

UNIVERZITA PALACKÉHO V OLMOUCI

Přírodovědecká fakulta

Katedra fyzikální chemie



Neobvyklé konformační stavy DNA

Bakalářská práce

Autor:

Dokládalová Kateřina

Studijní program:

B1407 Chemie

Studijní obor:

Aplikovaná chemie

Forma studia:

Prezenční

Vedoucí práce:

doc. RNDr. Jurečka Petr, Ph.D.

Olomouc 2020

Prohlašuji, že jsem tuto práci vypracovala samostatně po vedením doc. RNDr. Petra Jurečky, Ph.D. Veškeré literární zdroje, které byly použity při psaní této práce jsou uvedeny v seznamu použité literatury.

Souhlasím s tím, aby práce byla zpřístupněna v knihovně Katedry fyzikální chemie, Přírodovědecké fakulty, Univerzity Palackého.

V Olomouci

.....

Dokládalová Kateřina

Děkuji doc. RNDr. Petru Jurečkovi, Ph.D. za drahocenné odborné rady a připomínky, které mi poskytl během vypracování mé bakalářské práce. Dále děkuji za neocenitelnou ochotu, kterou mi po celou dobu věnoval. Rovněž bych chtěla poděkovat Mgr. Marii Zgarbové, Ph.D. za pomoc při tvoření mé práce.

Bibliografická identifikace

Autor:	Dokládalová Kateřina
Název práce:	Neobvyklé konformační stavy DNA
Typ práce:	Bakalářská
Pracoviště:	Katedra fyzikální chemie, Přírodovědecká fakulta UP
Vedoucí práce:	doc. RNDr. Jurečka Petr, Ph.D.
Rok obhajoby:	2020
Abstrakt:	<p>Cílem této bakalářské práce je probádat chování nekanonických konformerů DNA v molekulově dynamických simulacích, a to u čtyř struktur – kvadruplexu, Hollidayovy struktury a dvou protein-DNA komplexů. Simulace byly prováděny ve dvou silových polích, OL15 a bsc1, a výsledky byly porovnávány s cílem zjistit, zda jsou v těchto polích nekanonické konformace popsány s dostatečnou přesností, nebo zdali potřebují v budoucnu zpřesnění.</p>
Klíčová slova:	DNA, konformace, torzní úhly, silové pole, molekulová dynamika
Počet stran:	72
Jazyk:	český

Bibliographical identification

Author:	Dokládálová Kateřina
Title:	Unusual conformational states of DNA
Type of thesis:	Bachelor
Department:	Department of physical chemistry, UP
Supervisor:	doc. RNDr. Jurečka Petr, Ph.D.
The year of presentation:	2020
Abstract:	The aim of this bachelor thesis is to investigate the behavior of selected non-canonical DNA conformers in molecular dynamics simulations in four structures – quadruplex, Holliday junction and two protein DNA complexes. The simulations were performed in two force fields, OL15 and bsc1, and the results were compared to see if the non-canonical conformations were described with sufficient accuracy in these fields or if they needed refinement in the future.
Keywords:	DNA, conformation, torsion angles, force field, molecular dynamics
Number of pages:	72
Language:	czech

OBSAH

1	ÚVOD.....	7
2	TEORETICKÁ ČÁST.....	8
2.1	NUKLEOVÉ KYSELINY	8
2.1.1	STRUKTURA DNA.....	8
2.1.2	PÁROVÁNÍ BÁZÍ.....	10
2.2	PŘEHLED DNA FOREM.....	10
2.2.1	B-DNA	10
2.2.2	A-DNA	12
2.2.3	Z-DNA FORMA.....	13
2.2.4	TROJITÁ ŠROUBOVICE DNA	14
2.3	KLASIFIKACE KONFORMACÍ PÁTEŘE	14
2.3.1	KVADRUPLEXY.....	17
2.3.2	HOLLIDAYOVA STRUKTURA	19
2.3.3	PROTEIN-DNA KOMPLEXY	20
2.4	TEORETICKÁ CHEMIE.....	23
2.4.1	MOLEKULOVÁ MECHANIKA A DYNAMIKA.....	23
2.4.2	SILOVÁ POLE.....	25
3	EXPERIMENTÁLNÍ ČÁST	27
3.1	CÍLE PRÁCE.....	27
3.2	METODY	27
3.2.1	VYBRANÉ STRUKTURY A PŘÍPRAVA	27
3.3	VÝSLEDKY A DISKUZE.....	29
3.3.1	STABILITA A DYNAMIKA STUDOVANÝCH MOLEKUL V MD SIMULACÍCH.....	29
3.3.2	STABILITA NEKANONICKÝCH STAVŮ PÁTEŘE β a γ	37
3.3.3	POPIS KONFORMACE DEOXYRIBÓZY	58
4	ZÁVĚR.....	62
5	SUMMARY	64
6	LITERATURA	66
	SEZNAM GRAFŮ.....	69
	SEZNAM OBRÁZKŮ	71
	SEZNAM TABULEK.....	72

1 ÚVOD

V oboru chemie vše souvisí se vším, vše na sebe navazuje. Na základech se v chemii staví a tudíž bez základních znalostí člověk není schopen chemii zcela porozumět. Její krása, ale i příkořím, která chemie často nabízí. Souvislosti, kterých je v oboru chemie nespočet, nám dávají možnost nahlížet na obor chemie jako na obor, kde lze předvídat, jak budou reakce probíhat, jaké budou jejich mechanismy a jaké výsledky mohou být očekávány. Někdy se ale stává, že opak je pravdou. Obor chemie je obor nepředvídatelný a to s sebou přináší jisté kouzlo a půvab.

Příkladem mohou být některé DNA molekuly, které jsou velice vyjímečné a mají jedinečné struktury

To, že je DNA nositelkou genetické informace je známo již velmi dlouho. Někomu také může vytanout na mysli, jak taková molekula DNA vypadá. Nejznámější DNA strukturou je určitě pravotočivá dvojitá šroubovice. DNA ale může mít více různých forem, často velmi zajímavých.

O tyto další formy DNA se v posledních letech zajímá mnoho teoretických chemiků, kteří za pomoci výpočetní chemie simulují různé systémy a prohlubují tak znalosti o nekanonických formách DNA.

2 TEORETICKÁ ČÁST

2.1 NUKLEOVÉ KYSELINY

Nukleové kyseliny byly poprvé objeveny v roce 1868 Friedrichem Miescherem. Byly nalezeny v biologickém materiálu, a to v hnisu či spermích ryb. [1] [2]

Na základě toho, jaký cukr je obsažen, rozeznáváme dva typy nukleových kyselin. Ribonukleovou kyselinu (RNA), která obsahuje D-ribosu a deoxyribonukleovou kyselinu (DNA), ve které je obsažena 2-deoxy-D-ribosa. [2]

Nukleové kyseliny jsou biologické makromolekuly, které můžeme nalézt ve všech buňkách i virech. Jsou jejich nedílnou součástí. DNA je nositelkou genetické informace, a bez RNA by nedocházelo k jejímu přepisu. DNA uchovává veškeré informace např. jak se buňka chová, jaký má charakter, vlastnosti i to, jak se buňka bude dále dělit a růst.

Nukleové kyseliny nazýváme také biopolymery, podobně jako je tomu například u bílkovin. Nukleové kyseliny mají své stavební jednotky, kterým se říká nukleotidy, ty pak vytvářejí dlouhý řetězec. Tyto nukleotidy jsou složeny z nukleosidu, což je dusíkatá báze vázaná s aldopentosou, který je vázán na fosfátovou skupinu.

U eukaryotních organismů se DNA nachází v buněčném jádře či v mitochondriích a chloroplastech. V případě organismů prokaryotních se DNA nalézá v cytoplazmě.

