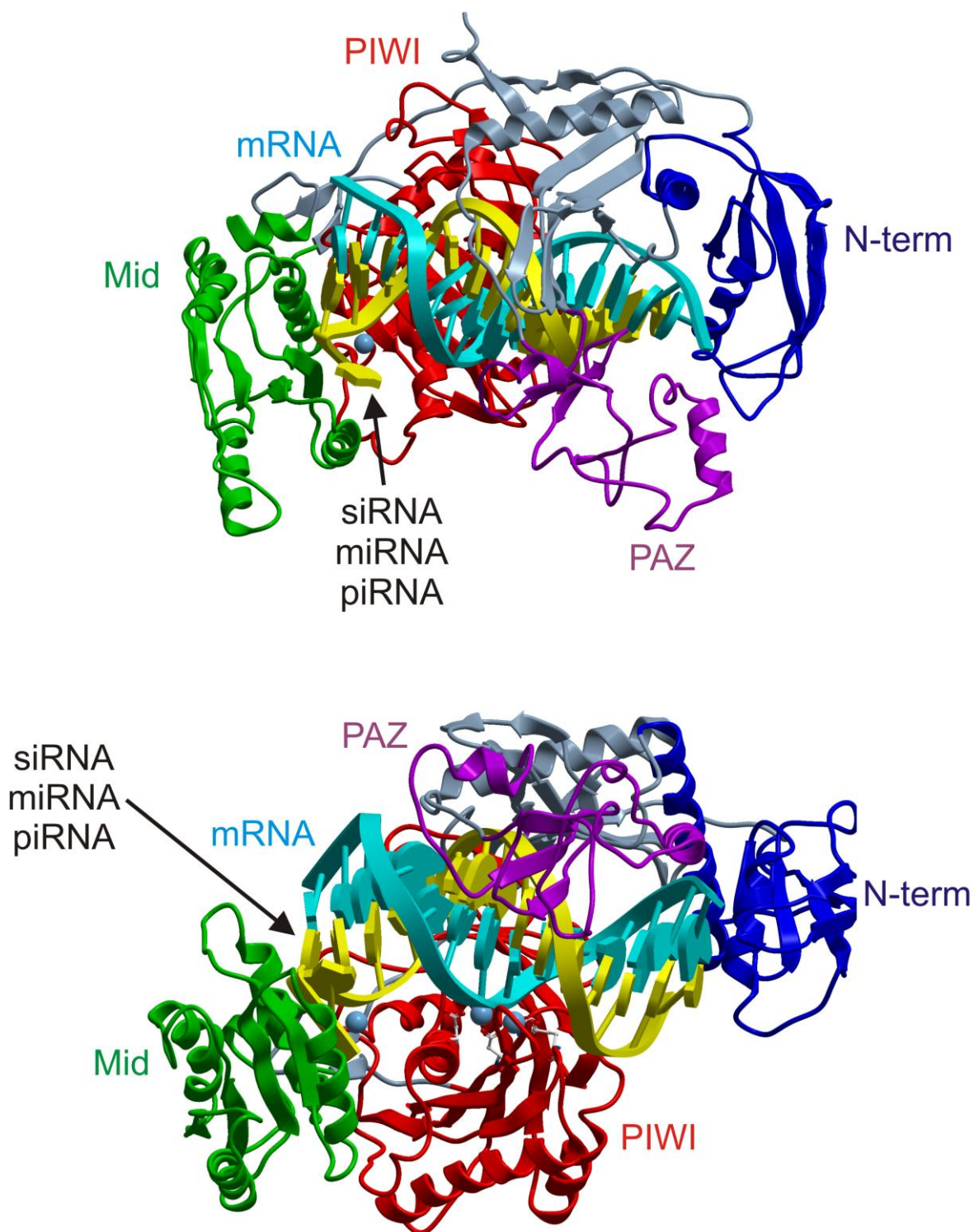
Doména PAZ se skládá ze dvou podoblastí [58-59]. První z nich je podobná proteinům s tzv. "Oligonucleotide-Binding" (OB) foldem [55]. Druhá podoblast je složena z β -hairpinu následovaného α -helixem. Štěrbina mezi těmito poddoménami interaguje s 3'-koncem krátké siRNA molekuly. Konzervovaná hydrofobní dutina v doméně PAZ tak rozpoznává duplexy siRNA či miRNA vytvořené endoribonukleázou Dicer se dvěma přesahujícími nukleotidy na obou 3'-konech duplexu [54-56]. Podoblast PAZ také může detekovat chemické modifikace 3'-konce krátkých siRNA molekul. Zvířecí piRNA či siRNA hmyzu jsou modifikované skupinou atomů 2'-O-CH₃. PAZ oblasti lidských PIWI proteinů (Hiwi1, Hiwi2, Hili) se vážou lépe ke dvěma nukleotidům přesahujícím 3'-konec dsRNA, jestliže tyto nesou 2'-O-metylovou skupinu. Naopak PAZ oblast lidského proteinu Ago1 preferuje 2'-hydroxylovou skupinu [56, 58].

Doména MID s velikostí cca. 150 aminokyselinových zbytků ukotvuje 5'-monofosfátový konec krátké RNA k proteinu Argonaute [53-54, 58-59]. *In vitro* studie naznačují, že vazba 5'-monofosfátu pomáhá zarovnat krátké siRNA na povrchu enzymu Argonaute tak, aby internukleotidová spojka cílové mRNA byla umístěna v jeho aktivním místě [55]. U lidského Ago2 stackuje 5'-báze s reziduem Tyr529, které společně s Tyr529, Lys533, Asn545 a Lys566 je také v kontaktu s 5'-fosfátem siRNA [58].

Doména MID se podílí na třídění krátkých RNA mezi různé paralogy enzymu Argonaute dle identity prvního nukleotidu. Rozpoznávání je možná díky drobným rozdílům v postranních řetězcích aminokyselin v okolí vazebné štěrby 5'-fosfátu. Krystalové struktury MID domény lidského Ago2 navázané na monofosfáty naznačují, že Ago2 upřednostňuje krátké RNA začínající uracilem nebo adenosinem. Rigidní smyčka ve vazebné oblasti totiž zamezuje navázání cytosinu nebo guanosinu [58].



Obrázek 2.9.: Komplex siRNA/miRNA/piRNA (žluté vlákno) s cílovou mRNA (bleděmodré vlákno). Dále jsou znázorněny hořčikové atomy a acidické aminokyseliny v aktivním místě enzymu Argonaute.



Obrázek 2.10.: Komplex siRNA/miRNA/piRNA (žluté vlákno) s cílovou mRNA (bleděmodré vlákno) a enzymem Argonaute (katalytická doména PIWI je obarvena červeně, N-terminální doména modře, Mid doména zeleně a PAZ doména fialově).

Vazba nukleové kyseliny k AGO

Téměř veškerou specifitu pro navázání na cílovou mRNA poskytují nukleotidy 2-7 siRNA molekuly. Ta je v komplexu RISC předpřipravena v kanonické A-konformaci [53, 58, 61, 64]. Protein Argonaute přitom váže negativně nabitou fosfordiesterovou páteř siRNA a obnažuje hrany jejích bází tak, aby se mohly spojit s bázemi cílové mRNA [62-63]. V případě lidského Ago2 jsou fosfátové skupiny nukleotidů 3-9 v kontaktu s residui Lys566, Lys709, His753, Tyr790, Arg792, Ser798, Tyr804, Ser220, Arg357, Arg714 a Arg761 [58].

AGO proteiny zpočátku komplexují s krátkým RNA:RNA duplexem, který se posléze přetvoří na jednovláknovou siRNA [53, 55, 60, 62-63]. Argonaute se přitom spojí se siRNA duplexem ve stanovené orientaci. Do vazebné štěrbině v oblasti MID se naváže termodynamicky méně stabilní 5'-konec siRNA duplexu [62]. Většina siRNA duplexů tak obsahuje snadno určitelné vlákno siRNA nebo miRNA. Někdy jsou ale dvoušroubovicovité RNA bifunkční - obě vlákna mohou ustavit komplex s proteiny AGO. Nicméně i u těchto „symetrických“ RNA:RNA duplexů zůstane v komplexu RISC nakonec pouze jedno vlákno [56]. Protein AGO svázaný s duplexem siRNA nebo miRNA se při tom nazývá pre-RISC až do té doby, než je druhé vlákno rozštěpeno mechanismem, který je stejný, jako kdyby se jednalo o cílovou mRNA [56, 60-61].

Medicínské využití AGO

Proteiny Argonaute je možné využít k selektivní inhibici exprese nežádoucí genetické informace na stejném principu, na jakém funguje RNáza H a antisense deoxyoligonukleotidy. Do buněk jsou dodávána krátká syntetická vlákna RNA imitující siRNA či miRNA molekuly, která stimulují proteiny Argonaute k degradaci cílové mRNA s komplementární sekvencí bází.

2.6. Ribonukleáza L

Funkce RNázy L

Ribonukleáza L (RNáza L) je latentní endoribonukleáza, která se podílí na antivirovém obranném mechanismu buněk [65-67]. Po detekci virové dsRNA stimuluje interferon typu I syntézu 2',5' oligoadenylátových syntetáz, které tvoří z ATP krátké β -anomerní 5'-fosforylované oligoadenyláty s 2',5' internukleotidovými spojkami (25A). 25A stimuluje přechod RNázy L z neaktivní monomerní do aktivní dimerní formy [65-68]. Aktivovaná RNáza L je schopna štěpit virovou ssRNA a zamezit tak tvorbě virových proteinů [66-67]. Aktivovaná RNáza L štěpí ssRNA v oblastech s vysokým obsahem uracilových nukleotidů. Typicky štěpí RNA po diádách UG, UA nebo UU a výsledkem tohoto štěpného procesu jsou produkty zakončené 5'-OH a 3'-monofosfátovými skupinami [67-69]. Samotné 2-5A oligoadenyláty jsou v buňkách štěpeny 2',5'-fosfodiesterázami, což moduluje antivirovou aktivitu RNázy L [65, 67].

25A

Je potřeba zdůraznit, že 2',5' internukleotidová spojka je jednoznačný stereochemický požadavek pro biologickou aktivitu RNázy L. Nahrazení 2',5'-spojky 3',5'-spojkou způsobuje narušení stackingu mezi bázemi oligonukleotidu a aromatickými aminokyselinami enzymu [66, 68].

RNáza L vyšších obratlovců je maximálně aktivována 25A typu $px5'Ap(Ap)_nA2'$, kde $x = 1-3$, $n = 1$ to 2 [66-67, 70]. 25A dimery se mohou k RNáze L vázat, ale nejsou schopné ji aktivovat [66-68, 71]. Oligonukleotidy delší než tetramery (2',5' oligoadenylátové syntetázy tvoří až 8-mery) nemají výrazně odlišné vazebné či aktivační vlastnosti. V biochemických studiích se proto obvykle používají 25A trimery a tetramery [68].

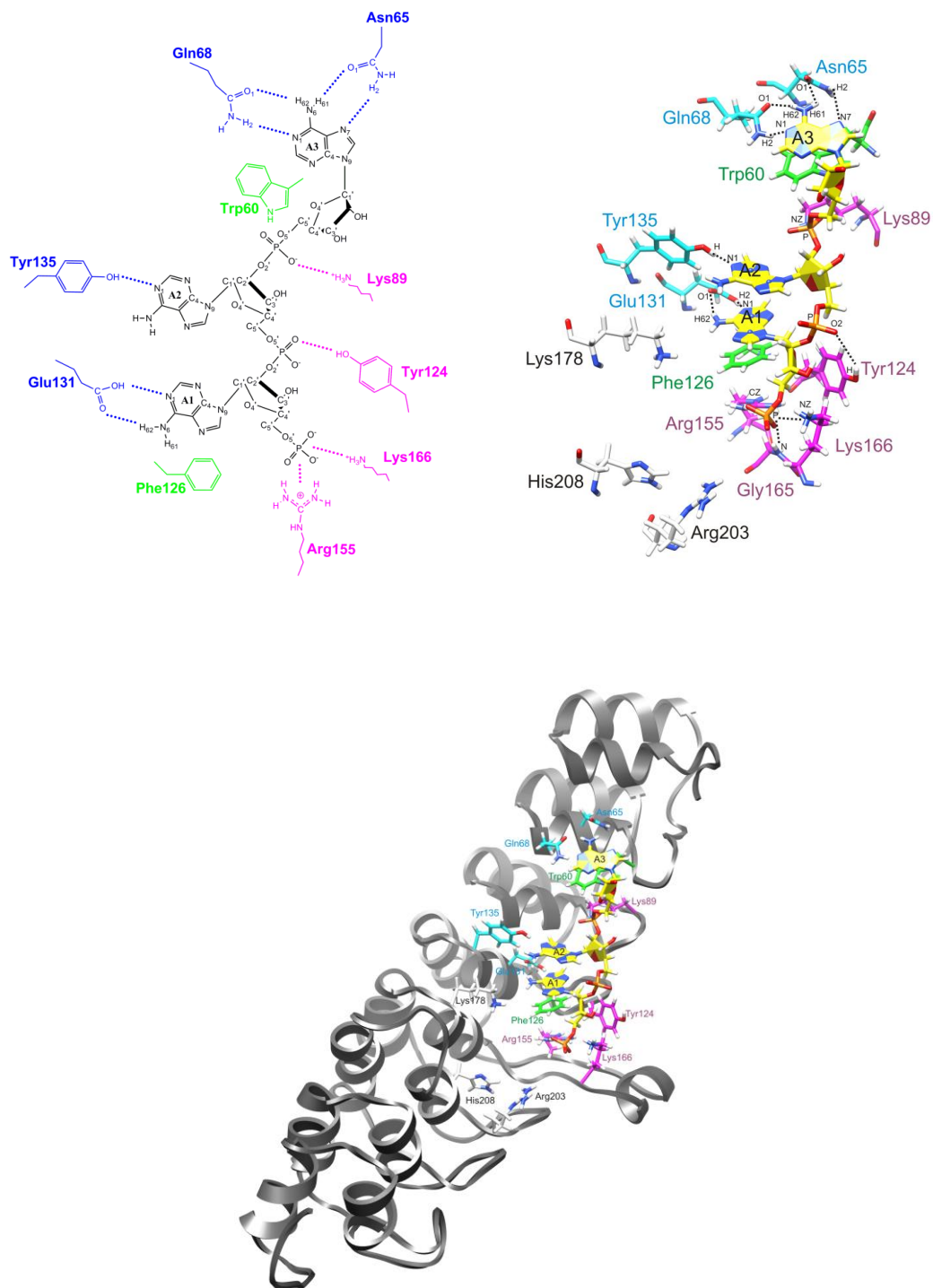
Pro efektivní aktivaci lidské RNázy L je nezbytné, aby na 5'-konci 25A byla přítomna alespoň jedna fosfátová skupina [66, 70]. Nicméně stabilita 5'-terminálního fosfátu je v buněčném prostředí velmi nízká. Obvykle je velmi rychle odštěpen fosfomonoesterázami [66].

Struktura RNázy L

Neaktivní monomer lidské RNázy L má molekulovou hmotnost 83543 Da a skládá se z 741 aminokyselinových zbytků, které tvoří tři podjednotky: ankyrinovou N-terminální, protein kináze-podobnou a katalytickou C-terminální [66-68, 71]. Krystalová struktura protein kináze-podobné a C-terminální oblasti doposud nebyla určena [66-67]. Nicméně 3D struktury ankyrinové oblasti lidské RNázy L jsou v proteinové databance dostupné [65, 68, 71-72].

Vazebné místo 25A

2-5A se váže k ankyrinové N-terminální oblasti RNázy L (Obrázek 2.11.) [65-66], která se skládá z osmi kompletních (sestavujících z 33 aminokyselinových zbytků) ankyrinových motivů, jednoho neúplného (a u monomerní RNázy L neuspořádaného) ankyrinového motivu a elementu aI [65-68].



Obrázek 2.11.: Ligand 2–5A je navázán na druhý až čtvrtý ankyrinový motiv RNázy L a interaguje s aminokyselinami prostřednictvím vodíkových vazeb (Asn65, Gln68, Glu131, Tyr135, Tyr124), solných můstků (Lys89, Arg155, Lys166) a stackingu (Trp60, Phe126).

RNáza L má dvě vazebná místa pro 25A. V prvním vazebném místě interagují oligoadenyláty s ankyrinovými motivy 2-4 a s elementem αI , jehož rezidua 159-164 tvoří dno prvního vazebného místa. Ve druhém vazebném místě interagují 25A s devátým neuspořádaným ankyrinovým motivem [65-66, 68, 71].

Ankyrinová N-terminální část lidské RNázy L sestává z aminokyselinových zbytků 24-335, nicméně vazebných interakcí s 25A se účastní jen několik málo z nich. Jmenovitě jde o Trp60, Asn65, Gln68, Phe126, Glu131, Lys89, Tyr135, Arg155, Arg310 a Tyr312. Tato rezidua interagují s oligoadenyláty prostřednictvím stackingu (Trp60, Phe126, Tyr312), vodíkových vazeb (Asn65, Gln68, Glu131, Tyr135, Arg310) nebo solných můstků (Lys89, Arg155) [65, 67-68, 71]. Vliv jednotlivých reziduí na vazebnou afinitu oligoadenylátu a aktivaci lidské RNázy L byl zkoumán prostřednictvím mutantních RNáz L, ve kterých byla rezidua interagující s 25A nahrazována alaninem [71]. Na základě výsledků těchto experimentů, mohou být interagující rezidua lidské RNázy L rozdělena do dvou skupin. V první z nich jsou ta, jež jsou pro vazbu 25A a aktivaci RNázy L kritická (Arg155, Phe126, Glu131, Trp60, Lys89) a ve druhé skupině jsou ta, která až tak esenciální nejsou (Tyr135, Gln68 a Asn65) (Obrázek 2.11.) [67-68, 71].

Terapeutické využití aktivátorů RNázy L

Pouhé dodání 25A trimerů či tetramerů do buněk vyvolá určitý terapeutický efekt, ten je však krátkodobý, neboť nukleázy přirozené 25A velmi rychle degradují [66, 70]. Proto je potřeba 25A stabilizovat prostřednictvím chemických modifikací fosfodiesterových spojek, bází či cukrů [73-74]. Nicméně struktura přirozeného 25A aktivátoru RNázy L je poměrně striktně definována a většina chemických modifikací způsobí redukci vazebných a aktivačních schopností [70-71, 73].

Např. fosfonátová analoga oligonukleotidů, která obsahují tzv. "bridging P-C" vazbu, jsou obecně velmi slibná z hlediska možného terapeutického využití. Vysoká metabolická a chemická stabilita, konformační adaptibilita a isopolarita, předurčují tyto látky k přípravě syntetických oligoribonukleotidů s modifikovanou internukleotidovou spojkou [70, 73]. Prodloužení $O2'-P-O5'$ spojek 25A modifikací typu $O2'-C-P-O5'$ nebo $O2'-P-C-O5'$ však vyvolává fundamentální změny v geometrii komplexu s RNázou L [70, 73]. Obzvláště, internukleotidová spojka mezi druhým a třetím residuem 25A je citlivá na tento druh modifikací [74].

Pro vzrůst enzymatické stability je přínosné nahradit nebo modifikovat 3'-hydroxylovou skupinu, protože 2',3'-cyklické-nukleotidové 2'-fosfodiesterázy štěpí oligonukleotidy přes intermediáty s cyklickým 2',3'-fosfátem. Na základě znalosti tohoto faktu byly provedeny komparativní studie přirozených a xylo-oligoadenylátů, které prokázaly, že pouhá inverze 3'-hydroxyly 25A poskytuje oligoadenyláty, které jsou enzymaticky pozoruhodně stabilní [74].

V současné době je již zřejmé, že modifikované aktivátory RNázy L reprezentují novou generaci potenciálních virostatik.

3. Metodika

Jednou z nejzajímavějších a pravděpodobně nejrychleji se rozvíjejících oblastí lidského vědění je makromolekulární chemie a biochemie. Zde je kvantová mechanika základním způsobem, jak popsat formování molekul a jejich reakce na změny vnějšího prostředí. Protože je detailní výpočet založený na kvantové mechanice velmi náročný na výpočetní čas, zůstává problémem jeho technická stránka. Z tohoto důvodu je nezbytné přikročit k (někdy až velmi hrubým) aproximacím redukujícím výpočetní náročnost a stále ještě poskytujícím určitý vhled do skutečného chování zkoumaných systému. Cílem této kapitoly je poskytnout stručný přehled těchto aproximací a poukázat na jejich vhodnost či nevhodnost pro daný problém, kterým je dynamický vývoj biomolekul složených ze stovek či tisíců atomů a obklopených vodní obálkou sestávající z desítek až stovek tisíc atomů [75-83].

3.1. Stacionární systémy

3.1.1. Popis stacionárních systémů

Kvantová mechanika

Začneme u Schrödingerovy rovnice samotné, její obecná forma je:

$$i\hbar \frac{d}{dt} \Phi = H\Phi \quad (3.1.1.1)$$

kde Φ je komplexní vlnová funkce popisující systém a H je Hamiltonova funkce. V mnoha případech je nutné studovat pouze stacionární jevy (tj. jevy, které jsou buď na čase zcela nezávislé, nebo se v čase opakují). V takovém případě je za předpokladu, že $\Phi(t) = \Phi e^{\frac{-iEt}{\hbar}}$ možné přejít k časově nezávislé Schrödingerově rovnici:

$$H\Phi = E\Phi \quad (3.1.1.2)$$

Toto je první zjednodušení, které značně usnadňuje výpočet vlnové funkce Φ pro daný systém.

Molekula vodíku

Ve většině případů, které jsou zajímavé jak pro aplikace, tak i pro základní výzkum, však není možné systém zredukovat pouze na jednu částici. Často je nutné pracovat s velmi velkým počtem částic, které se vzájemně ovlivňují. Snad nejjednodušším případem je molekula vodíku. Je to případ natolik ilustrativní a blízký k výpočtům složitějších molekul, že stojí za to se o něm krátce zmínit. Jedná se o systém se dvěma pozitivně nabitými jádry a dvěma elektrony, které obíhají okolo těchto jader. I takto jednoduchý systém však není možné řešit analyticky a tak nezbyvá, než zvolit jiný postup. První, kdo navrhl úspěšné řešení tohoto problému, byl britský matematik Douglas Rayner Hartree v roce 1928 [84]. Protože jsou jádra přibližně 2000x hmotnější než elektrony lze předpokládat, že se elektrony pohybují podstatně rychleji než jádra. To znamená, že je možné řešit pohyb elektronů v elektrostatickém poli

stacionárních jader. Z pohledu jader se naopak rychle se pohybující elektrony jeví jako nábojový oblak – efektivní pole – na jehož změny jádra jen pomalu reagují. Používá se proto tzv. Born-Oppenheimerova aproximace, kdy se řeší dvě Schrödingerovy rovnice. První pro pohyb elektronů v poli pevných jader a druhá pro pohyb jader v efektivním poli elektronů.

Po separaci pohybu elektronů a jader se úloha nalézt vlnovou funkci elektronů v molekule vodíku redukuje na řešení Schrödingerovy rovnice:

$$\left(-\frac{1}{2} \sum_i \Delta_i + \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}} - \sum_{i,A} \frac{Z_A}{r_{iA}} \right) \Phi = E\Phi, \quad i, j \leq 2 \quad (3.1.1.3)$$

kde první člen rovnice odpovídá součtu kinetických energií elektronů, druhý člen odpovídá potenciálu interakce elektron-elektron a třetí člen odpovídá coulombovské interakci elektronů a jader. Připomeňme, že u molekuly vodíku jedná o systém pouhých dvou elektronů a dvou jader. Pro úspěšné řešení této rovnice je nezbytné nahradit vliv druhého elektronu na pohyb zvoleného elektronu působením jistého efektivního pole. Elektron se pak pohybuje v poli vytvářeném jádry a efektivním polem ostatních elektronů. Pro každý z elektronů lze pak zavést jednoelektronovou vlnovou funkci, která je řešením jednoelektronové Schrödingerovy rovnice. Vlnovou funkci obou elektronů pak můžeme rozepsat jako součin vlnových funkcí jednotlivých elektronů:

$$\Phi(\mathbf{r}_1, \mathbf{r}_2) = \Phi_1(\mathbf{r}_1) \Phi_2(\mathbf{r}_2) \quad (3.1.1.4)$$

kde $\Phi(\mathbf{r}_1, \mathbf{r}_2)$ je vlnová funkce zahrnující oba elektrony a $\Phi_1(\mathbf{r}_1)$, $\Phi_2(\mathbf{r}_2)$ jsou vlnové funkce jednotlivých elektronů. Tento postup je samozřejmě možné rozšířit na libovolný počet částic.

K odvození rovnic, kterým vyhovují jednoelektronové funkce, se používá variační princip, který se aplikuje na výraz pro celkovou energii:

$$E = \langle \Phi | H | \Phi \rangle$$

$$E = \int \Phi^* \sum_i H^{core}(\mathbf{r}_i) \Phi d\mathbf{r} + \frac{1}{2} \int \Phi^* \sum_{i \neq j} \frac{1}{r_{ij}} \Phi d\mathbf{r} \quad (3.1.1.5)$$

přičemž

$$H^{core}(\mathbf{r}_i) = -\frac{1}{2} \Delta_i - \sum_A \frac{Z_A}{r_{iA}} \quad (3.1.1.6)$$

je jednočásticový operátor zahrnující operátor kinetické energie a coulombovskou interakci s jádry. Takto získáme Hartreeho rovnice:

$$\left(-\frac{1}{2} \Delta_1 - \sum_A \frac{Z_A}{r_{1A}} + \sum_{j(\neq i)} \int \frac{|\Phi_j(\mathbf{r}_2)|^2}{r_{12}} d\mathbf{r}_2 \right) \Phi_i(\mathbf{r}_1) = \epsilon_i \Phi_i(\mathbf{r}_1) \quad (3.1.1.7)$$

$$i = 1, 2, 3, \dots, N$$

kde index i zahrnuje všechny (v našem případě dva) elektrony systému a Lagrangeovy multiplikátory mají tvar:

$$\epsilon_i = \int \phi_i^*(\mathbf{r}_1) \left[-\frac{1}{2}\Delta_1 - \sum_A \frac{Z_A}{r_{1A}} + \sum_{j(\neq i)} \int \frac{|\phi_j(\mathbf{r}_2)|^2}{r_{12}} d\mathbf{r}_2 \right] \phi_i(\mathbf{r}_1) d\mathbf{r}_1 \quad (3.1.1.8)$$

Hartreeho rovnice představují soustavu nelineárních a nelokálních integrodiferenciálních rovnic, kterou lze řešit pouze numericky – iterační metodou. Zvolíme vhodné vlnové jednoelektronové funkce v nulté aproximaci, dosadíme je do integrálu v rovnici (3.1.1.7) a najdeme jednoelektronové funkce v první aproximaci. Tento postup opakujeme tak dlouho, až dosáhneme toho, že se vlnové funkce v následujících iteracích liší pouze zanedbatelně.

K uvedenému příkladu molekuly vodíku je nutné podotknout, že tento přístup nebere v úvahu fermionovou povahu elektronů a z ní vyplývající nezbytnost antisymetrie vlnových funkcí elektronů. Toto upřesnění je zahrnuto v Hartree-Fockově přístupu, kde je do vlnových funkcí elektronů zahrnut spin elektronů. Tento případ však zde nebudeme rozebírat. Uvedený příklad zanedbává také relativistické efekty, spin-orbitální interakce a spin jader.

3.1.2. Silové pole AMBER

Pro výpočet energie velkých biomolekulárních systémů není možné kvantově-chemický přístup aplikovaný výše na molekulu vodíku použít. Biomolekuly se proto reprezentují jako soustava VdW kuliček nesoucích parciální náboje a namísto kovalentních vazeb propojených pružinkami. Energie biomolekuly je pak spočtena prostřednictvím analytického výrazu, v němž konkrétní hodnoty parametrů - tzv. silových konstant - bývají určeny na základě experimentálních dat (empirické potenciály), nebo teoretických kvantově mechanických výpočtů založených na principech naznačených výše na příkladu molekuly vodíku a aplikovaných na malé fragmenty biomolekul.

Tzv. silové pole AMBER bylo původně parametrizováno na základě experimentálních dat získaných pro fragmenty nukleových kyselin a proteinů. V novějších verzích silového pole AMBER bylo pro jeho parametrizaci použito i *ab initio* výpočtů [85-87]. Funkční vyjádření použité pro parametrizaci silového pole AMBER vypadá takto:

$$E_{pot} = \sum_b K_b (b - b_0)^2 + \sum_\theta H_\theta (\theta - \theta_0)^2 + \sum_\phi \frac{V_n}{2} [1 + \cos(n\phi - \phi_0)] + \sum \varepsilon \left[\left(\frac{r^*}{r} \right)^{12} - \left(\frac{r^*}{r} \right)^6 \right] + \sum \frac{q_i q_j}{\varepsilon_{ij} r_{ij}} + \sum \left[\left(\frac{C_{ij}}{r_{ij}} \right)^{12} - \left(\frac{D_{ij}}{r_{ij}} \right)^{10} \right] \quad (3.1.2.1)$$

Tato rovnice obsahuje pouze minimální počet členů. V prvních dvou členech jsou zahrnuty změny délek jednotlivých vazeb a úhlů mezi vazbami. 4. a 5. člen popisují

Van der Waalsovou a elektrostatickou interakci. Poslední člen odpovídá zahrnutí vodíkových vazeb a v novějších verzích silového pole AMBER se již nepoužívá.

Uvedená analytická formulka nezahrnuje členy vyšších řádů (popisující anharmonicitu) nezbytné např. pro přesný výpočet vibračních spekter molekul. Nicméně je vhodná pro modelování struktur a interakcí biomolekul [88-92].

Doplnění chybějících parametrů silového pole AMBER

V této práci byly prostřednictvím molekulárně dynamických simulací zkoumány komplexy modifikovaných nukleových kyselin s RNázou H, RNázou L, případně NV RdRp. Silové pole AMBER, které bylo zvoleno jako nejvhodnější pro molekulárně dynamické simulace zmíněných komplexů, však neobsahuje silové konstanty nezbytné pro popis modifikovaných částí nukleových kyselin.

Silové konstanty pro vazby, úhly a van der Waalsovou interakci bylo většinou možné určit na základě analogie s parametry v silovém poli AMBER již obsaženými. Nicméně parciální náboje na atomech a silové konstanty pro některé torzní úhly bylo nezbytné určit na základě kvantově mechanických výpočtů.

Z tohoto důvodu byla programem Gaussian provedena řada *ab initio* výpočtů pro malé modelové systémy, na základě nichž byla silová pole doplněna o nové parametry – silové konstanty. Výpočty byly přitom provedeny konzistentně s originálním postupem aplikovaným při parametrizaci původního silového pole AMBER [87, 93].

Parciální bodové náboje na atomech užívané při výpočtu elektrostatické interakce

$$\sum \frac{q_i q_j}{\epsilon_{ij} r_{ij}} \quad (3.1.2.2)$$

byly vypočteny s použitím metody RESP (Restrained ElectroStatic Potential) [93], která je implementována v softwarovém balíku AMBER. Hodnoty nábojů umístěných v centrech atomů se fitují tak, aby výsledný molekulárně mechanický (MM) elektrostatický potenciál (ESP) v okolí molekuly odpovídal elektrostatickému potenciálu spočtenému kvantově mechanicky (QM). Vstupem pro metodu RESP jsou hodnoty V_i kvantově mechanického ESP spočtené v diskrétních bodech vně VdW povrchu molekuly. Pro jejich spočtení je nezbytné použít programy jako Gaussian, nebo GAMESS. V souladu s originálním přístupem byl ESP spočten pomocí HF metody a báze 6-31G* [87]. V metodě RESP se posléze používá metoda nejmenších čtverců k fitu náboje q_j j -tého atomového centra v molekule. Vzhledem k tomu, že molekulárně-mechanický ESP \hat{V}_i je dán rovnicí:

$$\hat{V}_i = \sum_j \frac{q_j}{r_{ij}} \quad (3.1.2.3)$$

tak funkce, která má být minimalizována v metodě nejmenších čtverců, má tvar:

$$\chi_{esp}^2 = \sum_i (V_i - \hat{V}_i)^2 \quad (3.1.2.4)$$

V minimu pro všechna j pak:

$$\frac{\partial(\chi_{esp}^2)}{\partial q_j} = -2 \sum_i \frac{V_i - \hat{V}_i}{r_{ij}} = 0 \quad (3.1.2.5)$$

Z toho plyne, že rovnice mohou být uspořádány a vyřešeny v maticové formě a že se součet všech nábojů musí rovnat celkovému molekulárnímu náboji. Přidáme-li do procesu fitování nábojů penalizační funkci ke členu χ_{esp}^2 , pak funkce, která má být minimalizována, získá tvar:

$$\chi^2 = \chi_{esp}^2 + \chi_{rstr}^2 \quad (3.1.2.6)$$

minimum nejmenších čtverců je nyní pro všechna j definováno takto:

$$\frac{\partial(\chi^2)}{\partial q_j} = \frac{\partial(\chi_{esp}^2)}{\partial q_j} + \frac{\partial(\chi_{rstr}^2)}{\partial q_j} = 0 \quad (3.1.2.7)$$

Obvykle se volí harmonická penalizační funkce:

$$\chi_{rstr}^2 = a \sum_j (q_j^\circ - q_j)^2 \quad (3.1.2.8)$$

kde a je škálovací faktor, který definuje sílu restrikce a q° určuje cílový náboj. Takto druhý člen z rovnice (3.1.2.7) nabývá tvar (3.1.2.9):

$$\partial(\chi_{rstr}^2)/\partial q_j = -2a(q_j^\circ - q_j) \quad (3.1.2.9)$$

Tímto jsme řešení problému převedli na řešení maticové rovnice pro náboje \mathbf{q} :

$$\mathbf{A}\mathbf{q} = \mathbf{B} \quad (3.1.2.10)$$

kde nediagonální prvky matice \mathbf{A} mají tvar:

$$A_{jk} = \sum_i \frac{1}{r_{ij}r_{ik}} \quad (3.1.2.11)$$

Diagonální prvky matice \mathbf{A} jsou ve tvaru:

$$A_{jj} = \sum_i \frac{1}{r_{ij}^2} + \frac{\partial(\chi_{rstr}^2)}{\partial q_j} \quad (3.1.2.12)$$

a řádky pravé strany \mathbf{B} jsou ve tvaru:

$$B_j = \sum_i \frac{V_i}{r_{ij}^2} + q_j^\circ \frac{\partial(\chi_{rstr}^2)}{\partial q_j} \quad (3.1.2.13)$$

Maticová rovnice pro náboje \mathbf{q} (3.1.2.10) se řeší běžnými postupy [95].

Dále bylo potřeba nafitovat silové konstanty pro torzní úhly. S využitím malých modelových systémů byly proto nejprve programem Gaussian pomocí metody MP2/6-31G* získány kvantově mechanické profily potenciální energie pro příslušné torzní úhly [87, 95].

Podobným postupem byly pro stejné modelové systémy vypočítány molekulárně-mechanické profily potenciální energie pomocí modulu *Sander* ze softwarového balíku AMBER. Neznámé silové konstanty pro torzní úhly byly přitom vynulovány.

Odečtením QM a MM energetických profilů byl získán energetický „příspěvek“ torzních úhlů k MM energii silového pole AMBER, který byl použit pro nafitování chybějících parametrů V_n a ϕ_0 [96]:

$$\sum_{\phi} \frac{V_n}{2} [1 + \cos(n\phi - \phi_0)] \quad (3.1.2.14)$$

3.2. Dynamické systémy

3.2.1. Aproximace dynamických systémů

Kvantově popsaná dynamika komplikovaného systému – TDSCF

Pokud zkoumáme dynamický vývoj systému sestávajícího z jader a elektronů je třeba začít u časové Schrödingerovy rovnice (3.1.1.1). Vhodnou aproximací vlnové funkce Φ , která separuje funkce jader a elektronů je její vyjádření ve tvaru:

$$\Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}; t) \approx \Psi(\{\mathbf{r}_i\}; t) \chi(\{\mathbf{R}_I\}; t) \exp \left[\frac{i}{\hbar} \int_{t_0}^t dt' \tilde{E}_e(t') \right] \quad (3.2.1.1)$$

Po dosazení rovnice (3.2.1.1) do časově závislé Schrödingerovy rovnice (za předpokladu zachování energie) získáme soustavu dvou provázaných rovnic:

$$i\hbar \frac{\partial \Psi}{\partial t} = - \sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 \Psi + \quad (3.2.1.2)$$

$$\begin{aligned}
& + \left\{ \int d\mathbf{R} \chi^* (\{\mathbf{R}_I\}; t) V_{n-e} (\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) \chi (\{\mathbf{R}_i\}; t) \right\} \Psi \\
i\hbar \frac{\partial \chi}{\partial t} & = - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 \chi + \left\{ \int d\mathbf{r} \Psi^* (\{\mathbf{r}_i\}; t) \mathcal{H}_e (\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) \Psi (\{\mathbf{r}_i\}; t) \right\} \chi
\end{aligned} \tag{3.2.1.3}$$

kteřá slouží jako základ pro metodu tzv. časově závislého selfkonzistentního pole (TDSCF - time-dependent self-consistent field) zavedenou v roce 1930 Diracem. Elektrony i jádra se pohybují kvantově-mechanicky v časově závislém efektivním potenciálu (neboli v selfkonzistentně získaném průměrném poli) [97].

Hybridní Ehrenfestova molekulární dynamika

Kvantový výpočet rovnic (3.2.1.2) a (3.2.1.3) je náročný na výpočetní čas a tudíž je použitelný pouze pro malý počet částic. Pro situace, kdy je třeba řešit systém s velkým počtem elektronů v poli mnoha jader, nezbyvá než přikročit k daleko zásadnějším zjednodušením [97]. Takovým zjednodušením je přechod od kvantového popisu k popisu semiklasickému, nebo klasickému, kde jsou jádra atomů pojímána jako klasické částice. Vlnovou funkci χ pro jádra je pak možné přepsat do tvaru:

$$\chi (\{\mathbf{R}_I\}; t) = A (\{\mathbf{R}_I\}; t) \exp \left\{ \frac{iS (\{\mathbf{R}_I\}; t)}{\hbar} \right\} \tag{3.2.1.4}$$

kde je S fáze této vlnové funkce a A je její amplituda. Dosazením do TDSCF rovnice pro jádra (3.2.1.3) a separováním reálné a imaginární části získáme:

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I} (\nabla_I S)^2 + \int d\mathbf{r} \Psi^* \mathcal{H}_e \Psi = \hbar^2 \sum_I \frac{1}{2M_I} \frac{\nabla_I^2 A}{A} \tag{3.2.1.5}$$

$$\frac{\partial A}{\partial t} + \sum_I \frac{1}{M_I} (\nabla_I A) (\nabla_I S) + \sum_I \frac{1}{2M_I} A (\nabla_I^2 S) = 0 \tag{3.2.1.6}$$

Tyto dvě rovnice se obvykle nazývají “dynamická reprezentace kvantové kapaliny” a mohou být též použity pro řešení časově závislé Schrödingerovy rovnice. Rovnice (3.2.1.6) může být chápána jako rovnice kontinuity, která lokálně zajišťuje zachování pravděpodobnosti nalezení částice. Důležitější je v danou chvíli rovnice (3.2.1.5). Pro $\hbar \rightarrow 0$ přejde na tvar:

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I} (\nabla_I S)^2 + \int d\mathbf{r} \Psi^* \mathcal{H}_e \Psi = 0 \tag{3.2.1.7}$$

což je rovnice ve stejném tvaru jako klasická pohybová rovnice Hamilton-Jakobiho:

$$\frac{\partial S}{\partial t} + \mathcal{H} (\{\mathbf{R}_I\}, \{\nabla_I S\}) = 0 \tag{3.2.1.8}$$

s klasickým Hamiltoniánem:

$$\mathcal{H}(\{\mathbf{R}_I\}, \{\mathbf{P}_I\}) = T(\{\mathbf{P}_I\}) + V(\{\mathbf{R}_I\}) \quad (3.2.1.9)$$

kde $T(\mathbf{P})$ je kinetická složka energie a $V(\mathbf{R})$ je potenciální složka energie. Připomeneme-li si navíc transformační vztah pro moment hybnosti:

$$\mathbf{P}_I \equiv \nabla_I S \quad (3.2.1.10)$$

získáme klasické Newtonovy pohybové rovnice odpovídající rovnici (3.2.1.7) ve tvaru:

$$\begin{aligned} M_I \ddot{\mathbf{R}}_I(t) &= -\nabla_I \int d\mathbf{r} \Psi^* \mathcal{H}_e \Psi \\ &= -\nabla_I V_e^E(\{\mathbf{R}_I(t)\}) \end{aligned} \quad (3.2.1.11)$$

Tento přechod nám umožňuje vypočítat klasickým způsobem pohyb jader (to je změny poloh atomů) v kvantově vypočteném efektivním poli elektronů.

Nicméně vlnová funkce jader se stále ještě vyskytuje v elektronové TDSCF rovnici. Zde dosáhneme přechodu ke klasickému popisu jednoduše:

$$\int d\mathbf{R} \chi^*(\{\mathbf{R}_I\}; t) \mathbf{R}_I \chi(\{\mathbf{R}_I\}; t) \xrightarrow{\hbar \rightarrow 0} \mathbf{R}_I(t) \quad (3.2.1.12)$$

Tato klasická limita vede k časově závislé vlnové rovnici pro elektrony:

$$\begin{aligned} i\hbar \frac{\partial \Psi}{\partial t} &= -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 \Psi + V_{n-e}(\{\mathbf{r}_i\}, \{\mathbf{R}_I(t)\}) \Psi \\ &= \mathcal{H}_e(\{\mathbf{r}_i\}, \{\mathbf{R}_I(t)\}) \Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}; t) \end{aligned} \quad (3.2.1.13)$$

kteřá se vyvíjí self-konzistentně s pohybem klasicky popsaných jader (3.2.1.11). Vlnová rovnice elektronů závisí parametricky na poloze jader. Přístup, který spočívá v řešení rovnic (3.2.1.11) a (3.2.1.13) se nazývá Ehrenfestova molekulární dynamika.

Klasická molekulární dynamika

Další důležité zjednodušení je možné, pokud se u elektronové vlnové funkce omezíme na základní stav systému [97]. To je dobrá aproximace v případě, kdy je termální energie $\mathbf{k}_B \mathbf{T}$ menší, než vzdálenost od základního stavu \mathbf{E}_0 (s nejnižší energií) k nejbližší excitované energetické hladině \mathbf{E}_1 . V takovém případě se jádra pohybují v souladu s rovnicí (3.2.1.11) po jednom povrchu potenciálové energie:

$$V_e^E = \int d\mathbf{r} \Psi_0^* \mathcal{H}_e \Psi_0 \equiv E_0(\{\mathbf{R}_I\}) \quad (3.2.1.14)$$

který může být vypočten z časově nezávislé Schrödingerovy rovnice pro elektrony v základním stavu:

$$\mathcal{H}_e \Psi_0 = E_0 \Psi_0 \quad (3.2.1.15)$$

Nyní můžeme oddělit generování dynamického vývoje poloh jader od výpočtu elektronového povrchu potenciálové energie:

1. V prvním kroku je E_0 vypočteno z rovnice (3.2.1.15) pro mnoho různých poloh jader.
2. Ve druhém kroku jsou tato data naitována na analytické funkční vyjádření, čímž je získán globální povrch potenciálové energie, jehož gradienty je možné též analyticky vypočítat.
3. Ve třetím kroku jsou řešeny Newtonovy pohybové rovnice na tomto povrchu potenciálové energie.

Pro systém s mnoha jádry může výpočet globálního potenciálového povrchu elektronů stále představovat komplikovanou úlohu. Proto se obvykle aproximuje zkráceným rozvojem mnohačasticových příspěvků:

$$\begin{aligned} V_e^E \approx V_e^{approx}(\{\mathbf{R}_I\}) = & \sum_{I=1}^N v_1(\mathbf{R}_I) + \sum_{I<J}^N v_2(\mathbf{R}_I, \mathbf{R}_J) + \\ & + \sum_{I<J<K}^N v_3(\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_K) + \dots \end{aligned} \quad (3.2.1.16)$$

V tomto rozvoji každý koeficient \mathbf{v}_i pak odpovídá určitému mnohačasticovému interakčnímu potenciálu (kupříkladu \mathbf{v}_2 je koeficient interakčního potenciálu mezi dvěma částicemi). Snížení náročnosti výpočtu pak dosáhneme zkrácením rozvoje (3.2.1.16) pouze na několik prvních členů. Smíšený kvantově/klasický problém je tak redukován na čistě klasický, pokud jsou určeny koeficienty \mathbf{v}_n . V klasické molekulární dynamice:

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I V_e^{approx}(\{\mathbf{R}_I(t)\}) \quad (3.2.1.17)$$

se do úvahy obvykle berou pouze dvou- a tří-časticové interakce [98-99].

3.2.2. Numerické metody vhodné pro řešení pohybových rovnic

V této kapitole budeme krátce diskutovat numerické metody řešení molekulárně-dynamických problémů. Očividně nejjednodušší přístup je diskretizace pohybových rovnic a jejich následné numerické řešení. Tento přístup se často nazývá metoda konečných diferencí (FDE = finite difference equation). Verletova metoda založená na principu konečných diferencí je symplektická (symplectic) metoda, což znamená, že je reverzibilní v čase (otočením směru toku času se systém bude vyvíjet zpět po stejné trajektorii ve fázovém prostoru, až na zaokrouhlovací a jiné numerické nepřesnosti). Příkladem ne-symplektických algoritmů jsou Gearovy algoritmy, což jsou algoritmy typu prediktor-korektor (PC). Hlavní přínosem algoritmů založených na Verletově metodě při použití v molekulárně-dynamických simulacích je jejich

enormní stabilita a to, že zrychlení částic (síly) je nutné vypočítat v každém časovém kroku nejvýše jednou. Nyní probereme základní rysy jednotlivých metod [100-102].

Verletova metoda

Začneme rozepsáním pohybových rovnic do diskrétního tvaru použitím Taylorova rozvoje

$$\begin{aligned} r(t + \delta t) &= r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \dots, \\ v(t + \delta t) &= v(t) + a(t)\delta t + \frac{1}{2}b(t)\delta t^2 + \dots, \\ a(t + \delta t) &= a(t) + b(t)\delta t + \dots \end{aligned} \tag{3.2.2.1}$$

kde jsou $\mathbf{r}(t + \delta t)$ pozice, $\mathbf{v}(t + \delta t)$ rychlost a $\mathbf{a}(t + \delta t)$ zrychlení v čase $t + \delta t$. Jak jsme se zmínili výše, tyto rovnice jsou v čase t reverzibilní. Pro systém pohybující se vpřed a zpět v čase je možné napsat:

$$\begin{aligned} r(t + \delta t) &= r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \dots \\ r(t - \delta t) &= r(t) - v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \dots \end{aligned} \tag{3.2.2.2}$$

sečtením těchto rovnic, získáme rovnici:

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + a(t)\delta t^2 \tag{3.2.2.3}$$

kteřá nám dává návod na výpočet pozice částic v čase $\mathbf{t+h}$. Je dobré si povšimnout, že v této metodě nejsou explicitně zahrnuty rychlosti částic. Výhodou Verletovy metody je přímočarost a nenáročnost na počítačovou paměť.

Rychlostní Verletova metoda (velocity Verlet)

Verletovu metodu můžeme přeformulovat na rovnice, které zahrnují i rychlosti \mathbf{v} . Tyto rovnice jsou odvozeny postupem analogickým k odvození (3.2.2.2):

$$\begin{aligned} r(t + \delta t) &= r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 \\ v(t + \delta t) &= v(t) + \frac{1}{2}[a(t) + a(t + \delta t)]\delta t \end{aligned} \tag{3.2.2.4}$$

Tato metoda modifikuje základní Verletovu metodu tím, že jsou explicitně vypočteny rychlosti částic. Výpočet rychlostí je nezbytný například u NVT souborů, ve kterých se konstantní teplota často udržuje přeskálováním rychlostí částic.

Leap frog algoritmus

Leap-frog algoritmus je velmi podobný výše popsanému rychlostnímu Verletovu algoritmu. Hlavní rozdíl spočívá v tom, že rychlosti a pozice částic jsou vypočteny v různých časech. Nejprve jsou rychlosti \mathbf{v} vypočteny v čase $\mathbf{t + 1/2h}$, hodnoty těchto rychlostí jsou pak použity k výpočtu pozic částic v čase $\mathbf{t + h}$.

$$\begin{aligned}
v\left(t + \frac{1}{2}\delta t\right) &= v\left(t - \frac{1}{2}\delta t\right) + a(t)\delta t \\
r(t + \delta t) &= r(t) + v\left(t + \frac{1}{2}\delta t\right)\delta t
\end{aligned}
\tag{3.2.2.5}$$

Nevýhodou je, že pozice a rychlosti nejsou vypočteny pro stejný čas t . Chceme-li vypočítat rychlosti částic v čase t , je možné použít následující vztah:

$$v(t) = \frac{1}{2} \left[v\left(t - \frac{1}{2}\delta t\right) + v\left(t + \frac{1}{2}\delta t\right) \right]
\tag{3.2.2.6}$$

Algoritmus Leap frog je implementován v softwarovém balíku AMBER.

Beemanův algoritmus

Beemanův algoritmus je velmi podobný Verletovu algoritmu:

$$\begin{aligned}
r(t + \delta t) &= r(t) + v(t)\delta t + \frac{2}{3}a(t)\delta t^2 - \frac{1}{6}a(t - \delta t)\delta t^2 \\
v(t + \delta t) &= v(t) + v(t)\delta t + \frac{1}{3}a(t)\delta t + \frac{5}{6}a(t)\delta t - \frac{1}{6}a(t - \delta t)
\end{aligned}
\tag{3.2.2.7}$$

Výhodou algoritmu je větší přesnost ve výpočtu rychlostí a lepší zachování celkové energie systému. Nevýhodou tohoto algoritmu je větší náročnost na výpočetní čas. I tento algoritmus je implementován v softwarovém balíku AMBER.

Brünger Brooks Karplusova (BBK) metoda – příklad rozšíření Verletovy metody

BBK metoda je příkladem rozšíření Verletovy metody na Langevinovu rovnici, která slouží k popisu Langevinovy dynamiky. Tato rovnice rozšiřuje Newtonovy pohybové zákony o přibližné výpočty vlivu prostředí na molekulární systém. Langevinova rovnice má následující obecný tvar:

$$ma(t) = F(t) - \gamma mv(t) + R(t)
\tag{3.2.2.8}$$

kde γ je třecí koeficient, m je hmotnost částice, $F(t)$ je síla působící na částici, $a(t)$ zrychlení částice a $R(t)$ náhodná síla. $R(t)$ představuje stochastickou sílu odpovědnou za náhodné srážky molekulárního systému s imaginárními částicemi prostředí. BBK integrátor využívá aproximaci (3.2.2.9). Vzhledem k přidaným silám v Langevinově dynamice už není zrychlení časovou derivací rychlosti a proto zde neplatí rovnost.

$$\dot{v}(t) \approx a(t) = \frac{1}{m} [F(t) - \gamma mv(t) + R(t)]
\tag{3.2.2.9}$$

Rozvineme-li pozici částice do Taylorovy řady a použijeme-li postup stejný jako v Verletově metodě, dostaneme pro pozici částice:

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \frac{\delta t^2}{m} [F(t) - \gamma m v(t) + R(t)] \quad (3.2.2.10)$$

ze které dále plyne:

$$r(t + \delta t) = \frac{1}{1 + \gamma \frac{\delta t}{2}} \left\{ 2r(t) - r(t - \delta t) + \gamma \frac{\delta t}{2} r(t - \delta t) + \frac{\delta t^2}{m} [F(t) + R(t)] \right\} \quad (3.2.2.11)$$

Verletova metoda nezahrnuje rovnice pro výpočet rychlostí \mathbf{v} . Vyjdeme-li z Taylorova rozvoje rychlosti:

$$v(t + \frac{\delta t}{2}) = v(t) + \dot{v}(t) \frac{\delta t}{2} = v(t) + \frac{\delta t}{2m} [F(t) - \gamma m v(t) + R(t)] \quad (3.2.2.12)$$

dostaneme pro vzorec rychlosti (3.2.2.13):

$$v(t + \delta t) = \frac{1}{1 + \gamma \frac{\delta t}{2}} \left\{ v(t + \frac{\delta t}{2}) + \frac{\delta t}{2m} [F(t + \delta t) + R(t + \delta t)] \right\} \quad (3.2.2.13)$$

Je však třeba zmínit, že tyto vzorce jsou pouhou aproximací, protože v Langevinově dynamice může částice náhle změnit trajektorii v důsledku náhodné srážky s imaginární částicí okolí. Také Taylorův rozvoj nemůže být použit za těchto podmínek, jelikož vyžaduje spojité derivace rozvíjených funkcí. Z tohoto důvodu je přesnost algoritmu BBK nižší, než je odhad získaný Taylorovým rozvojem. Přesnost integrátoru BBK stoupá s $\gamma \rightarrow 0$ a $\delta t \rightarrow 0$. Metoda BBK je implementována v softwarovém balíku NAMD pro molekulárně dynamické simulace NVT souborů a je jedním z možných rozšíření Verletovy metody [103-104].

Prediktor-korektor (PC) algoritmy

Zcela jinou třídu představují algoritmy prediktor-korektor (PC) [105-106]. Tyto algoritmy ukládají řešení vypočtená v předchozích krocích a používají je k výpočtu (pomocí explicitních metod) “předpovědi“ řešení v kroku příštím (predikční krok). Poté vypočtou (pomocí implicitních metod) “opravu“ řešení (korekční krok). Rozdíl mezi řešeními vypočtenými v predikčním a korekčním kroku může být použit k další korekci řešení.

Každý z výše probraných algoritmů má své výhody a nevýhody. Volba vhodného algoritmu záleží na požadované přesnosti výpočtu, velikosti studovaného systému, velikosti paměti a rychlosti počítače, který je k dispozici.

Přesnost výpočtu - porovnání zaokrouhlovacích chyb a chyb metod

V obecné rovině je potřeba sledovat dva základní zdroje nepřesností výpočtu: zaokrouhlovací chyby a nepřesnosti algoritmu. Zaokrouhlovací chyby jsou zjevně chyby způsobené zaokrouhlením výsledků numerických výpočtů na počítačem zpracovatelnou přesnost. Význam této chyby roste s počtem cyklů (časových kroků

Δt) výpočtu. S ohledem na zaokrouhlovací chyby je tudíž nejvýhodnější použít co nejméně kroků a použít co největší délku časového kroku Δt . Situace se zkomplikuje, vezme-li do úvahy chyby v přesnosti vlastního algoritmu. Algoritmy popsané výše jsou založeny na diskretizaci diferenciálních rovnic. Užijeme-li příliš dlouhý krok Δt v čase, pak toto přiblížení (diskretizace) pozbývá smyslu (lépe řečeno s velkým Δt nemá diferenční rovnice příliš společného s vlastní diferenciální rovnicí, kterou se snažíme řešit). Chceme-li sledovat systém po dobu τ ve smyslu celkové doby vývoje simulovaného systému v čase (nezaměňovat s výpočetním časem), pak s délkou kroku Δt je tento čas rozdělen na $\tau/\Delta t$ intervalů. Pro rozumný výpočet je nutné použít délku intervalu $\Delta t = (0,05-0,1) \Delta t_{\max}$, kde Δt_{\max} je maximální časový krok umožňující získat stabilní trajektorii. Nyní se pokusíme najít kvantitativní vyjádření pro Δt_{\max} .

Příklad vyšetření stability Verletovy metody

Nejjednodušším příkladem, na kterém můžeme studovat vývoj chyby metody, je harmonický oscilátor s lokální konstantou pružiny $k(x[n])$, jí odpovídající vlastní frekvencí $\omega[n]$ a potenciálem U [107]. Zde jsou pro připomenutí rovnice harmonického oscilátoru:

$$\begin{aligned} \frac{m d^2 x}{dt^2} &= f(x) = \frac{-dU(x)}{dx} = -k(x) x \\ a[n] &= -\frac{k(x[n])}{m} x[n] = -\omega[n]^2 x[n] \end{aligned} \quad (3.2.2.14)$$

Pro chybu metody $\varepsilon[n]$ platí stejný vztah jako pro polohu r .

$$\varepsilon[n+1] = (2 - \omega^2 \Delta t^2) \varepsilon[n] - \varepsilon[n-1] \quad (3.2.2.15)$$

Protože nás zajímá změna této chyby v závislosti na počtu kroků, najdeme řešení pro $\varepsilon[n]$ ve tvaru:

$$\varepsilon[n+1] = g \varepsilon[n] = g^n \varepsilon[0] \quad (3.2.2.16)$$

V tomto případě $|g| > 1$ odpovídá růstu chyby a $|g| < 1$ odpovídá zmenšování chyby. Nejlepší případ nastává pro $|g| = 1$. Rovnici (3.2.2.15) přepíšeme do tvaru

$$\begin{aligned} g \varepsilon[n] &= (2 - \omega^2 \Delta t^2) \varepsilon[n] - (1/g) \varepsilon[n] \\ g^2 - (2 - \omega^2 \Delta t^2) g + 1 &= 0 \end{aligned} \quad (3.2.2.17)$$

což znamená, že obě řešení polynomiální rovnice druhého řádu pro g (tj. g_1 a g_2) nemohou být současně < 1 (má-li platit, že $g_1 g_2 = 1$). Mají-li splňovat podmínku $g_1 g_2 = 1$, pak řešení jsou

$$\begin{aligned} g_1, g_2 = e^{i\theta}, e^{-i\theta} & \quad \text{pro} \quad \omega \Delta t < 2 \\ g_1 < -1, -1 < g_2 < 0 & \quad \text{pro} \quad \omega \Delta t > 2 \end{aligned} \quad (3.2.2.18)$$

V případě harmonického oscilátoru maximální délka časového kroku závisí na fyzikálních podmínkách (přesněji řečeno na tom, ve které oblasti potenciálu U se systém pohybuje).

Pro PC algoritmy je situace poněkud složitější. Poznamenejme pouze, že např. pro druhý a pátý řád Gearova algoritmu platí následující omezení:

$$\omega\Delta t < \sqrt{2}$$
$$\omega\Delta t < \frac{1}{2}$$
(3.2.2.19)

Z čehož plyne, že Verletův algoritmus dovoluje použití delšího časového kroku Δt než PC algoritmy. Pro malé Δt je PC algoritmus přesnější než algoritmy Verletovy. Nicméně PC algoritmy ztrácejí rychle přesnost, když se $\omega\Delta t$ začne přibližovat k hodnotě 0,5. To znamená, že PC algoritmy poskytují buď velmi přesné řešení, nebo velmi nestabilní řešení (v závislosti na Δt). Zatímco Verletovy algoritmy poskytují stabilní řešení s relativně velkou chybou. Pro daný problém je tudíž nezbytné zvolit správnou výpočetní (integrační) metodu v závislosti na požadované délce studovaného intervalu τ (a kroku Δt), počítači který je k dispozici a požadované přesnosti výpočtu. V molekulárně dynamických simulacích požadujeme metodu s co nejdelším stabilním časovým krokem a tento požadavek splňují Verletovy algoritmy.

Shrnutí

Na závěr je možné shrnout, že algoritmy rychlostní Verletův, Leap-frog a Beemanův jsou si velmi podobné. Je snadné ukázat, že Leap-frog a rychlostní verze Verletova algoritmu jsou ekvivalentní jeho základní verzi (v níž explicitně nevystupují rychlosti). Poskytují tedy stejnou přesnost výpočtu, výpočetní náročnost a stabilitu výpočtu. Metody prediktor-korektor umožňují dosáhnout vyšší přesnosti výpočtu než Verletova metoda, nicméně velkou výhodou Verletova algoritmu je jeho enormní stabilita a fakt, že zrychlení částic (síly) je nutné vypočítat pouze jednou pro každý časový krok Δt . Význam této výhody je patrný obzvlášť pokud je výpočet prováděn pro systém s velkým počtem částic, které jsou svázány dalekodosačovými interakcemi.

3.3. Použité modely

3.3.1. Termodynamické soubory – NVE, NVT, NPT

V předchozím textu jsme předpokládali, že je systém izolovaný od vnějšího okolí a že nedochází k výměně energie ani částic. V termodynamické terminologii se takový systém nazývá *mikrokanonický*. Lze snadno nahlédnout, že tento přístup není zcela realistický a je pravděpodobné, že si studovaný systém ve skutečnosti vyměňuje se svým okolím (lázní) přinejmenším energii, čímž je udržován na konstantní teplotě. Toto odpovídá kupříkladu situaci, kdy je daný systém (molekula, protein) studován ve vodném roztoku a tepelná kapacita roztoku je mnohonásobně vyšší, než energie uvolňovaná nebo přijímaná naším systémem (molekulou, proteinem). V termodynamickém názvosloví se takový systém nazývá *kanonický*. V některých případech je možné, že námi studovaný systém nejen vyměňuje energii s lázní (jsa udržován na konstantní teplotě), ale může také docházet k výměně částic (atomů či iontů v chemické reakci). V takovém případě nezbyvá, než zahrnout energii

(chemický potenciál) odpovídající výměně částic mezi studovaným systémem a lázní. V termodynamice se takový systém nazývá *grandkanonický* soubor a v našem pojednání se jím nebudeme zabývat. V některých případech je nutné studovat systémy za podmínek konstantního tlaku. Takový systém se nazývá *izobarický*.

Mikrokanonický soubor (NVE)

Mikrokanonický soubor, jak již bylo řečeno, nevyměňuje ani energii ani částice se svým okolím (lázní). V molekulárně-dynamických simulacích se tento případ realizuje, pokud numericky řešíme klasické pohybové rovnice. Sledujeme pak pohyb systému ve fázovém prostoru po ploše konstantní energie. Jde o často používaný přístup díky tomu, že je nejméně náročný na výpočetní čas. A to i přesto, že není zcela realistický.

Kanonický soubor (NVT)

Realističtější je popis, kde je systému umožněno vyměňovat si energii se svým okolím (termodynamickou lázní) a tak udržovat konstantní teplotu T . Samozřejmě takové vylepšení modelu není zadarmo a nezbyvá, než přikročit ke komplikovanějšímu modelu vyžadujícímu více výpočetního času. Jelikož pohybové rovnice popisují pouze systém na ploše konstantní energie, je nezbytné je modifikovat. Vhodným způsobem zahrnutí tepelné lázně je zavedení dodatečných vnějších parametrů [99]. K řešení takto upravených rovnic je poté použito stejného přístupu, jako v případě souboru mikrokanonického.

Jednou z možností je Berendsenova metoda, která je spolu s Nosé-Hooverovou metodou implementována v softwarovém balíku AMBER [98]. Berendsenova metoda spočívá ve vynásobení rychlostí v každém časovém kroku koeficientem λ . V Nosé-Hooverově metodě přidáme fiktivní stupeň volnosti s hmotností a termodynamickým třecím koeficientem. Tzv. Langevinovy rovnice jsou rozšířeny o člen pro tření úměrný rychlosti částice a o náhodnou sílu, které společně imitují molekulární kolize a viskozitu:

$$\dot{\mathbf{r}}_i = \frac{\dot{\mathbf{p}}_i}{m_i}$$

$$\dot{\mathbf{p}}_i = \mathbf{f}_i - \gamma \mathbf{p}_i + \mathbf{R}_i \quad (3.3.1.1)$$

kde γ je tlumící konstanta (kolizní frekvence). Vektor náhodné síly \mathbf{R}_i je stacionární Gaussovský proces se statistickými vlastnostmi:

$$\langle \mathbf{R}_i(t) \rangle = 0$$

$$\langle \mathbf{R}_i(t) \mathbf{R}_i(t')^T \rangle = 2\gamma k_B T m_i \delta(t - t') \quad (3.3.1.2)$$

kde k_B je Boltzmannova konstanta a T je teplota Langevinova termostatu. Tlumící konstanta γ ovládá jak velikost třecí síly, tak i odchylku náhodných sil. Díky tomu systém konverguje k Boltzmannovskému rozdělení charakterizovaném teplotou T . Čím větší je hodnota tlumící konstanty γ , tím větší je vliv okolí. S rostoucími

hodnotami γ přechází systém od inertního k difusnímu (Brownovskému) režimu. Hodnoty γ se určují pomocí hydrodynamické teorie. Ze Stokesova zákona popisujícího lineární závislost třecího odporu sférické částice v roztoku na jejím poloměru plyne, že velikost třecí síly je rovna $6\pi\eta a_r$ násobku rychlosti částice, kde a_r je hydrodynamický poloměr všech sférických částic a η je viskozita. V (3.3.1.1) lze hodnotu γ volit:

$$\gamma = 6\pi\eta a_r / m \quad (3.3.1.3)$$

kde m je hmotnost částice. Hodnoty kolizní frekvence γ lze také volit tak, aby výsledná translační difuzní konstanta D_t souhlasila s experimentálními hodnotami v difuzní oblasti (čili pro velká γ). Tento vztah se řídí vzorcem:

$$D_t = k_B T / \sum_i m_i \gamma \quad (3.3.1.4)$$

ve kterém suma probíhá přes všechny hmotnosti m_i obsažené v systému. Tento vzorec se v případě použití Stokesova zákona pro určení hodnot kolizní frekvence γ redukuje na tvar:

$$D_t = k_B T / 6\pi\eta a_r \quad (3.3.1.5)$$

což je Stokes-Einsteinův zákon pro difuzi sférických částic skrz kapalinu pro nízká Reynoldsova čísla. Langevinova dynamika je pro NVT soubory implementována v softwarových balících NAMD a ACEMD [108-111].

Izotermicko-Izobarický soubor (NPT)

Popis systémů za konstantního tlaku je poněkud složitější než v případě výše diskutovaného systému za konstantní teploty. V tomto případě totiž objem systému nezůstává konstantní, ale fluktuuje tak, aby tlak v systému zůstal konstantní. K fiktivním proměnným zavedeným pro zachování konstantní teploty v kanonickém souboru nám přibudou ještě další fiktivní proměnné odpovídající fiktivnímu zařízení - „pístu“ - zodpovědnému za vyrovnávání tlaku [99]. Tato metoda byla poprvé popsána Andersenem a Hooverem. Výše zmíněná Berendsenova metoda se dá použít i na nezbytné udržování konstantního tlaku [98]. Tato metoda spočívá ve vynásobení souřadnic atomů a rozměrů periodického boxu koeficientem μ v každém časovém kroku. Obě metody, jak metoda popsána Andersenem tak metoda popsána Hooverem a Berendsenem jsou opět implementovány v softwarovém balíku AMBER.

Nicméně tyto metody opět neobsahují stochastickou složku, která přibližuje modely k systémům nacházejícím se v reálném světě. Proto se Andersen-Hooverova metoda kombinuje s Langevinovou dynamikou. Deterministické rovnice jak pro rychlost částic tak i barostat jsou převedeny na Langevinovy rovnice. Rovnice pro i -tou částici v Langevinově barostatu mají následující tvar:

$$\dot{\mathbf{r}}_i = \frac{\dot{\mathbf{p}}_i}{m_i} + \frac{p_\epsilon}{W} \mathbf{r}_i$$

$$\dot{\mathbf{p}}_i = -\nabla_{\mathbf{r}_i} \Phi - \left(1 + \frac{3}{N_f}\right) \frac{p_\epsilon}{W} \mathbf{p}_i - \gamma \mathbf{p}_i + \mathbf{R}_i$$

$$\dot{V} = 3V p_\epsilon / W$$

$$\dot{p}_\epsilon = 3V(X - P_{ext}) + \frac{3}{N_f} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - \gamma_p p_\epsilon + R_p \quad (3.3.1.6)$$

kde R_p je stochastická síla působící na barostat. W je fiktivní hmotnost barostatu a V objem. N_f je počet stupňů volnosti. Dimenze je rovna třem. Je třeba připomenout, že v nehamiltonovském případě má zachovávaná veličina ve fázovém prostoru následující tvar:

$$H' = \mathcal{H} + PV + p_\epsilon^2 / 2W \quad (3.3.1.7)$$

ve kterém je první člen Hamiltonián částicového pod systému. V softwarovém balíku NAMD jsou pro NPT soubory implementovány pohybové rovnice, které se od (3.3.1.6) liší v členu $1 + 3/N_f$ druhé a čtvrté rovnice. Člen $3/N_f$ je ve většině případů zanedbatelný [103, 109-111].

3.3.2. Periodické okrajové podmínky, Ewaldova sumace, Particle mesh metoda, Particle mesh Ewaldova metoda

Periodické okrajové podmínky

Důvodem pro vytvoření myšlenkové konstrukce vedoucí k periodickým okrajovým podmínkám (periodic boundary conditions PBC) je získání realistického modelu systému částic [101-102]. Vyjdeme ze základní představy, při níž biomolekulu umístíme do základní buňky a obklopíme vodním roztokem. Jestliže bychom uvažovali základní buňku s pevnou hranicí (nepropustné stěny), pak by tato hranice ovlivňovala příliš vysoký počet částic. Zamezení nežádoucího ovlivňování částic hranicí lze dosáhnout dvěma způsoby. První způsob spočívá ve vytvoření souboru s velmi vysokým počtem částic, to je však technicky náročné. Druhý způsob spočívá v tom, že základní buňku obklopíme identickými kopiemi této buňky tak, aby byl pokryt celý uvažovaný prostor. Touto myšlenkovou konstrukcí jsme získali systém, v němž nemusíme požadovat nepropustné stěny. Opustí-li částice základní buňku, pak kopie této částice vstoupí do uvažované buňky ze sousední buňky. Celkový počet částic ve všech buňkách zůstane zachován.

Při použití PBC je interakční energie dána nekonečnou sumou přes všechny periodické obrazy. U potenciálů, které dostatečně rychle ubývají (např. Lennard-Jonesův), můžeme tento potenciál useknout („cut off“) v jisté vzdálenosti od částice a tím zamezit výpočtu nekonečných sum. Nejjednodušší způsob odseknutí spočívá v tom, že potenciál $u(\mathbf{r})$ postavíme roven nule ve vzdálenosti větší než $r_{\text{cut off}}$. Je vhodné volit $r_{\text{cut off}} \leq L/2$, kde L je délka hrany buňky. Pak leží v oblasti interakce maximálně jeden obraz dané částice.

Poněkud komplikovanější situace nastává, řešíme-li problém velkého systému částic, kde interakční potenciál neklesá dostatečně rychle se vzdáleností. Pak nezbyvá, než do výpočtu chování systému zahrnout všechny částice a jimi vytvářené potenciály. Příkladem takového systému je systém elektrostaticky interagujících částic:

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = q_i \nabla_i V(\mathbf{r}_i), \quad i = 1, 2, 3, \dots, N \quad (3.3.2.1)$$

kde m_i a r_i jsou hmotnosti a pozice jednotlivých částic, q_i je jejich elektrický náboj a (elektrostatický) potenciál je vyjádřen výrazem:

$$V(\mathbf{r}_i) = \sum_{j \neq i}^N \frac{q_j}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (3.3.2.2)$$

Naším úkolem je výpočet potenciální energie referenčního náboje umístěného v libovolném místě takového systému. Je několik způsobů jak tento problém řešit. Velmi často používaným přístupem je Ewaldova sumace a takzvaná “particle mesh” (PM) metoda [112-116].

Ewaldova sumace

Ewald v roce 1921 vyřešil problém výpočtu potenciální energie pro periodické systémy s dalekodosahovými interakcemi velmi elegantním způsobem [112, 114]. Mějme interakční potenciál $u_{qq}=q_1q_2/r$ mezi dvěma nabitými částicemi. V Ewaldově sumačním přiblížení je základní buňka obsahující $N/2$ kladně nabitých a $N/2$ záporně nabitých částic v nějakém prostorovém uspořádání interpretována jako izolovaný krystalografický element obklopený nekonečným počtem identických kopií sebe sama. Celý systém je neutrálně nabitý a obsahuje nekonečný počet nábojů umístěných v pozicích \mathbf{r}_{j+} a \mathbf{r}_{j-} .

Celkový potenciál libovolného iontu i v pozici \mathbf{r}_i v základní buňce je vyjádřen konečným rozdílem dvou nekonečných divergentních řad:

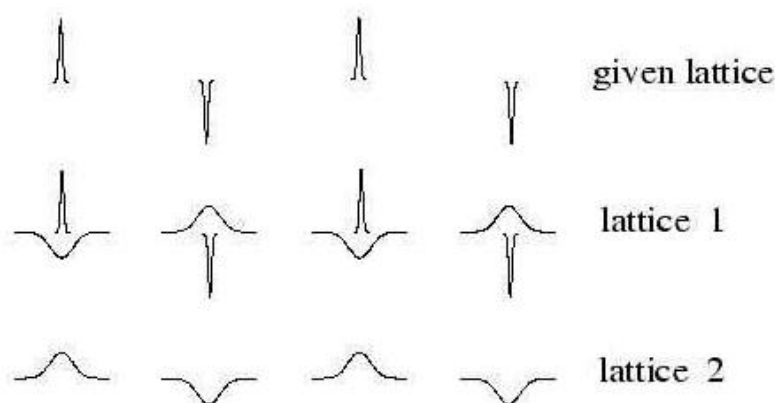
$$\Phi(\mathbf{r}_i) = q \sum_{j+=1}^{\infty} \frac{1}{|\mathbf{r}_i - \mathbf{r}_{j+}|} - q \sum_{j-=1}^{\infty} \frac{1}{|\mathbf{r}_i - \mathbf{r}_{j-}|} \quad (3.3.2.3)$$

Místo rozepsání potenciálu do řad přes bodové náboje můžeme nejprve přepsat tyto náboje jako δ funkce nábojových hustot:

$$\rho(\mathbf{r}) = q \sum_{j+=1}^{\infty} \delta(\mathbf{r} - \mathbf{r}_{j+}) - q \sum_{j-=1}^{\infty} \delta(\mathbf{r} - \mathbf{r}_{j-}) \quad (3.3.2.4)$$

a poté tyto δ funkce rozložíme do Fourierových řad, jejichž členy určují Fourierovy komponenty $\Phi(\mathbf{k})$ elektrostatického potenciálu. Jelikož Fourierova reprezentace δ funkcí požaduje nekonečně mnoho členů, mohou výpočty ve Fourierově prostoru vést k problémům s konvergencí. Řešením je rozdělení potenciálu na dvě „slušné“ se chovající části, z nichž jedna je reprezentována v \mathbf{r} -prostoru a druhá v \mathbf{k} -prostoru rychle konvergujícími řadami.

Tento postup si budeme demonstrovat na jednodimenzionální iontové mřížce s rozdělením nábojů zobrazeným v obrázku (obrázek 3.1.).



Obrázek 3.1.: Schématický náčrt rozdělení nábojů [112]

Přidáme k bodovým nábojům vyjádřeným δ funkcemi “mrak” Gaussovského náboje s opačným znaménkem (lattice 1 v obrázek 3.1.):

$$\rho'(\mathbf{r}) = -q_j \left(\frac{\eta^2}{\pi} \right)^{3/2} e^{-\eta^2(\mathbf{r}-\mathbf{r}_j)^2} \quad (3.3.2.5)$$

Další mřížka (lattice 2) je zavedena pro kompenzaci přidaných gaussovských nábojů tak, že “mřížka 1 + mřížka 2 = původní mřížka”.

Příspěvky těchto dvou mřížek k celkovému potenciálu jsou počítány odděleně:

- Mřížka 1: Při pohledu z větší dálky se Gaussovské nábojové mraky podobají bodovým nábojům δ funkcí, což doprovází efektivní kompenzace původních nábojů. Efekt mřížky 1 je tudíž nejlépe vypočitatelný v \mathbf{r} -prostoru, kde budou řady konvergovat poměrně rychle. Tato konvergence bude rychlejší, budou-li Gaussiány úzké (což nastane, bude-li parametr η veliký).

- Mřížka 2: Potenciálový součet je vyhodnocen v \mathbf{k} -prostoru. Jsou-li Gaussiány široké (je-li parametr η malý) potřebujeme menší počet Fourierových komponent.

Vhodným nastavením η tedy můžeme dosáhnout optimální konvergence obou řad.

Nyní přejdeme k třidimenzionálnímu modelovému systému s uvažovanou krychlovou základní buňkou o délce hrany L , která obsahuje N nabitých částic.

Zavedeme následující definice:

Fourierovy vektory $\mathbf{k} \equiv 2\pi(k_x, k_y, k_z)/L$, kde k_x jsou celá čísla. Dále definujeme mezičásticové vektory (obsahující všechny periodické obrazy základní buňky) $\mathbf{r}_{i,j,n} \equiv \mathbf{r}_j + n\mathbf{L} - \mathbf{r}_i$ pro $i, j=1, \dots, N$, kde $n\mathbf{L}$ je vektor posunutí v periodické mřížce.

Ewaldův součet pak má tento tvar:

$$\Phi(r_i) = \frac{2\pi}{L^3} \sum_{j=1}^N q_j \left[\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}_{ij}} k^{-2} e^{-k^2/4\eta^2} + \sum_{\mathbf{n}} F(\eta|\mathbf{r}_{i,j}, \mathbf{n}) \right] \quad (3.3.2.6)$$

kde

$$F(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt \quad (3.3.2.7)$$

Tato řada však obsahuje ještě dva problémy. První problémem je, že zahrnuje interakci Gaussovských nábojů se sebou samými a tím způsobuje vznik falešného příspěvku k funkci potenciální energie. Tento příspěvek musí být odečten. Druhým problémem je, že je potřeba vhodným způsobem přejít k nekonečně velkému systému. Tento problém vyřešíme zavedením pomyslné jednotkové buňky (přibližně kulového tvaru), obklopené médiem s dielektrickou konstantou ϵ . K nekonečně velkému systému pak přejdeme postupným zvětšováním počtu těchto buněk. Pro potenciální energii po zahrnutí těchto dvou oprav máme rovnici:

$$E_{pot} = \frac{1}{2} \sum_{i=1}^N q_i \Phi(\mathbf{r}_i) - \frac{\eta}{\sqrt{\pi}} \sum_{i=1}^N q_i^2 + \frac{2\pi}{3L^3} \left| \sum_{i=1}^N q_i \mathbf{r}_i \right|^2 \quad (3.3.2.8)$$

kteřá nám dává předpis pro výpočet potenciální energie nekonečně velkého systému částic svázaných dalekodosahovými interakcemi. Ze znalosti potenciální energie a z pohybových rovnic pak můžeme vypočítat polohy částic v čase. Pro rozumně malý (< 1000 částic) systém je tento přístup s patřičnými optimalizacemi sčítacího procesu (kupříkladu užitím operací symetrie) velmi efektivní a je často užíván. Je-li však naším úkolem vypočítat chování daleko většího systému, nezbyvá než přikročit k další aproximaci zrychlující výpočet. Takovou aproximací je particle mesh (PM) metoda.

Particle mesh (PM) metoda

PM je metoda [113-114], která značně zrychluje výpočet, ovšem za cenu snížení přesnosti výpočtu v mikroskopickém měřítku. Tato ztráta přesnosti nepředstavuje problém, jestliže nás zajímá především celkové, kolektivní chování systému na vzdálenostech přibližně stejných, jako jsou rozměry samotného systému. Tato metoda je používána pro výpočet tak rozdílných problémů, jako je chování proteinů, elektricky neutrálního plazmatu, nebo galaxií. Podstata metody spočívá v tom, že systém 10^N částic je modelován jako systém 10^A super částic, z nichž každá obsahuje 10^B částic (kde $A+B = N$). Při vhodné volbě A a B pro dané N je možné modelovat chování systému N částic s požadovanou přesností a se sníženými nároky na výpočetní čas. Existuje mnoho různých schémat jak rozdělit daný systém částic a jak (pokud je to vyžadováno) sledovat chování uvnitř super částic.

Ve čtvercových dvojdimenzionálních prostorech zavedeme buňky tak, že vedeme čáry rovnoběžně s hranicemi. Průsečíky těchto čar nám definují uzly. V metodě NGP (Nearest-grid-point - nejbližší uzel) se propojuje částice s uzlem tak, že k danému uzlu mřížky přiřadíme všechny částice, jejichž vzdálenost od uzlu je menší než $L/2$, kde L je délka hrany buňky. Počet částic by měl být mnohem vyšší než počet uzlů buněk. To nám zajistí, že každá buňka bude v průběhu simulace obsahovat průměrně několik částic.

Pohyb částic v elektrostatickém poli je popsán rovnicí:

$$\frac{d}{dt}(v_i) = \frac{q_i}{m_i} E, \quad i = 1, 2, 3, \dots, N \quad (3.3.2.9)$$

Nábojovou hustotu $\rho(\mathbf{r})$ získáme zobrazením částic na síti pomocí váhové funkce W a rovnice:

$$\rho(\mathbf{r}) = \sum_j q_j W(\mathbf{r}_j - \mathbf{r}), \quad j = 1, 2, 3, \dots, N_{cell} \quad (3.3.2.10)$$

Váhová funkce W popisuje tvar částic a můžeme ji vybrat specificky pro daný problém. Pro metodu NGP je $W \equiv 1$, jestliže je vzdálenost částice od uzlu menší než $L/2$.

Vlastní cyklus PM metody lze rozdělit na čtyři části. Předpokládejme, že byl úspěšně proveden n -tý časový krok:

1. máme jako vstup pro další časový krok polohy \mathbf{r}_i a rychlosti \mathbf{v}_i i -té částice. Částice přidělíme k uzlům mřížky podle zvolené metody (viz. NGP metoda uvedená výše). Náboj uzlu má potom hodnotu součtu nábojů jednotlivých částic příslušejících k danému uzlu vynásobených váhovou funkcí W . V jednodimenzionálním prostoru máme tento náboj vyjádřen rovnicí:

$$\rho_j^{n+1} = \sum_i q_i W(x_i - x_j) \quad (3.3.2.11)$$

kde v případě NGP metody je $W \equiv 1$ (pokud je vzdálenost částice od uzlu menší než $L/2$).

2. V tomto kroku se z hustot nábojů v uzlech vypočítá pole E_j v uzlech. V jednodimenzionálním prostoru použijeme vztah:

$$E_{j+\frac{1}{2}}^{n+1} = E_{j-\frac{1}{2}}^{n+1} + \rho_j^{n+1} \Delta x \quad (3.3.2.12)$$

neboť můžeme položit $E' = \rho$. Ve dvojdimenzionálním prostoru je situace poněkud složitější. Nejprve je potřeba z Poissonovy rovnice vypočítat potenciál Φ , což lze učinit pomocí rychlé Fourierovy transformace (FFT - fast fourier transform) a na základě tohoto potenciálu je potom nutné vypočítat elektrické pole E_j v uzlech.

3. V průběhu tohoto kroku je elektrické pole E_j působící v uzlech rozděleno na složky E_i působící na jednotlivé částice metodou inverzní k metodě, kterou jsme přidělili uzlu hustotu náboje vztaženou k nábojům jednotlivých částic.

4. V tomto kroku jsou získány nové souřadnice a rychlosti částic. Děje se tak pomocí vztahů:

$$v_i^{n+\frac{1}{2}} = v_i^{n-\frac{1}{2}} + \frac{q_i}{m_i} E_i^n \Delta t,$$

$$x_i^{n+1} = x_i^n + v_i^{n+\frac{1}{2}} \Delta t \quad (3.3.2.13)$$

Tím je cyklus uzavřen. Tyto kroky se pak cyklicky opakují, až je vypočteno chování systému na požadovaném časovém intervalu

Particle mesh Ewaldova metoda

Výše popsaná Ewaldova sumace aplikovaná v particle mesh aproximaci umožňuje zahrnout dalekodosahové elektrostatické interakce v molekulárně dynamických simulacích s přesností, která je dostatečná pro simulování chování velkých biomolekul po dostatečně dlouhou dobu (stovky nanosekund) [114-116]. Particle mesh Ewaldova metoda umožňuje úspěšně řešit problémy spojené s vysokým nábojem nukleových kyselin. Je to doposud jediný známý způsob, jak zabránit rozpadu systému studované nukleové kyseliny v průběhu molekulárně dynamické simulace. Metody využívané dříve používaly k zamezení rozpadů struktur nukleových kyselin řešení, které neodpovídalo přirozenému prostředí - např. předdefinované dodatečné síly. Tato metoda umožňuje také zahrnout do systémů nukleových kyselin vodní obálku a tím dosáhnout stavů systémů, jež jsou v požadované shodě s realitou. Podstatnou výhodou Particle Mesh Ewaldovy metody je, že pro výpočet Ewaldovy sumace lze použít FFT, čímž je značně snížena výpočetní náročnost. V případě, že jsou jako báze funkce při interpolaci použity B-spliny, hovoříme o takzvané smooth Particle Mesh Ewaldově metodě (PME). Využití B-splinů přináší další zrychlení výpočtu dalekodosahových elektrostatických interakcí a v důsledku toho i celé molekulárně dynamické simulace [117].

3.4. Použitý software

Zatímco v předchozích kapitolách jsme se zabývali teorií, v následujících odstavcích se budeme zabývat praktickými hledisky molekulárně dynamických simulací. Předmětem našeho zájmu budou výhradně softwarové balíky použité při řešení této práce. Vynecháme proto např. řadu programů užívaných v bioinformatice. Probereme nejen softwarové balíky potřebné pro vytváření struktur biomolekul a jejich molekulárně dynamické simulace, ale i software vhodný pro vyhodnocování a zobrazování výsledků těchto simulací. Softwarové balíky jsou probírány v pořadí, ve kterém byly použity v průběhu řešení této práce.

Softwarový balík AMBER

Silové pole a softwarový balík AMBER (Assisted Model Building with Energy Refinement) umožňují provádět molekulárně-dynamické simulace biomolekul - především nukleových kyselin a proteinů.

Silové pole AMBER bylo vyvinuto skupinou Petera Kollmana na University of Kaliforni v San Francisku. Jeho součástí jsou silové konstanty pro nukleové kyseliny a aminokyseliny (ff94, ff99SB, ff03 ...), karbohydráty (Glycam), fosfolipidy (Lipid11) a malé organické molekuly (Gaff). Softwarový balík AMBER navíc obsahuje nástroj pro parametrizaci silových konstant pro nové molekuly (např. léky).

Vlastní programový balík AMBER je tvořen přibližně 50 programy, které pokrývají široké spektrum metod výpočetní chemie. Softwarový balík AMBER je neustále rozvíjen díky spolupráci mezi celou řadou vědeckých pracovišť. Nejvýznamnější

jsou týmy z Rutgersovy University a University of Utah, které spolupracovaly i na poslední verzi softwarového balíku AMBER, distribuované ve dvou částech: Amber12 a AmberTools12. Moduly v Amber12 jsou soustředěny kolem dvou nejdůležitějších simulačních programů *Sander* a *Pmemd*. *Sander* je tradiční nástroj pro výpočty, kdežto *Pmemd* je zaměřen na maximalizaci výkonu. V AmberTools12 jsou mimo jiné implementovány důležité nástroje pro přípravu simulovaných systémů (*LEAP*) a analýzu trajektorií (*ptraj*), které byly hojně využívány v rámci této práce.

Sander (Simulated Annealing with NMR-Derived Energy Restraints) - je modul, který umožňuje provádět minimalizaci energie a vlastní molekulárně dynamické simulace. Základním výstupem jsou soubory souřadnic - trajektorie, ale je možné ukládat i rychlosti či energie.

Pmemd vznikl přepsáním programu *Sander* se zaměřením na vysoký výkon a paralelizaci výpočtů prostřednictvím MPI (Message Passing Interface). Vzhledem k nedávno se objevivší možnosti enormního urychlení výpočtů na špičkových grafických kartách (GPU - graphics processing units) byla vytvořena verze *Pmemd.cuda*, která využívá programovací jazyk CUDA (Compute Unified Device Architecture) vyvinutý firmou NVIDIA.

LEAP - je program pro výstavbu molekulárního modelu a tvorbu vstupních souborů pro molekulárně-dynamické simulace obsahující souřadnice atomů, topologii molekuly a silové konstanty. Vstupními daty jsou kartézské souřadnice biomolekuly ve formátu *pdb* (Brookhaven Protein DataBank format). Topologie jednotlivých reziduí jsou extrahovány ze standardních databází a z nich je posléze vytvořena topologie celého simulovaného systému. *LEAP* rozhodne, které vazby, torzní úhly, dihedrální úhly a typy atomů se nacházejí v systému a extrahuje příslušné parametry ze souborů silových polí. Výstupem jsou dva soubory: *prmtop* (topologie, silové konstanty) a *inpcrd* (souřadnice). Tímto je simulovaný systém připraven pro vlastní molekulárně-dynamickou simulaci. Existují přitom dvě varianty modulu *LEAP* a to *tleap* a *xleap*. V rvním případě jsou instrukce zadávány z příkazové řádky, ve druhém případě je k dispozici komfortnější uživatelské prostředí.

ptraj slouží k analýze výsledných dat z molekulárně dynamických simulací. Slouží k výpočtu korelačních funkcí, středních kvadratických odchylek od referenční struktury, pseudorotačních úhlů, analýze vodíkových vazeb atd. Výsledkem jsou soubory dat, ze kterých je možné snadno vytvořit grafy například prostřednictvím programu Gnuplot. Existuje také paralelizovaná verze tohoto modulu - *ptraj.MPI*.

Program NAMD

NAMD (Not (just) Another Molecular Dynamics) je program navržený pro velmi náročné molekulárně dynamické simulace velkých biomolekulárních systémů (sestavující řádově z milionů atomů).

Softwarový balík NAMD je založen na paralelních objektech vytvořených jazykem Charm++, který byl na základě C++ vyvinut ve spolupráci mezi Parallel Programming Laboratory a Theoretical and Computational Biophysics Group University of Illinois [103, 119-121]. Charm++ je implementován pomocí tzv. converse runtime systému. NAMD byl poprvé uveden v roce 1995 a od té doby je průběžně vyvíjen, přičemž poslední stabilní verze NAMD 2.9 pochází z dubna 2012.

NAMD výborně škáluje na stovkách procesorů na high-end paralelních platformách, nebo s využitím gigabitového Ethernetu na desítkách procesorů na běžných výpočetních klastrech. K dispozici je zdrojový kód, takže NAMD může být odladěn na kterékoliv platformě podporující Ethernet a MPI.