2.1.1 STRUKTURA DNA

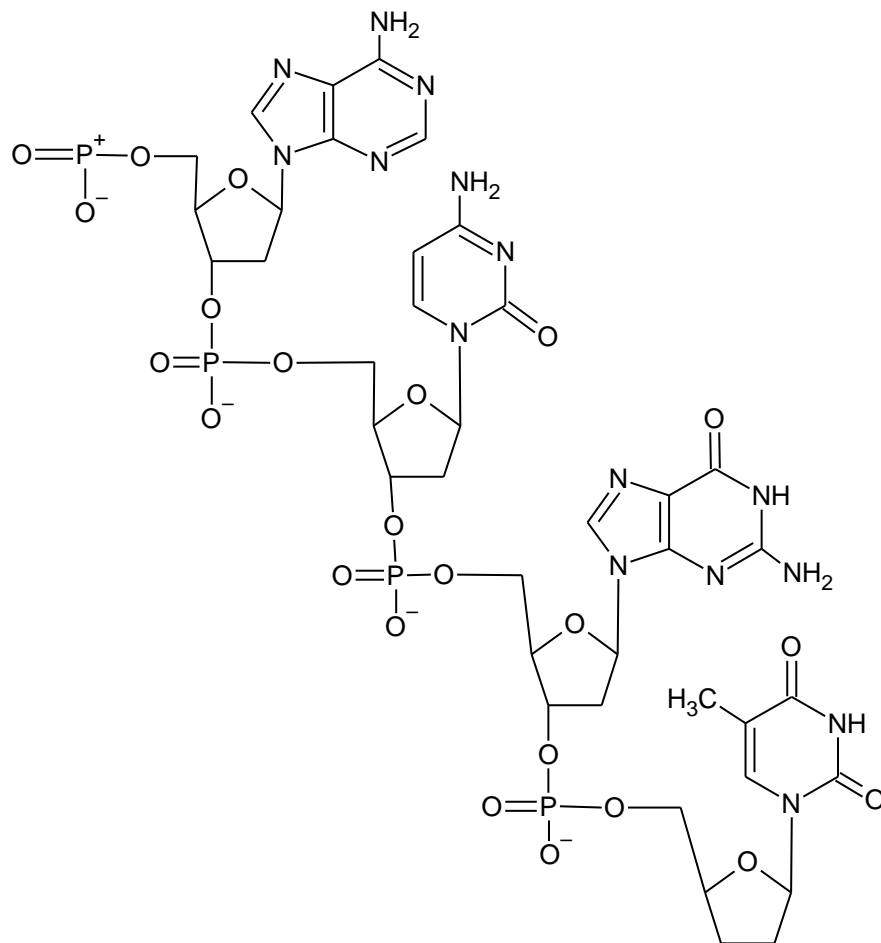
Dvoušroubovice DNA je tvořena dvěma polynukleotidovými vlákny deoxyribonukleové kyseliny s komplementární sekvencí (primární struktura). Tato vlákna mezi sebou interagují a jsou obtočena kolem společné osy. Helikální strukturu vymezuje struktura sekundární, která určuje prostorovou formu stočení.

Nukleotidy, které tvoří řetězce DNA, se skládají z cukru, konkrétně z deoxyribózy, dále z fosfátové skupiny a nukleové báze. Spojením těchto tří částí vzniká nukleotid. Polynukleotidový řetězec DNA je vyobrazen na Obrázku 1.

Do řetězce jsou nukleotidy spojeny za pomoci kovalentních vazeb, konkrétně s pomocí 3',5' diesterové vazby. Čísla 3' a 5' označují čísla atomů uhlíku na deoxyribóze, na které se pak vážou fosfátové skupiny.

Dusíkaté báze jsou pak ve dvoušroubovici DNA párovány na základě komplementarity bází.

Obrázek 1. Struktura polynukleotidového řetězce DNA (vytvořeno pomocí programu ChemSketch)

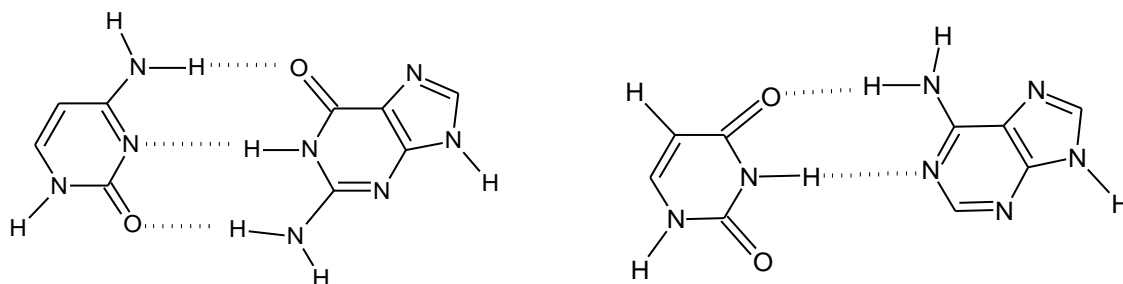


2.1.2 PÁROVÁNÍ BÁZÍ

DNA molekula se skládá ze dvou polydeoxyribonukleových řetězců. Za pomoci vodíkových vazeb jsou nukleové báze jednoho řetězce spojeny s komplementární bází v řetězci druhém. Nukleovou bází může být adenin (A, Ade), thymin (T, Thy), cytosin (C, Cyt) nebo guanin (G, Gua). Adenin s guaninem řadíme mezi takzvané puriny a cytosin s thyminem mezi pyrimidiny.

Na Obrázku 2 je znázorněno párování bází. Adenin je komplementární s thyminem a jsou spojeny dvěma vodíkovými vazbami. V případě guaninu, který se páruje s cytosinem jsou tyto báze spojeny třemi vodíkovými vazbami. Vlákná, která jsou navzájem propojena vodíkovými vazbami a tvoří tak dvoušroubovici DNA, jsou antiparalelní, což znamená, že každé vlákno směřuje opačným směrem. Směr vlákna je definován od 5' konce k 3' konci.

Obrázek 2. Párování bází – cytosin s guaninem (vlevo), thymin s adeninem (vpravo) (vytvořeno pomocí programu ChemSketch)



2.2 PŘEHLED DNA FOREM

2.2.1 B-DNA

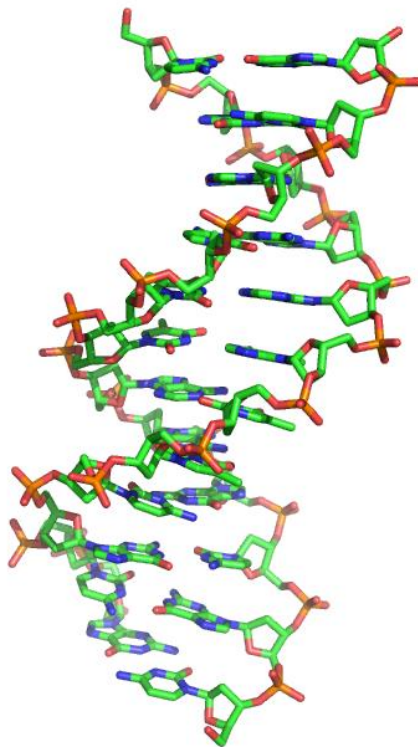
DNA se nejčastěji vyskytuje ve formě zvané B-DNA, kterou Crick a Watson popsali již v roce 1953. [3] Tato forma se někdy nazývá „kanonickou“ formou a je charakterizována tvarem ukázaným na Obrázku 3 a charakteristickými hodnotami páteřních úhlů (viz níže). Každý fosfát v páteři B-DNA se navíc může vyskytovat ve dvou konformacích, nazývaných BI a BII. Tyto dvě konformace B-DNA se od sebe odlišují hodnotami svých torzních úhlů. Existuje hypotéza, že konformace BII je klíčová pro rozpoznávání protein-DNA komplexů.[4] Ukázalo se, že výskyt této konformace závisí

na sekvenci. Zatímco dinukleotidy GC, CG, CA, TG a TA tvoří BII konformaci ochotně, dinukleotidy TT, TC, CT a CC tuto BII konformaci tvoří zřídka. [3] [4]

B-DNA forma není ale jedinou možnou formou helixu. Existují i další způsoby, jak mohou být řetězce uspořádány.

K dalším modelům pak patří například Z-DNA nebo A-DNA. A-DNA forma je stejně jako B-DNA forma pravotočivá, na rozdíl od Z-DNA, která je levotočivá.

Obrázek 3. Struktura B-DNA (vytvořeno pomocí programu PyMOL)

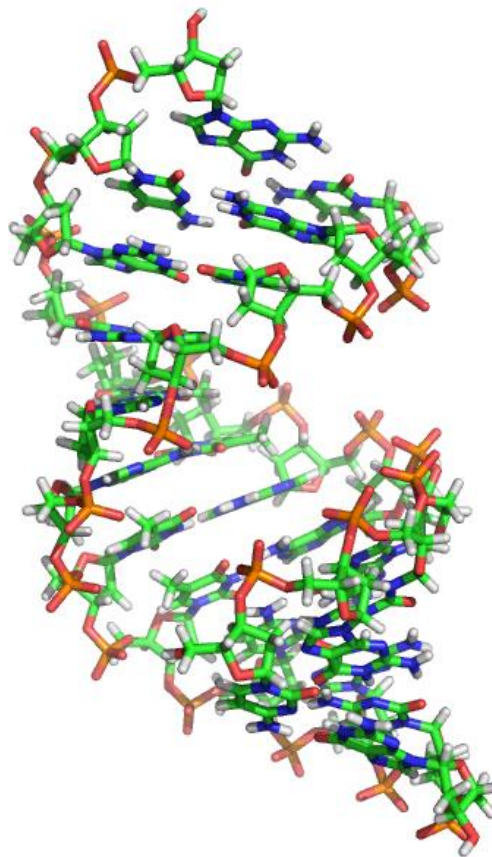


2.2.2 A-DNA

A-DNA je s B-DNA shodná v tom, že je také pravotočivá. Rozdíl mezi těmito dvěma strukturami však spočívá v tom, že A-DNA je širší a má plochý tvar (viz Obrázek 4). To je způsobeno rozdíly v úhlech páteře (viz níže).

Tato forma se tvoří ve vláknech DNA s náhodnou sekvencí za podmínek nízké vlhkosti nebo v roztoku, kde je aktivita vody redukována přidáním různých typů alkoholů. [3]

Obrázek 4. Struktura A-DNA (vytvořeno pomocí programu PyMOL)



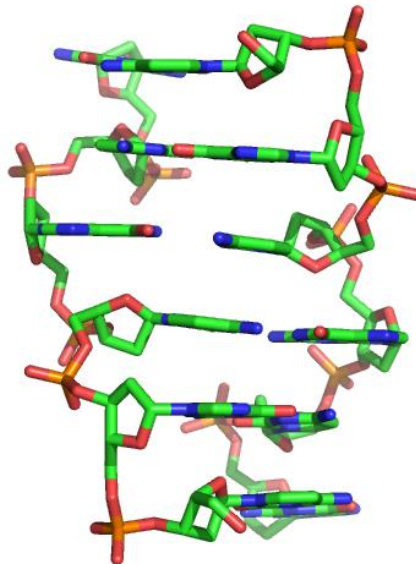
2.2.3 Z-DNA FORMA

Z-DNA molekula je na rozdíl od B-DNA levotočivá (viz Obrázek 5).

První oligonukleotidová duplexní monokrystalová struktura, která byla objevena v roce 1979, [3] byla sekvence d(CGCGCG). Tato struktura je velice zajímavá, a to díky tomu, že se skládá z duplexu, který je tvořen dvěma antiparalelními vlákny dvoušroubovice a je levotočivá. [3]

Páteř Z-DNA je ve srovnání s B-DNA či A-DNA nepravidelná, a to proto, že báze d(G) a d(C) mají odlišnou konformaci. [3]

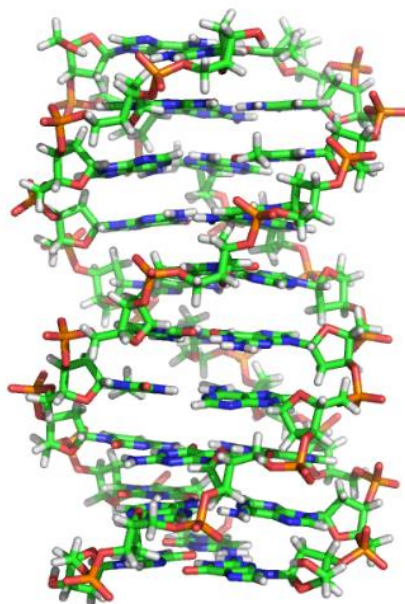
Obrázek 5. Struktura Z-DNA (vytvořeno pomocí programu PyMOL)



2.2.4 TROJITÁ ŠROUBOVICE DNA

Mezi další formy DNA mj. patří trojitá šroubovice (viz Obrázek 6), která byla objevena v roce 1957. [3] Trojitá šroubovice je pravotočivá a skládá se např. ze dvou pyrimidinových řetězců a jednoho purinového řetězce. Tato trojitá šroubovice vzniká z dvojité šroubovice, kdy se naváže další vlákno pomocí Hoogstenových vazeb. [3] Toto vlákno může pocházet buď z původní šroubovice nebo z jiné. V případě, že by vlákno pocházelo z původní šroubovice, pak by se šroubovice rozpojila. Jedno vlákno vytvoří s druhou částí šroubovice triplex a druhé vlákno zůstane samotné. [3]

Obrázek 6. Struktura triplexu (vytvořeno pomocí programu PyMOL)