Program NAMD umožňuje provádět různé typy molekulárně dynamických simulací - při konstantní teplotě udržované přeškálováním nebo pomocí Langevinovy dynamiky, či při konstantním tlaku pomocí Nose-Hoover-Langevinovy nebo Berendsenovy metody. Je možné využít také smooth-Particle-Mesh-Ewaldovu metodu pro sumaci dalekodosahových interakcí. Je možné vypočítat volnou energii simulovaného systému. Molekulární dynamika přitom slouží k vygenerování souboru konfigurací, ze kterého jsou následně vypočítány termodynamické veličiny simulovaného systému. Pro NVE soubory používá NAMD jako integrátor rychlostní Verletovu metodu, tato metoda je symplektická a vyžaduje pouze jeden výpočet sil v každém časovém kroku. Vzhledem k tomu, že je v NAMD u NVT souborů aplikována stochastická Langevinova rovnice, je této rovnici přizpůsoben i integrátor, kterým je Brunger-Brooks-Karplusova metoda, což je integrátor založený na Verletově algoritmu. Podobně je přizpůsobený integrátor i v případě NPT souborů. Pro kontrolu tlaku se používá Nose-Hoover-Langevinova metoda, což je kombinace Nose-Hooverovy metody pro kontrolu tlaku a Langevinovy dynamiky.

NAMD je kompatibilní s programy AMBER a CHARMM. Je schopen pracovat s jejich formáty vstupních souborů, silovými poli atd.

NAMD může běžet i v interaktivním módu, kdy je možné do simulovaného systému manuálně vnášet externí síly, k čemuž se využívá grafické rozhraní VMD, které je možné použít i pro analýzu trajektorií. Grafické rozhraní VMD bylo vyvinuto skupinou Theoretical and Computational Biophysics group na University of Illinois.

V roce 2002 obdrželi Prof. Kale, James C. Phillips, Gengbin Zheng, Sameer Kumar cenu Gordon Bell Award za paralelizaci programu NAMD.

Program ACEMD

ACEMD [108, 117, 122-124] je další výkonný softwarový balík pro molekulární dynamiku se zaměřením na biomolekuly. Na rozdíl od ostatních softwarových nástrojů, které byly primárně navrženy pro molekulárně dynamické simulace prováděné na vysokém počtu nezávislých procesorů a tím při následném vývoji v mnoha ohledech limitovány, ACEMD byl od počátku optimalizován pro použití na GPU od firmy NVIDIA. V současné době se pravděpodobně jedná o nejvýkonnější softwarový balík s tímto zaměřením. Umožňuje běžně provádět molekulárně dynamické simulace v délce mikrosekund. Celý výpočet probíhá na GPU pouze s minimální komunikací s CPU. Je plně kompatibilní s CUDA platformou vytvořenou firmou NVIDIA pro paralelní výpočty na jimi produkovaných GPU. Rovněž je kompatibilní se standardem pro paralelní programování heterogenních počítačů OpenCL (Open Computing Language). Je spustitelný na všech 64-bitových distribucích Linuxu.

V ACEMD není v současné době implementován barostat, proto není možné provádět molekulárně dynamické simulace v NPT souboru. Ekvilibrační běhy je proto nutné provést pomocí jiného vhodného softwarového balíku (například NAMD). Jinak ale má ACEMD implementovány všechny důležité algoritmy pro

molekulárně dynamické simulace v NVT souborech. Pro výpočet dalekodosahových elektrostatických sil je použita metoda PME. Jako integrátor je implementována rychlostní Verletova metoda. Pro kontrolu teploty je zaveden Langevinův termostat. Program ACEMD je schopen pracovat se silovými poli AMBER a CHARMM.

Výpočty jsou prováděny s jednoduchou přesností plovoucí desetinné čárky, což enormně urychluje výpočet a přitom poskytuje dostatečně přesné výsledky.

V ACEMD je také implementována metoda přerozdělení hmoty z těžkých atomů na vodíky (HMR - Hydrogen Mass Repartitioning), která umožňuje prodloužení časového kroku až na 4fs.

Vstupní a výstupní soubory ACEMD jsou stejné jako analogické soubory u programu NAMD. Simulované systémy tak lze budovat prostřednictvím AmberTools či programu VMD.

Software použitý na vyhodnocení simulací

Programy, které jsou popisovány níže, již nejsou používány přímo pro provádění molekulárně dynamických simulací, slouží pro následné vyhodnocení molekulárně dynamických trajektorií.

CURVES+ se používá k analýze tzv. helikálních parametrů u nukleových kyselin (včetně dvou, tří a čtyř vláknových komplexů či modifikovaných nukleových kyselin [125]).

VMD slouží především k vizualizaci nasimulovaných trajektorií. Je schopen zpracovat nejrůznější formáty: *pdb*, *prmtop* spolu s *inpcrd*, *psf* spolu s *dcd* atd. [126].

RASMOL je program sloužící k vizualizaci proteinů, nukleových kyselin a malých molekul. Formát *pdb* je jedním z podporovaných formátů.

CHIMERA je program pro vizualizaci a analýzu molekulárních struktur a molekulárně dynamických trajektorií. Kromě toho lze provádět např. sekvenční či strukturální alignment či vizualizovat cryoEM mapy elektronové hustoty. Podporuje také celou řadu formátů vstupních a výstupních souborů [127].

ICM MOLSOFT je softwarový balík, který obsahuje celou řadu nástrojů pro strukturální biologii. Je přímo propojen s online databázemi.

4. Výsledky a diskuze

4.1. Molekulárně dynamické simulace komplexů RNA-dependentní RNA polymerázy Noroviru s duplexem RNA:RNA a CTP či jeho analogy

Motivace

Všechny viry s genomem uloženým ve formě jednovláknové ssRNA(+) jsou v hostitelské buňce reprodukovány pomocí RNA-dependentní RNA polymerázy (RdRp). Příkladem jsou virus hepatitidy C (HCV) a virus Norwalk (NV). HCV je rozšířen mezi 3% světové populace, z nichž pouze čtvrtina má symptomy. HCV způsobuje cirhózu a rakovinu jater [23, 128-130]. NV vyvolává u lidí epidemickou nebakteriální akutní gastroenteritidu.

V současné době se pro léčbu HCV používá interferon- α a ribavirin. Tento léčebný postup má pouze omezenou účinnost a celou řadu vedlejších účinků. Snaha o celkové zlepšení léčebných výsledků a snížení vedlejších účinků léčby HCV vede k hledání nových cíleně působících antivirotik. Nedávno byla pro léčbu v kombinaci s interferonem a ribavirinem odsouhlasena dvě nová léčiva telaprevir a boceprevir [128]. Ta fungují jako inhibitory HCV proteázy. Nicméně přinášejí další vedlejší účinky. Z toho důvodů pokračuje hledání nových léčiv a léčebných mechanismů.

RdRp není obsažena v buňkách savců a proto obecně je vhodným cílem pro inhibici. Jedna z možností je alosterická inhibice RdRp pomocí non-nukleosidových inhibitorů (NNI), které zamezují dosažení aktivní konformace polymerázy [131]. Klinické testy prokázaly nevhodnost monoterapie pomocí NNI v důsledku vzniku rychlé rezistence viru vůči léčivu. Ve struktuře HCV RdRp se však nachází hned několik vhodných míst, která umožňují kombinovanou léčbu pomocí několika různých NNI [128]. Příkladem vhodného místa pro navázání NNI k enzymu HCV RdRp je dutina v palci, která je u neaktivního enzymu obsazena jednou ze „špiček prstů“. NNI navázaný v této dutině přerušuje interakce mezi prsty a palcem a udržuje tak enzym v neaktivní formě [129].

Hladina rezistence virů vůči tzv. nukleosidovým inhibitorům (NI) RdRp je ve srovnání s NNI nižší. Je totiž málo pravděpodobné, že dojde ke vzniku rezistentního kmenu s mutací aktivního místa RdRp, které by na jednu stranu efektivně bránila vazbě NI a přitom zásadně nenarušila funkčnost RdRp [128].

HCV RdRp mohou inhibovat chemicky modifikované dinukleotidy, které imitují produkt iniciace replikace virového genomu, kdy spojením dvou GTP vznikne GpG [128, 132-133].

Nejčastěji se ale u HCV RdRp jako NI používají 2'-C-metyl nukleosidy, které vykazují antivirovou aktivitu bez ohledu na typ báze [32, 128, 131-136]. Nukleosidové inhibitory jsou hostitelskou buňkou postupně fosforylovány na nukleotidy - nukleosidmonofosfáty (NMP), nukleosiddifosfáty (NDP) a nukleosidtrifosfáty (NTP). Inhibitory ve formě chemicky modifikovaných NTP pak

mohou být začleněny do syntetizovaného vlákna RNA, kde působí jako terminátory přepisu virového genomu.

V důsledku chemických modifikací ve struktuře NI je však někdy zejména jejich první fosforylace buněčnými enzymy (při které vzniká chemicky modifikovaný nukleotid) málo pravděpodobná. Ve studii [135], byla proto připravena série adenosin 5'-fosfonátových analog AMP, která se v buňce musí fosforylovat pouze dvakrát. Následná analýza ukázala, že coATP funguje jako terminátor RNA řetězce, kdežto cocATP jako klasický kompetitivní inhibitor [135].

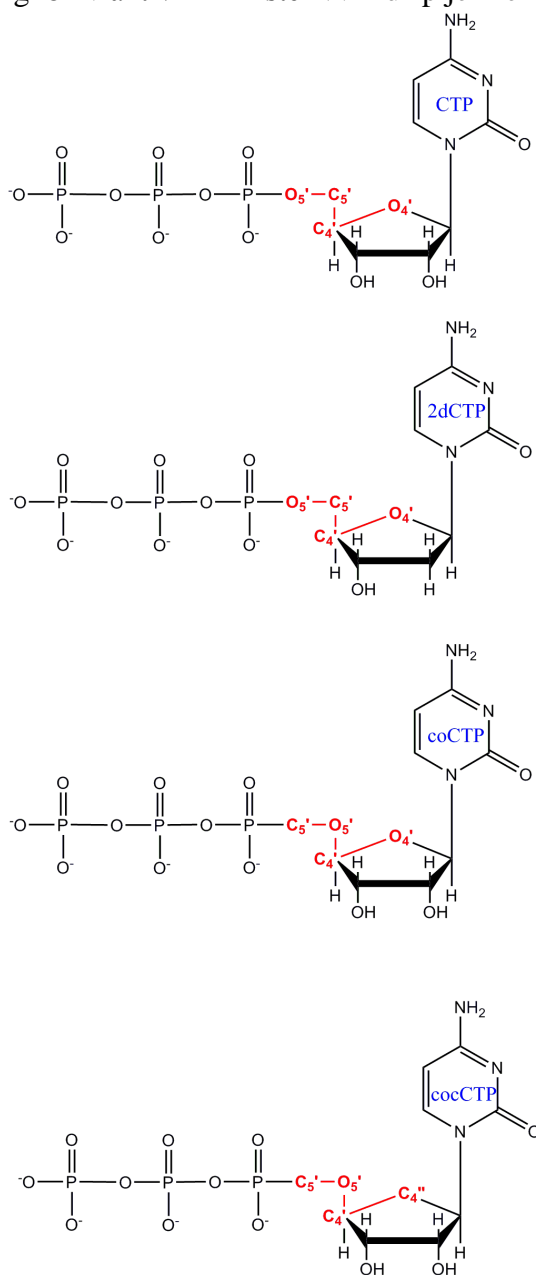
MD simulace jsou v současnosti již plně respektovanou metodou pro zkoumání biomolekul na atomární úrovni. Pomocí programovatelných grafických procesorů (GPU) lze dosáhnout délky MD trajektorií, která byla dříve dosažitelná pouze při použití velkých výpočetních klastrů. Díky dostatečnému výpočetnímu výkonu je tak nyní možné zkoumat i drobné rozdíly ve vazbě velmi podobných substrátů do aktivního místa RdRp. Mikrosekundové MD trajektorie jsou dostačující pro relaxaci simulovaného systému..

Zde byly provedeny srovnávací MD simulace CTP, 2dCTP, coCTP a cocCTP (obrázek 4.1.) v komplexu s NV RdRp (jejíž strukturní podobnost s HCV RdRp umožňuje přenositelnost mnoha pozorování). CTP a 2dCTP jsou nemodifikované nukleosid trifosfáty, v nichž je cukr a trifosfát spojen sekvencí atomů O5'-C5'-C4'-O4'. CoCTP (analogický k terminátoru coATP [135]) a cocCTP (analogický ke kompetitivnímu inhibitoru cocATP [135]) jsou modifikované nukleosid trifosfáty se sekvencemi atomů C5'-O5'-C4'-O4' (coCTP) a C5'-O5'-C4'-C4' (cocCTP). Dříve provedené MD/NMR studie dinukleotidů a oligonukleotidů s izosterní fosfonátovou spojkou zaznamenaly značné rozdíly v konformačních preferencích torzních úhlů O5'-C5'-C4'-O4' a C5'-O5'-C4'-O4' [137]. Proto byly i zde očekávány rozdíly v konformačních preferencích přirozených a modifikovaných CTP a z nich plynoucí dopady na geometrii aktivního místa NV RdRp.

Pro rozlišení 2'-deoxyribonukleotidů od ribonukleotidů slouží v NV RdRp síť vodíkových vazeb. Konzervovaná rezidua Asn309 a Ser300 se váží na 2'-OH skupinu přichozího NTP. Reziduum Ser300 interaguje také s reziduem Asp247. Rezidua Asn297, Ser288 a Asp238 u 3Dpol RdRp jsou homologní s rezidui Asn309, Ser300 a Asp247 u NV RdRp. U polymerázy 3Dpol RdRp s mutací postranního řetězce rezidua Asp238 je schopnost včlenění správného nukleotidu značně nižší, než je tomu u polymeráz s mutací postranního řetězce rezidua Asn297. Navíc 3Dpol RdRp, ve kterých byla rezidua Asp238 nebo Asn297 nahrazena alaninem, ztrácí schopnost rozlišit mezi ATP a 2'-dATP [138]. S ohledem na výše uvedená zjištění byly v MD trajektoriích detailně analyzovány interakce 2'-OH_CTP-Ser300-Asp247 (obrázek 4.2.a.).

V rentgenových strukturách NV RdRp [23] se báze CTP váže pomocí vodíkových vazeb ke komplementární bázi guanosinu v templátovém vlákne RNA. Báze CTP také stackuje mezi 3'-terminální bázi syntetizovaného vlákna RNA a reziduem Arg182. Reziduum Arg182 je přitom lokalizováno v blízkosti α -fosfátu CTP. Argininové reziduum v této pozici nalezneme i u jiných RdRp. U HCV RdRp je to Arg158. U mutantu HCV RdRp R158A jsou značně narušeny všechny jeho funkce,

pokud jako kofaktory slouží pouze hořčičkové ionty Mg^{2+} [139]. V postavení podobném jako má Arg182 v aktivním místě NV RdRp je i reziduum Arg72 v



Obrázek 4.1.: Struktura CTP, 2dCTP, coCTP a cocCTP [1].

aktivním místě HIV RT [140]. Také v DNA polymeráze β formuje reziduum Arg183 solný můstek s β -fosfátem přichozícího NTP [141-142]. Argininové „prsty“ hrají důležitou roli i v hydrolýze GTP prováděné tzv. GTPázami [143-145]. S ohledem na výše uvedené byly v MD trajektoriích detailně analyzovány interakce rezidua Arg186 z NV RdRp s trifosfátovými skupinami nukleosid trifosfátů.

Taktéž byla sledována vzdálenost ($\alpha P-O3'$) oddělující reakční substráty: α -fosfátovou skupinu NTP a $O3'$ -terminální atom primeru RNA. Nicméně je třeba mít na paměti, že chemický krok reakce katalyzované NV RdRp může být ovlivněn nahrazením fosfátu fosfonátem v modifikovaných nukleosidtrifosfonátech.

Simulovaný systém

V literatuře se nedávno objevily krystalové struktury s vysokým rozlišením pro komplexy NV RdRp, CTP a dvoušroubovice RNA:RNA [23, 130] (v proteinové databance PDB je lze nalézt pod kódovým označením 3BSN, 3BSO, 3H5X a 3H5Y). Zmíněné komplexy jsou zachyceny ve fázi těsně před transferem CMP části CTP na prodlužované vlákno RNA. Je zřejmé, že v aktivním místě NV RdRp [23, 130] se realizují vodíkové vazby spojující 2'-OH/3'-OH skupiny příchozího CTP s Asp247, Ser300 a Asn309. Analogickými rezidui u HCV RdRp jsou rezidua Asp225, Ser282 a Asn291. Jako výchozí struktura NV RdRp vhodná pro tvorbu simulovaných systémů byla vybrána struktura 3BSO [23].

Vstupní soubory pro MD simulace byly vytvořeny pomocí modulu TLEAP ze softwarového balíku AMBER [146]. K biomolekulám bylo v průběhu budování vstupních souborů *.inpcrd a *.prmtop přidáno cca. 15.565 molekul vody TIP3P [92]. Výsledný periodický box měl rozměry cca. 84 x 90 x 73 Å.

Silové pole

Chybějící parametry silového pole AMBER [87, 96] pro analoga CTP (obrázek 4.1.) byla doplněna standardním způsobem prostřednictvím *ab initio* výpočtů. Náboje atomů byly vypočítány metodou RESP (HF/6-31G*) [87]. Silové konstanty pro torzní úhly byly nařizovány tak, aby bylo dosaženo max. odchylky cca. 1kcal/mol mezi QM (MP2/6-31G*) a MM relativními energiemi jednotlivých konformerů CH₃-O-CH₂-P-O-CH₃.

Pro reprezentaci hořčikových iontů byl zvolen sedmibodový Åqvist-Warshelův (AW) model, který je založen na použití fiktivních pseudoatomů, na které se přerozdělí kladný náboj iontu [147]. AW model se osvědčil pro reprezentaci oktaedrálně koordinovaných divalentních iontů hořčíku, nebo tetraedrálně koordinovaných divalentních iontů zinku v mononukleárních a binukleárních aktivních místech metaloenzymů. Sedmibodový AW model má vždy lepší výsledky než tradiční jednoatomový model [147].

Metoda MD simulací

Ekvilibrační MD trajektorie o délce 5 ns byly získány prostřednictvím softwarového balíku NAMD 2.7 [103] a multiprocesorového systému osazeného šestnácti CPU typu Intel Itanium. Langevinova dynamika byla použita ke kontrole teploty při ohřívání simulovaných systémů z 0 na 310 K. Metoda tzv. Langevinova pístu byla použita pro udržení tlaku na úrovni 1 atm. Délka časového kroku byla nastavena na 2 fs [103].

Pro produkční MD běhy byl použit softwarový balík ACEMD v.2591 [108, 148], spouštěný na lokálních stanicích osazených GPU typu NVIDIA. ACEMD nepodporuje NPT simulace, proto byl použit soubor NVT. Po dostatečně dlouhých ekvilibračních MD bězích se objem velkých simulovaných systémů mění v NPT simulacích již jen zcela nepatrně. Proto je použití NVT souborů pro produkční MD simulace dobrou aproximací [108, 148]. Langevinovým termostatem byla udržována teplota na 310 K. Cutoff u nevazebných dalekodosahových sil byl nastaven na 9 Å.

Pro produkční MD běhy byly použity programovatelné GPU typu NVIDIA GTX-580 osazené 480 jádry. Časový krok byl nastaven na 3 fs. Délka MD trajektorií dosáhla 1050 ns.

Data byla ukládána každých 75 ps. Trajektorie byly následně analyzovány pomocí softwarových balíků CHIMERA 1.5.3, VMD 1.9, Curves+ a modulem PTRAJ ze softwarového balíku AMBER [125-127, 147]. K tvorbě obrázků byl použit programový balík ICM Molsoft 3.7.

Výsledky MD simulací

Celkem byly provedeny čtyři MD simulace. První dva MD běhy byly provedeny pro komplexy NV RdRp s přirozenými substráty CTP a 2dCTP (obrázek 4.1.). Ve dvou zbylých MD simulacích se jednalo o komplexy NV RdRp s modifikovanými trifosfáty coCTP (jehož analog coATP je u HCV RdRp terminátorem řetězce) a cocCTP (jehož analog cocATP je u HCV RdRp kompetitivním inhibitorem [135]).

RMSD

Stabilita simulovaných systémů v extrémně dlouhých mikrosekundových MD simulacích byla první otázkou, na kterou bylo nezbytné odpovědět. Nestability mohou nastat z mnoha různých důvodů. Významné konformační změny mohou doprovázet proces relaxace biomolekul, které jsou v krystalech deformovány interakcemi mezi svými identickými kopiemi. Tzv. uzavřený stav aktivního místa NV RdRp vznikl jako důsledek jeho interakce s CTP během krystalizace [23]. Nicméně analoga CTP mohou stimulovat přechod NV RdRp do otevřeného stavu. V neposlední řadě mohou nestability pocházet z artefaktů v současných silových polích [96], které se v dosavadních kratších MD trajektoriích nemusely dosud projevit. Proto byla stabilita simulovaných biomolekul analyzována pomocí tzv. střední kvadratické odchylky (RMSD) poloh jejich atomů vzhledem k výchozímu uspořádání. Časový vývoj RMSD prokázal stabilitu simulovaných systémů po celou dobu trvání mikrosekundových MD simulací.

Vzdálenost mezi ionty Mg_1^{2+} a Mg_2^{2+} v aktivním místě

Při standardní reprezentaci iontů jako atomů s bodovým nábojem +2 vznikají v průběhu MD simulací binukleárních aktivních míst metaloenzymů nestability v důsledku nepřiměřených odpuzivých sil. Tomuto artefaktu lze zamezit použitím sedmibodové AW reprezentace pro hořčikové ionty [147]. Nikdo (pokud je autorovy této práce známo) však neproověřil tuto metodu prostřednictvím mikrosekundových MD simulací. Vzdálenosti Mg_1^{2+} - Mg_2^{2+} a $Mg_{1,2}^{2+}$ -ligand ve zde prezentovaných MD simulacích byly stabilní. Nebyla tak zaznamenána ani snaha aktivního místa NV RdRp o přechod do otevřené konfigurace, která by započala rozpadem interakcí mezi iontem Mg_2^{2+} a reziduem Asp247.

Strukturní analýza

CTP

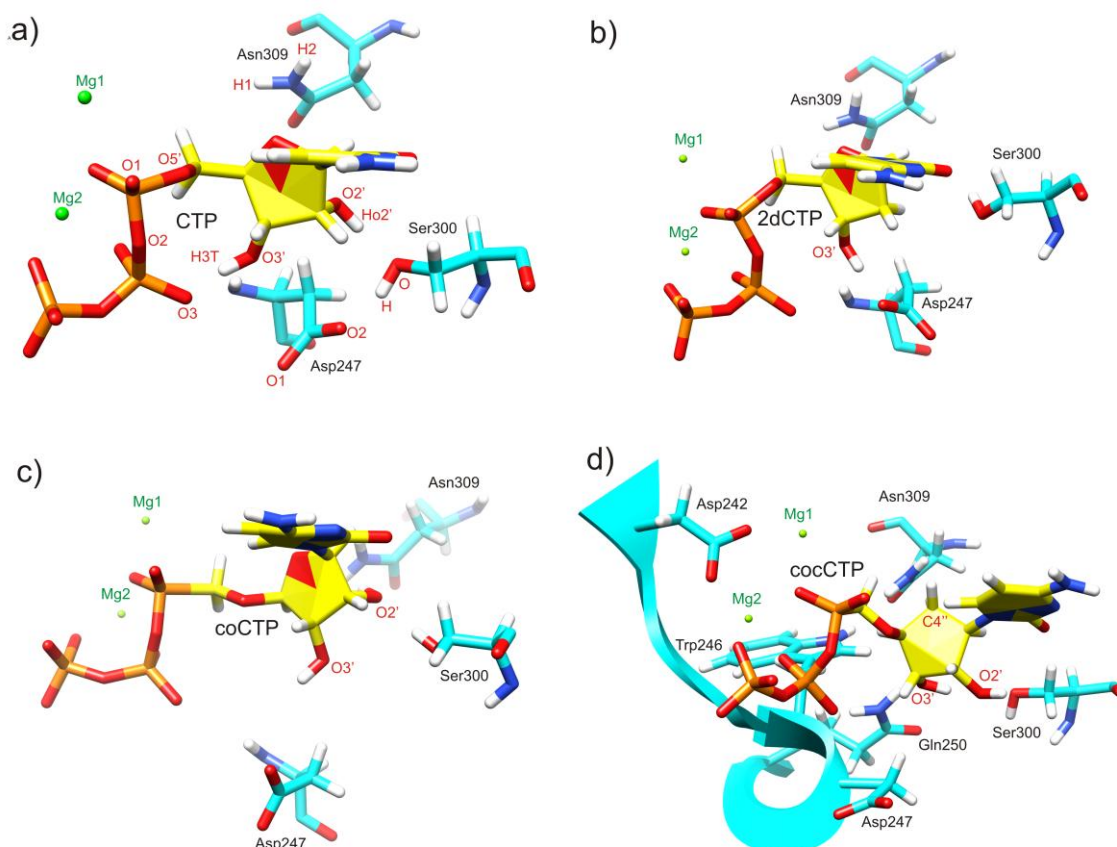
Po většinu MD simulace setrval CTP (obrázek 4.1.) v konformaci A-RNA. Torzní úhel $O5'-C5'-C4'-O4'$ preferoval konformaci **-gauche** a cukr CTP byl v konformaci **C3'-endo** s krátkými reverzibilními přechody do konformace **C2'-endo** (obrázek 4.4. graf A02.). Na konci MD běhu (v okamžiku reverzibilního narušení kontaktů mezi postranními řetězci aminokyselin a CTP), přešel torzní úhel $O5'-C5'-C4'-O4'$ do konformace **+gauche** (obrázek 4.4. graf A01.) a puckering cukru do konformace **C3'-exo** (obrázek 4.4. graf A02.). V krystalových strukturách [23] je 2'-OH skupina CTP navázána na Ser300, které je dále stabilizováno interakcí s Asp247. MD simulace naznačila, že ve vodném roztoku je síť vodíkových vazeb mezi těmito partnery komplikovanější. Uplatnily se totiž veškeré možné kombinace akceptorů a donorů vodíkových vazeb (obrázek 4.2.a.). Reziduum Arg182 po většinu MD běhu interagovalo s α -fosfátovou skupinou CTP (konkrétně s atomem O1 (obrázek 4.4. graf A08.)) (obrázek 4.3.a.). V této pozici bylo reziduum Arg182 stabilizováno prostřednictvím rezidua Glu168. Vzhledem k tomu, že je v průběhu reakce rozštěpena vazba mezi atomy O2 a αP , může výše popsaná interakce rezidua Arg182 s α -fosfátovou skupinou CTP reakci spíše brzdit. Zhruba jednou za 100ns přerušil postranní řetězec Arg182 kontakt s reziduem Glu168 a interagoval pouze s pyrofosfátovou skupinou CTP. Konkrétně interagoval s atomem O3 z β -fosfátové skupiny a s kyslíkem O2 (obrázek 4.4. graf A09.) spojujícím α - a β -fosfátovou skupinu CTP (obrázek 4.3.b.). Reziduum Arg182 tak v této konfiguraci interagovalo přímo s atomem kyslíku ze štěpené vazby, což by mělo přinést snížení energie přechodového stavu reakce. Vzdálenost oddělující reakční substráty, čili fosfor α -fosfátové skupiny CTP a O3' atom z konce primeru RNA, klesla dvakrát pod hranici 2.9 Å a osmnáctkrát pod 2.95 Å (obrázek 4.4. graf A10.).

2dCTP

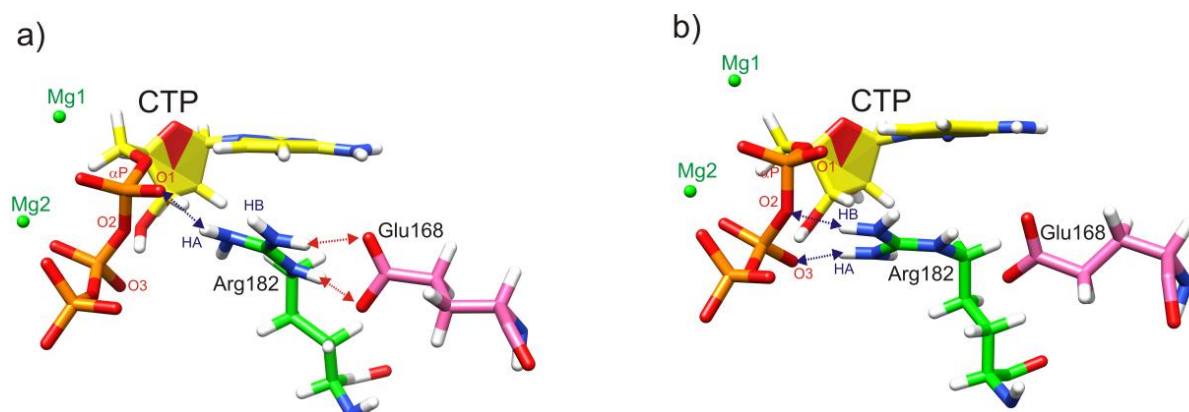
2dCTP (obrázek 4.1.) byl podstatnou část MD běhu ve stejné konformaci jako CTP. Podstatně se však lišila interakce 2dCTP s postranními řetězci aminokyselin aktivního místa NV RdRp. Do MD simulace vstoupilo 2dCTP v atypické A-RNA konfiguraci s deoxyribózou v konformaci **C3'-endo**, která převládla i na konci MD simulace. Zhruba po 250 ns se sice prosadila konformace **C2'-endo**, celkově však deoxyribóza inklinovala více ke konformaci **C3'-endo**. To zřejmě způsobovala intramolekulární vodíková vazba poutající 3'-OH a β -fosfátovou skupinu 2dCTP (zaznamenaná i u CTP). Torzní úhel $O5'-C5'-C4'-O4'$ byl v konformaci **-gauche** (stejně jako u CTP). Souběžně s repuckeringem deoxyribózy do konformace **C2'-endo** došlo k přerušení interakcí mezi rezidui Asp247 a Ser300 (obrázek 4.2.b.). Jejich kontakty již nebyly obnoveny patrně i díky absenci 2'-OH skupiny u 2dCTP, která by v interakcích s Asp247 a Ser300 participovala. Solný můstek mezi reziduem Arg182 a α -fosfátovou skupinou 2dCTP byl v důsledku konformačních změn deoxyribózy 2dCTP též přerušen. Pouze třikrát byla zaznamenána alternativní konfigurace s reziduem Arg182 navázaným na kyslík štípané vazby O2- αP . Horší výsledky než ty, které byly zaznamenány u CTP, byly pozorovány i pro vzdálenost oddělující reakční substráty. Vzdálenost mezi fosforem α -fosfátové skupiny 2dCTP a atomem O3' konce RNA primeru nikdy neklesla pod 2.9 Å a pouze sedmkrát spadla pod 2.95 Å.

coCTP

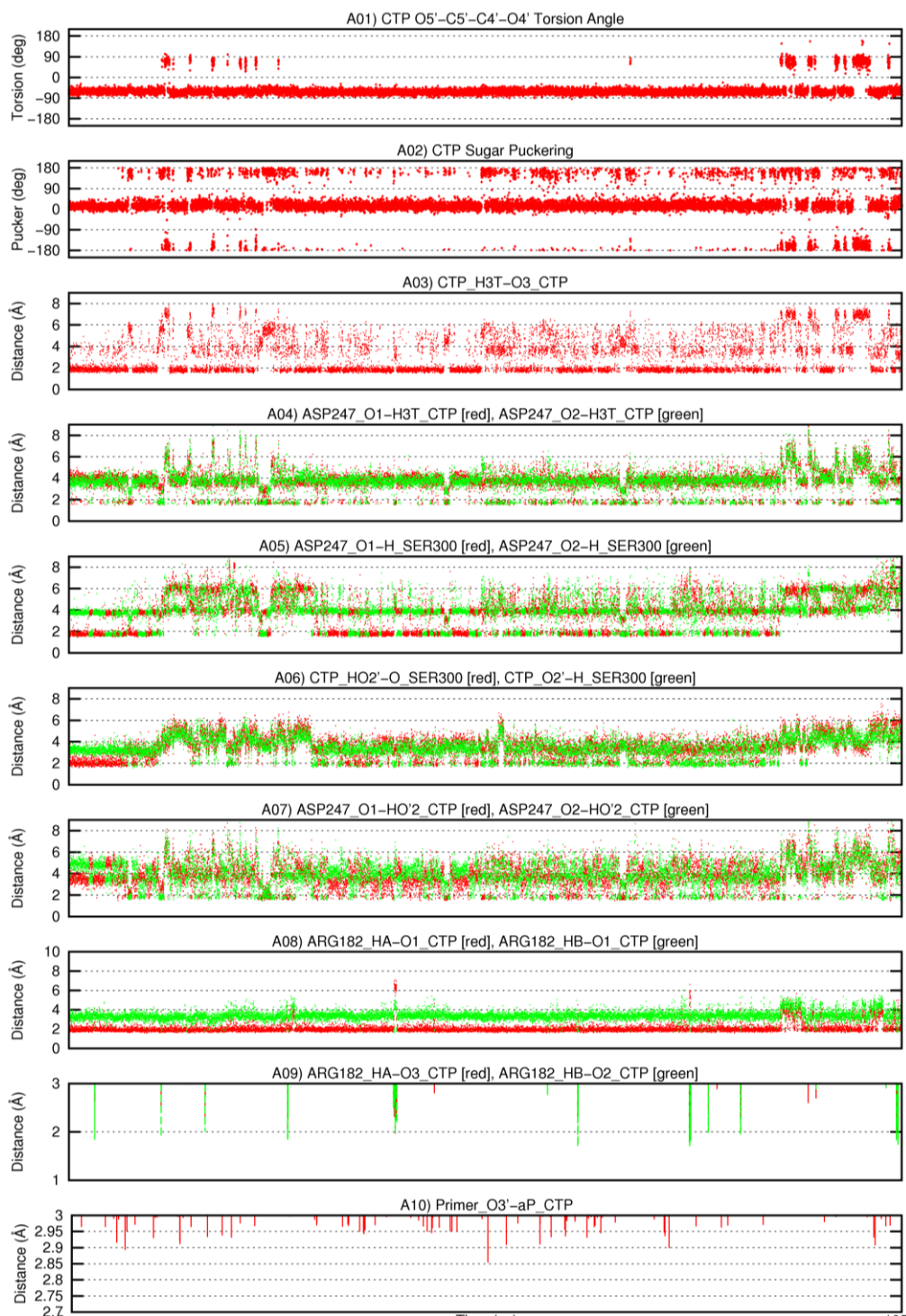
Vzhledem k jednoznačné preferenci atypické konformace **+gauche** torzním úhlem **C5'-O5'-C4'-O4'** byla po umístění coCTP (obrázek 4.1.) do aktivního místa NV RdRp očekávána určitá relaxace simulovaného systému. Překvapením bylo, že trvala několik stovek nanosekund. Ribóza coCTP byla vychýlena z konformace **C3'-endo** do konformace **C2'-exo**. Na počátku MD simulace přitom prošla řadou konformačních změn. Následně začal torzní úhel **C5'-O5'-C4'-O4'** fluktuovat mezi konformacemi



Obrázek 4.2.: CTP, 2dCTP, coCTP a cocCTP v aktivním místě NV RdRp [1].



Obrázek 4.3.: a) Arg182 interagující s α -fosfátovou skupinou CTP; b) Arg182 interagující s β -fosfátovou skupinou CTP [1].



Obrázek 4.4.: Konformační preference CTP a jeho interakce v aktivním místě NV RdRp: A01) CTP O5'-C5'-C4'-O4' torzní úhel; A02) CTP puckerung cukru; A03) Intramolekulární vodíkové vazby: CTP_3'OH - β-fosfát_CTP; A04) Asp247 - 3'OH_CTP; A05) Asp247 - Ser300; A06) Ser300 - 2'OH_CTP; A07) Asp247 - 2'OH_CTP; A08) Arg182 - α-fosfát_CTP; A09) Arg182 - β-fosfát_CTP; A10) Primer_O3'-α-fosfát_CTP; [1]

+gauche a **trans**. Pominutí těchto konformačních přechodů bylo spojeno s rekonfigurací vodíkových vazeb mezi coCTP a postranními řetězci aminokyselin v aktivním místě NV RdRp, která byla dokonána až koncem druhé třetiny MD simulace. Interakce mezi postranními řetězci Asp247 a Ser300 ustaly po cca. 300 ns. Poté Ser300 sloužilo jako donor v pozoruhodně stabilní vodíkové vazbě s 2'-OH skupinou coCTP (obrázek 4.2.c.). Předtím zde Ser300 figurovalo jako akceptor. Byl zaznamenán pouze jeden konformační přechod Arg182 ke kyslíku štěpené vazby O2- α P. Jinak bylo Arg182 po celou dobu MD simulace v kontaktu s α -fosfátovou skupinou coCTP. Vzdálenost mezi fosforem α -fosfátové skupiny coCTP a atomem O3' konce RNA primeru klesala pod hranici 2.9 a 2.95 Å s lehce vyšší frekvencí než tomu bylo u následující cocCTP MD simulace. Konkrétně sedmkrát pod hranici 2.9 Å a více jak čtyřicetkrát pod 2.95 Å

cocCTP

Relaxace cocCTP (obrázek 4.1.) v aktivním místě NV RdRp trvala ještě déle, než tomu bylo u coCTP. Byla dokončena až po repuckeringu modifikovaného cukru (tj. po cca. 500 ns). Torzní úhel **C5'-O5'-C4'-C4'** se většinou nacházel v konformaci **+gauche** (podobně jako v předchozí coCTP MD simulaci). Prvních 500 ns preferoval modifikovaný cukr cocCTP výhradně konformaci **C2'-exo**. Po tu dobu byly interakce cocCTP_2'-OH-Ser300-Asp247 mimořádně stálé. Interakce mezi Ser300 a Asp247 byla podstatně stabilnější než v ostatních MD simulacích. Částečně to bylo způsobeno interakcemi mezi 3'-OH skupinou cocCTP a reziduem Asp247. V MD simulacích s CTP a 2dCTP tato 3'-OH skupina participovala spíše v intramolekulárních interakcích s trifosfátovou skupinou. Veškeré interakce mezi rezidui Ser300 a Asp247 však byly přerušeny po přechodu modifikované ribózy do konformace **C3'-exo**. Modifikovaný cukr pak byl v kontaktu s postranním řetězcem rezidua Trp246. 3'-OH skupina cocCTP příležitostně vytvářela vodíkové vazby na rezidua Gln250 a Asn309. Celkově cocCTP pronikl o něco hlouběji do aktivního místa NV RdRp. Přesněji řečeno, až za reziduum Asn247 z motivu A (obrázek 4.2.d.). Penetrace je zřejmá z nárůstu vzdálenosti mezi atomem C2' 3'-terminálního rezidua primeru RNA a atomem C1' z cocCTP. Tato vzdálenost vzrostla kvůli příliš těsnému počátečnímu kontaktu mezi -CH₂- skupinou modifikovaného cukru a vodíkovým atomem H2' navázaným na C2' atom 3-terminálního nukleotidu RNA primeru. Výsledky výpočtů s cocCMP již zabudovaným do syntetizovaného vlákna RNA ukazují, že právě výše popsané sterické konflikty předurčují cocCTP k funkci kompetitivního inhibitoru (na rozdíl od terminátoru řetězce coCTP). Reziduum Arg182 interagovalo po celou dobu MD simulace s α -fosfátovou skupinou cocCTP a bylo v této pozici stabilizováno reziduem Glu168. Vzdálenost mezi reakčními substráty - fosforem α -fosfátové skupiny cocCTP a atomem O3' z 3'-konce primeru RNA klesla pětkrát pod 2.9 Å a třicetšestkrát pod hranici 2.95 Å.

Shrnutí

Byly provedeny mikrosekundové MD simulace komplexů NV RdRp s přirozenými (CTP, 2dCTP) a modifikovanými (coCTP, cocCTP) substráty (obrázek 4.1.). Na základě dříve prováděných studií [132-133, 137] s podobně modifikovanými di- a oligonukleotidy byly v prezentovaných MD simulacích očekávány rozdílné konformační preference modifikovaných substrátů. Zatímco totiž v nemodifikovaných dinukleotidech preferoval torzní úhel O5'-C5'-C4'-O4'

konformaci **-gauche**, torzní úhel $C5'-O5'-C4'-O4'$ v jejich fosfonátových analogích preferoval konformaci **+gauche** [137]. Ke stejným výsledkům dospěly zde provedené MD simulace komplexů NV RdRp s trifosfáty. V CTP a 2dCTP byl atom kyslíku $O5'$ situován nad ribózou, kdežto v coCTP a cocCTP směřoval atom $C5'$ mimo ribózu (obrázek 4.2.). Dále byly sledovány interakce xCTP s postranními řetězci aminokyselin, jejichž mutace snižuje katalytickou efektivitu RdRp. Příklady takových aminokyselin jsou Arg182 [139] interagující s trifosfátovou skupinou xCTP nebo Asp247 [138, 149], které je spolu se Ser300 a Asn309 zapojeno do mechanismu rozpoznávání $2'-OH$ skupin NTP.

Ribóza CTP preferovala konformaci **C3'-endo** (obrázek 4.4. graf A02.) a torzním úhlem $O5'-C5'-C4'-O4'$ byl preferován konformer **-gauche** (obrázek 4.4. graf A01.). Síť vodíkových vazeb mezi $2'-OH$ skupinou CTP, reziduem Ser300 a reziduem Asp247 byla vitální po celou časovou evoluci simulovaného systému. V této síti vodíkových vazeb se postupně uplatňovaly všechny možné kombinace párů donor-akceptor. Zhruba každých 100 ns se změnila prostorová konfigurace postranního řetězce rezidua Arg182, který místo s α -fosfátovou začal interagovat s β -fosfátovou skupinou CTP.

Stejně jako u CTP, byl i u 2dCTP preferován torzním úhlem $O5'-C5'-C4'-O4'$ konformer **-gauche**. Deoxyribóza 2dCTP preferovala konformaci **C2'-endo**, byly však zaznamenány i četné reverzibilní přechody do konformace **C3'-endo**. Ty byly z části důsledkem intramolekulárních interakcí mezi $3'-OH$ a β fosfátovou skupinou 2dCTP. Interakce mezi rezidui Asp247 a Ser300 byly trvale přerušeny. $3'-OH$ skupina 2dCTP příležitostně interagovala s postranním řetězcem rezidua Asp247 (obrázek 4.2.b.). Navíc byl přerušen kontakt mezi postranním řetězcem rezidua Arg182 a trifosfátovou skupinou 2dCTP. Výše zmíněná fakta jasně ukazují, že se vazba nepatřičného substrátu 2dCTP do aktivního místa NV RdRp v mnohém liší od vazby správného substrátu CTP.

V MD simulaci coCTP preferoval torzní úhel $C5'-O5'-C4'-O4'$ rotamer **+gauche**. Ribóza byla v konformaci **C3'-endo/C2'-exo**. Poziční relaxace coCTP v aktivním místě NV RdRp trvala řádově stovky nanosekund. Po jejím ukončení interagovala $2'-OH$ skupina coCTP s rezidui Ser300 a Asn309 (obrázek 4.2.c.). Po celou dobu MD simulace bylo reziduum Arg182 propojeno solnými můstky s α -fosfonátovou skupinou coCTP.

Od začátku cocCTP MD simulace preferoval torzní úhel $C5'-O5'-C4'-C4''$ rotamer **+gauche**. Modifikovaná ribóza preferovala dvě konformace: **C2'-exo** a **C3'-exo**. Konformační přechod mezi těmito dvěma konformacemi nastal zhruba v polovině MD simulace. Pokud se modifikovaná ribóza nacházela v konformaci **C2'-exo**, byla síť vodíkových vazeb $2'-OH_{cocCTP}$ -Ser300-Asp247 velmi stabilní. Důvodem vysoké stability bylo participace $3'-OH$ skupiny cocCTP v této síti vodíkových vazeb. Po repuckeringu modifikované ribózy do konformace **C3'-exo** pronikl cocCTP hlouběji do aktivního místa NV RdRp (až za reziduum Asn247 z motivu A (obrázek 4.2.d.)). Modifikovaný cukr pak interagoval s postranním řetězcem rezidua Trp246. Byly také zaznamenány příležitostné vodíkové vazby mezi rezidui Gln250, Asn309 a $3'-OH$ skupinou cocCTP. Příčinou této markantní konformační a poziční změny cocCTP byly příliš těsné počáteční kontakty mezi atomem $H2'$ z $3'$ -terminálního nukleotidu primeru RNA a skupinou $-CH_2-$ modifikovaného cukru

cocCTP. Výpočty provedené s cocCMP již zabudovaným do syntetizovaného vlákna RNA ukázaly, že právě tyto sterické konflikty jsou důvodem proč je cocCTP kompetitivním inhibítozem NV RdRp (na rozdíl od coCTP, který je terminátorem RNA řetězce).

Torzní úhly **C5'-O5'-C4'-O4'** a **C5'-O5'-C4'-C4'** v modifikovaných strukturách coCTP a cocCTP tedy jednoznačně preferovaly konformaci **+gauche** v souladu s tím, co bylo zaznamenáno ve studiích podobně modifikovaných systémů prováděných prostřednictvím NMR [137] či rentgenové krystalografie [150].

Poziční relaxace xCTP v aktivním místě NV RdRp trvala řádově stovky nanosekund. Z tohoto úhlu pohledu se jedna mikrosekunda jeví jako minimální délka MD trajektorií nezbytná pro dostatečnou evoluci komplexů RdRp a potenciálních inhibitorů.

U přirozených substrátů - CTP a 2dCTP - působí Arg182 jako katalytické reziduum [139]. Otázka, kterou mohou zodpovědět pouze budoucí QM/MM výpočty [151-153] je, zda intenzivní kontakt mezi α -fosfonátovou skupinou a reziduem Arg182 zaznamenaný u coCTP a cocCTP reakci spíše nebrzdí.

Celkově lze konstatovat, že velmi podobné nukleosidtrifosfáty se v MD simulacích do aktivního místa NV RdRp vážou velmi rozdílným způsobem. Což plně odpovídá tomu, že CTP je přirozený substrát, 2dCTP chabý substrát, coCTP terminátor RNA řetězce a cocCTP kompetitivní inhibitor RdRp.

4.2. Molekulárně dynamické simulace komplexů lidské RNázy H s fosfonátovými analogy oligonukleotidů

Motivace

V posledních dvaceti letech jsou intenzivně zkoumány možnosti využití tzv. antisense deoxyoligonukleotidů (AO) k selektivnímu zabránění exprese genetické informace z konkrétních genů [42, 49]. Například tzv. Fomivirsen/Vitavene již prošel klinickými testy a používá se jako léčivo.

Dvoušroubovicovité hybridní duplexy AO:mRNA mohou zamezit translaci genetické informace z mRNA prostým zablokováním ribosomů. Ty jsou však někdy schopné duplex rozvinout a v přepisu genetické informace pokračovat. V živých organismech se navíc nachází řada dalších enzymů (helikáz), které také jsou schopné vlákna hybridního duplexu AO:mRNA oddělit [154-161]. Proto je vhodné využít mechanismus, který cílovou mRNA rozloží a trvale tím zamezí translaci genetické informace z ní. Enzym RNáza H specificky štěpí pouze vlákno RNA v hybridních duplexech DNA:RNA. AO tak po degradaci mRNA zůstává neporušen. Může tedy vytvořit komplex s další kopií mRNA a opět vyvolat její degradaci prostřednictvím RNázy H [160].

Přirozené deoxyoligonukleotidy nelze jako AO použít pro jejich nízkou odolnost vůči buněčným nukleázám [50]. Tu lze zvýšit prostřednictvím chemických modifikací ve struktuře AO. Může jít o modifikace cukr-fosfátové páteře i bazí [50]. Zdaleka ne všechny modifikované AO jsou však schopné po navázání k mRNA aktivovat RNázu H. Z experimentu je známo [162], že AO, v nichž alternují fosfonátové internukleotidové spojky typu 3'-O-P-CH₂-O-5' (3pc5) s klasickými fosfodiesterovými spojkami 3'-O-P-O-5', jsou schopné aktivitu RNázy H stimulovat. Naproti tomu AO, v nichž alternují fosfonátové internukleotidové spojky typu 3'-O-CH₂-P-O-5' (3cp5) s klasickými fosfodiesterovými spojkami 3'-O-P-O-5', aktivní nejsou.

Prostřednictvím MD simulací [52] byly proto studovány 3pc5 a 3cp5 deoxyoligonukleotidy v komplexu s vlákny RNA a katalytickou doménou lidské verze (HS) enzymu RNáza H [44]. Díky použití grafických procesorů (graphics processing unit - GPU) byly získány trajektorie o délce 200 ns.

Stability binukleárního aktivního místa RNázy H bylo v těchto MD simulacích (obrázek 4.6., obrázek 4.7.) dosaženo díky modifikovanému jednoatomovému modelu iontu Mg²⁺ a polarizaci fosfátové skupiny štípané internukleotidové spojky ve vlákne RNA. Tato internukleotidová spojka je v aktivním místě RNázy H zaklesnuta mezi oba hořčíkové ionty. Adekvátnost použitého přístupu byla ověřena srovnáním s výsledky referenčního MD běhu o délce 1000 ns, ve kterém byla použita sedmibodová reprezentace pro oba hořčíkové ionty Mg²⁺ [147] (obrázek 4.5.).

Simulované systémy

Základem pro vytvoření simulovaných systémů - komplexů HS RNázy H s hybridními duplexy DNA:RNA - byla krystalová struktura 2QKK [44] z proteinové databanky PDB. Z té byly nejprve odstraněny páry nukleotidů, které se nepodílejí na

vazbě hybridního duplexu k enzymu. Na 3'-konci vlákna DNA byly modifikovány tři internukleotidové spojky oběma výše zmíněnými typy modifikací 3pc5 a 3cp5. Takto vznikly dva komplexy: HS RNáza H + RNA : 3cp5-DNA (3'-G-3cp5-C-3cp5-T-3cp5-G-T-G-G-A-5') a HS RNáza H + RNA : 3pc5-DNA (3'-G-3pc5-C-3pc5-T-3pc5-G-T-G-G-A-5'). Na základě takto upravených .pdb souborů byly vytvořeny vstupní soubory pro celý simulovaný systém pomocí modulu TLEAP ze softwarového balíku AMBER [146]. Konkrétně se jednalo o soubory *.inpcrd (počáteční souřadnice) a *.prmtop (silová pole, molekulární topologie atd.). V průběhu budování vstupních souborů byly k modelovaným strukturám přidány molekuly vody typu TIP3P [92]. Přidání molekul vody vedlo k rozšíření simulovaného systému o cca. 10 Å v každém směru. Výsledný periodický box měl rozměry cca. 82 Å x 81 Å x 61 Å. Simulovaný systém se skládal z 39.577 atomů.

Metoda MD simulací

Pomocí softwarového balíku NAMD 2.7 [103] byly provedeny ekvilibrační MD běhy v délce 5 ns. Pro ně byl použit izotermicko-izobarický soubor (NPT) a časový krok 2 fs. Použitím Langevinovy dynamiky pro kontrolu teploty byl simulovaný systém ohřát z 0 na 310 K. Metoda tzv. Langevinova pístu byla použita pro kontrolu tlaku nastaveného na 1 atmosféru [103]. Cut-off nevazebných sil byl nastaven na 9 Å. Pro výpočet dalekodosahové elektrostatických sil byla použita smooth Particle-Mesh Ewaldova (PME) sumace. Algoritmus SHAKE byl aplikován na vazby s atomy vodíku. Ekvilibrační MD běhy byly realizovány prostřednictvím multiprocessorového systému osazeného šestnácti procesory Intel Itanium s kmitočtem 1,5 GHz. Výpočet probíhal rychlostí 4 ns/den.

Dlouhé produkční MD simulace byly realizovány prostřednictvím softwarového balíku ACEMD v. 2591 [108, 148] a grafických procesorů firmy NVIDIA. Pro tyto MD simulace byl použit soubor NVT, neboť ACEMD použití NPT neumožňuje. To není nikterak na škodu, neboť po dostatečně dlouhých ekvilibračních MD bězích se objem velkých simulovaných systémů mění v NPT simulacích jen zcela nepatrně [108, 148]. Časový krok byl nastaven na 4 fs díky tzv. "mass repartitioning scheme", které umožňuje přerozdělit část hmotnosti těžkých atomů na vodíky na ně navázané. Jinak byly použity metody a parametry shodné s těmi, které byly použity v ekvilibračních bězích. Pro realizaci produkčních MD simulací byly použity programovatelné GPU (NVIDIA GTX-580) osazené 480 jádry. Rychlost výpočtu byla cca. 50 ns/den. Celková délka trajektorií byla 1400 ns.

Snímky z MD simulací byly ukládány každých 100 ps. Následně byly analyzovány pomocí softwarových balíků CHIMERA 1.5.3, VMD 1.9, Curves+ a PTRAJ/AMBER10 [125-127, 147]. Obrázky byly vytvořeny prostřednictvím softwarového balíku ICM Molsoft 3.7.

Doplnění silového pole - fosfonátové internukleotidové spojky

Standardní silové pole AMBER neobsahuje všechny silové konstanty nezbytné pro MD simulace 3cp5/3pc5 fosfonátových analog oligonukleotidů. Bylo potřeba je získat pomocí *ab initio* QM výpočtů způsobem konzistentním s metodikou použitou při vývoji standardního silového pole AMBER [87, 96]. Fit silových konstant byl proveden s cílem dosáhnout odchylky do 1 kcal/mol mezi QM (MP2/6-31G*) a MM

hodnotami relativních energií pro jednotlivé konformery modelového systému CH₃-O-PO₂-CH₂-O-CH₃.

Tradiční jednobodová reprezentace hořčkových iontů

V přípravných MD simulacích se standardně parametrizovanými hořčkovými ionty se v binukleárním aktivním místě RNázy H objevovaly nestability. Podobná pozorování lze nalézt i v literatuře (např. v aktivním místě polymerázy- β [147]). Problémy pramení z pomnutí polarizace a podhodnocení VdW poloměru hořčkového iontu.

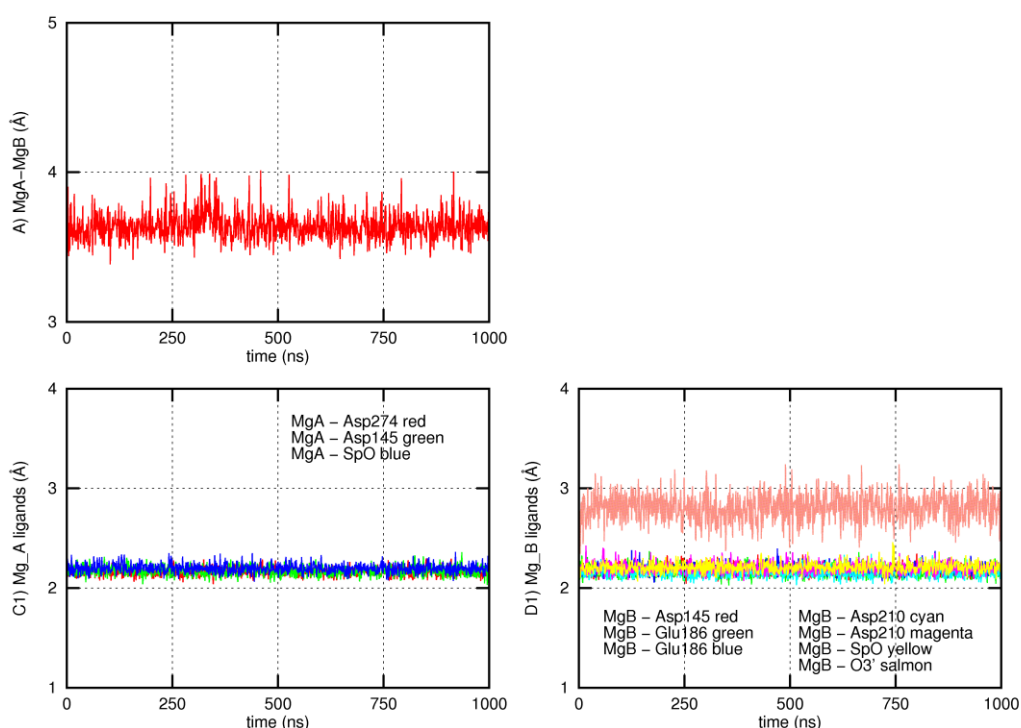
V MD simulacích interagují divalentní hořčkové ionty s biomolekulami prostřednictvím nevazebných interakcí. Jsou reprezentovány jako van der Waalsovské kuličky ($R^* = 0.7926 \text{ \AA}$, $\epsilon = 0.8947$) s bodovým nábojem +2. Takto vysoké náboje však při použití nepolarizovatelných silových polí zapříčiňují v binukleárních aktivních místech uměle chování způsobené nadměrnými odpudivými silami vyvstávajícími mezi oběma Mg²⁺ ionty [147, 163-164].

Analýza krystalových struktur uložených v proteinové databance PDB ukazuje, že ionty hořčíku jsou invariantně hexakoordinovány s oktahedrání geometrií a délka vazby Mg-O kolísá v rozpětí 2.05-2.25 \AA [147]. V MD simulacích se standardní parametrizaci iontů je však tato vzdálenost obvykle podhodnocena. Modelovaná aktivní místa enzymů jsou v důsledku toho příliš kompaktní [165-166]. Standardní parametrizace pro ionty byla získána Åqvistem na základě výpočtů hydratační volné energie pomocí metody FEP. Podhodnocení VdW poloměru u hořčíku (nejmenšího ze zkoumaných iontů) bylo v této průkopnické práci reflektováno [166]. I při vyšší hodnotě VdW poloměru ($R^* = 1.3 \text{ \AA}$, $\epsilon = 0.06$) použité v rámci tradičního jednoatomového modelu pro Mg²⁺ však nesoulad mezi modelovanými a krystalovými strukturami binukleárních aktivních míst přetrvává (např. zvětšená vzdálenost mezi oběma hořčíky [147]).

Åqvist-Warshelova sedmibodová reprezentace hořčkových iontů

V tzv. Åqvist-Warshelově (AW) sedmibodovém modelu jsou okolo centrálního iontu umístěny fiktivní pseudoatomy. Kladný náboj centrálního iontu je na ně přerozdělen. Výsledky MD simulací používajících AW model pro Mg²⁺ jsou v podstatně lepší shodě s realitou, než MD simulace využívající tradiční jednoatomový model [147].

Zde byl AW model [147] testován v mikrosekundové MD simulaci HS RNázy H v komplexu s nemodifikovaným hybridním duplexem DNA:RNA. Binukleární aktivní místo RNázy H bylo po celou dobu MD simulace stabilní (obrázek 4.5.). Nicméně bylo zřejmé, že rezidua Asp210 a Glu186 mají tendenci tvořit bidentátní kontakt s MgB, což vede k jeho překoordinování.



Obrázek 4.5.: Kontakty v aktivním místě RNázy H v MD simulaci se sedmibodovou reprezentací hořčíkových iontů - A) vzdálenost MgA-MgB (Graph A); C) vzdálenost MgA - ligand (Graf C1: Asp274 červená, Asp145 zelená, Sp modrá; D) vzdálenost MgB - ligand (Graf D1: Asp145 červená, Glu186 zelená, Glu186 modrá, Asp210 bleděmodrá, Asp210 fialová, Sp žlutá, O3' lososová) [2].

Kromě toho je budování simulovaného systému při použití sedmibodového AW modelu [147] poměrně zdlouhavé, což může být na překážku, např. pokud je cílem provést sérii MD simulací s využitím různých krystalových struktur daného enzymu. Proto byla výše zmíněná MD simulace využívající AW model použita jako referenční při rafinaci nového jednobodového modelu pro hořčíkové ionty umístěné v binukleárním aktivním místě RNázy H.

Nová jednobodová reprezentace hořčíkových iontů

Jednoatomový model iontu Mg^{2+} byl precizován tak, aby výsledky MD simulací (obrázek 4.6., obrázek 4.7.) byly ve shodě s výsledky přípravné MD simulace získané prostřednictvím sedmibodového AW modelu (obrázek 4.5.). Bylo vzato do úvahy, že zdrojem problémů je zanedbání polarizace a podhodnocení VdW poloměru

hořčikového iontu [165-166]. Proto byly parametry pro hořčikový iont změněny z ($R^* = 0.7926 \text{ \AA}$, $\epsilon = 0.8947$) na ($R^* = 1.1 \text{ \AA}$, $\epsilon = 0.8947$). Modifikovány byly ale také parciální náboje ($z = -0.7761/-0.7761$ na $-0.9761/-0.5761$) postranních kyslíků pro-Sp a pro-Rp v štípané internukleotidové spojení rA5-rC6, která se v aktivním místě RNázy H váže mezi oba hořčikové ionty.

MD simulace 3pc5 a 3cp5 - aktivní místo RNázy H

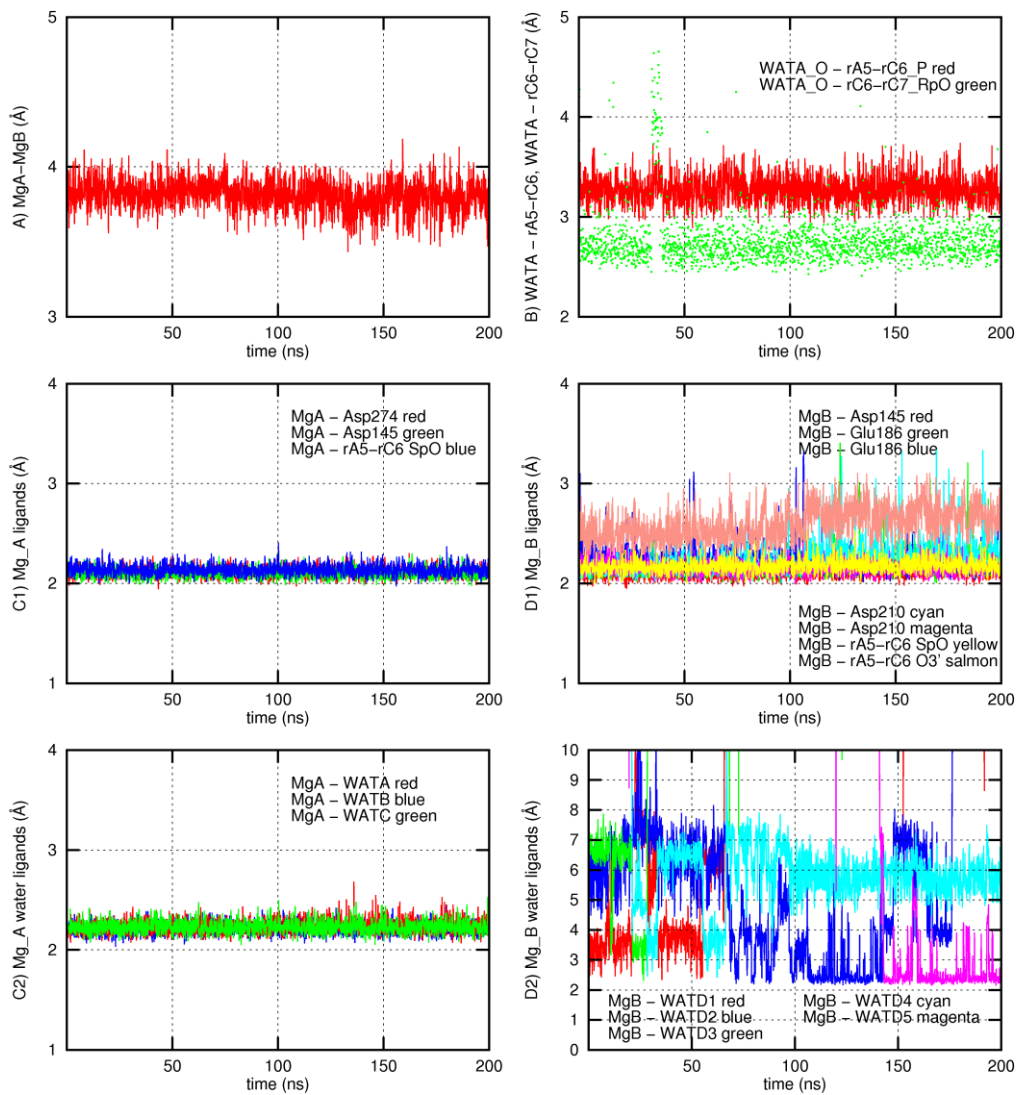
V hlavních MD simulacích provedených s využitím nové jednobodové reprezentace hořčikových iontů pak vzdálenost mezi ionty hořčíku Mg^{2+} v pozicích A a B zůstala konstantní po celých 200 ns (obrázek 4.6. graf A., obrázek 4.7. graf A.).

MgA byl v 3pc5 MD simulaci koordinován v oktahedrální konfiguraci tvořené rezidui Asp145, Asp274, postranním kyslíkem pro-Sp štípané spojky rA5-rC6 (obrázek 4.6. graf C1.) a třemi molekulami vody (obrázek 4.6. graf C2.). Kyslíkový atom molekuly vody WATA atakující štípanou internukleotidovou spojkou rA5-rC6 setrval po celou dobu 3pc5 MD simulace v blízkosti atomu fosforu této spojky. Často se přitom nacházel pod prahovou vzdáleností 3Å. Vodíky WATA přicházely do kontaktu se sousední internukleotidovou spojkou rC6-rC7 (obrázek 4.6. graf B.).

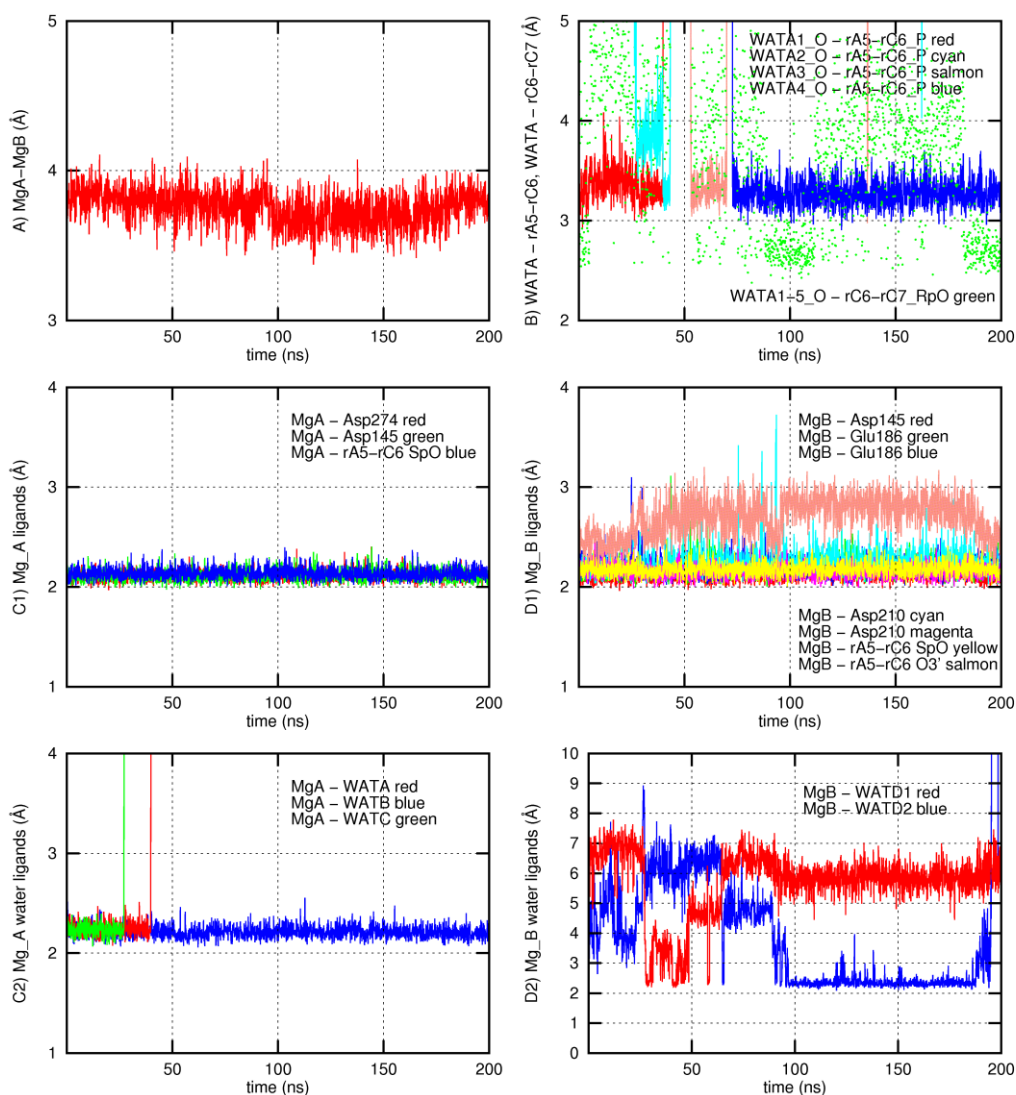
V 3cp5 MD simulaci se kyslík WATA dostával pod prahovou vzdálenost 3 Å méně často a také kontakty vodíků WATA se spojkou rC6-rC7 byly řidší (obrázek 4.7.). Byl to především důsledek rozdílných konformačních preferencí postranního řetězce rezidua Arg278 v MD simulacích 3pc5 a 3cp5.

V 3pc5 MD simulaci byl nevázaný pro-Sp kyslík štípané internukleotidové spojky rA5-rC6 navázán kromě MgA i k MgB. Podobně byl na MgA i MgB navázán postranní řetězec rezidua Asp145. Rezidua Glu186 a Asp210 měla tendenci poskytnout iontu MgB vždy oba kyslíky - tj. vytvářet tzv. bidentátní kontakt (obrázek 4.6. graf D1.).

Kyslík O3' z fosfátové skupiny štípané internukleotidové spojky rA5-rC6 byl po první polovinu MD simulace 3pc5 slabě navázán na MgB. Ve druhé polovině 3pc5 MD simulace se tato interakce ještě více rozvolnila. V 3cp5 MD simulaci byl pak tento proces ještě zřetelnější. V nitru HS RNázy H poblíž MgB se totiž nacházelo několik molekul vody. Zhruba v polovině obou 3pc5/3cp5 MD simulací tyto molekuly vody vytlačily O3' štípané internukleotidové spojky rA5-rC6 od MgB. V 3pc5 MD simulaci se pak v této pozici vystřídal několik různých molekul vody (obrázek 4.6., obrázek 4.7.).



Obrázek 4.6.: Kontakty v aktivním místě HS RNázy H v 3pc5 MD simulaci - A) vzdálenost MgA-MgB (Graf A); C) vzdálenost MgA - ligand (Graf C1: Asp274 červená, Asp145 zelená, Sp modrá; Graf C2: WATA-C); D) vzdálenost MgB - ligand (Graf D1: Asp145 červená, Glu186 zelená, Glu186 modrá, Asp210 bleděmodrá, Asp210 fialová, Sp žlutá, O3' lososová, Graf D2: WAT) [2].



Obrázek 4.7.: Kontakty v aktivním místě HS RNázy H v 3cp5 MD simulaci - A) vzdálenost MgA-MgB (Graf 6A); C) vzdálenost MgA - ligand (Graf C1: Asp274 červená, Asp145 zelená, Sp modrá; Graf C2: WATA-C); D) vzdálenost MgB - ligand - (Graf D1: Asp145 červená, Glu186 zelená, Glu186 modrá, Asp210 bleděmodrá, Asp210 fialová, Sp žlutá, O3' lososová, Graph D2: WAT) [2].

MD simulace 3pc5 a 3cp5 - interakce RNázy H s duplexem DNA:RNA

Ve vazebné oblasti pro vlákno DNA na povrchu HS RNázy H se střídají klastry hydrofobních a polárních reziduí podobně, jako se v cukrfofátové páteři DNA

střídají záporně nabitě fosfátové skupiny a spíše hydrofobní deoxyribózy. To umožňuje optimální vzájemnou interakci. S vazebným žlábkem pro vlákno DNA na povrchu HS RNázy H interaguje celkem pět internukleotidových spojek (obrázek 2.7.) (bráno od 3'-konce DNA vlákna). První dvě spojky (dG1-dC2 a dC2-dT3) interagují s jádrem HS RNázy H. Třetí spojka (dT3-dG4) se váže do tzv. vazebné kapsy pro fosfát. Poslední dvě spojky (dG4-dT5 a dT5-dG6) mohou případně interagovat s výrazně basickou smyčkou, která obsahuje rezidua Arg153-Arg155 a Arg157.

V 3pc5 MD simulaci DNA vlákno vyklouzlo velmi rychle (během prvních desítek ns) z vazebné štěrbiny na povrchu RNázy H. Zůstalo částečně navázáno pouze ve třetím vazebném místě pro DNA (obrázek 4.14.). To se zdá být v souladu s faktem, že 3pc5 oligonukleotidy jsou biologicky neaktivní [162]. Následující odstavce proto pojednávají o 3pc5 MD simulaci, kde byl komplex nukleové kyseliny a RNázy H stabilní.

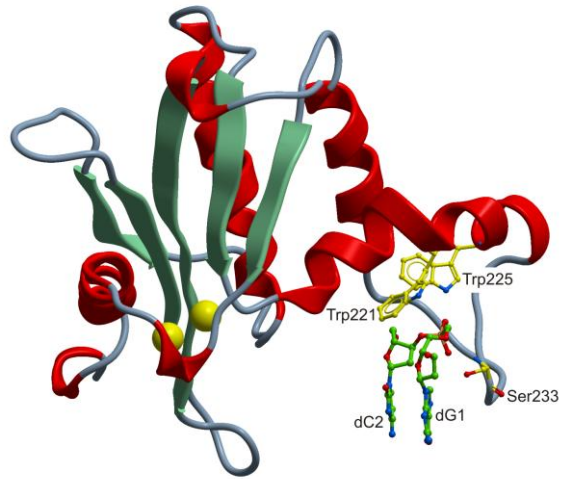
První vazebné místo pro DNA se nachází u tzv. basického výčnělku HS RNázy H. Je to žlábek zformovaný rezidui Trp221, Trp225 a Ser233, do kterého zapadá cukr-fosfátová páteř vlákna DNA [44, 48]. V MD simulaci 3pc5 byly vodíkové vazby páteře s reziduem Trp225 velmi stabilní. U vodíkových vazeb s reziduem Ser233 bylo zaznamenáno několik reverzibilních rozpadů. Nejdelší z těchto rozpadů nastal v čase cca. 80-100 ns. Reziduum Trp221 (které pravděpodobně napomáhá RNáze H odlišit duplexy DNA/RNA od RNA/RNA [44]) bylo v těsném kontaktu s fosfonátovou spojkou dG1-dC2 (obrázek 4.8., obrázek 4.13. graf A).

Vzhledem k tomu, že přirozená fosfátová spojka dC2-dT3 neproniká dostatečně hluboko k povrchu HS RNázy H, nebylo druhé vazebné místo pro DNA explicitně zmíněno ve studii o krystalové struktuře [44], která posloužila jako základ pro tvorbu zde prezentovaných simulovaných systémů.

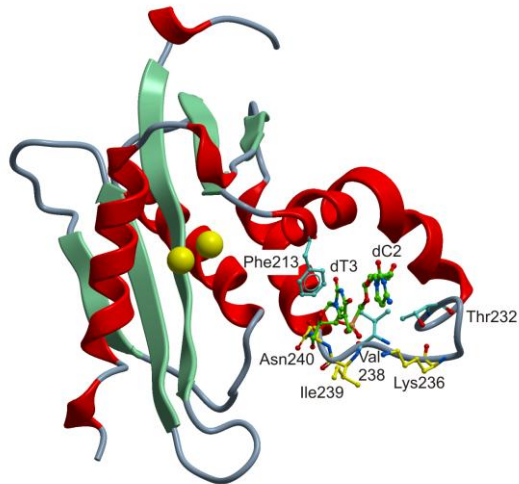
Fosfonátová spojka dC2-dT3 interagovala v 3pc5 MD simulaci s amidovými skupinami hlavního řetězce u reziduí Asn240 a Ile239. Tyto stabilní vazby byly doprovázeny krátkodobými kontakty s postranními řetězci reziduí Lys236 a Asn240. Obzvláště v případě Lys236 však byly tyto kontakty velmi letmé. Byly také zaznamenány interakce mezi hydrofobním reziduem Phe213 a deoxyribózou dT3 a dále mezi rezidui Thr232 a Val238 a -CH₂- skupinou modifikované spojky dC2-dT3 (obrázek 4.9., obrázek 4.13. graf B).

Třetí vazebné místo pro DNA (tzv. vazebná kapsa pro fosfát) se nachází dva nukleotidy před štípanou internukleotidovou spojkou. Je citlivé vůči chemickým modifikacím, které u DNA způsobují přechod do atypické A-formy. Tyto modifikace v konečném důsledku nepřímo zamezují štěpení RNA vlákna RNázou H. Odtud plyne mimořádný význam třetího vazebného místa pro rozpoznání správného substrátu RNázou H [44].

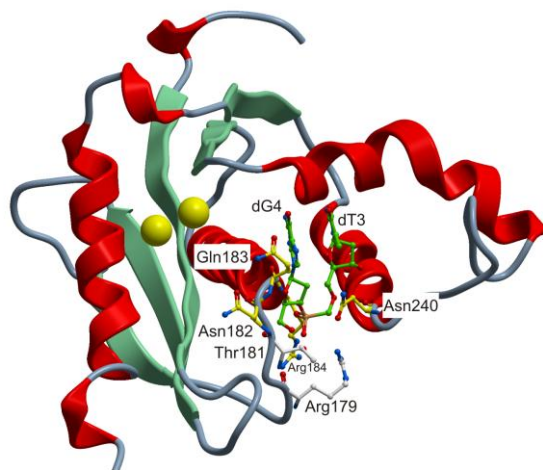
K efektivnímu navázání fosfátové spojky do vazebné kapsy je totiž potřeba přechod torzních úhlů α a γ v páteři vlákna DNA do konformace vzdálené cca. 150° od energetického minima fosfátové spojky. Toho je schopno dosáhnout vlákno DNA, nikoliv však vlákno RNA [44, 167]. Když byla v přípravném MD běhu spojce dT3-dG4



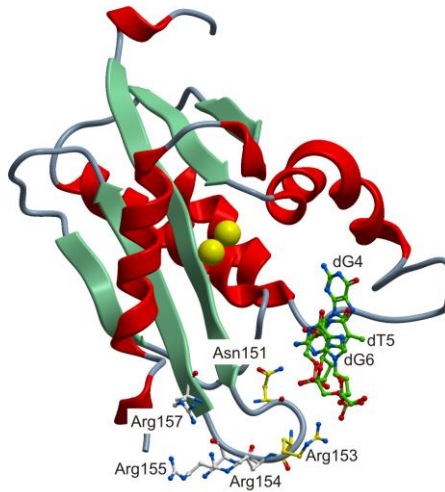
Obrázek 4.8.: Vazba 3pc5 dG1-dG2 internukleotidové spojky do prvního DNA vazebného místa na povrchu HS RNázy H [2].



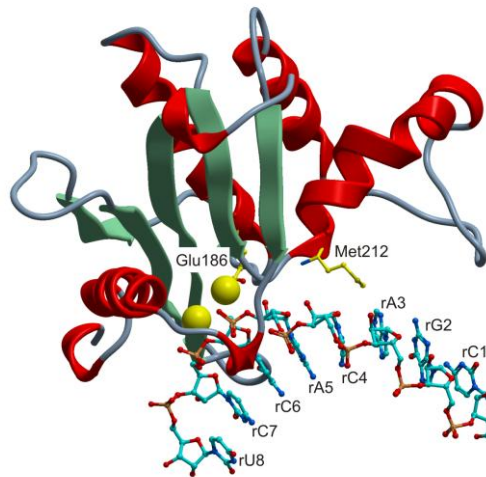
Obrázek 4.9.: Vazba 3pc5 dC2-dT3 internukleotidové spojky do druhého DNA vazebného místa na povrchu HS RNázy H [2].



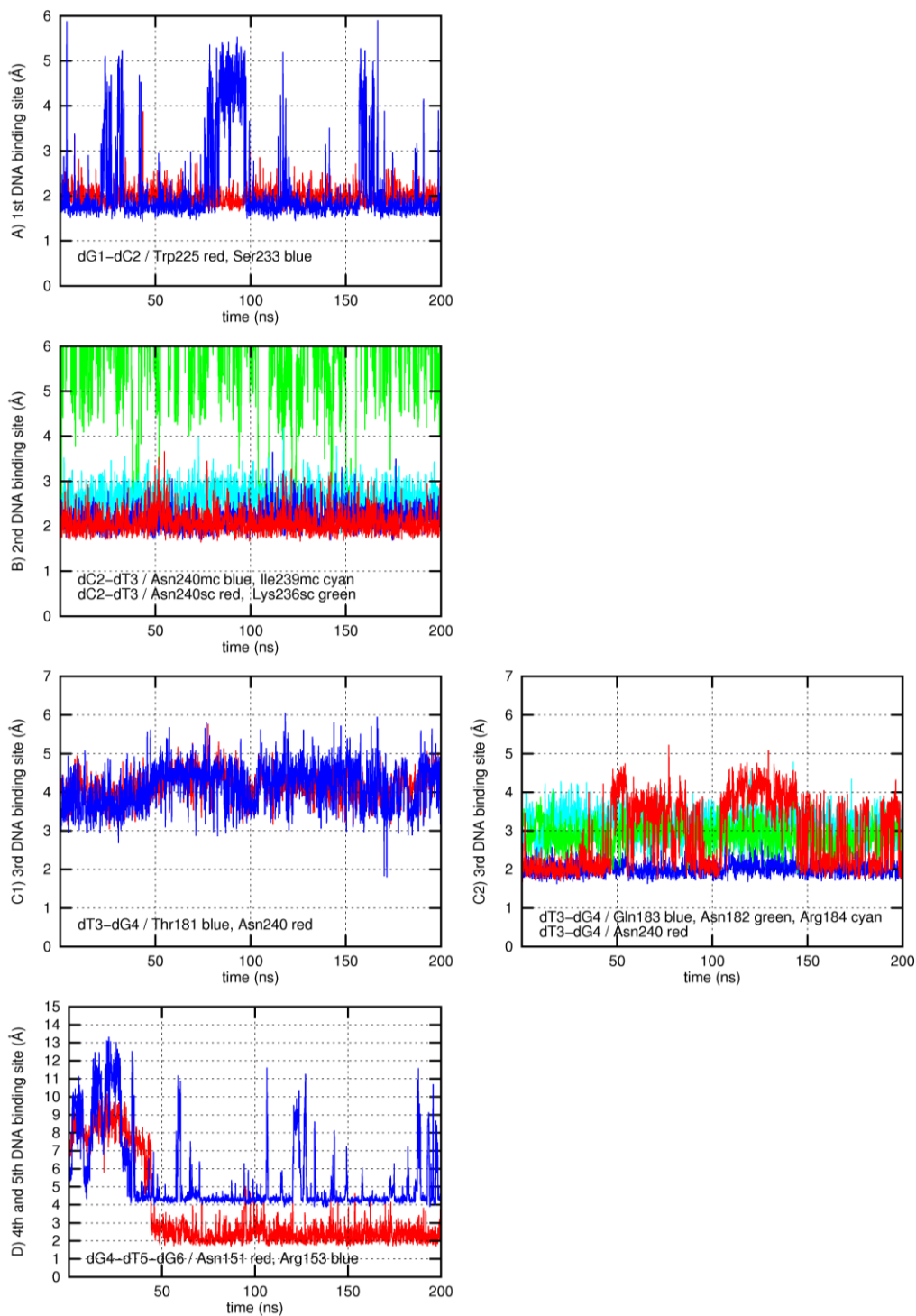
Obrázek 4.10.: Vazba 3pc5 dT3-dG4 internukleotidové spojky do třetího DNA vazebného místa na povrchu HS RNázy H [2].



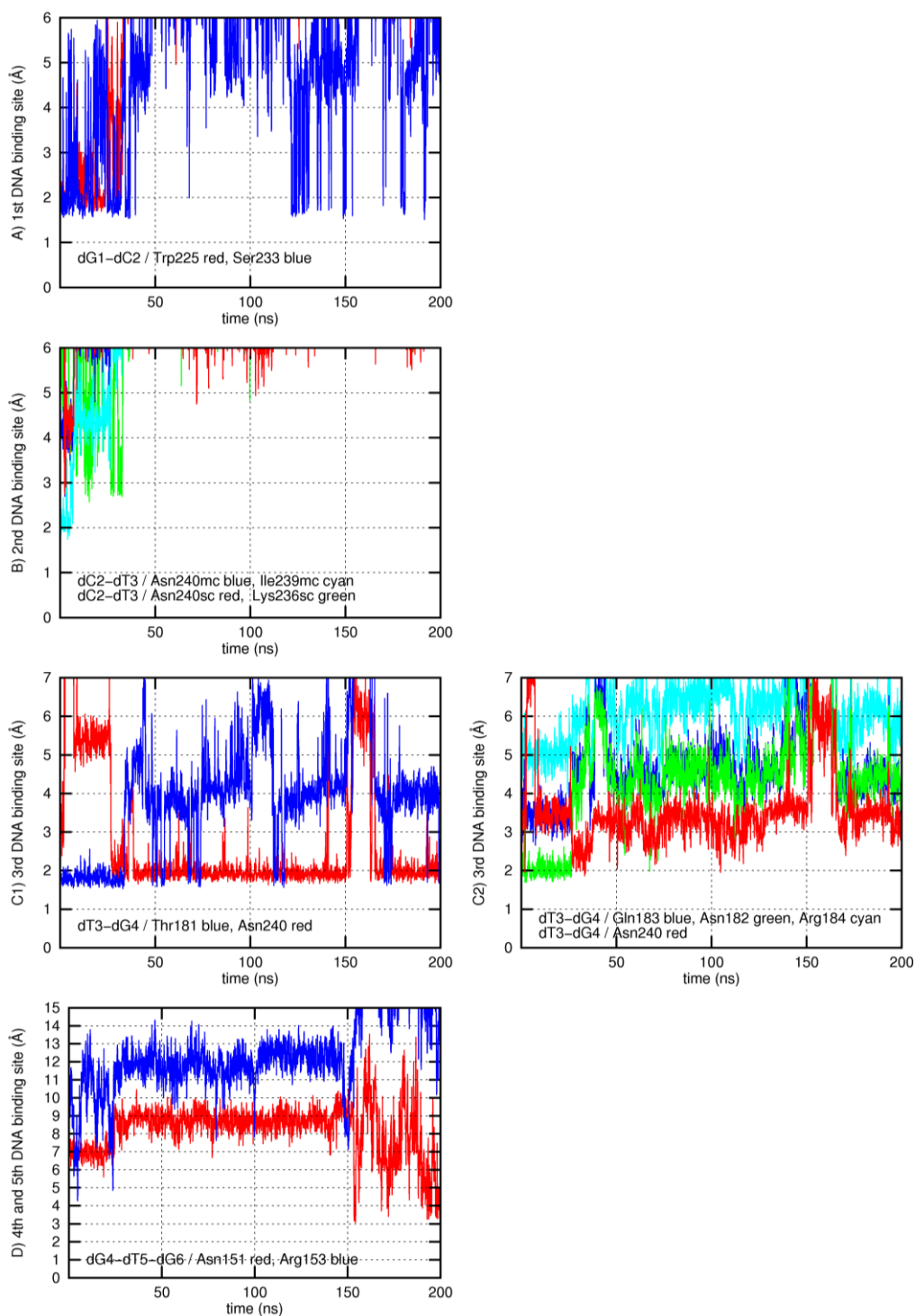
Obrázek 4.11.: Vazba internukleotidových spojek dG4-dT5 a dT5-dG6 do čtvrtého a pátého DNA vazebného místa na povrchu RNázy H [2].



Obrázek 4.12.: Interakce RNA vlákna s HS RNázou H [2].



Obrázek 4.13.: Interakce 3pc5 DNA s HS RNázou H - A) Trp225 – červená, Ser233 - modrá; B) Asn240 m.c. - modrá, Ile239 m.c. - bleděmodrá, Asn240 s.c. - červená, Lys236 s.c. - zelená; C1) Thr181 – modrá, Asn240 – červená; C2) Gln183 - modrá, Asn182 - zelená, Arg184 - bleděmodrá, Asn240 - červená; D) Asn151 - červená, Arg153 - modrá [2].



Obrázek 4.14.: Interakce 3cp5 DNA s HS RNázou H - A) Trp225 - červená, Ser233 - modrá; B) Asn240 - modrá, Ile239 - bleděmodrá, Asn240 - červená, Lys236 - zelená; C1) Thr181 - modrá, Asn240 - červená; C2) Gln183 - modrá, Asn182 - zelená, Arg184 - bleděmodrá, Asn240 - červená; D) Asn151 - červená, Arg153 - modrá [2].

vnucena obvyklá konformace **-g-g**, která odpovídá jejímu energetickému minimu, došlo k přerušení interakcí mezi reziduem Thr181 a fosfátovou skupinou dT3-dG4.

V případě AO je možné striktní požadavky kladené tímto vazebným místem na konformační preference substrátu obejít prostřednictvím tzv. gapmerů - modifikovaných oligonukleotidů, do jejichž centrální části je vloženo několik přirozených deoxynukleotidů. Další možností je použití oligonukleotidů se střídajícími se modifikovanými a nemodifikovanými spojkami.

Což je pravděpodobně i případ 3pc5 fosfonátů [162]. V 3pc5 MD simulaci se totiž nerealizovala vazba postranního řetězce rezidua Thr181 na fosfonátovou skupinu dT3-dG4 [44]. Namísto toho interagovala fosfonátová spojka dT3-dG4 s amidovou skupinou hlavního řetězce u rezidua Gln183. Kromě této stabilní vazby byla zaznamenána i ne zcela stabilní vazba na postranní řetězec rezidua Asn240 (obrázek 4.10., obrázek 4.13.) (s 50 ns reverzibilním rozpadem).

Díky suboptimální interakci 3pc5 fosfonátové spojky dT3-dG4 s vazebnou kapsou pro fosfát se zdá, že oligonukleotidy s alternujícími přirozenými a 3pc5 modifikovanými spojkami jsou aktivní spíše v okamžiku navázání přirozené spojky do vazebné kapsy pro fosfát, než naopak.

Ve čtvrtém vazebném místě pro DNA navázala internukleotidová spojka dG4-dT5 kontakt s reziduem Asn151 až po cca. 30 ns MD simulace. Podobně i v pátém vazebném místě navázala spojka dT5-dG6 kontakt s reziduem Arg153 (obrázek 4.11., obrázek 4.13. graf D). Tyto dvě interakce byly poté stabilní (s výjimkou reverzibilního rozpadu vazby s reziduem Arg153 v cca. 120 ns). Reziduum Arg153 spolu s Arg154, Arg155 a Arg157 tvoří výrazně kladně nabitou smyčku, která je pravděpodobně zapojena do počátečního rozpoznávání mezi enzymem a substrátem - duplexem RNA:DNA se záporně nabitou páteří. V krystalových strukturách byly u této smyčky zaznamenány různé konformery [44]. Přípravné MD simulace ukázaly, že ani trajektorie dlouhá 1 μ s nemusí být k relaxaci smyčky dostatečná.

Interakce HS RNázy H s RNA vláknem duplexu DNA:RNA

Aktivní místo RNázy H leží ve vazebném žlábků pro vlákno RNA z duplexu DNA:RNA [44]. Dostupné krystalové struktury naznačují, že mechanismus rozeznávání RNA vlákna enzymem je stejný u RNáz H z různých organismů. RNA vlákno s enzymem interaguje prostřednictvím 2'-OH skupin čtyř po sobě jdoucích nukleotidů. Dvě interagující 2'-OH skupiny se přitom nacházejí před štípanou internukleotidovou spojkou a dvě za ní. Do interakcí s 2'-OH skupinami jsou zapojena rezidua Glu186, Cys148, Ser150, Asn151 a Met212 [44].

V provedených MD simulacích byl kontakt mezi RNA a HS RNázou H zprostředkovan především postranním řetězcem Glu186 (obrázek 4.12.). Toto reziduum se přitom také váže na MgB v aktivním místě HS RNázy H [44]. Navíc leží na stejném helixu jako nedaleké Thr181, které hraje ústřední úlohu ve fosfát vázící kapse - třetím DNA vazebném místě. Nabízí se tedy hypotéza, že vazba substrátu do tohoto DNA vazebného místa citlivého na chemické modifikace by mohla alostericky ovlivňovat dění v aktivním místě po lince Thr181-Glu186.

Reziduum Met212 bylo v kontaktu s rC4, zejména ve druhé polovině 3pc5 MD simulace. Nevýrazné interakce mezi rezidui Cys148, Ser150, Asn151 a 2'-OH skupinami rC6 a rC7 byly zřejmě částečně způsobeny nedostatečnou relaxací nedaleké basické smyčky 153-157.

Funkční význam by mohla mít interakce mezi reziduem Arg278 a spojkou rC6-rC7, která byla v MD simulaci s výjimkou krátkých reverzibilních rozpadů stabilní. Spojka rC6-rC7 přitom sousedí se štípanou spojkou rA5-rC6. Arg278 by tak například mohlo napomáhat stabilizaci produktů reakce. Internukleotidová spojka rC6-rC7 může navíc přijmout proton uvolněný z katalytické molekuly vody WATA.

Duplex DNA:RNA - strukturní parametry

Vazba 3pc5/3cp5 duplexů DNA:RNA k RNáze H ovlivnila jejich geometrii. Kvantitativně je možné to postihnout prostřednictvím tzv. helikálních parametrů z nichž nejvíce byly ovlivněny Buckle a Shift. Patrné jsou také rozdílné šířky malého žlábků v různých místech podél duplexu DNA:RNA. Tyto distorze přetrvávaly i po uvolnění 3cp5 DNA z vazebného žlábků RNázy H. K plné relaxaci duplexu 3cp5 tedy nestačil ani MD běh trvající celých 200 ns. Z toho je zřejmé, že výše zmíněné deformace jsou způsobeny spíše celkovým tvarem HS RNázy H, než detailními interakcemi v jednotlivých vazebných místech.

Párování bází bylo stabilní (kromě terminálních párů).

Cukry RNA vlákna byly v konformaci **C3'-endo**, což odpovídá klasické A-formě. V okolí tzv. fosfátové vazebné kapsy [44] však byly v obou strukturách 3pc5/3cp5 znamenány četné přechody do konformace **C2'-exo**. Konformační variabilita cukrů v DNA vláknech byla větší. Převládaly konformery **C2'-endo** a **O4'-endo**. Výjimkou byla rezidua dC2 a dT3 ve vazebném kanálu DNA. V 3pc5 MD simulaci byly v konformaci **C4'-exo**. V 3cp5 MD simulaci bylo v této konformaci pouze dT3 (v důsledku částečné relaxace 3cp5 DNA po uvolnění od HS RNázy H).

Všechny fosfonátové internukleotidové spojky prokazovaly značnou konformační variabilitu. Fosfátové spojky, jak ve vláknech DNA tak i RNA, preferovaly konformaci odpovídající energetickému minimu **-g-g**. Dokonce i spojka rC6-rC7 (interagující s Arg278) preferovala konformaci **-g-g** s oscilacemi do konformace **t-g**.

Shrnutí

DNA vlákno s modifikacemi internukleotidových spojek typu 3'-O-P-CH₂-O-5' (3pc5) bylo v 200 ns MD schopné vytvořit stabilní komplex s RNA vláknem a HS RNázou H - na rozdíl od DNA vlákna s modifikacemi typu 3'-O-CH₂-P-O-5' (3cp5). Vyšší afinita 3pc5 DNA k HS RNáze H vysvětluje schopnost těchto oligonukleotidů vyvolat enzymatické štěpení RNA [162] (na rozdíl od 3cp5 DNA).

Artifciální repulze mezi hořčikovými ionty v binukleárních aktivních místech způsobuje během MD simulací zásadní změny v koordinaci ligandů [147]. Ve štípané spojkě rA5-rC6 (která je v aktivním místě vetknuta mezi oba hořčikové ionty) byla proto použita polarizovaná fosfátová skupina. Standardní parametrizace hořčikového iontu vede v MD simulacích také k podcenění vzdáleností iont-ligand. Proto byl VdW rádius R* pro Mg²⁺ změněn z 0.7926 Å na 1.1Å. Díky výše

zmíněným korekcím byly vzdálenosti MgA–MgB v aktivním místě RNázy H stabilní během MD simulací o délce 200 ns (obrázek 4.6. graf A, obrázek 4.7. graf A). Referenční 1000 ns MD simulace se sedmibodovou AW reprezentací hořčkových iontů [147] vedla k téměř shodným výsledkům (obrázek 4.5. graf A).

Kyslík O3' štípané spojky rA5-rC6 byl v průběhu všech MD simulací vypuzen od iontu MgB molekulami vody. Je to pravděpodobně způsobeno rozdílnými metodikami pro odvození parciálních nábojů u biomolekul (AMBER/RESP) a molekul vody TIP3P [92]. Tento artefakt byly zřetelnější u MD simulace se sedmibodovou AW reprezentací iontu hořčíku [147].

Reziduum Glu186 (interagující s MGB v aktivním místě HS RNázy H [44]) se nachází v blízkosti konzervovaného rezidua Thr181 (v tzv. fosfátové vazebné kapse pro DNA). V případě suboptimální vazby modifikovaných internukleotidových spojek AO do této fosfátové vazebné kapsy je proto možné, že jsou od Thr181 přenášeny alosterické signály přes Glu186 do aktivního místa RNázy H, které v konečném důsledku zabrání degradaci RNA vlákna.

4. 3. Molekulárně dynamické simulace komplexu enzymu Argonaute s duplexem DNA:RNA

Motivace

Proteiny Argonaute (AGO) (obrázek 4.15.) jsou zapojeny do procesu tzv. RNA interference (RNAi) [54-55, 58, 59-60], což je způsob posttranskripčního utlumení exprese genetické informace prostřednictvím krátkých dvoušroubovic RNA (dsRNA) [168-170]. RNAi se uplatňuje u eukaryotů při regulaci minimálně 1/3 genů, vývoji buněk, obraně proti virové nákaze atd.

V první fázi RNAi endoribonukleáza Dicer rozštěpí dlouhou dsRNA virového původu na fragmenty - tzv. krátké interferující RNA (siRNA) [53]. Ty následně vytvoří komplexy s efektorovými enzymy AGO, které jedno z vláken siRNA duplexu rozštěpí. Výsledný ribonukleoproteinový komplex (RISC) [60-61] používá zbývající „řídící“ vlákno siRNA jako templát pro specifické interakce s cílovou RNA (virového původu či mRNA). V případě dostatečné komplementarity řídícího vlákna siRNA k cílové RNA provede enzym AGO přeštipnutí vlákna cílové RNA. Řídící siRNA vlákno zůstává zachováno a v komplexu s proteinem AGO může štěpit další kopie cílové RNA.

Proteiny AGO jsou analogické RNáze H, pouze namísto antisense deoxyoligonukleotidů jsou aktivované krátkými vlákny RNA [53]. Velikost konformačních změn v průběhu štěpení cílové RNA je ale u RNázy H a AGO rozdílná. Aktivní místo RNázy H je pro okamžité štěpení RNA stále předpřipraveno.

U AGO proteinů tomu tak není. Vyplývá to pravděpodobně z vysoké specifity těchto enzymů, která je nezbytná pro jejich správnou funkci. V důsledku toho probíhají u AGO konformační změny v průběhu celého procesu štěpení cílové RNA. Mezi těmito konformačními změnami je i rotace PAZ domény a také pohyb smyček L1, L2 umístěných na povrchu podjednotky PIWI. V průběhu štěpení cílové RNA se mění také konformace řídícího siRNA vlákna natolik, že jeho 3'-konec je dočasně uvolněn z vazebné kapsy v doméně PAZ. 5'-konec siRNA vlákna však zůstává ukotven ve vazebné kapse podjednotky Mid po celou dobu štěpení [171-172].

A jak bylo zjištěno nedávno, konformační změny se odehrávají i v aktivním místě AGO, kde krystalové struktury odhalovaly přítomnost katalytické triády DDX. Vzhledem k tomu, že podjednotka PIWI AGO má stejný fold jako RNáza H [53, 59-60] (obrázek 4.15.) (s katalytickou tetradou DEDD), spekulovalo se o tom, že i u AGO proteinů ve skutečnosti existuje katalytická tetraáda.

V nedávné krystalové struktuře *Kluyveromyces polysporus* Argonaute (KpAGO) bylo identifikováno reziduum Glu1013 zasahující do aktivního místa. Glu1013 je součástí mobilní smyčky, která může být stabilizovaná v blízkosti aktivního místa pomocí sítě vodíkových vazeb [171]. Mutace rezidua Glu1013 narušuje funkci KpAGO ve stejném rozsahu jako mutace katalytického Asp1046.

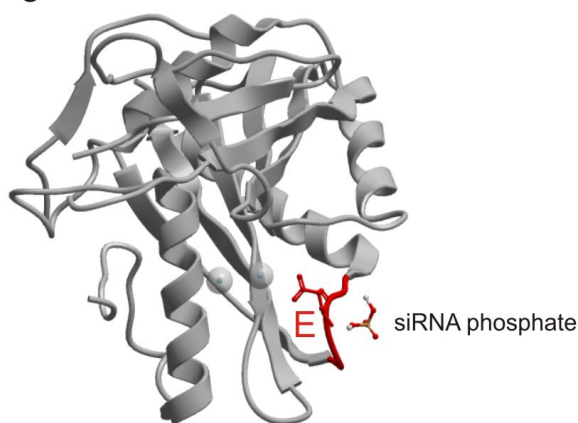
V krystalové struktuře *Thermus thermophilus* Argonaute (TtAGO) je analogické Glu512 součástí smyčky L2, která přímo interaguje s nukleotidy 11-15 řídícího

vlákna DNA. Kromě toho se Glu512 nachází v blízkosti jednoho z divalentních iontů aktivního místa. Navíc je Glu512 lokalizováno i v blízkosti 2'-OH skupiny nukleotidu sousedícího se štípanou fosfátovou skupinou. Glu512 by tak mohlo být jedním z katalytických reziduí TtAGO [171-172].

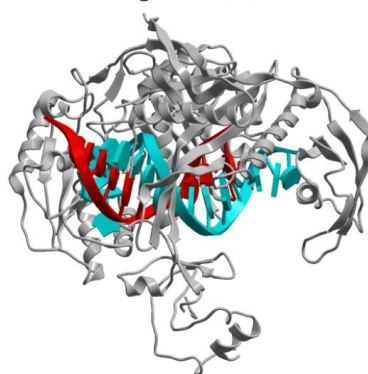
Konzervovaná a stabilní katalytická tetráda DEDD RNáz H (obrázek 4.17.) by tak byla téměř izosterická s labilní katalytickou tetrádou DEDX obou výše zmíněných proteinů AGO, která se díky konformační variabilitě smyček, jejichž součástí jsou výše zmíněná Glu rezidua, může snadno změnit na neaktivní triádu DDX. Přechody DEDX-DDX zřejmě stojí za vysokou specifitou proteinů AGO [171].

Cílem zde prezentované MD simulace bylo ověřit, že s novou jednočásticovou reprezentací hořčkových iontů bude DEDX uspořádání aktivního místa TtAGO stabilní.

Argonaute "Rnase H" domain



Argonaute



Obrázek 4.15.: Doména enzymu Argonaute strukturně homologní s RNázou H. Červeně je vyznačenou Glu reziduum interagující zároveň s aktivním místem i fosfátovou skupinou jedné z internukleotidových spojek siRNA vlákna [2].

Simulovaný systém

Simulovaný systém byl vytvořen na základě krystalové struktury komplexu TtAGO s nukleovou kyselinou, která má v proteinové databázi PDB kódové označení 3HVR

[172]. Vstupní soubory pro celý simulovaný systém byly vytvořeny pomocí modulu TLEAP ze softwarového balíku AMBER [146]. Konkrétně se jednalo o soubory *.inpcrd (počáteční souřadnice) a *.prmtop (silová pole, molekulární topologie atd.). V průběhu budování vstupních souborů byly k biomolekulám přidány molekuly vody typu TIP3P [92]. Tak vznikl simulovaný systém s 88.603 atomy vytvářející periodický box o rozměrech cca. 117 Å x 106 Å x 87 Å.

Metoda MD simulací

Pomocí softwarového balíku NAMD 2.7 [103] byly provedeny ekvilibrační MD běhy v délce 5 ns. Pro ně byl použit izotermicko-izobarický soubor (NPT) a časový krok 2 fs [103]. Použitím Langevinovy dynamiky pro kontrolu teploty byl simulovaný soubor ohřát z 0 na 310 K. Metoda tzv. Langevinova pístu byla použita pro kontrolu tlaku nastaveného na 1 atmosféru [103]. Cut-off nevazebných sil byl nastaven na 9 Å. Pro výpočet dalekodosahových elektrostatických sil byla použita smooth Particle-Mesh Ewaldova (PME) sumace. Algoritmus SHAKE byl aplikován na vazby s atomy vodíku. Ekvilibrační MD běhy byly realizovány prostřednictvím multiprocesorového systému osazeného šestnácti procesory Intel Itanium s kmitočtem 1,5 GHz. Výpočet probíhal rychlostí 2 ns/den.

Dlouhé produkční MD simulace byly realizovány prostřednictvím softwarového balíku ACEMD v. 2591 [108, 143] a grafických procesorů firmy NVIDIA. Pro tyto MD simulace byl použit soubor NVT, neboť ACEMD použití NPT neumožňuje [108, 143]. To není nikterak na škodu, neboť po dostatečně dlouhých ekvilibračních MD bězích se objem velkých simulovaných systémů mění v NPT simulacích jen zcela nepatrně [108, 143]. Časový krok byl nastaven na 4 fs díky tzv. "mass repartitioning scheme", které umožňuje přerozdělit část hmotnosti těžkých atomů na vodíky na ně navázané. Jinak byly použity metody a parametry shodné s těmi, které byly použity v ekvilibračních bězích. Pro realizaci produkčních MD simulací byly použity programovatelné GPU (NVIDIA GTX-580) osazené 480 jádry. Rychlost výpočtu byla cca. 20 ns/den. Celková délka trajektorie byla 500 ns.

Snímky z MD simulací byly ukládány každých 100 ps. Následně byly analyzovány pomocí softwarových balíčků CHIMERA 1.5.3, VMD 1.9 a PTRAJ/AMBER10 [125-127, 147]. Obrázky byly vytvořeny prostřednictvím softwarového balíku ICM Molsoft 3.7.

Nová jednobodová reprezentace hořčíkových iontů

V přípravných MD simulacích se standardně parametrizovanými hořčíkovými ionty se v binukleárním aktivním místě AGO objevovaly nestability. Podobná pozorování byla učiněna i u RNázy H (viz. předchozí kapitola) a lze je nalézt i v literatuře (např. v aktivním místě polymerázy- β [147]). Jednoatomový model iontu Mg^{2+} byl tedy precizován tak, aby výsledky MD simulací byly ve shodě s výsledky přípravné MD simulace získané prostřednictvím sedmibodového AW modelu (podrobněji viz. předchozí kapitola). Bylo vzato do úvahy, že zdrojem problémů je zanedbání polarizace a podhodnocení VdW poloměru hořčíkového iontu [165-166]. Proto byly parametry pro hořčíkový iont změněny z ($R^* = 0.7926$ Å, $\epsilon = 0.8947$) na ($R^* = 1.1$ Å, $\epsilon = 0.8947$). Modifikovány byly ale také parciální náboje postranních

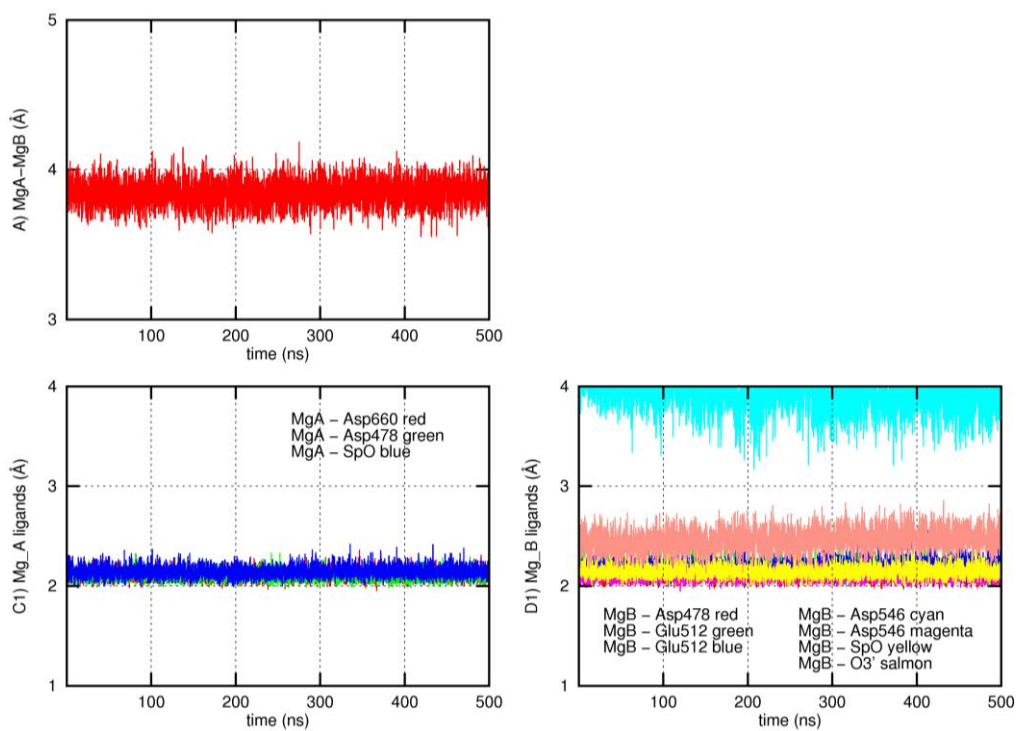
kyslíků (z $-0.7761/-0.7761$ na $-0.9761/-0.5761$) ve štípané internukleotidové spojce, která se v aktivním místě AGO váže mezi oba hořčíkové ionty.

Výsledky MD simulací

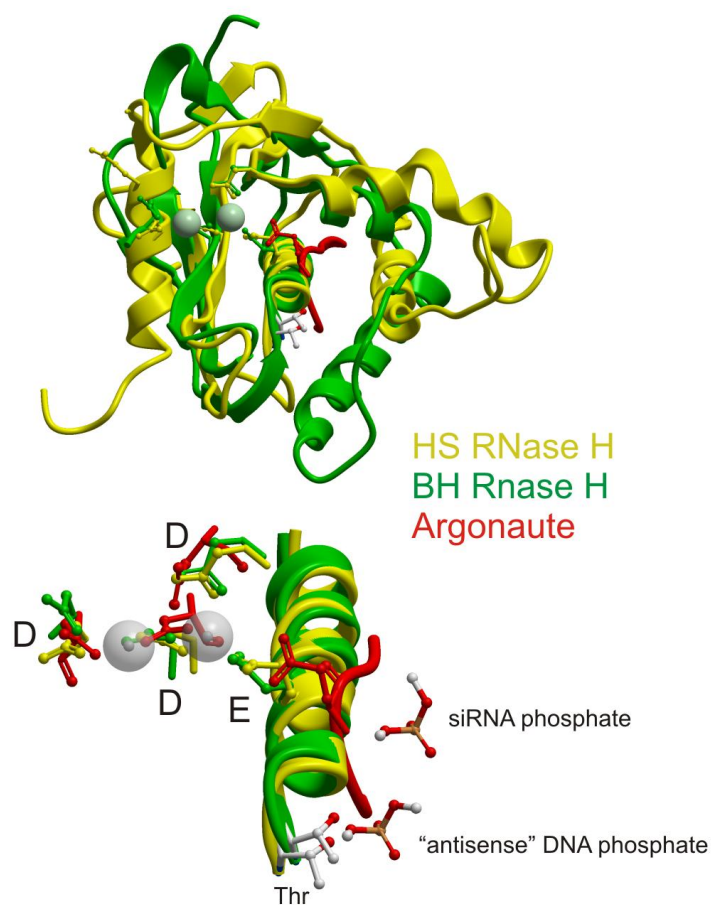
Pro simulovaný systém TtAGO v komplexu s nukleovou kyselinou byla provedena MD simulace trvající 500 ns. Uspořádání aktivního místa bylo po celou dobu MD simulace stabilní (obrázek 4.16.). To potvrdilo transferabilitu nové reprezentace binukleárního aktivního místa vytvořené původně pro RNázu H (viz. předchozí kapitola).

Postranní řetězec Glu512 se v aktivním místě TtAGO vázal jak k MgB, tak i k 2'-OH skupině nukleotidu sousedícího se štípanou fosfátovou skupinou (přesně stejně jako analogické reziduum Glu186 v HS RNáze H [44]). Zajímavé je, že atomy hlavního řetězce proteinu u Glu512 přitom přímo interagují s fosfátovou skupinou siRNA na pozici 13-14. siRNA molekuly s modifikovanou páteří tak mohou zcela přímočaře ovlivňovat aktivní místo AGO. Připomeňme, že u RNázy H se reziduum Glu186 (analogické Glu512 u TtAGO) nachází pět pozic od konzervovaného rezidua Thr181 situovaného v tzv. fosfátové vazebné kapse pro DNA [44]. V případě suboptimální vazby modifikovaných internukleotidových spojek AO do této fosfátové vazebné kapsy jsou proto od Thr181 přenášeny alosterické signály přes Glu186 do aktivního místa RNázy H, které v konečném důsledku zabrání degradaci RNA vlákna. Přímočařejší způsob alosterického ovlivňování aktivního místa AGO zřejmě souvisí s podstatně většími nátoky na specificku působení proteinů AGO ve srovnání s RNázou H.

Pro kvantitativní postihnutí dopadů chemických modifikací ve struktuře siRNA přes Glu512 na geometrii aktivního místa AGO bude zřejmě potřeba dále rafinovat způsob, jakým jsou v MD simulacích reprezentována binukleární aktivní místa. Dost možná bude nezbytné použít QMMM metody. Ve zde prezentované MD simulaci totiž tvořil Asp546 (analogický Asp210 u RNázy H), monodentátní kontakt s MgB, kdežto kruciální Glu512 (jehož ekvivalentem je v HS RNáze H Glu186 [44]) kontakt bidentátní (obrázek 4.16. graf D1) (tj. Glu512 mělo tendenci interagovat s MgB prostřednictvím obou kyslíků). Vzhledem k rozlišení krystalových struktur není možné posoudit, zda se jedná o artefakt či realitu. Způsob interakce Glu512 s MgB může přitom mít zásadní vliv na průběh reakce a dost možná může této reakci zabránit.



Obrázek 4.16.: Kontakty v aktivním místě v TtAGO MD simulaci - A) vzdálenost MgA-MgB (Graph A); C) vzdálenost MgA - ligand (Graph C1: Asp660 červená, Asp478 zelená, Sp modrá; D) vzdálenost MgB - ligand (Graf 6D1: Asp478 červená, Glu512 zelená, Glu512 modrá, Asp546 bleděmodrá, Asp546 fialová, Sp žlutá, O3' lososová) [2].



Obrázek 4.17.: Strukturní alignment HS RNázy H, BH RNázy H a reziduí v aktivním místě enzymu Argonaute. Dole je znázorněno alosterické ovlivňování aktivního místa RNáz H prostřednictvím vazby antisense oligonukleotidu a přímé ovlivňování aktivního místa enzymu Argonaute vláknem siRNA [2]

4.4. Molekulárně dynamické simulace komplexů ankyrinové domény Ribonukleázy L s 2'-5' oligoadenyláty a jejich analogy

Motivace

Přirozená odezva imunitního systému savců na začínající virovou infekci je zprostředkována převážně pomocí interferonů typu I (IFN). Ty řídí přepis genů řady proteinů bojujících s replikací virů pomocí široké plejády rozdílných mechanismů [67-68, 173-177]. IFN aktivují mimo jiné geny kódující několik 2',5' oligoadenylátových syntetáz (OAS) a jeden gen kódující ribonukleázu L (RNázu L).

OAS jsou aktivovány navázáním na virovou dsRNA (duplexy RNA se jinak v buňkách savců běžně nevyskytují). Aktivované OAS tvoří z ATP oligoadenyláty s 2',5'-internukleotidovými spojkami (2-5A) [67-68, 173-177]. 2-5A se váží na RNázu L, která následně dimerizuje do aktivní formy. Aktivní forma RNázy L je schopna štěpit ssRNA [67] obsahující diády UU, UA, AU, AA a UG [178].

Aktivace RNázy L je v současné době jediná známá funkce 2-5A. Zásadní roli přitom hraje délka oligoadenylátu. Zatímco trimery a tetramery 2-5A dosahují maximální aktivace RNázy L, dimery jsou neúčinné [67]. Nahrzení jedné z 2',5' internukleotidových spojek v trimerní 2-5A klasickou 3',5' fosfodiesterovou spojkou podstatně snižuje vazebnou schopnost tohoto trimerního 2-5A k RNáze L. RNáza L má také nižší afinitu k 2-5A oligoadenylátům s alternativními bázemi cytidinu, uridinu či inosinu. Nezbytný strukturní požadavek pro plnou aktivaci lidské RNázy L je přítomnost alespoň jedné 5'-fosfátové skupiny v 2-5A [68].

Existuje řada mechanismů, kterými RNáza L společně s OAS přispívá k funkci přirozeného imunitního systému [67]. Jedná se zejména o štěpení virové ssRNA obsahující genom viru, nebo štěpení virových mRNA. Oba tyto mechanismy zamezují syntéze virových proteinů [175]. RNáza L v aktivní formě nicméně štěpí veškerou ssRNA obsaženou v buňce, bez rozlišování je-li virového původu nebo ne. Narušení ribozomální 28S rRNA a 18S rRNA je charakteristický projev působení RNázy L v infikované buňce [175]. Poškození buněčného aparátu nezbytného pro replikaci virů přispívá k antivirovému působení RNázy L. Konstitutivně aktivní zkrácená forma RNázy L je považována za biologický indikátor chronického únavového syndromu [71]. Trvalá aktivace RNázy L, nebo překročení mezních hodnot její aktivace, může vyvrcholit buněčnou apoptózou. To naznačuje možnost využití OAS/RNázy L jako supresorů nádorů [175]. Vzhledem k výše uvedeným skutečnostem je zřejmé, že přímá aktivace RNázy L pomocí 2-5A je novým atraktivním přístupem v léčbě virových, bakteriálních a nádorových onemocnění.

Lidskou RNázu L tvoří 741 aminokyselinových zbytků a má celkovou molekulovou hmotnost 83.543 Da. Celkem se skládá ze tří domén. Jsou to N-terminální ankyrinová (ANK), pseudokinázová a C-terminální ribonukleázová oblast. N-terminální ankyrinová oblast RNázy L (rezidua 24-335) se skládá z devíti ankyrinových motivů zodpovědných za vázání 2-5A. Je to přitom jediná známá ANK, která je schopna vázat oligonukleotidy [67]. Přestože je RNáza L řazena do široké rodiny lidských kináz, její pseudokinázová doména je katalyticky neaktivní. C-terminální oblast RNázy L (rezidua 583-720) je zodpovědná za její katalytickou aktivitu [71].

Ligand 2-5A se váže na druhý až čtvrtý ankyrinový motiv RNázy L, kde interaguje s rezidui Arg155, Tyr135, Glu131, Phe126, Lys89, Gln68, Asn65 a Trp60 [68]. Relativní význam jednotlivých reziduí pro navázání 2-5A byl určen na základě studia mutantů RNázy L, v nichž byla tato vazebná rezidua postupně nahrazena alaninem. Bylo zjištěno, že interakce nezbytné pro navázání 2-5A a aktivaci RNázy L jsou: π - π stacking s rezidui Phe126 a Trp60, elektrostatické interakce s Arg155 a Lys89 a vodíkové vazby s Glu131 [71]. Vzhledem k tomu, že ANK část RNázy L byla původně vykrytalizována jako monomer, nemohly výše zmíněné studie plně vysvětlit mechanismus dimerizace a aktivace RNázy L [68]. To se podařilo až zcela nedávno. Dvě molekuly 2-5A se navážou do dvou vazebných míst na povrchu ANK vzdálených od sebe 50 Å, čímž vyvolají dimerizaci, v jejímž průběhu nenastávají u ANK žádné konformační změny [65].

Přirozené 2-5A jsou v buňkách velmi rychle degradovány díky přítomnosti fosfatáz a nukleáz. Ve snaze vyvinout stabilní analoga 2-5A byla mimo jiné uplatněna i tzv. „bridging“ strategie zavedením 3'-O, 4'-C-alkylenových spojek do adenosinových jednotek. Posléze byly testovány vlastnosti těchto analog včetně jejich odolnosti vůči buněčným nukleázám a schopnosti aktivovat RNázu L [179].

Zde byla provedena série MD simulací ve snaze porozumět dopadu různých bodových mutací lidské RNázy L (R155A, F126A, W60A, K89A) na vazbu 2-5A k enzymu. MD simulace série analog 2-5A s 3'-O, 4'-C-alkylenovou spojkou byly provedeny s cílem zjistit, zda lze pomocí MD simulací v délce stovek nanosekund rozlišit agonisty RNázy L (pB₂₅A₂₅A, pA₂₅A₂₅B a pB₂₅A₂₅B) od inaktivních analog (pA₂₅B₂₅A, pB₂₅B₂₅A, pA₂₅B₂₅B a pB₂₅B₂₅B) [179].

Simulované systémy

Simulované komplexy trimerů 2-5A a jejich analog s HS RNázou L (Tab. 1) byly vytvořeny na základě krystalové struktury 1wdy [68] z proteinové databanky PDB. Modulem TLEAP ze softwarového balíku AMBER [146] byly vytvořeny vstupní soubory pro MD simulace typu .inpcrd (vstupní souřadnice) a .prmtop (sílové pole AMBER [87], molekulární topologie atd.). Přidání molekul vody TIP3P [92] vedlo k rozšíření systému o zhruba 10 Å v každém směru. Výsledný periodický box měl rozměry cca. 121 x 77 x 77 Å. Simulovaný systém sestával z cca. 70.000 atomů.

Metoda MD simulací

Pro ekvilibrační MD běhy o délce 5 ns byl použit izotermicko-izobarický soubor (NPT), standardní simulační protokoly a softwarový balík NAMD 2.7 [103]. Časový krok byl zvolen 2 fs. Pomocí metody tzv. Langevinova pístu byly simulované soubory udržovány na konstantním tlaku 1 atmosféry. Ohřev z 0 K na 310 K byl kontrolován pomocí Langevinovy dynamiky [103]. Cut-off pro nevazebné dalekodosahové síly byl nastaven na 9 Å. Pro výpočet elektrostatických sil byla použita smooth Particle-Mesh Ewaldova (PME) metoda. Na vazby s atomy vodíku byl aplikován algoritmus SHAKE. K provedení ekvilibračních MD běhů byl použit multiprocessorový systém osazený šestnácti procesory Intel Itanium s kmitočtem 1,5 GHz.

Produkční MD simulace byly provedeny na lokálních stanicích osazených grafickými procesory (GPU) pomocí softwarového balíku ACEMD v.2591 [108, 148]. ACEMD nepodporuje MD simulace v NPT souboru, proto byl použit soubor NVT (po ekvilibrizaci se beztak objem velkých simulovaných systémů s biomolekulami mění jen minimálně). Konstantní teplota (290 K či 310 K) byla udržována pomocí Langevinova termostatu [108, 148]. Cut-off byl u dalekodosahových interakcí nastaven na 9 Å. Časový krok 4 fs umožnilo přerozdělení části hmoty těžkých atomů na vodíkové atomy na ně navázané. Pro provedení produkčních MD simulací byly použity programovatelné GPU NVIDIA GTX-580 osazené 480 jádry. Výpočty byly prováděny rychlostí cca. 25 ns/den. Celková délka jednotlivých běhů se pohybovala v rozsahu od 100 do 200 ns.

Data byla ukládána každých 100 ps. Trajektorie byly analyzovány prostřednictvím softwarových balíků CHIMERA 1.5.3, VMD 1.9, Curves+ a PTRAJ ze softwarového balíku AMBER10 [125-127, 147]. Softwarový balík ICM Molsoft 3.7 byl použit na vytvoření obrázků.

Výsledky MD simulací

Zde prezentované MD simulace lze rozdělit do několika skupin. Jednak byla testována a porovnávána různá silová pole AMBER (MD simulace č. 1-3 a 4-6 - viz. Tab. 1). Kromě toho byly porovnávány výsledky MD simulací komplexů s protonovaným a neprotonovaným reziduem Glu131 (MD simulace č. 1-3 a 4-6). Další skupinu tvořily MD simulace zaměřené na zkoumání dopadu bodových mutací RNázy L na vazbu 2-5A k enzymu (MD simulace č. 7-12, konkrétně MD č. 7, 8 - R155A; MD č. 9 - F126A; MD č. 10 - W60A; MD č. 11, 12 - K89A). Nakonec bylo zkoumáno, zda lze použitím MD běhů v řádu stovek nanosekund tozlišit u analog 2-5A agonisty RNázy L (MD simulace č. 13-15 s ligandy pB₂₅A₂₅A, pA₂₅A₂₅B, pB₂₅A₂₅B, kde B obsahují 3'-O,4'-C-alkylenovou spojku) od neaktivních analog (MD simulace č. 16-19 s ligandy pA₂₅B₂₅A, pB₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅B) [179].

Pro snazší orientaci je v následujícím textu zavedeno značení pro trvalou konformační změnu znak -> (např. syn -> anti). Pro reverzibilní konformační změnu znak <->.

Interakce RNázy L s 25A a konformační preference 25A jsou podrobněji popsány u první MD simulace, kde jsou definovány sledované parametry. Dále pak u druhé MD simulace, protože simulovaný systém zde zůstal nejbližší počátečnímu uspořádání z krystalu a v MD simulaci č. 10 s RNázou L_E131H_W60A, protože tam 25A přešla do šroubovicovitého uspořádání, ve kterém se zřejmě nachází před a na samém počátku interakce s RNázou L.

01/ RNáza_L_E131H; pA₂₅A₂₅A; FF_94; T = 310 K; t < 200 ns

V MD simulaci č. 1 zůstala molekula pA₂₅A₂₅A stabilně navázána na enzym RNáza L_E131H a to i přes konformační změny, které způsobil trvalý přechod báze třetího AMP z konformace **syn** do konformace **high-anti**.

V molekule pA₂₅A₂₅A, která se nacházela v částečně rozvinuté konformaci, byl pozorován stacking bází prvního AMP (Ade1) a druhého AMP (Ade2). Báze třetího

AMP (Ade3) se do vzájemných interakcí bází nezapojila, namísto toho se realizovala π - π interakce s postranním řetězcem rezidua Trp60. Báze Ade1 obdobně interagovala s postranním řetězcem rezidua Phe126. Vodíkové vazby mezi protonovaným reziduem Glu131 a Ade1 (vazby Glu131_O1-H62_Ade1 a Glu131_H2-N1_Ade1) byly stabilní po celých 200 ns MD běhu s výjimkou krátkých reverzibilních rozpadů. Stejně chování bylo pozorováno u interakcí mezi Ade2 a postranním řetězcem rezidua Tyr135 (vazba Tyr135_H-N1_Ade2). Naproti tomu systém vodíkových vazeb mezi Ade3 a postranními řetězci reziduí Gln68 (Gln68_O1-H62_Ade3 a Gln68_H2-N1_Ade3) a Asn65 (Asn65_O1-H61_Ade3) byl rozpadlý již od počátku MD simulace v důsledku konformačního přechodu Ade3 (obrázek 4.20. MD č. 1 Ade3).

Mezi 5'-fosfátovou skupinou prvního AMP (Phos1) a postranními řetězci reziduí Arg155 a Lys166 byly zformovány solné můstky. U 5'-fosfátové skupiny druhého AMP (Phos2), nebyla zaznamenána dlouhodobá přímá interakce s RNázou L. Jedinou výjimkou byly krátkodobé kontakty Phos2 s reziduem Tyr124 v okamžicích narušení interakcí mezi Phos1 a reziduem Arg155, které nastávaly po konformační změně Ade3. Nakonec si rezidua Lys166 a Tyr124 vyměnili své vazebné protějšky (Phos1 a Phos2). 5'-fosfátová skupina třetího AMP (Phos3) zformovala solný můstek s postranním řetězcem rezidua Lys89.

Ade1 byl v konformaci **high-anti** (odpovídající hodnota N-glykosidického torzního úhlu O4'-C1'-N9-C4 je cca. -90°) a Ade2 v konformaci **anti** (cca. 180°). U Ade3 byl zaznamenán konformační přechod z konformace **syn** (cca. 60°) do konformace **high-anti** (obrázek 4.20. MD č. 1 Glycosidic). Puckering prvních dvou cukrů byl stabilně **C3'-endo**, zatímco u třetí ribózy proběhla konformační změna **C4'-exo/O4'-endo** -> **C2'-endo**. U atomů C4'-C5' pojících ribózu a internukleotidovou spojku byly zaznamenány následující konformace (kvantifikované prostřednictvím torzního úhlu C3'-C4'-C5'-O5'): **trans** -> **gauche** (Ade1), **-gauche** (Ade2), **-gauche** -> **gauche** (Ade3). Internukleotidová spojka Ade1-Ade2 byla trvale v konformaci **-g-g** (vyjádřeno prostřednictvím konformačních preferencí torzních úhlů C5'-O5'-P-O2' a O5'-P-O2'-C2'). U Internukleotidové spojky Ade2-Ade3 došlo ke konformačnímu přechodu **gg** -> **-g-g**.

02/ RNáza_L_E131H; pA₂₅A₂₅A; FF_99SB; T = 310 K; t = 200 ns

V MD simulaci č. 2 zůstala molekula pA₂₅A₂₅A stabilně (obrázek 4.20. MD č. 2) navázána k povrchu RNázy L_E131H v původní poloze (obrázek 4.18.).

Molekula pA₂₅A₂₅A přitom zaujímal částečně rozvinutou konformaci s π - π stackingem pouze mezi bázemi Ade1 a Ade2. Báze Ade3 stackovala s postranním řetězcem rezidua Trp60. Báze Ade1 takto interagovala s postranním řetězcem rezidua Phe126. Báze Ade1 byla ve vazebném místě stabilizována prostřednictvím vodíkových vazeb, které ji po celou 200 ns MD simulaci poutaly k protonovanému reziduu Glu131. Stejně tak i báze Ade2 interagovala s postranním řetězcem rezidua Tyr135 prostřednictvím vodíkové vazby. Báze Ade3 byla stabilizována sítí vodíkových vazeb s účastí postranních řetězců reziduí Gln68 a Asn65.

Solné můstky zformované mezi Phos1 a postranními řetězci reziduí Arg155 a Lys166 byly stabilní po celou dobu MD simulace. Nebyl pozorován podstatný

kontakt mezi Phos2 a postranním řetězcem rezidua Tyr124. Phos3 zformovala stabilní solný můstek s postranním řetězcem rezidua Lys89 a vodíkovou vazbu s postranním řetězcem rezidua Trp60.

Ade1 byl v **high-anti** a Ade2 v **anti** konformaci. Ade3 setrval v konformaci **syn**. Puckering prvního a druhého cukru byl **C3'-endo**, kdežto u třetí ribózy **C4'-exo/O4'-endo**. Torzní úhly C3'-C4'-C5'-O5' se nacházely v konformacích **trans** (Ade1), **-gauche** (Ade2), **-gauche** (Ade3). Internukleotidová spojka Ade1-Ade2 byla trvale v konformaci **-g-g**, spojka Ade2-Ade3 preferovala konformaci **gg**.

03/ RNáza_L_E131H; pA₂₅A₂₅A; FF_Olo; T = 310 K; t = 200 ns

V MD simulaci č. 3 zůstala molekula pA₂₅A₂₅A navázána k povrchu RNázy L_E131H v konformaci blízké počátečnímu uspořádání. Báze Ade3 stackovala s postranním řetězcem Trp60 v poněkud vychýlené pozici, která byla způsobena rozpadem systému vodíkových vazeb mezi bází Ade3 a postranními řetězci reziduí Gln68 a Asn65 (obrázek 4.20. MD č. 3 Ade3).

04/ RNáza_L_E131; pA₂₅A₂₅A; FF_94; T = 290 K; t < 200 ns

05/ RNáza_L_E131; pA₂₅A₂₅A; FF_99SB; T = 290 K; t = 200 ns

06/ RNáza_L_E131; pA₂₅A₂₅A; FF_Olo; T = 290 K; t = 200 ns

V MD simulaci č. 4 zůstal pouze Ade3 navázán v původní vazebné pozici na povrchu RNázy L_E131 (obrázek 4.20. MD č. 4 Ade3). Báze Ade3 byla přitom ze všech bází molekuly pA₂₅A₂₅A nejvíce vzdálená od neprotonovaného rezidua Glu131. V zásadě to samé bylo pozorováno i v následujících dvou MD simulacích s neprotonovaným Glu131 (obrázek 4.20. MD č. 5 Ade3, obrázek 4.20. MD č. 6 Ade3).

Vzhledem k nestabilitě všech simulovaných systémů s neprotonovaným Glu131 (MD č. 4-6) bylo ve všech následujících MD simulacích toto residuum protonované. Bylo přitom použito silové pole AMBER_99SB, se kterým uspořádání simulovaného systému v MD simulaci č. 2 zůstalo nejbližší tomu, co bylo pozorováno v krystalové struktuře.

07/ RNáza_L_E131H_R155A; pA₂₅A₂₅A; FF_99SB; T = 290 K; t = 200 ns

V průběhu MD simulace č. 7 se molekula pA₂₅A₂₅A téměř odpoutala od povrchu RNázy_L_E131H_R155A. V původní pozici setrval pouze Ade3 nejvíce vzdálený od místa bodové mutace R155A (obrázek 4.21. MD č. 7 Ade3). Postranní řetězec rezidua Tyr124 byl prostřednictvím vodíkové vazby trvale navázán na Phos2 a vychyloval tak pA₂₅A₂₅A z původní vazebné pozice.

08/ RNáza_L_E131H_R155A; pA₂₅A₂₅A; FF_99SB; T = 310 K; t = 200 ns

Ačkoliv byl v MD simulaci č. 8 simulovaný systém naprosto stejný jako v předchozí MD simulaci, molekula pA₂₅A₂₅A tentokrát zůstala navázána na povrchu RNázy_L_E131H_R155A v téměř nezměněné pozici. Postranní řetězce reziduí Lys166 a Lys178 totiž vytvořily solné můstky s Phos1, čímž byla částečně vykompenzována absence postranního řetězce R155 (obrázek 4.21. MD č. 8 Phos1).

Navíc nebyly zaznamenány významnější interakce postranních řetězců rezidua Tyr124 s Phos2, které by vychylovaly pA₂₅A₂₅A z původní vazebné pozice.

09/ RNáza_L_E131H_F126A; pA₂₅A₂₅A; FF_99SB; T=310K; t=200 ns

V MD simulaci č. 9 se molekula pA₂₅A₂₅A téměř odpoutala od povrchu RNázy L_E131H_F126A. V původní vazebné pozici zůstal navázaný pouze Ade3. Ade1 neměl v důsledku bodové mutace obvyklého partnera (Phe126) pro stacking a vodíkové vazby s Glu131 též zanikly. Následně byla přerušena i vodíková vazba mezi bází Ade2 a postranním řetězcem rezidua Tyr135 (obrázek 4.21. MD č. 9 Ade2).

10/ RNáza_L_E131H_W60A; pA₂₅A₂₅A; FF_99SB; T = 310 K; t = 200 ns

V MD simulaci č. 10 zůstala molekula pA₂₅A₂₅A navázána k RNáze L_E131H_W60A pouze prostřednictvím Ade1 - AMP nejbližšího od bodové mutace W60A (obrázek 4.19.).

Uvolněná molekula pA₂₅A₂₅A přešla do šroubovicovité konformace, ve které docházelo k π - π stackingu všech bází Ade1-Ade2-Ade3. Stejně finální uspořádání bylo pozorováno i v následující MD simulaci. Dále byl zaznamenán π - π stacking mezi bází Ade1 a postranním řetězcem rezidua Phe126. Trp60 bylo nahrazeno alaninem, jehož postranní řetězec nemohl stackovat s bází Ade3. Báze Ade1 interagovala prostřednictvím vodíkových vazeb s postranním řetězcem protonovaného rezidua Glu131. Tyr135 interagovalo s Ade3 namísto Ade2. Zcela přerušen byl systém vodíkových vazeb, do kterého byly kromě Ade3 původně zapojeny také postranní řetězce reziduí Gln68 a Asn65 (obrázek 4.21. MD č. 10 Ade3).

Byl také přerušen solný můstek mezi Phos1 a postranním řetězcem rezidua Arg155 nikoliv však Lys166. Nebyly pozorovány žádné podstatné interakce mezi postranním řetězcem rezidua Tyr124 a Phos2. Mezi postranním řetězcem rezidua Lys89 a Phos3 byl zformován solný můstek, který se reverzibilně rozpadal (obrázek 4.21. MD č. 10 Phos3).

Všechny AMP ukončily MD simulaci v konformaci **anti**. Puckering první a druhé ribózy byl **C3'-endo/C4'-exo** a třetí ribózy **C2'-endo**. U všech C4'-C5' spojek ribóz a internukleotidových spojek byl preferován konformer **-gauche**. Internukleotidové spojky Ade1-Ade2 a Ade2-Ade3 byly v konformaci **-g-g**.

11/ RNáza_L_E131H_K89A; pA₂₅A₂₅A; FF_99SB; T = 290 K; t = 200 ns

Během prvních cca. 150 ns MD simulace č. 11 byla molekula pA₂₅A₂₅A stabilně navázána k RNáze L_E131H_K89A. Potom se ale z většiny vazebných míst odpoutala a zůstala navázána pouze prostřednictvím Ade1 - AMP nejvíce vzdáleného od bodové mutace K89A (obrázek 4.21. MD č. 11 Ade1). Báze Ade1 stackovala s postranním řetězcem rezidua Phe126 a byla stabilizována vodíkovými vazbami vytvářenými s postranním řetězcem protonovaného rezidua Glu131. Stejně jako v předchozí MD simulaci zaujala uvolněná molekula pA₂₅A₂₅A šroubovicovitou konformaci, ve které docházelo k π - π interakcím všech bází Ade1-Ade2-Ade3.

12/ RNáza_L_E131H_K89A; pA₂₅A₂₅A; FF_99SB; T = 310 K; t = 200 ns

V MD simulaci č. 12 se molekula pA₂₅A₂₅A odpoutala od enzymu RNázy L_E131H_K89A (obrázek 4.21. MD č. 12). Její koncová orientace byla kolmá vůči orientaci počáteční. V důsledku toho byl přerušen π - π stacking jak mezi bází Ade1 a postranním řetězcem Phe126, tak i mezi bází Ade3 a postranním řetězcem rezidua Trp60.

13/ RNáza_L_E131H; pA₂₅A₂₅B; FF_99SB; T = 310 K; t = 200 ns

14/ RNáza_L_E131H; pB₂₅A₂₅A; FF_99SB; T = 310 K; t = 200 ns

15/ RNáza_L_E131H; pB₂₅A₂₅B; FF_99SB; T = 310K; t = 200 ns

V případě MD simulace č. 13 se molekula pA₂₅A₂₅B stabilně vážala k povrchu enzymu RNázy L_E131H (obrázek 4.22. MD č. 13). Stejně pozorování bylo učiněno i v následujících dvou MD simulacích s pB₂₅A₂₅A a pB₂₅A₂₅B (obrázek 4.22. MD č. 14, obrázek 4.22. MD č. 15). To je v souladu s experimentem, který prokázal pozoruhodnou schopnost pA₂₅A₂₅B, pB₂₅A₂₅A a pB₂₅A₂₅B aktivovat RNázu L [179].

16/ RNáza_L_E131H; pA₂₅B₂₅A; FF_99SB; T = 310 K; t = 100 ns

V MD simulaci č. 16 byla molekula pA₂₅B₂₅A navázána k RNáze L_E131H po téměř celých 100 ns. To se změnilo v samém závěru MD simulace, aby nakonec pouze Ade3 zůstal navázaný ve svém původním vazebném místě (obrázek 4.22. MD č. 16 Ade3). To je v souladu s experimentálně prokázanou malou schopností molekuly pA₂₅B₂₅A aktivovat RNázu L [179].

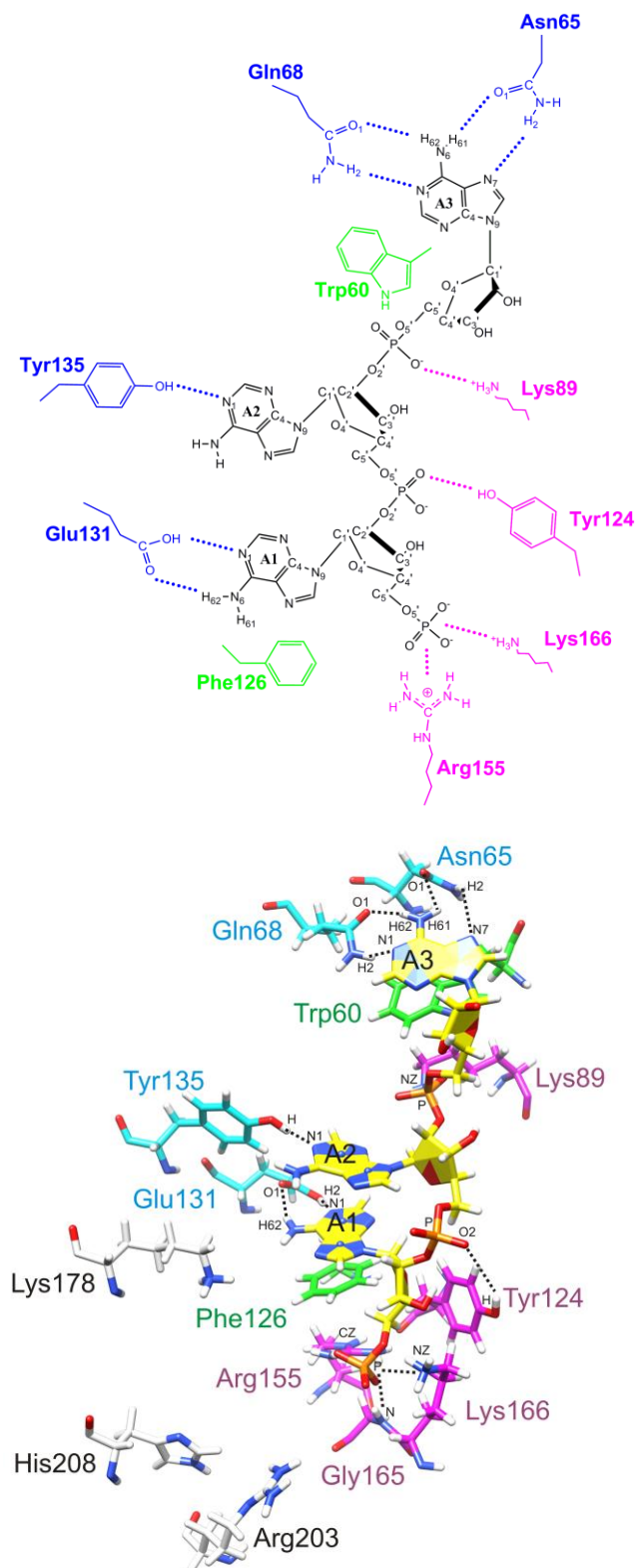
17/ RNáza_L_E131H; pA₂₅B₂₅B; FF_99SB; T = 310 K; t = 100 ns

V MD simulaci č. 17 byla molekula pA₂₅B₂₅B navázána k RNáze L_E131H. Nedochovalo však k π - π interakci mezi bázemi Ade1 a Ade2 v důsledku konformačních změn u druhého AMP. Internukleotidová spojka Ade1-Ade2 přitom přecházela mezi konformacemi -gg -> -gt -> -g-g (obrázek 4.22. MD č. 17 PL 1-2). Druhá internukleotidová spojka Ade2-Ade3 pak g-g -> -gg. Konformační změny Ade2 a obou internukleotidových spojek, které ho obklopují, by nepochybně znemožnily vazbu Tyr312 z druhého monomeru RNázy L k Phos2, což by vysvětlovalo neschopnost molekuly pA₂₅B₂₅B aktivovat RNázu L [179].

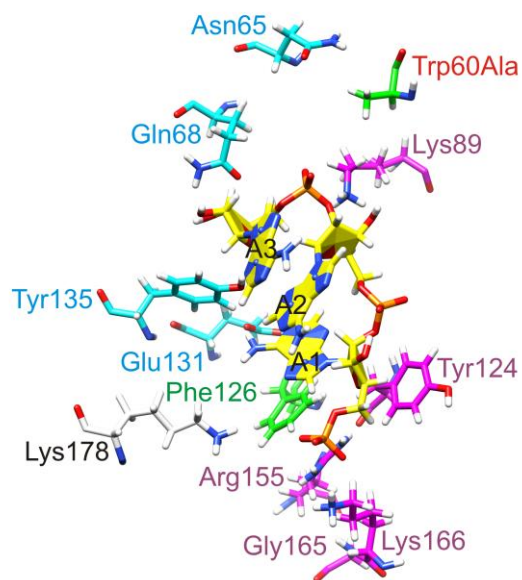
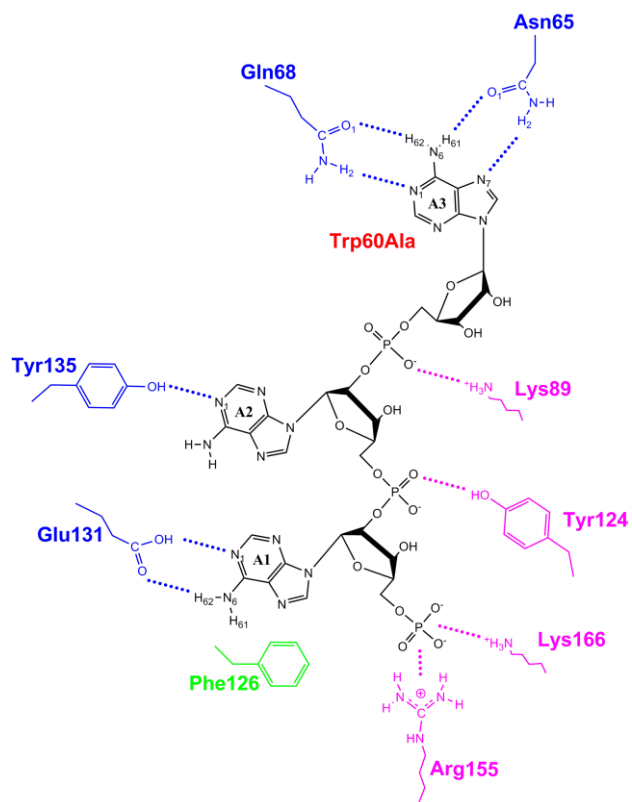
18/ RNáza_L_E131H; pB₂₅B₂₅A; FF_99SB; T = 310 K; t = 100 ns

19/ RNáza_L_E131H; pB₂₅B₂₅B; FF_99SB; T = 310 K; t = 100 ns

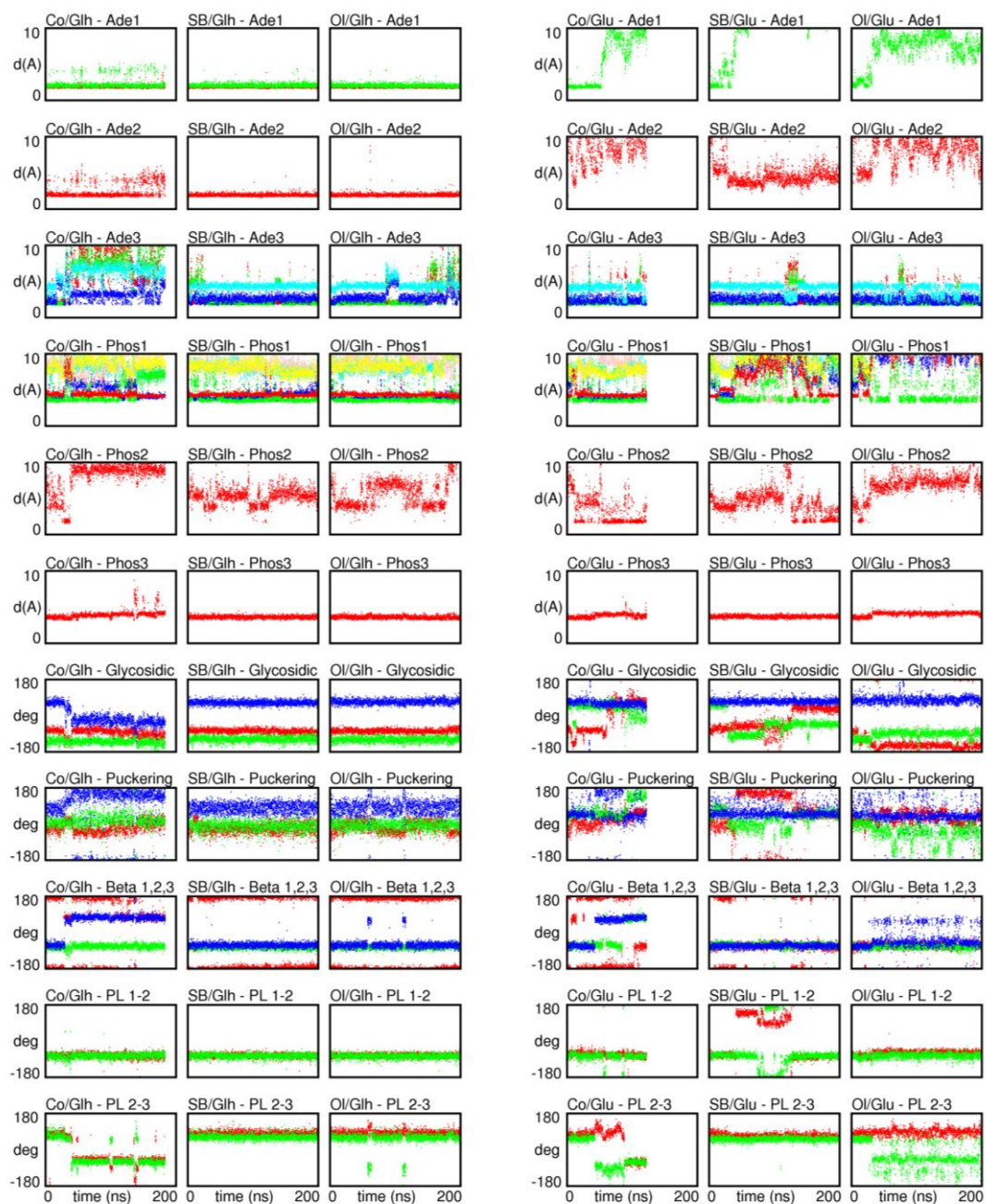
Molekuly pB₂₅B₂₅A a pB₂₅B₂₅B byly na konci MD simulací č. 18 a 19 navázané k povrchu enzymu RNázy L_E131H pouze prostřednictvím Ade3. To vysvětluje jejich neschopnost aktivovat RNázu L [179].



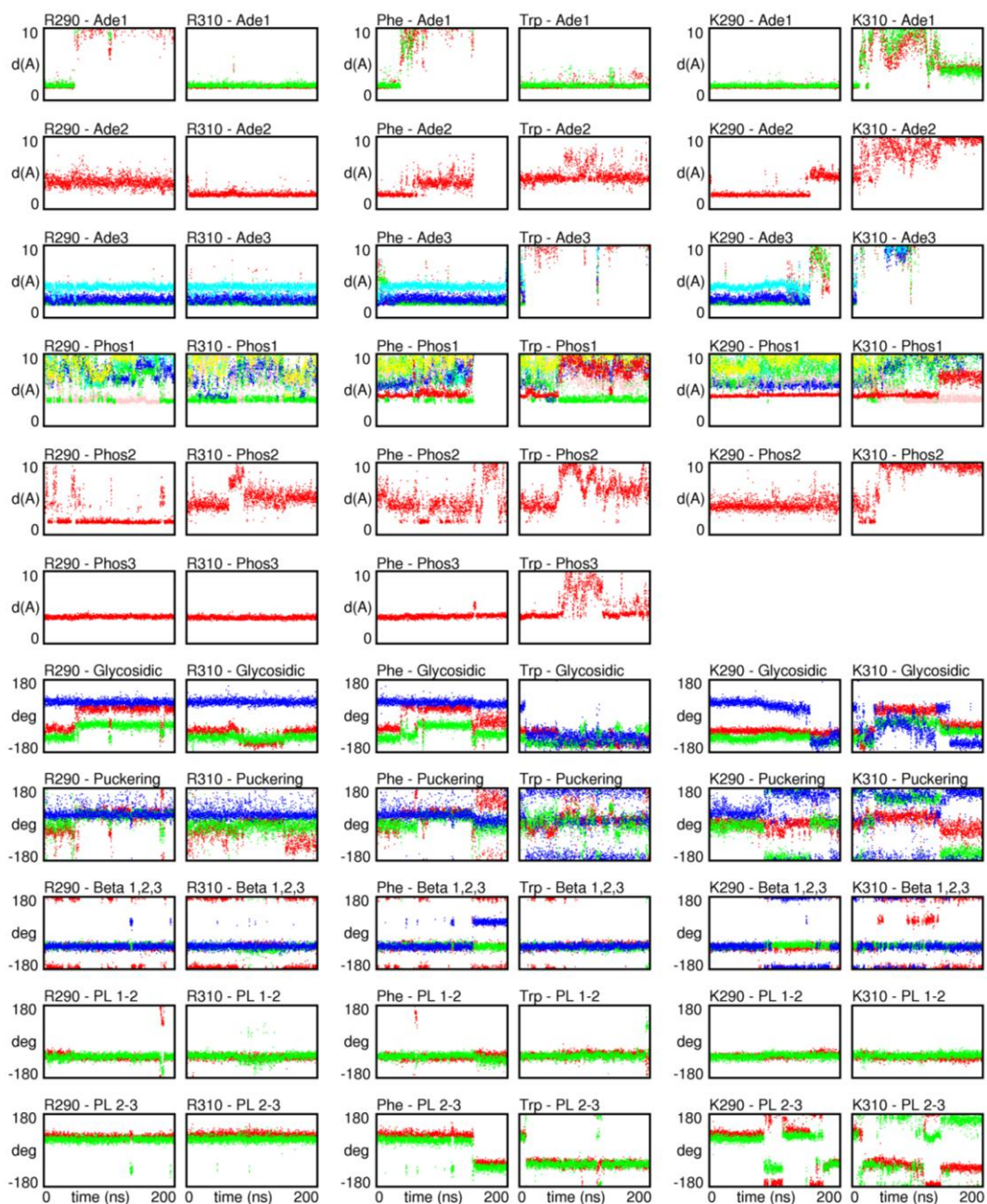
Obrázek 4.18.: Molekula pA₂₅A₂₅A navázaná k povrchu RNázy L_E131H během MD simulace č.2 způsobem [180], který odpovídá tomu, co bylo zaznamenáno ve výchozí krystalové struktuře.



Obrázek 4.19.: Molekula pA₂₅A₂₅A změnila v MD simulaci č. 10 s mutantní RNaseL_E131H_W60A svou konformaci tak, že pouze Ade1 zůstalo ve své původní vazebné pozici [180].



Obrázek 4.20.: Výsledky MD simulací č. 1-6 ve sloupcích grafů zleva doprava. Časový vývoj různých strukturálních parametrů v řádcích grafů shora dolů. Vazba **Ade1** prostřednictvím vodíkových vazeb k protonovanému/neprotonovanému Glu131: Glu131_O-H62_Ade1 (červená čára) a Glu131_OH-N1_Ade1 (zelená čára). Kontakty **Ade2** s postranním řetězcem Tyr135: Tyr135_OH-N1_Ade2 (červená čára). Sítí vodíkových vazeb zahrnujících **Ade3** a postranní řetězce Gln68: Gln68_O1-H62_Ade3 (červená čára) a Gln68_H2-N1_Ade3 (zelená čára) a Asn65: Asn65_O1-H61_Ade3 (modrá čára). Postranní řetězce Arg155 (červená čára) a Lys166 (zelená čára) formují solné můstky s 5'-fosfátovou skupinou prvního AMP (**Phos1**). 5'-fosfátová skupina druhého AMP (**Phos2**) formovala příležitostně vodíkové vazby s postranním řetězcem Tyr124 (červená čára). 5'-fosfátová skupina třetího AMP (**Phos3**) formovala příležitostně solné můstky s postranním řetězcem Lys89 (červená čára). Hodnoty **glykosidického** torzního úhlu ($O4'-C1'-N9-C4$) pro Ade1 (červená čára), Ade2 (zelená čára), Ade3 (modrá čára). **Puckering** ribózy (Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace segmentů $C4'-C5'$ spojujících ribózy a internukleotidové spojky vyjádřené prostřednictvím torzního úhlu **beta** ($C3'-C4'-C5'-O5'$): Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace fosfodiesterových internukleotidových spojek Ade1-Ade2 a Ade2-Ade3 (tj. **PL 1-2** and **PL 2-3**) vyjádřené prostřednictvím konformací torzních úhlů $C5'-O5'-P-O2'$ (červená čára) a $O5'-P-O2'-C2'$ (zelená čára) [180].



Obrázek 4.21.: Výsledky MD simulací č. 7-12 ve sloupcích grafů zleva doprava. Časový vývoj různých strukturálních parametrů v řádcích grafů shora dolů. Vazba **Ade1** prostřednictvím vodíkových vazeb k protonovanému/neprotonovanému Glu131: Glu131_O-H62_Ade1 (červená čára) a Glu131_OH-N1_Ade1 (zelená čára). Kontakty **Ade2** s postranním řetězcem Tyr135: Tyr135_OH-N1_Ade2 (červená čára). Sít vodíkových vazeb zahrnujících **Ade3** a postranní řetězce Gln68: Gln68_O1-H62_Ade3 (červená čára) a Gln68_H2-N1_Ade3 (zelená čára) a Asn65: Asn65_O1-H61_Ade3 (modrá čára). Postranní řetězce Arg155 (červená čára) a Lys166 (zelená čára) formují solné můstky s 5'-fosfátovou skupinou prvního AMP (**Phos1**). 5'-fosfátová skupina druhého AMP (**Phos2**) formovala příležitostně vodíkové vazby s postranním řetězcem Tyr124 (červená čára). 5'-fosfátová skupina třetího AMP (**Phos3**) formovala příležitostně solné můstky s postranním řetězcem Lys89 (červená čára). Hodnoty **glykosidického** torzního úhlu (O4'-C1'-N9-C4) pro Ade1 (červená čára), Ade2 (zelená čára), Ade3 (modrá čára). **Puckering** ribózy (Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace segmentů C4'-C5' spojujících ribózy a internukleotidové spojky vyjádřené prostřednictvím torzního úhlu **beta** (C3'-C4'-C5'-O5'): Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace fosfodiesterových internukleotidových spojek Ade1-Ade2 a Ade2-Ade3 (tj. **PL 1-2** and **PL 2-3**) vyjádřené prostřednictvím konformací torzních úhlů C5'-O5'-P-O2' (červená čára) a O5'-P-O2'-C2' (zelená čára) [180].



Obrázek 4.22.: Výsledky MD simulací č. 13-19 ve sloupcích grafů zleva doprava. Časový vývoj různých strukturálních parametrů v řádcích grafů shora dolů. Vazba **Ade1** prostřednictvím vodíkových vazeb k protonovanému/neprotonovanému Glu131: Glu131_O-H62_Ade1 (červená čára) a Glu131_OH-N1_Ade1 (zelená čára). Kontakty **Ade2** s postranním řetězcem Tyr135: Tyr135_OH-N1_Ade2 (červená čára). Sít vodíkových vazeb zahrnujících **Ade3** a postranní řetězce Gln68: Gln68_O1-H62_Ade3 (červená čára) a Gln68_H2-N1_Ade3 (zelená čára) a Asn65: Asn65_O1-H61_Ade3 (modrá čára). Postranní řetězce Arg155 (červená čára) a Lys166 (zelená čára) formují solné můstky s 5'-fosfátovou skupinou prvního AMP (**Phos1**). 5'-fosfátová skupina druhého AMP (**Phos2**) formovala příležitostně vodíkové vazby s postranním řetězcem Tyr124 (červená čára). 5'-fosfátová skupina třetího AMP (**Phos3**) formovala příležitostně solné můstky s postranním řetězcem Lys89 (červená čára). Hodnoty **glykosidického** torzního úhlu ($O4'-C1'-N9-C4$) pro Ade1 (červená čára), Ade2 (zelená čára), Ade3 (modrá čára). **Puckering** ribózy (Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace segmentů $C4'-C5'$ spojujících ribózy a internukleotidové spojky vyjádřené prostřednictvím torzního úhlu **beta** ($C3'-C4'-C5'-O5'$): Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace fosfodiesterových internukleotidových spojek Ade1-Ade2 a Ade2-Ade3 (tj. **PL 1-2** and **PL 2-3**) vyjádřené prostřednictvím konformací torzních úhlů $C5'-O5'-P-O2'$ (červená čára) a $O5'-P-O2'-C2'$ (zelená čára) [180].

	RNaseL	RNaseL	25A	FF	Temp	Length	EC50 [x]	Crel [x]
01	E131 H	-	pA ₂₅ A ₂₅ A	94	310K	<200ns		
02	E131 H	-	pA ₂₅ A ₂₅ A	99SB	310K	200ns		
03	E131 H	-	pA ₂₅ A ₂₅ A	Olo	310K	200ns		
04	E131	-	pA ₂₅ A ₂₅ A	94	290K	<200ns		
05	E131	-	pA ₂₅ A ₂₅ A	99SB	290K	200ns		
06	E131	-	pA ₂₅ A ₂₅ A	Olo	290K	200ns		
07	E131 H	R155A	pA ₂₅ A ₂₅ A	99SB	290K	200ns		
08	E131 H	R155A	pA ₂₅ A ₂₅ A	99SB	310K	200ns		
09	E131 H	F126A	pA ₂₅ A ₂₅ A	99SB	310K	200ns		
10	E131 H	W60A	pA ₂₅ A ₂₅ A	99SB	310K	200ns		
11	E131 H	K89A	pA ₂₅ A ₂₅ A	99SB	290K	200ns		
12	E131 H	K89A	pA ₂₅ A ₂₅ A	99SB	310K	200ns	3.5	1.0
13	E131 H	-	p B ₂₅ A ₂₅ A	99SB	310K	200ns	77	22
14	E131 H	-	pA ₂₅ A ₂₅ B	99SB	310K	200ns	4.9	1.4
15	E131 H	-	p B ₂₅ A ₂₅ B	99SB	310K	200ns	130	37
16	E131 H	-	pA ₂₅ B ₂₅ A	99SB	310K	100ns	530	150
17	E131 H	-	p B ₂₅ B ₂₅ A	99SB	310K	100ns	> 1000	> 300
18	E131 H	-	pA ₂₅ B ₂₅ B	99SB	310K	100ns	> 1000	> 300
19	E131 H	-	p B ₂₅ B ₂₅ B	99SB	310K	100ns	> 1000	> 300

Tab.1.: Knihovna MD simulací použitých komplexů s HS RNase L [180].

Shrnutí

V provedených MD simulacích bylo testováno několik verzí silového pole AMBER (MD simulace č. 1-3 a 4-6 - viz. Tab. 1), které prochází opakovanými revizemi [87, 146]. Zejména u nukleových kyselin je totiž obtížné správně určit a vybalancovat všechny silové konstanty. Například ty, které ovlivňují konformační preference u N-glykosidického torzního úhlu, který může nabývat hodnot **anti**, **high-anti**, **syn**. Přičemž u trimeru 2-5A se shodou okolností realizují všechny tyto konformery. Protože v MD běhu č. 2 (obrázek 4.18.) byla konformace 2-5A a způsob jeho interakce s RNázou L nejbližší tomu, co bylo zachyceno ve výchozí krystalové struktuře 1wdy [68], bylo v následných MD simulacích s mutantní RNázou L (MD simulace č. 7-12 - Tab. 1) a analogy 2-5A (MD simulace č. 13-19 - Tab. 1) použito silové pole 99SB [181].

Další otevřenou otázkou bylo, zda v MD simulacích použít protonované nebo neprotonované reziduum Glu131. Studium krystalové struktury 1wdy [68] naznačilo, že vzhledem ke vzdálenostem mezi atomy Ade1 a Glu131 by reziduum Glu131 mělo být protonované a tvořit dvě vodíkové vazby s Ade1. To bylo následně potvrzeno v MD simulacích (MD simulace č. 1-3 vs. 4-6 - Tab. 1), které prokázaly nestabilitu komplexu RNázy L a 2-5A při použití neprotonovaného rezidua Glu131. Proto v následujících MD simulacích s mutantní RNázou L (MD simulace č. 7-12 - Tab. 1) a analogy 2-5A (MD simulace č. 13-19 - Tab. 1) bylo použito protonované Glu131.

Další série MD simulací byla provedena s cílem zobrazit nežádoucí vliv bodových mutací RNázy L na vazbu 2-5A. Celkem bylo provedeno šest MD simulací (viz. Tab. 1: R155A - MD č. 7-8; F126A - MD č. 9; W60A - MD č. 10; K89A - MD č. 11-12) o délce 200 ns. Markantní dopad bodových mutací na vazbu 25A se projevil u všech studovaných struktur, pouze u R155A se tak stalo pouze v jedné ze dvou realizovaných MD simulací. Bylo to způsobeno tím, že ztráta kontaktu mezi reziduem Arg155 a terminální fosfátovou skupinou 2-5A může být krátkodobě a do jisté míry kompenzována interakcemi s flexibilními postranními řetězci reziduí Lys166 a Lys178. Porovnání obou MD simulací mutantu R155A určilo jako rozhodující faktor pro (ne)stabilitu vazby 25A skutečnost, zda byl či nebyl na počátku MD běhu Ade2 navázán na Tyr135. Tato interakce může být narušena kontakty mezi reziduem Tyr124 a Phos2. I v některých dalších MD simulacích se reziduum Tyr124 ukázalo jako stabilizátor struktury 2-5A vychýlené ze své původní vazebné pozice. Tato skutečnost by měla být brána do úvahy při navrhování chemicky modifikovaných 2-5A agonistů RNázy L.

Poslední série MD simulací byla provedena na analozích 2-5A vytvořených pomocí „bridging“ strategie zavedením 3'-O,4'-C-alkylenové spojky do adenosinových jednotek. Cílem bylo posoudit schopnost MD simulací rozlišit na časové škále stovek nanosekund agonisty RNázy L (viz. Tab. 1: MD simulace č. 13-15 - pB₂₅A₂₅A, pA₂₅A₂₅B, pB₂₅A₂₅B) a inaktivní analoga (viz. Tab.1: MD simulace č. 16-19: pA₂₅B₂₅A, pB₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅B) [179]. Což se podařilo. Agonisté RNázy L se v průběhu 200 ns MD simulací stabilně vážaly k povrchu RNázy L. Pouze báze Ade2 z hlediska interakce s Tyr135 občas oscillovala mezi dvěma pozicemi. Nicméně i v tomto případě převládalo obvyklé uspořádání. Naproti tomu u neaktivních analog 2-5A byly pozorovány zásadní poruchy v jejich interakcích s RNázou L. A to již v průběhu 100 ns MD běhů.

Tímto provedené MD simulace naplnily očekávání, že by mohly být vhodným nástrojem pro návrh a optimalizaci struktur agonistů RNázy L. V literatuře se přitom nachází velké množství experimentálních dat čekajících na interpretaci. V minulosti bylo vysyntetizováno a následně *in vitro* testováno mnoho analog 2-5A s modifikovanými bázemi, ribózami, internukleotidovými spojkami a 5'-fosfátovými skupinami. Kromě fosforothioátovaných 2-5A však doposud nebyl nalezen žádný další analog dostatečně odolný vůči degradaci buněčnými enzymy a schopný *in vivo* aktivovat RNázu L [179, 182].

5. Závěr

RdRp

RNA-dependentní-RNA polymeráza (RdRp) replikuje virovou RNA. Děje se tak v tzv. binukleárním aktivním místě, kde se dva hořčikové ionty váží k několika acidickým aminokyselinám. Hořčikové ionty stabilizují O3' konec syntetizovaného vlákna RNA a fosfátové skupiny přístupujícího NTP v aktivním místě RdRp a v konečném důsledku umožňují atak O3' konce RNA na α -fosfátovou skupinu NTP. V mnoha přípravných MD simulacích s klasicky parametrizovanými nepolarizovatelnými silovými poli byla opakovaně zaznamenána nestabilita binukleárního aktivního místa RdRp v důsledku nadměrné repulze mezi oběma hořčikovými ionty. Východiskem se ukázalo použití tzv. sedmibodové reprezentace hořčikových iontů, kdy je jejich kladný náboj přerozdělen na několik fiktivních pseudočástic. Stabilizace aktivního místa RdRp otevřela cestu k realistickým výpočtům. Další otázkou bylo, zda použití výkonných grafických karet pro molekulárně dynamické výpočty, jež dovolilo získat až mikrosekundové trajektorie, umožní postihnout difference ve vazbě velmi podobných substrátů do aktivního místa RdRp. Byly proto provedeny mikrosekundové MD simulace komplexů NV RdRp s přirozenými (CTP, 2dCTP) a modifikovanými (coCTP, cocCTP) substráty. Navzdory téměř stejné struktuře je CTP přirozený substrát enzymu, 2dCTP chabý substrát, coCTP terminátor syntetizovaného RNA vlákna a cocCTP klasický kompetitivní inhibitor. Molekulárně-dynamické simulace ukázaly markantně odlišné konformační preference CTP a 2dCTP vs. coCTP a cocCTP. Torzní úhel O5'-C5'-C4'-O4' v nemodifikovaných strukturách preferoval konformer **-gauche**, kdežto analogické torzní úhly C5'-O5'-C4'-O4' a C5'-O5'-C4'-C4' v modifikovaných strukturách coCTP a cocCTP jednoznačně preferovaly konformaci **+gauche**. V CTP a 2dCTP tak byl atom kyslíku O5' situován nad ribózou, kdežto v coCTP a cocCTP směřoval atom C5' mimo ribózu. Zajímavé bylo, že relaxace aktivního místa RdRp v důsledku odlišných konformačních preferencí coCTP a cocCTP trvala řádově stovky nanosekund. Z tohoto úhlu pohledu se jedna mikrosekunda jeví jako minimální délka MD trajektorií nezbytná pro dostatečnou evoluci komplexů RdRp a potenciálních inhibitorů. Dále byly sledovány interakce xCTP s postranními řetězci aminokyselin, jejichž mutace snižuje katalytickou efektivitu RdRp (Arg182 interagujícím s trifosfátovou skupinou xCTP nebo Asp247, které je spolu se Ser300 a Asn309 zapojeno do mechanismu rozpoznávání 2'-OH skupin NTP). U CTP se zhruba každých 100 ns změnila prostorová konfigurace postranního řetězce rezidua Arg182, který místo s α -fosfátovou začal interagovat s β -fosfátovou skupinou CTP. U 2dCTP byl kontakt mezi postranním řetězcem rezidua Arg182 a trifosfátovou skupinou 2dCTP přerušen. U coCTP a cocCTP poutaly solné můstky Arg182 s α -fosfonátovou skupinou po celou dobu MD simulace. U přirozených substrátů - CTP a 2dCTP - působí Arg182 jako katalytické reziduum. Otázka, kterou mohou zodpovědět pouze budoucí QM/MM výpočty je, zda intenzivní kontakt mezi α -fosfonátovou skupinou a reziduem Arg182 zaznamenaný u coCTP a cocCTP reakci spíše nebrzdí. Síť vodíkových vazeb mezi 2'-OH skupinou CTP, reziduem Ser300 a reziduem Asp247 byla vitální po celou časovou evoluci simulovaného systému. V této síti vodíkových vazeb se postupně uplatňovaly všechny možné kombinace párů donor-akceptor. V MD simulaci s 2dCTP byly interakce mezi rezidui Asp247 a Ser300 trvale přerušeny. V coCTP MD simulaci interagovala 2'-OH skupina s rezidui Ser300 a Asn309. Pouze budoucí QMMM výpočty vycházející z uspořádání

aktivního místa NV RdRP dosaženého v těchto MD simulacích mohou odpovědět na to, jak zjištěné rozdíly v rozpoznávání 2'OH skupiny ovlivňují průběh reakce. Celkově lze konstatovat, že velmi podobné nukleosidtrifosfáty se v MD simulacích vázaly do aktivního místa NV RdRp velmi rozdílným způsobem. Což plně odpovídá tomu, že CTP je přirozený substrát, 2dCTP chabý substrát, coCTP terminátor RNA řetězce a cocCTP kompetitivní inhibitor RdRp.

RNáza H

RNáza H selektivně degraduje vlákno mRNA v duplexech mRNA:DNA. Ačkoliv je tedy její funkce přesně opačná než funkce RdRp (viz. výše), aktivní místa obou enzymů jsou si velmi podobná v tom smyslu, že obsahují dva hořčikové ionty. Nicméně RdRp poutá dva substráty reakce (O3' konec RNA vlákna a fosfátové skupiny NTP), které přetváří na dva produkty reakce (prodloužené RNA vlákno a zbytkový pyrofosfát). Kdežto RNáza H štípe jedno vlákno mRNA na dvě kratší. Na první pohled se tedy zdálo, že problémy se stabilitou binukleárního aktivního místa RNázy H by nemusely být tak naléhavé, neboť oba ionty k sobě poutá jedno jediné souvislé vlákno nukleové kyseliny. Navíc hlavní oblastí zájmu byla vazba vlákna DNA k povrchu RNázy H v místě relativně hodně vzdáleném od aktivního místa. Nicméně přípravné MD simulace s klasicky parametrizovanými hořčikovými ionty opět ukázaly, že geometrie aktivního místa je nestabilní, a že tato nestabilita má vliv i na vazbu DNA vlákna. Ve štípané spojce rA5-rC6 (která je v aktivním místě vetknuta mezi oba hořčikové ionty) byla proto použita polarizovaná fosfátová skupina. Standardní parametrizace hořčikového iontu vede v MD simulacích také k podcenění vzdáleností iont-ligand. Proto byl VdW rádius R^* pro Mg^{2+} změněn z 0.7926 Å na 1.1 Å. Díky výše zmíněným korekcím byly vzdálenosti MgA–MgB v aktivním místě RNázy H stabilní během MD simulací o délce 200 ns. Referenční 1000 ns MD simulace se sedmibodovou reprezentací hořčikových iontů přitom vedla k téměř shodným výsledkům. Další MD simulace ukázaly, že DNA vlákno s modifikacemi internukleotidových spojek typu 3'-O-P-CH₂-O-5' (3pc5) bylo v 200 ns MD schopné vytvořit stabilní komplex s RNA vláknem a HS RNázou H - na rozdíl od DNA vlákna s modifikacemi typu 3'-O-CH₂-P-O-5' (3cp5). Vyšší afinita 3pc5 DNA k HS RNáze H vysvětluje schopnost těchto oligonukleotidů vyvolat enzymatické štěpení RNA (na rozdíl od 3cp5 DNA) pozorované v experimentech.

Argonaute

V roce 1998 byl náhodně objeven mechanismus tzv. RNA interference, kdy dodání dvoušroubovicovité RNA do buněk vede k inhibici homologních genů. Postupně se zjistilo, že dsRNA je v buňce štěpena na krátké duplexy siRNA. Jedno ze siRNA vláken je posléze ukotveno k enzymu Argonaute. Tento komplex pak na základě komplementarity bází váže vlákna mRNA, která jsou enzymem Argonaute štěpena. Buňky se takto brání proti virové nákaze a prostřednictvím endogenních miRNA molekul regulují expresi genetické informace většiny svých genů. Navzdory obrovskému biologickému významu enzymu Argonaute nejsou krystalové struktury jeho komplexů s nukleovými kyselinami příliš kvalitní. Donedávna se díky tomu mělo za to, že v jeho aktivním místě je pouze katalytická triáda acidických reziduí poutajících hořčikové ionty. Teprve zcela nedávno se objevily spekulace o tom, že v aktivním místě enzymu Argonaute se ve skutečnosti nachází obvyklá katalytická tetráda, neboť čtvrté Glu reziduum je umístěné na velmi mobilní smyčce a díky tomu

se může nacházet jak v aktivním místě tak i velmi daleko od něj. Enzym Argonaute díky tomu může působit velmi specificky a například mRNA nedokonale komplementární k siRNA molekule neštěpit, ale pouze dočasně blokovat atd. Cílem zde prezentované MD simulace bylo potvrdit, že příslušné Glu reziduum se může stabilně vázat do aktivního místa enzymu Argonaute způsobem obvyklým (tj. analogickým tomu, který pozorujeme u RNázy H). Zároveň tato molekulárně-dynamická simulace posloužila i k ověření toho, že nová parametrizace aktivního místa vyvinutá původně pro RNázu H se stabilně předzformovaným aktivním místem stabilizuje i podstatně labilnější binukleární aktivní místo enzymu Argonaute. Pro simulovaný systém TtAGO v komplexu s nukleovou kyselinou byla proto provedena MD simulace trvající 500 ns. Uspořádání aktivního místa bylo po celou dobu MD simulace stabilní. Postranní řetězec Glu512 se v aktivním místě TtAGO vázal jak k MgB, tak i k 2'-OH skupině nukleotidu sousedícího se štípanou fosfátovou skupinou (přesně stejně jako analogické reziduum Glu186 v HS RNáze H). Zajímavé je, že atomy hlavního řetězce proteinu u Glu512 přitom přímo interagují s fosfátovou skupinou siRNA na pozici 13-14. siRNA molekuly s modifikovanou páteří tak mohou zcela přímočaře ovlivňovat aktivní místo AGO. Připomeňme, že u RNázy H se reziduum Glu186 (analogické Glu512 u TtAGO) nachází pět pozic od konzervovaného rezidua Thr181 situovaného v tzv. fosfátové vazebné kapse pro DNA). V případě suboptimální vazby modifikovaných internukleotidových spojek AO do této fosfátové vazebné kapsy jsou proto od Thr181 přenášeny alosterické signály přes Glu186 do aktivního místa RNázy H, které v konečném důsledku zabrání degradaci RNA vlákna. Přímočařejší způsob alosterického ovlivňování aktivního místa AGO zřejmě souvisí s podstatně většími nátoky na specifitu působení proteinů AGO ve srovnání s RNázou H. Pro kvantitativní postihnutí dopadů chemických modifikací ve struktuře siRNA přes Glu512 na geometrii aktivního místa AGO bude zřejmě potřeba dále rafinovat způsob, jakým jsou v MD simulacích reprezentována binukleární aktivní místa. Dost možná bude nezbytné použít QMMM metody. Ve zde prezentované MD simulaci totiž tvořil Asp546 (analogický Asp210 u RNázy H), monodentátní kontakt s MgB, kdežto kruciální Glu512 (jehož ekvivalentem je v HS RNáze H Glu186) kontakt bidentátní (tj. Glu512 mělo tendenci interagovat s MgB prostřednictvím obou kyslíků). Vzhledem k rozlišení krystalových struktur není možné posoudit, zda se jedná o artefakt či realitu. Způsob interakce Glu512 s MgB může přitom mít zásadní vliv na průběh reakce a dost možná může této reakci zabránit.