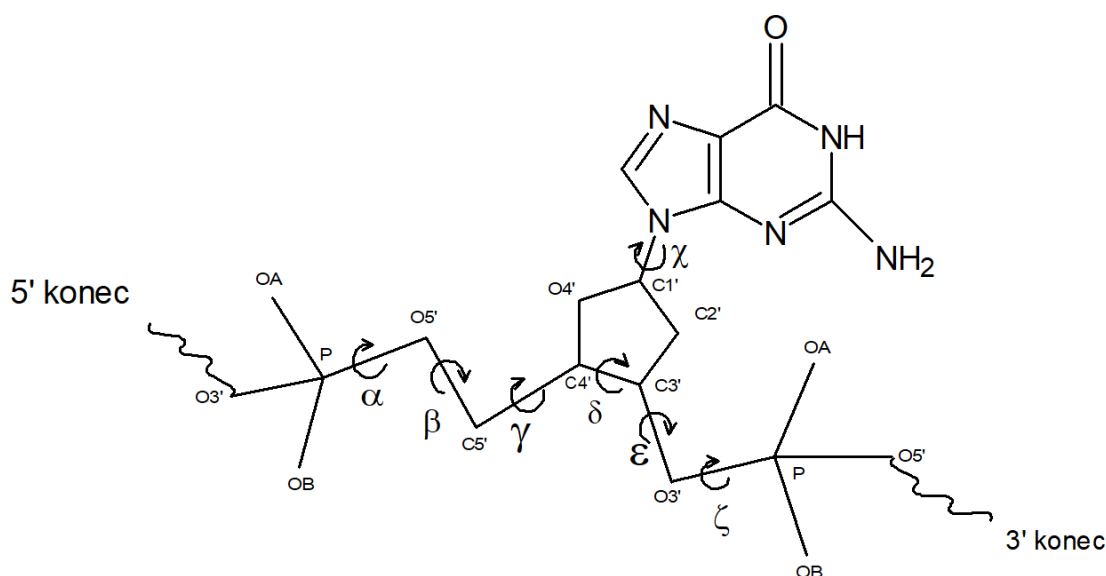
2.3 KLASIFIKACE KONFORMACÍ PÁTEŘE

DNA je sama o sobě velmi variabilní. Mnohotvárnost DNA dokazuje zvyšující se počet nalezených nekanonických krystalových a NMR struktur. Pro úplné pochopení polymorfismu DNA bylo nutné provést analýzu páteře DNA za účelem prohloubení znalostí o konformačním chování DNA. [5] Detailní klasifikace DNA konformací se postupem času vyvíjí a stává se preciznější. Díky tomu přibývají nové klasifikované konformery. Zatímco dříve byly konformery nalézány a přiřazovány manuálně, v nedávné době byla vyvinuta automatická klasifikační metoda založená na použití umělé inteligence (artificial intelligence, machine learning) a neuronových sítí. [5]

Pomocí této metody bylo v roce 2013 identifikováno šest nových konformerů. Čtyři z těchto konformerů byly přiřazeny do dvou funkčně odlišných typů DNA šroubovice, do dvou velice důležitých DNA skupin. Patří sem guaninové kvadruplexy a Hollidayovy struktury. Studium dvou z těchto nově objevených konformerů bude tématem této práce.

Konformace páteře DNA jsou nejčastěji popisovány pomocí torzních (dihedrálních) úhlů. Torzní úhly cukrfosfátové páteře jsou definovány každý pomocí čtyř atomů (viz Obrázek 7 a Tabulka 1). V případě glykosidické vazby torze χ však nastávají dva případy. Pro puriny je tato torze definována jako O4' – C1' – N9 – C4, pro pyrimidiny jako O4' – C1' – N1 – C2.

Obrázek 7. Nukleotidová jednotka (vytvořeno pomocí programu ChemSketch)



Tabulka 1. Definice torzních úhlů (tabulka převzata a upravena) [4]

Torzní úhel	Definice úhlu
α	O3' – P – O5' – C5'
β	P – O5' – C5' – C4'
γ	O5' – C5' – C4' – C3'
δ	C5' – C4' – C3' – O3'
ϵ	C4' – C3' – O3' – P
ζ	C3' – O3' – P – O5'

Práce se bude věnovat hlavně dvěma konformerům označeným čísly 35 a 117. [5] Konformery 35 a 117 se vyskytují např. v různých protein-DNA komplexech, ale také v některých kvadruplexech nebo Hollidayově struktuře. A právě tyto struktury, tedy zmíněné kvadruplexy, Hollidayova struktura a vybrané protein-DNA komplexy, budou v práci zkoumány. V Tabulce 2 se nachází hodnoty dihedrálních úhlů konformerů 35 a 117 ve srovnání s hlavními formami DNA.

Tabulka 2. Hodnoty dihedrálních úhlů nekanonických konformací (převzato a upraveno) [5]

Konformace	Velikosti úhlů [°]								
	δ	ε	ζ	$\alpha+1$	$\beta+1$	$\gamma+1$	$\delta+1$	χ	$\chi+1$
Kanonická B forma	136 (S)	183 (t)	259 (g ⁻)	303 (g ⁻)	181 (t)	44 (g ⁺)	138 (S)	252 (high- anti)	259 (high- anti)
A forma	83 (N)	205 (t)	287 (g ⁻)	294 (g ⁻)	174 (t)	54 (g ⁺)	83 (N)	199 (anti)	202 (anti)
Konformer 35	136 (S)	199 (t)	288 (g ⁻)	253 (g ⁻)	73 (g ⁺)	168 (t)	87 (N)	264 (high- anti)	187 (anti)
Konformer 117	139 (S)	196 (t)	286 (g ⁻)	249 (g ⁻)	73 (g ⁺)	172 (t)	145 (S)	263 (high- anti)	211 (anti)

Běžnou konvencí pro popis torzních úhlů je pojmenování minim potenciálu za pomoci speciálních termínů. Jsou to např. *gauche*⁺ (g⁺) pro hodnoty ~ 60°, *gauche*⁻ (g⁻) pro hodnoty ~ - 60° (300°) a *trans* (t) pro ~ 180°. Některé úhly se mohou vyskytovat ve více než jedné konformaci a tak třeba úhel α či úhel γ mohou existovat v konformacích g⁺, g⁻ nebo t. [3] Konformace deoxyribózy je popisována s pomocí pseudorotačního úhlu P (pucker). Nejčastějšími konformacemi jsou jižní (south, P ~ 160°), označovaná S, a severní (north, P ~ 18°), označovaná N (viz Tabulka 2). Glykosidický úhel má minima potenciálu označena *syn* (~ 60°), *anti* (~180°) a někdy *high-anti* (~260°).

Jedním z cílů této práce je prozkoumat chování vybraných nekanonických konformerů DNA v molekulově dynamických (MD) simulacích a jejich stabilitu

v různých empirických potenciálech (viz níže). Konformery 117 a 35 byly vybrány hlavně proto, že obsahují úhel γ v nekanonické oblasti $\gamma \sim 180^\circ$. Právě u této konformace bylo nedávno ukázáno, že je pravděpodobně nesprávně popsána v současných silových polích. [6]

Z databáze krystalových struktur analyzovaných v článku Čech et. al. [5] byly proto vybrány čtyři významné struktury, které tuto nekanonickou konformaci obsahují - guaninový kvadruplex, Hollidayova struktura a dva komplexy DNA s proteinem. Tyto struktury budou dále popsány detailněji.

2.3.1 KVADRUPLEXY

Kvadruplexy jsou složeny z guaninových tetrad. Tuto guaninovou tetradu tvoří čtyři guaniny, jež jsou cyklicky uspořádané, ty jsou pak spojeny vodíkovými vazbami.

Ve středu každé tetrady se nachází volný prostor se záporným potenciálem, který je zde díky orientaci karbonylové skupiny od oblasti volného prostoru. Tento volný prostor tvoří při navrstvení těchto tetrad tzv. kanál, který je zaplněn jednomocnými ionty, sodnými či draselnými. Rozdíl spočívá v tom, že sodné ionty se nachází v rovině tetrady, na rozdíl od draselných iontů, které se nachází v prostoru mezi jednotlivými tetradami. [7]