RNáza L

RNáza L nespécificky štěpí jednovláknovou virovou RNA. Musí být ovšem aktivována krátkými 2'-5' oligoadenyláty. Vzhledem k tomu, že RNáza L zároveň štěpí i buněčnou mRNA či rRNA může mít její stimulace i protinádorové účinky. Struktura katalytické domény RNázy L není zatím známa. MD simulace byly proto zaměřeny na ankyrinovou doménu, jejíž dimerizace po vazbě 2-5A vede k přechodu RNázy L do aktivní formy. V provedených MD simulacích bylo testováno několik verzí silového pole AMBER, které prochází opakovanými revizemi. Protože v MD běhu č. 2 byla konformace 2-5A a způsob jeho interakce s RNázou L nejbližší tomu, co bylo zachyceno ve výchozí krystalové struktuře 1wdy, bylo v následných MD simulacích s mutantní RNázou L a analogy 2-5A použito silové pole AMBER_99SB. Další otevřenou otázkou bylo, zda v MD simulacích použít protonované nebo neprotonované reziduum Glu131. Studium krystalové struktury 1wdy naznačilo, že

vzhledem ke vzdálenostem mezi atomy Ade1 a Glu131 by reziduum Glu131 mělo být protonované a tvořit dvě vodíkové vazby s Ade1. To bylo následně potvrzeno v MD simulacích, které prokázaly nestabilitu komplexu RNázy L a 2-5A při použití neprotonovaného rezidua Glu131. Proto v následujících MD simulacích s mutantní RNázou L a analogy 2-5A bylo použito protonované Glu131. Další série MD simulací byla provedena s cílem zobrazit nežádoucí vliv bodových mutací RNázy L na vazbu 2-5A. Celkem bylo provedeno šest MD simulací (2xR155A, F126A, W60A, 2xK89A) o délce 200 ns. Markantní dopad bodových mutací na vazbu 2-5A se projevil u všech studovaných struktur, pouze u R155A se tak stalo pouze v jedné ze dvou realizovaných MD simulací. Bylo to způsobeno tím, že ztráta kontaktu mezi reziduem Arg155 a terminální fosfátovou skupinou 2-5A může být krátkodobě a do jisté míry kompenzována interakcemi s flexibilními postranními řetězci reziduí Lys166 a Lys178. Poslední série MD simulací byla provedena na analozích 2-5A vytvořených pomocí „bridging“ strategie zavedením 3'-O,4'-C-alkylenové spojky do adenosinových jednotek. Cílem bylo posoudit schopnost MD simulací rozlišit na časové škále stovek nanosekund agonisty RNázy L (pB₂₅A₂₅A, pA₂₅A₂₅B, pB₂₅A₂₅B) a inaktivní analoga (pA₂₅B₂₅A, pB₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅B). Což se podařilo.

Klasické molekulárně-dynamické simulace provedené prostřednictvím výkonných grafických karet umožnily získat trajektorie o délce stovek nanosekund až jedné mikrosekundy, což umožnilo postihnout rozdíly ve vazbě i zcela nepatrně odlišně modifikovaných nukleových kyselin k enzymům RdRp, RNáza H a RNáza L. Zjištěné rozdíly přitom odpovídají experimentálním výsledkům, což otevírá prostor pro racionální návrh struktury potenciálních chemoterapeutik na bázi chemicky modifikovaných nukleových kyselin.

Seznam použité literatury

- [1] K. Maláč, I. Barvík Jr., „Substrate recognition by norovirus polymerase: microsecond molecular dynamics study“, *Journal of Computer-Aided Molecular Design*, 27(4):373-388, (2013).
- [2] K. Maláč, I. Barvík Jr., „Complex between Human RNase HI and the phosphonate-DNA/RNA duplex: Molecular dynamics study“, *Journal of Molecular Graphics and Modelling*, 44:81–90, (2013).
- [3] J. M. Berg, J. L. Tymoczko, L. Stryer, „Biochemistry“, W. H. Freeman, New York, (2002).
- [4] K. S. Murakami, S. A. Darst, „Bacterial RNA polymerases: the whole story“, *Current Opinion in Structural Biology*, 13:31-39, (2003).
- [5] K. Geszvain, R. Landick, „The structure of bacterial RNA polymerase.“ In „The Bacterial Chromosome.“, N. P. Higgins (ed.), American Society for Microbiology Press, Washington DC, 283–296, (2005).
- [6] D. G. Vassylyev, M. N. Vassylyeva, A. Perederina, T. H. Tahirov, I. Artsimovitch, „Structural basis for transcription elongation by bacterial RNA polymerase“, *Nature*, 448:157-162, (2007).
- [7] O. Laptenko, J. Lee, I. Lomakin, S. Borukhov, „Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase“, *The European Molecular Biology Organization Journal*, 22(23):6322–6334, (2003).
- [8] N. Opalka, M. Chlenov, P. Chacon, W. J. Rice, W. Wriggers, S. A. Darst, „Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase.“, *Cell*, 114(3):335-345, (2003).
- [9] S. T. Estrem, W. Ross, T. Gaal, Z. W. S. Chen, W. Niu, R. H. Ebright, R. L. Gourse, „Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase α subunit“, *Genes & Development*, 13(16):2134-47, (1999).
- [10] V. Sosunov, E. Sosunova, A. Mustaev, I. Bass, V. Nikiforov, A. Goldfarb, „Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase“, *The European Molecular Biology Organization Journal*, 22(9):2234–2244, (2003).
- [11] I. Artsimovitch, R. Landick, „Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals“, *Proceedings of the National Academy of Sciences of the United States of America*, 97(13):7090-7095, (2000).
- [12] P. Ahlquist, „RNA-Dependent RNA Polymerases, Viruses, and RNA Silencing“, *Science*, 296(5571):1270-1273, (2002).
- [13] C. C. Kao, P. Singh, D. J. Ecker, „De Novo Initiation of Viral RNA-Dependent RNA Synthesis“, *Virology*, 287(2):251-60, (2001).
- [14] B. D. Lindenbach, C. M. Rice, „Unravelling hepatitis C virus replication from genome to function“, *Nature*, 436(7053):933-938, (2005).
- [15] C. T. Ranjith-Kumar, C. C. Kao., „Biochemical Activities of the HCV NS5B RNA-Dependent RNA Polymerase“, In *The „Hepatitis C Viruses: Genomes and Molecular Biology.“*, S. L. Tan (ed.), Horizon Bioscience, Norfolk (UK), (2006).
- [16] R. R. Deore, J. W. Chern, „NS5B RNA dependent RNA polymerase inhibitors: the promising approach to treat hepatitis C virus infections.“, *Current Medicinal Chemistry*, 17(32):3806-3826, (2010).
- [17] R. T. Mosley, T. E. Edwards, E. Murakami, A. M. Lam, R. L. Grice, J. Du, M. J. Sofia, P. A. Furman, M. J. Otto, „Structure of Hepatitis C Virus Polymerase in

Complex with Primer-Template RNA.“, *Journal of Virology*, 86(12): 6503–6511, (2012).

[18] S. Bressanelli, L. Tomei, A. Roussel, I. Incitti, R. L. Vitale, M. Mathieu, R. De Francesco, F. A. Rey, „Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus“, *Proceedings of the National Academy of Sciences of the United States of America*, 96(23):13034-13039, (1999).

[19] T. Yamashita, S. Kaneko, Y. Shirota, W. Qin, T. Nomura, K. Kobayashi, S. Murakami, „RNA-dependent RNA polymerase activity of the soluble recombinant hepatitis C virus NS5B protein truncated at the C-terminal region.“, *The Journal of Biological Chemistry*, 273(25):15479-15486, (1998).

[20] A. A. van Dijk, E. V. Makeyev, D. H. Bamford, „Initiation of viral RNA-dependent RNA polymerization“, *Journal of General Virology*, 85(5):1077–1093, (2004).

[21] S. Bressanelli, L. Tomei, F. A. Rey, R. De Francesco, „Structural Analysis of the Hepatitis C Virus RNA Polymerase in Complex with Ribonucleotides.“, *Journal of Virology*, 76(7):3482-3492, (2002).

[22] P. Gong, O. B. Peersen, „Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase.“, *Proceedings of the National Academy of Sciences of the United States of America* 107(52):22505–22510, (2010).

[23] D. F. Zamyatkin, F. Parra, J. M. M. Alonso, D. A. Harki, B. R. Peterson, P. Grochulski, K. K. Ng., „Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase.“, *The Journal of Biological Chemistry*, 283(12):7705-7712, (2008).

[24] K. K. Ng, N. Pendás-Franco, J. Rojo, J. A. Boga, Á. Machín, J. M. M. Alonso, F. Parra, „Crystal structure of Norwalk virus polymerase reveals the carboxyl terminus in the active site cleft.“, *The Journal of Biological Chemistry*, 279(16):16638-16645, (2004).

[25] Á. Machín, J. M. M. Alonso, F. Parra, „Identification of the Amino Acid Residue Involved in Rabbit Hemorrhagic Disease Virus VPg Uridylylation“, *The Journal of Biological Chemistry*, 276(30):27787-27792, (2001).

[26] T. Takegami, R. J. Kuhn, C. W. Anderson, E. Wimmer, „Membrane-dependent uridylylation of the genome-linked protein VPg of poliovirus.“, *Proceedings of the National Academy of Sciences of the United States of America*, 80(24):7447–7451, (1983).

[27] B. P. Steil, D. J. Barton, „Poliovirus *cis*-Acting Replication Element-Dependent VPg Uridylylation Lowers the K_m of the Initiating Nucleoside Triphosphate for Viral RNA Replication.“, *Journal of Virology*, 82(19):9400–9408, (2008).

[28] C. Ferrer-Orta, A. Arias, R. Agudo, R. Pérez-Luque, C. Escarmís, E. Domingo, N. Verdager, „The structure of a protein primer–polymerase complex in the initiation of genome replication“, *The European Molecular Biology Organization Journal*, 25(4):880 – 888, (2006).

[29] D. Duijsings, R. Kormelink, R. Goldbach, „In vivo analysis of the TSWV cap-snatching mechanism: single base complementarity and primer length requirements.“, *The European Molecular Biology Organization Journal*, 20(10):2545-2552, (2001).

[30] A. J. Caton, S. Robertson, „Structure of the host-derived sequences present at the 5' ends of influenza virus mRNA“, *Nucleic Acids Research*, 8(12):2591-2604, (1980).

[31] M. H. Powdrill, J. A. Bernatchez, M. Götte, „Inhibitors of the Hepatitis C Virus RNA-Dependent RNA Polymerase NS5B“, *Viruses*, 2(10):2169-2195, (2010).

- [32] G. Migliaccio, J. E. Tomassini, S. S. Carroll, L. Tomei, S. Altamura, B. Bhat, L. Bartholomew, M. R. Bosserman, A. Ceccacci, L. F. Colwell, R. Cortese, R. De Francesco, A. B. Eldrup, K. L. Getty, X. S. Hou, R. L. LaFemina, S. W. Ludmerer, M. MacCoss, D. R. McMasters, M. W. Stahlhut, D. B. Olsen, D. J. Hazuda, O. A. Flores, „Characterization of resistance to non-obligate chain-terminating ribonucleoside analogs that inhibit hepatitis C virus replication in vitro.“, *The Journal of Biological Chemistry*, 278(49):49164-49170, (2003).
- [33] B. Harpstead, „HIV-1 reverse transcriptase structure, activities, and inhibitors“, *MMG 445 Basic Biotechnology eJournal*, 1(1), (2005).
- [34] H. Jonckheere, J. Anné, E. De Clercq, „The HIV-1 reverse transcription (RT) process as target for RT inhibitors.“, *Medicinal Research Reviews*, 20(2):129-154., (2000).
- [35] A. D. Frankel, J. A. T. Young, „HIV-1: fifteen proteins and an RNA.“, *Annual Review of Biochemistry*, 67:1-25., (1998).
- [36] S. G. Sarafianos, B. Marchand, K. Das, D. Himmel, M. A. Parniak, S. H. Hughes, E. Arnold, „Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition“, *Journal of Molecular Biology*, 385(3):693–713, (2009).
- [37] D. W. Rodgers, S. J. Gamblin, B. A. Harris, S. Ray, J. S. Culp, B. Hellmig, D. J. Woolf, C. Debouck, S. C. Harrison, „The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1.“, *Proceedings of the National Academy of Sciences of the United States of America*, 92(4):1222–1226, (1995).
- [38] K. Klumpp, J. Q. Hang, S. Rajendran, Y. Yang, A. Derosier, P. W. K. In, H. Overton, K. E. B. Parkes, N. Cammack, J. A. Martin, „Two-metal ion mechanism of RNA cleavage by HIV RNase H and mechanism-based design of selective HIV RNase H inhibitors“, *Nucleic Acids Research*, 31(23):6852-6859, (2003).
- [39] B. Julg, F. D. Goebel, „HIV Genetic Diversity: Any Implications for Drug Resistance?“, *Infection*, 33(4):299-301, (2005).
- [40] J. D. Reeves, R. W. Doms, „Human immunodeficiency virus type 2“, *Journal of General Virology*, 83(6):1253-1265, (2002).
- [41] <http://www.chem.qmul.ac.uk/iubmb/enzyme/>
- [42] M. Nowotny, S. A. Gaidamakov, R. J. Crouch, W. Yang, „Crystal Structures of RNase H Bound to an RNA/DNA Hybrid: Substrate Specificity and Metal-Dependent Catalysis“, *Cell*, 121(7):1005–1016, (2005).
- [43] S. J. Schultz, J. J. Champoux, „RNase H Activity: Structure, Specificity, and Function in Reverse Transcription“, *Virus Research*, 134(1-2):86–103, (2008).
- [44] M. Nowotny, S. A. Gaidamakov, R. Ghirlando, S. M. Cerritelli, R. J. Crouch, W. Yang, „Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription“, *Molecular Cell*, 28(2):264-276, (2007).
- [45] G. L. Beilhartz, M. Götte, „HIV-1 Ribonuclease H: Structure, Catalytic Mechanism and Inhibitors“, *Viruses*, 2(4):900-926, (2010).
- [46] S. M. Cerritelli, R. J. Crouch, „Ribonuclease H: the enzymes in Eukaryotes“, *Federation of European Biochemical Societies Journal*, 276(6):1494–1505, (2009).
- [47] M. P. Rychlik, H. Chon, S. M. Cerritelli, P. Klimek, R. J. Crouch, M. Nowotny, „Crystal Structures of RNase H2 in Complex with Nucleic Acid Reveal the Mechanism of RNA-DNA Junction Recognition and Cleavage“, *Molecular Cell*, 40(4):658–670, (2010).

- [48] T. Tadokoro, S. Kanaya, „Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes.“, *Federation of European Biochemical Societies Journal*, 276(6):1482-1493, (2009).
- [49] N. Dias, C. A. Stein, „Antisense oligonucleotides: basic concepts and mechanisms.“, *Molecular Cancer Therapeutics*, 1(5):347-355, (2002).
- [50] J. H. P. Chan, S. Lim, W. S. F. Wong, „Antisense oligonucleotides: from design to therapeutic application.“, *Clinical and Experimental Pharmacology and Physiology*, 33(5-6):533–540, (2006).
- [51] I. Barvík Jr., J. Štěpánek, J. Bok, „Fully solvated molecular dynamics simulations of duplexes formed by modified oligonucleotides with xylo/phosphodiesteric and xylo/phosphonate internucleoside linkages and their natural counterpart.“, *Computer Physics Communications*, 147(1-2):158-161, (2002).
- [52] I. Barvík Jr., J. Štěpánek, J. Bok, „Explicit solvent molecular dynamics simulation of duplex formed by the modified oligonucleotide with alternating phosphate/phosphonate internucleoside linkages and its natural counterpart.“, *Journal of Biomolecular Structure and Dynamics*, 19(5):863-875, (2002).
- [53] C. Ender, G. Meister, „Argonaute proteins at a glance“, *Journal of Cell Science*, 123(11):1819-1823, (2010).
- [54] J. Höck, G. Meister, „The Argonaute protein family“, *Genome Biology*, 9(2):210, (2008).
- [55] G. Hutvagner, M. J. Simard, „Argonaute proteins: key players in RNA silencing“, *Nature Reviews Molecular Cell Biology*, 9(1):22–32, (2008).
- [56] M. Ghildiyal, P. D. Zamore, „Small silencing RNAs: an expanding universe“, *Nature Reviews Genetics*, 10(2):94–108, (2009).
- [57] V. Anantharaman, E. V. Koonin, L. Aravind, „Comparative genomics and evolution of proteins involved in RNA metabolism“, *Nucleic Acids Research*, 30(7):1427-1464, (2002).
- [58] N. T. Schirle, I. J. MacRae, „The Crystal Structure of Human Argonaute2“, *Science*, 336(6084):1037-1040, (2012).
- [59] A. Boland, E. Huntzinger, S. Schmidt, E. Izaurralde, O. Weichenrieder, „Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein“, *Proceedings of the National Academy of Sciences of the United States of America*, 108(26):10466-10471, (2011).
- [60] B. Czech, G. J. Hannon, „Small RNA sorting: matchmaking for Argonautes“, *Nature Reviews Genetics*, 12:19–31, (2011).
- [61] R. W. Carthew, E. J. Sontheimer, „Origins and Mechanisms of miRNAs and siRNAs“, *Cell*, 136(4):642–655, (2009).
- [62] W. Filipowicz, „RNAi: The Nuts and Bolts of the RISC Machine“, *Cell*, 122(1):17–20, (2005).
- [63] D. P. Bartel, „MicroRNAs: Target Recognition and Regulatory Functions“, *Cell*, 136(2):215-233, (2009).
- [64] W. Saenger, „Principles of nucleic acid structure“, Springer-Verlag, New York, (1984).
- [65] Y. Han, G. Whitney, J. Donovan, A. Korennykh, „Innate Immune Messenger 2-5A Tethers Human RNase L into Active High-Order Complexes“, *Cell Reports*, 2(4):902-913, (2012).
- [66] S. L. Liang, D. Quirk, A. Zhou, „RNase L: Its Biological Roles and Regulation“, *International Union of Biochemistry and Molecular Biology Life*, 58(9):508 – 514, (2006).

- [67] A. Chakrabarti, B. K. Jha, R. H. Silverman, „New Insights into the Role of RNase L in Innate Immunity“, *Journal of Interferon & Cytokine Research*, 31(1):49-57, (2011).
- [68] N. Tanaka, M. Nakanishi, Y. Kusakabe, Y. Goto, Y. Kitade, K. T. Nakamura, „Structural basis for recognition of 2',5'-linked oligoadenylates by human ribonuclease L“, *The European Molecular Biology Organization Journal*, 23(20):3929-3938, (2004).
- [69] B. Dong, M. Niwa, P. Walter, R. H. Silverman, „Basis for regulated RNA cleavage by functional analysis of RNase L and Ire1p“, *RNA*, 7(3):361-373, (2001).
- [70] O. Pav, E. Protivinska, M. Pressova, M. Collinsova, J. Jiracek, J. Snasel, M. Masojdkova, M. Budesinsky, I. Rosenberg, „Activation of murine RNase L by isopolar 2'-phosphonate analogues of 2',5' oligoadenylates.“, *Journal of Medicinal Chemistry*, 49(13):3955-3962, (2006).
- [71] M. Nakanishi, N. Tanaka, Y. Mizutani, M. Mochizuki, Y. Ueno, K. T. Nakamura, Y. Kitade, „Functional Characterization of 2',5'-Linked Oligoadenylate Binding Determinant of Human RNase L“, *The Journal of Biological Chemistry*, 280(50):41694-41699, (2005).
- [72] <http://www.rcsb.org/pdb/home/home.do>
- [73] O. Páv, N. Panova, J. Snášel, E. Zborníková, I. Rosenberg, „Activation of human RNase L by 2'- and 5'-O-methylphosphonate-modified oligoadenylates“, *Bioorganic & Medicinal Chemistry Letters*, 22(1):181-185, (2012).
- [74] O. Páv, „Fosfonátová analoga ribonukleotidů a oligoribonukleotidů“, *Vysoká škola chemicko-technologická, Disertační práce*, (2006).
- [75] Y. N. Vorobjev, „In Silico Modeling and Conformational Mobility of DNA Duplexes“, *Molecular Biology*, 37(2):210-222, (2003).
- [76] M. Karplus, J. A. McCammon, „Molecular dynamics simulations of biomolecules“, *Nature Structural Biology*, 9(9):646-652, (2002).
- [77] T. E. Cheatham 3rd., M. A. Young, „Molecular dynamics simulation of nucleic acids: successes, limitations, and promises.“, *Biopolymers*, 56(4):232-256, (2001).
- [78] W. Wang, O. Donini, C. M. Reyes, P. A. Kollman, „Biomolecular Simulations: Recent Developments in Force Fields, Simulations of Enzyme Catalysis, Protein-Ligand, Protein-Protein, and Protein-Nucleic Acid Noncovalent Interactions“, *Annual Review of Biophysics and Biomolecular Structure*, 30:211-243, (2001).
- [79] D. L. Beveridge, K. J. McConnell, „Nucleic acids: theory and computer simulation Y2K“, *Current Opinion in Structural Biology*, 10(2):182-196, (2000).
- [80] P. Auffinger, E. Westhof, „Simulations of the molecular dynamics of nucleic acids“ *Current Opinion in Structural Biology*, 8(2):227-236, (1998).
- [81] G. Ravishanker, P. Auffinger, D. R. Langley, B. Jayaram, M. A. Young, D. L. Beveridge, „Treatment of counterions in computer simulations of DNA.“, In „Reviews in Computational Chemistry“, K. B. Lipkowitz, D. B. Boyd (eds), John Wiley & Sons, Hoboken, 11:317-327, (1997).
- [82] T. E. Cheatham III, P. A. Kollman, „Molecular Dynamics Simulations Highlight the Structural Differences among DNA:DNA, RNA:RNA, and DNA:RNA Hybrid Duplexes“ *Journal of the American Chemical Society*, 119(21):4805-4825, (1997).
- [83] Y. Duan, P. A. Kollman „Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution“, *Science*, 282(5389):740-744, (1998).
- [84] L. Skála, „Kvantová teorie molekul“, *Univerzita Karlova, Praha*, (1994).
- [85] S. J. Weiner, P. A. Kollman, D. A. Case, U. Ch. Singh, C. Ghio, G. Alagona, S. Profeta, Jr., P. Weiner, „A New Force Field for Molecular Mechanical Simulation of

Nucleic Acids and Proteins“, Journal of the American Chemical Society, 106(3):765-784, (1984).

[86] S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, „An All Atom Force Field for Simulations of Proteins and Nucleic Acids“, Journal of Computational Chemistry, 7(2):230-252, (1986).

[87] W. D. Cornell, P. Cieplak, Ch. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, „A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules“, Journal of the American Chemical Society, 117(19):5179-5197, (1995).

[88] J. Šponer, R. Burel, P. Hobza, „Interactions between Amino Groups in DNA. An Ab Initio Study and a Comparison with Empirical Potentials“, Journal of Biomolecular Structure and Dynamics, 11(6):1357-1376, (1994).

[89] J. Šponer, J. Kypr, „Theoretical Analysis of the Base Stacking in DNA: Choice of the Force Field and a Comparison with the Oligonucleotide Crystal Structures“, Journal of Biomolecular Structure and Dynamics, 11(2):277-292, (1993).

[90] P. Hobza, F. Hubálek, M. Kabelác, P. Mejzlík, J. Šponer, J. Vondrášek, „Ability of empirical potentials (AMBER, CHARMM, CVFF, OPLS, Poltev) and semi-empirical quantum chemical methods (AM1, MNDO/M, PM3) to describe H-bonding in DNA base pairs, comparison with ab initio results“, Chemical Physics Letters, 257(1):31-35, (1996).

[91] J. Šponer, J. Lesyczynski, and P. Hobza „Nature of Nucleic Acid-Base Stacking: Nonempirical ab Initio and Empirical Potential Characterization of 10 Stacked Base Dimers. Comparison of Stacked and H-Bonded Base Pairs“, The Journal of Physical Chemistry, 100(13): 5590-5596, (1996).

[92] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, „Comparison of simple potential functions for simulating liquid water“, The Journal of Chemical Physics, 79:926-935, (1983).

[93] P. Cieplak, W. D. Cornell, C. Bayly, P. A. Kollman, „Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA and proteins.“ Journal of Computational Chemistry, 16(11):1357-1377, (1995).

[94] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, „A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model.“, The Journal of Physical Chemistry, 97(40):10269–10280, (1993).

[95] K. L. Meagher, L. T. Redman, H. A. Carlson, „Development of Polyphosphate Parameters for Use with the AMBER Force Field“, Journal of Computational Chemistry, 24(9):1016–1025, (2003).

[96] A. Perez, I. Marchan, D. Svozil, J. Šponer, T. E. Cheatham III, C. A. Laughton, M. Orozco, „Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers“, Biophysical Journal, 92(11):3817–3829, (2007).

[97] D. Marx, J. Hutter „Ab Initio Molecular Dynamics: Theory and Implementation“, John von Neumann Institute for Computing Series, 3:329-477, (2000).

[98] <http://ambermd.org/>

[99] MSI „Force-field based MD simulations“, manual.

[100] D. W. Heerman, “Computer Simulations Methods in Theoretical Physics”, Springer-Verlag, New York, (1986).

[101] P. Jungwirth „Klasická a kvantová molekulová dynamika“, Praha, skripta.

- [102] I. Nezbeda, J. Kolafa, M. Kotrla, „Úvod do počítačových simulací – Metody Monte Carlo a molekulární dynamiky“, Karolinum, Praha (1998).
- [103] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, “Scalable Molecular Dynamics with NAMD”, *Journal of Computational Chemistry*, 26(16):1781–1802, (2005).
- [104] A. Brünger, C. L. Brooks III, M. Karplus, „Stochastic boundary conditions for molecular dynamics simulations of ST2 water“, *Chemical Physics Letters*, 105(5):495-500, (1984).
- [105] E. Vitásek, „Numerické metody“, Státní nakladatelství technické literatury, Praha, (1987).
- [106] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, „Numerical Recipes in C“, Cambridge University Press, Cambridge, (1992).
- [107] K. Cho, „Atomistic Simulations: Nanomaterials and Device Modeling“, Lecture notes.
- [108] M. J. Harvey, G. Giupponi, G. De Fabritiis, „ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale“, *Journal of Chemical Theory and Computation*, 5(6):1632–1639, (2009).
- [109] T. Schlick, „Molecular Modeling and Simulation - An Interdisciplinary Guide“, Springer, New York, (2002).
- [110] D. Quigley, M. I. J. Probert, „Langevin Dynamics in Constant Pressure Extended Systems“, *The Journal of Chemical Physics*, 120(24):11432-11442, (2004).
- [111] D. Quigley, „Constant Pressure Langevin Dynamics: Theory and Application to the Study of Phase Behaviour in Core-Softened Systems.“, University of York, York, (2005).
- [112] <http://homepage.univie.ac.at/franz.vesely/simsp/dx/node47.html>
- [113] P. Gibbon, G. Sutmann, „Long-range Interactions in Many-Particle Simulations“, *John von Neumann Institute for Computing Series*, 10:467-506, (2002).
- [114] C. Sagui, T. A. Darden, „Molecular dynamics simulations of biomolecules – Long-Range Electrostatic Effects“, *Annual Review of Biophysics and Biomolecular Structures*, 28:155-179, (1999).
- [115] T. Darden, D. York, L. Pedersen, „Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems“, *The Journal of Chemical Physics*, 98(12):10089-10092, (1993).
- [116] D. M. York, W. Yang, H. Lee, T. Darden, L. G. Pedersen, „Toward the Accurate Modeling of DNA: The Importance of Long-Range Electrostatics“, *Journal of the American Chemical Society*, 117(17):5001-5002, (1995).
- [117] M. J. Harvey, G. De Fabritiis, „An Implementation of the Smooth Particle Mesh Ewald Method on GPU Hardware“, *Journal of Chemical Theory and Computation*, 5(9):2371–2377, (2009).
- [118] R. Salomon-Ferrer, D. A. Case, R. C. Walker, “An overview of the Amber biomolecular simulation package”, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2):198–210, (2012).
- [119] M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R. D. Skeel, K. Schulten. „NAMD: A parallel, object-oriented molecular dynamics program.“ *International Journal of Supercomputer Applications and High Performance Computing*, 10(4):251-268, (1996).
- [120] <http://www.ks.uiuc.edu/>
- [121] R. Brunner, L. Kalé, J. Philips „Flexibility and Interoperability in a Parallel Biomolecular Dynamics code“, In M. E. Henderson, C. R. Anderson, S. L. Lyons,

“Object Oriented Methods for Interoperable Scientific and Engineering Computing: Proceedings of the 1998 SIAM Workshop”, Society for Industrial and Applied Mathematics, New York, 78-87, (1999).

[122] <http://www.multiscalelab.org>

[123] I. Buch, M. J. Harvey, T. Giorgino, D. P. Anderson, G. De Fabritiis, “High-Throughput All-Atom Molecular Dynamics Simulations Using Distributed Computing”, *Journal of Chemical Information and Modeling*, 50(3):397–403, (2010).

[124] K. A. Feenstra, B. Hess, H. J. C. Berendsen, “Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems”, *Journal of Computational Chemistry*, 20(8):786–798, (1999).

[125] R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, K. Zakrzewska, “Conformational analysis of nucleic acids revisited: Curves+”, *Nucleic Acids Research*, 37(17):5917–5929, (2009).

[126] W. Humphrey, A. Dalke, K. Schulten, „VMD – visual molecular dynamics“, *Journal of Molecular Graphics*, 14(1):33–38, (1996).

[127] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, „UCSF Chimera – a visualization system for exploratory research and Analysis“, *Journal of Computational Chemistry*, 25(13):1605–1612, (2004).

[128] M. J. Sofia, W. Chang, P. A. Furman, R. T. Mosley, B. S. Ross, „Nucleoside, nucleotide, and non nucleoside inhibitors of hepatitis C virus NS5B RNA-dependent RNA-polymerase.“, *Journal of Medicinal Chemistry*, 55(6):2481–2531, (2012).

[129] R. De Francesco, G. Migliaccio, „Challenges and successes in developing new therapies for hepatitis C.“, *Nature*, 436:953–960, (2005).

[130] D. F. Zamyatkin, F. Parra, A. Machín, P. Grochulski, K. K. Ng, „Binding of 2'-amino-2' deoxycytidine-5'-triphosphate to norovirus polymerase induces rearrangement of the active site.“, *Journal of Molecular Biology*, 390(1):10–16, (2009).

[131] J. Paeshuyse, I. Vliegen, L. Coelmont, P. Leyssen, O. Tabarrini, P. Herdewijn, H. Mittendorfer, J. Easmon, V. Cecchetti, R. Bartenschlager, G. Puerstinger, J. Neyts, „Comparative in vitro anti hepatitis c virus activities of selected series of polymerase, protease, and helicase inhibitors.“, *Antimicrobial Agents and Chemotherapy*, 52(9):3433–3437, (2008).

[132] I. Zlatev, H. Dutartre, I. Barvík, J. Neyts, B. Canard, J. J. Vasseur, K. Alvarez, F. Morvan, „Phosphoramidate dinucleosides as hepatitis C virus polymerase inhibitors.“, *Journal of Medicinal Chemistry*, 51(18):5745–5757, (2008).

[133] S. Priet, I. Zlatev, I. Barvík Jr, K. Geerts, P. Leyssen, J. Neyts, H. Dutartre, B. Canard, J. J. Vasseur, F. Morvan, K. Alvarez, „3'-Deoxy phosphoramidate dinucleosides as improved inhibitors of hepatitis C virus subgenomic replicon and NS5B polymerase activity.“, *Journal of Medicinal Chemistry*, 53(18):6608–6617, (2010).

[134] D. L. Wales, R. T. Schooley, K. A. Kaihara, J. R. Beadle, K. Y. Hostetler, „Anti-hepatitis C virus replicon activity of alkoxyalkyl esters of (S)-HPMPA and other acyclic nucleoside phosphonates.“, *Antiviral Research*, 78(2):A21, (2008).

[135] Y. Koh, J. H. Shim, J. Z. Wu, W. Zhong, Z. Hong, J. L. Girardet, „Design, synthesis, and antiviral activity of adenosine 5'-phosphonate analogues as chain terminators against hepatitis C virus.“, *Journal of Medicinal Chemistry*, 48(8):2867–2875, (2005).

- [136] C. Sheridan, „Calamitous HCV trial casts shadow over nucleoside drugs.“, *Nature Biotechnology*, 30:1015–1016, (2012).
- [137] Z. Točík, M. Buděšínský, I. Barvík Jr, I. Rosenberg, „Conformational evaluation of labeled C3'-O-P-(13)CH(2)-O-C4'' phosphonate internucleotide linkage, a phosphodiester isostere.“, *Biopolymers*, 91(7):514–529, (2009).
- [138] D. W. Gohara, S. Crotty, J. J. Arnold, J. D. Yoder, R. Andino, C. E. Cameron, „Poliovirus RNA-dependent RNA polymerase (3Dpol).“, *The Journal of Biological Chemistry*, 275(33):25523–25532, (2000).
- [139] C. T. Ranjith-Kumar, R. T. Sarisky, L. Gutshall, M. Thomson, C. C. Kao, „De Novo initiation pocket mutations have multiple effects on hepatitis C virus RNA-dependent RNA polymerase activities.“, *Journal of Virology*, 78(22):12207–12217, (2004).
- [140] S. G. Sarafianos, V. N. Pandey, N. Kaushik, M. J. Modak, „Sitedirected mutagenesis of arginine 72 of HIV-1 reverse transcriptase.“, *The Journal of Biological Chemistry*, 270(34):19729–19735, (1995).
- [141] P. Lin, L. C. Pedersen, V. K. Batra, W. A. Beard, S. H. Wilson, L. G. Pedersen, „Energy analysis of chemistry for correct insertion by DNA polymerase β .“, *Proceedings of the National Academy of Sciences of the United States of America*, 103(36):13294–13299, (2006).
- [142] P. Lin, V. K. Batra, L. C. Pedersen, W. A. Beard, S. H. Wilson, L. G. Pedersen, „Incorrect nucleotide insertion at the active site of a G:A mismatch catalyzed by DNA polymerase β .“, *Proceedings of the National Academy of Sciences of the United States of America*, 105(15):5670–5674, (2008).
- [143] T. M. Glennon, J. Villa, A. Warshel, „How does GAP catalyze the GTPase reaction of Ras?: a computer simulation study.“, *Biochemistry*, 39(32):9641–9651, (2000).
- [144] H. te Heesen, K. Gerwert, J. Schlitter, „Role of the arginine finger in Ras. RasGAP revealed by QM/MM calculations.“, *Federation of the European Biochemical Societies Letters*, 581(29):5677–5684, (2007).
- [145] W. M. Oldham, H. E. Hamm, „Structural basis of function in heterotrimeric G proteins.“, *Quarterly reviews of Biophysics*, 39(2):117–166, (2006).
- [146] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. R. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, „AMBER, a package of computer programs for applying molecular mechanics normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules“, *Computer Physics Communications*, 91(1-3):1–41, (1995).
- [147] P. Oelschlaegner, M. Klahn, W.A. Beard, S.H. Wilson, A. Warshel, „Magnesiumcationic dummy atom molecules enhance representation of DNA polymerase β in molecular dynamics simulations: improved accuracy in studies of structural features and mutational effects“, *Journal of Molecular Biology*, 366(2):687–701, (2007).
- [148] ACEMD. <http://www.acellera.com/>
- [149] J. J. Arnold, C. E. Cameron, „Poliovirus RNA-dependent RNA polymerase (3Dpol): pre steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg^{2+} .“, *Biochemistry*, 43(18):5126–5137, (2004).
- [150] E. B. Lansdon, D. Samuel, L. Lagpacan, K. M. Brendla, K. L. White, M. Hung, X. Liu, C. G. Boojamra, R. L. Mackman, T. Cihlar, A. S. Ray, M. E. McGrath, S. Swaminathan, „Visualizing the molecular interactions of a nucleotide analog, GS-9148, with HIV-1 reverse transcriptase-DNA complex.“, *Journal of Molecular Biology*, 397(4):967–978, (2010).

- [151] J. Sgrignani, A. Magistrato, „First-principles modeling of biological systems and structure-based drug-design.“, *Current Comput Aided Drug Design* 9(1):15–34, (2013).
- [152] H. M. Senn, W. Thiel, „QM/MM methods for biomolecular systems.“, *Angewandte Chemie International Edition*, 48(7):1198–1229, (2009).
- [153] G. Groenhof, „Introduction to QM/MM simulations.“ *Methods in Molecular Biology*, 924:43–66, (2013).
- [154] M. L. Jain, P. Y. Bruice, I. E. Szabo, T. C. Bruice, „Incorporation of positively charged linkages into DNA and RNA backbones: a novel strategy for antigene and antisense agents“, *Chemical Reviews*, 112(3):1284–1309, (2012).
- [155] J. Kurreck, „Antisense technologies. Improvement through novel chemical modifications“, *European Journal of Biochemistry*, 270(8):1628–1644, (2003).
- [156] Y. S. Cho-Chung, „Antisense DNAs as targeted genetic medicine to treat cancer“, *Archives of Pharmacal Research*, 26(3):183–191, (2003).
- [157] J. B. Opalinska, A. M. Gewirtz, „Nucleic-acid therapeutics: basic principles and recent applications“, *Nature Reviews Drug Discovery*, 1(7):503–514, (2002).
- [158] A. Dove, „Antisense and sensibility“, *Nature Biotechnology*, 20(2):121–124, (2002).
- [159] D. B. Braasch, D. R. Corey, „Novel antisense and peptide nucleic acid strategies for controlling gene expression“, *Biochemistry*, 41(14):4503–4510, (2002).
- [160] A. M. Gewirtz, D. L. Sokol, M. Z. Ratajczak, „Nucleic acid therapeutics: state of the art and future prospects“, *Blood*, 92(3):712–736, (1998).
- [161] E. Uhlmann, A. Peyman, „Antisense oligonucleotides a new therapeutic principle“, *Chemical Reviews*, 90(4):544–584, (1990).
- [162] D. Rejman, J. Snasel, R. Liboska, Z. Tocik, O. Paces, S. Kralikova, M. Rinnova, P. Kois, I. Rosenberg, „Oligonucleotides with isopolar phosphonate internucleotide linkage: a new perspective for antisense compounds?“, *Nucleosides, Nucleotides and Nucleic Acids*, 20(4-7):819–823, (2001).
- [163] Y. Xiang, P. Oelschlaeger, J. Florian, M. F. Goodman, A. Warshel, „Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: evaluating amino acid group contribution and allosteric coupling for ionized residues in Human Pol β “, *Biochemistry*, 45(23):7036–7048, (2006).
- [164] J. Florian, M. F. Goodman, A. Warshel, „Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase“, *Journal of the American Chemical Society*, 125(27):8163–8177, (2003).
- [165] M. Dal Peraro, K. Spiegel, G. Lamoureux, M. De Vivo, W. F. De Grado, M. L. Klein, „Modeling the charge distribution at metal sites in proteins for molecular dynamics simulations“, *Journal of Structural Biology*, 157(3):444–453, (2007).
- [166] J. Aqvist, „Ion–water interaction potentials derived from free energy perturbation Simulations“, *Journal of Physical Chemistry*, 94(21):8021–8024, (1990).
- [167] H. Wu, W. F. Lima, S. T. Crooke, „Investigating the structure of human RNase H1 by site-directed mutagenesis“, *Journal of Biological Chemistry*, 276(26):23547–23553, (2001).
- [168] A. Fire, S. Q. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, C. C. Mello, „Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*“, *Nature*, 391(6669):806–811, (1998).
- [169] A. Fire, „Gene silencing by double-stranded RNA (Nobel Lecture)“, *Angewandte Chemie International Edition*, 46:6966–6984, (2007).

- [170] C. C. Mello, „Return to the RNAi world: rethinking gene expression and evolution (Nobel Lecture)“, *Angewandte Chemie International Edition*, 46:6985–6994, (2007).
- [171] K. Nakanishi, D. E. Weinberg, D. P. Bartel, D. J. Patel, „Structure of yeast Argonaute with guide RNA“, *Nature*, 486:368–374, (2012).
- [172] Y. Wang, S. Juranek, H. Li, G. Sheng, G. S. Wardle, T. Tuschl, D. J. Patel, „Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes“, *Nature*, 461:754–762, (2009).
- [173] I. M. Kerr, R. E. Brown, „pppA2'p5'A2'p5'A: An inhibitor of protein synthesis synthesized with an enzyme fraction from interferon-treated cells“, *Proceedings of the National Academy of Sciences of the United States of America*, 75(1):256-260, (1978).
- [174] M. R. Player, P. F. Torrence, „The 2-5A System: Modulation of Viral and Cellular Processes Through Acceleration of RNA Degradation“, *Pharmacology & Therapeutics*, 78(2):55-114, (1998).
- [175] R. H. Silverman, „A scientific journey through the 2-5A/RNase L system“, *Cytokine & Growth Factor Reviews*, 18(5-6):381-388, (2007).
- [176] R. H. Silverman, „Viral encounters with OAS and RNase L during the IFN antiviral response.“, *Journal of Virology*, 81(23):12720–12729, (2007).
- [177] X. L. Li, H. J. Ezelle, T. J. Kang, L. Zhang, K. A. Shirey, J. Harro, J. D. Hasday, S. K. Mohapatra, O. R. Crasta, S. N. Vogel, A. S. Cross, B. A. Hassel, „An essential role for the antiviral endoribonuclease, RNase-L, in antibacterial immunity.“, *Proceedings of the National Academy of Sciences of the United States of America*, 105(52):20816–20821, (2008).
- [178] K. Lesiak, J. Imai, G. Floyd-Smith, P. F. Torrence, „Biological Activities of Phosphodiester Linkage Isomers of 2-5A“, *The Journal of Biological Chemistry*, 258(21):13082-13088, (1983).
- [179] K. Morita, M. Kaneko, S. Obika, T. Imanishi, Y. Kitade, M. Koizumi, „Biologically Stable 2-5A Analogues containing 3'-O,4'-C-bridged Adenosine as Potent RNase L Agonists“, *ChemMedChem*, 2(12):1703-1707, (2007).
- [180] K. Maláč, I. Barvík Jr., „Recognition of 2',5'-linked oligoadenylates by human ribonuclease L: Molecular dynamics study“, manuskript.
- [181] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, „Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters“, *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, (2006).
- [182] R. W. Sobol, E. E. Henderson, N. Kon, J. Shao, P. Hitzges, E. Mordechai, N. L. Reichenbach, R. Charubala, H. Schirmeister, W. Pfliederer, R. J. Suhadolnik, „Inhibition of HIV-1 Replication and Activation of RNase L by Phosphorothioate/Phosphodiester 2',5'-Oligoadenylate Derivatives“, *The Journal of Biological Chemistry*, 270(11):5963-5978, (1995).

Seznam tabulek

Tab.1. Knihovna MD simulací použitých komplexů s HS RNase L

Seznam obrázků

- Obrázek 1.1.:** Schéma exprese genetické informace.1
- Obrázek 1.2.:** Virové polymerázy (RdRp, HIV RT) a enzymy provádějící degradaci RNA (HIV RT - RNA v duplexech RNA:DNA; RNáza L - jednovláknová RNA; RNáza H - RNA v duplexech DNA:RNA; Argonaute - RNA v duplexech RNA:RNA). Tato práce se zabývá enzymy RdRp, RNázou L, RNázou H a Argonaute. 3
- Obrázek 1.3.:** Porovnání aktivních míst Norovirové RNA-dependentní-RNA polymerázy (*nahoře*) a lidské RNázy H (*dole*). Ačkoliv obě aktivní místa mají opačnou funkci (syntéza vs. degradace nukleové kyseliny), jejich uspořádání je velmi podobné. Několik záporně nabitých acidických Asp/Glu reziduí poutá dva kladně nabitě hořčičkové ionty, které přímo interagují se záporně nabitou nukleovou kyselinou. Aktivní místa všech polymeráz (mnohapodjednotkové DNA-dependentní-RNA polymerázy, virových RNA-dependentních-RNA polymeráz, HIV reverzní transkriptázy) jsou si přitom velmi podobná. Stejně tak jsou si podobná aktivní místa enzymů provádějících degradaci RNA (RNázové domény HIV RT, lidské RNázy H a enzymu Argonaute) 4
- Obrázek 2.1.:** Mnohapodjednotková DNA-dependentní-RNA polymeráza. Barevně odlišeny jsou největší podjednotky β (světlomodrá), β' (světležlutá) a σ (zelená). Dále je vyznačena poloha podjednotek α a ω . Dále jsou znázorněna vlákna nukleových kyselin vytvářející tzv. transkripční bublinu. Netemplátové vlákno DNA je obarveno žlutě, templátové vlákno DNA modře a vlákno syntetizované RNA červeně. Poloha aktivního místa je naznačena prostřednictvím symbolu AS..... 6
- Obrázek 2.2.:** Interakce σ podjednotky (zelená) DNA-dependentní-RNA polymerázy s tzv. transkripční bublinou. Netemplátové vlákno DNA je obarveno žlutě, templátové vlákno DNA modře a vlákno syntetizované RNA červeně. Poloha aktivního místa je naznačena prostřednictvím symbolu AS. Jsou v něm patrné dva hořčičkové ionty (modré kuličky) a přístupující nukleosidtrifosfát (NTP)..... 7
- Obrázek 2.3.:** Interakce iniciační σ podjednotky (zelená) s RNA (červená) v aktivním místě DNA-dependentní-RNA polymerázy. Poloha aktivního místa je naznačena prostřednictvím symbolu AS. Jsou v něm patrné dva hořčičkové ionty (modré kuličky) a přístupující nukleosidtrifosfát (NTP). Dále jsou schematicky vyznačeny helixy tzv. Bridge Helixu (BH) a Trigger Loopu (TL), jejichž pohyb otevírá a uzavírá aktivní místo RNAP a přispívá k extrémní přesnosti přepisu genetické informace. Antibakteriální látka Streptolydigin inhibuje RNAP sterickou blokadou BH a TL. Antibiotikum Rifampicin narušuje interakci mezi σ podjednotkou RNAP a nukleovými kyselinami vytvářejícími transkripční bublinu. V konečném důsledku tak zabraňuje prodloužení RNA na délku větší než 4-5 nt....9
- Obrázek 2.4.:** Norovirová RNA-dependentní-RNA polymeráza (NV RdRp) v komplexu s nukleovou kyselinou (PDB id: 3BSO). NV RdRp má typickou strukturu polymeráz připomínající pravou ruku (dlaň - šedá, palec - modrá, prsty - červená). V RdRp bývají palec a prsty spojené a tak nemohou měnit své konformace nezávisle..... 11

Obrázek 2.5.: CTP v aktivním místě NV RdRP v kontextu primerového a templátového vlákna RNA. Mg ²⁺ ionty se váží na konzervovaná rezidua Asp344, Asp343 a Asp242 v doméně dlaně NV RdRp	12
Obrázek 2.6.: HIV reverzní transkriptáza sestávající ze dvou podjednotek (p51 a p66). Templátové vlákno virové RNA je obarveno modře. Syntetizované vlákno DNA je obarveno žlutě. V dolní části je polymerázové aktivní místo (syntetizující DNA) s přístupujícím dNTP. V horní části je aktivní místo domény RNázy H degradující vlákno virové RNA.	14
Obrázek 2.7.: Pět internukleotidových spojek z 3' konce DNA se podílí na stabilizujících interakcích s povrchem HS RNázy H	18
Obrázek 2.8.: Aktivní místo HS RNázy H	18
Obrázek 2.9.: Komplex siRNA/miRNA/piRNA (žluté vlákno) s cílovou mRNA (bleděmodré vlákno). Dále jsou znázorněny hořčičkové atomy a acidické aminokyseliny v aktivním místě enzymu Argonaute.	22
Obrázek 2.10.: Komplex siRNA/miRNA/piRNA (žluté vlákno) s cílovou mRNA (bleděmodré vlákno) a enzymem Argonaute (katalytická doména PIWI je obarvena červeně, N-terminální doména modře, Mid doména zeleně a PAZ doména fialově).	23
Obrázek 2.11.: Ligand 2–5A je navázán na druhý až čtvrtý ankyrinový motiv RNázy L a interaguje s aminokyselinami prostřednictvím vodíkových vazeb (Asn65, Gln68, Glu131, Tyr135, Tyr124), solných můstků (Lys89, Arg155, Lys166) a stackingu (Trp60, Phe126)	26
Obrázek 3.1.: Schématický náčrt rozdělení nábojů	46
Obrázek 4.1.: Struktura CTP, 2dCTP, coCTP a cocCTP	55
Obrázek 4.2.: CTP, 2dCTP, coCTP a cocCTP v aktivním místě NV RdRp.....	59
Obrázek 4.3.: a) Arg182 interagující s α-fosfátovou skupinou CTP; b) Arg182 interagující s β-fosfátovou skupinou CTP	59
Obrázek 4.4.: Konformační preference CTP a jeho interakce v aktivním místě NV RdRp: A01) CTP O5'-C5'-C4'-O4' torzní úhel; A02) CTP puckering cukru; A03) Intramolekulární vodíkové vazby: CTP_3'OH - β-fosfát_CTP; A04) Asp247 - 3'OH_CTP; A05) Asp247 - Ser300; A06) Ser300 - 2'OH_CTP; A07) Asp247 - 2'OH_CTP; A08) Arg182 - α-fosfát_CTP; A09) Arg182 - β-fosfát_CTP; A10) Primer_O3' - α-fosfát_CTP;	60
Obrázek 4.5.: Kontakty v aktivním místě RNázy H v MD simulaci se sedmibodovou reprezentací hořčičkových iontů - A) vzdálenost MgA-MgB (Graph A); C) vzdálenost MgA - ligand (Graf C1: Asp274 červená, Asp145 zelená, Sp modrá; D) vzdálenost MgB - ligand (Graf D1: Asp145 červená, Glu186 zelená, Glu186 modrá, Asp210 bleděmodrá, Asp210 fialová, Sp žlutá, O3' lososová).	67

- Obrázek 4.6.:** Kontakty v aktivním místě HS RNázy H v 3pc5 MD simulaci - A) vzdálenost MgA-MgB (Graf A); C) vzdálenost MgA - ligand (Graf C1: Asp274 červená, Asp145 zelená, Sp modrá; Graf C2: WATA-C); D) vzdálenost MgB - ligand (Graf D1: Asp145 červená, Glu186 zelená, Glu186 modrá, Asp210 bleděmodrá, Asp210 fialová, Sp žlutá, O3' lososová, Graf D2: WAT). 69
- Obrázek 4.7.:** Kontakty v aktivním místě HS RNázy H v 3pc5 MD simulaci - A) vzdálenost MgA-MgB (Graf 6A); C) vzdálenost MgA - ligand (Graf C1: Asp274 červená, Asp145 zelená, Sp modrá; Graf C2: WATA-C); D) vzdálenost MgB - ligand - (Graf D1: Asp145 červená, Glu186 zelená, Glu186 modrá, Asp210 bleděmodrá, Asp210 fialová, Sp žlutá, O3' lososová, Graph D2: WAT). 70
- Obrázek 4.8.:** Vazba 3pc5 dG1-dG2 internukleotidové spojky do prvního DNA vazebného místa na povrchu HS RNázy H. 72
- Obrázek 4.9.:** Vazba 3pc5 dC2-dT3 internukleotidové spojky do druhého DNA vazebného místa na povrchu HS RNázy H. 72
- Obrázek 4.10.:** Vazba 3pc5 dT3-dG4 internukleotidové spojky do třetího DNA vazebného místa na povrchu HS RNázy H. 72
- Obrázek 4.11.:** Vazba internukleotidových spojek dG4-dT5 a dT5-dG6 do čtvrtého a pátého DNA vazebného místa na povrchu RNázy H. 73
- Obrázek 4.12.:** Interakce RNA vlákna s HS RNázou H. 73
- Obrázek 4.13.:** Interakce 3pc5 DNA s HS RNázou H - A) Trp225 – červená, Ser233 - modrá; B) Asn240 m.c. - modrá, Ile239 m.c. - bleděmodrá, Asn240 s.c. - červená, Lys236 s.c. - zelená; C1) Thr181 – modrá, Asn240 – červená; C2) Gln183 - modrá, Asn182 - zelená, Arg184 - bleděmodrá, Asn240 - červená; D) Asn151 - červená, Arg153 – modrá 74
- Obrázek 4.14.:** Interakce 3pc5 DNA s HS RNázou H - A) Trp225 - červená, Ser233 - modrá; B) Asn240 - modrá, Ile239 - bleděmodrá, Asn240 - červená, Lys236 - zelená; C1) Thr181 – modrá, Asn240 - červená; C2) Gln183 - modrá, Asn182 - zelená, Arg184 - bleděmodrá, Asn240 - červená; D) Asn151 - červená, Arg153 - modrá 75
- Obrázek 4.15.:** Doména enzymu Argonaute strukturně homologní s RNázou H. Červeně je vyznačenou Glu reziduum interagující zároveň s aktivním místem i fosfátovou skupinou jedné z internukleotidových spojek siRNA vlákna 80
- Obrázek 4.16.:** Kontakty v aktivním místě v TtAGO MD simulaci - A) vzdálenost MgA-MgB (Graph A); C) vzdálenost MgA - ligand (Graph C1: Asp660 červená, Asp478 zelená, Sp modrá; D) vzdálenost MgB - ligand (Graf 6D1: Asp478 červená, Glu512 zelená, Glu512 modrá, Asp546 bleděmodrá, Asp546 fialová, Sp žlutá, O3' lososová) 83
- Obrázek 4.17.:** Strukturní alignment HS RNázy H, BH RNázy H a reziduí v aktivním místě enzymu Argonaute. Dole je znázorněno alosterické ovlivňování aktivního místa RNáz H prostřednictvím vazby antisense oligonukleotidu a přímé ovlivňování aktivního místa enzymu Argonaute vláknem siRNA. 84

Obrázek 4.18.: Molekula pA₂₅A₂₅A navázaná k povrchu RNázy L_E131H během MD simulace č.2 způsobem, který odpovídá tomu, co bylo zaznamenáno ve výchozí krystalové struktuře. 92

Obrázek 4.19.: Molekula pA₂₅A₂₅A změnila v MD simulaci č. 10 s mutantní RNaseL_E131H_W60A svou konformaci tak, že pouze Ade1 zůstalo ve své původní vazebné pozici 93

Obrázek 4.20.: Výsledky MD simulací č. 1-6 ve sloupcích grafů zleva doprava. Časový vývoj různých strukturních parametrů v řádcích grafů shora dolů. Vazba **Ade1** prostřednictvím vodíkových vazeb k protonovanému/neprotonovanému Glu131: Glu131_O-H62_Ade1 (červená čára) a Glu131_OH-N1_Ade1 (zelená čára). Kontakty **Ade2** s postranním řetězcem Tyr135: Tyr135_OH-N1_Ade2 (červená čára). Síť vodíkových vazeb zahrnujících **Ade3** a postranní řetězce Gln68: Gln68_O1-H62_Ade3 (červená čára) a Gln68_H2-N1_Ade3 (zelená čára) a Asn65: Asn65_O1-H61_Ade3 (modrá čára). Postranní řetězce Arg155 (červená čára) a Lys166 (zelená čára) formují solné můstky s 5'-fosfátovou skupinou prvního AMP (**Phos1**). 5'-fosfátová skupina druhého AMP (**Phos2**) formovala příležitostně vodíkové vazby s postranním řetězcem Tyr124 (červená čára). 5'-fosfátová skupina třetího AMP (**Phos3**) formovala příležitostně solné můstky s postranním řetězcem Lys89 (červená čára). Hodnoty **glykosidického** torzního úhlu (O4'-C1'-N9-C4) pro Ade1 (červená čára), Ade2 (zelená čára), Ade3 (modrá čára). **Puckering** ribózy (Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace segmentů C4'-C5' spojujících ribózy a internukleotidové spojky vyjádřené prostřednictvím torzního úhlu **beta** (C3'-C4'-C5'-O5'): Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace fosfodiesterových internukleotidových spojek Ade1-Ade2 a Ade2-Ade3 (tj. **PL 1-2** and **PL 2-3**) vyjádřené prostřednictvím konformací torzních úhlů C5'-O5'-P-O2' (červená čára) a O5'-P-O2'-C2' (zelená čára) 94

Obrázek 4.21.: Výsledky MD simulací č. 7-12 ve sloupcích grafů zleva doprava. Časový vývoj různých strukturních parametrů v řádcích grafů shora dolů. Vazba **Ade1** prostřednictvím vodíkových vazeb k protonovanému/neprotonovanému Glu131: Glu131_O-H62_Ade1 (červená čára) a Glu131_OH-N1_Ade1 (zelená čára). Kontakty **Ade2** s postranním řetězcem Tyr135: Tyr135_OH-N1_Ade2 (červená čára). Síť vodíkových vazeb zahrnujících **Ade3** a postranní řetězce Gln68: Gln68_O1-H62_Ade3 (červená čára) a Gln68_H2-N1_Ade3 (zelená čára) a Asn65: Asn65_O1-H61_Ade3 (modrá čára). Postranní řetězce Arg155 (červená čára) a Lys166 (zelená čára) formují solné můstky s 5'-fosfátovou skupinou prvního AMP (**Phos1**). 5'-fosfátová skupina druhého AMP (**Phos2**) formovala příležitostně vodíkové vazby s postranním řetězcem Tyr124 (červená čára). 5'-fosfátová skupina třetího AMP (**Phos3**) formovala příležitostně solné můstky s postranním řetězcem Lys89 (červená čára). Hodnoty **glykosidického** torzního úhlu (O4'-C1'-N9-C4) pro Ade1 (červená čára), Ade2 (zelená čára), Ade3 (modrá čára). **Puckering** ribózy (Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace segmentů C4'-C5' spojujících ribózy a internukleotidové spojky vyjádřené prostřednictvím torzního úhlu **beta** (C3'-C4'-C5'-O5'): Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace fosfodiesterových internukleotidových spojek Ade1-Ade2 a Ade2-Ade3 (tj. **PL 1-2** and **PL 2-3**) vyjádřené prostřednictvím konformací torzních úhlů C5'-O5'-P-O2' (červená čára) a O5'-P-O2'-C2' (zelená čára). 95

Obrázek 4.22.: Výsledky MD simulací č. 13-19 ve sloupcích grafů zleva doprava. Časový vývoj různých strukturních parametrů v řádcích grafů shora dolů. Vazba **Ade1** prostřednictvím vodíkových vazeb k protonovanému/neprotonovanému Glu131: Glu131_O-H62_Ade1 (červená čára) a Glu131_OH-N1_Ade1 (zelená čára). Kontakty **Ade2** s postranním řetězcem Tyr135: Tyr135_OH-N1_Ade2 (červená čára). Síť vodíkových vazeb zahrnujících **Ade3** a postranní řetězce Gln68: Gln68_O1-H62_Ade3 (červená čára) a Gln68_H2-N1_Ade3 (zelená čára) a Asn65: Asn65_O1-H61_Ade3 (modrá čára). Postranní řetězce Arg155 (červená čára) a Lys166 (zelená čára) formují solné můstky s 5'-fosfátovou skupinou prvního AMP (**Phos1**). 5'-fosfátová skupina druhého AMP (**Phos2**) formovala příležitostně vodíkové vazby s postranním řetězcem Tyr124 (červená čára). 5'-fosfátová skupina třetího AMP (**Phos3**) formovala příležitostně solné můstky s postranním řetězcem Lys89 (červená čára). Hodnoty **glykosidického** torzního úhlu (O4'-C1'-N9-C4) pro Ade1 (červená čára), Ade2 (zelená čára), Ade3 (modrá čára). **Puckering** ribózy (Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace segmentů C4'-C5' spojujících ribózy a internukleotidové spojky vyjádřené

prostřednictvím torzního úhlu **beta** ($C3'-C4'-C5'-O5'$): Ade1 – červená čára, Ade2 – zelená čára, Ade3 – modrá čára). Konformace fosfodiesterových internukleotidových spojek Ade1-Ade2 a Ade2-Ade3 (tj. **PL 1-2** and **PL 2-3**) vyjádřené prostřednictvím konformací torzních úhlů $C5'-O5'-P-O2'$ (červená čára) a $O5'-P-O2'-C2'$ (zelená čára) 96

Seznam použitých zkratk

siRNA	small interfering RNA
RdRp	RNA dependentní RNA polymeráza
HIV RT	HIV reverzní transkriptáza
RNAP	DNA dependentní RNA polymeráza
NTP	nukleosidtrifosfát
BH	Bridge Helix
TL	Trigger Loop
HCV	virus hepatitidy typu C
NV	virus Norwalk
NV RdRp	Norovirová RNA dependentní RNA polymeráza
NI	nukleosidové inhibitory
NNI	non-nukleosidové inhibitory
PPi	analoga pyrofosfátu
RT	reverzní transkriptáza
PBS	vazebné místo primeru
RNáza H	Ribonukleáza H
HBD	oblast pro navázání hybridů
eIF2	eukaryotický Iniciační Faktor 2
RNAi	RNA interference
RISC	RNA-induced silencing complex
miRNA	micro RNA
siRNA	small interfering RNA
piRNA	Piwi-interacting RNA
OB	Oligonucleotide Binding
RNáza L	Ribonukleáza L
25A	β -anomerní 5'-fosforylované oligoadenyláty s 2',5' internukleotidovými spojkami
RESP	Restrained ElectroStatic Potential
ESP	elektrostatický potenciál
TDSCF	time dependent self consistent field
FDE	finite difference equation
PC	algoritmy typu prediktor-korektor
BBK	Brünger Brooks Karplus
NVE	Mikrokanonický soubor
NVT	Kanonický soubor
NPT	Izotermicko-Izobarický soubor
PBC	Periodic boundary conditions
PM	Particle mesh (metoda)
NGP	Nearest grid point
FFT	fast fourier transform
PME	smooth Particle Mesh Ewaldova metoda
AMBER	Assisted Model Building with Energy Refinement
Sander	Simulated Annealing with NMR-Derived Energy Restraints
MPI	Message Passing Interface
GPU	graphics processing units
CUDA	Compute Unified Device Architecture
pdb	Brookhaven Protein DataBank format

NAMD	Not (just) Another Molecular Dynamics
OpenCL	Open Computing Language
HMR	Hydrogen Mass Repartitioning
NMP	nukleosidmonofosfát
NDP	nukleosiddifosfát
AW	Åqvist-Warshel
RMSD	střední kvadratická odchylka
HS	lidská
3pc5	3'-O-P-CH ₂ -O-5'
3cp5	3'-O-CH ₂ -P-O-5'
3cp5-DNA	3'-G-3cp5-C-3cp5-T-3cp5-G-T-G-G-A-5'
3pc5-DNA	3'-G-3pc5-C-3pc5-T-3pc5-G-T-G-G-A-5'
AGO	proteiny Argonaute
KpAGO	Kluyveromyces polysporus Argonaute
TtAGO	Thermus thermophilus Argonaute
IFN	interferony typu I
OAS	2',5' oligoadenylátové syntetázy
ANK	N-terminální ankyrinová
Ade1	první AMP
Ade2	druhý AMP
Ade3	třetí AMP
Phos1	5'-fosfátová skupina prvního AMP
Phos2	5'-fosfátová skupina druhého AMP
Phos3	5'-fosfátová skupina třetího AMP

Přílohy

- [1] K. Maláč, I. Barvík Jr., „Substrate recognition by norovirus polymerase: microsecond molecular dynamics study“, *Journal of Computer-Aided Molecular Design*, 27(4):373-388, (2013).
- [2] K. Maláč, I. Barvík Jr., „Complex between Human RNase HI and the phosphonate-DNA/RNA duplex: Molecular dynamics study“, *Journal of Molecular Graphics and Modelling*, 44:81–90, (2013).
- [180] K. Maláč, I. Barvík Jr., „Recognition of 2',5'-linked oligoadenylates by human ribonuclease L: Molecular dynamics study“, manuskript, odesláno do *Journal of Molecular Modeling*.

Substrate recognition by norovirus polymerase: microsecond molecular dynamics study

Kamil Maláč · Ivan Barvík

Received: 7 October 2012 / Accepted: 18 April 2013 / Published online: 26 April 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Molecular dynamics simulations of complexes between Norwalk virus RNA dependent RNA polymerase and its natural CTP and 2dCTP (both containing the O5′–C5′–C4′–O4′ sequence of atoms bridging the triphosphate and sugar moiety) or modified coCTP (C5′–O5′–C4′–O4′), cocCTP (C5′–O5′–C4′–C4′′) substrates were produced by means of CUDA programmable graphical processing units and the ACEMD software package. It enabled us to gain microsecond MD trajectories clearly showing that similar nucleoside triphosphates can bind surprisingly differently into the active site of the Norwalk virus RNA dependent RNA polymerase. It corresponds to their different modes of action (CTP—substrate, 2dCTP—poor substrate, coCTP—chain terminator, cocCTP—inhibitor). Moreover, extremely rare events—as repetitive pervasion of Arg182 into a potentially reaction promoting arrangement—were captured.

Keywords Norwalk virus · Hepatitis C virus · RNA dependent RNA polymerase · Molecular dynamics · Phosphonate · Competitive inhibitor · Chain terminator · NAMD · GPU · CUDA · ACEMD

Electronic supplementary material The online version of this article (doi:10.1007/s10822-013-9652-8) contains supplementary material, which is available to authorized users.

K. Maláč · I. Barvík (✉)
Faculty of Mathematics and Physics, Institute of Physics,
Charles University, Ke Karlovu 5, Prague 2 121 16,
Czech Republic
e-mail: ibarvik@karlov.mff.cuni.cz

Introduction

Some viruses containing just the RNA genome are reproduced in host cells by RNA dependent RNA polymerase (RdRp) [1, 2]. Norwalk virus (NV) is a major cause of gastroenteritis outbreaks in developed countries [3, 4]. Unfortunately, effective treatments are not currently available. Hepatitis C virus (HCV) is spread among 3 %, or 170 million, of world population, of which only about 25 % are having symptoms [2]. Current treatment with interferon- α (immunomodulation) and ribavirin (nucleotide inhibitor) has limited efficiency and numerous side effects. Only 40–50 % of people exposed to HCV outbreak fully recover while rest is affected by chronic liver problems, with possibility of developing cirrhosis or even liver cancer [2]. Those patients undergoing interferon/ribavirin treatment also face the prospects of significant adverse effects that include fatigue, hemolytic anemia, depression, and flulike symptoms which ultimately lead to high rates of drug discontinuations [1].

In attempts to improve on the overall HCV cure rate and reduce or eliminate the serious side effects associated with interferon/ribavirin therapy, the search for direct acting antivirals (DAAs) that are safe and effective has become an urgent endeavor [1]. Recently, two protease inhibitors, telaprevir and boceprevir, in combination with interferon/ribavirin have been approved for treating HCV infection. However, they bring additional side effects that require further medical management [1]. Therefore, the search continues for novel agents with complementary mechanisms of action.

RdRp is not present in mammalian cells. Therefore, it is a suitable target for inhibition [5]. Mostly, the RdRp enzyme has four subunits. One of them has an active site for replication. The other three subunits help replicase to

locate and bind the 3' end of RNA. RdRps contain the classic structural domains, denoted fingers, palm, and thumb, of other single-chain polynucleotide polymerases [6] (Fig. S1). Crystal structure of NV RdRp reveals the carboxyl terminus in the active site cleft, which is reminiscent of the carboxyl-terminal extension domain and “plough” loop in bacteriophage ϕ 6 RdRp, as well as the carboxyl-terminal segment and β -loop insertion in the thumb domain of HCV RdRp [7]. RNA binding displaces the C-terminal tail from the active site cleft of NV RdRp. The binding of RNA also cause the central helix of the thumb domain (residues 435–449) to rotate by 22° , thus inducing the formation of a binding groove for the primer strand [3] (Fig. S2).

In contrast to most other single-chain polymerases of known structure, the finger and thumb domains are connected in the RdRps and are therefore not free to change conformation independently of each other [6]. The encircled active site structure makes it unlikely that major swinging motions of the fingers domain act to reposition the nascent base pair from an initial preinsertion site into the catalytic site, as is seen in other classes of single-subunit polymerases. Therefore, RdRps use a fully prepositioned templating base for nucleotide recognition [8] (Fig. S2).

Generally, nucleotide substrate binds to the polymerase-primer/template complex (step 1), and this complex isomerizes into a catalytically active complex (step 2). Phosphoryl transfer occurs (step 3), followed by a second conformational change (step 4) and pyrophosphate release (step 5) [9]. In HIV reverse transcriptase (HIV RT), step 2 involves an inward bending of fingertips toward the palm by about 20° resulting in a displacement of the outermost parts of the domain by more than 5 \AA . As a result, a number of residues are brought into contact with the incoming nucleoside triphosphate [10]. Closure of the fingers is also a noteworthy feature of dNTP binding in members of both the DNA polymerase I and terminal transferase families [10].

RdRps, in contrast, close their encircled active sites for catalysis using a subtle structural rearrangement in the palm domain. The key element of 3Dpol/NV active site closure is a subtle shift of motif A (depicted in Fig. 1d as a blue ribbon) to position the essential Asp233/Asp242 for metal binding and subsequent catalysis [3, 8].

One possibility of HCV treatment is non-nucleoside inhibition (NNI) of RdRp using allosteric inhibitors. They prevent HCV RdRp from reaching an active conformation [5]. In fact more than 80 crystal complexes of HCV RdRp

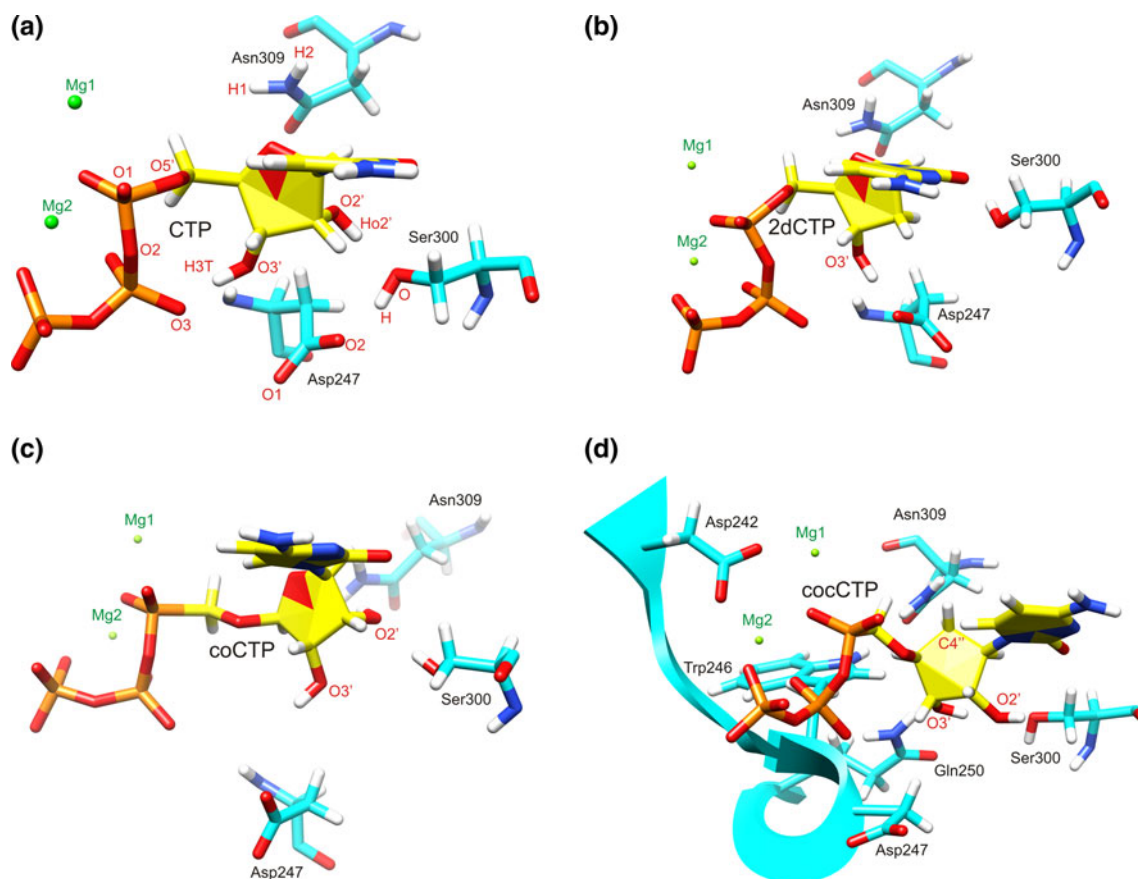


Fig. 1 CTP, 2dCTP, coCTP and cocCTP in the NV RdRp active site

with small molecule allosteric inhibitors have been reported [1]. Several suitable binding sites for NNIs on the enzyme allow combination therapy. Inhibitors bind to the thumb cavity, which is crucial for “fingertip-thumb” interaction, thus forcing enzyme to remain in the inactive form [2]. Human clinical trials evaluating HCV NNIs as monotherapy have all shown viral breakthrough resulting from rapid emergence of resistant virus [1].

In contrast, nucleoside-based HCV RdRp inhibitors (nucleosides and nucleotide prodrugs—to enhance cellular uptake) have the potential to become the backbone of HCV therapeutic regimens. These nucleoside-based agents have demonstrated a high barrier to resistance. No other small molecule DAA developed to date has been able to mirror these characteristics [1]. This high barrier to resistance suggests that there is a high degree of amino acid conservation in the active site of HCV RdRp and that only subtle amino acid changes in the active site, which do not impair the replication fitness of the virus, are tolerated [1].

Nucleoside inhibitors of HCV RdRp have been explored in recent years [1, 5, 11–16]. They are converted by host cells into nucleotides, nucleoside diphosphates and nucleoside triphosphates. During the viral replication, they are incorporated into the synthesized RNA and act as chain terminators. In the case of HCV, 2'-C-methyl nucleosides are mostly used. They confer antiviral activity irrespectively to the base [11]. The methyl group added to C2' of ribose transforms a nucleotide substrate into a chain terminator. Several proof of concept clinical studies have shown that the combination of a nucleoside (2'- α -F-2'- β -C-methylpyridine) with a HCV protease inhibitor and the combination of two nucleotide prodrugs (2'- α -F-2'- β -C-methyluridine 5'-phosphoramidate and 3', 5'-cyclic phosphate prodrug of 2'- α -F-2'- β -C-methylguanosine) have the potential to deliver an interferon/ribavirin free combination therapy [1]. Nevertheless, Bristol-Myers Squibb's disastrous phase 2 combination trial of the HCV drugs BMS-986094 and daclatasvir, which has caused the death of one patient and severe heart and kidney damage to another eight, has raised questions about this class of nucleoside inhibitors [16].

But there are other promising structural motives. Recent experimental studies show that (S)-9-[3-hydroxy-2-(phosphonomethoxy)propyl]-adenine (S-HPMPA—a substance known for more than 20 years) in a suitable prodrug form—might be able to fight HCV too [14]. Nucleosides and nucleotides are known to inhibit HCV during the elongation phase of viral RNA synthesis via a nonobligate chain termination mechanism. Dinucleotides have the potential to inhibit RNA replication at the initiation phase by behaving as primer molecules [1, 12, 13].

Molecular dynamics (MD) simulations are a well-established tool for the investigation of structural features of biomolecules at atomic detail level. Nucleotide

incorporation efficiency, fidelity and the catalytic mechanism of different polymerases were studied by means of MD simulations in past [17–24]. We used MD simulations to investigate binding of phosphoramidate dinucleotides toward HCV RdRp [12, 13].

CUDA programmable graphical processing units (GPUs) enable now to extend MD trajectories on time scales, which were until recently available only with supercomputing facilities. It inspired us to try to find subtle differences in binding of quite similar substrates into the active site of RdRp. Microsecond MD trajectories could provide time for sufficient relaxation of complexes and capturing of even very slow conformational changes.

A series of adenosine 5'-phosphonate analogues were designed in [15] to mimic naturally occurring adenosine monophosphate. These compounds were synthesized and evaluated in a cellular hepatitis C virus replication assay. To elucidate the mechanism of action of these novel adenosine phosphonates, their diphosphate derivatives were synthesized. Further nucleotide incorporation assays by HCV RdRp revealed that coATP (analogous to coCTP in Fig. S3) can serve as a chain terminator, whereas co-cATP (analogous to cocCTP in Fig. S3) compound is a competitive inhibitor.

Here, MD simulations of complexes between NV RdRp and its natural CTP and 2dCTP (both containing the O5'-C5'-C4'-O4' sequence of atoms bridging the triphosphate and sugar moiety—see Fig. S3) or modified coCTP (C5'-O5'-C4'-O4') and cocCTP (C5'-O5'-C4'-C4'') substrates were produced and analyzed in detail.

Methodology

There is an apparent structural similarity of NV/HCV RdRps, which allows straightforward transferability of many observations. NV RdRp complexes with double helical nucleic acids and CTP appeared in literature recently (3BSN, 3BSO [3], 3H5X, 3H5Y [4]). The structures of the NV RdRp complexes provided the first high resolution views of RdRp complexes trapped immediately before nucleotidyl transfer. Indirect effects from crystal packing appear to have fortuitously trapped these complexes at this stage of reaction cycle. The distal end of the primer-template duplex is located near a symmetry-related polymerase molecule, and there is insufficient space for primer extension to occur in this crystal form. As a result, polymerase molecules in this crystal form bind substrates and assemble active complexes that can only proceed up to the point immediately prior to nucleotidyl transfer. Therefore, the geometry of the NV RdRp active site was determined very precisely in [3, 4] uncovering clearly a hydrogen bonding connecting the 2'-OH/3'-OH groups of the incoming NTP with Asp247, Ser300 and Asn309 of NV

RdRp (their counterparts in HCV RdRp are Asp225, Ser282 and Asn291).

Therefore, 3BSO [3] was chosen as a suitable representative of RdRps in our MD simulations. Simulated systems were surrounded by ~15565 TIP3P [25] water molecules which extended to a distance of approximately 10 Å (in each direction) from the RdRp atoms. This gives a periodic box size of ~84, ~90, ~73 Å. New *.inpcrd (initial coordinates) and *.prmtop (molecular topology, force field etc.) files for the whole simulated system including modified residues, were created by use of the TLEAP module (AMBER software package [26]).

The AMBER force field used in this study [27, 28] does not contain force constants needed to describe modified parts of the phosphonate analogs. The force field parameters were therefore completed with those obtained from ab initio calculations consistently with the original approach [27] (Figs. S4–S5). RESP charges were computed at the HF/6-31G* level. Force constants were fitted to achieve the agreement of ~1 kcal/mol between QM (MP2/6-31G*) and MM relative energies of different conformers of the CH₃–O–PO₂–CH₂–O–CH₃ model system.

Almost half of the proteome of living organisms is constituted of metalloproteins. Unfortunately, the ability of the current generation of molecular dynamics pairwise-additive force fields to properly describe metal pockets is severely lacking due to the intrinsic difficulty of handling polarization and charge transfer contributions [29]. Magnesium ions are usually represented as atoms with a formal point charge +2 that interact with the protein environment and the substrate through non-bonded interactions. In binuclear active sites, however, high charges can lead to repulsion between two Mg²⁺ and instabilities, resulting in an alteration of the proper coordination of the Mg²⁺ ions by the ligands [24, 30, 31]. For example, using the standard AMBER Mg²⁺ ions ($R^* = 0.7926$ Å, $\epsilon = 0.8947$), the coordination of both Mg(b) and Mg(c) in the β -polymerase active site changed [24]: Mg(b) lost contact with Asp192 and O2B (a non-bridging β -phosphate oxygen of the dNTP) and was instead coordinated by O3B (the oxygen bridging the β and γ phosphates of the dNTP); Mg(c) lost contact with the O3' of the primer and O1A (a non-bridging α -phosphate oxygen of the dNTP coordinating both metal ions in the crystal structure).

Further, a general tendency to underestimate slightly magnesium-ligand distances, leading to a more compact active site, is usually observed in MD simulations [29, 32]. A consistent set of empirical interaction parameters for the alkali and alkaline-earth metal cations was derived by performing free energy perturbation simulations in aqueous solution [32]. The resulting parameters reproduce the observed hydration free energies as well as radial distribution functions of these ions. However, the value of radius

for Mg²⁺ ions shows a less good agreement with the observed distance than for any of the other ions and appears to be slightly too small [32] ($R^* = 0.7926$ used in the AMBER force-field). Even though using Mg²⁺ with $R^* = 1.300$ Å and $\epsilon = 0.06$ yielded the best structural results among the traditional one-atom Mg²⁺ models, there were some significant discrepancies between the crystal structures and the modeled structures, i.e. increased Mg(b)–Mg(c), O1A–Mg(b) distances [24].

The Aqvist-Warshel (AW) model that uses cationic dummy atoms and has proven very useful in accurately representing the energetics and structures of systems with transition metals such as octahedrally coordinated bivalent manganese ions and tetrahedrally coordinated bivalent zinc ions in mononuclear and binuclear metalloenzymes [24]. The AW model allow for stable coordination geometry by placing the cationic dummy atoms at the defined positions around the central metal atom and for a smaller repulsion between the metal ions in binuclear sites by distributing the positive charge over the cationic dummy atoms. Attaching cationic dummy atoms to the central metal atom at defined positions leads to more directed electrostatic interactions with ligands and delocalization of the charge to prevent repulsion in binuclear sites. In [24], the AW model always performed better than the best traditional one-atom Mg²⁺ model. In particular, the Mg(b)–Mg(c) distance and the distance between Mg(b) and O1A of the incoming dNTP were much closer to the X-ray crystallographic distances. Further, in the case of the RNA ligase ribozyme [33], the Mg–O bond lengths obtained via a dummy approach [24] were more similar to those measured in QM/MM MD simulations and in the X-ray structures of other Mg²⁺ containing enzymes. In the PDB database, magnesium ions are invariably hexacoordinated with octahedral geometry. An analysis of structures available in the PDB to the 2002 indicated that the Mg–O bond length can range between 2.05 and 2.25 Å [24].

Equilibration MD trajectories lasting for 5 ns were computed with the aid of the NAMD 2.7 software package [34]. The smooth Particle-mesh Ewald (PME) method was employed for long-range electrostatic forces [35]. The non-bonded cutoff was set to 9 Å. The SHAKE algorithm (tolerance 0.0005) was applied to constrain bonds where the hydrogen atoms were involved [36]. Simulated systems were energy minimized, then the Langevin dynamics was used for a temperature control [34]. The simulated systems were heated from 0 to 310 K. The Langevin piston method was applied to reach an efficient pressure control with target pressure set to 1 atm [34]. The integration timestep was set to 2 fs. A multiprocessor system equipped with 16 Intel Itanium CPUs was used for equilibration MD simulations.

For production runs (reaching up to 1,050 ns) we have used ACEMD v. 2591 [37, 38] running on a local workstation

equipped with Nvidia GPU. ACEMD is a new generation molecular dynamics software which runs exclusively on graphics processing units (GPUs) at the equivalent speed of tens to hundred of standard processors. ACEMD implements all features of an MD simulation on a CUDA-compatible GPU device, including those usually required for production simulations in the NVT ensemble (i.e. bonded and non-bonded force term computation, velocity-Verlet integration, Langevin thermostatic control, smooth Ewald long-range electrostatics (PME), and hydrogen bond constraints implemented using the M-shake algorithm and RATTLE for velocity constraints within the velocity Verlet integration scheme [39, 40]). ACEMD does not presently contain a barostat for production runs, so simulations in the NPT ensemble are not possible. However, it is noted that with large molecular systems, changes in volume due to the pressure control are very limited after an initial equilibration making NVT simulations viable for production runs [37, 38]. It was proven on multiple biomolecular systems including membrane proteins [41–43]. Therefore, our production runs have been performed in the NVT ensemble, Langevin thermostat at 310 K, computing the electrostatic interactions with the particle-mesh Ewalds algorithm. The non-bonded cutoff was set to 9 Å. Also implemented in ACEMD is the hydrogen mass repartitioning scheme (the mass of the bonded heavy atoms to hydrogen is repartitioned among hydrogen atoms, leaving the total mass of the system unchanged) used, for instance, in codes such as Gromacs, which allows an increased time step of up to 4 fs [37, 44]. Therefore, the integration timestep in our production MD runs was set to 3 fs. The total energy was conserved within MD simulations (Fig. S6). For production MD runs CUDA programmable NVIDIA GTX-580 graphical processing units (GPUs, equipped with 480 cores) were used.

Data were recorded every 75 ps. MD trajectories were analyzed with the aid of the CHIMERA 1.5.3, VMD 1.9, Curves+ and AMBER10/ptraj software packages [26, 45–47]. Figures were produced by means of the ICM Molsoft 3.7 software package.

Results

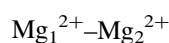
MD simulations of complexes consisting of NV RdRp and its natural (CTP, 2dCTP—Fig. S3) or modified (coCTP, cocCTP—Fig. S3) substrates were carried out. The closed active site of NV RdRp at the moment just before the phosphoryl transfer step was studied.

RMSD

Structural instabilities could appear in MD simulations from many reasons. Generally, crystal structures are

affected by crystal packing forces. Therefore, remarkable conformational changes may accompany the process of relaxation. Further, the closed state of NV RdRp was established in crystals due to its interactions with CTP. Analogs of CTP used in this study could stimulate a transition of NV RdRp to the open state. Moreover, instabilities can simply arise due to artifacts or imbalances in contemporary force fields [28].

Therefore, the stability of simulated biomolecules was assessed by the root-mean-square deviation (RMSD). Time evolution of RMSD for NV RdRp as well as for nucleic acids is provided in Fig. S7. It clearly shows that all simulated systems were stable over microsecond MD simulations.



In MD simulations of binuclear active sites [24], remarkable instabilities due to the excessive repulsion arising between magnesium ions (represented as VdW spheres with non-polarizable point charges +2) were observed. The seven-point representation (with dummy atoms) of magnesium ions should prevent these instabilities [24], but according to the best of our knowledge, no one has verified whether it will remain true on a microsecond time scale. The time evolution of $\text{Mg}_1^{2+}-\text{Mg}_2^{2+}$ and $\text{Mg}_{1,2}^{2+}$ -ligand distances shows (Fig. S8) that stable NV RdRp active sites were gained here.

Moreover, no tendency of NV RdRp to move from close to open state was found here. It would have to start through erosion of contacts between Mg_2^{2+} and Asp242 from Motif A (Figs. 1d, S2).

Nevertheless, remarkable differences in recognition of the 2'OH group of CTP, coCTP and cocCTP by Asp247 (which is part of Motif A as well—Fig. 1d) were observed (see below). Its potential impacts on transitions of NV RdRp either from open to close or from close to open state will be the subject for further investigation.

Structural analysis

Conformational preferences of CTP, 2dCTP, coCTP and cocCTP (Fig. S3) and their contacts with side chains of amino acids whose mutation significantly decreases the catalytic efficiency of RdRp were studied in detail here:

1. In past, one of us studied isosteric phosphonate linkages in the context of dinucleotides and oligonucleotides by means of ab initio calculations and MD simulations [48]. Large differences in conformational preferences of the torsion angles $\text{O5}'-\text{C5}'-\text{C4}'-\text{O4}'$ and $\text{C5}'-\text{O5}'-\text{C4}'-\text{O4}'$ were found (and confirmed by NMR [48]). It significantly influenced melting temperatures of whole double

helical structures. Interestingly, the same conformational preference (as observed for C5′–O5′–C4′–O4′ in [48]) was seen in the crystal structure of an isosteric phosphonate analogue GS-9148 bound in the active site of HIV RT (PDB id: 3kk1 [49]). Therefore, different conformational preferences of natural and modified NTPs (Fig. S3) were expected to appear here with serious implications on the geometry of the NV RdRp active site.

- In poliovirus 3Dpol RdRp, incorporation of a 2′-deoxyribonucleotide with a correct base should occur at a frequency 1/7,500 [50]. Therefore, conformational preferences of sugars (sugar pucker) were quantified here using the pseudorotation angle.
- In NV RdRp [3], the 2′-OH group of the incoming nucleoside triphosphate accepts a hydrogen bond from Asn309 and donates a hydrogen bond to Ser300, which in turn donates a hydrogen bond to Asp247. This pattern of hydrogen bonding shows how these highly conserved residues distinguish ribonucleotides from 2′-deoxyribonucleotides. Homologous residues in poliovirus 3Dpol RdRp are Asn297, Ser288, and Asp238 respectively. 3Dpol derivatives that contain alanine at position Asp238 or Asn297 are incapable of distinguishing between ATP and 2′-dATP. While the loss of the Asn297 side chain caused a 20-fold reduction in the observed rate constant for correct nucleotide incorporation, the loss of the Asp-238 side chain caused a 2,000-fold reduction [51]. Therefore, the 2′OH-Ser300–Asp247 mutual contacts were monitored here in detail (Fig. 1).
- In NV RdRp X-ray structures [3], the CTP bases form hydrogen bonds with the complementary guanosine base of the template and stack against the 3′-terminal base of the primer, as well as the Arg182 guanidino group. Arg182 is near the NTP α -phosphate. This arginine is highly conserved in RdRps. Arg158 is the homologous residue in HCV RdRp. In the presence of

Fig. 3 CTP conformational preference and its interactions in the NV RdRp active site: **A01** CTP O5′–C5′–C4′–O4′ torsion angle; **A02** CTP Sugar Puckering; **A03** Intramolecular hydrogen bond: CTP_3′OH– β -phosphate_CTP; **A04** Asp247–3′OH_CTP; **A05** Asp247–Ser300; **A06** Ser300–2′OH_CTP; **A07** Asp247–2′OH_CTP; **A08** Arg182– α -phosphate_CTP; **A09** Arg182– β -phosphate_CTP; **A10** Primer_O3′– α -phosphate_CTP

only Mg²⁺, the HCV RdRp R158A mutant was severely affected for all activities: initiation of RNA synthesis by a de novo mechanism, extension from a primed template, non-templated nucleotide addition, and synthesis of a recombinant RNA product from two or more noncovalently linked templates (template switch) [52]. Arg residues occupy a similar position in distantly related enzymes like HIV RT. The HIV RT R72A has the *k*_{cat} of the polymerase reaction reduced by nearly 100-fold with poly(rA) template [53]. In DNA polymerase β , Arg183 is found to form a salt bridge with the β -phosphate of an incoming NTP. A mutation of Arg183 to alanine significantly lowers the catalytic efficiency, which agrees well with the results from energy decomposition analysis, where the residues Arg183 (–7.1 kcal/mol) and Arg149 (–4.6 kcal/mol) were found to have the most significant contributions [20, 21]. The so-called arginine fingers play a crucial role in the GTP hydrolysis promoted by various GTPases [54–56]. The effect of Arg789 in Ras/GAP complexes is associated with its direct interaction with the transition state charge distribution [54]. Therefore, Arg182 contacts with triphosphate moieties of our NTPs were studied here in detail as well (Fig. 2).

- The distance separating reaction substrates—i.e. the phosphorus atom of the α -phosphate group of NTPs and the O3′-terminal atom of the RNA primer was followed here (Fig. S2b). However, the phosphate/phosphonate substitution in coCTP and cocCTP could affect the chemical step of the NV RdRp reaction. A direct comparison of the α P–O3′ distances thus

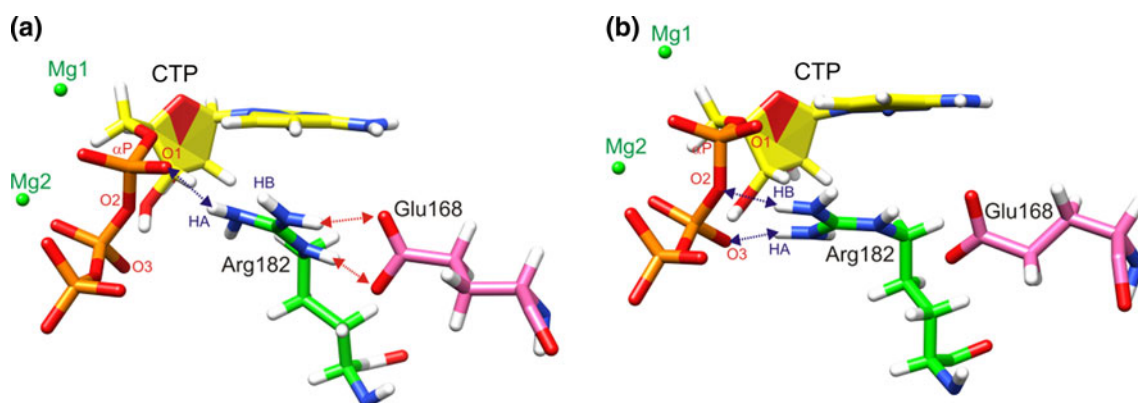
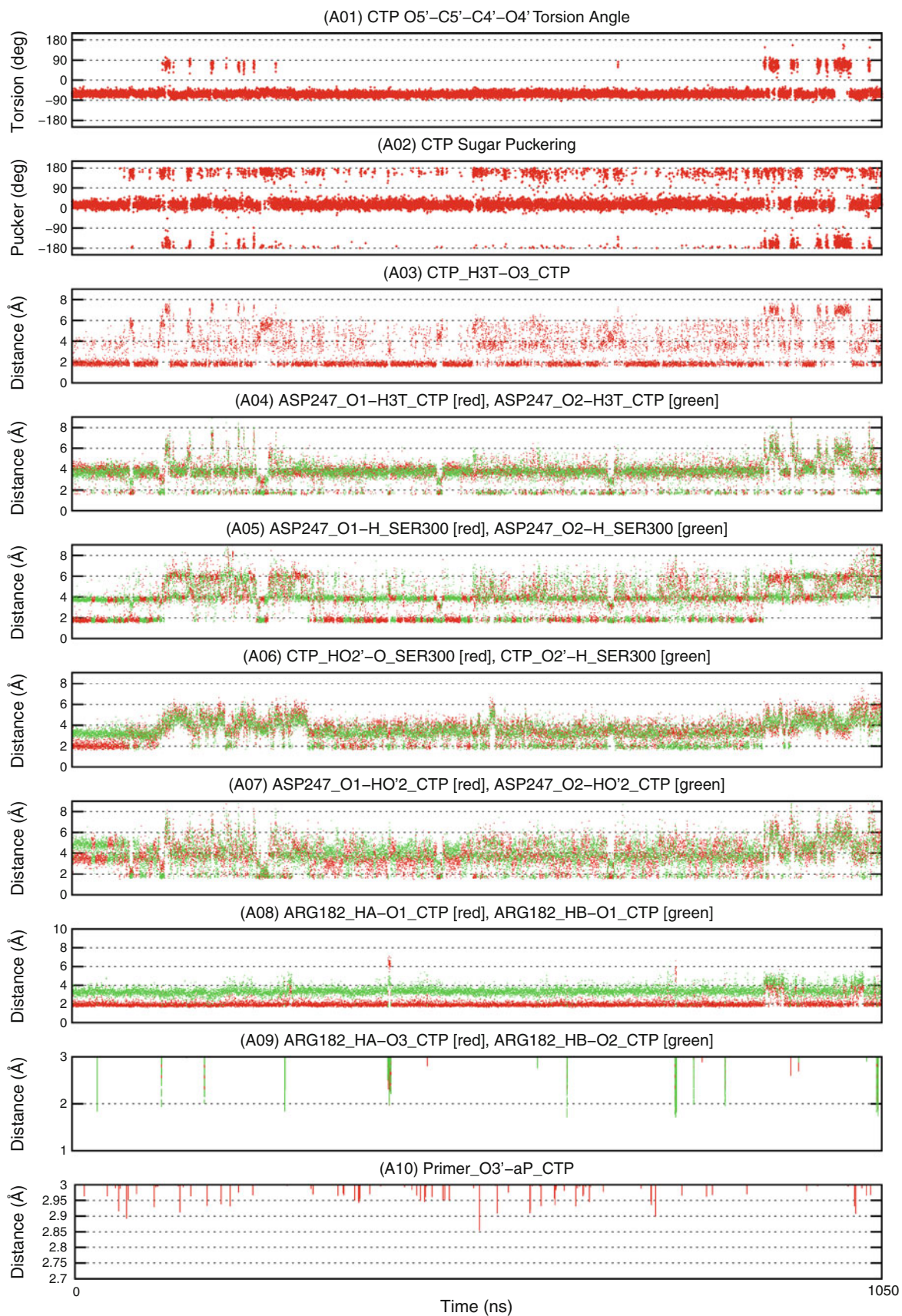


Fig. 2 a Arg182 interacting with the α -phosphate group of CTP; b Arg182 interacting with the β -phosphate group of CTP



probably makes sense especially in the case of closely related structures (CTP vs. 2dCTP) and (coCTP vs. cocCTP—where it could provide the most straightforward explanation of their different modes of action—competitive inhibition vs. chain termination).

CTP: Figs. 1a, 2, 3—Graph A01–A10

CTP was mostly in the canonical A-RNA conformation within the MD simulation. The O5′–C5′–C4′–O4′ torsion angle (Fig. S3) was in the **-gauche** conformation (Fig. 3—Graph A1) and the ribose moiety in the **C3′-endo** conformation (Fig. 3—Graph A2). Nevertheless, reversible short-living conformational transitions of the ribose moiety to **C2′-endo** were captured repeatedly within the microsecond MD trajectory. In moments, when contacts of CTP with amino acid side chains in the NV RdRp active site were temporarily disrupted (see below), the ribose passed even to **C3′-exo** and the O5′–C5′–C4′–O4′ torsion angle was in **+gauche**.

For most of the time, the 2′OH group of CTP was hydrogen bonded with Asp247 (Fig. 3—Graph A7) and Ser300 (Fig. 3—Graph A6). Asp247 and Ser300 interacted with each other as well (Fig. 3—Graph A5). In fact, all possible combinations of hydrogen bond donors (group 2′OH/atom HO2′; Ser300/H—see Fig. 1a) and acceptors (group 2′OH/atom O2′; Asp247/O1, O2; Ser300/O—see Fig. 1a) were applied. Where necessary, green and red lines in charts distinguish multiple hydrogen bonds of the same amino acid side chain. In crystal structures [3], the 2′-OH group of CTP donates a hydrogen bond to Ser300, which in turn donates a hydrogen bond to Asp247. MD simulations thus show that the hydrogen bond network is much more dynamic in solution.

Arg182 (atom HA—Fig. 2a) was in close contacts with the non-bridging oxygen atom of the α -phosphate group of CTP (atom O1—Fig. 2a) for most of the time (Fig. 3—Graph A8—red line). Arg182 was stabilized in this position by Glu168 (Fig. 2a). The polymerase reaction event would break the bond between the α P and O2 atoms (Fig. 2a) that should move away each from other. Therefore, we speculate that Arg182 binding to the α -phosphate group of CTP is more likely to hinder the reaction (of course—only future QM/MM calculations can be conclusive).

Nevertheless, about once every 100 ns, the Arg182 side chain lost any contacts with Glu168 and instead interacted with the β -phosphate group of CTP (atom O3—Fig. 2b) and with the oxygen atom bridging the α and β phosphate groups of CTP (atom O2—Fig. 2b, Fig. 3—Graph A9). It means that Arg182 interacted exclusively with the pyrophosphate moiety of CTP. The role of analogous Arg72

from HIV RT in the pyrophosphate release was demonstrated experimentally [53]. Moreover, the latter configuration (Fig. 2b) seems to be optimal for lowering of the transition state energy, because Arg182 interacts directly with the oxygen atom of the scissile bond (atom O2—Fig. 2b). Nevertheless, only future QM/MM calculations can be conclusive.

The distance separating reaction substrates—i.e. the phosphorus atom of the α -phosphate group of CTP and the O3′ atom in the RNA primer terminus (Fig. S2b) fell $2 \times / 18 \times$ below the 2.9/2.95 Å threshold (Fig. 3—Graph A10).

2dCTP: Figs. 1b, 4—Graph B01–B10

2dCTP started in the A-RNA-like conformation with the deoxyribose in the **C3′-endo** conformation (Fig. 4—Graph B2). **C2′-endo** was predominant after elapsing of about 250 ns. However, up to the end of the MD simulation repuckerings into the **C3′-endo** conformer lasting for several tens of ns were observed. It means that the minor **C3′-endo** conformers of 2dCTP survived substantially longer than the minor conformers **C2′-endo** of CTP.

This bias of 2dCTP toward **C3′-endo** seems to be partly caused by hydrogen bonding of the 3′OH group (atom H3T—Fig. 1a) toward the β -phosphate group of 2dCTP (atom O3—Fig. 1a)—compare (Fig. 4—Graph B2 with B3). This interaction was also applied in CTP (unlike all phosphonate analogs of CTP).

The O5′–C5′–C4′–O4′ torsion angle (Fig. S3) was mostly in the **-gauche** conformation (Fig. 4—Graph B1) that was preferred in CTP as well (Fig. 3—Graph A1).

Overall, 2dCTP was in the same conformation as CTP within a significant part of the MD simulation (i.e. the deoxyribose in the **C3′-endo** conformation and the O5′–C5′–C4′–O4′ torsion angle in the **-gauche** conformation). Despite of it, 2dCTP/CTP contacts with amino acid side chains in the NV RdRp active site were remarkably different.

Contacts between Asp247 and Ser300 residues were lost irreversibly once the deoxyribose repuckering into **C2′-endo** appeared (compare Fig. 4—Graph B5 with Graph B2). These contacts were not re-established again, apparently due to the absence of the 2′OH group, which could participate in a dynamic network of hydrogen bonds with Asp247 and Ser300.

Further, conformational transitions of the deoxyribose moiety disturbed salt bridges between Arg182 and the α -phosphate group of 2dCTP (Figs. 2a, 4—Graph B8). The alternative arrangement with Arg182 bound to the oxygen atom involved in the α P–O2 scissile bond, was observed just three times (Figs. 2b, 4—Graph B9)—i.e. less frequently than for CTP.

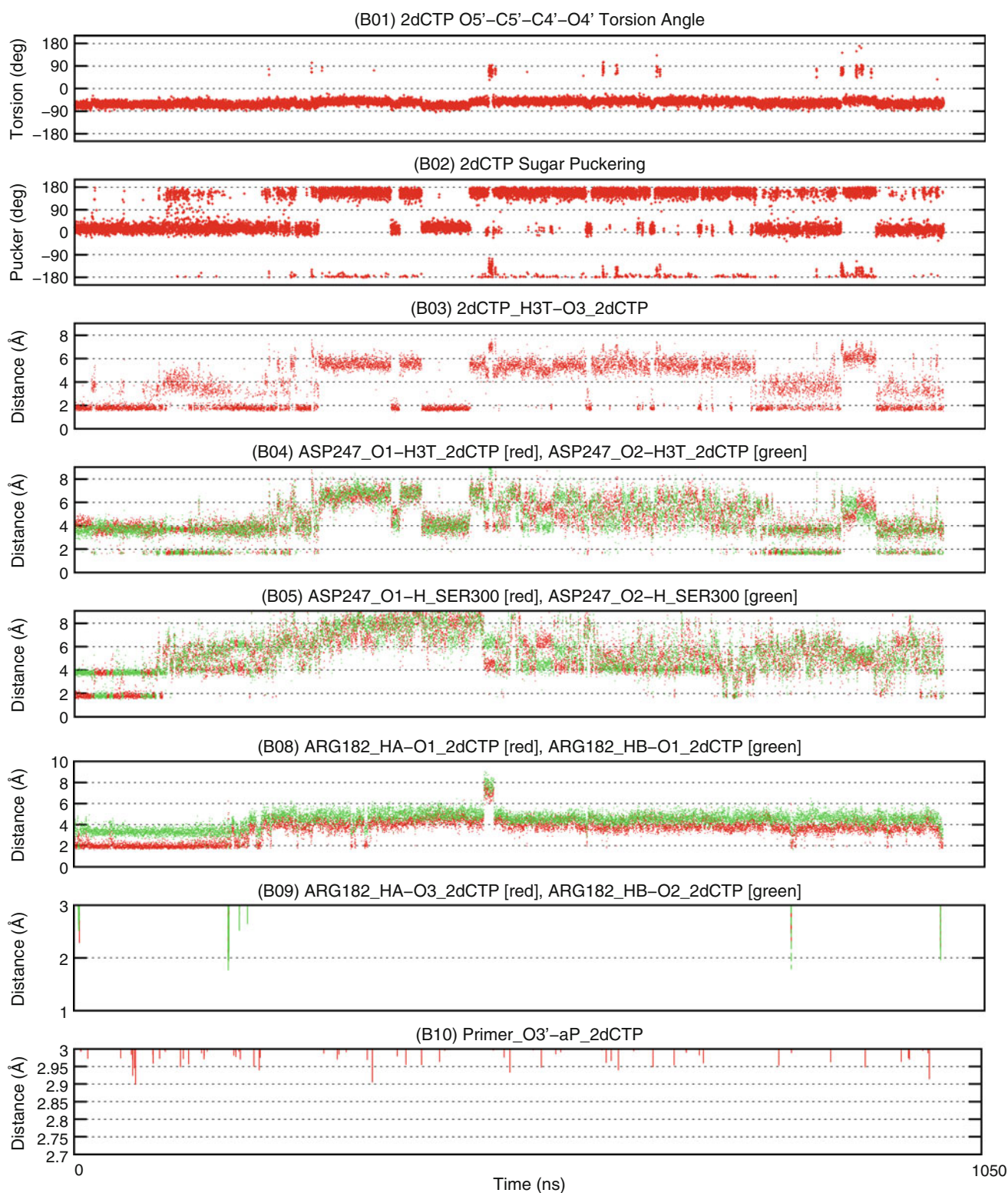


Fig. 4 2dCTP conformational preference and its interactions in the NV RdRp active site: **B01** 2dCTP O5'-C5'-C4'-O4' torsion angle; **B02** 2dCTP Sugar Puckering; **B03** Intramolecular hydrogen bond:

2dCTP_3'OH- β -phosphate_2dCTP; **B04** Asp247-3'OH_2dCTP; **B05** Asp247-Ser300; **B08** Arg182- α -phosphate_2dCTP; **B09** Arg182- β -phosphate_2dCTP; **B10** Primer_O3'- α -phosphate_2dCTP

The crucial substrate–substrate distance (between the phosphorus atom in the α -phosphate group of 2dCTP and the O3' atom of the RNA primer terminus) was not dropping below 2.9 Å. It fell seven times below the 2.95 Å threshold (Fig. 4—Graph B10)—i.e. less frequently than for CTP.

coCTP: Figs. 1c, 5—Graph C01–C10

The torsion angle O5'–C5'–C4'–O4' in CTP and 2dCTP adopted the **-gauche** conformer (Fig. 3—Graph A1, Fig. 4—Graph B1). In contrast, the torsion angle C5'–O5'–C4'–O4' in coCTP (Fig. S3) preferred **+gauche** (Fig. 5—Graph C1). Therefore, it was reasonable to expect a relaxation of the coCTP positioning in the active site of NV RdRp. Surprisingly, it took hundreds of nanoseconds.

The coCTP ribose was distorted from **C3'-endo** to **C2'-exo** (Fig. 5—Graph C2) Nevertheless, short-living transitions into many different conformers were observed at the beginning of the MD simulation. Then conformational fluctuations of the C5'–O5'–C4'–O4' torsion angle from **+gauche** to **trans** appeared (Fig. 5—Graph C1). Passover of these transitions was clearly associated with a substantial rearrangement of a network of hydrogen bonds in the NV RdRp active site. It was not completed until the end of the second third of the MD simulation (Fig. 5—Graph C6).

Contacts between residues Asp247 and Ser300 disappeared after elapsing of roughly 300 ns (Fig. 5—Graph C5). Until Asp247–Ser300 contacts were lost, Ser300 served rather as a hydrogen bond acceptor from the 2'OH of coCTP. After that, Ser300 was a hydrogen bond donor for 2'OH. This contact was remarkably stable (Fig. 5—Graph C6).

Arg182 (stabilized by Glu168) was steadily in contact with the α -phosphonate group of coCTP throughout the whole MD simulation (Figs. 2a, 5—Graph C08). Just one conformational transition of the Arg182 side chain toward the oxygen atom in the O2- α P scissile bond was observed (Fig. 5—Graph C09).

The phosphorus atom of the α -phosphonate group of coCTP and the activated oxygen atom O3' of the RNA primer approached repeatedly ($7 \times / > 40 \times$) below the 2.9/2.95 Å threshold (Fig. 5—Graph C10)—i.e. more a little bit more frequently than for cocCTP.

cocCTP: Figs. 1d, 6—Graph D01–D10

The C5'–O5'–C4'–C4'' torsion angle (Fig. S3) was mostly in the **+gauche** conformation (Fig. 6—Graph D1) as in coCTP. The relaxation of the cocCTP position in the NV RdRp active site was even slower now (roughly 500 ns)

Fig. 5 coCTP conformational preference and its interactions in the NV RdRp active site: **C01** coCTP C5'–O5'–C4'–O4' torsion angle; **C02** coCTP Sugar Puckering; **C03** Intramolecular hydrogen bond: coCTP_3'OH– β -phosphate_coCTP; **C04** Asp247–3'OH_coCTP; **C05** Asp247–Ser300; **C06** Ser300–2'OH_coCTP; **C07** Asp247–2'OH_coCTP; **C08** Arg182– α -phosphonate_coCTP; **C09** Arg182– β -phosphate_coCTP; **C10** Primer_O3'– α -phosphonate_coCTP

and was made up after repuckering of the cocCTP modified sugar (the O4' atom was replaced by the $-\text{CH}_2-$ group—Fig. S3).

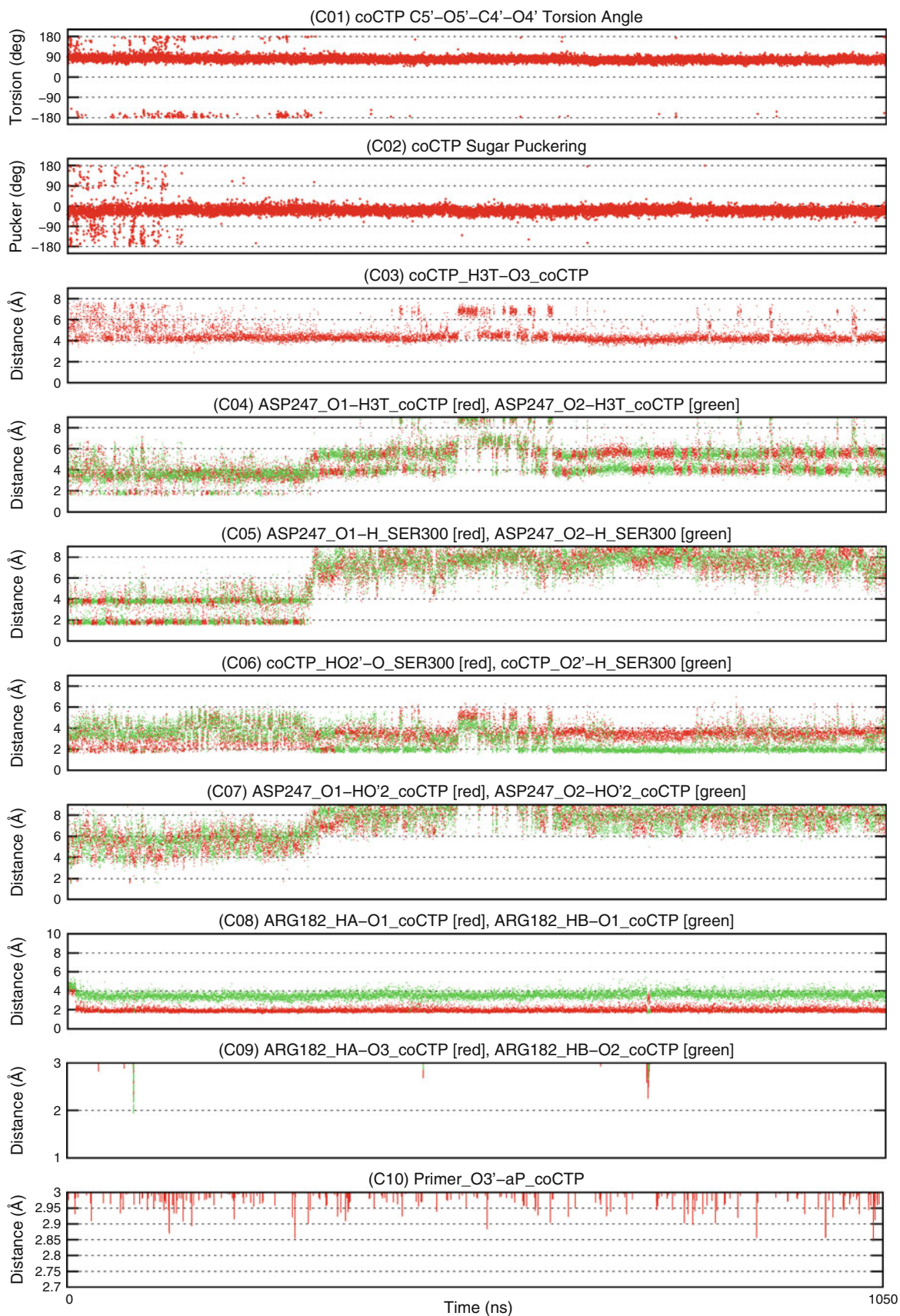
The modified sugar of cocCTP preferred the **C2'-exo** conformation almost exclusively up to 500 ns (Fig. 6—Graph D2).

While the modified sugar adopted the **C2'-exo** conformation, the Asp247–Ser300–2'OH_cocCTP mutual contacts were remarkably stable. In fact, the Asp247–Ser300 hydrogen bonding was the most stable of all MD simulations carried out in this study (Figs. 3, 4, 5, 6—Graph A5, B5, C5, D5). It's partly due to the 3'OH group interactions with Asp247, which were considerably more stable comparing to what was observed in other MD simulations (Fig. 6—Graph D4). In CTP and 2dCTP, the 3'OH group was involved rather in the intramolecular hydrogen bonding toward the triphosphate moiety (Fig. 3—Graph A3, Fig. 4—Graph B3).

While the modified ribose have adopted the **C3'-exo** conformation, Asp247–Ser300 contacts were broken. The modified sugar of cocCTP was then in contact with the bulky side-chain of Trp246 and the 3'OH group of cocCTP was occasionally hydrogen bonded to Gln250 and Asn309. In fact, cocCTP penetrated a little bit deeper into the active site of NV RdRp beyond Asn247 in Motif A (represented as ribbon in Fig. 1).

Figure S9 shows the time evolution of the distance between the C2' atom in the 3'-terminal residue of the RNA primer (Fig. S2b) and the C1' atom in xCTPs (Fig. S2b). A significant increase of the distance is evident in the second half of the cocCTP MD simulation. This change was caused by too close contacts between the $-\text{CH}_2-$ group (which replaced the O4' atom in the modified sugar moiety of cocCTP—Fig. S3) and a hydrogen atom bound to the C2' atom in the 3'-terminal residue of the RNA primer (Fig. S2b).

Results of preliminary calculations indicate that a steric conflict between these groups of atoms could cause that cocCTP is an inhibitor and not a chain terminator as coCTP. The topology of existing models was modified, that a bond between the α -phosphate/phosphonate group of CTP (coCTP, cocCTP) and its β/γ -pyrophosphate moiety was canceled producing cytidine monophosphate (CMP) or its phosphonate analogs (coCMP, cocCMP). Instead a linkage between the terminal O3' atom of the RNA primer and the α -phosphate/phosphonate group of CMP (coCMP,



cocCMP) was created. The simulated system adjusted so that it corresponds to the state immediately after the phosphoryl transfer step of the polymerase reaction (instead of the state just before the reaction) without causing a conformational change in the CMP moiety. In coCMP, the torsion angle $C5'-O5'-C4'-O4'$ smoothly switched from **+gauche** to **trans** during this process. In contrast, in the case of cocCMP, the $C5'-O5'-C4'-C4''$ torsion angle transition from **+gauche** to **trans** was slowed down by a steric conflict between the above mentioned groups of atoms. The question is whether this process was done, if the newly formed bond was not represented by a simple harmonic oscillator. Only future empirical valence bond (EVB) [57] and quantum mechanical/molecular mechanical (QM/MM) [58–61] calculations could resolve it.

Potentially catalytic Arg182 (stabilized by Glu168) was all the time in contact with the α -phosphonate group of cocCTP (Fig. 6—Graph D8).

The distance that separates the phosphorus atom of the α -phosphonate group from the $O3'$ -end of the RNA primer drop $5 \times / 36 \times$ below $2.9/2.95 \text{ \AA}$ (Fig. 6—Graph D10)—i.e. a little bit less frequently than for coCTP.

Conclusions

Microsecond MD simulations of complexes formed between NV RdRp and its natural (CTP, 2dCTP—Fig. S3) or chemically modified (coCTP, cocCTP—Fig. S3) substrates were produced. Different conformational preferences of natural versus modified substrates were expected to appear here based on our previous experience with similarly modified dinucleotides and oligonucleotides. NMR evidences [48] that while the $O5'-C5'-C4'-O4'$ torsion angle of natural dinucleotides prefers the **-gauche** conformer, the $C5'-O5'-C4'-O4'$ torsion angle of phosphonate dinucleotides adopts the **+gauche** rotamer. Indeed, the $O5'$ oxygen atom of CTP and 2dCTP was located above the ribose ring (Fig. 1a,b), whereas the $C5'$ atom of coCTP and cocCTP analogs was directed outside this ring (Fig. 1c,d). The same conformational preference was seen in the crystal structure of a similar isosteric phosphonate analogue GS-9148 in the active site of HIV RT (PDB id: 3kk1 [49]). Further, we looked at how xCTPs interact with the side chains of amino acids whose mutation significantly decreases the catalytic efficiency of RdRp (i.e. Arg183 [52] interacting with the triphosphate moiety and Asp247 [50, 51] involved together with Ser300 and Asn309 in recognition of the $2'OH$ group). Major differences in conformational preferences and binding of the natural substrates (CTP, 2dCTP) and of their isosteric phosphonate

Fig. 6 cocCTP conformational preference and its interactions in the NV RdRp active site: **D01** cocCTP $C5'-O5'-C4'-C4''$ torsion angle; **D02** cocCTP Sugar Puckering; **D03** Intramolecular hydrogen bond: cocCTP_3'OH- β -phosphate_cocCTP; **D04** Asp247-3'OH_cocCTP; **D05** Asp247-Ser300; **D06** Ser300-2'OH_cocCTP; **D07** Asp247-2'OH_cocCTP; **D08** Arg182- α -phosphonate_cocCTP; **D09** Arg182- β -phosphate_cocCTP; **D10** Primer_3'- α -phosphonate_cocCTP

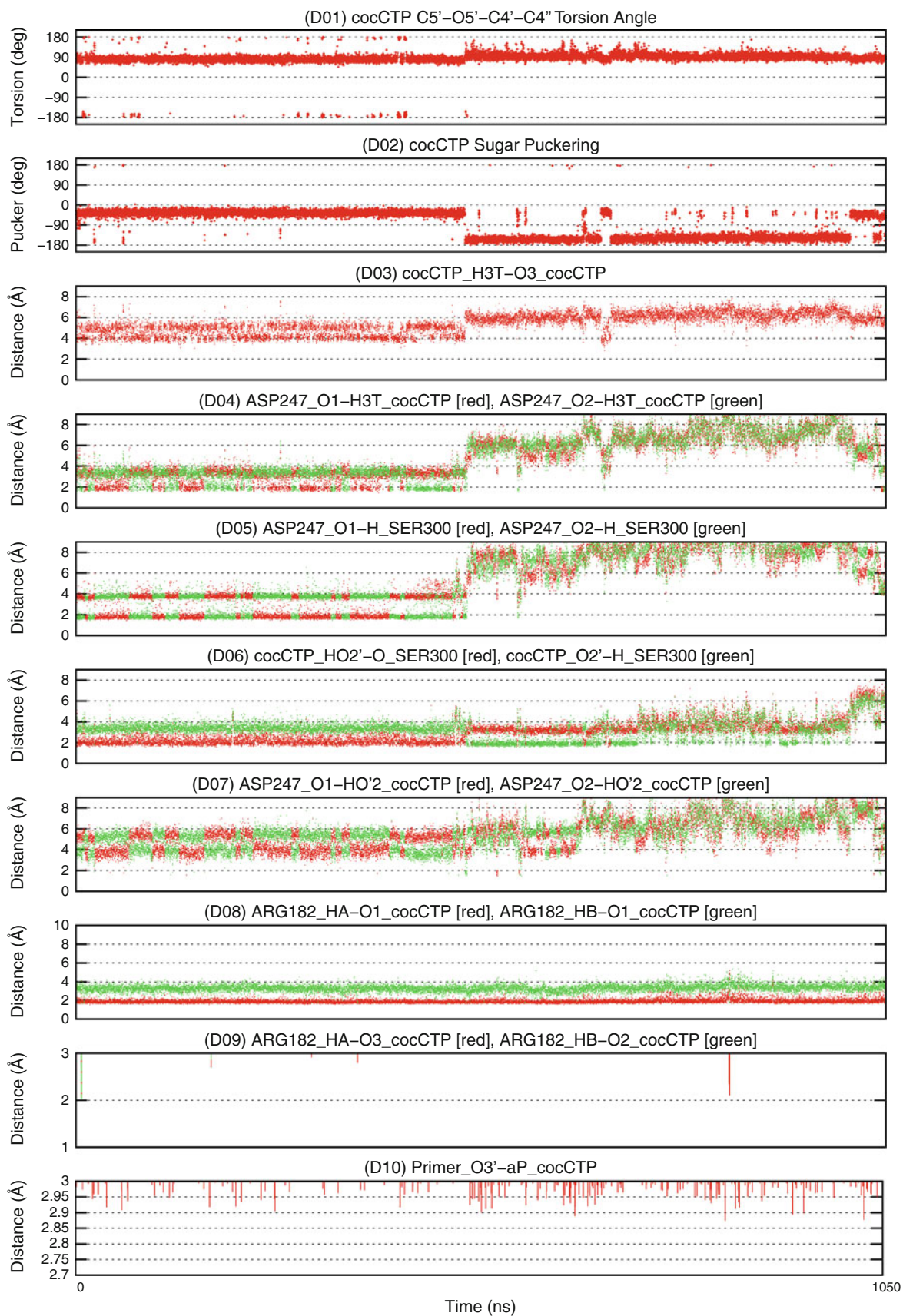
analog (coCTP, cocCTP) into the active site of NV RdRp are depicted on Fig. 1.

The CTP ribose was in the **C3'-endo** conformation and the $O5'-C5'-C4'-O4'$ torsion angle adopted the **-gauche** conformer. The Asp247–Ser300–2'OH_CTP hydrogen bond network (known from crystal structures [3, 4], Fig. 1a) was vital. Nevertheless, all imaginable combinations of hydrogen bond donors and acceptors were realized in the flow of time. About once every 100 ns, a substantial conformational change of the Arg182 side chain was observed. Arg182 then interacted with the β -phosphate instead of the α -phosphate group of CTP (compare Fig. 2a, b).

The 2dCTP deoxyribose preferred the **C2'-endo** conformation. However, **C3'-endo** was populated too—partly due to the intramolecular bonding between the 3'OH and β -phosphate group of 2dCTP (also observed in CTP—but not in CTP analogs). Contacts between Asp247 and Ser300 side chains were irreversibly interrupted (recall that 2dCTP lacks the 2'OH hydroxyl group that could participate in mutual contacts). The 3'OH group of 2dCTP occasionally interacted with the side chain of Asp247 (Fig. 1b). Further, the Arg182 side chain lost contacts with the triphosphate moiety of 2dCTP. It clearly shows that binding of proper CTP and undesired 2dCTP building blocks into the NV RdRp active site differs in many aspects.

In coCTP, the $C5'-O5'-C4'-O4'$ torsion angle preferred the **+gauche** rotamer from the very beginning of the MD simulation (Fig. 1c). The ribose was in the **C3'-endo/C2'-exo** conformation. Yet positional relaxations of coCTP in the NV RdRp active site took place on the timescale of hundreds of ns. After their completion, the 2'OH group of coCTP was in close contacts with Ser300 and also near to Asn309. Arg178 formed salt bridges with the α -phosphonate group of coCTP during the whole MD simulation.

In cocCTP, the $C5'-O5'-C4'-C4''$ torsion angle preferred again the **+gauche** rotamer from the very beginning of the MD simulation. Initially, the modified sugar of cocCTP was in the **C2'-exo** conformation (as in coCTP). At that time the Asp247–Ser300–2'OH_cocCTP hydrogen-bond network was extremely stable, since the 3'OH group of cocCTP was involved too. After of about 500 ns, the modified sugar moiety repuckered to **C3'-exo**. cocCTP consequently penetrated a little bit deeper (Fig. S9) into the active site of NV RdRp beyond Asn 247 in Motif A (represented as ribbon in Fig. 1d). The modified sugar of cocCTP was then in contact with the bulky side-chain of



Trp246 and the 3'OH group of cocCTP was occasionally hydrogen bonded to Gln250 and Asn309. This conformational transition was caused by too close contacts between the -CH₂- group of atoms (which replaced the O4' atom in the modified sugar moiety of cocCTP—Fig. S3) and a hydrogen atom bound to the C2' atom in the 3'-terminal residue of the RNA primer (Fig. S2b). Results of preliminary calculations indicate that a steric conflict between these atoms could cause that cocCTP is an inhibitor and not a chain terminator as coCTP.

Summarized, the **+gauche** conformer of the C5'–O5'–C4'–O4' and C5'–O5'–C4'–C4'' torsion angles was preferred in coCTP and cocCTP from the beginning of MD simulations (as expected based on NMR [48] and X-ray [49] studies of closely related systems). Surprisingly, the full positional relaxation of xCTP analogs in the NV RdRp active site took hundreds of nanoseconds. In hindsight, the microsecond length of MD trajectories appears to be rather a necessary minimum.

In here presented MD simulations, the NV RdRp active site in a state just before the phosphoryl-transfer step was studied. To address the question what is happening before and after this moment, different computational approaches will be required in which, however, conformers from current MD simulations will serve as starting points. Without extensive sampling of conformational space carried out in this study, those subsequent calculations were based on misleading initial structures especially in the case of CTP analogs.

Future EVB [57] and QM/MM [58–61] calculations will deal with the phosphoryl transfer step. In this case, snapshots from current MD simulations, when the O3'– α P substrate–substrate distance was minimal, will be used. Further, here presented MD simulations raised several questions, which evidently can be addressed only by QM/MM calculations. In the case of natural NTPs, Arg182 is a catalytic residue [52]. Nevertheless, which from the alternative conformers of Arg182 observed here is responsible for it? Could contacts of Arg182 with the α -phosphonate group of coCTP and cocCTP rather hinder the phosphoryl transfer step?

A complete process of a ligand binding into protein active sites was studied recently by atomic MD simulations [43]. For polymerases, analogous calculations are unfeasible using conventional hardware resources especially due to the strong binding of magnesium ions to ligands resulting in a very slow sampling of configurational space in MD simulations. Here presented MD trajectories starting with fully formed coordination shells of magnesium ions show that merely minor positional relaxations of xCTP analogues in the NV RdRp active site can take hundreds of nanoseconds. Anyway, snapshots from here presented MD trajectories can serve as starting points for MD

simulations showing how CTP (and its analogues) leaves the NV RdRp active site, which is moving from close to open state. This reverse process can be effectively accelerated by means of metadynamics [62, 63], replica exchange molecular dynamics [64, 65] or steered molecular dynamics [66, 67].

Acknowledgments This work was supported by the Grant Agency of the Czech Republic (202/09/0193). The access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme “Projects of Large Infrastructure for Research, Development, and Innovations” (LM2010005) is highly acknowledged.

References

- Sofia MJ, Chang W, Furman PA, Mosley RT, Ross BS (2012) Nucleoside, nucleotide, and non-nucleoside inhibitors of hepatitis C virus NS5B RNA-dependent RNA-polymerase. *J Med Chem* 55:2481–2531
- De Francesco R, Migliaccio G (2005) Challenges and successes in developing new therapies for hepatitis C. *Nature* 436:953–960
- Zamyatkin DF, Pirra F, Alonso JM, Harki DA, Peterson BR, Grochulski P, Ng KK-S (2008) Structural insights into mechanisms of catalysis and inhibition in norwalk virus polymerase. *J Biol Chem* 283:7705–7712
- Zamyatkin DF, Parra F, Machín A, Grochulski P, Ng KK-S (2009) Binding of 2'-amino-2'-deoxycytidine-5'-triphosphate to norovirus polymerase induces rearrangement of the active site. *J Mol Biol* 390:10–16
- Paeshuyse J, Vliegen I, Coelmont L, Leyssen P, Tabarrini O, Herdewijn P, Mittendorfer H, Easmon J, Cecchetti V, Bartenschlager R, Puerstinger G, Neyts J (2008) Comparative in vitro anti-hepatitis c virus activities of selected series of polymerase, protease, and helicase inhibitors. *Antimicrob Agents Chemother* 52(9):3433–3437
- Bressanelli S, Tomei L, Rey FA, De Francesco R (2002) Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. *J Virol* 76:3482–3492
- Ng KK, Pendas-Franco N, Rojo J, Boga JA, Machin A, Alonso JM, Parra F (2004) Crystal structure of norwalk virus polymerase reveals the carboxyl terminus in the active site cleft. *J Biol Chem* 279:16638–16645
- Gong P, Peersen OB (2010) Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc Natl Acad Sci* 107:22505–22510
- Arnold JJ, Gohara DW, Cameron CE (2004) Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mn²⁺. *Biochemistry* 43:5138–5148
- Huang HF, Chopra R, Verdine GL, Harrison SC (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* 282:1669–1675
- Migliaccio G, Tomassini JE, Carroll SS, Tomei L, Altamura S, Bhat B, Bartholomew L, Bosserman MR, Ceccacci A, Colwell LF, Cortese R, De Francesco R, Eldrup AB, Getty KL, Hou XS, LaFemina RL, Ludmerer SW, MacCoss M, McMasters DR, Stahlhut MW, Olsen DB, Hazuda DJ, Flores OA (2003) Characterization of resistance to non-obligate chain-terminating ribonucleoside analogs that inhibit hepatitis C virus replication in vitro. *J Biol Chem* 278(49):49164–49170

12. Zlatev I, Dutartre H, Barvik I, Neyts J, Canard B, Vasseur JJ, Alvarez K, Morvan F (2008) Phosphoramidate dinucleosides as hepatitis C virus polymerase inhibitors. *J Med Chem* 51(18): 5745–5757
13. Priet S, Zlatev I, Barvik I Jr, Geerts K, Leyssen P, Neyts J, Dutartre H, Canard B, Vasseur JJ, Morvan F, Alvarez K (2010) 3'-Deoxy phosphoramidate dinucleosides as improved inhibitors of hepatitis C virus subgenomic replicon and NS5B polymerase activity. *J Med Chem* 53:6608–6617
14. Wales DL, Schooley RT, Kaihara KA, Beadle JR, Hostetler KY (2008) Anti-hepatitis C virus replicon activity of alkoxyalkyl esters of (S)-HPMPA and other acyclic nucleoside phosphonates. *Antiviral Res* 78(2):A21
15. Koh Y, Shim JH, Wu JZ, Zhong W, Hong Z, Girardet JL (2005) Design, synthesis, and antiviral activity of adenosine 5'-phosphonate analogues as chain terminators against hepatitis C virus. *J Med Chem* 48:2867–2875
16. Sheridan C (2012) Calamitous HCV trial casts shadow over nucleoside drugs. *Nature Biotechnol* 30:1015–1016
17. Radhakrishnan R, Schlick T (2006) Correct and incorrect nucleotide incorporation pathways in DNA polymerase beta. *Biochem Biophys Res Commun* 24:521–529
18. Radhakrishnan R, Schlick T (2005) Fidelity discrimination in DNA polymerase beta: differing closing profiles for a mismatched (G:A) versus matched (G:C) base pair. *J Am Chem Soc* 127:13245–13253
19. Radhakrishnan R, Schlick T (2005) Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase beta's closing. *Proc Natl Acad Sci* 101:5970–5975
20. Lin P, Pedersen LC, Batra VK, Beard WA, Wilson SH, Pedersen LG (2006) Energy analysis of chemistry for correct insertion by DNA polymerase beta. *Proc Natl Acad Sci* 103:13294–13299
21. Lin P, Batra VK, Pedersen LC, Beard WA, Wilson SH, Pedersen LG (2008) Incorrect nucleotide insertion at the active site of a G:A mismatch catalyzed by DNA polymerase beta. *Proc Natl Acad Sci* 105:5670–5674
22. Wang L, Yu X, Hu P, Brodey S, Zhang Y (2007) A water-mediated and substrate-assisted catalytic mechanism for *Sulfolobus solfataricus* DNA polymerase IV. *J Am Chem Soc* 129: 4731–4737
23. Sucato CA, Upton TG, Kashemirov BA, Batra VK, Martinek V, Xiang Y, Beard WA, Pedersen LC, Wilson SH, McKenna CE, Florian J, Warshel A, Goodman MF (2007) Modifying the beta, gamma leaving-group bridging oxygen alters nucleotide incorporation efficiency, fidelity and the catalytic mechanism of DNA polymerase beta. *Biochemistry* 46:461–471
24. Oelschlaeger P, Klahn M, Beard WA, Wilson SH, Warshel A (2007) Magnesium-cationic dummy atom molecules enhance representation of DNA polymerase beta in molecular dynamics simulations: improved accuracy in studies of structural features and mutational effects. *J Mol Biol* 366:687–701
25. Jorgensen WL, Chandrasekhar J, Madura J, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
26. Pearlman DA, Case DA, Caldwell JW, Ross WR, Cheatham TE, DeBolt S, Ferguson D, Seibel G, Kollman P (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput Phys Commun* 91:1–41
27. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic acids, and organic-molecules. *J Am Chem Soc* 117:5179–5197
28. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE III, Laughton ChA, Orozco M (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys J* 92:3817–3829
29. Dal Peraro M, Spiegel K, Lamoureux G, De Vivo M, De Grado WF, Klein ML (2007) Modeling the charge distribution at metal sites in proteins for molecular dynamics simulations. *J Struct Biol* 157:444–453
30. Xiang Y, Oelschlaeger P, Florian J, Goodman MF, Warshel A (2006) Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: evaluating amino acid group contribution and allosteric coupling for ionized residues in human pol beta. *Biochemistry* 45:7036–7048
31. Florian J, Goodman MF, Warshel A (2003) Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. *J Am Chem Soc* 125:8163–8177
32. Aqvist J (1990) Ion-water interaction potentials derived from free energy perturbation simulations. *J Phys Chem* 94:8021–8024
33. Sgrignani J, Magistrato A (2012) The structural role of Mg²⁺ + ions in a class I RNA polymerase ribozyme: a molecular simulation study. *J Phys Chem B* 116:2259–2268
34. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
35. Cheatham TE, Miller JL, Fox T, Darden TA, Kollman PA (1994) Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J Am Chem Soc* 117:4193–4194
36. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341
37. Harvey MJ, Giupponi G, De Fabritiis G (2009) ACEMD: accelerating biomolecular dynamics in the microsecond time scale. *J Chem Theory Comput* 5:1632–1639
38. ACEMD <http://www.acellera.com/>
39. Andersen HC (1983) RATTLE: a “velocity” version of the SHAKE algorithm for molecular dynamics calculations. *J Comput Phys* 52:24–34
40. Lambrakos SG, Boris JP, Oran ES, Chandrasekhar I, Nagumo M (1989) A modified SHAKE algorithm for maintaining rigid bonds in molecular dynamics simulations of large molecules. *J Comput Phys* 85:473–486
41. Giorgino T, Gianni de Fabritiis (2011) A high-throughput steered molecular dynamics study on the free energy profile of ion permeation through gramicidin A. *J Chem Theor Comput* 7:1943–1950
42. Selent M, Sanz F, Pastor M, De Fabritiis G (2010) Induced effects of sodium ions on dopaminergic G-protein coupled receptors. *PLOS Comput Biol* 6:e1000884
43. Buch I, Giorgino T, De Fabritiis G (2011) Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc Natl Acad Sci* 108:10184–10189
44. Feenstra KA, Hess B, Berendsen HJC (1999) Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J Comput Chem* 20:786–798
45. Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. *J Mol Graph* 14:33–38
46. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612
47. Lavery R, Moakher M, Maddocks JH, Petkeviciute D, Zakrzewska K (2009) Conformational analysis of nucleic acids revisited: curves+. *Nucleic Acids Res* 37:5917–5929

48. Točík Z, Buděšínský M, Barvík I Jr, Rosenberg I (2009) Conformational evaluation of labeled C3'-O-P-(13)CH(2)-O-C4'' phosphonate internucleotide linkage, a phosphodiester isostere. *Biopolymers* 91:514–529
49. Lansdon EB, Samuel D, Lagpacan L, Brendla KM, White KL, Hung M, Liu X, Boojamra CG, Mackman RL, Cihlar T, Ray AS, McGrath ME, Swaminathan S (2010) Visualizing the molecular interactions of a nucleotide analog, GS-9148, with HIV-1 reverse transcriptase-DNA complex. *J Mol Biol* 397:967–978
50. Arnold JJ, Cameron CE (2004) Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg²⁺. *Biochemistry* 43:5126–5137
51. Gohara DW, Crotty S, Arnold JJ, Yoder JD, Andino R, Cameron CE (2000) Poliovirus RNA-dependent RNA polymerase (3Dpol). *J Biol Chem* 275:25523–25532
52. Ranjith-Kumar CT, Sarisky RT, Gutshall L, Thomson M, Kao CC (2004) De Novo initiation pocket mutations have multiple effects on hepatitis C virus RNA-dependent RNA polymerase activities. *J Virol* 78:12207–12217
53. Sarafianos SG, Pandey VN, Kaushik N, Modak MJ (1995) Site-directed mutagenesis of arginine 72 of HIV-1 reverse transcriptase. *J Biol Chem* 270:19729–19735
54. Glennon TM, Villa J, Warshel A (2000) How does GAP catalyze the GTPase reaction of Ras?: a computer simulation study. *Biochemistry* 39:9641–9651
55. te Heesen H, Gerwert K, Schlitter J (2007) Role of the arginine finger in Ras. RasGAP revealed by QM/MM calculations. *FEBS Lett* 581:5677–5684
56. Oldham WM, Hamm HE (2006) Structural basis of function in heterotrimeric G proteins. *Q Rev Biophys* 39:117–166
57. Kamerlin SCL, Warshel A (2011) The empirical valence bond model: theory and applications. *Wiley Interdiscip Rev Comput Mol Sci* 1:30–45
58. Sgrignani J, Magistrato A (2013) First-principles modeling of biological systems and structure-based drug-design. *Current Comput Aided Drug Design* 9:15–34
59. Senn HM, Thiel W (2009) QM/MM methods for biomolecular systems. *Angew Chem Int Ed* 48:1198–1229
60. Groenhof G (2013) Introduction to QM/MM simulations. *Methods Mol Biol* 924:43–66
61. Vreven T, Byun KS, Komaromi I, Dapprich S, Montgomery JA Jr, Morokuma K, Frisch MJ (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J Chem Theory Comput* 2:815–826
62. Laio A, Gervasio FL (2008) Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep Progr Phys* 71:126601
63. Gervasio FL, Laio A, Parrinello M (2005) Flexible docking in solution using metadynamics. *J Am Chem Soc* 127:2600–2607
64. Grater F, de Groot BL, Jiang H, Grubmuller H (2006) Ligand-release pathways in the pheromone-binding protein of bombyx mori. *Structure* 14:1567–1576
65. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151
66. Li MS, Mai BK (2012) Steered molecular dynamics—a promising tool for drug design. *Curr Bioinform* 7:342–351
67. Grubmuller H, Heymann B, Tavan P (1996) Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271:997–999

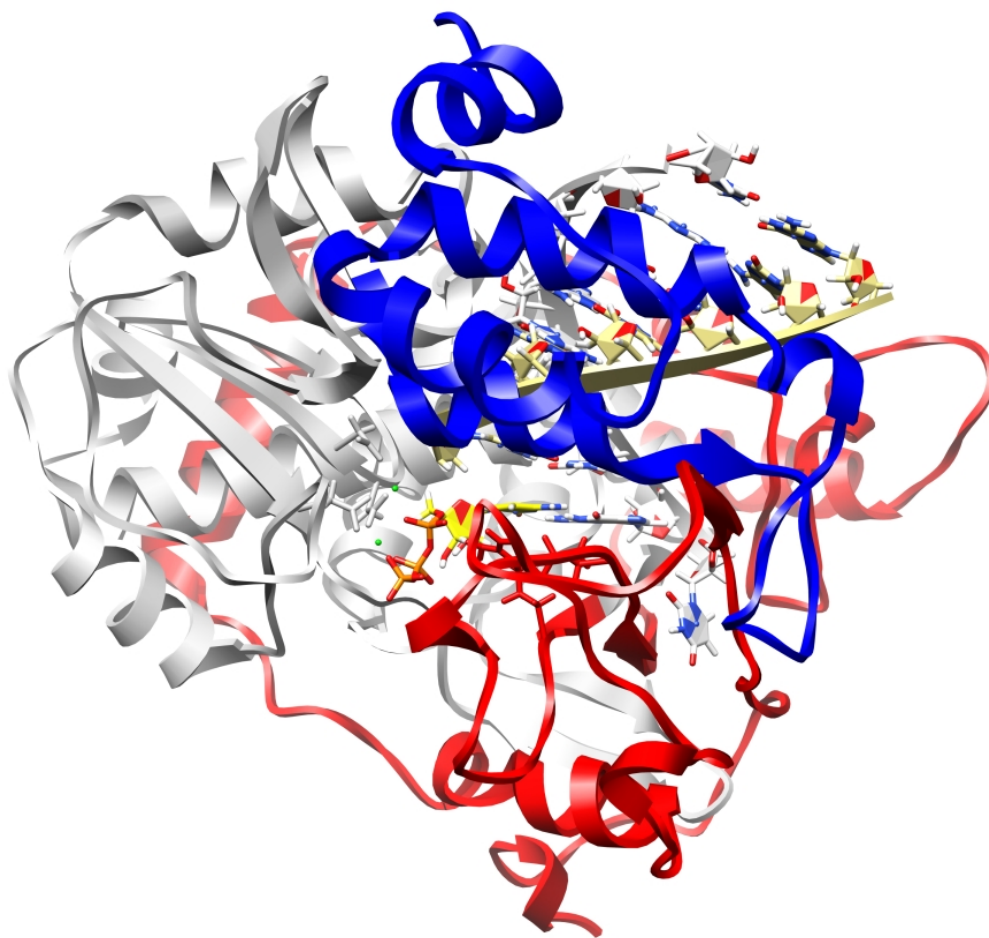


Figure S1: NV RdRp in complex with nucleic acids (PDB id: 3BSO [3-4]). NV RdRp has a typical structure of polymerases resembling the right hand (palm - grey, thumb - blue, fingers - red). In RdRps, thumb and fingers are bound, and hence not allowed to change conformation independently.

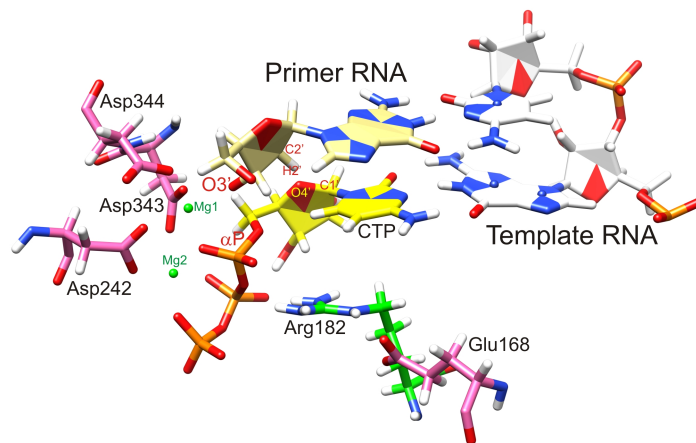
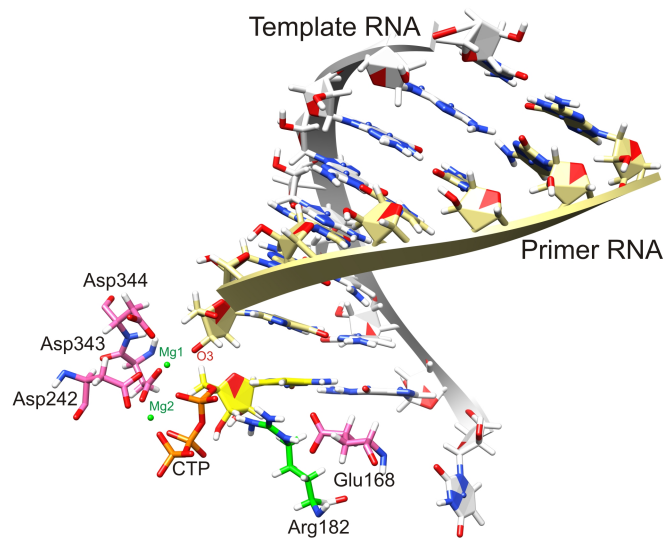


Figure S2: CTP in the active site of NV RdRp shown in the context of Primer and Template RNA. Mg^{2+} ions are bound by conserved aspartic acids (Asp344, Asp343, Asp242) in the palm domain of NV RdRp.

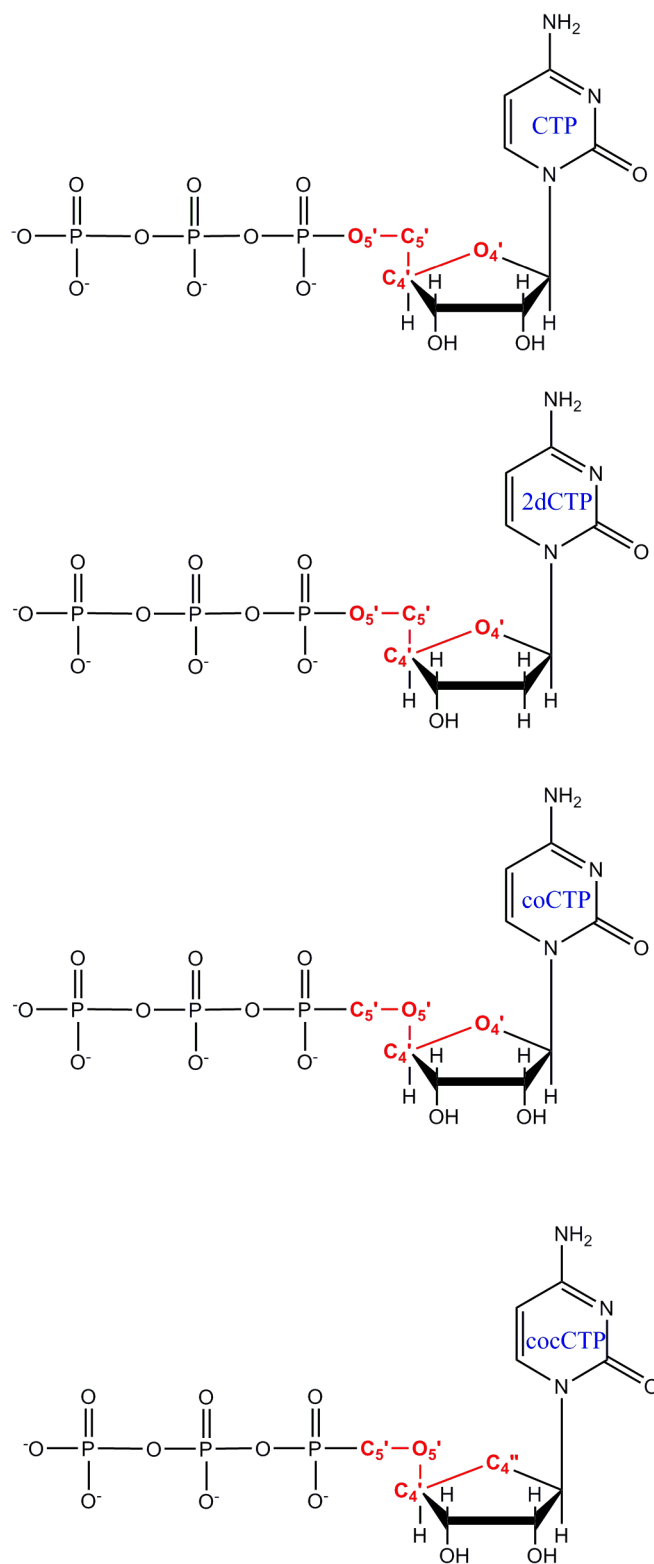


Figure S3: Structure of CTP, 2dCTP, coCTP and cocCTP.

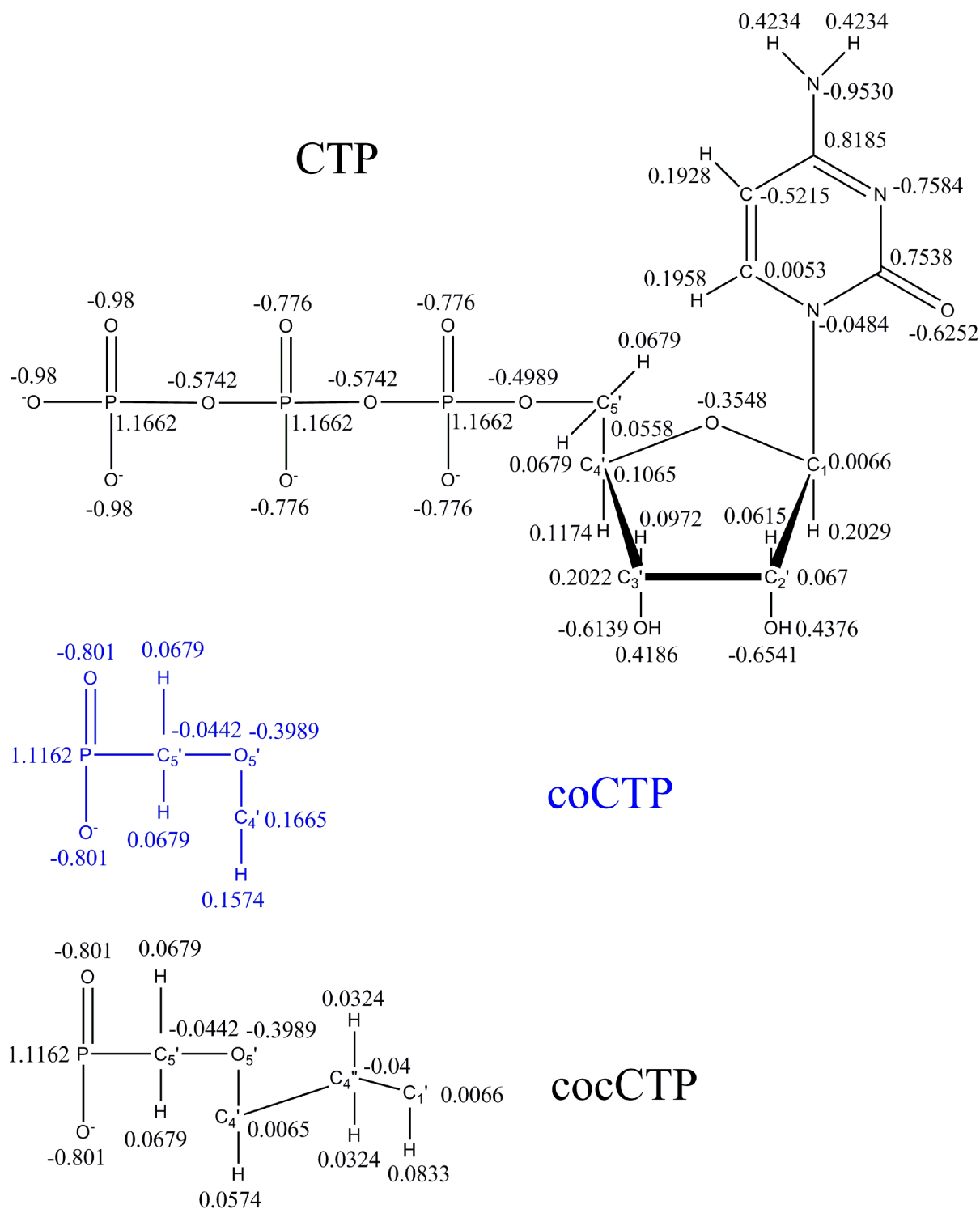


Figure S4: Point charges

BOND				
CT-P	186.0	1.846		
ANGLE				
OS-P -CT	45.0	102.60		
P -CT-HC	50.0	110.00		
P -CT-OS	50.0	110.00		
O2-P -CT	100.0	108.23		
HC-CT-OS	50.0	110.00		
OS-CT-OS	50.0	109.50		
DIHEDRAL				
X -CT-P -X	9	1.20	0.0	3.
CT-P -OS-CT	1	0.00	0.0	-3.
CT-P -OS-CT	1	0.00	0.0	-2.
CT-P -OS-CT	1	1.00	180.0	1.
OS-CT-P -OS	1	0.00	0.0	-3.
OS-CT-P -OS	1	0.00	0.0	-2.
OS-CT-P -OS	1	0.00	0.0	1.
P -CT-OS-CT	1	0.00	0.0	3.

Figure S5: Force constants for modified parts of CTP analogs.

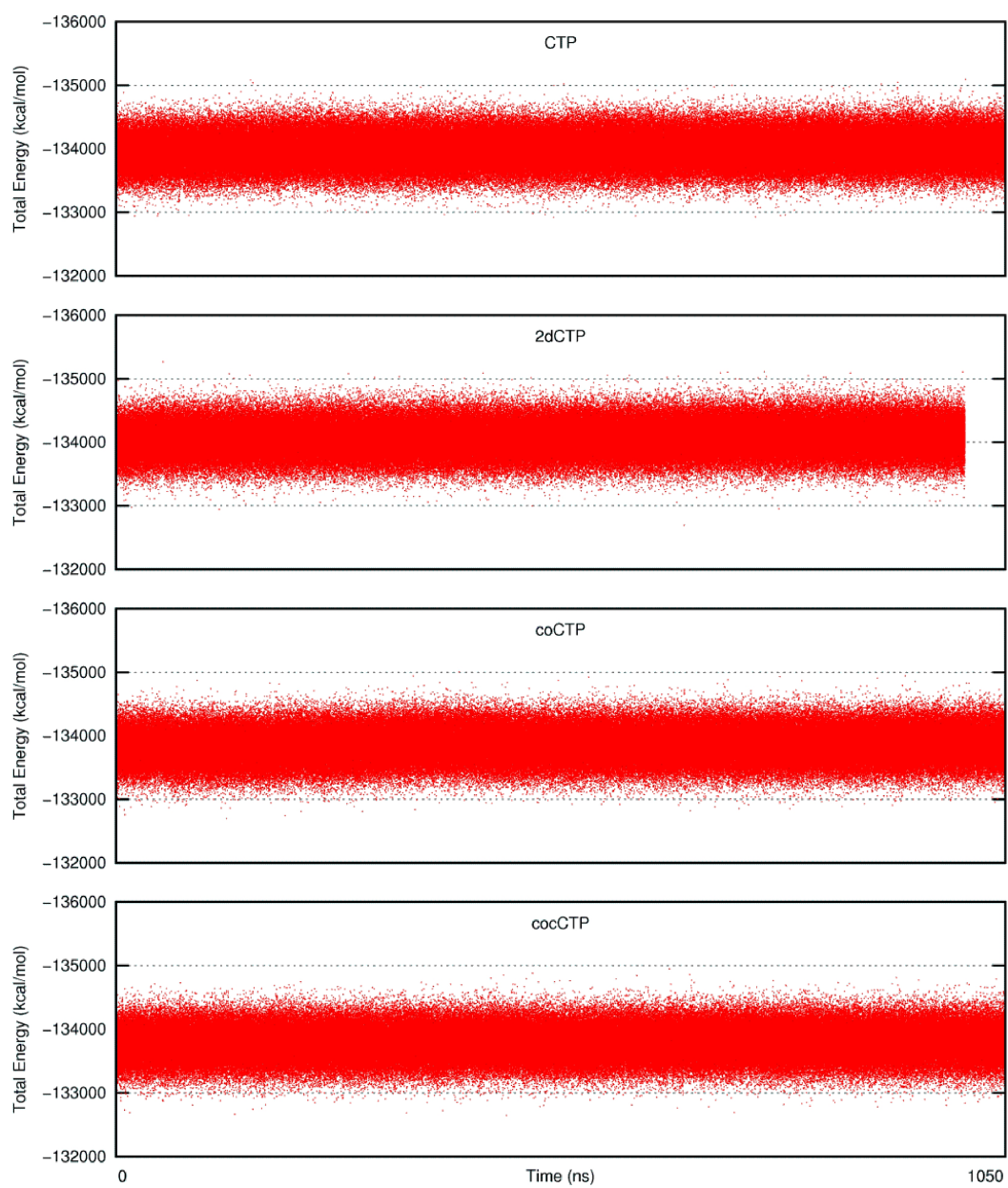


Figure S6: Time evolution of the total energy within MD simulations.

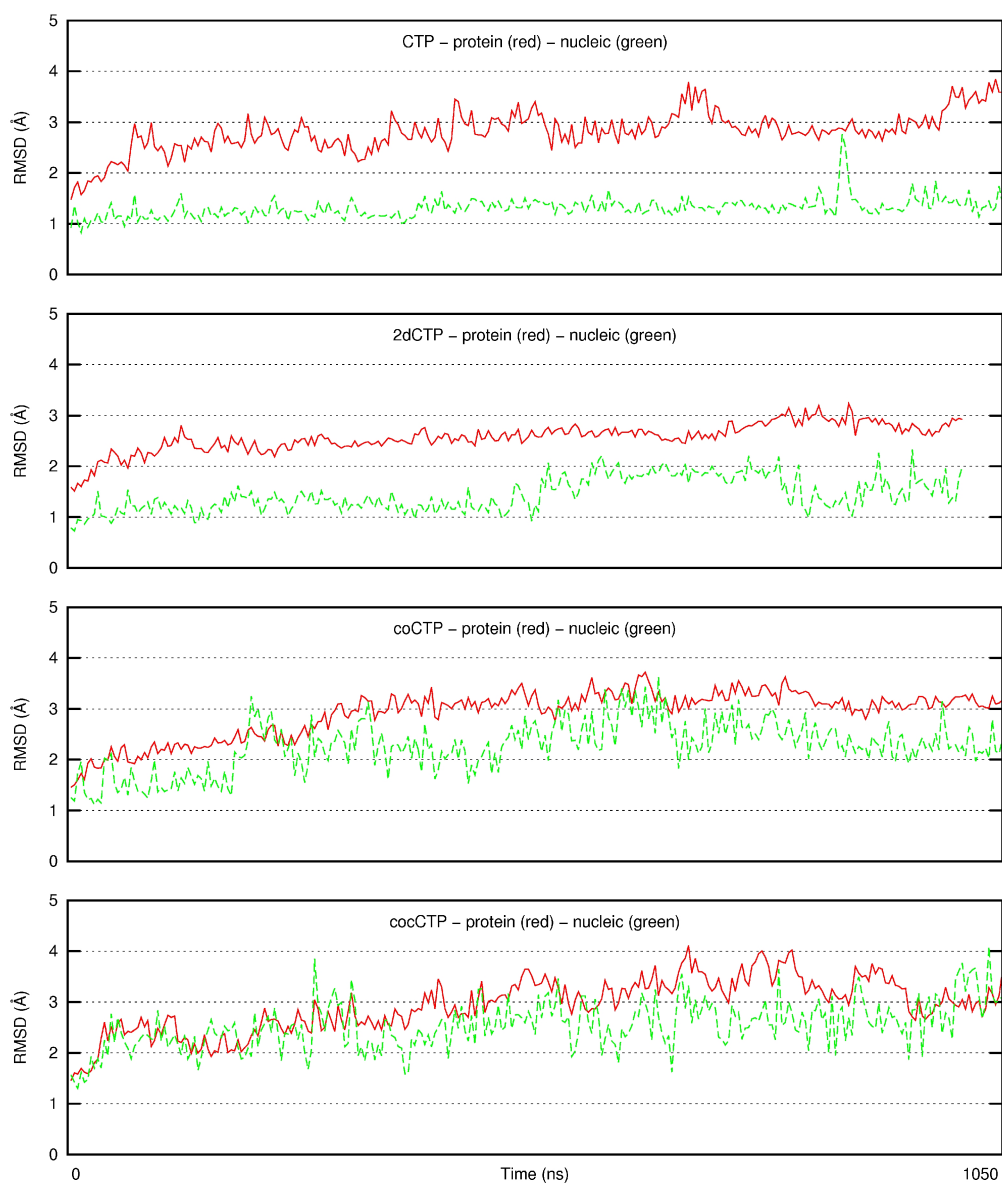


Figure S7: Time evolution of RMSD for NV RdRp and nucleic acids.

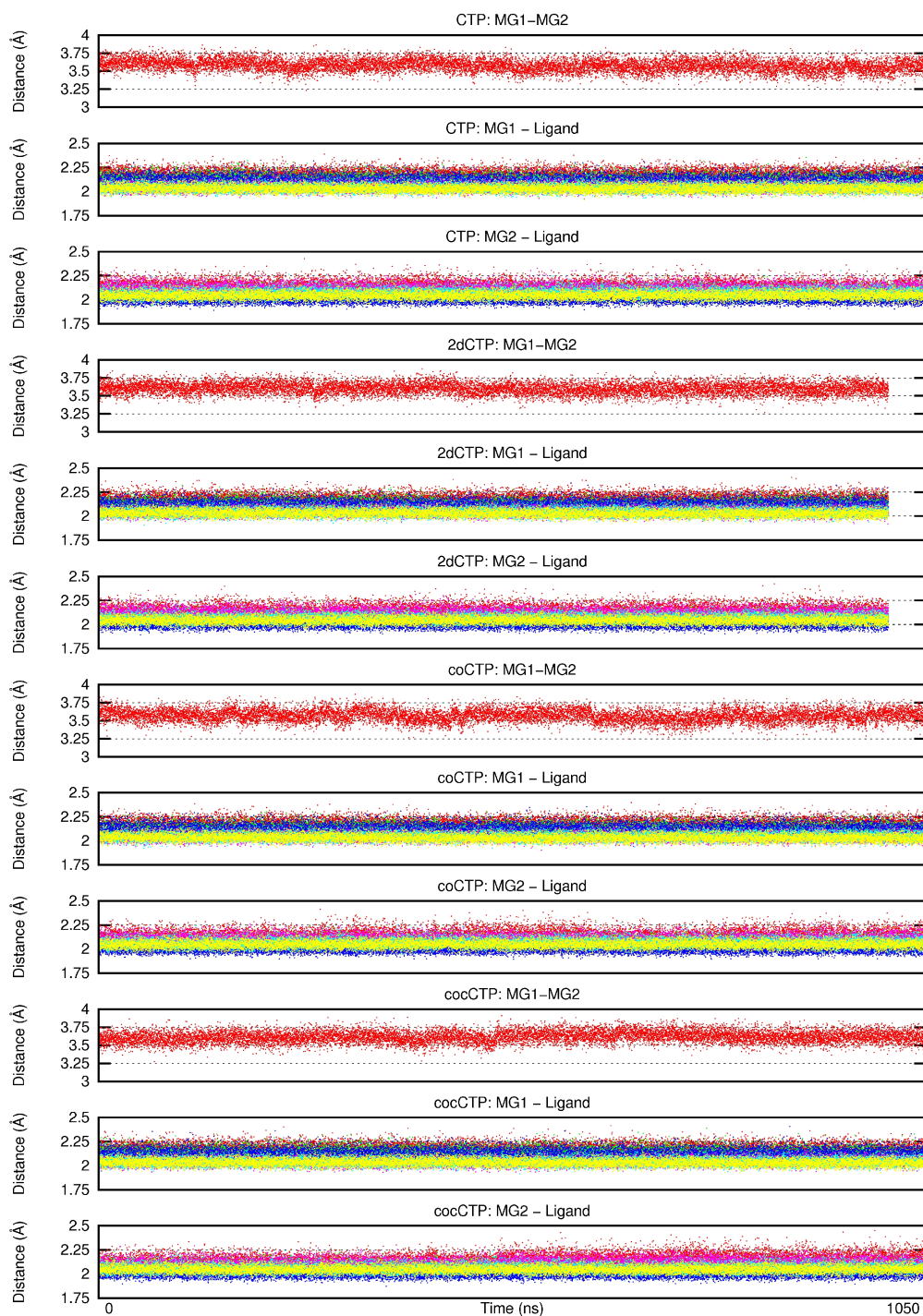


Figure S8: Time evolution of Mg_1^{2+} - Mg_2^{2+} , Mg_1^{2+} - ligand and Mg_2^{2+} - ligand distances in the NV RdRp active site. Mg_1^{2+} ligands (see Figure S2b) are distinguished in charts by the following color code: water molecule (red line), 3' RNA terminus (green), α -phosphate/phosphonate group of CTP/2dCTP/coCTP/cocCTP (blue), Asp344 (magenta), Asp343 (cyan), Asp242 (yellow). Mg_2^{2+} ligands are distinguished in charts by the following color code: α -phosphate/phosphonate group of CTP/2dCTP/coCTP/cocCTP (red), β -phosphate group (green), γ -phosphate group (blue), Tyr243 (magenta), Asp242 (cyan), Asp343 (yellow)

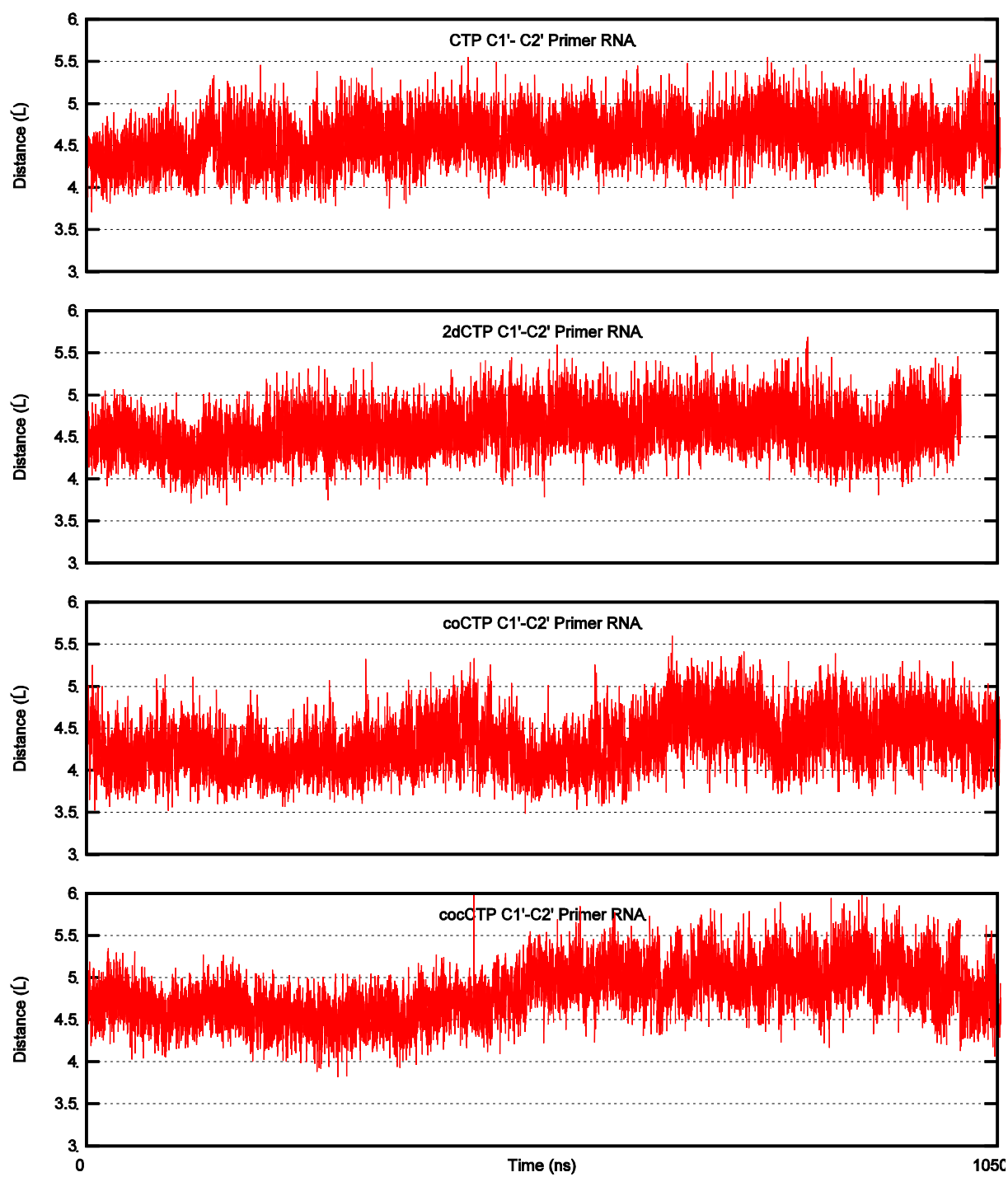
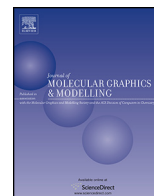


Figure S9: Time evolution of distances between RNA_primer_C2' and C1'_xCTP atoms.



Complex between *Human* RNase H1 and the phosphonate-DNA/RNA duplex: Molecular dynamics study



Kamil Maláč, Ivan Barvík*

Charles University, Faculty of Mathematics and Physics, Institute of Physics, Ke Karlovu 5, Prague 2 121 16, Czech Republic

ARTICLE INFO

Article history:

Received 7 October 2012
Received in revised form 30 April 2013
Accepted 5 May 2013
Available online 17 May 2013

Keywords:

RNase H
Antisense oligonucleotides
Argonaute
RNAi
Phosphonate
Molecular dynamics
ACEMD

ABSTRACT

Our 200 ns MD simulations show that even fully modified oligonucleotides bearing the 3'-O-P-CH₂-O-5' (but not 3'-O-CH₂-P-O-5') phosphonate linkages can be successfully attached to the surface of *Human* RNase H. It enables to explain that oligonucleotides consisting of the alternating 3'-O-P-CH₂-O-5' phosphonate and phosphodiester linkages are capable to elicit the RNase H activity (while the 3'-O-CH₂-P-O-5' phosphonates are completely inactive). Stability of the binuclear active site of *Human* RNase H was achieved using the one-atom model for Mg²⁺ in conjunction with a polarized phosphate group of the scissile bond, which is wedged between both magnesium ions. The reference MD simulation (lasting for 1000 ns), which was produced using a well-established seven-point (with dummy atoms) model for Mg²⁺ led to essentially the same results. The MD run (lasting for 500 ns) produced for the *Thermus thermophilus* Argonaute enzyme shows the transferability of our approach for the stabilization of a binuclear active site. Glu512 was bound in the *T. thermophilus* Argonaute active site to the 2'-OH of the nucleotide adjacent to the scissile phosphate and one of the two active-site divalent metal ions in exactly the same way as Glu186 in *Human* RNase H. Glu512 thus completes the catalytic tetrad of Argonaute.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

So called “antisense” oligonucleotides represent a perspective approach in chemotherapy, promising to inhibit selectively unwanted gene expression by creation of a helical complex with target mRNA (carrying “sense” genetic information) [1–8]. Several antisense oligonucleotide constructs, selected from many hundreds, were able to proceed to clinical trials and use [9,10]. If hybridization between the target mRNA and the exogenous nucleotide sequence occurs, a duplex is created, which forms a jam that prevents the ribosomal complex from reading along the message. If the ribosomal complex cannot read the message, the appropriate tRNAs are not assembled and the encoded peptide is not made. However, RNA–DNA duplexes can be unwound by a variety of repair/editing enzymes. In addition, the ribosomal complex itself has an unwindase activity that likely permits reading of the complexly fold mRNA [7]. Fortunately, some deoxyoligonucleotides support the binding of ribonuclease H (RNase H) at sites of RNA–DNA duplex formation. Such binding is thought to be an important effector of antisense actions because once bound, RNase H, a ubiquitous nuclear enzyme required for DNA synthesis, functions as an endonuclease that recognizes and cleaves the RNA

in the duplex. Once initiated, destruction of the message by cleavage is assured. Of significant interest also is the fact that the DNA comprising in the duplex is undamaged by the enzymatic attack. Therefore, it is free to hybridize with multiple RNA molecules, leading their destruction in a catalytic manner [7].

RNA interference (RNAi, discovered in 1998 [11–13]) is a eukaryotic-specific gene-silencing pathway triggered by double-stranded RNA (dsRNA) [14]. In this pathway, the RNase III enzyme Dicer first cleaves the dsRNA trigger into small interfering RNAs (siRNAs). The siRNA duplex is incorporated into the effector protein Argonaute (AGO) – analogous to RNase H, at which point one of the strands is cleaved. After the cleaved strand is discarded, the resulting ribonucleoprotein complex uses the remaining siRNA strand to specify interactions with target RNAs. If sequence complementarity between quide and target strands is extensive, AGO again catalyzes cleavage, resulting in ‘slicing’ of the target RNA [14].

The oligonucleotides with natural chemical composition, the first candidates for the “antisense/siRNA” drugs, have been, however, found as unsuitable for in vivo applications because of their insufficient resistance against nucleases. Oligonucleotides consisting of both the 3'-O-P-CH₂-O-5' and phosphodiester linkages are capable to elicit RNase H activity [15]. The oligonucleotides with either the 3'-O-P-CH₂-O-5' or 3'-O-CH₂-P-O-5' linkage has been investigated by means of molecular dynamics (MD) simulations [16–18]. Binding of DNA (and of its phosphonate analogs) toward EC RNase H was modeled [19]. Plausibility of this model

* Corresponding author. Tel.: +420 221 911 450; fax: +420 224 922 797.
E-mail address: ibarvik@karlov.mff.cuni.cz (I. Barvík).

was confirmed later by crystal structures of HS RNase H [20]. There are many analogous residues (EC/HS RNase H – Trp85/Tr225, Trp81/Trp221, Thr43/Thr181, etc.) bound to nucleic acids with truly perfect compliance at the atomic level. Further, impacts of phosphonate internucleotide linkages placed in various positions along DNA on its binding toward EC RNase H were investigated in past by means of short (1.5 ns) MD simulations [19].

Here, graphics processing units (GPUs) were used to extend such MD trajectories up to 200 ns. The ability of 3'-O-P-CH₂-O-5' and 3'-O-CH₂-P-O-5' phosphonate oligonucleotides to bind toward HS RNase H was studied in detail. Stability of the binuclear active site of HS RNase H was achieved using the one-atom model for Mg²⁺ in conjunction with a polarized phosphate group of the scissile bond (wedged between both magnesium ions). The reference MD simulation (lasting for 1000 ns and produced using a well-established seven-point model for Mg²⁺ [21]) led to essentially the same results. In future, we intend to compare MD trajectories gained in this study with those carried out for differently modified DNA chains.

RNase H forms its active site during initial folding. This active site is well-suited to its role in nonspecifically cleaving RNA–DNA hybrids. In contrast, proper function of AGO requires high specificity, therefore a conformational change occurs to form its active site [14]. Although the PIWI domain of AGO has an RNase H fold, only a conserved 'DDX' catalytic triad (where 'X' is generally Asp or His) had been recognized in AGOs with slicer activity. Nevertheless, the recent crystal structure of *Kluyveromyces polysporus* Argonaute (KpAGO) [14] uncovered a hydrogen-bond network that stabilizes an expanded and repositioned loop, which inserts an invariant glutamate into the catalytic pocket. The structures of ternary *Thermus thermophilus* Argonaute (TtAGO) complexes in the glutamate plugged-in conformation show the position of loop L2 in the context of guide and target strands [22]. TtAGO loop L2 interacts with the guide DNA at positions 11–15. Moreover, the carboxyl group of the glutamate finger Glu512 approaches both the 2'-OH of the nucleotide adjacent to the scissile phosphate and one of the two active-site divalent metal ions, which indicates that the glutamate finger might act as a catalytic residue [14,22]. The putative DEDD catalytic tetrads in conformations of both TtAGO and KpAGO are nearly isosteric with the RNase H DEDD tetrad. Moreover, only mutation of KpAGO Glu1013 (homologous to TtAGO Glu512) abrogated RNAi to the extent observed for mutation of KpAGO Asp1046, a previously identified active-site residue [14]. Thus, it was proposed that the glutamate finger constitutes the second residue of a universally conserved RNase H-like DEDX catalytic tetrad at the active site of slicing AGOs [14].

The MD run (lasting for 500 ns) produced here for TtAGO shows the transferability of our approach for the stabilization of a binuclear active site. Glu512 was bound in the TtAGO active site to the 2'-OH of the nucleotide adjacent to the scissile phosphate and one of the two active-site divalent metal ions in exactly the same way as Glu186 in HS RNase H. Glu512 thus completes the catalytic tetrad of TtAGO.

2. Methods

The model of HS RNase H complexed with a phosphonate-DNA:RNA hybrid was constructed using the recent crystal structure [20] (Protein Data Bank entry 2QKK). Base pairs not participating in the DNA:RNA binding toward the enzyme were omitted from the simulated system. Three modified internucleotide linkages (either the 3'-O-P-CH₂-O-5'/3pc5 or 3'-O-CH₂-P-O-5'/3pc5 type) were incorporated into the 3'-end of DNA (i.e. 3'-G-3pc5-C-3pc5-T-3pc5-G-T-G-G-A-5' or 3'-G-3pc5-C-3pc5-T-3pc5-G-T-G-G-A-5'). Both simulated systems (HS RNase H + 3pc5-DNA:RNA and HS

RNase H + 3cp5-DNA:RNA) were surrounded by TIP3P [23] water molecules which extended to a distance of approximately 10 Å (in each direction) from the enzyme-nucleic acids atoms. This gives a periodic box size of ~82 Å, ~81 Å, ~61 Å for a simulated system consisting of 39,577 atoms. Similarly, the TtAGO enzyme [22] (Protein Data Bank entry 3HVR) was surrounded by TIP3P water molecules (periodic box size ~117 Å, ~106 Å, ~87 Å, 88,603 atoms). New *.inpcrd (initial coordinates) and *.prmtop (molecular topology, force field etc.) files for the whole simulated system, were created by means of the TLEAP module (the AMBER software package [24]).

The AMBER force field used in this study [25,26] does not contain force constants needed to describe modified parts of the 3cp5/3pc5 phosphonate analogs. The force field parameters were therefore completed with those obtained from ab initio calculations consistently with the original approach [25]. Force constants (see Table S1) were fitted to achieve the agreement of ~1 kcal/mol between QM (MP2/6-31G*) and MM relative energies of different conformers of the CH₃-O-PO₂-CH₂-O-CH₃ model system.

Almost half of the proteome of living organisms is constituted of metalloproteins. In the PDB database, magnesium ions are invariably hexacoordinated with octahedral geometry. An analysis of structures available in the PDB to the 2002 indicated that the Mg–O bond length can range between 2.05 and 2.25 Å [21]. Unfortunately, the ability of the current generation of molecular dynamics pairwise-additive force fields to properly describe metal pockets is severely lacking due to the intrinsic difficulty of handling polarization and charge transfer contributions [27]. Magnesium ions are usually represented as van der Waals (VdW) spheres with a formal point charge +2 that interact with the protein environment and the substrate through non-bonded interactions. In binuclear active sites, however, high charges can lead to repulsion between two Mg²⁺ and instabilities, resulting in an alteration of the proper coordination of Mg²⁺ ions by ligands [21,28,29].

A consistent set of empirical interaction parameters for the alkali and alkaline-earth metal cations was derived by Aqvist using free energy perturbation simulations in aqueous solution [30]. When used these parameters for Mg²⁺ (as rescaled for AMBER: R* = 0.7926 Å, epsilon = 0.8947), the coordination of both Mg(b) and Mg(c) in the β-polymerase active site changed [21]: Mg(b) lost contact with Asp192 and O2B (a non-bridging β-phosphate oxygen of the dNTP) and was instead coordinated by O3B (the oxygen bridging the β and γ phosphates of the dNTP); Mg(c) lost contact with the O3' of the primer and O1A (a non-bridging α-phosphate oxygen of the dNTP coordinating both metal ions in the crystal structure).

Further, a general tendency to underestimate slightly magnesium-ligand distances, leading to a more compact active site, is usually observed in MD simulations [27,30]. The Aqvist parameters reproduce the observed hydration free energies as well as radial distribution functions of ions reasonably. However, the value of radius for Mg²⁺ ions (i.e. R* = 0.7926 used in the AMBER force-filed) shows a less good agreement with the observed distance than for any of the other ions and appears to be slightly too small [30]. Even though using Mg²⁺ with R* = 1.300 Å and epsilon 0.06 yielded the best structural results among the traditional one-atom Mg²⁺ models, there were some significant discrepancies between the crystal structures and the modeled structures, i.e. increased Mg(b)–Mg(c), O1A–Mg(b) distances [21].

The Aqvist–Warshel (AW) model that uses cationic dummy atoms and has proven very useful in accurately representing the energetics and structures of systems with transition metals such as octahedrally coordinated bivalent manganese ions and tetrahedrally coordinated bivalent zinc ions in mononuclear and binuclear metalloenzymes [21]. The AW model allows for stable coordination geometry by placing the cationic dummy atoms at the defined positions around the central metal atom and for a smaller repulsion

between the metal ions in binuclear sites by distributing the positive charge over the cationic dummy atoms. In [21], the AW model always performed better than the best traditional one-atom Mg^{2+} model. In particular, the $Mg(b)$ – $Mg(c)$ distance and the distance between $Mg(b)$ and O1A of the incoming dNTP were much closer to the X-ray crystallographic distances. Further, in the case of the RNA ligase ribozyme [31], the Mg –O bond lengths obtained via a dummy approach [21] were more similar to those measured in QM/MM MD simulations and in the X-ray structures of other Mg^{2+} containing enzymes.