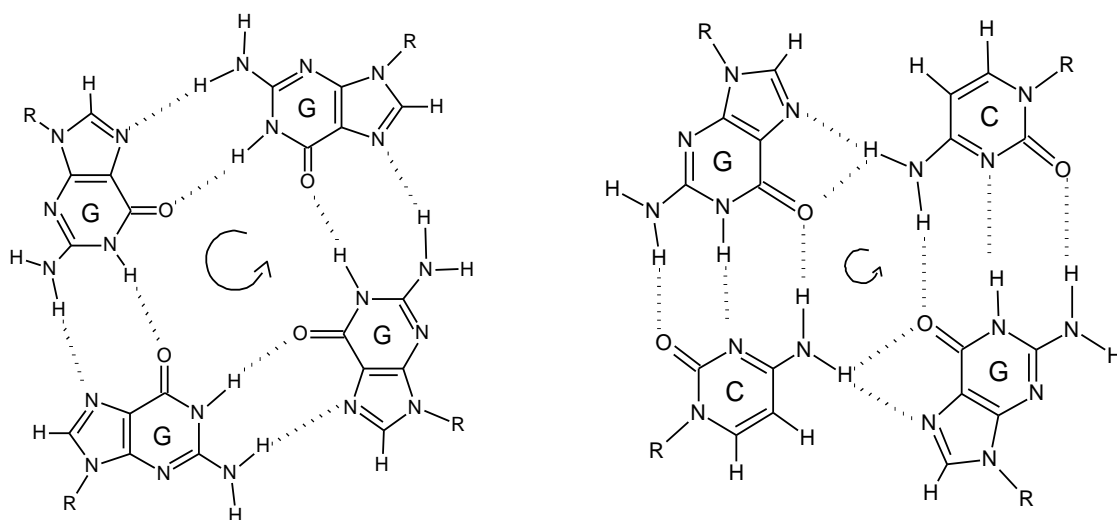
Smyčky (loops) jsou takzvané spojovací segmenty, které musí obsahovat bimolekulární kvadruplexy. Smyčky umožňují zapojení dalších sekvencí guaninů do struktury kvadruplexu.

Jednotlivé smyčky jsou tvořeny thyminy, ale může se zde vyskytnout například i adenin. Smyčky můžeme rozdělit na diagonální a laterální. Diagonální mohou spojit dva protější řetězce. Orientace těchto řetězců je antiparalelní. Tady smyčka přechází přes oblast ústí centrálního kanálu. Laterální smyčka také spojuje dva sousední řetězce, jež jsou stejně jako v předchozím případě antiparalelní a vytváří struktury známé jako vlásenka (hairpin). Externí smyčka taktéž spojuje dva sousední řetězce, ale jsou paralelní a smyčka je zde vedena přes boční část molekuly v oblasti žlábků. Externí smyčka se vyskytuje především v bimolekulárních kvadruplexech. [7] [8]

Kvadruplexy vznikají v průběhu replikace a transkripce, kdy dochází k rozpletení dvoušroubovice. Vzhledem k velké stabilitě kvadruplexů je nelze vrátit do původního stavu, protože by to bylo velmi energeticky náročné. [8]

Pro účely této bakalářské práce jsou kvadruplexy zajímavé tím, že jejich páteř často obsahuje zajímavé nekanonické konformace. K dalšímu zkoumání byl vybrán kvadruplex (PDB ID 1A6H). V tomto případě se kvadruplex vyskytuje s ionty sodnými. Další jeho zajímavostí je to, že struktura obsahuje nejen tetrády (G·G·G·G), ale i tetrády obsahující cytosin (viz Obrázek 8).

Obrázek 8. Guaninová tetráda G·G·G·G (vlevo) a tetráda obsahující cytosin C·G·C·G (vpravo) obrázek převzat a upraven v programu ChemSketch) [9]



2.3.2 HOLLIDAYOVA STRUKTURA

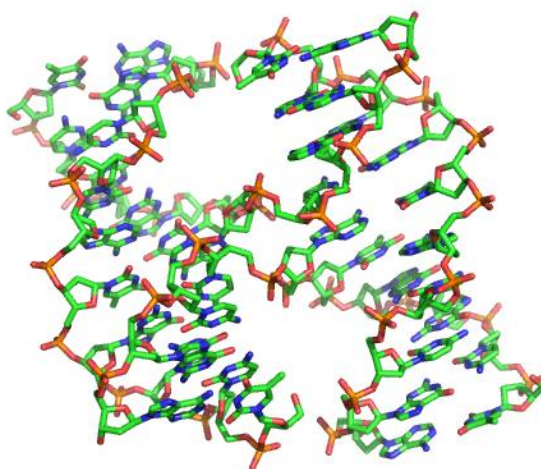
Jak již bylo výše zmíněno, konkrétně konformer 117 se může vyskytovat v Hollidayově struktuře, jinak zvané také Hollidayův spoj.

Do procesu genetické rekombinace je zahrnuto křížení řetězců ze dvou sekvencí DNA. Klíčovým meziproduktem v tomto procesu je právě Hollidayův spoj. Skládá se ze čtyř řetězců, z nichž každý je součástí čtyř antiparalelních párů (viz Obrázek 9). Pro definování trojrozměrné struktury Hollidayova spoje bylo použito mnoho biofyzikálních technik, jako je například gelová elektroforéza. [3]

Studie, která se zabývala vlivem trinukleotidů v sekvenci (CC_nnnN₆N₇N₈GG) se snažila určit krystalovou strukturu 64 permutací, kde N₆, N₇ a N₈ jsou libovolné volně se vyskytující nukleotidy. Ve studii bylo zjištěno, že Hollidayovu strukturu tvoří trinukleotidy ACC, GCC a ATC. Pouze trinukleotid ACC tvoří čistě Holliday junction, zbylé dva mohou vytvořit A-formu i B-formu DNA. Společným rysem pro tyto trinukleotidy je cytosin v poloze N₈. [9]

Ne všechny trinukleotidy NNC tvoří Hollidayovu strukturu, tudíž se cytosin sice jeví jako nezbytný pro tvorbu Holliday junction, ale jeho přítomnost není dostačující. [10] V této práci byla jako výchozí struktura použita Hollidayova struktura (CCGGTACCGA) s PDB ID 1NVY a rozlišením 1,5 Å, která byla zatím publikována pouze na stránkách PDB databáze.

Obrázek 9. Hollidayova struktura (vytvořeno pomocí programu PyMOL)



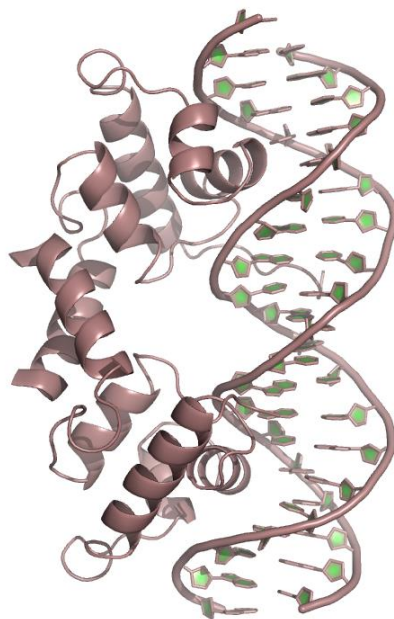
2.3.3 PROTEIN-DNA KOMPLEXY

Komplexy DNA s proteiny jsou mimořádně důležité pro mnoho biologických procesů, jako jsou např. transkripce, replikace, translace a další. DNA v komplexu často podléhá konformačním změnám (přízpusobení), které zahrnují změnu konformace deoxyribózy z S (běžná pro B-DNA) směrem k N (běžná pro A-DNA). Někdy zřejmě ale také dochází k rekonformaci dalších páteřních úhlů. V této práci se budu zajímat hlavně o změny úhlů β a γ v konformeru č. 35 (viz výše). [11]

První zmíněný protein DNA komplex je λ represor – operátor (Obrázek 10), jehož krystalografická studie s rozlišením 1.8 Å detailně popsala nové protein-DNA interakce [12]. λ represor, je sám o sobě členem obrovské skupiny vazebných proteinů prokaryotické DNA, která je charakterizována strukturálním motivem helix – turn – helix (HTH). [12] Represor používá jak konzervovaný motiv HTH, tak i flexibilní tzv. N rameno k rozpoznávání párových bází v hlavních žlábcích DNA. Společně pak rezidua z HTH a N ramena tvoří specifická místa pro interakce vodíkovými vazbami s páry bází. Také mnoho dalších párů bází v tomto komplexu se podílí na van der Waalsových nebo hydrofobních interakcích s proteinem.[12]

Zajímavostí je také to, že vodíkové vazby s páry bází ve velkých žlábcích zahrnují také interakce, které jsou zprostředkovány vodou a fosfáty páteře. [12] [13] Pro tuto práci je důležitá přítomnost β/γ konformeru č. 35 mezi guaniny G13 a G14, které jsou v tomto komplexu v přímém kontaktu s proteinem.

Obrázek 10. Struktura λ represor opeátor protein-DNA komplexu (vytvořeno pomocí programu PyMOL)



Dalším zmíněným protein-DNA komplexem je RAP1 (repressor aktivátor protein 1, Obrázek 11). RAP1 je multifunkční protein, který se také podílí na regulaci transkripce, ale kromě toho je jeho pomocí vázána telomerová DNA v kvasinkách *Saccharomyces cerevisiae*. [13] RAP1 je velký protein, který se skládá z 827 aminokyselinových reziduí. Ty ale obsahují odlišné funkční domény. Například minimální doména DBD byla ve studii definována deleční analýzou [13] a nachází se uprostřed proteinu. Tato doména je až nezvykle velká, avšak nese žádnou významnou sekvenční shodu s žádným ze známých motivů, které se vážou v DNA. Oproti tomu tzv. C-terminální doména je nezbytná pro transaktivaci. [13] Také tento komplex obsahuje β/γ konformer č. 35 u bází, které jsou v přímém kontaktu s proteinem.

Obrázek 11. Struktura RAP1 protein-DNA komplexu (vytvořeno pomocí programu PyMOL)



2.4 TEORETICKÁ CHEMIE

Teoretická neboli výpočetní chemie je odvětví chemie, které umožňuje získávání dat fyzikální i chemické povahy za pomoci teoretických metod, pod které spadá molekulární modelování, kvantově-chemické výpočty a počítačové simulace. Časté jsou aplikace počítačového modelování na systémy jako jsou bílkoviny, DNA i RNA, polymery či molekuly organického původu či anorganické povahy. Díky výpočetní chemii se dá zjistit například to, jakou má molekula energii či strukturu. Dále také to, jaké jsou vazebné energie či molekulové orbitály a mnoho dalších užitečných informací. [14]

2.4.1 MOLEKULOVÁ MECHANIKA A DYNAMIKA

Molekulárně-mechanická molekulová dynamika je výpočetně relativně nenáročná a díky tomu je možno ji aplikovat na studium relativně velkých systémů, jakými jsou např. nukleové kyseliny.

Molekulová dynamika je jedna z metod, která se zabývá tím, jak se molekula pohybuje a vyvíjí v čase. K propagaci souřadnic se v molekulové dynamice využívá newtonovská mechanika. Díky této metodě je možné porozumět a předvídat chování molekul. Pro popis molekul se využívá klasická mechanika a slouží k popisu malých molekul až po molekuly s miliony atomů. [14]

Molekulová mechanika využívá metody silového pole neboli empirického potenciálu. V molekulové mechanice je molekula popisována jako seskupení, které je tvořeno bodovými náboji a vdW kuličkami (ty mají reprezentovat atomy) a vazbami mezi nimi. Vazby jsou v molekulové mechanice chápány jako pružiny, které mají svoji určitou délku a tuhost. [14]

Rovnice 1 popisuje vztah pohybu a vývoje studovaného systému v závislosti na čase.

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} = m \frac{d^2 r_i}{dt^2} \quad [\text{Rovnice 1}]$$

F je síla, působící na částici, m je její hmotnost a a je zrychlení, tedy derivace rychlosti v podle času t . Pro výpočet sil a energie celého systému se musí nejprve spočítat veškeré vazebné i nevazebné interakce. Vazebnými interakcemi se myslí například délka vazby (E_{vazba}), vazební úhly ($E_{\text{úhel}}$) nebo torzní úhly ($E_{\text{torzní úhel}}$). Mezi ty nevazebné jsou zahrnuty

van der Wallsovy síly (E_{vdw}) a Coulombické síly neboli elektrostatické síly (E_C) (viz Rovnice 2).

Aby bylo možné energetické složky spočítat, využívá se empirického potenciálu neboli silového pole. [14]

$$E_{pot} = E_{vazba} + E_{úhel} + E_{torzní\ úhel} + E_{vdw} + E_C \quad [\text{Rovnice 2}]$$

2.4.1.1 VAZEBNÉ A NEVAZEBNÉ INTERAKCE

Výpočet jednotlivých energií vazebných i nevazebných interakcí je vyobrazen v následujících rovnicích. Vztah pro výpočet vazebné energie udává Rovnice 3, kde r je vazebná vzdálenost, r_0 je hodnota rovnovážné vzdálenosti a K_r je silová konstanta vyjadřující tuhost vazby.

$$E_{vazba} = \sum K_r (r - r_0)^2 \quad [\text{Rovnice 3}]$$

Pro výpočet energie vazebných úhlů je dán vztah (viz Rovnice 4), kde α je vyjádřením vazebného úhlu, α_0 je rovnovážnou hodnotou vazebného úhlu a K_α je silovou konstantou vazebného úhlu.

$$E_{úhel} = \sum K_\alpha (\alpha - \alpha_0)^2 \quad [\text{Rovnice 4}]$$

Pro torzní úhel existuje vztah vyjádřený v Rovnici 5, kde E_t vyjadřuje torzní příspěvek k potenciální energii, ϕ_{n0} vyjadřuje fázový posun k n -té periodě, ϕ znázorňuje vazebnou torzi, E_n výšku potenciální bariéry v n -té periodě a n je právě perioda jedné harmonické složky.

$$E_{torzní\ úhel} = \sum_t \sum_n \frac{E_n}{2} [1 + \cos(n\phi - \phi_{n0})] \quad [\text{Rovnice 5}]$$

Torzní úhel nabývá hodnot od 0° do 360° . Torzní úhly říkají, jak se k sobě orientují jednotlivé segmenty molekul, např. nukleotidů. V nukleových kyselinách se jedná o úhly α , β , γ , δ , ϵ , ζ a χ (viz výše). Tyto úhly nabývají v deoxyribonukleové kyselině charakteristických hodnot.

Rovnice 6 zobrazuje vztah pro výpočet Van der Waalsovy interakční energie, kde ε_{ij} znázorňuje hloubku potenciálové jámy, r_{ij} vyjadřuje vzdálenost mezi dvěma atomy a R_{ij} vyjadřuje rovnovážnou vzdálenost. [15]

$$E_{vdw} = \sum_{i < j} \varepsilon_{ij} \left[\left(\frac{R_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}}{r_{ij}} \right)^6 \right] \quad [\text{Rovnice 6}]$$

Elektrostatická energie lze vyjádřit vzájemným působením dvou parciálních nábojů. Tento vztah vyjadřuje Rovnice 7, kde $q_i q_j$ je součin parciálních nábojů, ε_0 znázorňuje permitivitu vakua a r_{ij} představuje vzdálenost mezi atomy.

$$E_C = \sum_{i < j} \frac{1}{4\pi\varepsilon_0} \cdot \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad [\text{Rovnice 7}]$$

2.4.1.2 EXPLICITNÍ A IMPLICITNÍ ROZPOUŠTĚDLO

Do modelování v molekulové dynamice se často zahrnuje vodné prostředí molekuly, a to buď jako explicitní či implicitní rozpouštědlo. Pokud je použito explicitní rozpouštědlo, tak se v boxu molekula nachází i s molekulami vody. Mezi explicitní rozpouštědla patří například SPC/E [16] nebo TIP3P [17], TIP4P-Ew [18], které se liší svými parametry a chováním. [19] V případě implicitního rozpouštědla se hovoří o tzv. kontinuu, ve kterém dochází k pohybu studovaných systémů. [20]

2.4.2 SILOVÁ POLE

Pro začátek je důležité si vybrat vhodná silová pole pro provedení simulací. Pro studium nukleových kyselin existují různá silová pole. Mezi ty nejpoužívanější patří silová pole AMBER[21] a CHARMM[22]. V této bakalářské práci je použit AMBER (Assisted Model Building with Energy Refinement).