The AW model was used here for a preliminary MD simulation (lasting for 1000 ns) of HS RNase H enzyme in complex with unmodified nucleic acids. The active site of HS RNase H was stable (see Graph S14). Nevertheless, we would like to carry out comparative MD simulations of various RNase H/Argonaute enzymes with the use of multiple crystal structures for each of them. From this perspective, however, we felt the AW model a little bit complicated and uncomfortable for fast building of simulated systems. Especially when the enzyme active site is not our primary concern and we want just to ensure its stability during MD simulations.

Therefore, we decided to fine-tune an empirical one-atom model of Mg^{2+} so that the results of MD simulations were close to what was observed with the AW model in our preliminary MD simulation. We have taken into account that problems with binuclear active sites arise because of the neglect of polarization. One possibility would be to change the point charges on both ions and all of its ligands (which again is not practical if you want to work with many crystal structures). In [21], in the best one-atom Mg^{2+} model both the $Mg(b)$ – $Mg(c)$ distance and the $Mg(b)$ –O1A distance deviated significantly from the crystal structure by up to 13% and 18%, respectively. It was recognized that these distances are likely correlated, because O1A as a bridging ligand keeps the $Mg(b)$ – $Mg(c)$ distance short. The point charges of the non-bridging pro-Sp (analogous to O1A in [21]) and pro-Rp oxygen of the scissile rA5–rC6 phosphate group which undergoes contacts with both magnesium ions in the HS RNase H active site were modified here (from $-0.7761/-0.7761$ to $-0.9761/-0.5761$). Further, a general tendency to underestimate slightly magnesium–ligand distances, leading to a more compact active site, which is usually observed in MD simulations [27,30], was considered. Therefore, the radius R^* of Mg^{2+} was changed from 0.7926 to 1.1 Å (recall that $R^* = 1.3$ Å was used in the best one-atom model in [21]).

Equilibration MD trajectories lasting for 5 ns were computed with the aid of the NAMD 2.7 software package [32]. The smooth Particle-mesh Ewald (PME) method was employed for long-range electrostatic forces [33]. The non-bonded cutoff was set to 9 Å. The SHAKE algorithm (tolerance 0.0005) was applied to constrain bonds where the hydrogen atoms were involved [34]. Simulated systems were energy minimized, then the Langevin dynamics was used for a temperature control [32]. The simulated systems were heated from 0 K to 310 K. The Langevin piston method was applied to reach an efficient pressure control with target pressure set to 1 atm [32]. The integration timestep was set to 2 fs. A multiprocessor system equipped with 16 Intel Itanium CPUs (1.5 GHz) was used for equilibration MD simulations (computed at ~ 4 ns/day for RNase H and < 2 ns/day for TtAGO).

For production runs we have used ACEMD v. 2591 [35,36] running on a local workstation equipped with an NVIDIA graphics processing unit (GPU). ACEMD is a new generation molecular dynamics software which runs exclusively on GPUs at the equivalent speed of tens to hundred of standard processors. ACEMD implements all features of an MD simulation on a CUDA-compatible GPU device, including those usually required for production simulations in the NVT ensemble (i.e. bonded and nonbonded force term computation, velocity-Verlet integration, Langevin thermostatic control, smooth PME, and hydrogen bond constraints

implemented using the M-shake algorithm and RATTLE for velocity constraints within the velocity Verlet integration scheme [37,38]). ACEMD does not presently contain a barostat for production runs, so simulations in the NPT ensemble are not possible. However, it is noted that with large molecular systems, changes in volume due to the pressure control are very limited after an initial equilibration making NVT simulations viable for production runs [35,36]. It was proven on multiple biomolecular systems including membrane proteins [39–41]. Therefore, our production runs have been performed in the NVT ensemble, Langevin thermostat at 310 K, computing the electrostatic interactions with PME. Also implemented in ACEMD is the hydrogen mass repartitioning scheme (the mass of the bonded heavy atoms to hydrogen is repartitioned among hydrogen atoms, leaving the total mass of the system unchanged) used, for instance, in codes such as Gromacs, which allows an increased time step of up to 4 fs [35,42]. Therefore, the integration timestep in our production MD runs was set to 4 fs. For production MD runs (1000 ns + 2×200 ns = 1400 ns for RNase H, 500 ns for TtAGO) CUDA programmable NVIDIA GTX-580 GPU equipped with 480 cores were used. MD trajectories were computed at ~ 50 ns/day for RNase H and ~ 20 ns/day for TtAGO (i.e. ~ 53 days of GPU time in total).

Data were recorded every 100 ps. MD trajectories were analyzed with the aid of the CHIMERA 1.5.3, VMD 1.9, Curves+ and AMBER10/ptraj software packages [24,43–45]. Figures were produced by means of the ICM Molsoft 3.7 software package.

3. Results

3.1. HS RNase H – DNA recognition

Five internucleotide linkages (taken from the 3' end of DNA) are involved in stabilizing interactions with the HS RNase H surface (see Fig. 1). The first two internucleotide linkages (dG1–dC2–dT3) interact with the core of HS RNase H that is formed by either hydrophobic (Ile, Val, Phe) or bulky (Trp) side chains. The third internucleotide linkage (dT3–dG4) binds to the so called phosphate binding pocket (see Fig. S1). The fourth (dG4–dT5) and fifth (dT5–dG6) internucleotide linkages could potentially interact with Arg153–5, Arg157 residues in the extremely basic loop of HS RNase H.

In the sugar-phosphate backbone of DNA, highly charged phosphate groups alternate with rather hydrophobic deoxyribose moieties. The DNA binding surface of HS RNase H mirrors nucleic acids. Polar regions alternate with hydrophobic clusters as well. In the HS RNase H + DNA:RNA complex, phosphate groups and amino acids side chains are bound by intermolecular hydrogen bonds (Trp, Asn, Thr) or salt bridges (Lys, Arg). In addition, clusters of hydrophobic (Ile, Val, Phe) or partly hydrophobic (Thr) amino acids side chains establish favorable VdW contacts with rather hydrophobic deoxyribose moieties of DNA.

3.2. 1st DNA-binding site – 3pc5 MD (Fig. 2, Graph 1A: Trp225 red line, Ser233 blue line)

The first DNA-binding site is a channel formed by Trp221, Trp225, and Ser233 in the basic protrusion of HS RNase H [20] (homologous residues in EC RNase H are Trp81 and Trp85 [19]; BH RNase H lacks the basic protrusion completely [46]).

In our MD simulations, the 3pc5 phosphonate group of the first dG1–dC2 internucleotide linkage was hydrogen bonded toward Trp225 and Ser233. The interaction with Trp225 was remarkably stable despite of conformational promiscuity of the dG1–dC2 internucleotide linkage and instabilities found in the rC1.dG1 Watson–Crick base pair (see below). In contrast, there were several

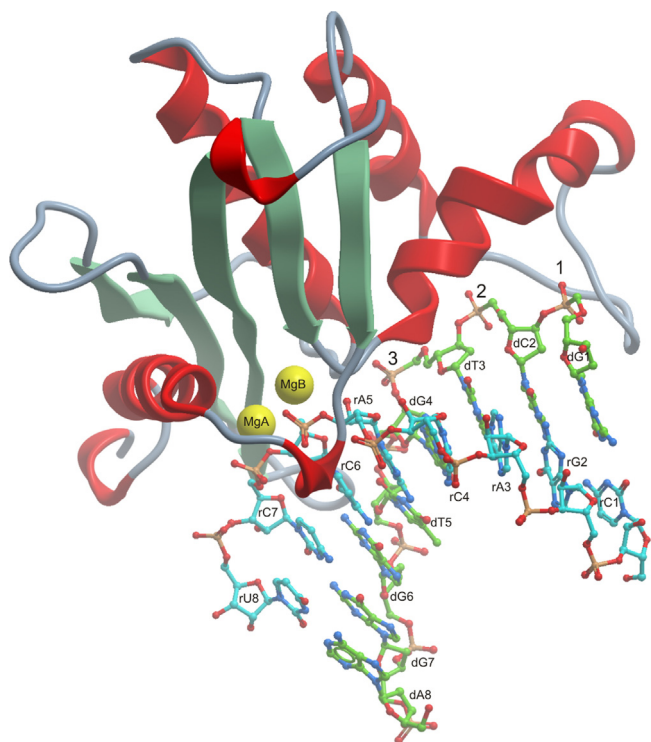


Fig. 1. Five internucleotide linkages from the 3' end of DNA are involved in stabilizing interactions with the HS RNase H surface.

short living interruptions in mutual contacts of dG1–dC2 with Ser233.

The dG1 and dC2 deoxyribose moieties as well as the $-\text{CH}_2-$ group wedged in the dG1–dC2 phosphonate internucleotide linkage were in close contacts with the Trp221 side chain, which probably helps to distinguish DNA:RNA and RNA:RNA hybrids (the 2'-OH group of RNA modeled into the DNA binding channel of HS RNase H clashes with the indole ring of Trp221 [20]).

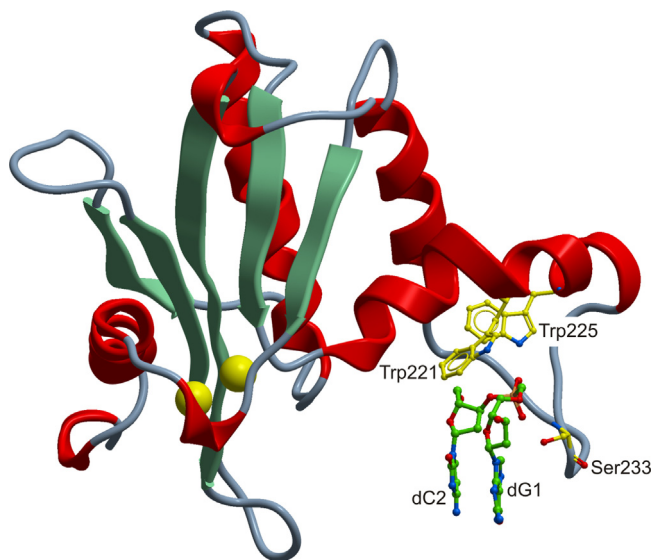


Fig. 2. The 3pc5 dG1–dG2 internucleotide linkage binding into the 1st DNA binding site.

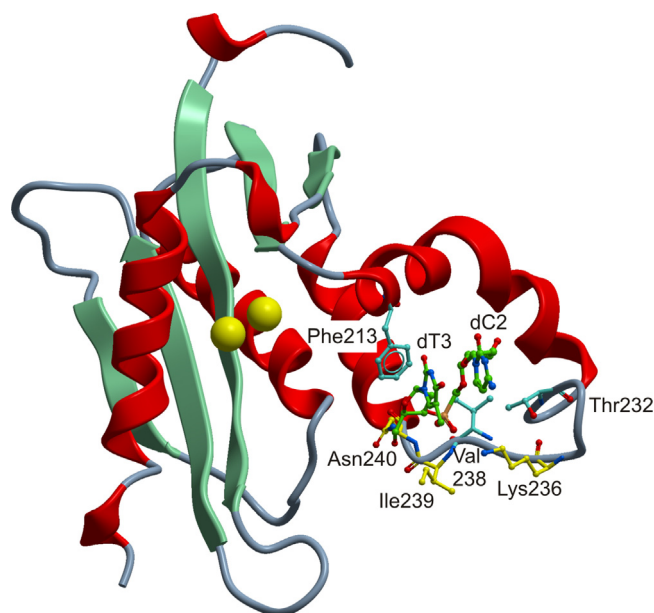


Fig. 3. The 3pc5 dC2–dT3 internucleotide linkage binding into the 2nd DNA binding site.

3.3. 2nd DNA-binding site – 3pc5 MD (Fig. 3, Graph 1B): Asn240 m.c. blue line, Ile239 m.c. cyan line, Asn240 s.c. red, Lys236 s.c. green line)

The second DNA-binding site, where the 3pc5 phosphonate group of dC2–dT3 was anchored, consisted of main chain amide groups of Asn240 and Ile239, which were occasionally accompanied by side chains of Asn240 and Lys236.

Further, there were observed VdW contacts between the bulky side chain of Phe213 and the dT3 deoxyribose moiety and between the Thr232, Val238 side chains and the $-\text{CH}_2-$ group of the dC2–dT3 phosphonate internucleotide linkage.

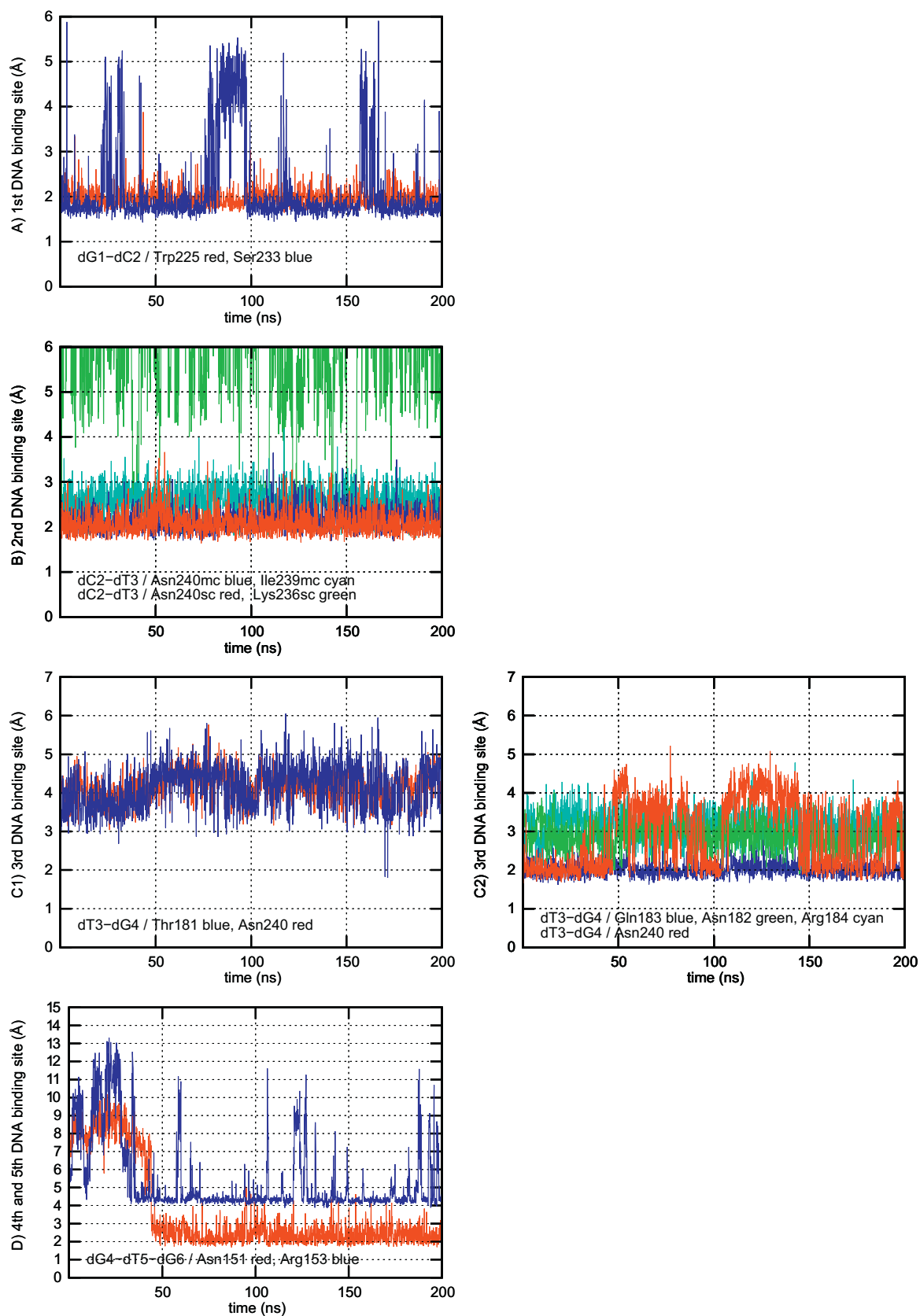
In [20], the second DNA-binding site was not explicitly mentioned as the natural phosphate internucleotide linkage was buried substantially less deeply into the HS RNase H surface. In the case of EC RNase H, the phosphate/phosphonate groups were potentially stabilized by polar interactions with side chains of Tyr73, Trp104 and Lys99 [19] (replacing Phe213, Phe244, Ile239 of HS RNase H).

3.4. 3rd DNA binding site – 3pc5 MD (Fig. 4, Graph 1C1: Thr181 blue line, Asn240 red line; Graph 1C2: Gln183 blue line, Asn182 green line, Arg184 cyan line, Asn240 red line)

The 3pc5 phosphonate group of the dT3–dG4 internucleotide linkage established contacts with the main chain amide group of Gln183. Moreover, the Asn240 side chain was loosely bound to the atom O5' of the 3pc5 phosphonate group.

In [20], the dT3–dG4 phosphate group was bound to Thr181, Arg179 and Asn240. A similar phosphate-binding pocket was found in BH RNase H [46]. The pockets are spatially conserved but share only one conserved residue, Thr181 in HS RNase H and Thr104 in BH RNase H. Interestingly, analogous Thr43 in EC RNase H was identified as interacting with DNA by one of us before [19].

Sensitivity of the phosphate binding pocket to modified internucleotide linkages was evidenced by experimental results [47]. This can be explained on the basis of structural information. In HS RNase H, Both Thr181 (in the phosphate binding pocket) and Glu186 (in the active site) are located on the same α -helix (Fig. S1). Therefore, any suboptimal binding of modified internucleotide linkages into



Graph 1. 3pc5 DNA recognition by DNA binding sites of HS RNase H: (A) Trp225 – red line, Ser233 – blue line; (B) Asn240 m.c. – blue line, Ile239 m.c. – cyan line, Asn240 s.c. – red line, Lys236 s.c. – green line; (C1) Thr181 – blue line, Asn240 – red line; (C2) Gln183 – blue line, Asn182 – green line, Arg184 – cyan line, Asn240 – red line; (D) Asn151 – red line, Arg153 – blue line (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.).

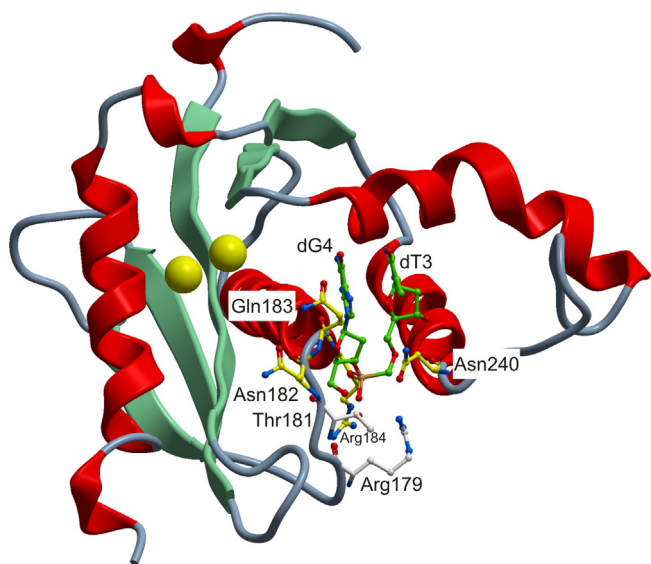


Fig. 4. The 3pc5 dT3–dG4 internucleotide linkage binding into the 3rd DNA binding site.

the phosphate binding pocket can lead to an allosteric modulation of the HS RNase H active site.

Crystal structures show that to place just a natural dT3–dG4 internucleotide linkage into the phosphate binding pocket requires large distortions of the backbone torsion angles (γ and α) by $\sim 150^\circ$ from the ideal values [20]. Indeed, in our preparatory MD simulations with natural oligonucleotides, the Thr181 – dT3–dG4 contacts were lost when the lowest energy conformer **-g-g** (in terms of the C3'–O3'–P–O5', O3'–P–O5'–C5' torsion angles) was realized. It imposes increased demands on potential antisense oligonucleotides, as modified internucleotide linkages must be able to mimic the phosphate group adopting not only its natural lowest-energy conformer **-g-g**. This hurdle can be by-passed by oligonucleotide gapmers or oligonucleotides with alternating natural and modified internucleotide linkages.

It seems to be the case of the 3pc5 phosphonates, which elicit the RNase H activity only if they alternate in oligonucleotide with natural internucleotide linkages [15]. In our MD simulations, the 3pc5 phosphonate group of the dT3–dG4 internucleotide linkage lacked contacts with Thr181 in the phosphate binding pocket. Therefore, it seems likely that oligonucleotides with alternating phosphate/3pc5-phosphonate internucleotide linkages will be rather active when the 3pc5 phosphonate will bind to the second DNA binding site and the first and third (phosphate binding pocket) DNA binding site will be occupied by natural internucleotide linkages.

3.5. 4–5th DNA binding site – 3pc5 MD (Fig. S2, Graph 1D: Asn151 red line, Arg153 blue line)

The phosphate group of the natural dG4–dT5 internucleotide linkage was in contact with Asn151. The natural dT5–dG6 internucleotide linkage interacted with Arg153. These contacts were established after elapsing of about 30 ns. Arg153 altogether with Arg154, Arg155 and Arg157 forms the basic loop in HS RNase H structures, where significant differences among HS RNase H crystal structures were observed [20]. Our preliminary MD simulations with natural substrates showed that neither the trajectories lasting for 1 μ s are sufficient for a complete conformational relaxation of this basic loop. Therefore, it cannot be excluded that the dG4–dT5 and dT5–dG6 internucleotide linkage come to contact with Arg154, Arg155 and Arg157 on longer time scales. Anyway, the Arg residues

would be probably involved in the initial enzyme–substrate recognition (as proposed in [20]).

3.6. 1st – 5th DNA binding site – 3cp5 MD (Graph 2A–D)

In contrast to what was observed for 3pc5 phosphonates, the 3cp5 phosphonate internucleotide linkages lost contacts with HS RNase H DNA binding sites within several tens of ns. It's in agreement with known experimental data as the 3cp5 phosphonate oligonucleotides are not able to elicit the RNase H activity [15].

3.7. RNA–HS RNase H interactions – (Fig. S3, Graph S1: Glu186 blue and salmon line, Met212 red line, Arg278 green line)

The HS RNase H active site is located in the RNA-binding groove, which interacts with the 2'–OH groups of four consecutive ribonucleotides (rC4, rA5, rC6, rC7 – two on each side of the scissile phosphate). The overall mechanism for RNA strand recognition is conserved [20] (as RNA strands complexed with HS, EC and BH RNase H superimpose well).

In our MD simulations, the RNA–HS RNase H contacts were mediated mostly by the side chain of Glu186 interacting with the 2'–OH group of rA5. Recall that Glu186 is bound to Mg^{2+} in the HS RNase H active site. Further, as we speculate above, Glu186 could transmit allosteric signals from the phosphate binding pocket to active site of HS RNase H.

The main chain atoms of Met212 were involved in contacts with rC4 – especially in the second half of the 3pc5 MD simulation.

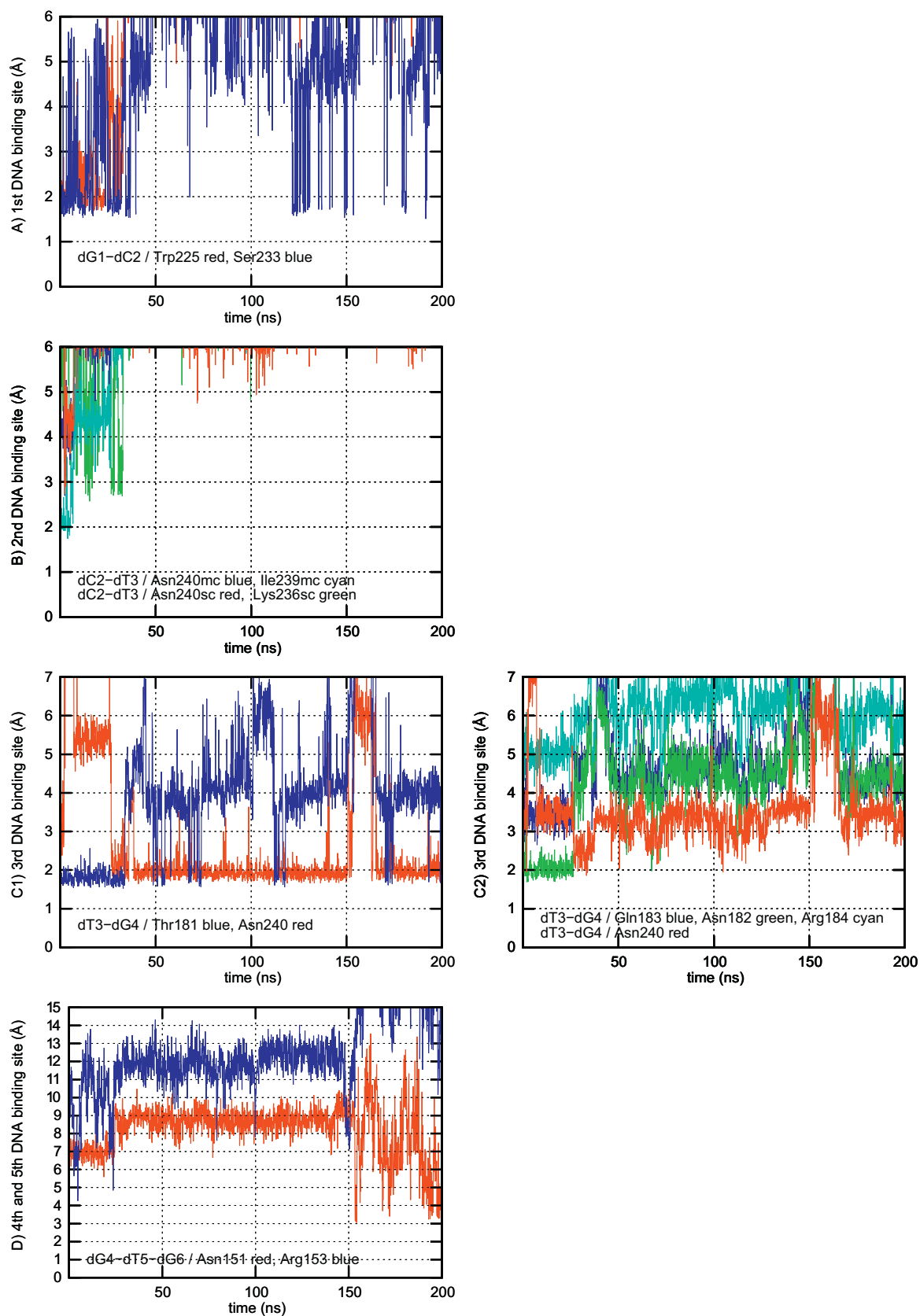
Contacts of the Cys148, Ser150 and Asn151 backbone atoms with the rC6 and rC7 2'–OH groups were rather poor, however, it could be partly rooted in the insufficient relaxation of the 153–157 basic loop, where significant differences among individual HS RNase H structures were observed [20].

Interactions between RNA and Arg residues were very sporadic. The only exception was the rC6–rC7 internucleotide linkage and Arg278. rC6–rC7 is located just behind the scissile rA5–rC6. Therefore, the Arg278–rC6–rC7 contacts could help to stabilize the reaction product. Moreover, the rC6–rC7 internucleotide linkage could trap the proton released from the catalytic water molecule. It could facilitate formation of the hydroxide ion that attacks the scissile rA5–rC6 internucleotide linkage. The Arg278 side chain was situated differently in the 3pc5/3cp5 MD simulations. Therefore, remarkable differences were observed in terms of contacts of the catalytic water molecule with both the rC6–rC7 as well as rA5–rC6 internucleotide linkages (see below).

3.8. Helical parameters (Graph S2–S5)

Helical parameters of both 3pc5/3cp5–DNA:RNA duplexes were calculated by means of the Curves+ program [45]. The most affected helical parameters were Buckle, Shift, Slide and Rise. As a result of the distorted DNA strand, the minor groove width of the DNA:RNA duplex varies from 7.36 Å (3pc5)/7.76 Å (3cp5) at the phosphate-binding pocket to 10.92/11.34 Å (3pc5)/11.14/11.48 Å (3cp5) at terminal parts of the DNA/RNA hybrid.

It means, that both 3pc5/3cp5–DNA:RNA duplexes were remarkably distorted despite of the 3cp5 DNA strand dissociation from the HS RNase H surface. Neither the 200 ns MD simulation was sufficient to achieve a full relaxation of the 3cp5–DNA:RNA duplex geometry. Apparently, deformations are dictated by the overall shape of HS RNase H rather than by detail interactions of phosphate/phosphonate groups with specific amino acids in DNA binding sites. In accordance with that no such variations were



Graph 2. 3cp5 DNA recognition by DNA binding sites of HS RNase H: (A) Trp225 – red line, Ser233 – blue line; (B) Asn240 – blue line, Ile239 – cyan line, Asn240 – red line, Lys236 – green line; (C1) Thr181 – blue line, Asn240 – red line; (C2) Gln183 – blue line, Asn182 – green line, Arg184 – cyan line, Asn240 – red line; (D) Asn151 – red line, Arg153 – blue line (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.).

observed in complexes of BH RNase H, which lacks the basic protrusion [46] (Fig. S1).

3.9. Watson–Crick hydrogen bonding (Graph S6)

In both MD simulations, the Watson–Crick hydrogen bonds were stable with the exception of terminal base pairs. Nevertheless, such occasional fraying of nucleic acids ends is commonly observed in MD simulations even in the case of double helical structures, which are unaffected by interactions with enzymes.

3.10. Sugar pucker (Graph S7–S10)

The RNA strand was in the regular A-form with sugars in **C3'-endo**. Nevertheless, **C2'-exo** was remarkably populated in the phosphate binding pocket and **C4'-exo** in terminal parts of both 3pc5/3cp5-DNA:RNA structures.

The 3pc5 DNA strand was mostly in the B-form (with sugars between **C2'-endo** and **O4'-endo**) except for dC2 and dT3 in the DNA-binding channel, which were in **C4'-exo** (i.e. between the A and B form). In the 3cp5 DNA only dT3 was in **C4'-exo**. It indicates a partial relaxation of 3cp5 DNA after its dissociation from the HS RNase H surface.

3.11. Phosphonate internucleotide linkages (Graph S11)

All 3pc5/3cp5 phosphonate internucleotide linkages were conformationally promiscuous (in terms of either the **C3'-O3'-P-C**, **O3'-P-C-O5'**, **P-C-O5'-C5'** or **C3'-O3'-C-P**, **O3'-C-P-O5'**, **C-P-O5'-C5'** torsion angles). This is evident especially in the case of the 3cp5 internucleotide linkages, which lost contacts with the DNA binding sites on the HS RNase H surface. However, even the 3pc5 internucleotide linkages that were steadily bound to all DNA binding sites fluctuated between several conformational basins.

3.12. Phosphate internucleotide linkages (Graph S7–S10)

Phosphate internucleotide linkages in DNA (i.e. dG4-dT5-dG6-dG7-dA8) adopted the lowest energy conformer **-g-g** (i.e. **-gauche** in terms of the **C3'-O3'-P-O5'**, **O3'-P-O5'-C5'** torsion angles). Phosphate internucleotide linkages in RNA preferred **-g-g** as well.

Higher populations of **t-g** in rA3-rC4 indicate a suboptimal binding of DNA:RNA to the HS RNase H surface at the beginning of the 3cp5 MD simulation.

In X-ray structures [20], the rC4-rA5 internucleotide linkage was distorted from **-g-g**, which was preferred in both MD simulations. Similarly, the rC6-rC7 internucleotide linkage adopted an atypical conformation (or maybe an averaged mixture of conformers was observed) in crystals [20]. In MD simulations, the rC6-rC7 phosphate preferred the lowest energy conformer **-g-g** with occasional oscillations to **t-g**. Recall that rC6-rC7 was in contact with Arg274.

3.13. Active site – MgA and MgB (Fig. 5, 3pc5 Graph S12A, 3cp5 Graph S13A, 7p AW model Graph S14A, TtAGO Graph S15A)

The HS RNase H active site is superimposable with that of either BH RNase H or EC RNase H, including the scissile phosphate. Two Mg^{2+} ions occupy the canonical A and B positions. It's assumed that the two-metal ion mechanism for catalysis is universally shared [48]. Both 3pc5/3cp5 MD simulations produced trajectories that were stable consider the distance separating both magnesium ions. That is not entirely obvious in binuclear active sites [21] (see Section 4).

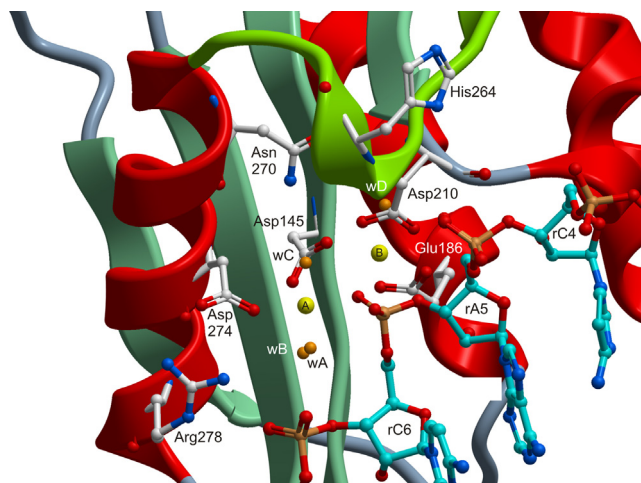


Fig. 5. HS RNase H active site.

3.14. MgA coordination shell – 3pc5 MD (Fig. 5, Graph S12C1: Asp274 red line, Asp145 green line, Sp blue line; Graph S12C2: WATA-C; Graph S12B: WATA)

Magnesium ion A was coordinated in the octahedral configuration by Asp145, Asp274, the pro-Sp non-bridging oxygen atom of the scissile rA5-rC6 phosphate and three water molecules. The water molecule that attacks the scissile rA5-rC6 phosphate will be referred as WATA in the following text.

Throughout the whole 3pc5 MD run, the oxygen atom of WATA protruded into the immediate vicinity (below the 3 Å threshold) of the phosphorus atom of the scissile rA5-rC6 internucleotide linkage. Hydrogen atoms of WATA made contacts with the adjacent rC6-rC7 internucleotide linkage. Both events may be favorable for the reaction.

3.15. MgA coordination shell – 3cp5 MD (Fig. 5, Graph S13C1: Asp274 red line, Asp145 green line, Sp blue line; Graph S13C2: WATA-C; Graph S13B: WATA)

Water molecules interchanged repeatedly in the WATA position within the 3cp5 MD run. The oxygen atom of WATA protruded less frequently below the 3 Å threshold consider the distance from the phosphorus atom of the scissile rA5-rC6 internucleotide linkage. Hydrogen atoms of WATA made contacts with the adjacent rC6-rC7 internucleotide linkage only occasionally.

This was mainly due to different (comparing to 3pc5 MD) conformational preferences of the Arg278 side chain interacting with the rC6-rC7 internucleotide linkage.

3.16. MgB coordination shell-

3pc5 MD (Graph S12D1: Asp145 red line, Glu186 green line, Glu186 blue line, Asp210 cyan line, Asp210 magenta line, Sp yellow line, O3' salmon line, Graph S12D2: WAT).

3cp5 MD (Graph S13D1: Asp145 red line, Glu186 green line, Glu186 blue line, Asp210 cyan line, Asp210 magenta line, Sp yellow line, O3' salmon line, Graph S13D2: WAT).

In X-ray structures [20], MgB is coordinated by five ligands: Asp145, Glu186, Asp210 (or Asn210), the O3' and pro-Sp non-bridging oxygen atoms of the scissile rA5-rC6 phosphate group. In our MD simulations, subtle differences were observed consider the 3rd–6th ligands:

1st ligand: The non-bridge pro-Sp oxygen atom of the rA5-rC6 scissile internucleotide linkage was tied quite symmetrically between MgA and MgB.

2nd ligand: Similarly, the side chain of Asp145 was bound simultaneously to MgA and MgB.

3rd ligand: The side chain of Asp210 had a tendency to provide both oxygen atoms toward MgB and thus create a so-called bi-dentate contact.

4th ligand: Similarly, Glu186 tend to form bi-dentate contacts with MgB.

5th ligand: The O3' atom of the rA5-rC6 scissile internucleotide linkage was weakly bound to MgB in the first half of the 3pc5 MD run. Then even more loosened contacts prevailed. In the 3cp5 MD simulation, this process was faster and even more pronounced.

6th ligand: Several water molecules reside close to MgB in the RNase H interior (see WD in Fig. 5). Roughly in the half of both MD simulations water molecules expelled the O3' atom of the rA5-rC6 scissile internucleotide linkage from MgB. It was more pronounced in the 3cp5 MD simulation, in fact, there were many molecules interchanging in this position in the 3pc5 MD simulation.

Altogether, Asp210 and Glu186 tendency to form bi-dentate contacts with MgB led to rather over-coordinated (7 ligands) MgB. This tendency was even more pronounced in reference MD simulations with the seven-point representation of magnesium ions (compare Graph S12 and Graph S14).

In the TtAGO MD simulation (Fig. S1, Graph S15), Asp546 (Asp210 in HS RNase H) was mono-dentate and Glu512 (Glu186 in HS RNase H) bi-dentate.

4. Discussion and conclusions

Our 200 ns MD simulations show that fully modified oligonucleotides bearing the 3'-O-P-CH₂-O-5' (but not 3'-O-CH₂-P-O-5') phosphonate linkages can be attached to the HS RNase H surface (compare Graph 1 with Graph 2). It enables to explain that oligonucleotides consisting of the alternating 3'-O-P-CH₂-O-5'-phosphonate and phosphodiester linkages are capable to elicit the RNase H activity (while the 3'-O-CH₂-P-O-5'-phosphonates are completely inactive) [15].

In binuclear active sites, which are modeled using non-polarizable force fields, the artificial repulsion arising between both magnesium ions can lead to severe instabilities and alterations of the proper coordination of magnesium ions by ligands [21]. Therefore, the polarized rA5-rC6 scissile phosphate (with the non-bridging oxygen atom wedged between magnesium ions) was used in our MD simulations. Further, a general tendency to underestimate slightly magnesium-ligand distances, leading to a more compact active site, is usually observed in MD simulations [27,30]. Therefore, the radius R^* of Mg²⁺ was changed from 0.7926 to 1.1 Å (recall that even greater value of 1.3 Å was used in the best one-atom model of Mg²⁺ in [21]). The resulting MgA-MgB distances found in our MD simulations were stable on the time scale of hundreds of nanoseconds (Graph S12–13). A reference MD simulation (lasting for 1000 ns) produced using the seven-point (with dummy atoms) AW representation of magnesium ions [21] led to almost the same results (compare Graph S12 and Graph S14). Therefore, the usage of a partial polarization of the scissile internucleotide linkage (which is easier for implementation than the AW model [21]) is justified especially in studies, where the enzyme active site is not a primary concern. Further, fluctuations around average values were a little bit more pronounced in our model (Graph S12) comparing to the AW model [21] (Graph S14). In fact, larger radius of Mg²⁺ led to a modest diminishing of stabilizing coulombic interactions acting between ions and ligands. Therefore, active sites in our MD simulations were inevitably a little bit under-stabilized. In subsequent studies, we want to destabilize ion-ligand interactions even much more to enable a proper relaxation (at

an acceptable time scale) of loops around the RNase H/Argonaute active site, which could be the subject of crystal packing forces or artifacts due to point mutations applied in X-ray structures.

Further, to demonstrate its transferability, we applied our approach chosen for the stabilization of binuclear active sites to TtAGO [22], where the metal-binding residue Glu512 (analogous to Glu186 in HS RNase H) is in a very labile position. In fact, only quite recently was recognized that Argonaute enzymes use the catalytic tetrad rather than catalytic triad of residues [14] (Fig. S1). The Argonaute active site was completely stable in the MD run lasting for 500 ns (Graph S15). As proposed in [14], Glu512 was bound in the TtAGO active site in exactly the same way as Glu186 was bound in the HS RNase H active site (i.e. Glu186/512 interacted with both the 2'-OH of the nucleotide adjacent to the scissile phosphate and one of the two active-site divalent metal ions, Fig. 5).

Glu186 in the HS RNase H active site is close to conserved Thr181 in the so called phosphate-binding pocket (Fig. S1) sensitive to chemical modifications of antisense oligonucleotides [47]. We speculate that an allosteric signal could be transmitted from Thr181 to Glu186 in the case of a suboptimal binding of modified internucleotide linkages (including here presented 3pc5 phosphonates – Graph 1C1) into the phosphate-binding pocket.

In the case of TtAGO, main chain atoms of Glu512 interact directly with the phosphate group of siRNA at position 13–14 (Fig. S1). Therefore, chemical modifications of siRNA backbones can act on the TtAGO active site quite straightforwardly. Indeed, significantly lower AGO2 cleavage activities were observed for the antisense RNAs containing a 2'-MOE substitution at positions 2, 13, and 14 [49,50]. Comparable binding affinities were observed for the antisense RNA containing the 2'-MOE substitutions at positions 12–14 compared with the unmodified antisense RNA. Taken together, these data suggest that the loss in AGO2 cleavage activity was due to the 2'-MOE residues interfering with catalysis [49,50].

In order to investigate an allosteric modulation of the RNase H/AGO active sites by Glu fingers precisely, it will need to address further, what exactly is the coordination of MgB in the moment immediately preceding the reaction event. It was speculated in [48] (based on crystal structures of BH RNase H) that under-coordinated MgB provides a driving force for the reaction. However, crystal structures are subjected to crystal packing forces, point mutations are carried out in their active sites (in order to prevent the reaction), different (Mg²⁺/Ca²⁺) ions are used for crystallization etc. Then, such as different form of bi-dentate/mono-dentate binding to MgB as observed here for analogous Asp210/Asp546 residues in HS RNase H/TtAGO can be found. In fact, Glu186 and Asp210 tendency to form bi-dentate contacts led to rather over-coordinated MgB in HS RNase H. It was even more pronounced with the AW model [21] applied in the reference MD simulation (compare Graph S12D1 with S14D1).

The balance between point charges on oxygen atoms of bio- and water molecules appears to be another moot point. Charges in TIP3P water molecules [23] were determined based on a completely different philosophy than RESP charges of biomolecules in the AMBER force field [25]. Therefore, probably, the O3' atom of the scissile rA5-rC6 internucleotide linkage was expelled from MgB by water molecules residing in the HS RNase H interior. Interestingly, this artifact was again more pronounced in the reference MD simulation with the seven-point AW model for Mg²⁺ [21] (compare Graph S12D1 with S14D1).

Acknowledgments

This work was supported by the Grant Agency of the Czech Republic (202/09/0193). The access to computing and storage facilities owned by parties and projects contributing to the National Grid

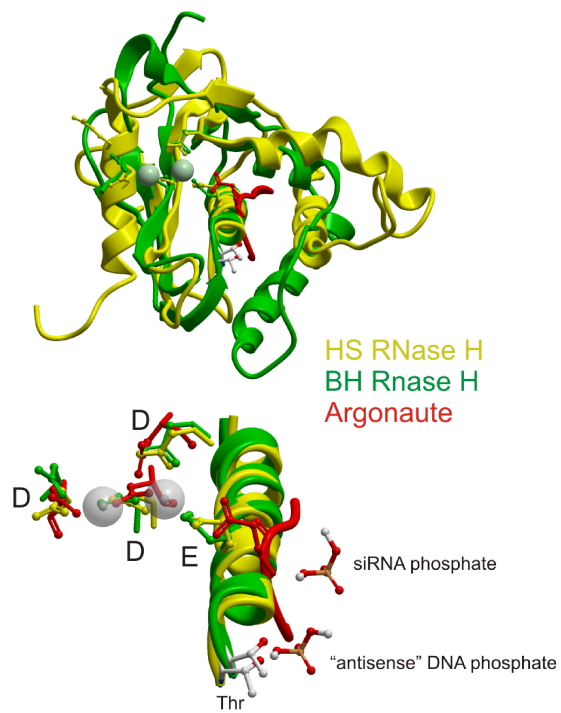
Infrastructure MetaCentrum, provided under the program “Projects of Large Infrastructure for Research, Development, and Innovations” (LM2010005) is highly acknowledged.

Appendix A. Supplementary data

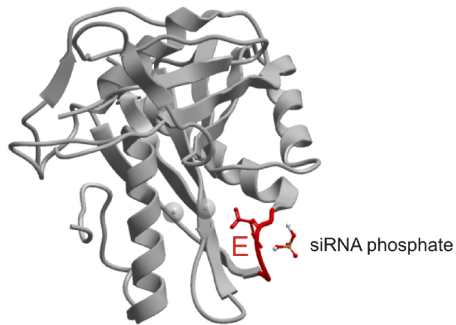
Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jmgs.2013.05.004>.

References

- [1] M.L. Jain, P.Y. Bruice, I.E. Szabo, T.C. Bruice, Incorporation of positively charged linkages into DNA and RNA backbones: a novel strategy for antigense and antisense agents, *Chemical Reviews* 112 (2012) 1284–1309.
- [2] J. Kurreck, Antisense technologies. Improvement through novel chemical modifications, *European Journal of Biochemistry* 270 (2003) 1628–1644.
- [3] Y.S. Cho-Chung, Antisense DNAs as targeted genetic medicine to treat cancer, *Archives of Pharmacological Research* 26 (2003) 183–191.
- [4] J.B. Opalinska, A.M. Gewirtz, Nucleic-acid therapeutics: basic principles and recent applications, *Nature Reviews Drug Discovery* 1 (2002) 503–514.
- [5] A. Dove, Antisense and sensibility, *Nature Biotechnology* 20 (2002) 121–124.
- [6] D.B. Braasch, D.R. Corey, Novel antisense and peptide nucleic acid strategies for controlling gene expression, *Biochemistry* 41 (2002) 4503–4510.
- [7] A.M. Gewirtz, D.L. Sokol, M.Z. Ratajczak, Nucleic acid therapeutics: state of the art and future prospects, *Blood* 92 (1998) 712–736.
- [8] E. Uhlmann, A. Peyman, Antisense oligonucleotides a new therapeutic principle, *Chemical Reviews* 90 (1990) 544–584.
- [9] S.T. Crooke, Vitravene – another piece in the mosaic, *Antisense and Nucleic Acid Drug Development* 8 (1998) vii–viii.
- [10] S.T. Crooke, An overview of progress in antisense therapeutics, *Antisense and Nucleic Acid Drug Development* 8 (1998) 115–122.
- [11] A. Fire, S.Q. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, C.C. Mello, Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature* 391 (1998) 806–811.
- [12] A.Z. Fire, Gene silencing by double-stranded RNA (Nobel Lecture), *Angewandte Chemie International Edition* 46 (2007) 6966–6984.
- [13] Craig C. Mello, Return to the RNAi world: rethinking gene expression and evolution (Nobel Lecture), *Angewandte Chemie International Edition* 46 (2007) 6985–6994.
- [14] K. Nakanishi, D.E. Weinberg, D.P. Bartel, D.J. Patel, Structure of yeast Argonaute with guide RNA, *Nature* 486 (2012) 368–374.
- [15] D. Rejman, J. Snášel, R. Liboska, Z. Točík, O. Pačes, S. Králíková, M. Rinnová, P. Kois, I. Rosenberg, Oligonucleotides with isopolar phosphonate internucleotide linkage: a new perspective for antisense compounds? *Nucleosides, Nucleotides and Nucleic Acids* 20 (2001) 819–823.
- [16] I. Barvík Jr., J. Štěpánek, J. Bok, Explicit solvent molecular dynamics simulation of duplex formed by the modified oligonucleotide with alternating phosphate/phosphonate internucleoside linkages and its natural counterpart, *Journal of Biomolecular Structure and Dynamics* 19 (2002) 863–875.
- [17] J. Hanuš, I. Barvík Jr., J. Štěpánek, P.-Y. Turpin, J. Bok, I. Rosenberg, M. Petrová, –CH₂– lengthening of the internucleotide linkage in the ApA dimer can improve its conformational compatibility with its natural polynucleotide counterpart, *Nucleic Acids Research* 29 (2001) 5182–5194.
- [18] O. Páv, I. Košiová, I. Barvík, R. Pohl, M. Budyšínský, I. Rosenberg, Synthesis of oligoribonucleotides with phosphonate-modified linkages, *Organic and Biomolecular Chemistry* 9 (2011) 6120–6126.
- [19] I. Barvík, *E. coli* RNase H1 and the phosphonate-DNA/RNA hybrid: molecular dynamics simulations, *Nucleosides, Nucleotides and Nucleic Acids* 24 (2005) 435–441.
- [20] M. Nowotny, S.A. Gaidamakov, R. Ghirlando, S.M. Cerritelli, R.J. Crouch, W. Yang, Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription, *Molecular Cell* 28 (2007) 264–276.
- [21] P. Oelschlaeger, M. Klahn, W.A. Beard, S.H. Wilson, A. Warshel, Magnesium-cationic dummy atom molecules enhance representation of DNA polymerase beta in molecular dynamics simulations: improved accuracy in studies of structural features and mutational effects, *Journal of Molecular Biology* 366 (2007) 687–701.
- [22] Y. Wang, S. Juránek, H. Li, G. Sheng, G.S. Wardle, T. Tuschl, D.J. Patel, Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes, *Nature* 461 (2009) 754–762.
- [23] W.L. Jorgensen, J. Chandrasekhar, J. Madura, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *Journal of Chemical Physics* 79 (1983) 926–935.
- [24] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.R. Ross, T.E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, AMBER, a package of computer programs for applying molecular mechanics normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, *Computer Physics Communications* 91 (1995) 1–41.
- [25] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules, *Journal of the American Chemical Society* 117 (1995) 5179–5197.
- [26] A. Perez, I. Marchan, D. Svozil, J. Sponek, T.E. Cheatham III, Ch A. Loughton, M. Orozco, Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers, *Biophysical Journal* 92 (2007) 3817–3829.
- [27] M. Dal Peraro, K. Spiegel, G. Lamoureux, M. De Vivo, W.F. De Grado, M.L. Klein, Modeling the charge distribution at metal sites in proteins for molecular dynamics simulations, *Journal of Structural Biology* 157 (2007) 444–453.
- [28] Y. Xiang, P. Oelschlaeger, J. Florian, M.F. Goodman, A. Warshel, Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: evaluating amino acid group contribution and allosteric coupling for ionized residues in Human Pol Beta, *Biochemistry* 45 (2006) 7036–7048.
- [29] J. Florian, M.F. Goodman, A. Warshel, Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase, *Journal of the American Chemical Society* 125 (2003) 8163–8177.
- [30] J. Aqvist, Ion–water interaction potentials derived from free energy perturbation simulations, *Journal of Physical Chemistry* 94 (1990) 8021–8024.
- [31] J. Sgrignani, A. Magistrato, The structural role of Mg²⁺ ions in a class I RNA polymerase ribozyme: a molecular simulation study, *The Journal of Physical Chemistry B* 116 (2012) 2259–2268.
- [32] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, *Journal of Computational Chemistry* 26 (2005) 1781–1802.
- [33] T.E. Cheatham, J.L. Miller, T. Fox, T.A. Darden, P.A. Kollman, Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA RNA, and proteins, *Journal of the American Chemical Society* 117 (1994) 4193–4194.
- [34] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *Journal of Computational Physics* 23 (1977) 327–341.
- [35] M.J. Harvey, G. Giupponi, G. De Fabritiis, ACEMD: accelerating biomolecular dynamics in the microsecond time scale, *Journal of Chemical Theory and Computation* 5 (2009) 1632–1639.
- [36] ACEMD. <http://www.acellera.com/>
- [37] H.C. Andersen, RATTLE A: “Velocity” version of the SHAKE algorithm for molecular dynamics calculations, *Journal of Computational Physics* 52 (1983) 24–34.
- [38] S.G. Lambrakos, J.P. Boris, E.S. Oran, I. Chandrasekhar, M. Nagumo, A modified SHAKE algorithm for maintaining rigid bonds in molecular dynamics simulations of large molecules, *Journal of Computational Physics* 85 (1989) 473–486.
- [39] T. Giorgino, Gianni de Fabritiis, A high-throughput steered molecular dynamics study on the free energy profile of ion permeation through gramicidin A, *Journal of Chemical Theory and Computation* 7 (2011) 1943–1950.
- [40] M. Selent, F. Sanz, M. Pastor, G. De Fabritiis, Induced effects of sodium ions on dopaminergic G-protein coupled receptors, *PLOS Computational Biology* 6 (2010) e1000884.
- [41] I. Buch, T. Giorgino, G. De Fabritiis, Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations, *Proceedings of the National Academy of Science* 108 (2011) 10184–10189.
- [42] K.A. Feenstra, B. Hess, H.J.C. Berendsen, Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems, *Journal of Computational Chemistry* 20 (1999) 786–798.
- [43] W. Humphrey, A. Dalke, K. Schulten, VMD – visual molecular dynamics, *Journal of Molecular Graphics* 14 (1996) 33–38.
- [44] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera – a visualization system for exploratory research and analysis, *Journal of Computational Chemistry* 25 (2004) 1605–1612.
- [45] R. Lavery, M. Moakher, J.H. Maddocks, D. Petkeviciute, K. Zakrzewska, Conformational analysis of nucleic acids revisited: Curves+, *Nucleic Acids Research* 37 (2009) 5917–5929.
- [46] M. Nowotny, S.A. Gaidamakov, R.J. Crouch, W. Yang, Crystal structures of RNase H bound to and RNA/DNA hybrid: substrate specificity and metal-dependent catalysis, *Cell* 121 (2005) 1005–1016.
- [47] H. Wu, W.F. Lima, S.T. Crooke, Investigating the structure of human RNase H1 by site-directed mutagenesis, *Journal of Biological Chemistry* 276 (2001) 23547–23553.
- [48] M. Nowotny, W. Yang, Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release, *The EMBO Journal* 25 (2006) 1924–1933.
- [49] W.F. Lima, H. Wu, J.G. Nichols, H. Sun, H.M. Murray, S.T. Crooke, Binding and cleavage specificities of human argonaute 2, *The Journal of Biological Chemistry* 284 (2009) 26017–26028.
- [50] D. M: Kenski, G. Butora, A.T. Willingham, A.J. Cooper, W. Fu, N. Qi, F. Soriano, I.W. Davies, W.M. Flanagan, siRNA-optimized modifications for enhanced in vivo Activity, *Molecular Therapy – Nucleic Acids* 1 (2012) e5.



Argonaute "Rnase H" domain



Argonaute

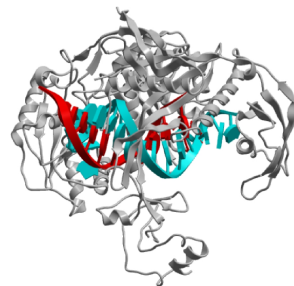


Figure S1: Crystal structures of HS RNase H [17], BH RNase H [40] and Argonaute [19] enzymes.

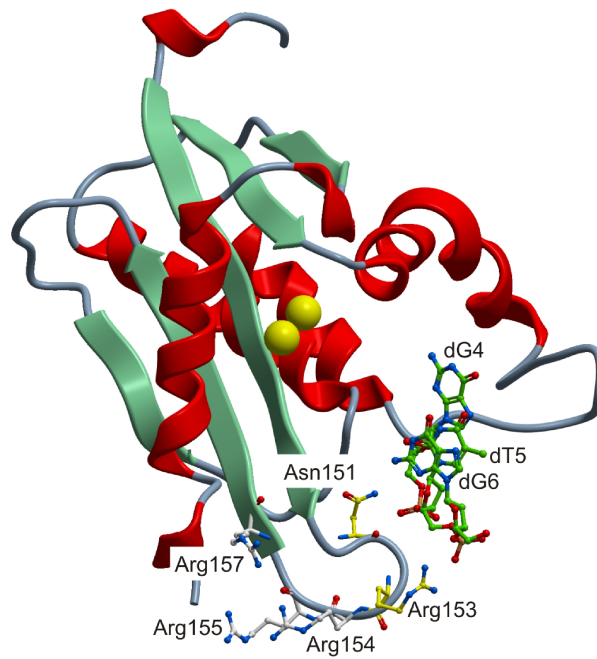


Figure S2: The dG4-dT5 and dT5-dG6 internucleotide linkage binding into the 4-5th DNA binding site

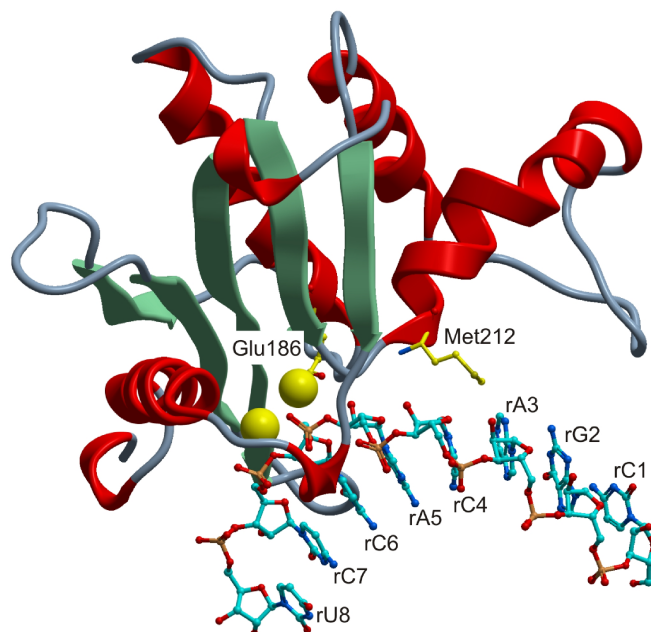
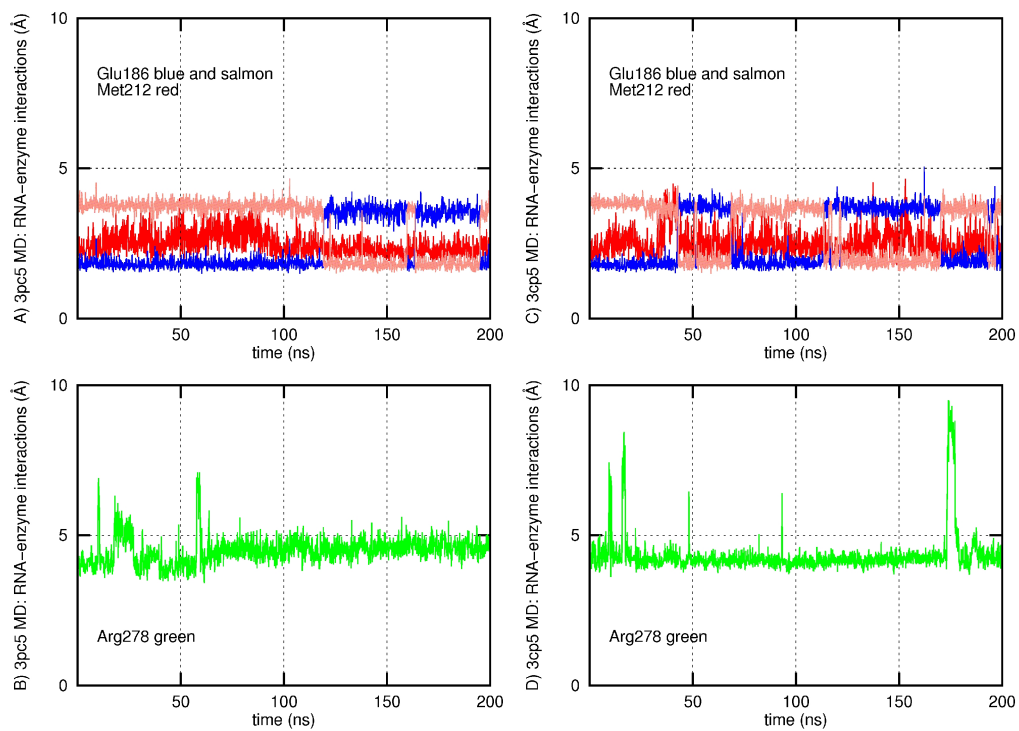


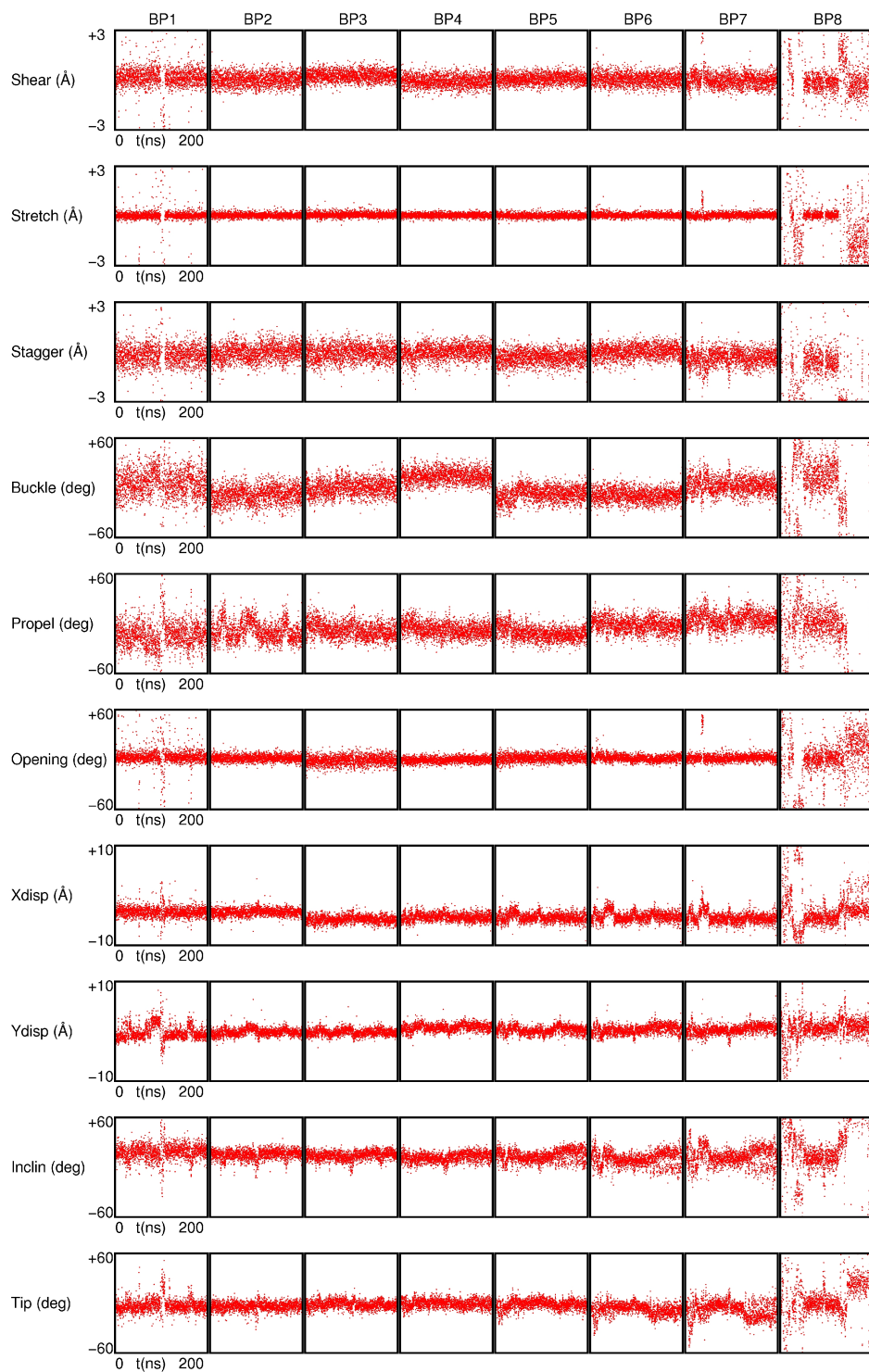
Figure S3: RNA recognition by HS RNase H

BOND				
CT-P	186.0	1.846		
ANGLE				
OS-P -CT	45.0	102.60		
P -CT-HC	50.0	110.00		
P -CT-OS	50.0	110.00		
O2-P -CT	100.0	108.23		
HC-CT-OS	50.0	110.00		
OS-CT-OS	50.0	109.50		
DIHEDRAL				
X -CT-P -X	9	1.20	0.0	3.
CT-P -OS-CT	1	0.00	0.0	-3.
CT-P -OS-CT	1	0.00	0.0	-2.
CT-P -OS-CT	1	1.00	180.0	1.
OS-CT-P -OS	1	0.00	0.0	-3.
OS-CT-P -OS	1	0.00	0.0	-2.
OS-CT-P -OS	1	0.00	0.0	1.
P -CT-OS-CT	1	0.00	0.0	3.

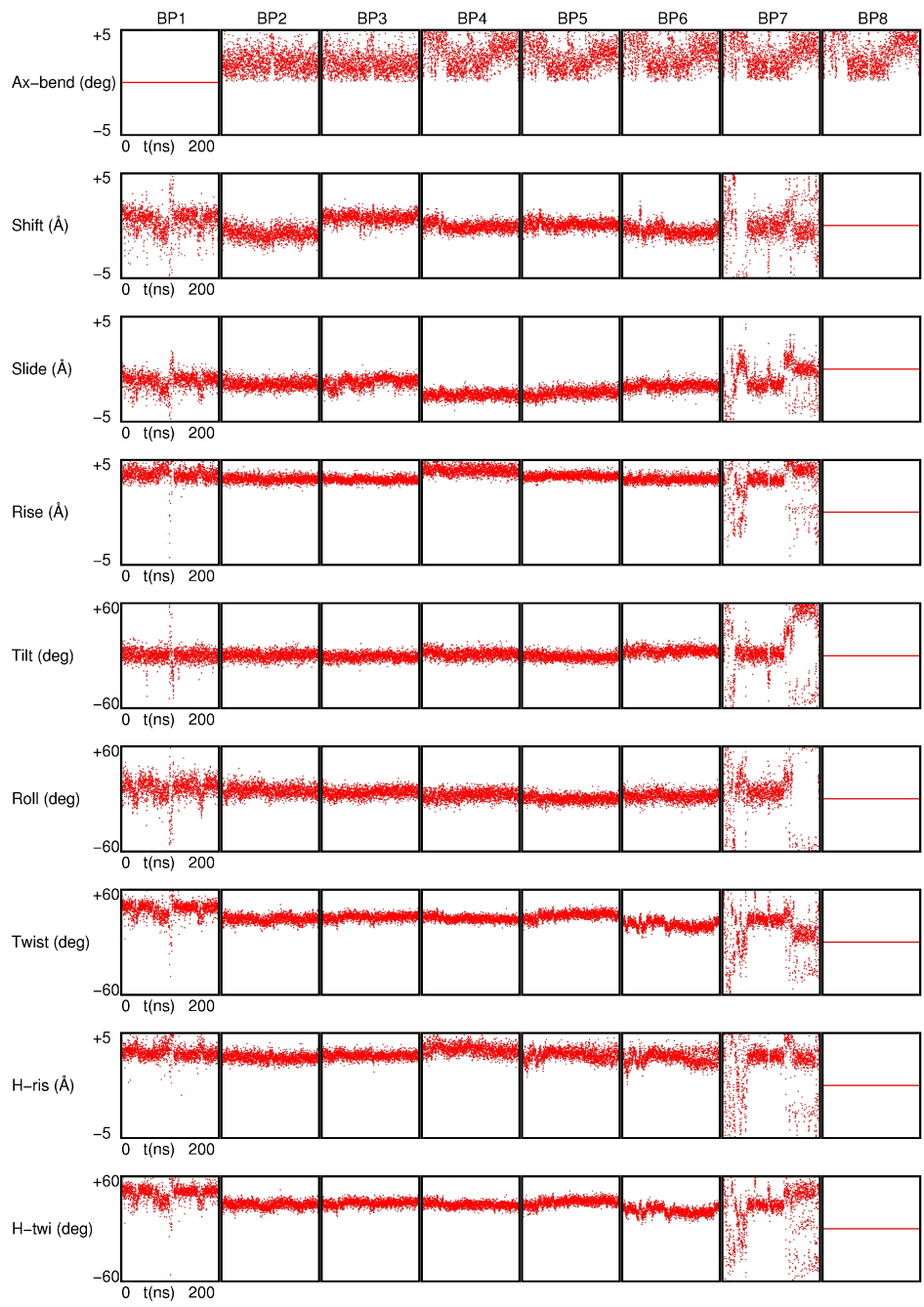
Table S1: Force constants for phosphonate internucleotide linkages.



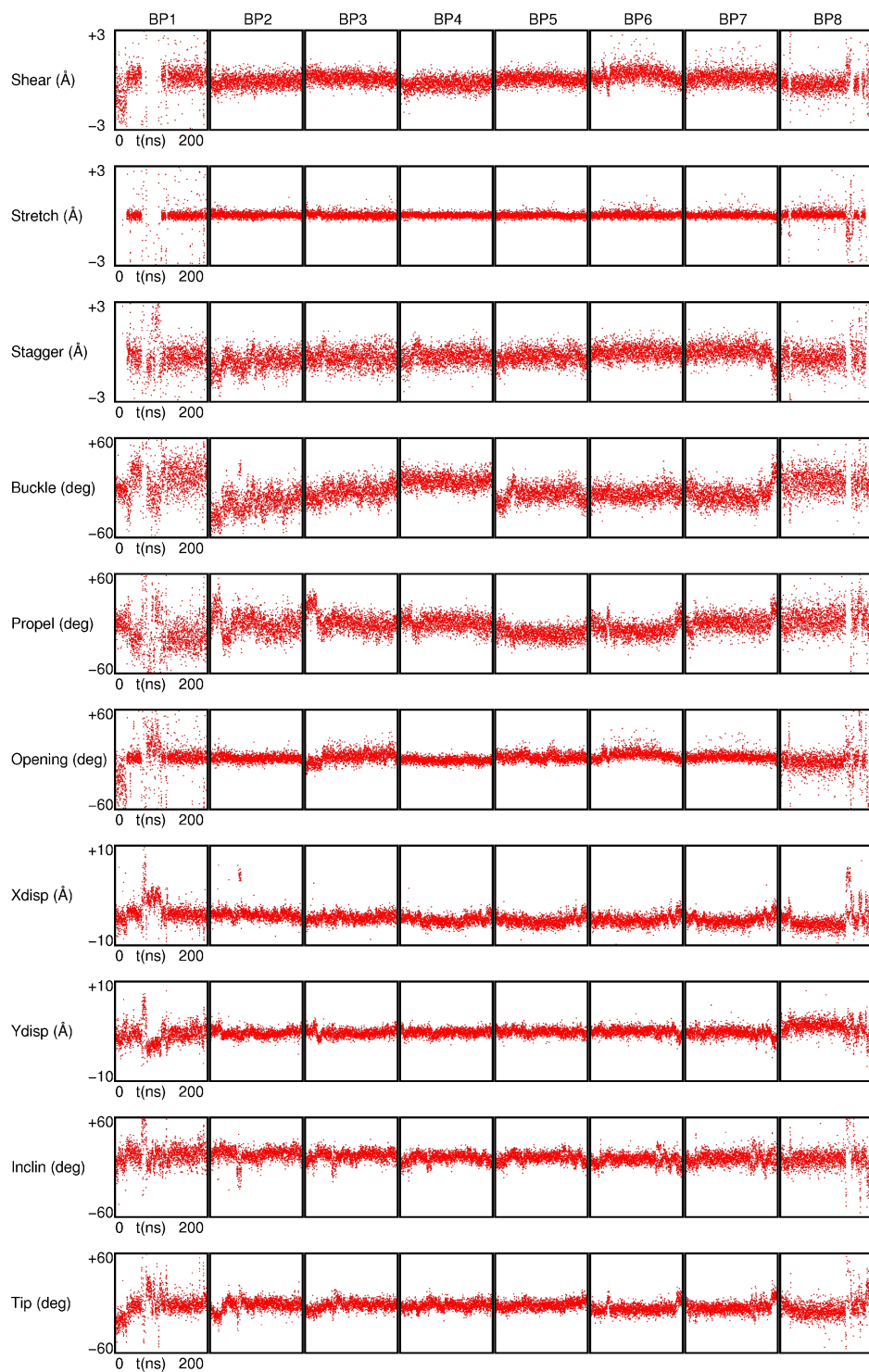
Graph S1: RNA recognition by HS RNase H - Glu186 blue and salmon line, Met212 red line, Arg278 green line



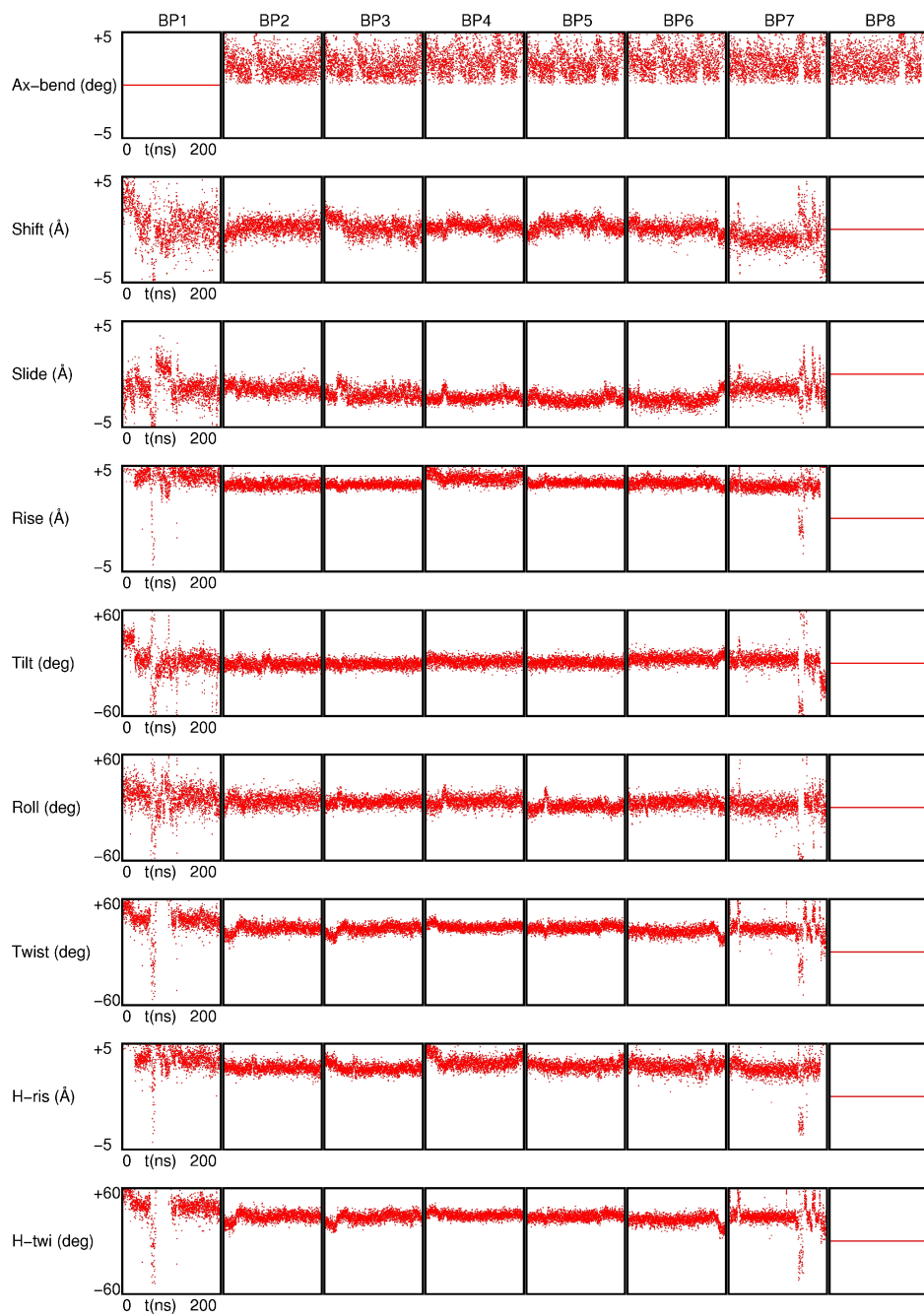
Graph S2: Helical parameters in the 3pc5 DNA:RNA duplex structure



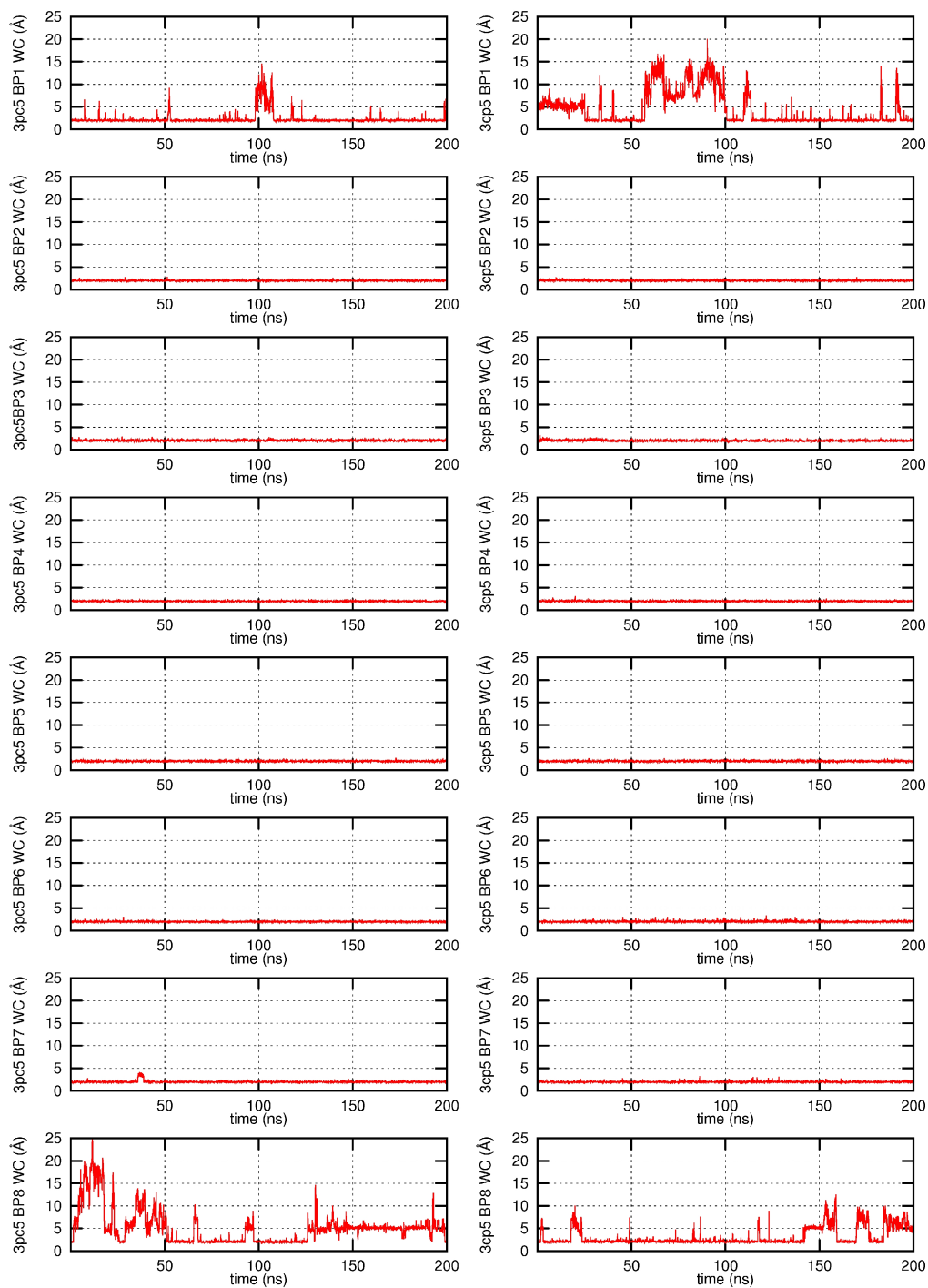
Graph S3: Helical parameters in the 3pc5 DNA:RNA duplex structure



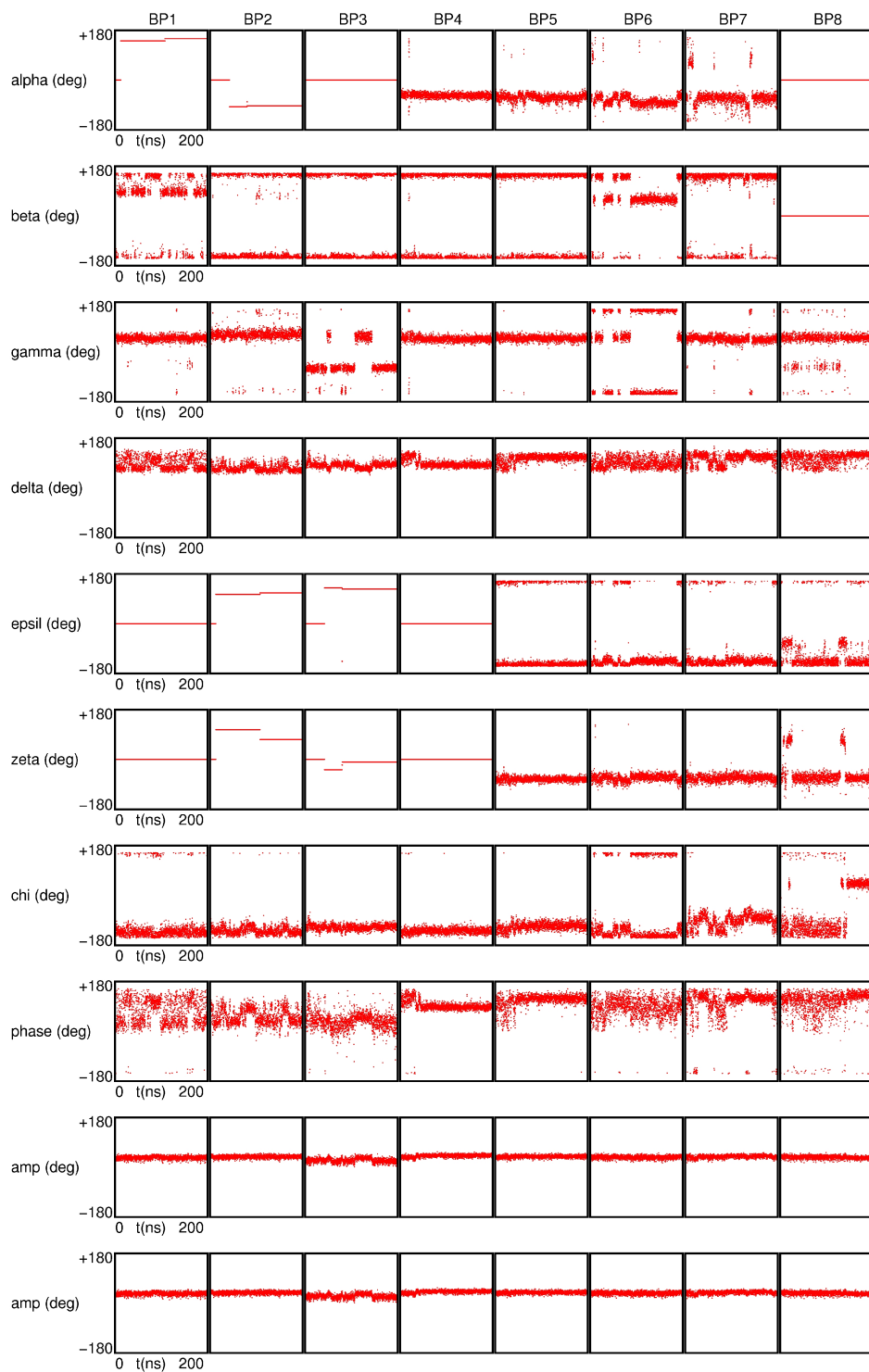
Graph S4: Helical parameters in the 3cp5 DNA:RNA duplex structure



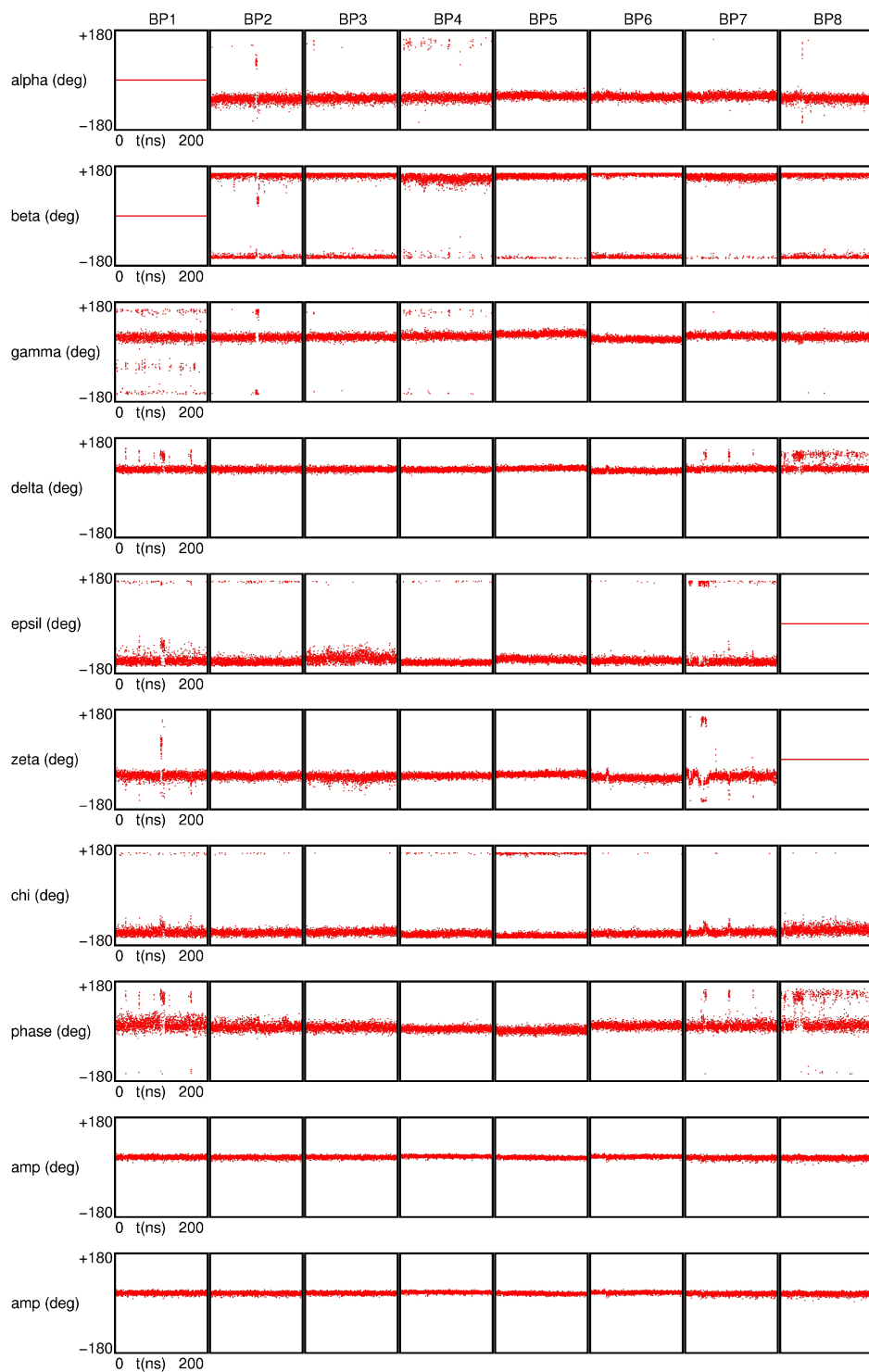
Graph S5: Helical parameters in the 3cp5 DNA:RNA duplex structure



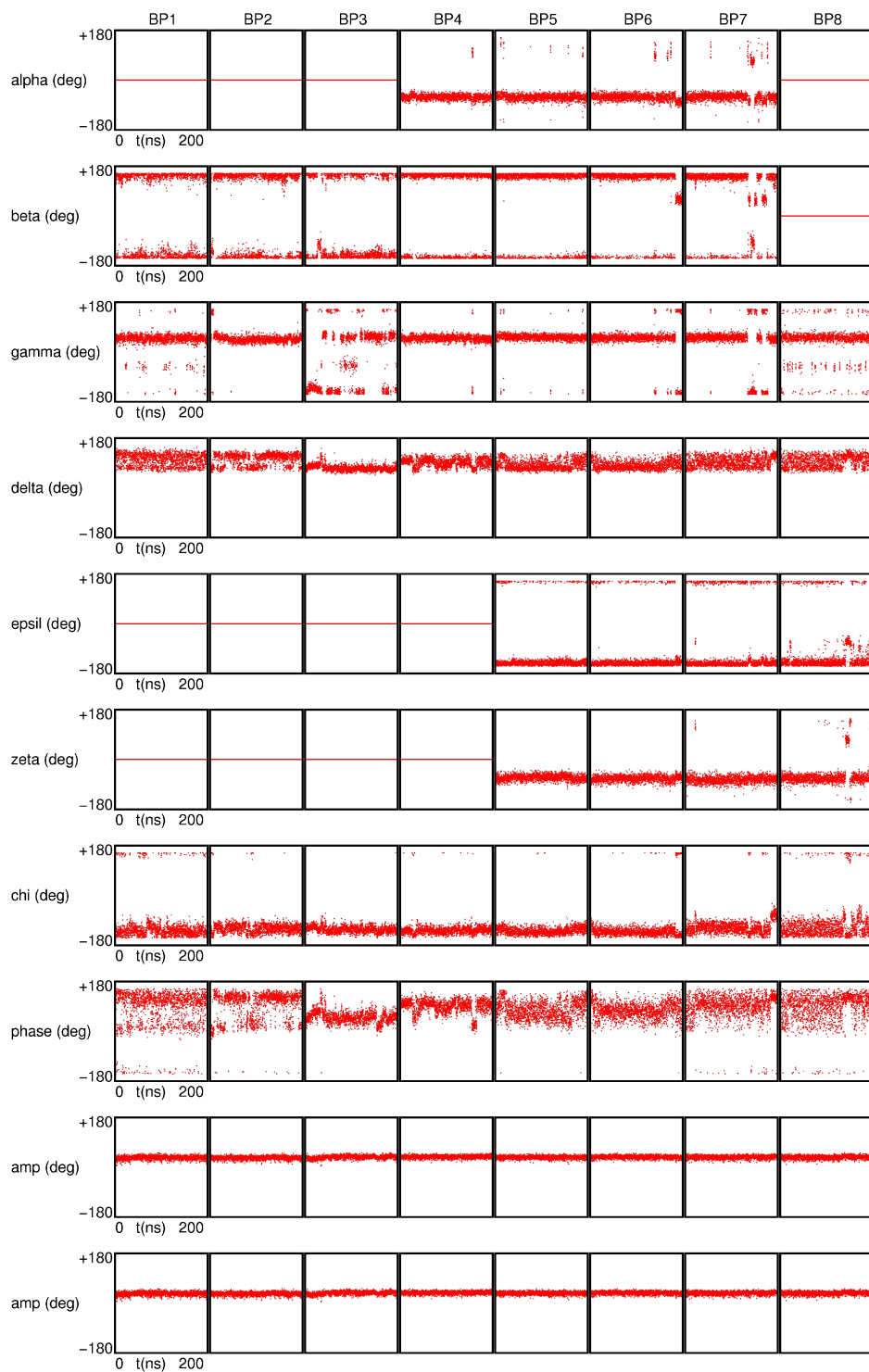
Graph S6: Watson-Crick hydrogen bonding in the 3pc5 and 3cp5 DNA:RNA duplex structure