AMBER je programový balík, ve kterém lze použít více variant silových polí, která jsou využívána při studiu nukleových kyselin. Dále se v tomto balíku nachází vše, co je třeba pro realizaci simulací a pro analýzu výsledků. [21]

Za zmínku stojí program LEAP, který se využívá pro přípravu nového systému. Dále programy, které provádějí minimalizace a molekulárně dynamické simulace. Mezi programy, které se využívají pro analýzu výsledků patří program PTRAJ. Díky tomuto programu je možno analyzovat trajektorie či souhrn souřadnic, které byly vytvořeny díky molekulově dynamickým simulacím.

Při studiu nukleových kyselin se nejvíce využívají silová pole, která vycházejí ze silového pole Cornell et al. [23] označovaného zkratkou ff94. Následné úpravy tohoto silového pole vedly ke vzniku novějších verzí silových polí – ff98 [24], ff99 [25]. Postupně se ale ukázalo, že ani v poli ff99 nejsou simulace duplexu DNA na delších časových intervalech stabilní. Z tohoto důvodu byly nutné další úpravy potenciálu. V roce 2007 vznikla nová verze silového pole ff99parmbsc0 [26], ve které se objevily nové parametry pro torze α a γ . V následujících několika letech byly představeny další verze silových polí, jež obsahovaly další změny v torzních potenciálech. Například verze silového pole parmbsc1, která navazovala na verzi parmbsc0 obsahovala modifikace torzí χ , ε a ζ .

Modifikace empirických potenciálů byly vyvíjeny také v České republice, přesněji v Olomouci. Základem pro návrh dihedrálních potenciálů se stala metoda, která zahrnuje konformačně závislý vliv rozpouštědla [27], pomocí které byly postupně vyvinuty korekce úhlů χ pro RNA [28], χ pro DNA [29], ε/ζ pro DNA [30] a β pro DNA [31]. Všechny navržené korekce jsou součástí silového pole OL15 [31]. V Tabulce 3 jsou přehledně uvedeny modifikace jednotlivých silových polí.

Tabulka 3. Přehled silových polí pro DNA

Silové pole	Rok modifikace	Modifikace vzhledem k ff99
ff94 [23]	1995	
ff98 [24]	1999	
ff99 [25]	2000	
ff99parmbsc0 [26]	2007	α/γ
ff99parmbsc0 χ OL4 [29]	2010	$\alpha/\gamma, \chi$
ff99parmbsc0- χ OL4 $\varepsilon\zeta$ OL1 [30]	2013	$\alpha/\gamma, \chi, \varepsilon/\zeta$
ff99parmbsc0- χ OL4 $\varepsilon\zeta$ OL1 β OL1 [31]	2015	$\alpha/\gamma, \chi, \varepsilon/\zeta, \beta$
ff99parmbsc1 [31]	2015	$\alpha/\gamma, \chi, \varepsilon/\zeta$

Přesto, že uvedené úpravy vedly k výrazným zlepšením pro simulace kanonické DNA i nekanonických forem, jako jsou Z-DNA a guaninové kvadruplexy, začíná se ukazovat, že některé torze jsou stále nepřesné, např. α/γ [6]. Studium těchto nepřesností v molekulách DNA je cílem této práce.

3 EXPERIMENTÁLNÍ ČÁST

3.1 CÍLE PRÁCE

Cílem této bakalářské práce bylo vyhodnotit chování nekanonických stavů páteřních úhlů β i γ v MD simulacích několika různých kanonických i nekanonických molekul DNA. Bylo posuzováno, zda jsou nekanonické stavy v simulacích stabilní a zdali hrají v systému strukturní roli, a to v souvislosti s dalšími jevy (rekonformacemi) v molekule, které byly zjišťovány vyhodnocením dalších páteřních úhlů v blízkém okolí studovaných úhlů β i γ . Cílem bylo také porovnat, jak dobře jednotlivá dnes používaná silová pole (OL15 a bsc1) popisují nekanonické stavy páteře.

3.2 METODY

3.2.1 VYBRANÉ STRUKTURY A PŘÍPRAVA

Simulace byly prováděny ve dvou silových polích doporučených pro simulace DNA struktur v programu AMBER, a to konkrétně pro pole OL15 a bsc1. Pro samotnou analýzu byly vybrány čtyři odlišné DNA systémy kvadruplex, Hollidayova struktura a dva protein-DNA komplexy.

Všechny vstupní systémy, které byly použité v této práci je možné nalézt pod svým PDB kódem v proteinové databázi www.rscb.org.

Pro simulaci kvadruplexu byla jako výchozí použita NMR struktura s PDB ID 1A6H (první model ze čtyř obsažených v PDB). Protože NMR struktura neobsahovala kationty v kanálu kvadruplexu, byly dodány 3 kationty Na^+ do prostoru mezi roviny tetrad a po jednom nad a pod krajní roviny tetrad, tedy na místa, kde bývají kationty obvykle v kvadruplexech přítomny. Dále byly dodány kationty Na^+ k neutralizaci náboje páteře DNA a další ionty Na^+ Cl^- , aby výsledná koncentrace v systému byla 0,15M NaCl.

V případě Hollidayovy struktury byla jako výchozí použita krystalová struktura s PDB ID 1NVY s rozlišením 1,50 Å. Ve struktuře bylo ponecháno deset Sr^{2+} ve svých krystalografických polohách a náboj páteře byl neutralizován draselnými ionty K^+ a dále byly přidány ionty K^+ a Cl^- pro dosažení výsledné koncentrace 0,15M KCl. Protože v simulacích docházelo k rozpojování konců helixů a koncové báze exponované do

roztoku způsobovaly v simulaci artefakty (vytvářely stabilní kontakty se sousedním helixem), bylo nutné provést další simulace, u kterých byly použity restry na vodíkové vazby koncových párů bází. Parametry restry byly $r_1 = 2.0 \text{ \AA}$, $r_2 = 2.5 \text{ \AA}$, $r_3 = 3.2 \text{ \AA}$, $r_4 = 5.0 \text{ \AA}$ a silové konstanty $rk_2 = 20.0 \text{ kcal/mol/\AA}^2$, $rk_3 = 30.0 \text{ kcal/mol/\AA}^2$. K dalším analýzám byly použity simulace s těmito restry.