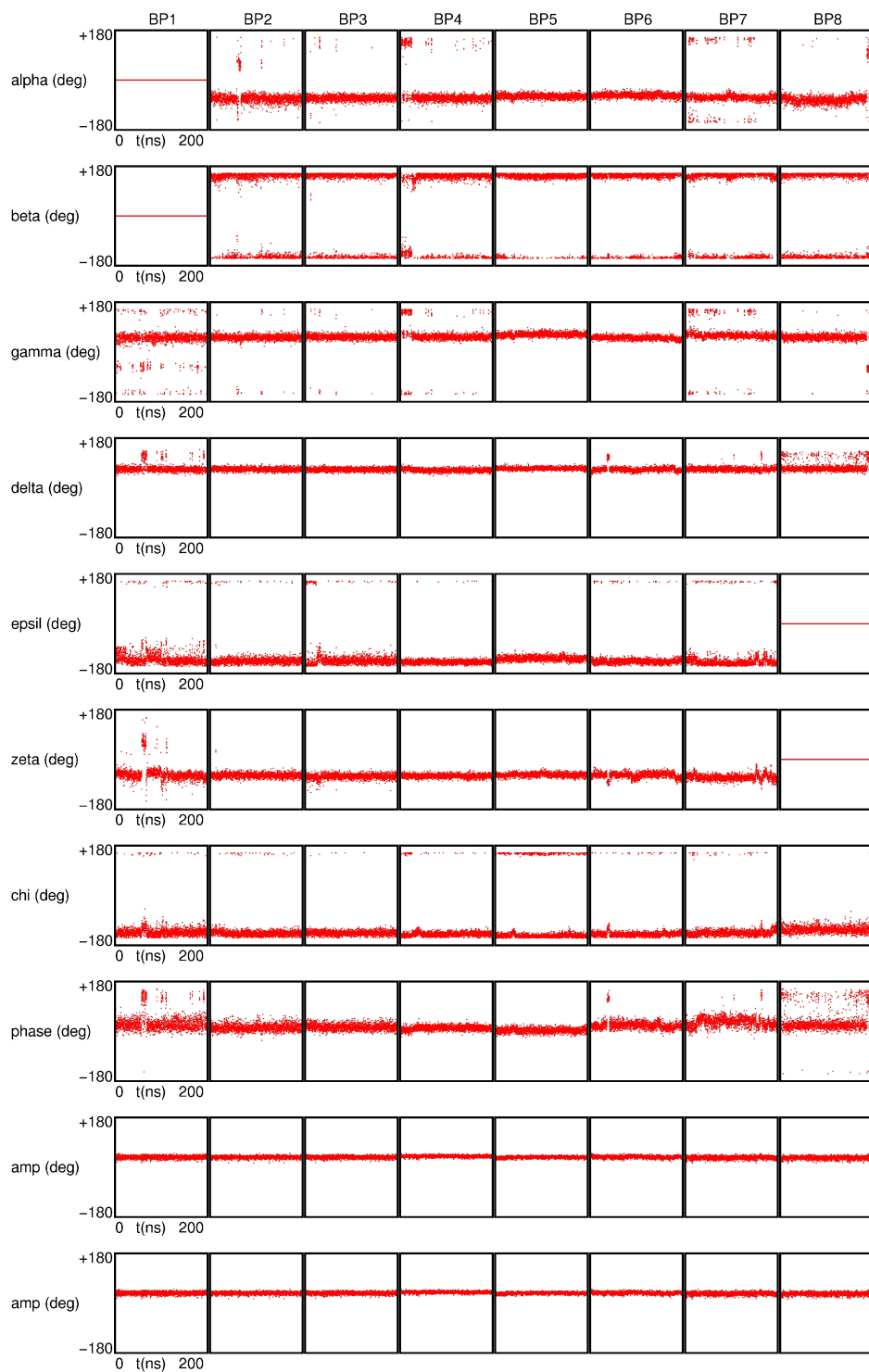
Graph S7: Torsion angles and sugar pucker along 3pc5 DNA



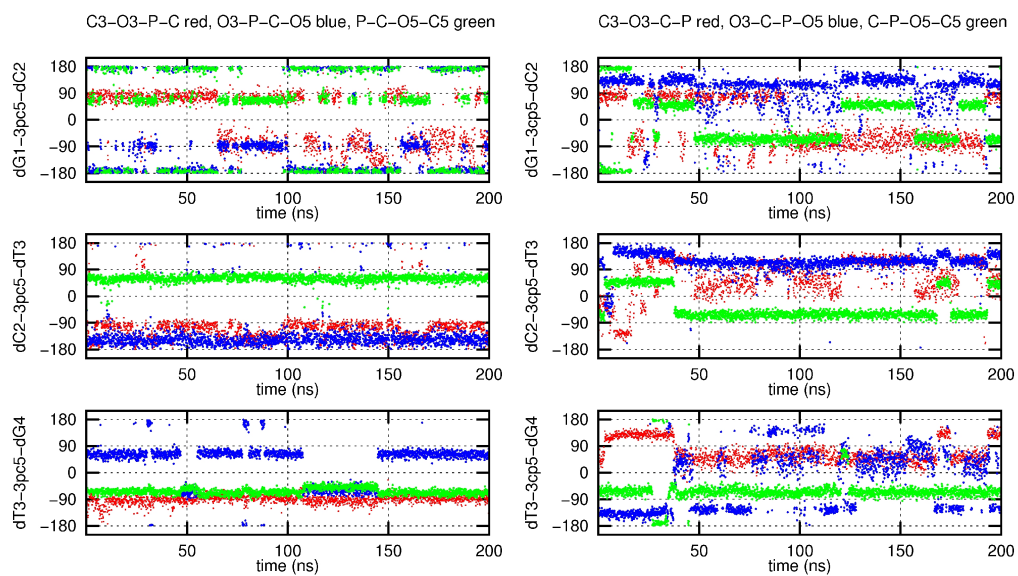
Graph S8: Torsion angles and sugar pucker along 3pc5 DNA



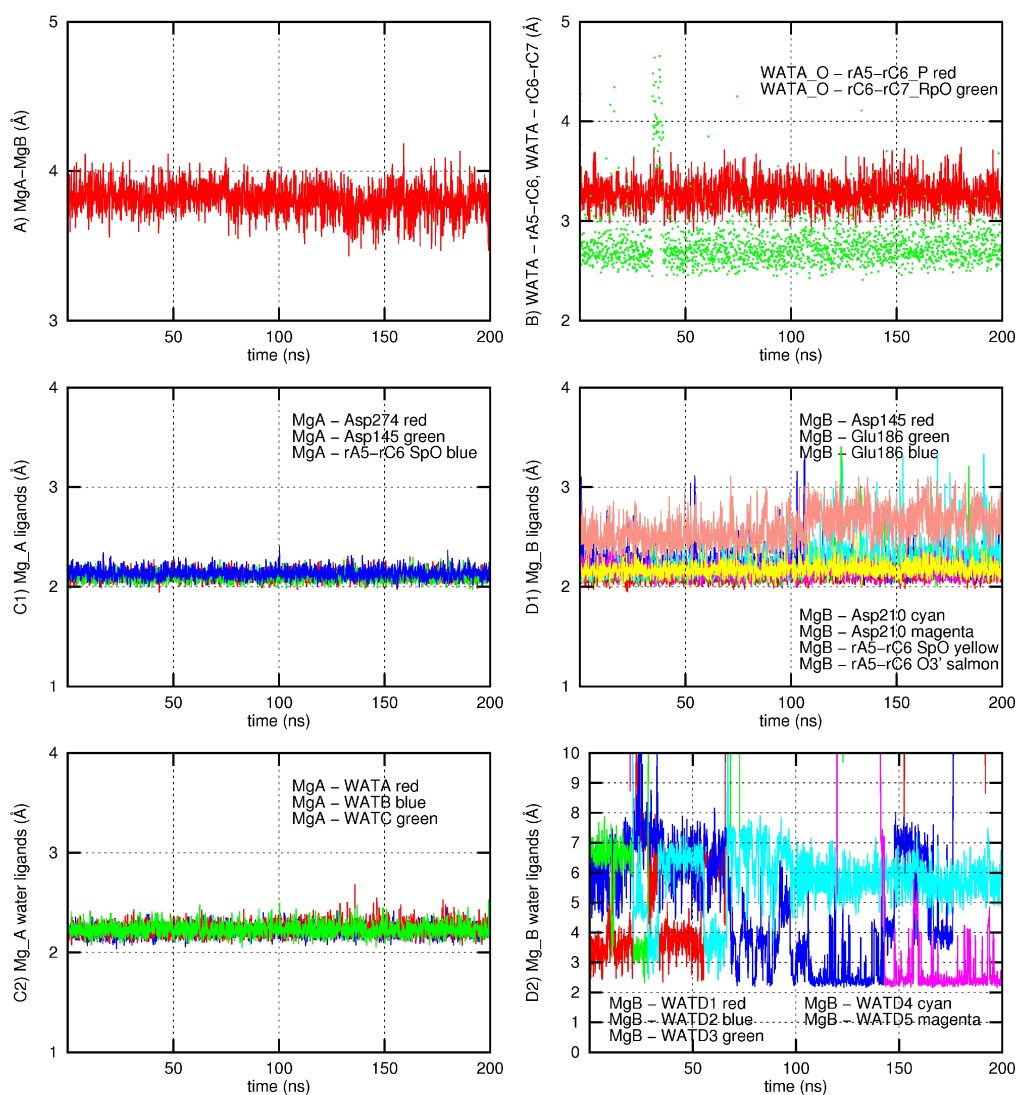
Graph S9: Torsion angles and sugar pucker along 3cp5 DNA



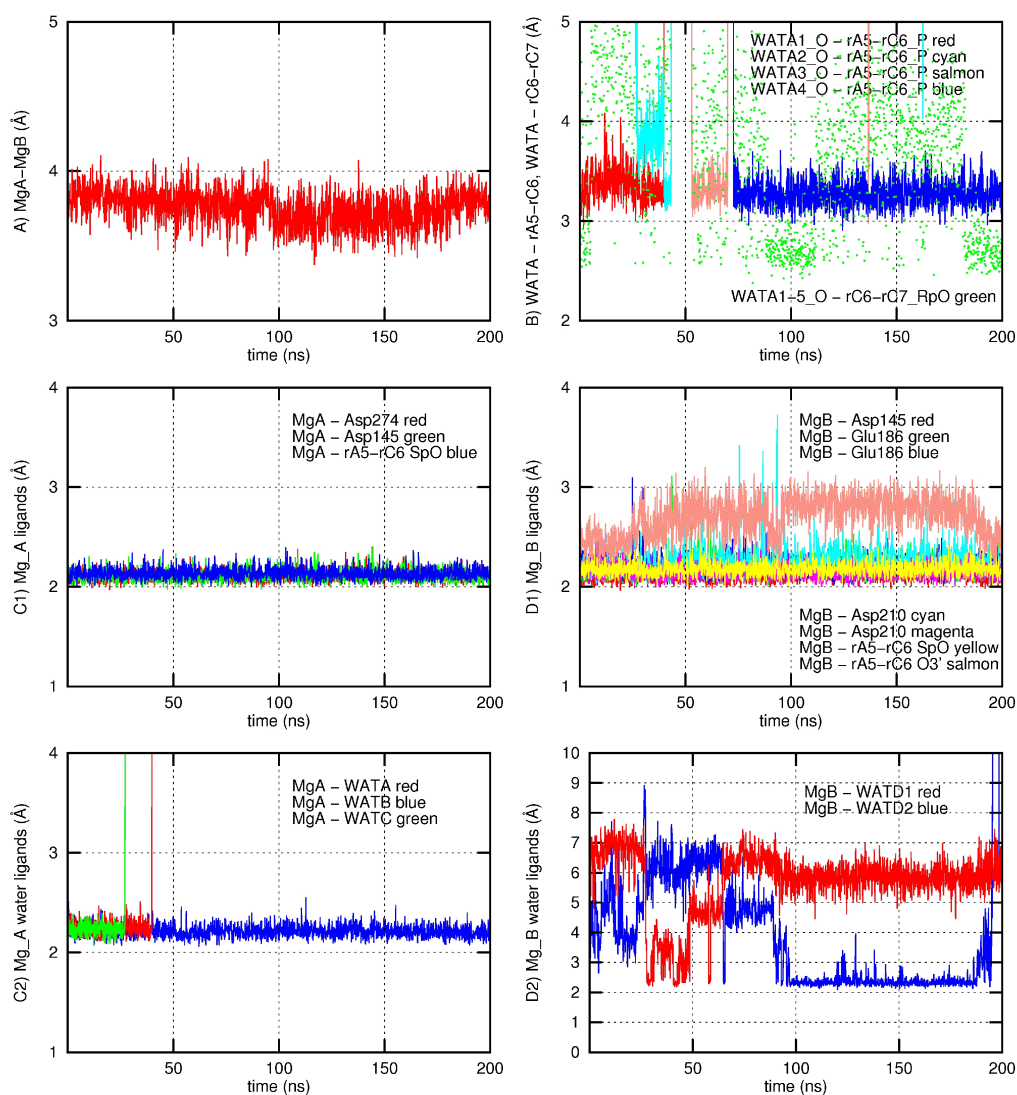
Graph S10: Torsion angles and sugar pucker along 3cp5 DNA



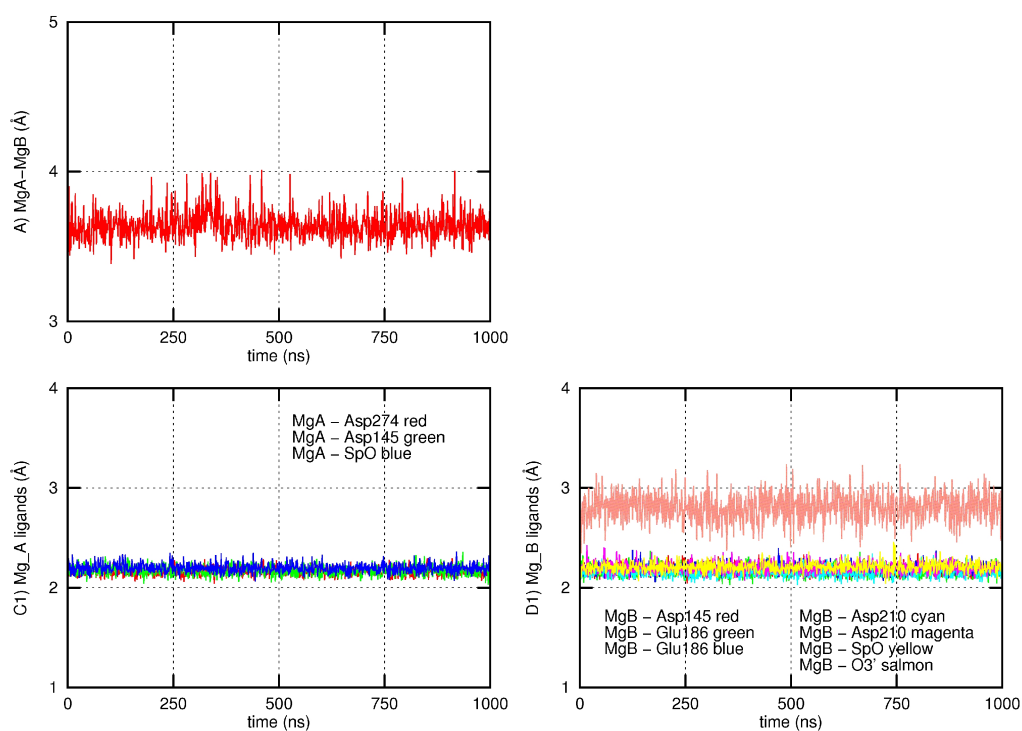
Graph S11: Conformational preferences of the 3pc5 and 3cp5 internucleotide linkages in terms of either the C3'-O3'-P-C, O3'-P-C-O5', P-C-O5'-C5' or C3'-O3'-C-P, O3'-C-P-O5', C-P-O5'-C5' torsion angles.



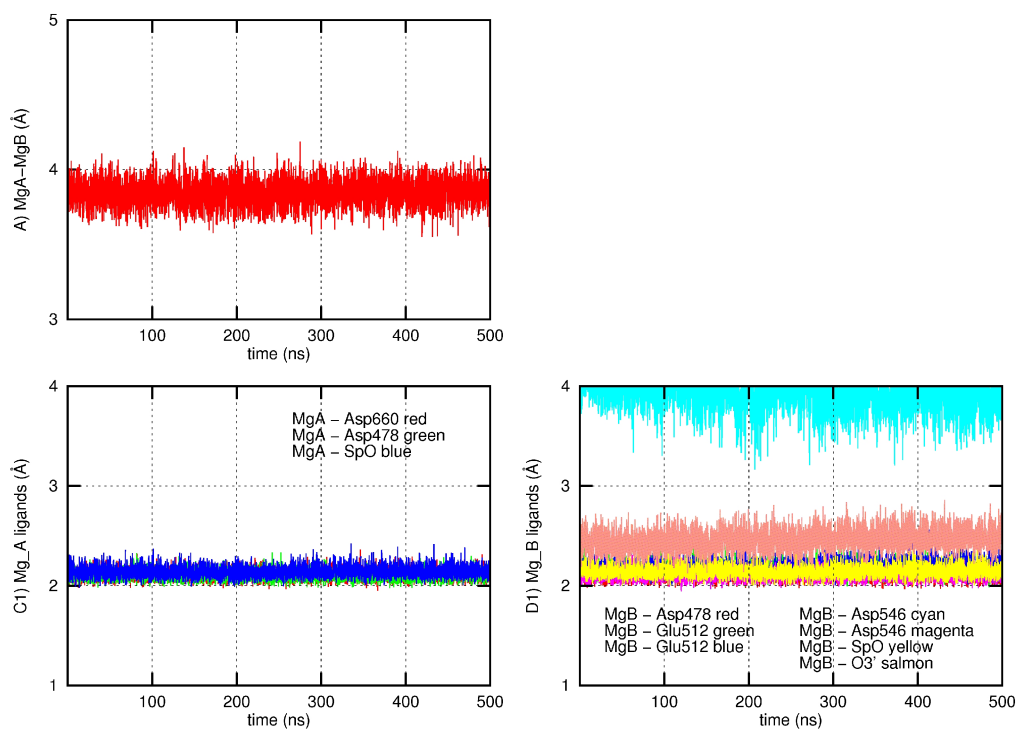
Graph S12: Active site contacts in 3pc5 MD - A) MgA-MgB distance (Graph 5A); C) MgA coordination shell (Graph 5C1: Asp274 red line, Asp145 green line, Sp blue line; Graph 5C2: WATA-C); D) MgB coordination shell - (Graph 5D1: Asp145 red line, Glu186 green line, Glu186 blue line, Asp210 cyan line, Asp210 magenta line, Sp yellow line, O3' salmon line, Graph 5D2: WAT)



Graph S13: Active site contacts in 3cp5 MD - A) MgA-MgB distance (Graph 6A); C) MgA coordination shell (Graph 6C1: Asp274 red line, Asp145 green line, Sp blue line; Graph 6C2: WATA-C); D) MgB coordination shell - (Graph 6D1: Asp145 red line, Glu186 green line, Glu186 blue line, Asp210 cyan line, Asp210 magenta line, Sp yellow line, O3' salmon line, Graph 6D2: WAT)



Graph S14: Active site contacts in 7p RNase H MD - A) MgA-MgB distance (Graph 6A); C) MgA coordination shell (Graph 6C1: Asp274 red line, Asp145 green line, Sp blue line; D) MgB coordination shell - (Graph 6D1: Asp145 red line, Glu186 green line, Glu186 blue line, Asp210 cyan line, Asp210 magenta line, Sp yellow line, O3' salmon line)



Graph S15: Active site contacts in TtAGO MD - A) MgA-MgB distance (Graph 6A); C) MgA coordination shell (Graph 6C1: Asp660 red line, Asp478 green line, Sp blue line; D) MgB coordination shell - (Graph 6D1: Asp478 red line, Glu512 green line, Glu512 blue line, Asp546 cyan line, Asp546 magenta line, Sp yellow line, O3' salmon line)

1 **Recognition of 2',5'-linked oligoadenylates by human ribonuclease**

2 **L: molecular dynamics study**

3
4 Kamil Malá , Ivan Barvík

5 *Charles University, Faculty of Mathematics and Physics, Institute of Physics, Ke Karlovu 5,*
6 *Prague 2, 121 16, Czech Republic*

7
8 *Corresponding author: Ivan Barvík*

9 *Fax: +420 224 922 797*

10 *Phone: +420 221 911 450*

11 *Email: ibarvik@karlov.mff.cuni.cz*

12 13 14 15 **ACKNOWLEDGMENTS**

16
17
18 This work was supported by the Grant Agency of the Czech Republic (13-26526S) and Grant
19 Agency of Charles University (426311). The access to computing and storage facilities owned
20 by parties and projects contributing to the National Grid Infrastructure MetaCentrum,
21 provided under the programme "Projects of Large Infrastructure for Research, Development,
22 and Innovations" (LM2010005) is highly acknowledged.

23
24 **KEYWORDS:** RNase L, ankyrin, 2',5' oligoadenylate, molecular dynamics, force field,
25 NAMD, ACEMD, GPU, CUDA

26 27 28 **ABSTRACT**

29
30
31 Molecular dynamics simulations have shown that non-protonated Glu131 leads to dissociation
32 of 2',5'-linked oligoadenylates (pA₂₅A₂₅A; 2-5A) from monomeric human RNase L. In
33 contrast, in MD simulations with protonated Glu131, the RNase L + 2-5A complex was
34 stable. Further, MD simulations were used to capture detrimental impacts of various point
35 mutations of RNase L (R155A, F126A, W60A, K89A) on 2-5A binding. These
36 conformational transitions were observed on a time scale of 200 ns. Finally, 2-5A analogues
37 with a bridged 3'-O,4'-C-alkylene linkage (B) introduced into the adenosine units (A) were
38 used to assess ability of MD simulations to distinguish on the time scale of hundreds of
39 nanoseconds between agonists of RNase L (pA₂₅A₂₅B, pB₂₅A₂₅A, pB₂₅A₂₅B) and inactive
40 analogs (pA₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅A, pB₂₅B₂₅B). Agonists were potently bound to RNase
41 L during 200 ns MD runs. For inactive 2-5A analogs, by contrast, there were found significant
42 disruptions of their interactions with RNase L already within 100 ns MD runs.

1 INTRODUCTION

2
3 In mammals, viral infections initiate an innate immune response predominantly mediated by
4 type I interferons (IFNs). Type I IFNs regulate the transcription of a number of genes that
5 inhibit or block viral replication by diverse mechanisms [1-8]. Combination therapy with IFN
6 and ribavirin can produce sustained responses against Hepatitis C virus (HCV) in some
7 patients [9] (HCV has infected about four million adults in the U.S. and is a major cause of
8 chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma).

9
10 Treatment of cells with IFN activates genes encoding several 2',5'-linked oligoadenylate
11 synthetases (OASs) and a single gene encoding ribonuclease L (RNase L). The OASs are
12 activated by binding to dsRNA, a frequent byproduct of viral infection. The activated OASs
13 generate 2',5'-linked oligoadenylates (2-5A) from ATP [1-8]. Upon binding 2-5A, RNase L is
14 converted from an inactive monomer (Figure 1) to a potently active dimeric endoribonuclease
15 for single-stranded RNA [5] containing dyads of UU, UA, AU, AA, and UG [10].

16
17 RNase L and OASs contribute to innate immunity in many ways [5]. Even a single cleavage
18 event in a viral ssRNA genome will prevent that genome from replicating [3-4]. Cleavage of
19 viral mRNA by RNase L contributes to inhibition of viral protein synthesis as well [3-4].
20 Activated RNase L within cells cleaves both viral and nonviral ssRNA substrates. Damage to
21 the host cell machinery required for viral replication, in particular ribosomes (28S and 18S
22 rRNA), contributes to the antiviral effects of RNase L. Interestingly, RNase L perpetuates and
23 amplifies IFN production [11-12]. RNase L cleaves single-stranded regions of RNA, leaving
24 as cleavage products short duplex RNAs with 3'-phosphoryl groups. These RNase L-
25 generated small RNAs stimulate innate immunity. Therefore, RNase L deficient mice produce
26 significantly less IFN- α in response to viral infections than identically infected WT mice. A
27 truncated form of RNase L, which is thought to be constitutively active, may be a biochemical
28 marker of chronic fatigue syndrome [13]. Sustained activation of RNase L or its activation
29 beyond a threshold level causes the cell to spiral into an RNA damage stress response that
30 culminates in apoptosis. It suggests a tumor suppressor function of the OAS/RNase L
31 pathway [3-4]. Further, RNase L is involved in cytokine induction and endosomal pathways
32 that suppress bacterial infections [6]. Therefore, the direct activation of RNase L by 2-5A
33 might be a novel chemotherapeutic approach to viral and bacterial diseases as well as cancer.

34
35 The human form of RNase L is a 741-amino-acid protein with a molecular mass of 83 543
36 Da. RNase L consists of three domains, namely the N-terminal ankyrin repeat domain (ARD),
37 the protein kinase homology domain (RNase L belongs to the family of >500 human kinases),
38 and the C-terminal ribonuclease domain. The N-terminal ARD sensor domain (residues 24-
39 335) containing nine ankyrin-like macromolecular recognition repeats, is responsible for 2-5A
40 binding (see Figure 1). The pseudokinase domain (residues 336-582) is catalytically inactive.
41 The C-terminal ribonuclease domain of RNase L (residues 583-720) is responsible for its
42 catalytic activity.

43
44 In fact, RNase L has the only ARD known to bind to an oligonucleotide [5] (ARDs in
45 transient receptor potential vanilloid channel proteins have a multi-ligand binding site for
46 ATP and calmodulin [14-17]). Interestingly, RNase L is the only enzyme that recognizes and
47 binds 2-5A, whereas the only known function of 2-5A is to activate RNase L [13]. The
48 length of the 2-5A species is critical for the activation of RNase L [2, 8, 41]. While the dimer
49 failed to activate RNase L, the trimeric and tetrameric 2-5As maximally activated RNase L at

1 the same concentration. Replacing either of the 2',5'-linkages in a 2-5A trimer by 3',5'
2 phosphodiester decreases binding to RNase L by 1 and 2 orders of magnitude, however,
3 substituting both linkages reduces binding drastically. RNase L has lower affinity for cytidine,
4 uridine, or inosine analogs of the adenine rings in 2-5A. Only a single 5'-phosphoryl group on
5 the trimer of 2',5'-linked oligoadenylate is required for the full activation of human RNase L.
6

7 The structural basis of 2-5A binding to RNase L was established by the crystal structure of the
8 sensor domain in complex with 2-5A [8]. The ligand 2-5A is bound by the second to the
9 fourth ankyrin repeats and interacts with the RNase L residues Trp60, Asn65, Gln68, Lys89,
10 Phe126, Glu131, Tyr135, and Arg155 by forming hydrogen bonds, salt bridges, or stacking
11 interactions [8] (see Figure 1). To determine the contributions of amino acid residues
12 surrounding the 2-5A binding site, mutant proteins were analyzed for their interaction with 2-
13 5A using a steady-state fluorescence technique. In addition, full-length RNase L mutants were
14 tested for their activation by 2-5A [13]. The results reveal that π - π stacking interactions of
15 Trp60 and Phe126, electrostatic interactions of Lys89 and Arg155, and hydrogen bonding by
16 Glu131 make crucial contributions to 2-5A binding [13].
17

18 Paradoxically, the first crystal structure of RNase L ARD [8] could not explain the
19 mechanism of 2-5A-dependent RNase L dimerization because the sensor domain crystallized
20 as a monomer [8]. Recently was found that the sensor domain undergoes 2-5A-mediated self-
21 association without any conformational changes [18]. Two molecules of 2-5A at a time tether
22 RNase L monomers via their ANK domains. Each ANK domain harbors two distinct sites for
23 2-5A recognition that reside 50 Å apart.
24

25 The direct activation of RNase L by 2-5A might be a novel chemotherapeutic approach both
26 to viral and bacterial diseases and cancer. It needs, however, to solve the problem with
27 transporting of highly charged 2-5A oligoadenylates into cells [19-22]. It could be resolved
28 thanks to unremitting effort to develop potent synthetic nucleic acids lipid particles in the
29 context of antisense oligonucleotides, siRNA molecules, oligonucleotide aptamers etc. And
30 there are also unexpected breakthroughs like exosomes, naturally occurring virus-like
31 particles produced by cells and transporting nucleic acids between them [23].
32

33 Moreover, natural 2-5As have a short biological half-life because of its rapid degradation by
34 phosphatases and nuclease activities in cells and in serum [24]. Thus far, a number of 2-5A
35 analogues with modification of bases, riboses, internucleotide linkages and a 5'-phosphoryl
36 group have been synthesized to test RNase L activating ability in vitro [25-44]. In [24], a
37 bridging strategy was applied to develop novel 2-5A analogues by introducing a 3'-O,4'-C-
38 alkylene linkage into the adenosine units, and then examining their properties including their
39 RNase L activating ability and their nuclease stability.
40

41 Here, MD simulations were used to understand detrimental impacts of various point
42 mutations of RNase L (R155A, F126A, W60A, K89A) on 2-5A binding [13]. Further, 2-5A
43 analogues with a bridged 3'-O,4'-C-alkylene linkage [24] introduced into the adenosine units
44 were used to assess ability of MD simulations to distinguish on the time scale of hundreds of
45 nanoseconds between agonists of RNase L (pA₂₅A₂₅B, pB₂₅A₂₅A, pB₂₅A₂₅B) and inactive
46 analogs (pA₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅A, pB₂₅B₂₅B).
47
48
49

1 METHODS

2
3 The model systems of human RNase L complexed with pA₂₅A₂₅A or its analogs (see [Table S1](#)) were constructed using the crystal structure [\[8\]](#) (Protein Data Bank entry 1wdy). All
4 simulated systems were surrounded by TIP3P [\[45\]](#) water molecules which extended to a
5 distance of approximately 10 Å (in each direction) from the enzyme-nucleic acids atoms. This
6 gives a periodic box size of ~121 Å, ~77 Å, ~77 Å for a simulated system consisting of
7 ~70.000 atoms. New *.inpcrd (initial coordinates) and *.prmtop (molecular topology, force
8 field [\[46-50\]](#)) files for the whole simulated system, were created by means of the TLEAP
9 module (the AMBER software package [\[51\]](#)).

10
11
12 Equilibration MD trajectories lasting for 5 ns were computed with the aid of the NAMD 2.7
13 software package [\[52\]](#). The smooth Particle-mesh Ewald (PME) method was employed for
14 long-range electrostatic forces [\[53\]](#). The non-bonded cutoff was set to 9 Å. The SHAKE
15 algorithm (tolerance 0.0005) was applied to constrain bonds where the hydrogen atoms were
16 involved [\[54\]](#). Simulated systems were energy minimized then the Langevin dynamics was
17 used for a temperature control [\[52\]](#). The simulated systems were heated from 0 K to 290/310
18 K. The Langevin piston method was applied to reach an efficient pressure control with target
19 pressure set to 1 atm [\[52\]](#). The integration timestep was set to 2 fs. A multiprocessor system
20 equipped with 16 Intel Itanium CPUs (1.5 GHz) was used for equilibration MD simulations.

21
22 For production MD runs we have used ACEMD v. 2591 [\[55-56\]](#) running on a local
23 workstation equipped with an NVIDIA graphics processing unit (GPU). ACEMD is a new
24 generation molecular dynamics software which runs exclusively on GPUs at the equivalent
25 speed of tens to hundred of standard processors. ACEMD implements all features of an MD
26 simulation on a CUDA-compatible GPU device, including those usually required for
27 production simulations in the NVT ensemble (i.e., bonded and nonbonded force term
28 computation, velocity-Verlet integration, Langevin thermostatic control, smooth PME, and
29 hydrogen bond constraints implemented using the M-shake algorithm and RATTLE for
30 velocity constraints within the velocity Verlet integration scheme [\[57-58\]](#)). ACEMD does not
31 presently contain a barostat for production runs, so simulations in the NPT ensemble are not
32 possible. However, it is noted that with large molecular systems, changes in volume due to the
33 pressure control are very limited after an initial equilibration making NVT simulations viable
34 for production runs [\[55-56\]](#). It was proven on multiple biomolecular systems including
35 membrane proteins [\[59-61\]](#). Therefore, our production runs have been performed in the NVT
36 ensemble, Langevin thermostat at 290/310 K, computing the electrostatic interactions with
37 PME. Also implemented in ACEMD is the hydrogen mass repartitioning scheme (the mass of
38 the bonded heavy atoms to hydrogen is repartitioned among hydrogen atoms, leaving the total
39 mass of the system unchanged) used, for instance, in codes such as Gromacs, which allows an
40 increased time step [\[55, 62\]](#). Therefore, the integration time-step in our production MD runs
41 was set to 4 fs. For production 100-200 ns MD runs CUDA programmable NVIDIA GTX-
42 580 GPU equipped with 480 cores was used. MD trajectories were computed at ~25 ns per
43 day (i.e. ~136 days of GPU time in total).

44
45 Data were recorded every 100 ps. MD trajectories were analyzed with the aid of the
46 CHIMERA 1.5.3, VMD 1.9 and AMBER10/ptraj software packages [\[63-64, 51\]](#). Figures
47 were produced by means of the CHIMERA 1.5.3 software package.

1 RESULTS

2
3 In this study, several versions of the AMBER force field were compared (MD simulations no.
4 1, 4 [46]; MD no. 2, 5 [47-49]; MD no. 3, 6 [50] - see Table S1). Further, RNase L with either
5 protonated or unprotonated Glu131 was tested in this subset of MD simulations (i.e. MD
6 simulations no. 1-3 vs. 4-6). MD simulations were then used to capture detrimental impacts of
7 various point mutations of RNase L (R155A - MD no. 7-8; F126A - MD no. 9; W60A - MD
8 no. 10; K89A - MD no. 11-12 - see Table S1) on 2-5A binding. Finally, 2-5A analogues with
9 a bridged 3'-O,4'-C-alkylene linkage [24] introduced into the adenosine units were used to
10 assess ability of MD simulations to distinguish on the time scale of hundreds of nanoseconds
11 between agonists of RNase L (pA₂₅A₂₅B, pB₂₅A₂₅A, pB₂₅A₂₅B - MD no. 13-15 - see Table
12 S1) and inactive analogs (pA₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅A, pB₂₅B₂₅B - MD no. 16-19 - see
13 Table S1).

14
15 *01) RNase L_E131H --- pA₂₅A₂₅A --- FF_94 [46] --- T=310K --- t<200ns*

16
17 The pA₂₅A₂₅A molecule was bound to the surface of RNase L_E131H despite of
18 conformational transitions of the adenine ring of the third AMP (Figure S1, Figure 2).

19
20 The bound pA₂₅A₂₅A molecule adopted an extended conformation in which the adenine ring
21 of the second AMP (Ade2) was stacked with the adenine ring of the first AMP (Ade1),
22 whereas the adenine ring of the third AMP (Ade3) was separated from them. Ade1 was
23 stacked with the side chain of Phe126. Ade3 (regardless of its conformational transition - see
24 below) was stacked with the side chain of Trp60. Ade1 was fixed by hydrogen bonds with the
25 side chain of protonated Glu131 (i.e. Glu131_O-H62_Ade1 and Glu131_OH-N1_Ade1 - see
26 Figure 1) throughout the whole MD simulation. Nevertheless, short-living disruptions of
27 those hydrogen bonds occurred from time to time. The same was observed in the case of
28 contacts of Ade2 with the side chain of Tyr135 (i.e. Tyr135_OH-N1_Ade2 - see Figure 1).
29 The hydrogen bond network involving Ade3 and the side chains of Gln68 (i.e. Gln68_O1-
30 H62_Ade3 and Gln68_H2-N1_Ade3 - see Figure 1) and Asn65 (Asn65_O1-H61_Ade3 - see
31 Figure 1) was broken at the very beginning as a result of the Ade3 transition.

32
33 The side chains of Arg155 and Lys166 formed salt bridges with the 5'-phosphate group of the
34 first AMP (Phos1). The 5'-phosphate group of the second AMP (Phos2) was exposed to
35 solvent, and no direct interactions were found between Phos2 and the surface of ANK in
36 crystal structures [8, 18]. Nevertheless, here were observed short-lasting contacts of Phos2
37 with the Tyr124 side chain in the moment when Arg155 transiently lost contacts with Phos1
38 and Ade3 underwent the conformational transition. Eventually, Lys166 and Tyr124
39 exchanged their Phos1/Phos2 counterparts. The 5'-phosphate group of the third AMP (Phos3)
40 formed a salt bridge with the side chain of Lys89.

41
42 Ade1 and Ade2, respectively, were in the **high anti** (i.e. $\sim -90^\circ$ in terms of the glycosidic
43 torsion angle O4'-C1'-N9-C4 - see Figure 1) and **anti** ($\sim 180^\circ$) conformation, whereas Ade3
44 jumped from **syn** ($\sim 60^\circ$) to **high anti** conformation. In the following text such non-reversible
45 conformational transitions will be marked with the symbol \rightarrow (i.e. **syn** \rightarrow **high anti**), whereas
46 reversible transitions will be marked with the symbol \leftrightarrow . The first and second ribose rings
47 had **C3'-endo** puckering, whereas the third ribose had **C4'-exo/O4'-endo** \rightarrow **C2'-endo**
48 puckering. The C4'-C5' moieties connecting ribose and internucleotide linkage segments
49 adopted (in terms of the β : C3'-C4'-C5'-O5' torsion angle - see Figure 1) conformations

1 **trans** -> **gauche** (Ade1), **-gauche** (Ade2), **-gauche** -> **gauche** (Ade3). The internucleotide
2 linkage Ade1-Ade2 adopted conformation **-g-g** (in terms of torsion angles C5'-O5'-P-O2' and
3 O5'-P-O2'-C2' - see **Figure 1**). The internucleotide linkage Ade2-Ade3 preferred
4 conformation **gg** -> **-g-g**.

5
6 02) RNase L_E131H --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=310K --- t=200ns

7
8 The pA₂₅A₂₅A molecule was bound to the surface of RNase L_E131H in the original
9 arrangement (**Figure 1-2**, **Figure S2**). Therefore, we used the AMBER force field v. [47-49] in
10 all other MD simulations with mutant RNase L (MD no. 7-12 - see **Table S1**) or 2-5A analogs
11 (MD no. 13-19 - see **Table S1**).

12
13 The bound pA₂₅A₂₅A adopted an extended conformation in which Ade2 was stacked with
14 Ade1, whereas Ade3 was separated from them. Ade1 was stacked with the side chain of
15 Phe126, whereas Ade3 was stacked with the side chain of Trp60. Ade1 was fixed by
16 hydrogen bonds with the side chain of protonated Glu131. Ade2 was stabilized by a single
17 hydrogen bond with the side chain of Tyr135. Ade3 was fixed by a hydrogen bond network
18 involving the side chains of Gln68 and Asn65.

19
20 The side chains of Arg155 and Lys166 formed salt bridges with Phos1. There were not
21 observed significant contacts of Phos2 with the Tyr124 side chain. Phos3 formed steadily a
22 salt bridge with the side chain of Lys89.

23
24 Ade1 was in the **high anti** conformation, Ade2 was in the **anti** conformation, whereas Ade3
25 was in the **syn** conformation. The first and second ribose rings had **C3'-endo** puckering,
26 whereas the third ribose had **C4'-exo/O4'-endo** puckering. The C4'-C5' moieties adopted
27 conformations **trans** (Ade1), **-gauche** (Ade2), **-gauche** (Ade3). The internucleotide linkage
28 Ade1-Ade2 adopted conformation **-g-g**. The internucleotide linkage Ade2-Ade3 preferred
29 conformation **gg**.

30
31 03) RNase L_E131H --- pA₂₅A₂₅A --- FF_Olo [50] --- T=310K --- t=200ns

32
33 The pA₂₅A₂₅A molecule stayed bound to the surface of RNase L_E131H close to the original
34 arrangement (**Figure S3**, **Figure 2**).

35
36 The bound pA₂₅A₂₅A adopted an extended conformation in which Ade2 was stacked with
37 Ade1, whereas Ade3 was separated from them. Ade1 was stacked with the side chain of
38 Phe126. Ade3 was stacked with the side chain of Trp60, nevertheless in a little bit distorted
39 position, which resulted in disruption of hydrogen bonds formed between Ade3 and the side
40 chains of Gln68 and Asn65. Ade1 was fixed by hydrogen bonds with the side chain of
41 protonated Glu131 throughout the whole MD simulation lasting for 200 ns. Ade2 was fixed
42 by a single hydrogen bond with the side chain of Tyr135.

43
44 The side chains of Arg155 and Lys166 formed salt bridges with Phos1. There were found no
45 significant contacts of Phos2 with the Tyr124 side chain. Phos3 formed a salt bridge with the
46 side chain of Lys89.

47
48 Ade1 was in the **high anti** conformation, Ade2 was in the **anti** conformation, whereas Ade3
49 was in the **syn** conformation. The first and second ribose rings had **C3'-endo** puckering,

1 whereas the third ribose had **C4'-exo/O4'-endo** puckering. The C4'-C5' moieties adopted
2 conformations **trans** (Ade1), **-gauche** (Ade2), **-gauche** (Ade3). The internucleotide linkage
3 Ade1-Ade2 adopted conformation **-g-g**. The internucleotide linkage Ade2-Ade3 preferred
4 conformation **gg** (with two short living oscillations to **g-g**).

5
6 *04) RNase L_E131 --- pA₂₅A₂₅A --- FF_94 [46] --- T=290K --- t<200ns*

7
8 The pA₂₅A₂₅A molecule rearranged substantially with only Ade3 (which was the most distant
9 adenine ring from the unprotonated E131 residue) bound in its original position on the surface
10 of RNase L_E131 (Figure S4, Figure 2). Roughly the same was observed in subsequent MD
11 simulations no. 5-6 with unprotonated RNase L_E131. Therefore, protonated RNase
12 L_E131H was used in all other MD simulations with mutant RNase L (MD no. 7-12 - see
13 Table S1) or 2-5A analogs (MD no. 13-19 - see Table S1).

14
15 The partly unbound pA₂₅A₂₅A adopted an extended conformation in which Ade2 was stacked
16 with Ade1 (despite of their rearrangement), whereas Ade3 was separated from them. Ade1
17 lost stacking contacts with the side chain of Phe126, whereas stacking of Ade3 with the side
18 chain of Trp60 was preserved. Ade1 lost contacts with the side chain of non-protonated
19 Glu131 (i.e. Glu131_O-H62_Ade1 - see Figure S4). A single hydrogen bond between Ade2
20 and the side chain of Tyr135 was broken from the very beginning. In contrast, Ade3 was
21 potentially fixed by a hydrogen bond network involving the side chains of Gln68 and Asn65.

22
23 The side chains of Arg155 and Lys166 formed salt bridges with Phos1. There were found
24 increasingly more pronounced interactions of Phos2 with the Tyr124 side chain, which
25 stabilized distorted pA₂₅A₂₅A. Phos3 formed a salt bridge with the side chain of Lys89.

26
27 Glycosidic torsion angles were in conformations high **anti** -> **syn** (Ade1), **anti** -> **syn** (Ade2),
28 **syn** (Ade3). The first ribose ring had **C4'-exo** puckering, the second ribose had **C2'-endo**
29 puckering, whereas the third ribose had **C4'-exo** puckering. The C4'-C5' moieties adopted
30 conformations **trans** -> **gauche** -> **trans** -> **-gauche** (Ade1), **-gauche** -> **gauche** (Ade2), -
31 **gauche** -> **gauche** (Ade3). The internucleotide linkage Ade1-Ade2 adopted conformation **-g-**
32 **g**. The internucleotide linkage Ade2-Ade3 preferred conformations **gg** -> **g-g** -> **-g-g**.

33
34 *05) RNase L_E131 --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=290K --- t=200ns*

35
36 Again, the pA₂₅A₂₅A molecule changed its conformation within a 200 ns MD run. Only Ade3
37 stayed in its original binding site on the surface of RNase L_E131 (Figure S5, Figure 2).

38
39 The partly unbound pA₂₅A₂₅A molecule adopted an extended conformation in which Ade2
40 was stacked with Ade1 (despite of their rearrangements), whereas Ade3 was separated from
41 them. Ade1 lost stacking contacts with the side chain of Phe126, whereas stacking of Ade3
42 with the side chain of Trp60 was preserved. Ade1 lost contacts with the side chain of non-
43 protonated Glu131. Similarly, a single hydrogen bond between Ade2 and the side chain of
44 Tyr135 was broken from the very beginning. In contrast, Ade3 was quite potentially fixed by a
45 hydrogen bond network involving the side chains of Gln68 and Asn65.

46
47 Phos1 formed salt bridges with side chains of Arg155 (at the beginning and end of the MD
48 trajectory) and Lys166 and Lys178 (in the meantime). The Tyr124 side chain again stabilized

1 distorted pA₂₅A₂₅A by binding to Phos2. Phos3 formed a salt bridge with the side chain of
2 Lys89.

3
4 Ade1 was in the **syn** -> **high anti** -> **syn** conformations, Ade2 was in the **anti** -> **high anti**
5 conformations, whereas Ade3 was in the **syn** conformation. Riboses were conformationally
6 promiscuous, nevertheless ended up in the **C4'-exo** conformation. The C4'-C5' moieties
7 adopted conformations **-gauche/trans** (Ade1), **-gauche** (Ade2), **-gauche** (Ade3). The
8 internucleotide linkage Ade1-Ade2 preferred conformation **-g-g** -> **t-g** -> **gt** -> **-g-g**. The
9 internucleotide linkage Ade2-Ade3 adopted conformation **gg**.

10
11 06) RNase L_E131 --- pA₂₅A₂₅A --- FF_Olo [50] --- T=290K --- t=200 ns

12
13 The pA₂₅A₂₅A molecule dissociated from ANK leaving only Ade3 in its original position on
14 the surface of RNase L_E131 (Figure S6, Figure 2).

15
16 The partly unbound pA₂₅A₂₅A molecule adopted an extended conformation in which all bases
17 were separated. Ade1 lost stacking contacts with the side chain of Phe126 (and stacked
18 instead with Tyr124), whereas stacking of Ade3 with the side chain of Trp60 was preserved.
19 Ade1 lost contacts with the side chain of non-protonated Glu131. Similarly, a single hydrogen
20 bond between Ade2 and the side chain of Tyr135 was broken from the very beginning. In
21 contrast, Ade3 was potentially fixed (with occasional interruptions) by a hydrogen bond network
22 involving the side chains of Gln68 and Asn65.

23
24 The side chain of Lys166 formed salt bridges with Phos1. There were again found contacts of
25 Phos2 with the Tyr124 side chain helping to destabilize pA₂₅A₂₅A. Phos3 formed a salt bridge
26 with the side chain of Lys89.

27
28 Ade1 was in the **high anti** -> **anti** conformations, Ade2 was in the **anti** -> **high anti**
29 conformations, whereas Ade3 was in the **syn** conformation. The first ribose ring had **C4'-exo**
30 puckering, the second ribose had **C3'-endo** puckering, whereas the third ribose had **C4'-exo**
31 puckering. The C4'-C5' moieties preferred conformations **-gauche** (Ade1-3). The
32 internucleotide linkage Ade1-Ade2 adopted conformation **-g-g**. The internucleotide linkage
33 Ade2-Ade3 preferred conformations **gg** -> **g-g**.

34
35 07) RNase L_E131H_R155A --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=290K --- t=200 ns

36
37 The pA₂₅A₂₅A molecule rearranged substantially within this MD run. Nevertheless, Ade3
38 (which was the most distant adenine ring from R155A) persisted in its original position on the
39 surface of RNaseL_E131H_R155A (Figure S7, Figure 3).

40
41 The unbound pA₂₅A₂₅A molecule adopted an extended conformation in which Ade2 was
42 stacked with Ade1 (despite of their rearrangements), whereas Ade3 was separated from them.
43 Ade1 lost stacking contacts with the side chain of Phe126, whereas stacking of Ade3 with the
44 side chain of Trp60 was preserved. Ade1 lost contacts with the side chain of protonated
45 Glu131. Similarly, a hydrogen bond between Ade2 and the side chain of Tyr135 was not
46 observed probably due to atypical interactions of Phos1 and Phos2 with amino acid side
47 chains (see below). In contrast, Ade3 was potentially fixed by a hydrogen bond network
48 involving the side chains of Gln68 and Asn65.

1 Remaining basic side chains of Lys166 and Lys178 formed alternately salt bridges with
2 Phos1. The Tyr124 side chain was steadily bound to Phos2 distorting pA₂₅A₂₅A. Phos3
3 formed a salt bridge with the side chain of Lys89.

4
5 Ade1 was in the high **anti** -> **syn** conformations, Ade2 was in the **anti** -> **high anti**
6 conformations, whereas Ade3 was in the **syn** conformation. All riboses ended up with the
7 **C4'-exo** puckering. The C4'-C5' moieties adopted conformations **trans** <-> **-gauche** (Ade1),
8 **-gauche** (Ade2), **-gauche** (Ade3). The internucleotide linkage Ade1-Ade2 adopted
9 conformation **-g-g**. The internucleotide linkage Ade2-Ade3 preferred conformation **gg**.

10
11 08) RNase L_E131H_R155A --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=310K --- t=200 ns

12
13 The pA₂₅A₂₅A molecule was steadily bound to the surface of RNaseL_E131H_R155A
14 (Figure S8, Figure 3).

15
16 The bound pA₂₅A₂₅A molecule adopted an extended conformation in which Ade2 was stacked
17 with Ade1, whereas Ade3 was separated from them. Ade1 was stacked with the side chain of
18 Phe126. Ade3 was stacked with the side chain of Trp60. Ade1 was fixed by hydrogen bonds
19 with the side chain of protonated Glu131 throughout the whole MD simulation lasting for 200
20 ns. Ade2 was fixed by a single hydrogen bond with the side chain of Tyr135 (in contrast to
21 previous MD no. 7). Ade3 was fixed by a hydrogen bond network involving the side chains of
22 Gln68 and Asn65.

23
24 The side chains of Lys166 and Lys178 formed salt bridges with Phos1. There were not found
25 significant interactions of Phos2 with the Tyr124 side chain (in contrast to previous MD no.
26 7). Phos3 formed a salt bridge with the side chain of Lys89 and a hydrogen bond with the side
27 chain of Trp60.

28
29 Ade1 and Ade2 were in the **high anti** conformation, whereas Ade3 was in the **syn**
30 conformation. The first ribose ring had changing puckering, the second ribose had **C3'-endo**
31 puckering, whereas the third ribose had **C4'-exo** puckering. The C4'-C5' moieties adopted
32 conformations **trans** -> **gauche** (Ade1), **-gauche** (Ade2), **-gauche** (Ade3). The
33 internucleotide linkage Ade1-Ade2 adopted conformation **-g-g**. The internucleotide linkage
34 Ade2-Ade3 preferred conformation **gg**.

35
36 09) RNase L_E131H_F126A --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=310K --- t=200 ns

37
38 The pA₂₅A₂₅A molecule dissociated from the surface of RNaseL_E131H_F126A with only
39 Ade3 (which was the most distant adenine ring from F126A) staying bound in its original
40 position (Figure S9, Figure 3).

41
42 The unbound pA₂₅A₂₅A molecule adopted an extended conformation in which Ade2 was
43 stacked with Ade1 (despite of their rearrangements), whereas Ade3 was separated from them.
44 Ade1 lost stacking contacts with the side chain of Phe126, whereas stacking of Ade3 with the
45 side chain of Trp60 was preserved. Ade1 lost contacts with the side chain of protonated
46 Glu131. Similarly, a single hydrogen bond between Ade2 and the side chain of Tyr135 was
47 subsequently broken. In contrast, Ade3 was potently fixed by a hydrogen bond network
48 involving the side chains of Gln68 and Asn65.

1 The side chains of Arg155, Lys166 and Lys178 lost salt bridging contacts with Phos1. There
2 were observed occasional contacts of Phos2 with the Tyr124 side chain stabilizing distorted
3 pA₂₅A₂₅A. Phos3 formed a salt bridge with the side chain of Lys89.

4
5 Ade1 was in the **high anti** -> **syn** -> **high anti** conformations, Ade2 was in the **anti** -> **high**
6 **anti** -> **anti** conformations, whereas Ade3 was in the **syn** conformation. The first ribose ring
7 had **C3'-endo** -> **C4'-exo** -> **O4'-endo** -> **C2'-endo** puckering, the second ribose had **C3'-**
8 **endo** -> **C4'-endo** -> **C3'-endo** puckering, whereas the third ribose had **C4'-exo** puckering.
9 The C4'-C5' moieties adopted conformations **trans** <-> **-gauche** (Ade1), **-gauche** (Ade2), **-**
10 **gauche** -> **gauche** (Ade3). The internucleotide linkage Ade1-Ade2 adopted conformation **-g-**
11 **g**. The internucleotide linkage Ade2-Ade3 preferred conformations **gg** -> **-g-g**.

12
13 10) RNase L_E131H_W60A --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=310K --- t=200 ns

14
15 The pA₂₅A₂₅A molecule rearranged substantially with only Ade1 (which was the most distant
16 adenine ring from W60A) staying in its original binding site on the surface of
17 RNaseL_E131H_W60A (Figure S10, Figure 3).

18
19 The unbound pA₂₅A₂₅A molecule adopted a helical-like conformation in which Ade1-Ade2-
20 Ade3 were stacked (the same was observed in the subsequent MD no. 11 simulation). Ade1
21 stacked with the side chain of Phe126, whereas Ade3 lacked its obligate stacking partner
22 Trp60. Ade1 interacted with the side chain of protonated Glu131. Tyr135 was hydrogen
23 bonded toward Ade3 (instead of Ade2). A hydrogen bond network involving Ade3 and the
24 side chains of Gln68 and Asn65 was disrupted.

25
26 Phos1 lost contacts with Arg155 and was bound to Lys166. There were no significant
27 interactions of Phos2 with the Tyr124 side chain. Phos3 lost and re-established again a salt
28 bridge with the side chain of Lys89.

29
30 All adenine rings (Ade1-3) ended up in the **anti** conformation. The first and second ribose
31 ring had **C3'-endo/C4'-exo** puckering, whereas the third ribose had **C2'-endo** puckering. The
32 C4'-C5' moieties adopted conformations **-gauche** (Ade1-3). Both internucleotide linkages
33 (Ade1-Ade2 and Ade2-Ade3) preferred conformation **-g-g**.

34
35 11) RNase L_E131H_K89A --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=290K --- t=200 ns

36
37 The pA₂₅A₂₅A molecule was steadily bound for ~150 ns. Then it dissociated leaving only
38 Ade1 (which was the most distant adenine ring from the K89A) in its original position on the
39 surface of RNaseL_E131H_K89A (Figure S11, Figure 3).

40
41 The unbound pA₂₅A₂₅A molecule adopted a helical-like conformation in which Ade1-Ade2-
42 Ade3 were stacked. Moreover, Ade1 stacked with the side chain of Phe126, whereas stacking
43 of Ade3 with the side chain of Trp60 was lost. Ade1 was fixed by hydrogen bonds with the
44 side chain of protonated Glu131 throughout the whole MD simulation lasting for 200 ns. The
45 usual hydrogen bond between Ade2 and the side chain of Tyr135 was broken. The hydrogen
46 bond network involving Ade3 and the side chains of Gln68 and Asn65 disappeared.

47
48 Arg155 formed salt bridges with Phos1. There were not found significant contacts of Phos2
49 with the Tyr124 side chain.

1
2 Ade1 was in the **high anti** conformation, Ade2 was in the **anti** conformation, whereas Ade3
3 was in the **syn** -> **anti** conformations. The first ribose ring had **C3'-endo** puckering, the
4 second ribose had **C3'-endo** -> **C2'-endo** -> **C3'-endo** puckering, whereas the third ribose
5 had **C4'-exo** -> **C2'-endo** puckering. The C4'-C5' moieties adopted conformations **-gauche**
6 (Ade1), **-gauche** (Ade2), **-gauche** -> **trans** -> **-gauche** (Ade3). The internucleotide linkage
7 Ade1-Ade2 preferred conformation **-g-g**. The internucleotide linkage Ade2-Ade3 adopted
8 conformations **gg** -> **t-g** -> **gg** -> **-g-g**.

9
10 *12) RNase L_E131H_K89A --- pA₂₅A₂₅A --- FF_99SB [47-49] --- T=310K --- t=200 ns*

11
12 The pA₂₅A₂₅A molecule dissociated from the surface of RNaseL_E131H_K89A (Figure S12,
13 Figure 3). Finally, pA₂₅A₂₅A was directed perpendicularly with respect to its original
14 orientation.

15
16 The unbound pA₂₅A₂₅A molecule adopted an extended conformation in which Ade2 was
17 stacked with Ade1, whereas Ade3 was separated from them. Ade1 lost stacking contacts with
18 the side chain of Phe126 and Ade3 lost stacking interactions with the side chain of Trp60.
19 Further, Ade1 lost contacts with the side chain of protonated Glu131 and the hydrogen bond
20 between Ade2 and the side chain of Tyr135 was broken as well. The hydrogen bond network
21 involving Ade3 and the side chains of Gln68 and Asn65 disappeared as well.

22
23 The side chain of Arg155 replaced by Lys178 formed salt bridges with Phos1. Again, there
24 were found contacts of Phos2 with the Tyr124 side chain coinciding with dissociation of
25 pA₂₅A₂₅A.

26
27 Ade1 was in the **high anti** -> **syn** -> **high anti** conformations, Ade2 was in the **high anti**
28 conformation, whereas Ade3 was in the **syn** -> **high anti** -> **syn** -> **anti** conformation. The
29 first ribose ring had **C3'-endo/C4'-exo** puckering, whereas second and third ribose had **C2'-**
30 **endo** puckering. The C4'-C5' moieties adopted conformations **-gauche** -> **trans** -> **-gauche**
31 (Ade1), **-gauche** (Ade2), **-gauche** (Ade3). The internucleotide linkage Ade1-Ade2 adopted
32 conformation **-g-g**. The internucleotide linkage Ade2-Ade3 preferred conformations **gg** -> **-g-**
33 **g** -> **tg** -> **-gt**.

34
35 *13) RNase L_E131H --- pA₂₅A₂₅B --- FF_99SB [47-49] -- T=310K --- t=200ns*

36
37 The pA₂₅A₂₅B molecule was steadily bound to the surface of RNaseL_E131H (Figure S13,
38 Figure 4). It's in agreement with remarkable ability of pA₂₅A₂₅B to activate RNase L (EC50 =
39 4.9 nM; Crel 1.4x [24] - where EC50 was defined as the molarity (M) of 2-5A analogues
40 required to cause a 50% degradation of 5S rRNA; Crel represents the ratio of EC50 of
41 analogues as compared to EC50 = 3.5 nM of natural 2-5A; i.e. the smaller Crel the more
42 effectively the analogue was able to activate RNase L).

43
44 The bound pA₂₅A₂₅B molecule adopted an extended conformation in which Ade2 was stacked
45 with Ade1, whereas Ade3 was separated from them. Ade1 was stacked with the side chain of
46 Phe126. Ade3 was stacked with the side chain of Trp60. Ade1 was fixed by hydrogen bonds
47 with the side chain of protonated Glu131 throughout the whole MD simulation lasting for 200
48 ns. Ade2 was fixed by a single hydrogen bond with the side chain of Tyr135. Ade3 was fixed
49 by a hydrogen bond network involving the side chains of Gln68 and Asn65.

1
2 The side chains of Arg155 and Lys166 formed salt bridges with Phos1. There were not found
3 significant contacts of Phos2 with the Tyr124 side chain. Phos3 formed a salt bridge with the
4 side chain of Lys89.

5
6 Ade1 was in the high **anti** conformation, Ade2 was in the **anti** conformation, whereas Ade3
7 was in the **syn** conformation. The first and second ribose ring had **C3'-endo** puckering,
8 whereas the third ribose had **C2'-endo** puckering. The C4'-C5' moieties adopted
9 conformations **-gauche** <-> **trans** (Ade1), **-gauche** (Ade2), **gauche** (Ade3). The
10 internucleotide linkage Ade1-Ade2 adopted conformation **-g-g**. The internucleotide linkage
11 Ade2-Ade3 preferred conformation **g-g**.

12
13 *14) RNase L_E131H --- pB₂₅A₂₅A --- FF_99SB [47-49] --- T=310K --- t=200 ns*

14
15 The pB₂₅A₂₅A molecule was steadily bound to the surface of RNaseL_E131H (Figure S14,
16 Figure 4). It seems to be in accordance with remarkable ability of pB₂₅A₂₅A to activate RNase
17 L (EC50 = 77 nM; Crel 22x [24]).

18
19 The bound pB₂₅A₂₅A molecule adopted an extended conformation in which Ade2 was stacked
20 with Ade1, whereas Ade3 was separated from them. Ade1 was stacked with the side chain of
21 Phe126. Ade3 was stacked with the side chain of Trp60. Ade1 was fixed by hydrogen bonds
22 with the side chain of protonated Glu131 throughout the whole MD run lasting for 200 ns.
23 Ade2 was fixed by a single hydrogen bond with the side chain of Tyr135. Ade3 was fixed by
24 a hydrogen bond network involving the side chains of Gln68 and Asn65.

25
26 The side chains of Arg155 and Lys166 formed salt bridges with Phos1. There were not found
27 contacts of Phos2 with the Tyr124 side chain. Phos3 formed a salt bridge with the side chain
28 of Lys89.

29
30 Ade1 was in the **high anti** conformation, Ade2 was in the **anti** conformation, whereas Ade3
31 was in the **syn** conformation. The first ribose ring had **C1'-endo** puckering, the second ribose
32 had **C3'-endo** puckering, whereas the third ribose had **C4'-exo** puckering. The C4'-C5'
33 moieties adopted conformations **trans** (Ade1), **-gauche** (Ade2), **-gauche** <-> **gauche** (Ade3).
34 The internucleotide linkage Ade1-Ade2 adopted conformation **-g-g**. The internucleotide
35 linkage Ade2-Ade3 preferred conformations **gg** <-> **g-g**.

36
37 *15) RNase L_E131H --- pB₂₅A₂₅B --- FF_99SB [47-49] --- T=310K --- t=200ns*

38
39 The pB₂₅A₂₅B molecule stayed close to its original position on the surface of RNaseL_E131H
40 (Figure S15, Figure 4). It seems to be in agreement with fairly well ability of pB₂₅A₂₅B to
41 activate RNase L (EC50 > 130 nM; Crel 37x [24])

42
43 The bound pB₂₅A₂₅B adopted an extended conformation in which Ade2 was stacked with
44 Ade1, whereas Ade3 was separated from them. Ade1 was stacked with the side chain of
45 Phe126. Ade3 was stacked with the side chain of Trp60. Ade1 was fixed by hydrogen bonds
46 with the side chain of protonated Glu131 throughout the entire MD simulation lasting for 200
47 ns. Ade2 was fixed by a single hydrogen bond with the side chain of Tyr135. Ade3 was fixed
48 by a hydrogen bond network involving the side chains of Gln68 and Asn65.

1 The side chains of Arg155 and (to some extent) Lys166 formed salt bridges with Phos1.
2 There were almost no contacts of Phos2 with the Tyr124 side chain. Phos3 formed a salt
3 bridge with the side chain of Lys89.

4
5 Ade1 was in the **high anti** conformation, Ade2 was in the **anti** conformation, whereas Ade3
6 was in the **syn** conformation. The first ribose ring had **C1'-endo** puckering, the second ribose
7 had **C3'-endo** puckering, whereas the third ribose transitioned from **C1'-endo** to **C2'-endo**
8 puckering. The C4'-C5' moieties adopted conformations **trans** (Ade1), **-gauche** (Ade2),
9 **trans** -> **-gauche** -> **gauche** (Ade3). The internucleotide linkage Ade1-Ade2 adopted
10 conformation **-g-g**. The internucleotide linkage Ade2-Ade3 preferred conformation **g-g** (the
11 same was observed for pA₂₅A₂₅B).

12
13 *16) RNase L_E131H --- pA₂₅B₂₅A --- FF_99SB [47-49] --- T=310K --- t=100ns*

14
15 The pA₂₅B₂₅A molecule was stable for quite some time, nevertheless at the very end of this
16 MD run it rearranged substantially with just Ade3 persisting in its original binding site on the
17 surface of RNaseL_E131H (Figure S16, Figure 4). It seems to be in accord with poor ability
18 of pA₂₅B₂₅A to activate RNase L (EC₅₀ > 530 nM; Crel 150x [24]).

19
20 The partly unbound pA₂₅B₂₅A molecule adopted an extended conformation in which Ade2
21 was stacked with Ade1 (regardless of all their rearrangements), whereas Ade3 was separated
22 from them. Ade1 lost stacking contacts with the side chain of Phe126, whereas stacking of
23 Ade3 with the side chain of Trp60 was well preserved. Ade1 lost contacts with the side chain
24 of protonated Glu131. Similarly, a single hydrogen bond between Ade2 and the side chain of
25 Tyr135 was broken. In contrast, Ade3 was potently fixed by a hydrogen bond network
26 involving the side chains of Gln68 and Asn65.

27
28 Phos1 formed salt bridges with the side chains of Arg155 and to some extent Lys166. There
29 were not found contacts of Phos2 with the Tyr124 side chain. Phos3 formed a salt bridge with
30 the side chain of Lys89.

31
32 Ade1 and Ade2 finished in the **high anti** conformation, whereas Ade3 was in the **syn**
33 conformation. Ribose rings underwent conformational transitions in the moment of
34 dissociation: **C3'-endo** -> **C2'-endo** (Ade1), **C3'-endo** -> **C2'-endo** (Ade2), **C4'-exo** (Ade3).
35 The C4'-C5' moieties adopted conformations **-gauche** <-> **trans** (Ade1), **-gauche** <->
36 **gauche** (Ade2), **-gauche** <-> **gauche** (Ade3). The internucleotide linkage Ade1-Ade2 adopted
37 gradually conformations **-g-g** -> **tg**. The internucleotide linkage Ade2-Ade3 preferred
38 conformation **gg**.

39
40 *17) RNase L_E131H --- pA₂₅B₂₅B --- FF_99SB [47-49] --- T=310K --- t=100 ns*

41
42 The pA₂₅B₂₅B molecule was bound to the surface of RNaseL_E131H. Ade1 was stacked with
43 the side chain of Phe126. Ade3 was stacked with the side chain of Trp60. Nevertheless, the
44 second AMP significantly changed its conformation (Figure S17, Figure 4). As a
45 consequence, all bases were separated from each other. It's in agreement with an apparent
46 inability of pA₂₅B₂₅B to activate RNase L (EC₅₀ > 1000 nM; Crel >300x [24]).

1 Ade1 was fixed by hydrogen bonds with the side chain of protonated Glu131 throughout the
2 whole MD run lasting for 100 ns. Ade2 lost any contacts with the side chain of Tyr135. Ade3
3 was fixed by a hydrogen bond network involving the side chains of Gln68 and Asn65.

4
5 Phos1 formed salt bridges with side chains of Arg155 and Lys166. Phos2 wasn't disturbed by
6 the Tyr124 side chain. Phos3 formed a salt bridge with the side chain of Lys89 regardless of
7 conformational transitions of pA₂₅B₂₅B - see below.

8
9 Ade1 and Ade2 were in the **high anti** conformation, whereas Ade3 was in the **syn**
10 conformation. All ribose rings preferred the **C2'-endo** puckering. The C4'-C5' moieties
11 adopted conformations **-gauche** <-> **trans** (Ade1), **trans** -> **gauche** (Ade2), **gauche** (Ade3).
12 The internucleotide linkage Ade1-Ade2 adopted conformations **-gg** -> **-gt** -> **-g-g**. The
13 internucleotide linkage Ade2-Ade3 preferred conformations **g-g** -> **-gg**.

14
15 18) RNase L_E131H --- pB₂₅B₂₅A --- FF_99SB [47-49] --- T=310K --- t=100ns

16
17 The pB₂₅B₂₅A molecule changed substantially its conformation - just Ade1 stayed partly
18 bound in its original position on the surface of RNase L_E131H (Figure S18, Figure 4). It's in
19 accordance with inability of pB₂₅B₂₅A to activate RNase L (EC50 > 1000 nM; Crel >300x
20 [24]).

21
22 The unbound pB₂₅B₂₅A 2-5A molecule adopted an extended conformation in which Ade2 was
23 stacked with Ade1 (despite of their rearrangements), whereas Ade3 was separated from them.
24 Ade1 lost partly stacking contacts with the side chain of Phe126, whereas stacking of Ade3
25 with the side chain of Trp60 was preserved. Ade1 lost one of two contacts with the side chain
26 of Glu131. Similarly, a single hydrogen bond between Ade2 and the side chain of Tyr135 was
27 broken. A hydrogen bond network between Ade3 and the side chains of Gln68 and Asn65
28 was broken.

29
30 The side chains of Arg155 and Lys166 formed salt bridges with Phos1. There were not found
31 contacts of Phos2 with the Tyr124 side chain. Phos3 lost contacts with Lys89.

32
33 Ade1 was in the **high anti** conformation, Ade2 was in the **syn** conformation, whereas Ade3
34 was in the **anti** conformation. The first and second ribose ring had **C2'-endo** puckering,
35 whereas the third ribose had **C4'-exo** puckering. The C4'-C5' moieties adopted
36 conformations **trans** (Ade1), **gauche** (Ade2), **-gauche** (Ade3). The internucleotide linkage
37 Ade1-Ade2 adopted conformation **-g-g** -> **-gt**. The internucleotide linkage Ade2-Ade3
38 preferred conformation **-g-g**.

39
40 19) RNase L_E131H --- pB₂₅B₂₅B --- FF_99SB [47-49] --- T=310K --- t=100 ns

41
42 The pB₂₅B₂₅B 2-5A molecule was remarkably rearranged within this MD run. Finally, it was
43 bound to the surface of RNaseL_E131H only through Ade3 (Figure S19, Figure 4). It's in
44 agreement with inability of pB₂₅B₂₅B to activate RNase L (EC50 > 1000 nM; Crel >300x
45 [24]).

46
47 The almost unbound pB₂₅B₂₅B molecule adopted an extended conformation in which Ade2
48 was stacked with Ade1 (regardless of all the rearrangements of both AMPs), whereas Ade3
49 was separated from them. Ade1 lost stacking contacts with the side chain of Phe126. Ade3

1 stacked with the side chain of Trp60, although this base overturned as already observed in the
2 first MD simulation (see above). Ade1 lost contacts with the side chain of protonated Glu131
3 and Ade3 lost contacts with Gln68 and Asn65. A hydrogen bond between Ade2 and the side
4 chain of Tyr135 was created only for a moment, when pB₂₅B₂₅B dissociated from the surface
5 of RNase L_E131H.

6
7 Phos1 formed salt bridges with the side chains of Arg155 and Lys166. There were no contacts
8 of Phos2 with the Tyr124 side chain. Phos3 formed a salt bridge with the side chain of Lys89.

9
10 All (i.e. Ade1, Ade2, Ade3) glycosidic torsion angles finished in the **high anti** conformation.
11 The first/second/third ribose ring was in the **C1'-endo/C2'-endo/C2'-endo** conformation.
12 The C4'-C5' moieties adopted conformations **trans** (Ade1), **-gauche** (Ade2), **-gauche**
13 (Ade3). The internucleotide linkage Ade1-Ade2 adopted conformation **-g-g**. The
14 internucleotide linkage Ade2-Ade3 preferred conformation **gg**.

15 16 **DISCUSSION and CONCLUSIONS**

17
18 The model systems of monomeric human RNase L complexed with pA₂₅A₂₅A or its analogs
19 were constructed using the crystal structure [8] (PDB id 1wdy). In principle, dimeric structure
20 of RNase L [18] (PDB id 4g8k, 4g8l), where two molecules of 2-5A at a time tether RNase L
21 monomers via their ANK domains, could be used for MD simulations. Nevertheless, the
22 sensor domain undergoes 2-5A-mediated self-association without conformational changes
23 [18]. Each ANK domain harbors two distinct sites for 2-5A recognition that reside 50 Å apart.
24 Nevertheless, 2-5A is bound to the primary binding site (with a variety of interacting amino
25 acids side chains - see Figure 1) in the same way in the monomeric as well as dimeric RNase
26 L. Binding of second monomer of RNase L to 2-5A involves just interactions of Tyr312 with
27 internucleotide linkage Ade2-Ade3. It means that if we used dimeric RNase L in MD
28 simulations, we did not gain any added value. It would decrease the chances that the expected
29 conformational transitions will be captured on the time scale of a few hundred nanoseconds.
30 Moreover, MD simulations with larger systems would be more time consuming (and we had
31 moved in this matter already to the edge of endurance). Nevertheless, in MD simulations with
32 monomeric RNase L, it was necessary to pay attention to residues Tyr135 and Ade2. Even
33 relatively subtle changes in their positions could obstruct a possible binding of Tyr312 from
34 the second monomer of RNase L to 2-5A with fatal consequences for its endoribonuclease
35 activity.

36
37 Several versions of the AMBER force field [46, 47-49, 50] were compared in this study (MD
38 simulations no. 1-6 - see Table S1). The AMBER force field for nucleic acids passes through
39 repeated revisions. It's caused by complexity of nucleic acids comparing to protein backbones.
40 It's so problematic to balance all force-field constants, such as those that may affect
41 conformational preferences of the glycosidic torsion angle, which can adopt **anti**, **high-anti**
42 and **syn** conformations. Interestingly, trimeric 2-5A carries all these conformers. Since 2-5A
43 in MD simulation no. 2 remained closest to what was observed in the crystal structure, we
44 used the AMBER force field v. [47-49] in all other MD simulations with mutant RNase L
45 (MD no. 7-12 - see Table S1) or 2-5A analogs (MD no. 13-19 - see Table S1).

46
47 Further, RNase L with either protonated or unprotonated Glu131 was tested in this subset of
48 MD simulations (i.e. MD simulations no. 1-3 vs. 4-6 - see Table S1). Based on the distance
49 between atoms of Glu131 and Ade1 found in the crystal structure it seemed likely that Glu131

1 must be protonated creating two hydrogen bonds with Ade1. MD simulations indeed
2 confirmed that non-protonated Glu131 leads to dissociation of 2-5A from RNase L (MD no.
3 4-6). In contrast, in MD simulations with protonated Glu131, the RNase L / 2-5A complex
4 was obviously stable (MD no. 1-3). We therefore used protonated Glu131 in all MD
5 simulations with mutant RNase L (MD no. 7-12 - see [Table S1](#)) or chemically modified 2-5A
6 (MD no. 13-19).

7
8 Further, MD simulations were used to capture detrimental impacts of various point mutations
9 of RNase L (R155A - MD no. 7-8; F126A - MD no. 9; W60A - MD no. 10; K89A - MD no.
10 11-12 - see [Table S1](#)) on 2-5A binding. These conformational transitions were captured on a
11 time scale of 200 ns. Only for R155A it happened just in one of two MD trajectories. It was
12 due to the fact that the loss of interactions of the terminal phosphate group of 2-5A with
13 Arg155 may be temporary and to a certain extent compensated by interactions with flexible
14 basic side chains of Lys166 and Lys178. When comparing both R155A MD simulations, it
15 proved decisive whether Ade2 was or was not bound to Tyr135 in the beginning of a MD
16 simulation. This interaction can be disrupted by contacts of Tyr124 with Phos2. Tyr124
17 showed as stabilizer of distorted structures of 2-5A in some other MD simulations as well (for
18 example MD no. 4-6). It should be taken into account when designing chemically modified 2-
19 5A agonists of RNase L.

20
21 Finally, 2-5A analogues with a bridged 3'-O,4'-C-alkylene linkage [\[24\]](#) introduced into the
22 adenosine units were used to assess ability of MD simulations to distinguish on the time scale
23 of hundreds of nanoseconds between agonists of RNase L (pA₂₅A₂₅B, pB₂₅A₂₅A, pB₂₅A₂₅B -
24 MD no. 13-15 - see [Table S1](#)) and inactive analogs (pA₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅A,
25 pB₂₅B₂₅B - MD no. 16-19 - see [Table S1](#)). And indeed it did. Agonists were potently bound to
26 RNase L during 200 ns MD runs. Just Ade2 sometimes oscillated between two positions with
27 respect to interactions with Tyr 135. In doing so, the usual arrangement prevailed. Taking into
28 account the number of these transitions, it would require multimicrosecond MD trajectories to
29 assess whether there is any correlation with varying degrees of activity (0-150 nM) of these
30 agonists of RNase L. For inactive 2-5A analogs, by contrast, there were observed significant
31 disruptions of their interactions with RNase L already during 100 ns MD runs.

32
33 Summarized, produced MD simulations met our expectations consider possible use of this
34 methodology as a tool for a rational design of RNase L agonists. With that as the number of
35 cores in high-end CUDA-programmable GPUs increases, we will soon be able to routinely
36 carry out multimicrosecond MD simulations and to capture tiny effects of subtle
37 modifications of 2-5A.

38
39 There is plenty of experimental data waiting for interpretation in literature. Thus far, a number
40 of 2-5A analogues with modification of bases, riboses, internucleotide linkages and a 5'-
41 phosphoryl group have been synthesized to test RNase L activating ability in vitro [\[24-44\]](#). In
42 addition, there is inspiration in the form of hundreds or thousands of modifications that have
43 been studied in the context of the so-called antisense or siRNA oligonucleotides [\[65\]](#).

1 **REFERENCES**

- 2
- 3 01. I. M. Kerr, R. E. Brown,
4 **pppA2'p5'A2'p5'A: an inhibitor of protein synthesis synthesized with an enzyme**
5 **fraction from interferon-treated cells.**
6 Proceedings of the National Academy of Sciences of the USA 75 (1978) 256-260.
7
- 8 02. M. R. Player and P. F. Torrence,
9 **The 2-5A System: Modulation of Viral and Cellular Processes Through Acceleration of**
10 **RNA Degradation.**
11 Pharmacology & Therapeutics 78 (1998) 55-114.
12
- 13 03. R. H. Silverman,
14 **A scientific journey through the 2-5A/RNase L system.**
15 Cytokine & Growth Factor Reviews 18 (2007) 381-388.
16
- 17 04. R. H. Silverman,
18 **Viral encounters with OAS and RNase L during the IFN antiviral response.**
19 Journal of Virology 81 (2007) 12720-12729.
20
- 21 05. A. Chakrabarti, B. K. Jha, and R. H. Silverman,
22 **New insights into the role of RNase L in innate immunity.**
23 Journal of Interferon and Cytokine Research 31 (2011) 49-57.
24
- 25 06. X. L. Li, H. J. Ezelle, T. J. Kang, L. Zhang, K. A. Shirey, J. Harro, J. D. Hasday, S. K.
26 Mohapatra, O. R. Crasta, S. N. Vogel, A. S. Cross, B. A. Hassel,
27 **An essential role for the antiviral endoribonuclease, RNase-L, in antibacterial immunity.**
28 Proceedings of the National Academy of Sciences of the USA 105 (2008) 20816-20821.
29
- 30 07. B. Dong, L. Xu, A. Zhou, B. A. Hassel, X. Lee, P. F. Torrence, and R. H. Silverman,
31 **Intrinsic Molecular Activities of the Interferon-induced 2-5A-dependent RNase L.**
32 The Journal of Biological Chemistry 269 (1994) 14153-14158.
33
- 34 08. N. Tanaka, M. Nakanishi, Y. Kusabake, Y. Goto, Y. Kitade, K. T. Nakamura,
35 **Structural basis for recognition of 2',5'-linked oligoadenylates by human ribonuclease**
36 **L.**
37 The EMBO Journal 23 (2004) 3929-3938.
38
- 39 09. S. Chevaliez, J. M. Pawlotsky,
40 **Interferon-based therapy of hepatitis C.**
41 Advanced Drug Delivery Reviews 59 (2007) 1222-1241.
42
- 43 10. C. L. Washenberger, J. Q. Han, K. J. Kechris, B. K. Jha, R. H. Silverman, D. J. Barton,
44 **Hepatitis C virus RNA: Dinucleotide frequencies and cleavage by RNase L.**
45 Virus Research 130 (2007) 85-95.
46
- 47 11. K. Malathi, T. Saito, N. Crochet, D. J. Barton, M. Gale, R. H. Silverman,
48 **RNase L releases a small RNA from HCV RNA that refolds into a potent PAMP.**
49 RNA 16 (2010) 2108-19.

- 1
2 12. K. Malathi, B. Dong, M. Gale, Jr., and R. H. Silverman,
3 **Small self-RNA generated by RNase L amplifies antiviral innate immunity.**
4 Nature 448 (2007) 816–819.
5
- 6 13. M. Nakanishi, N. Tanaka, Y. Mizutani, M. Mochizuki, Y. Ueno, K. T. Nakamura, Y.
7 Kitade,
8 **Functional Characterization of 2',5'-Linked Oligoadenylate Binding Determinant of**
9 **Human RNase L.**
10 The Journal of Biological Chemistry 280 (2005) 41694-41699.
11
- 12 14. R. Gaudet,
13 **A primer on ankyrin repeat function in TRP channels and beyond.**
14 Molecular BioSystems 4 (2008) 372-379.
15
- 16 15. P. V. Lishko, E. Procko, X. Jin, C. B. Phelps, R. Gaudet,
17 **The ankyrin repeats of TRPV1 bind multiple ligands and modulate channel sensitivity.**
18 Neuron 54 (2007) 905–918.
19
- 20 16. C. B. Phelps, R. R. Wang, S. S. Choo, R. Gaudet,
21 **Differential regulation of TRPV1, TRPV3, and TRPV4 sensitivity through a conserved**
22 **binding site on the ankyrin repeat domain.**
23 The Journal of Biological Chemistry 285 (2010) 731–740.
24
- 25 17. S. Y. Lau, E. Procko, and R. Gaudet,
26 **Distinct properties of Ca²⁺-calmodulin binding to N- and C- terminal regulatory**
27 **regions of the TRPV1 channel.**
28 The Journal of General Physiology 140 (2012) 541-555.
29
- 30 18. Y. Han, G. Whitney, J. Donovan, and A. Korennykh,
31 **Innate Immune Messenger 2-5A Tethers Human RNase L into Active High-Order**
32 **Complexes.**
33 Cell Reports 2 (2012) 902-913.
34
- 35 19. B. Bayard, L. D. Leserman, C. Bisbal, and B. Lebleu,
36 **Antiviral activity in L1210 cells of liposome-encapsulated (2'-5')oligo(adenylate)**
37 **analogues.**
38 European Journal of Biochemistry 151 (1985) 319-325.
39
- 40 20. L. Zhou, C. S. Thakur, R. J. Molinaro, J. M. Paranjape, R. Hoppes, K. T. Jeang, R. H.
41 Silverman and P. F. Torrence,
42 **Delivery of 2-5A cargo into living cells using the Tat cell penetrating peptide: 2-5A-tat.**
43 Bioorganic & Medicinal Chemistry 14 (2006) 7862-7874.
44
- 45 21. E. Kiuru, M. Ora, L. Beigelman, L. Blatt, H. Lonnberg,
46 **Synthesis and Enzymatic Deprotection of Fully Protected 2'-5' Oligoadenylates (2-5):**
47 **Towards a Prodrug Strategy for Short 2-5A.**
48 Chemistry & Biodiversity 9 (2012) 669-688.
49

- 1 22. C. S. Thakur, B. K. Jha, B. Dong, J. D. Gupta, K. M. Silverman, H. Mao, H. Sawai, A. O.
2 Nakamura, A. K. Banerjee, A. Gudkov, and R. H. Silverman,
3 **Small-molecule activators of RNase L with broad-spectrum antiviral activity.**
4 Proceedings of the National Academy of Sciences of the USA 104 (2007) 9585-9590.
5
- 6 23. J. G. van den Boorn, J. Dasler, C. Coch, M. Schlee, G. Hartmann,
7 **Exosomes as nucleic acid nanocarriers.**
8 Advanced Drug Delivery Reviews 65 (2013), 331-335.
9
- 10 24. K. Morita, M. Kaneko, S. Obika, T. Imanishi, Y. Kitade, M. Koizumi,
11 **Biologically Stable 2-5A Analogues containing 3'-O,4'-C-bridged Adenosine as Potent**
12 **RNase L Agonists.**
13 ChemMedChem 2 (2007) 1703-1707.
14
- 15 25. C. L. Hersh, R. E. Brown, W. K. Roberts, E. A. Swyryd, I. M. Kerr, and G. R. Stark,
16 **Simian Virus 40-infected, Interferon-treated Cells Contain 2',5'-Oligoadenylates Which**
17 **Do Not Activate Cleavage of RNA.**
18 The Journal of Biological Chemistry 259 (1984) 1731-1737.
19
- 20 26. J. Imai, M. I. Johnston, and P. F. Torrence,
21 **Chemical Modification Potentiates the Biological Activities of 2-5A and Its Congeners.**
22 The Journal of Biological Chemistry 257 (1982) 12739-12745.
23
- 24 27. M. C. Haugh, P. J. Cayley, H. T. Serafinowska, D. G. Norman, C. B. Reese, and I. M.
25 Kerr,
26 **Analogues and Analogue Inhibitors of ppp(A2'p)nA.**
27 European Journal of Biochemistry 132 (1983) 77-84.
28
- 29 28. K. Lesiak, J. Imai, G. Floyd-Smith, and P. F. Torrence,
30 **Biological Activities of Phosphodiester Linkage Isomers of 2-5A.**
31 The Journal of Biological Chemistry 258 (1983) 13082-13088.
32
- 33 29. H. Sawai, J. Imai, K. Lesiak, M. I. Johnston, and P. F. Torrence,
34 **Cordycepin Analogues of 2-5A and Its Derivatives.**
35 The Journal of Biological Chemistry 258 (1983) 1671-1677.
36
- 37 30. B. Bayard, C. Bisbal, M. Silhol, J. Cnockaert, G. Huez, and B. Lebleu,
38 **Increased stability and antiviral activity of 2'-O-phosphoglyceryl derivatives of (2'-**
39 **5')oligo(adenylate).**
40 European Journal of Biochemistry 142 (1984) 291-298.
41
- 42 31. P. J. Cayley, J. A. Davies, K. G. McCullagh, and I. M. Kerr,
43 **Activation of the ppp(A2'p)nA system in interferon-treated, Herpes simplex virus-**
44 **infected cells and evidence for novel inhibitors of the ppp(A2'p)nA-dependent RNase.**
45 European Journal of Biochemistry 143 (1984) 165-174.
46
- 47 32. J. Imai and P. F. Torrence,
48 **Synthesis and Biological Activity of 5'-Capped Derivatives of 5'-Triphosphoadenylyl(2'-**
49 **5')adenylyl(2'-5')adenosine.**

- 1 Biochemistry 23 (1984) 766-774.
2
- 3 33. P. F. Torrence, J. Imai, K. Lesiak, J. C. Jamouille, and H. Sawai,
4 **Oligonucleotide Structural Parameters That Influence Binding of 5'-O-**
5 **Triphosphoadenylyl-(2'-5')-adenylyl-(2'-5')-adenosine to the 5'-O-Triphosphoadenylyl-**
6 **(2'-5')-adenylyl-(2'-5')-adenosine Dependent Endoribonuclease: Chain Length,**
7 **Phosphorylation State, and Heterocyclic Base.**
8 Journal of Medicinal Chemistry 27 (1984) 726-733.
9
- 10 34. D. A. Eppstein, M. A. Van Der Pas, and B. B. Schryver,
11 **Cordycepin Analogs of ppp5'A2'p5'A2'p5'A (2-5A) Inhibit Protein Synthesis through**
12 **Activation of the 2-5A-dependent Endonuclease.**
13 The Journal of Biological Chemistry 260 (1985) 3666-3671.
14
- 15 35. J. Imai, K. Lesiak, and P. F. Torrence,
16 **Respective Role of Each of the Purine N-6 Amino Groups of 5'-O-**
17 **Triphosphoryladenyl(2'-5')adenyl(2'-5')adenosine in Binding to and Activation of**
18 **RNase L.**
19 The Journal of Biological Chemistry 260 (1985) 1390-1393.
20
- 21 36. C. Bisbal, M. Silhol, M. Lemaitre, B. Bayard, T. Salehzada, and B. Lebleu,
22 **5'-Modified Agonist and Antagonist (2'-5')(A)n Analogues: Synthesis and Biological**
23 **Activity.**
24 Biochemistry 26 (1987) 5172-5178.
25
- 26 37. P. F. Torrence, D. Brozda, D. Alster, R. Charubala, and W. Pfeiderer,
27 **Only One 3'-Hydroxyl Group of ppp5'A2'p5'A2'p5'A (2-5A) Is required for Activation**
28 **of the 2-5A-dependent Endonuclease.**
29 The Journal of Biological Chemistry 263 (1988) 1131-1139.
30
- 31 38. G. Charachon, R. W. Sobol, C. Bisbal, T. Salehzada, M. Silhol, R. Charubala, W.
32 Pfeiderer, B. Lebleu, and R. J. Suhadolnik,
33 **Phosphorothioate Analogues of (2'-5')(A)4: Agonist and Antagonist Activities in Intact**
34 **Cells.**
35 Biochemistry 29 (1990) 2550-2556.
36
- 37 39. R. W. Sobol, E. E. Henderson, N. Kon, J. Shao, P. Hitzges, E. Mordechai, N. L.
38 Reichenbach, R. Charubala, H. Shirmeister, W. Pfeiderer, and R. J. Suhadolnik,
39 **Inhibition of HIV-1 Replication and Activation of RNase L by**
40 **Phosphorothioate/Phosphodiester 2',5'-Oligoadenylate Derivatives.**
41 The Journal of Biological Chemistry 270 (1995) 5963-5978.
42
- 43 40. Y. Kitade, M. Wakana, T. Tsuboi, C. Yatome, S. F. Bayly, M. R. Player, and P. F.
44 Torrence,
45 **2-Methyladenosine-Substituted 2',5'-Oligoadenylates: Conformations, 2-5A Binding and**
46 **Catalytic Activities with Human Ribonuclease L.**
47 Bioorganic & Medicinal Chemistry Letters 10 (2000) 329-331.
48
- 49 41. A. Yoshimura, M. Nakanishi, C. Yatome, and Y. Kitade,

- 1 **Comparative Study of the Biological Properties of 2',5'-Oligoadenylate Derivatives with**
2 **Purified Human RNase L Expressed in E. coli.**
3 Journal of Biochemistry 132 (2002) 643-648.
4
- 5 42. O. Páv, E. Protivinska, M. Pressova, M. Collinsova, J. Jiracek, J. Snasel, M. Masojdkova,
6 M. Budesinsky, and I. Rosenberg,
7 **Activation of Murine RNase L by Isopolar 2'-Phosphonate Analogues of 2',5'**
8 **Oligoadenylates.**
9 Journal of Medicinal Chemistry 49 (2006) 3955-3962.
10
- 11 43. O. Páv, N. Panova, J. Snášel, E. Sborníková, I. Rosenberg,
12 **Activation of human RNase L by 2'- and 5'-O-methylphosphonate-modified**
13 **oligoadenylates.**
14 Bioorganic & Medicinal Chemistry Letters 22 (2012) 181-185.
15
- 16 44. K. Nagaoka, Y. Kitamura, Y. Ueno, Y. Kitade,
17 **5'-O-Dephosphorylated 2',5'-oligoadenylate (2-5A) with 8-methyladenosine at the 2'-**
18 **terminus activates human RNase L.**
19 Bioorganic & Medicinal Chemistry Letters 20 (2010) 1186-1188.
20
- 21 45. W. L. Jorgensen, J. Chandrasekhar, J. Madura, and M. L. Klein,
22 **Comparison of simple potential functions for simulating liquid water.**
23 Journal of Chemical Physics. 79 (1983) 926-935.
24
- 25 46. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C.
26 Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman,
27 **A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and**
28 **Organic-Molecules.**
29 Journal of the American Chemical Society 117 (1995) 5179-5197.
30
- 31 47. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling,
32 **Comparison of multiple Amber force fields and development of improved protein**
33 **backbone parameters.**
34 Proteins 65 (2006) 712-725.
35
- 36 48. T. E. Cheatham III, P. Cieplak and Peter A. Kollman,
37 **A Modified Version of the Cornell et al. Force Field with Improved Sugar Pucker**
38 **Phases and Helical Repeat.**
39 Journal of Biomolecular Structure & Dynamics 16 (1999) 845-862.
40
- 41 49. A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham III, Ch. A. Laughton and M.
42 Orozco,
43 **Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of**
44 **alpha/gamma Conformers.**
45 Biophysical Journal 92 (2007) 3817-3829.
46
- 47 50. M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. Cheatham III, and P.
48 Jure ka,

- 1 **Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum**
2 **Chemical Calculations of Glycosidic Torsion Profiles.**
3 Journal of Chemical Theory and Computation 7 (2011) 2886-2902.
4
- 5 51. D. A. Pearlman, D. A. Case, J. W. Caldwell, W. R. Ross, T. E. Cheatham, S. DeBolt, D.
6 Ferguson, G. Seibel and P. Kollman,
7 **AMBER, a package of computer programs for applying molecular mechanics, normal**
8 **mode analysis, molecular dynamics and free energy calculations to simulate the**
9 **structural and energetic properties of molecules.**
10 Computer Physics Communications 91 (1995) 1–41.
11
- 12 52. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D.
13 Skeel, L. Kale, and K. Schulten,
14 **Scalable molecular dynamics with NAMD.**
15 Journal of Computational Chemistry 26 (2005) 1781-1802.
16
- 17 53. T. E. Cheatham, J. L. Miller, T. Fox, T. A. Darden and P. A. Kollman,
18 **Molecular dynamics simulations on solvated biomolecular systems: the particle mesh**
19 **Ewald method leads to stable trajectories of DNA, RNA, and proteins.**
20 Journal of the American Chemical Society. 117 (1994) 4193–4194.
21
- 22 54. J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen,
23 **Numerical Integration of the Cartesian Equations of Motion of a System with**
24 **Constraints: Molecular Dynamics of n-Alkanes.**
25 Journal of Computational Physics 23 (1977) 327–341.
26
- 27 55. M. J. Harvey, G. Giupponi, and G. De Fabritiis,
28 **ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale.**
29 Journal of Chemical Theory and Computation 5 (2009) 1632–1639.
30
- 31 56. ACEMD <http://www.acellera.com/>
32
- 33 57. H. C. Andersen,
34 **RATTLE: A "Velocity" Version of the SHAKE Algorithm for Molecular Dynamics**
35 **Calculations.**
36 Journal of Computational Physics 52 (1983) 24–34.
37
- 38 58. S. G. Lambrakos, J. P. Boris, E. S. Oran, I. Chandrasekhar, M. Nagumo,
39 **A Modified SHAKE algorithm for Maintaining Rigid Bonds in Molecular Dynamics**
40 **Simulations of Large Molecules.**
41 Journal of Computational Physics 85 (1989) 473–486.
42
- 43 59. T. Giorgino, and Gianni de Fabritiis,
44 **A high-throughput steered molecular dynamics study on the free energy profile of ion**
45 **permeation through gramicidin A.**
46 Journal of Chemical Theory and Computation 7 (2011) 1943-1950.
47
- 48 60. M. Selent, F. Sanz, M. Pastor, G. De Fabritiis,
49 **Induced Effects of Sodium Ions on Dopaminergic G-Protein Coupled Receptors.**

- 1 PLOS Computational Biology 6 (2010) e1000884.
2
- 3 61. I. Buch, T. Giorgino, G. De Fabritiis,
4 **Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics**
5 **simulations.**
6 Proceedings of the National Academy of Sciences of the USA 108 (2011) 10184-10189.
7
- 8 62. K. A. Feenstra, B. Hess, H. J. C. Berendsen,
9 **Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-**
10 **rich systems.**
11 Journal of Computational Chemistry 20 (1999) 786-798.
12
- 13 63. W. Humphrey, A. Dalke and K. Schulten,
14 **VMD - Visual Molecular Dynamics.**
15 Journal of Molecular Graphics 14 (1996) 33-38.
16
- 17 64. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng
18 and T. E. Ferrin,
19 **UCSF Chimera - A Visualization System for Exploratory Research and Analysis.**
20 Journal of Computational Chemistry 25 (2004) 1605-1612.
21
- 22 65. A. Van Aerschot,
23 **Oligonucleotides as antivirals: Dream or realistic perspective?**
24 Antiviral Research 71 (2006) 307-316.
25
26

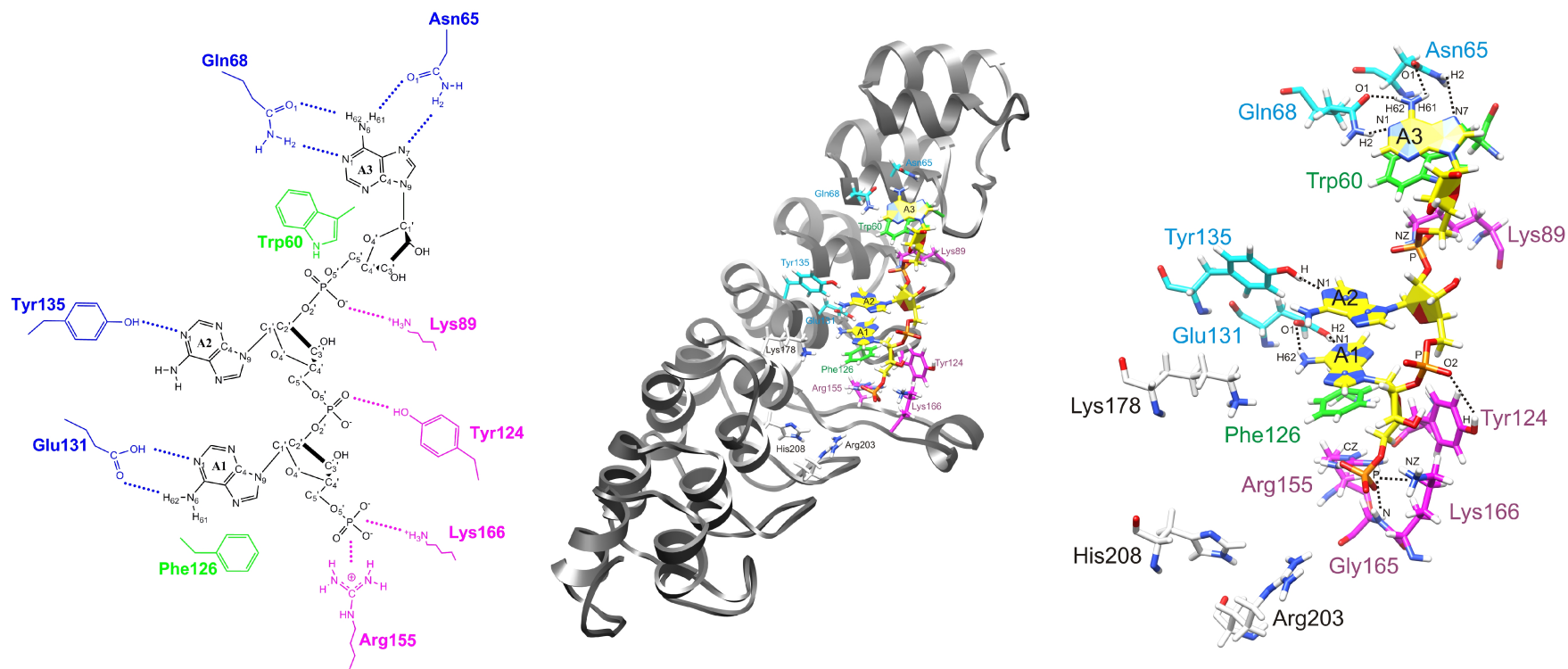


Figure 1: The ligand 2–5A is bound by the second to the fourth ankyrin repeats and interacts with the RNase L residues by forming hydrogen bonds (*Asn65*, *Gln68*, *Glu131*, *Tyr135*, *Tyr124*), salt bridges (*Lys89*, *Arg155*, *Lys166*) or stacking interactions (*Trp60*, *Phe126*) [8].

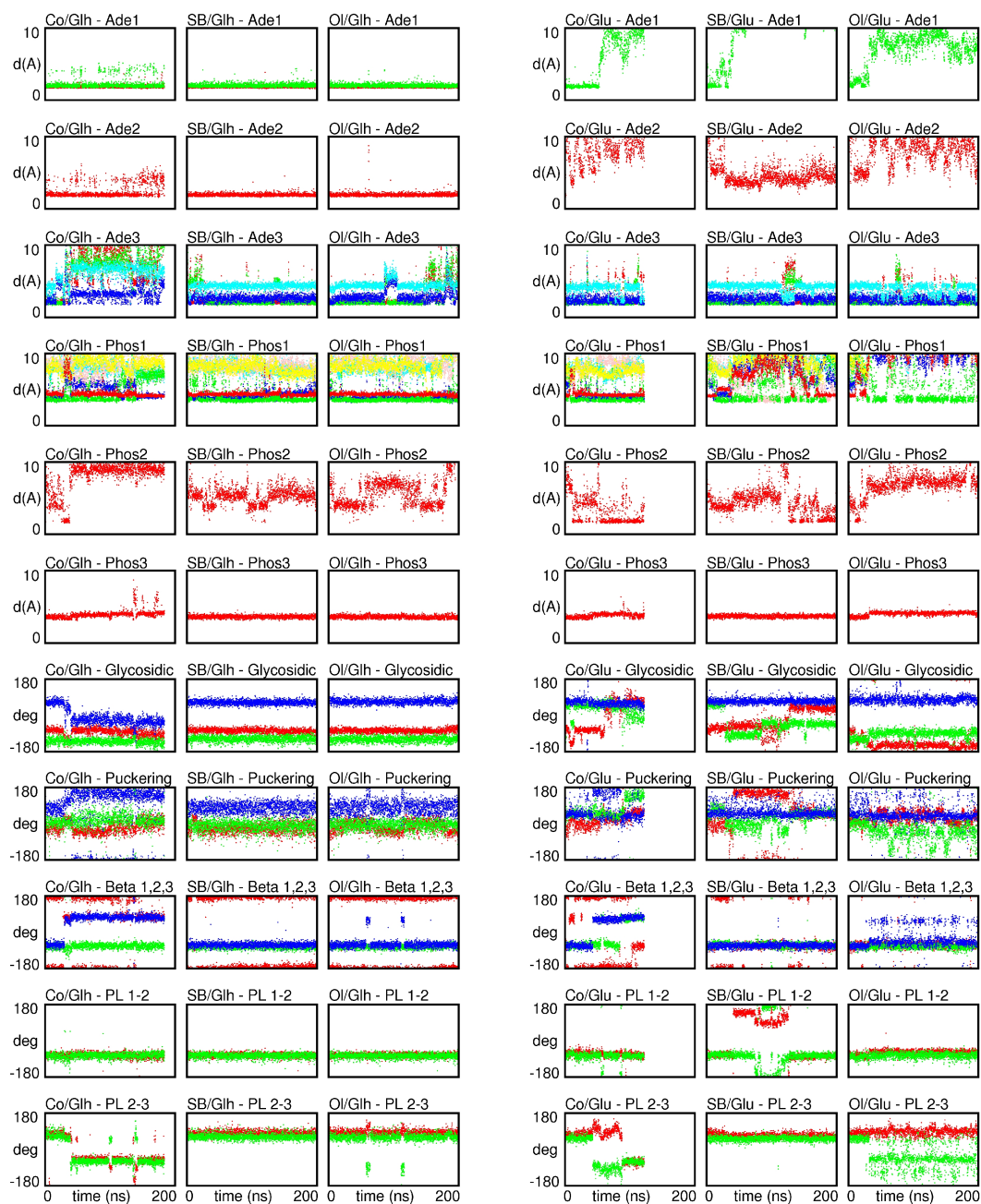


Figure 2: MD simulations no. 1-6 – columns of charts taken from left to right: Cornell FF [46] / RNase L_E131H (**Co/Glh**), 99SB FF [47-49] / RNase L_E131H (**SB/Glh**), Olo FF [50] / RNase L_E131H (**OI/Glh**), Cornell FF [46] / RNase L_E131 (**Co/Glu**), 99SB FF [47-49] / RNase L_E131 (**SB/Glu**), Olo FF [50] / RNase L_E131 (**OI/Glu**). Time evolution of various structural parameters – rows of charts taken from top to down. **Ade1** binding by hydrogen bonds with the side chain of protonated/unprotonated Glu131: Glu131_O-H62_Ade1 (red line) and Glu131_OH-N1_Ade1 (green line). **Ade2** contacts with the side chain of Tyr135: Tyr135_OH-N1_Ade2 (red line). The hydrogen bond network involving **Ade3** and the side chains of Gln68: Gln68_O1-H62_Ade3 (red line) and Gln68_H2-N1_Ade3 (green line) and Asn65: Asn65_O1-H61_Ade3 (blue line). The side chains of Arg155 (red line) and Lys166 (green line) formed salt bridges with the 5'-phosphate group of the first AMP (**Phos1**). The 5'-phosphate group of the second AMP (**Phos2**) formed occasional hydrogen bonds with the Tyr124 side chain (red line). The 5'-phosphate group of the third AMP (**Phos3**) formed a salt bridge with the side chain of Lys89 (red line). **Glycosidic** torsion angle values (in terms of torsion angle O4'-C1'-N9-C4) of Ade1 (red line), Ade2 (green line), Ade3 (blue line). **Puckering** of ribose rings (Ade1 - red line, Ade2 - green line, Ade3 - blue line). Conformations of the C4'-C5' moieties connecting ribose and internucleotide linkage segments expressed in terms of the **beta**: C3'-C4'-C5'-O5' torsion angle (Ade1 - red line, Ade2 - green line, Ade3 - blue line). Conformations of Ade1-Ade2 and Ade2-Ade3 internucleotide phosphodiester linkages (**PL 1-2** and **PL 2-3**) in terms of torsion angles C5'-O5'-P-O2' (red line) and O5'-P-O2'-C2' (green line).

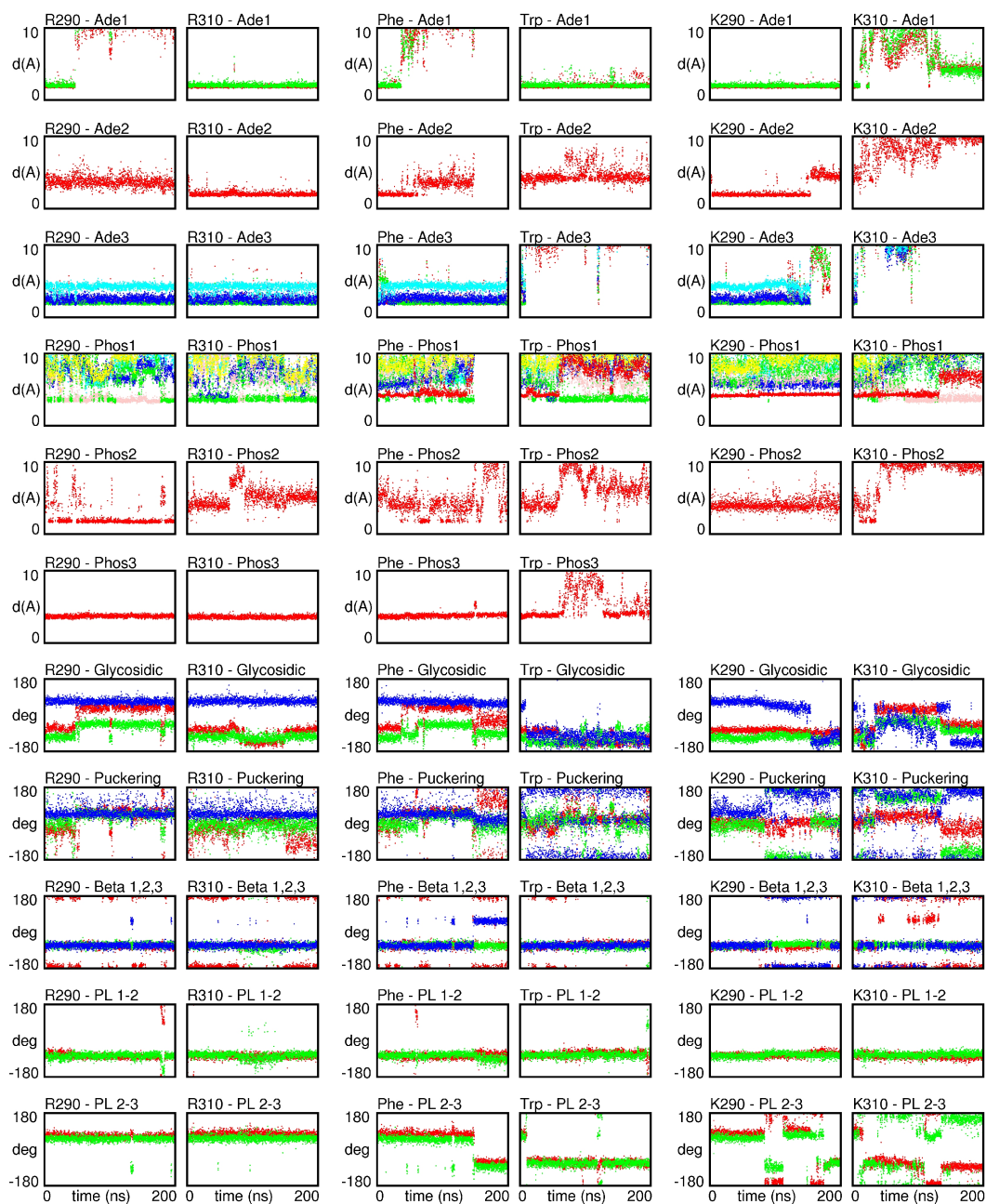


Figure 3: MD simulations no. 7-12 – columns of charts taken from left to right: R155A - MD no. 7-8; F126A - MD no. 9; W60A - MD no. 10; K89A - MD no. 11-12 - see Table S1. Time evolution of various structural parameters – rows of charts taken from top to down. **Ade1** binding by hydrogen bonds with the side chain of protonated/unprotonated Glu131: Glu131_O-H62_Ade1 (red line) and Glu131_OH-N1_Ade1 (green line). **Ade2** contacts with the side chain of Tyr135: Tyr135_OH-N1_Ade2 (red line). The hydrogen bond network involving **Ade3** and the side chains of Gln68: Gln68_O1-H62_Ade3 (red line) and Gln68_H2-N1_Ade3 (green line) and Asn65: Asn65_O1-H61_Ade3 (blue line). The side chains of Arg155 (red line) and Lys166 (green line) formed salt bridges with the 5'-phosphate group of the first AMP (**Phos1**). The 5'-phosphate group of the second AMP (**Phos2**) formed occasional hydrogen bonds with the Tyr124 side chain (red line). The 5'-phosphate group of the third AMP (**Phos3**) formed a salt bridge with the side chain of Lys89 (red line). **Glycosidic** torsion angle values (in terms of torsion angle O4'-C1'-N9-C4) of Ade1 (red line), Ade2 (green line), Ade3 (blue line). **Puckering** of ribose rings (Ade1 - red line, Ade2 - green line, Ade3 - blue line). Conformations of the C4'-C5' moieties connecting ribose and internucleotide linkage segments expressed in terms of the **beta**: C3'-C4'-C5'-O5' torsion angle (Ade1 - red line, Ade2 - green line, Ade3 - blue line). Conformations of Ade1-Ade2 and Ade2-Ade3 internucleotide phosphodiester linkages (**PL 1-2** and **PL 2-3**) in terms of torsion angles C5'-O5'-P-O2' (red line) and O5'-P-O2'-C2' (green line).



Figure 4: MD simulations no. 13-19 – columns of charts taken from left to right: agonists (pA₂₅A₂₅B, pB₂₅A₂₅A, pB₂₅A₂₅B - MD no. 13-15) and inactive analogs (pA₂₅B₂₅A, pA₂₅B₂₅B, pB₂₅B₂₅A, pB₂₅B₂₅B - MD no. 16-19). Time evolution of various structural parameters – rows of charts taken from top to down. **Ade1** binding by hydrogen bonds with the side chain of protonated/unprotonated Glu131: Glu131_O-H62_Ade1 (red line) and Glu131_OH-N1_Ade1 (green line). **Ade2** contacts with the side chain of Tyr135: Tyr135_OH-N1_Ade2 (red line). The hydrogen bond network involving **Ade3** and the side chains of Gln68: Gln68_O1-H62_Ade3 (red line) and Gln68_H2-N1_Ade3 (green line) and Asn65: Asn65_O1-H61_Ade3 (blue line). The side chains of Arg155 (red line) and Lys166 (green line) formed salt bridges with the 5'-phosphate group of the first AMP (**Phos1**). The 5'-phosphate group of the second AMP (**Phos2**) formed occasional hydrogen bonds with the Tyr124 side chain (red line). The 5'-phosphate group of the third AMP (**Phos3**) formed a salt bridge with the side chain of Lys89 (red line). **Glycosidic** torsion angle values (in terms of torsion angle O4'-C1'-N9-C4) of Ade1 (red line), Ade2 (green line), Ade3 (blue line). **Puckering** of ribose rings (Ade1 - red line, Ade2 - green line, Ade3 - blue line). Conformations of the C4'-C5' moieties connecting ribose and internucleotide linkage segments expressed in terms of the **beta**: C3'-C4'-C5'-O5' torsion angle (Ade1 - red line, Ade2 - green line, Ade3 - blue line). Conformations of Ade1-Ade2 and Ade2-Ade3 internucleotide phosphodiester linkages (**PL 1-2** and **PL 2-3**) in terms of torsion angles C5'-O5'-P-O2' (red line) and O5'-P-O2'-C2' (green line).

	RNase L	RNase L	25A	Force Field	Temp (K)	Length (ns)	EC50 (nM) [24]	Crel [24]
01	E131 H	-	pA ₂₅ A ₂₅ A	94 [46]	310	<200		
02	E131 H	-	pA ₂₅ A ₂₅ A	99SB [47-49]	310	200		
03	E131 H	-	pA ₂₅ A ₂₅ A	Olo [50]	310	200		
04	E131	-	pA ₂₅ A ₂₅ A	94 [46]	290	<200		
05	E131	-	pA ₂₅ A ₂₅ A	99SB [47-49]	290	200		
06	E131	-	pA ₂₅ A ₂₅ A	Olo [50]	290	200		
07	E131 H	R155A	pA ₂₅ A ₂₅ A	99SB [47-49]	290	200		
08	E131 H	R155A	pA ₂₅ A ₂₅ A	99SB [47-49]	310	200		
09	E131 H	F126A	pA ₂₅ A ₂₅ A	99SB [47-49]	310	200		
10	E131 H	W60A	pA ₂₅ A ₂₅ A	99SB [47-49]	310	200		
11	E131 H	K89A	pA ₂₅ A ₂₅ A	99SB [47-49]	290	200		
12	E131 H	K89A	pA ₂₅ A ₂₅ A	99SB [47-49]	310	200	3.5	1.0
14	E131 H	-	pA ₂₅ A ₂₅ B	99SB [47-49]	310	200	4.9	1.4
13	E131 H	-	p B ₂₅ A ₂₅ A	99SB [47-49]	310	200	77	22
15	E131 H	-	p B ₂₅ A ₂₅ B	99SB [47-49]	310	200	130	37
16	E131 H	-	pA ₂₅ B ₂₅ A	99SB [47-49]	310	100	530	150
18	E131 H	-	pA ₂₅ B ₂₅ B	99SB [47-49]	310	100	> 1000	> 300
17	E131 H	-	p B ₂₅ B ₂₅ A	99SB [47-49]	310	100	> 1000	> 300
19	E131 H	-	p B ₂₅ B ₂₅ B	99SB [47-49]	310	>100	> 1000	> 300

Table S1: Molecular dynamics simulations produced in this study.

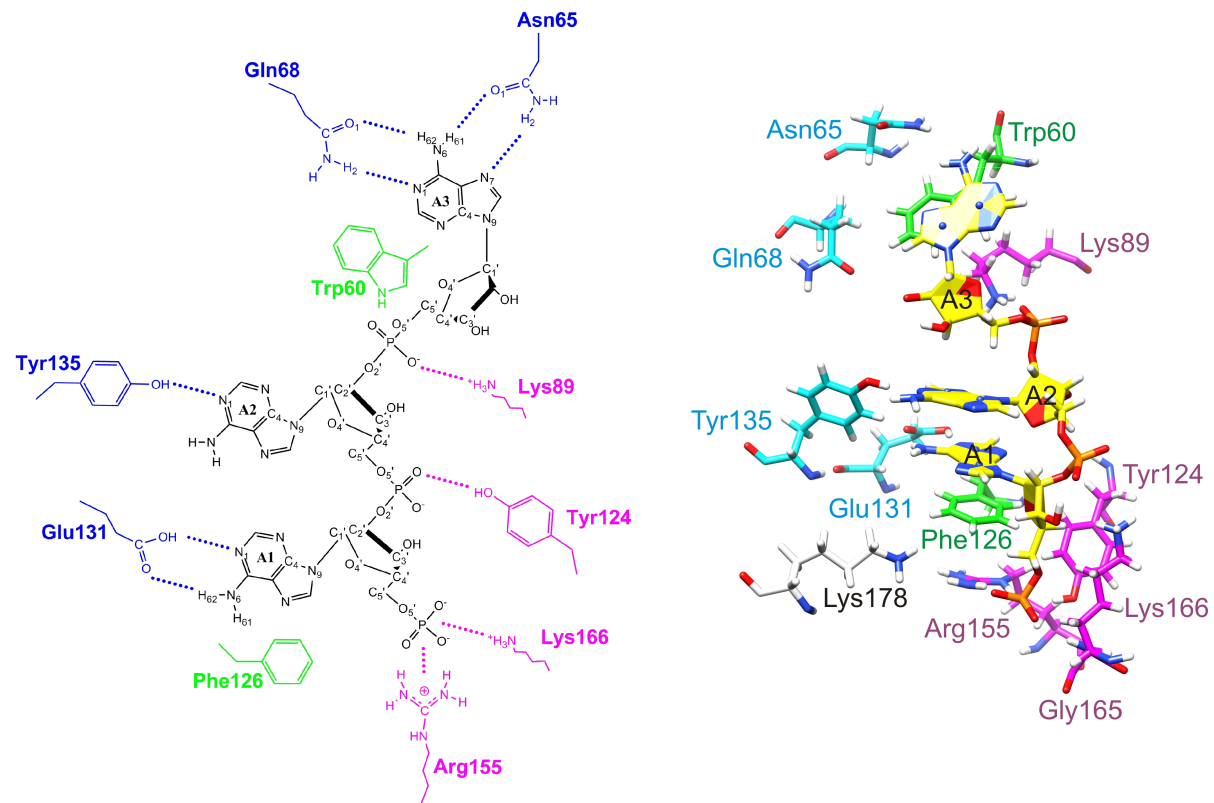


Figure S1: The pA₂₅A₂₅A molecule was bound to the surface of RNase L_E131H despite of conformational transitions of the adenine ring of the third AMP.

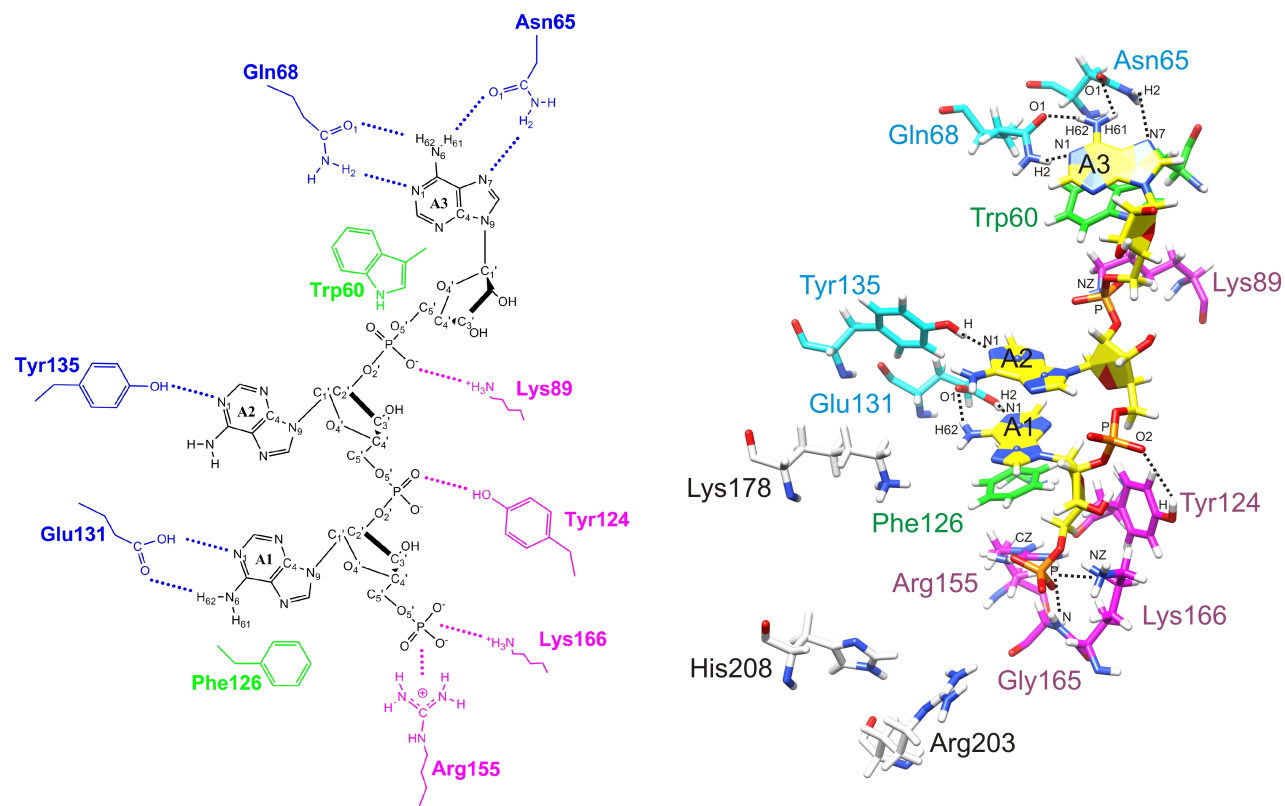


Figure S2: The pA₂₅A₂₅A molecule was bound to the surface of RNase L_{E131H} in the original arrangement.

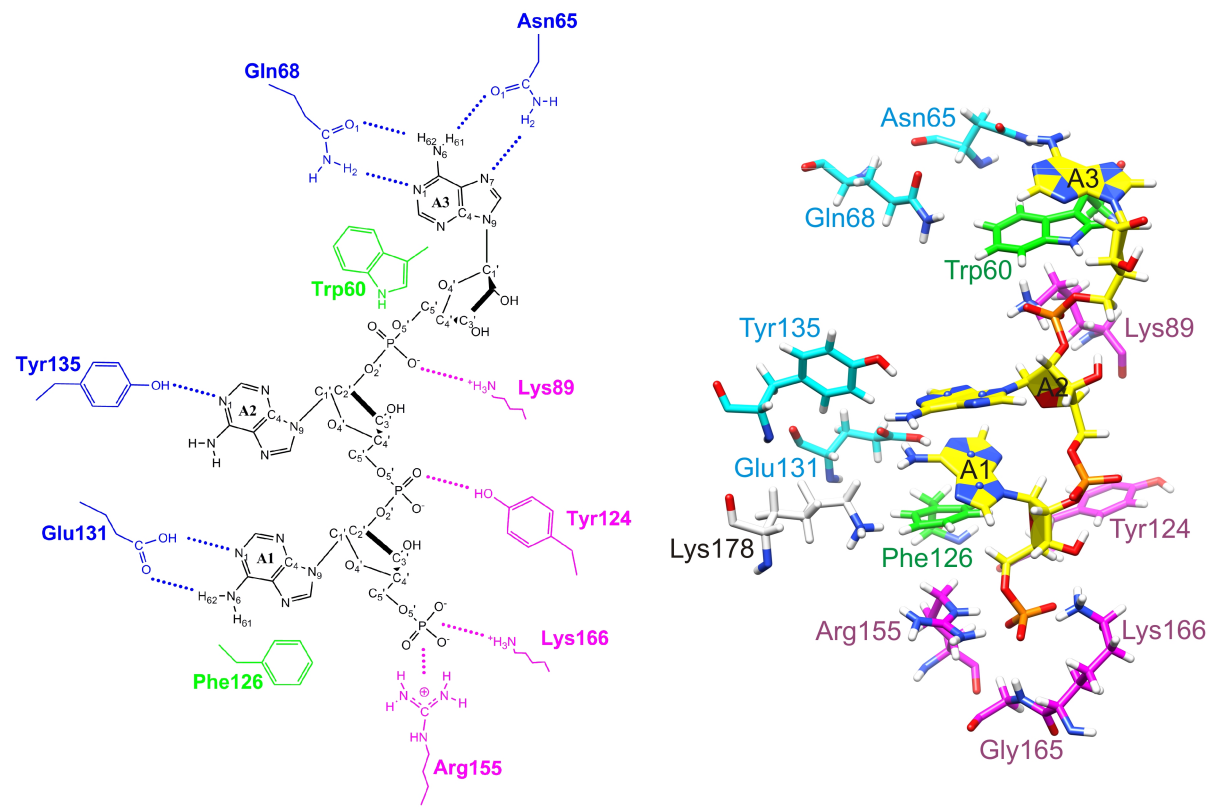


Figure S3: The pA₂₅A₂₅A molecule stayed bound to the surface of RNase L_{E131H} close to the original arrangement.

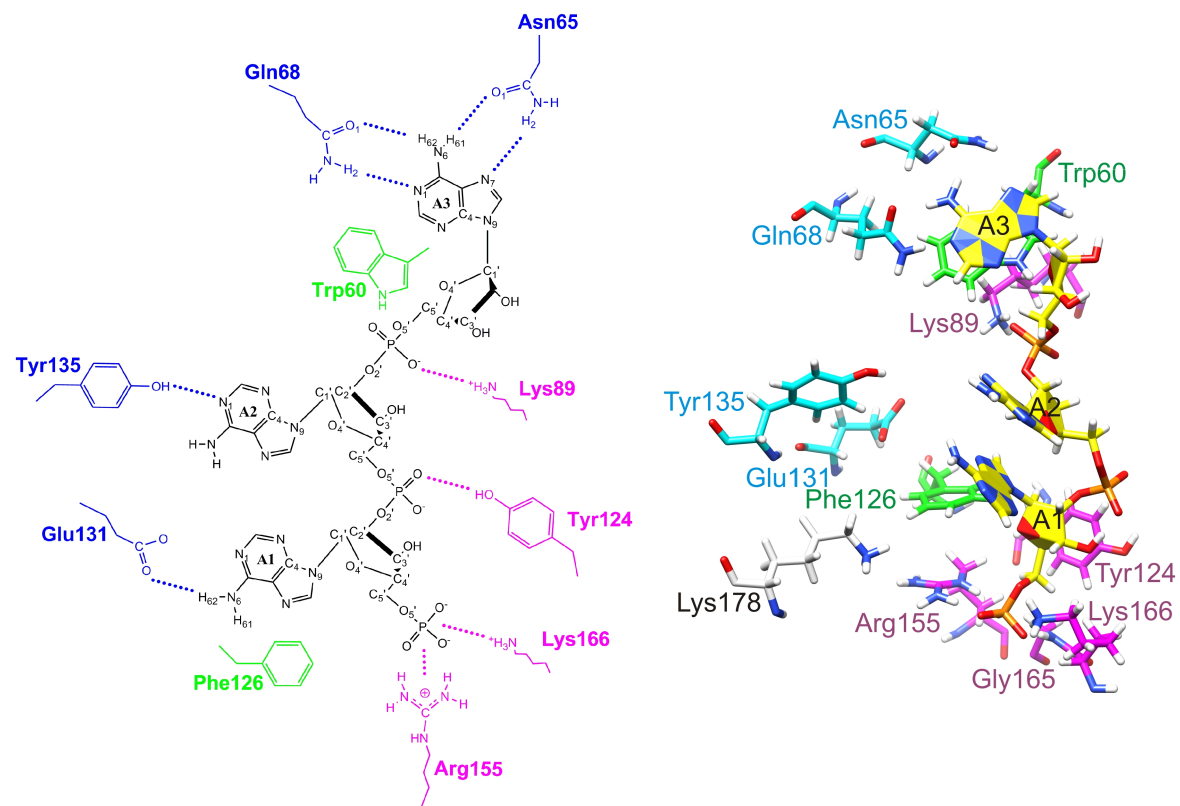


Figure S4: The pA₂₅A₂₅A molecule rearranged substantially with only Ade3 (which was the most distant adenine ring from the unprotonated E131 residue) bound in its original position on the surface of RNase L_E131.

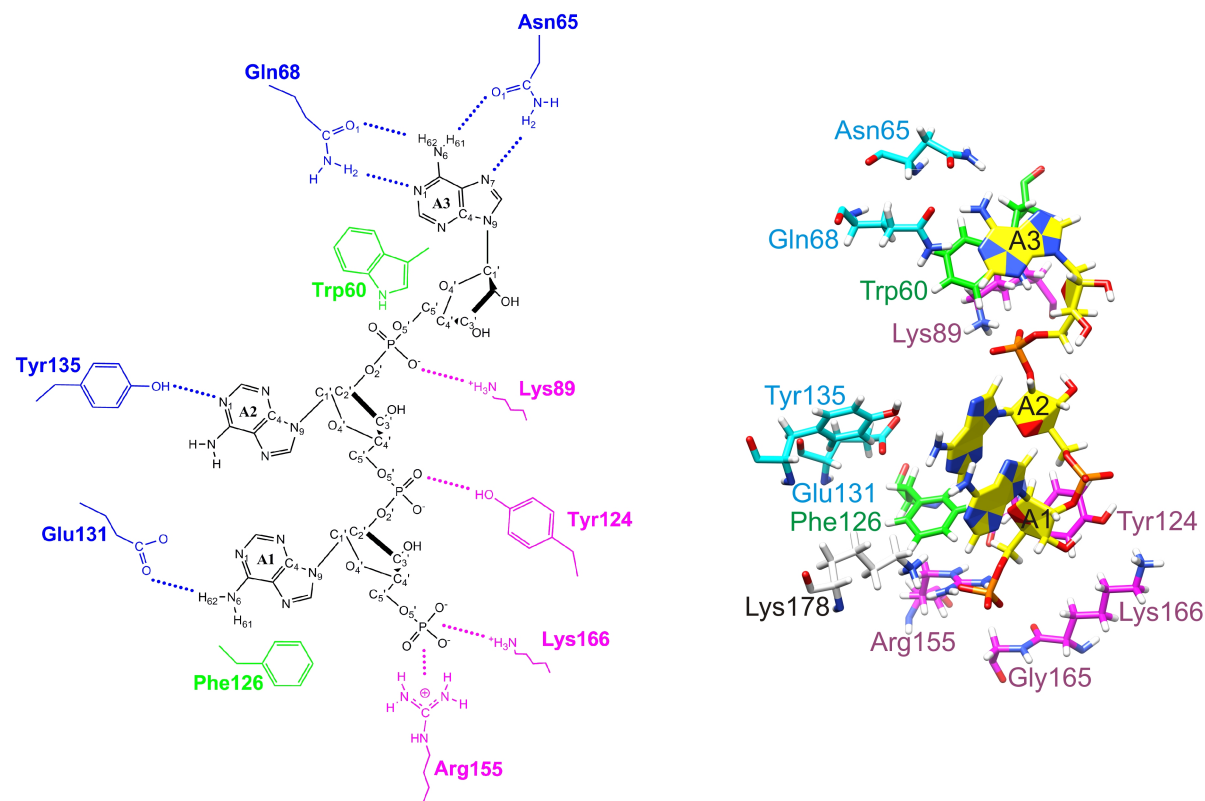


Figure S5: Again, the pA₂₅A₂₅A molecule changed its conformation within a 200 ns MD run. Only Ade3 stayed in its original binding site on the surface of RNase L_E131.

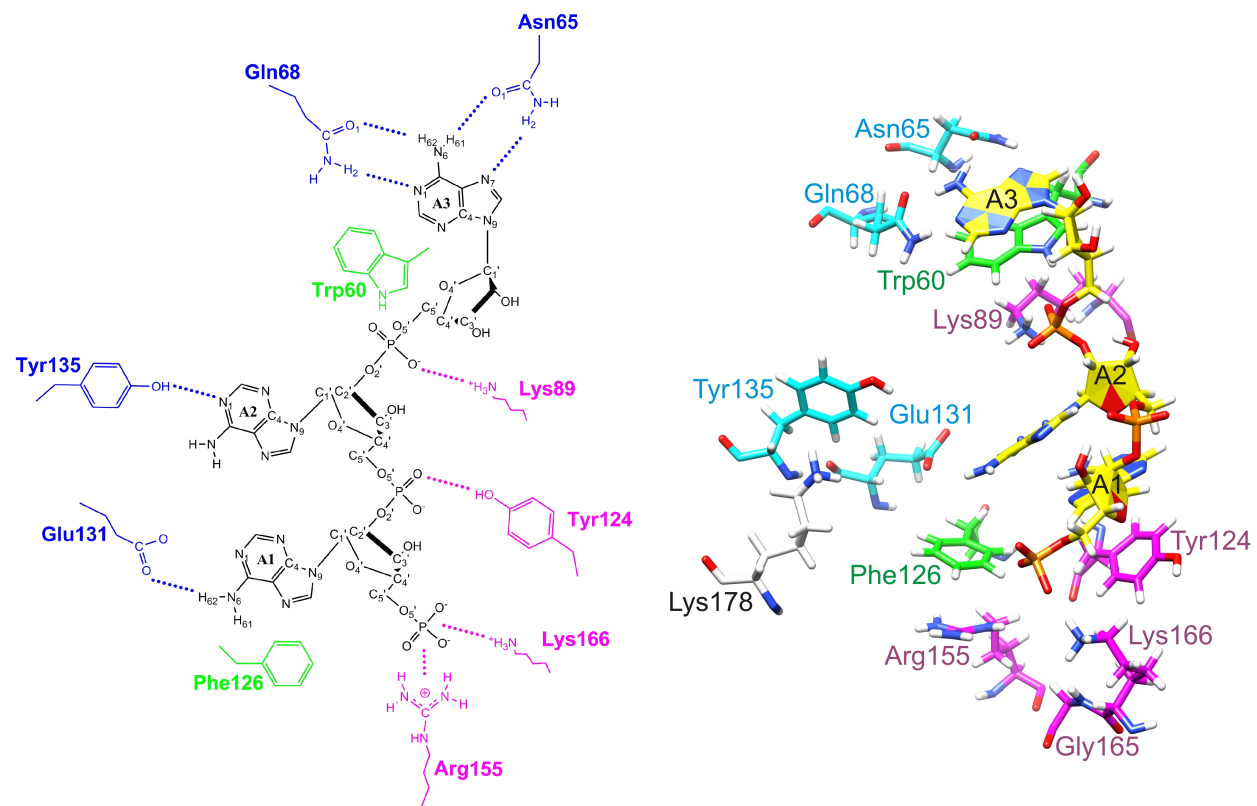


Figure S6: The pA₂₅A₂₅A molecule dissociated from ANK leaving only Ade3 in its original position on the surface of RNase L_E131.

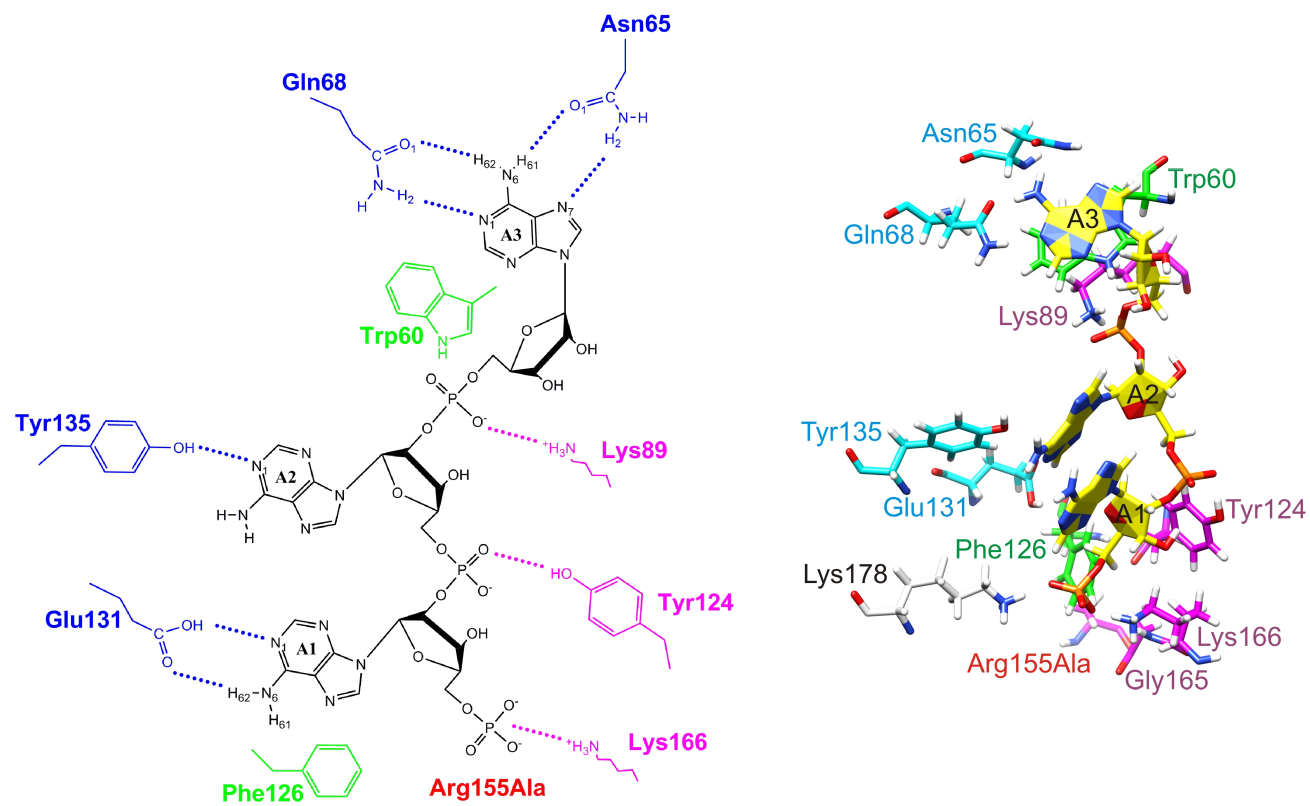


Figure S7: The pA₂₅A₂₅A molecule rearranged substantially within this MD run. Nevertheless, Ade3 (which was the most distant adenine ring from R155A) persisted in its original position on the surface of **RNaseL_E131H_R155A**.

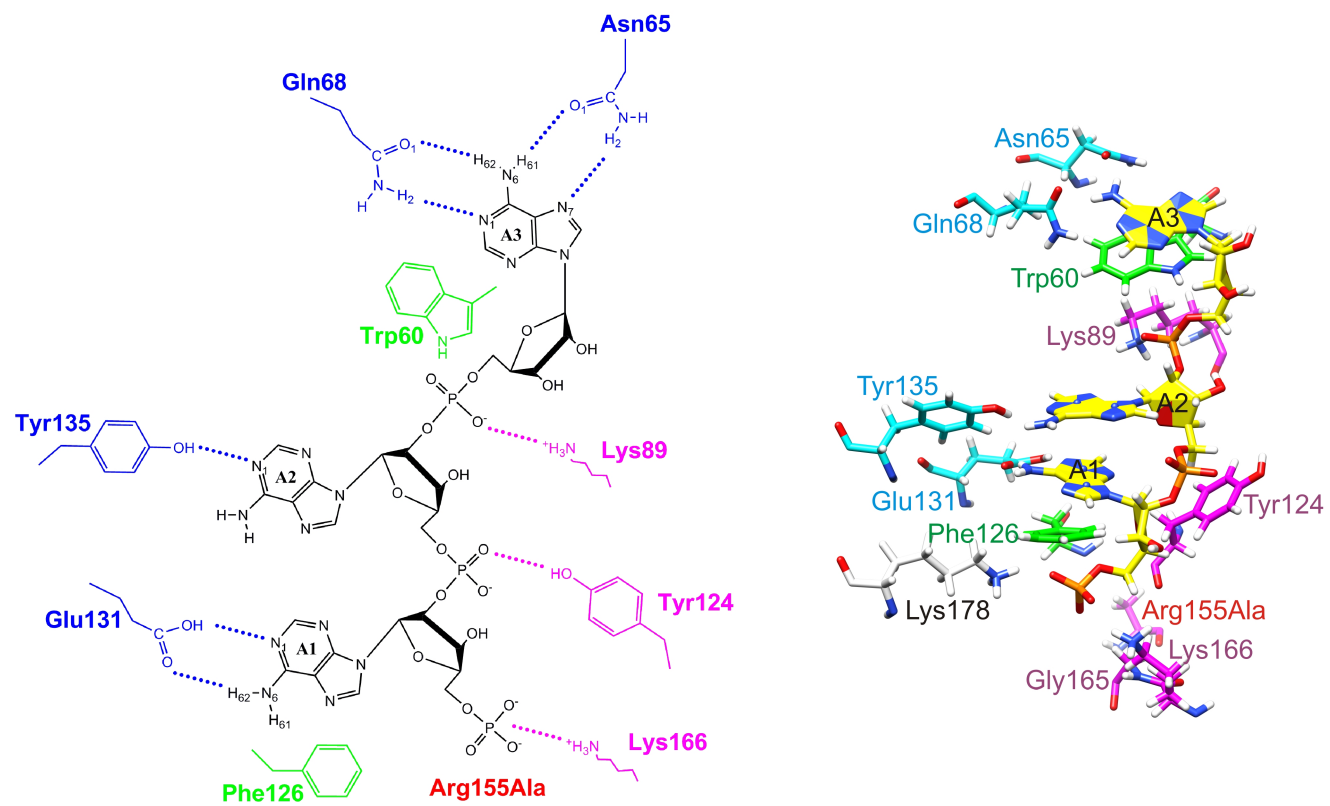


Figure S8: The pA₂₅A₂₅A molecule was bound steadily to the surface of RNaseL_E131H_R155A.

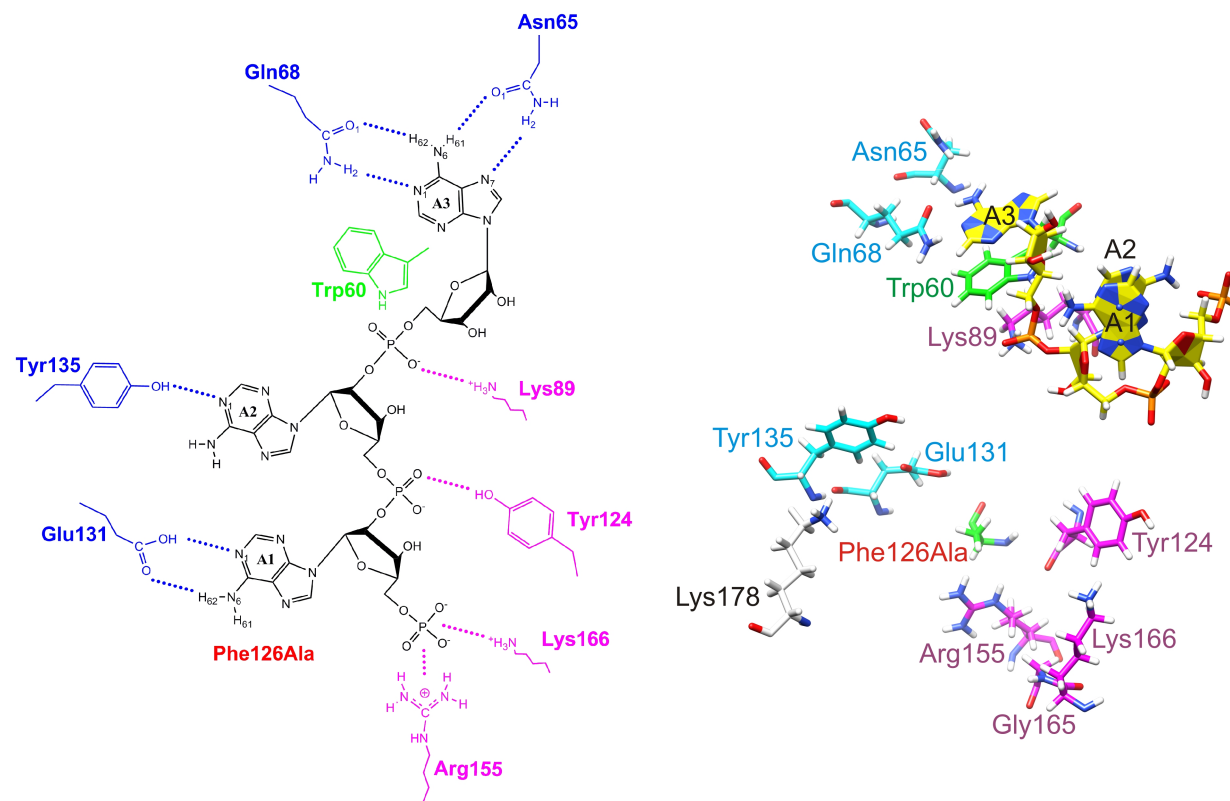


Figure S9: The pA₂₅A₂₅A molecule dissociated from the surface of RNaseL_E131H_F126A with only Ade3 (which was the most distant adenine ring from F126A) staying bound in its original position.

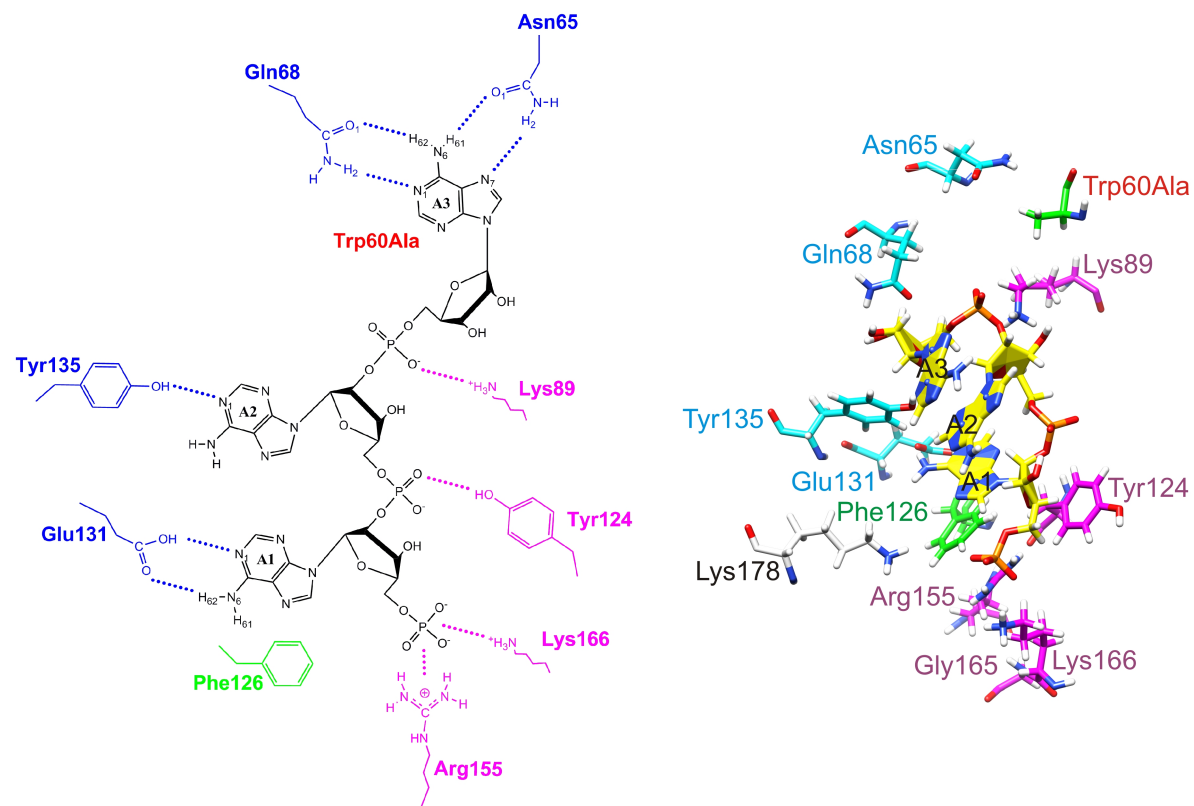


Figure S10: The pA₂₅A₂₅A molecule rearranged substantially with only Ade1 (which was the most distant adenine ring from the W60A) staying in its original binding site on the surface of **RNaseL_E131H_W60A**.

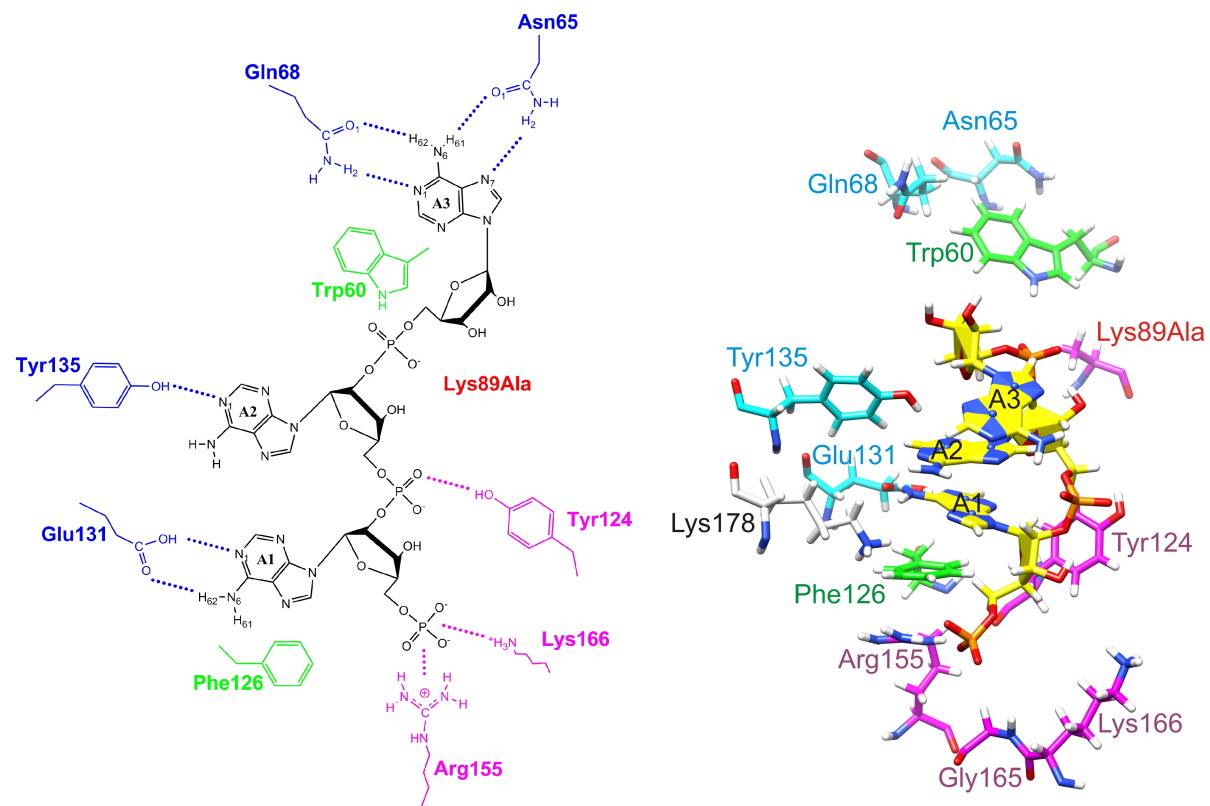


Figure S11: The pA₂₅A₂₅A molecule was steadily bound for ~150 ns. Then it dissociated leaving only Ade1 (which was the most distant adenine ring from the K89A) in its original position on the surface of RNaseL_E131H_K89A.

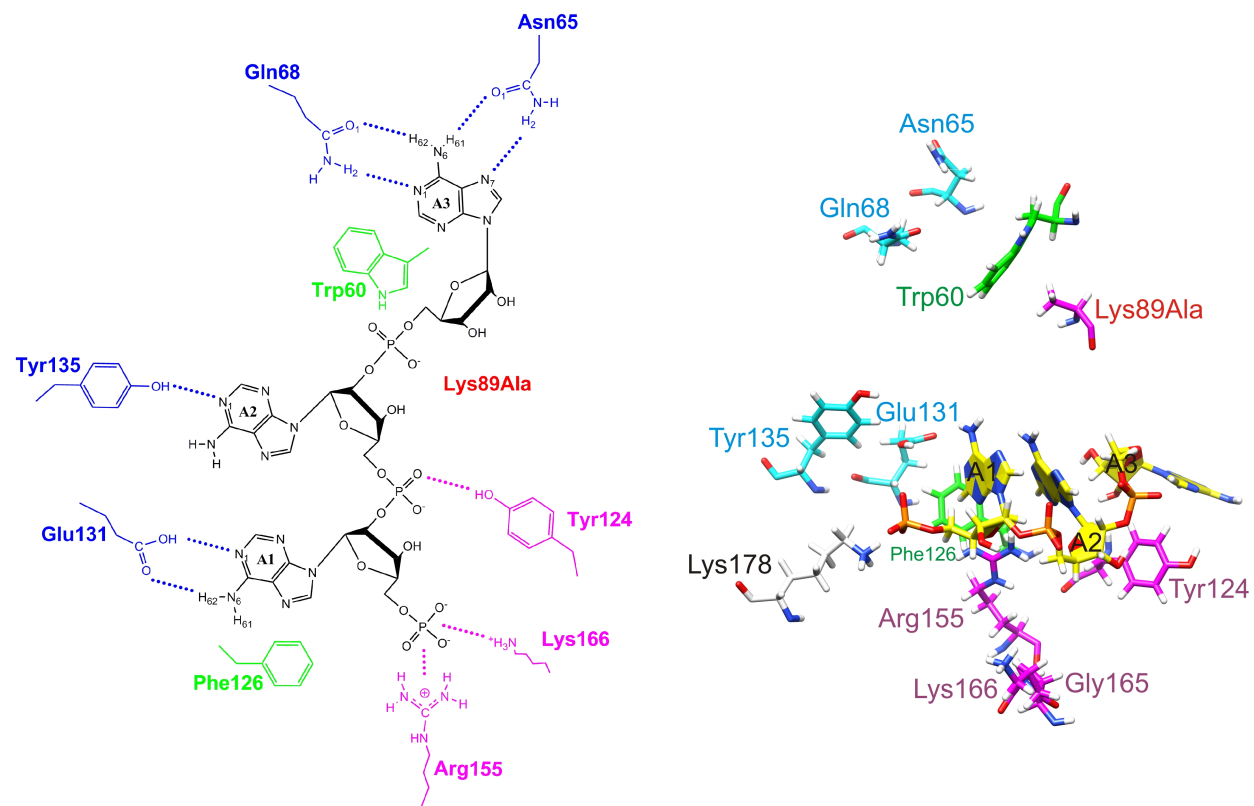


Figure S12: The pA₂₅A₂₅A molecule dissociated from the surface of RNaseL_E131H_K89A. Finally, pA₂₅A₂₅A was directed perpendicularly with respect to its original orientation.

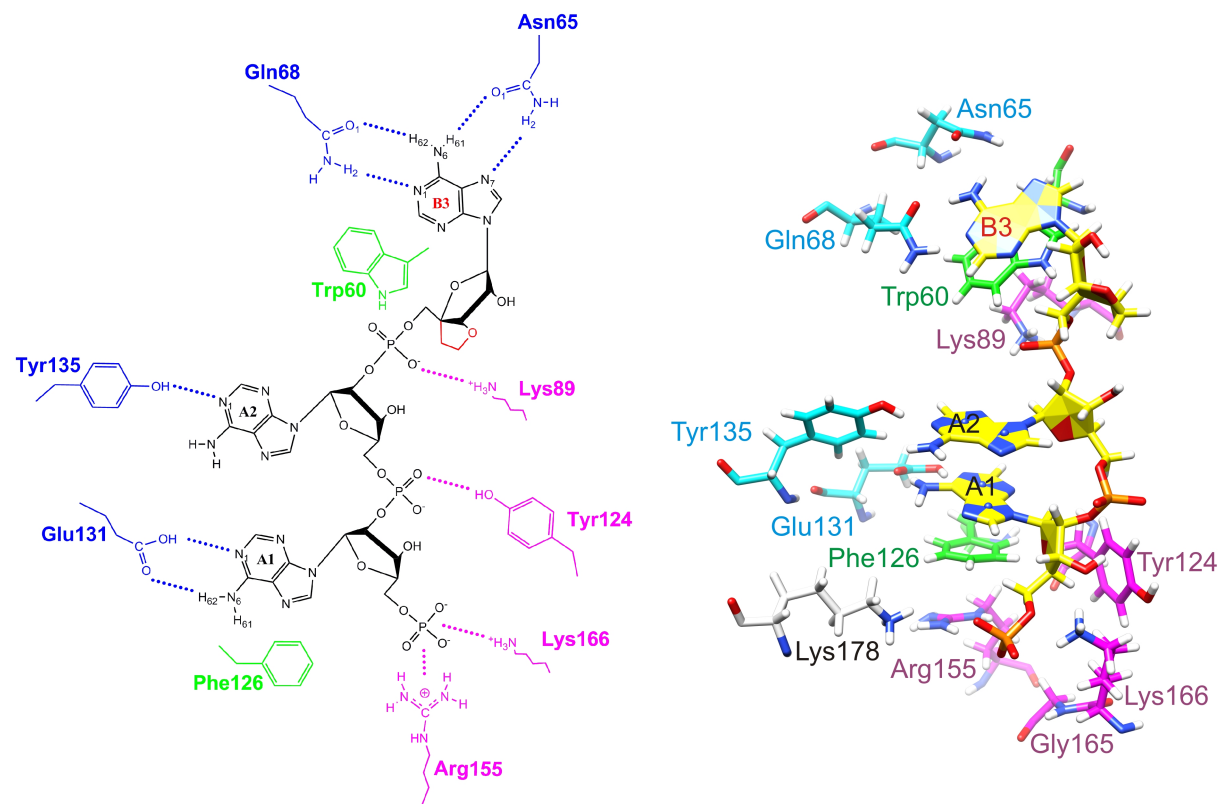


Figure S13: The pA₂₅A₂₅B molecule was steadily bound to the surface of RNaseL_E131H. It's in agreement with remarkable ability of pA₂₅A₂₅B to activate RNase L (EC₅₀ = 4.9 nM; Crel 1.4x [24] - where EC₅₀ was defined as the molarity (M) of 2-5A analogues required to cause a 50% degradation of 5S rRNA; Crel represents the ratio of EC₅₀ of analogues as compared to EC₅₀ = 3.5 nM of natural 2-5A; i.e. the smaller Crel the more effectively the analogue was able to activate RNase L).

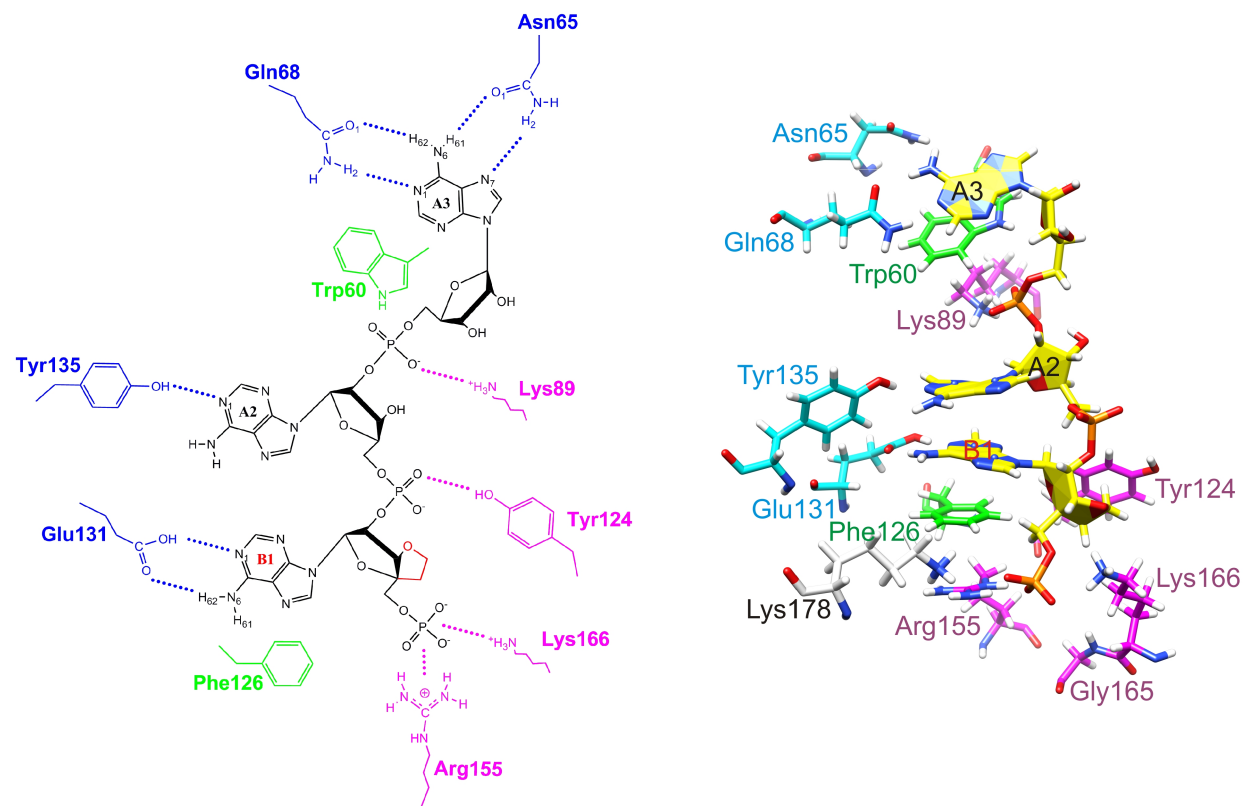


Figure S14: The pB₂₅A₂₅A molecule was steadily bound to the surface of RNaseL_E131H. It seems to be in accordance with remarkable ability of pB₂₅A₂₅A to activate RNase L (EC₅₀ = 77 nM; Crel 22x [24]).

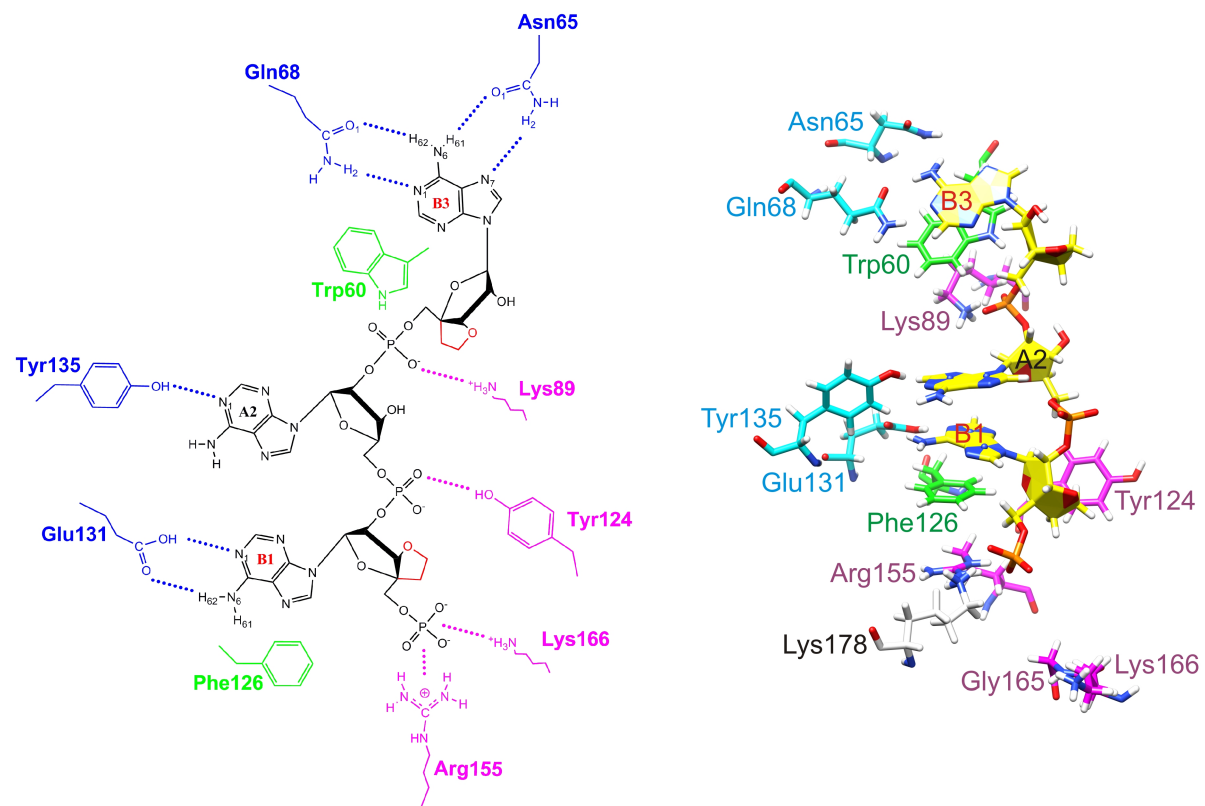


Figure S15: The p_B₂₅A₂₅B molecule stayed close to its original conformation/position on the surface of RNaseL_E131H. It seems to be in agreement with fairly well ability of p_B₂₅A₂₅B to activate RNase L (EC₅₀ > 130 nM; Crel 37x [24]).

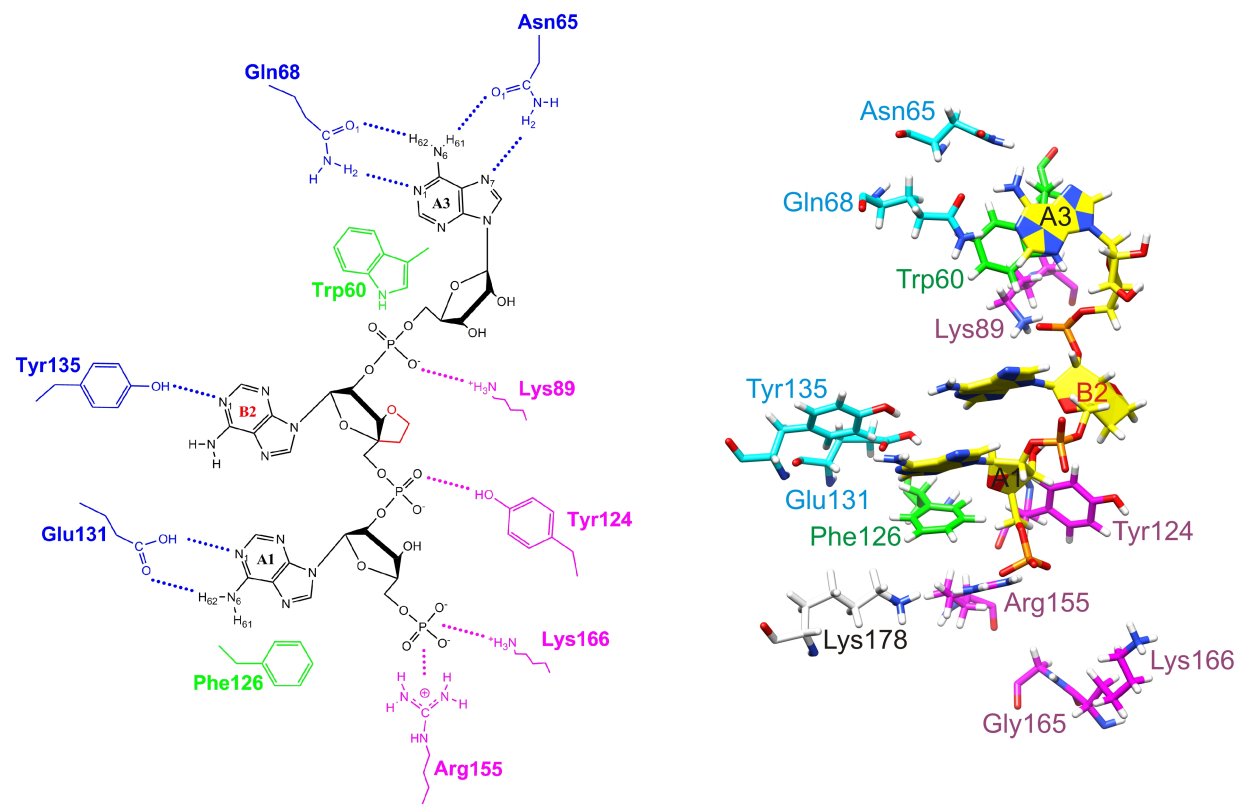


Figure S16: The pA₂₅B₂₅A molecule was stable for quite some time, nevertheless at the very end of a MD run it rearranged substantially with just only Ade3 persisting in its original binding site on the surface of RNaseL_E131H. It seems to be in accord with poor ability of pA₂₅B₂₅A to activate RNase L (EC₅₀ > 530 nM; Crel 150x [24]).

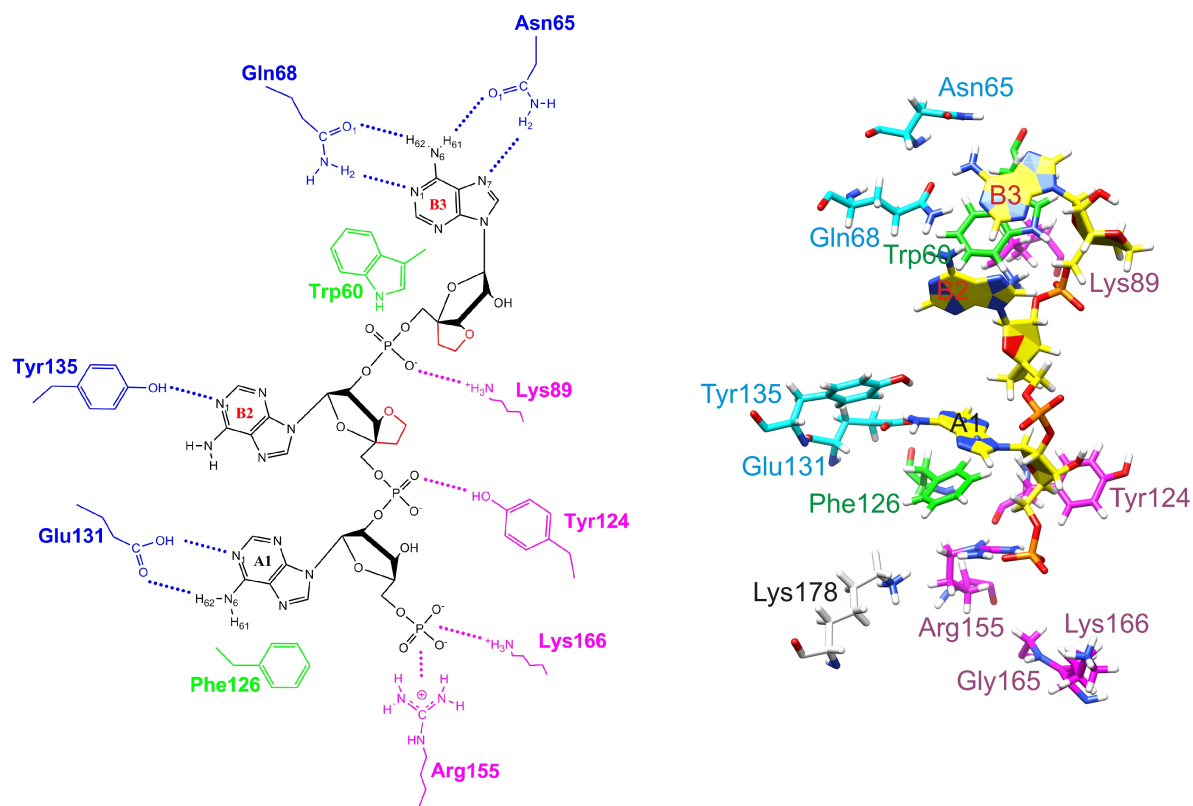


Figure S17: The pA₂₅B₂₅B molecule was bound to the surface of RNaseL_E131H. Ade1 was stacked with the side chain of Phe126. Ade3 was stacked with the side chain of Trp60. Nevertheless, the second AMP significantly changed its conformation. As a consequence, all bases were separated from each other. It's in agreement with an apparent inability of pA₂₅B₂₅B to activate RNase L (EC₅₀ > 1000 nM; Crel >300x [24]).

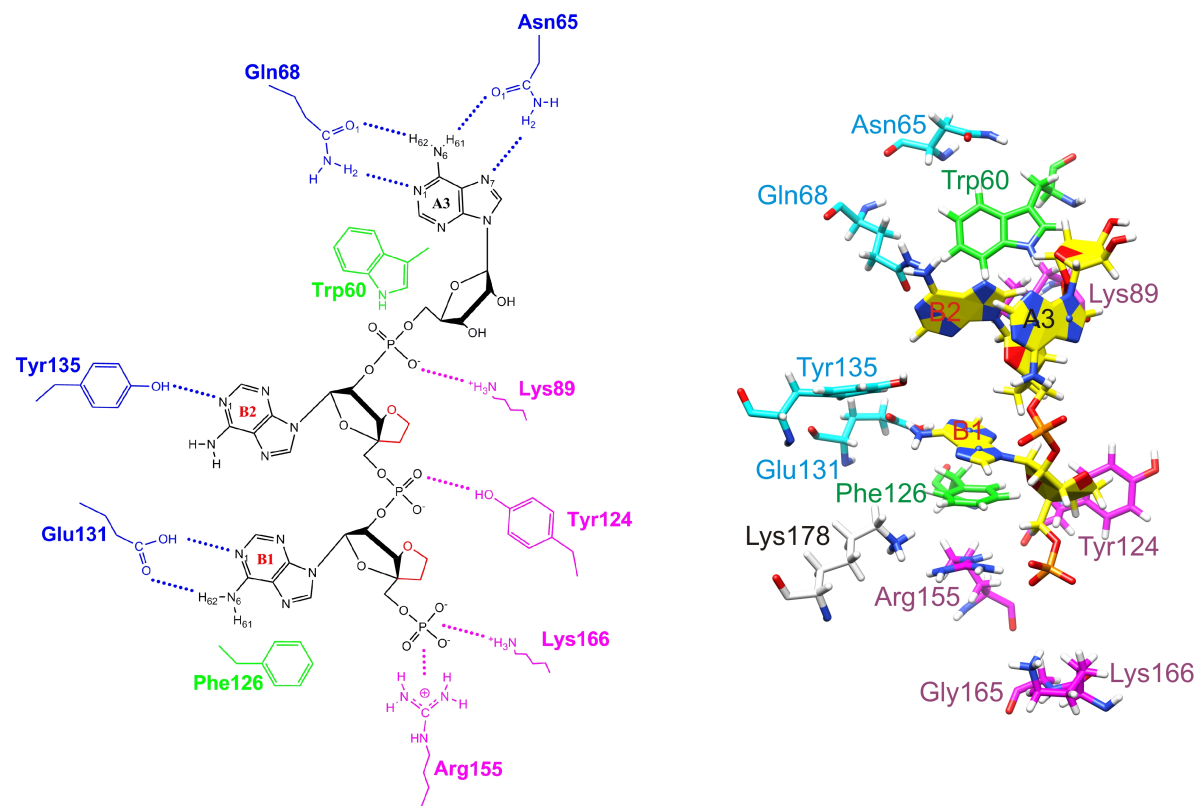


Figure S18: The pB₂₅B₂₅A molecule changed substantially its conformation - just only Ade1 stayed in its original position on the surface of RNase L_{E131H}. It's in accordance with inability of pB₂₅B₂₅A to activate RNase L (EC₅₀ > 1000 nM; Crel >300x [24]).

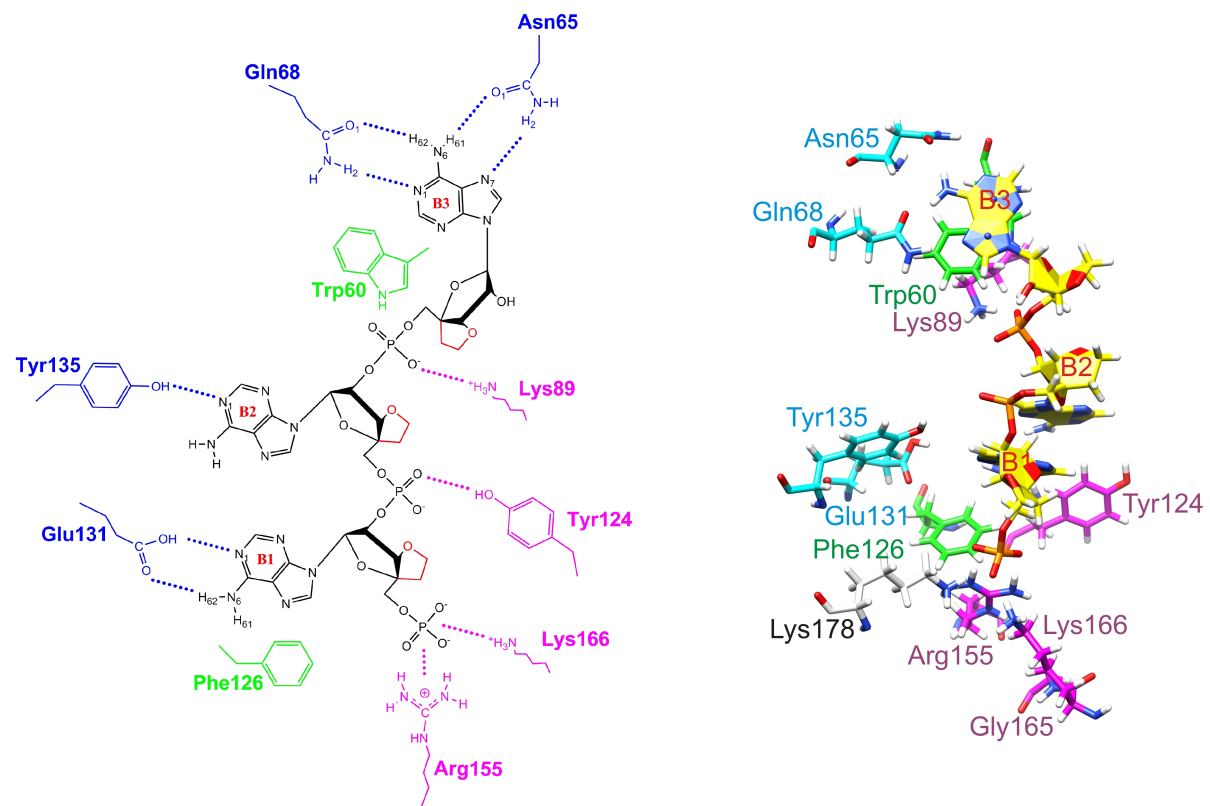


Figure S19: The pB₂₅B₂₅B molecule was remarkably rearranged within this MD run. Finally, it was bound to the surface of RNaseL_E131H only through Ade3. It's in agreement with inability of pB₂₅B₂₅B to activate RNase L (EC₅₀ > 1000 nM; CreI >300x [24]).