Pro simulaci prvního protein-DNA komplexu byla využita krystalová struktura RAP1 protein-DNA komplexu s PDB ID 1IGN s rozlišením $2,25 \text{ \AA}$. Krystalová struktura obsahovala v krystalografické buňce dva stejné protein-DNA komplexy a k simulaci byla použita pouze první struktura. Obě struktury obsahovaly stejný studovaný nekanonický stav. Krystalová struktura neobsahovala všechna rezidua RAP1 (řetězec byl na třech místech přerušen). Z toho důvodu byl na koncová rezidua v místech přerušení dodán restry na vzdálenost, která byla vzata z krystalové struktury. Parametry restry byly $r_1 = 17.1 \text{ \AA}$, $r_2 = 20.1 \text{ \AA}$, $r_3 = 22.1 \text{ \AA}$, $r_4 = 25.1 \text{ \AA}$ pro Ile 123 a Arg 126, dále $r_1 = 7.5 \text{ \AA}$, $r_2 = 10.5 \text{ \AA}$, $r_3 = 12.5 \text{ \AA}$, $r_4 = 15.5 \text{ \AA}$ pro Glu 177 a Glu 180 a $r_1 = 14.0 \text{ \AA}$, $r_2 = 17.0 \text{ \AA}$, $r_3 = 19.0 \text{ \AA}$, $r_4 = 22.0 \text{ \AA}$ pro Thr 186 a Phe 189. Silové konstanty byly $rk_2 = 10.0 \text{ kcal/mol/\AA}^2$, $rk_3 = 10.0 \text{ kcal/mol/\AA}^2$. Řetězce byly zakončeny rezidui acetátu (ACE) a N-methylaminu (NME). Protonační stavy histidinů byly nejdříve určeny s pomocí programu PROPKA [32] a poté byly v krystalové struktuře vizuálně zkontrolovány přítomnosti vodíkových vazeb odpovídajících předpovězeným protonačním stavům. Všech sedm histidinů bylo dvakrát protonováno (HIP) (pH krystalizačního roztoku bylo 6.0). Náboj páteře byl neutralizován draselnými ionty K^+ a následně byly přidány ionty K^+ a Cl^- pro dosažení výsledné koncentrace $0,15M$ KCl.

Pro simulaci λ represor protein-DNA komplexu byla jako výchozí použita krystalová struktura s PDB ID 1LMB s rozlišením $1,80 \text{ \AA}$. Aby se zabránilo třepení konců duplexů, (fraying) byly u terminálních párů použity restry na vodíkové vazby s parametry $r_1 = 2.0 \text{ \AA}$, $r_2 = 2.5 \text{ \AA}$, $r_3 = 3.2 \text{ \AA}$, $r_4 = 5.0 \text{ \AA}$ a silové konstanty $rk_2 = 20.0 \text{ kcal/mol/\AA}^2$, $rk_3 = 30.0 \text{ kcal/mol/\AA}^2$. Náboj páteře byl neutralizován pomocí draselných iontů K^+ a poté byly přidány ionty K^+ a Cl^- , aby výsledná koncentrace v systému byla $0,15M$ KCl.

Vstupní struktury byly připraveny pomocí programu LEAP z programového balíku AMBER. Simulace byly prováděny kódem PMEMD z programového balíku AMBER za izobaricko-izotermických podmínek (NPT). Teplota byla nastavena na 300 K ($22 \text{ }^\circ\text{C}$)

a tlak na 1 Atm (101,325 kPa). Časový krok byl 4 fs. Pro prodloužení časového kroku byl použit hydrogen mass repartitioning (část hmotnosti těžkých atomů byla přesunuta na vodíkové atomy). Cutoff pro ne vazebné interakce byl 9 Å. Vazby k vodíkovým atomům v molekulách rozpuštěné látky byly zamrazeny metodou SHAKE s výchozím nastavením tolerance na 1×10^{-5} . Ewaldova sumace (PME) byla použita s výchozím nastavením mřížky (1 Å) a tolerance (1×10^{-5}). Souřadnice byly ukládány každých 10 ps. K solvataci molekulárních systémů byl použit vodní box SPC/E. Všechny analýzy byly prováděny programem CPPTRAJ z balíku AMBER 18.

3.3 VÝSLEDKY A DISKUZE

3.3.1 STABILITA A DYNAMIKA STUDOVANÝCH MOLEKUL V MD SIMULACÍCH

3.3.1.1 KVADRUPLEX

Výchozí struktura kvadruplexu s PDB ID 1A6H pochází z NMR experimentu a v databázi byly uloženy čtyři mírně odlišné struktury (modely). Po shlédnutí těchto modelů bylo zjištěno, že hlavní odlišnosti se nacházejí v oblasti smyček tvořených thyminy. V případě prvního modelu jsou thyminové smyčky zčásti exponovány do roztoku a samotnému stemu se příliš neblíží. V druhém případě se thyminové smyčky přibližují stemu. V případě třetího modelu jsou thyminové smyčky opět spíše exponovány do roztoku. Smyčky nemají tendenci se přibližovat ke stemu. Tato situace je na obou koncích velmi podobná. V posledním případě čtvrtého modelu jsou thyminové smyčky spíše nakloněny ke stemu kvadruplexu. Naproti smyčkám je stem kvadruplexu, který je tvořený tetradami guaninu a cytosinu relativně stabilní. Důležité ale je, že na místě, kde se vyskytuje studovaná konformace 117 s $\beta/\gamma = g^+/t$ (C2-G3), se konformace páteře v různých modelech liší (viz Tabulka 4).

Tabulka 4. Konformace páteře ve 4 modelech NMR struktury 1A6H

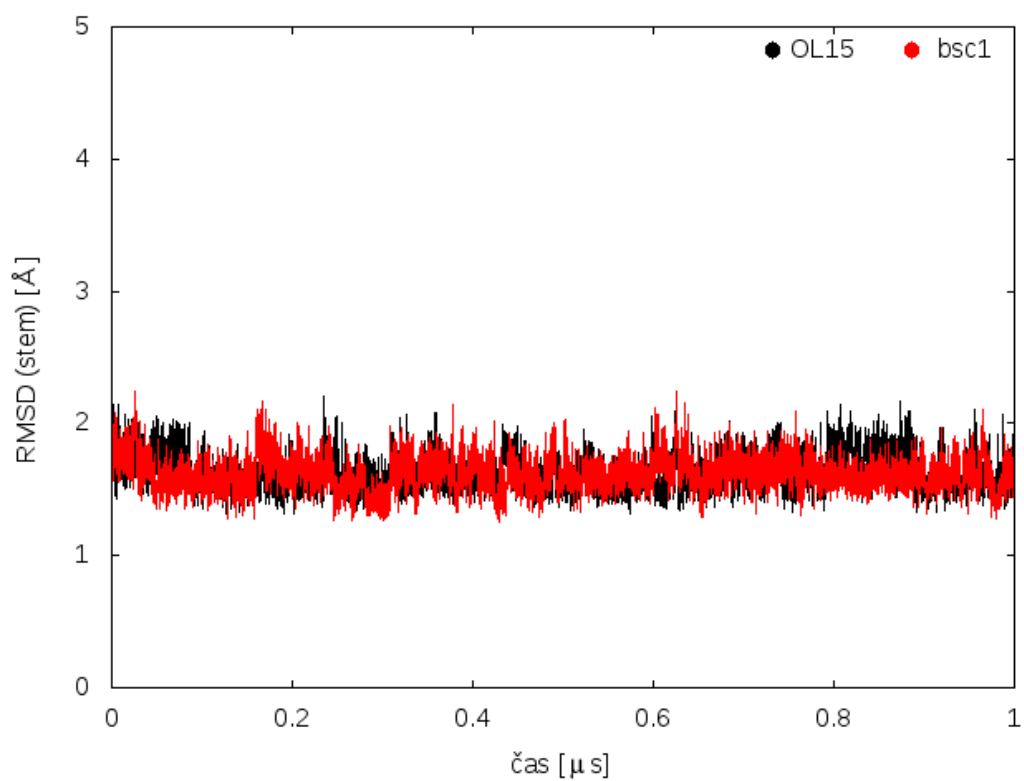
	Model 1	Model 2	Model 3	Model 4
konformace	β/γ	β/γ	β/γ	β/γ
C2-G3_A	62/186	182/180	62/185	173/190
C2-G3_B	62/186	181/180	62/184	187/190

Rozdílné hodnoty v různých modelech naznačují, že v páteři může existovat rovnováha mezi dvěma stavy uvedenými v Tabulce 4: konformací s $\beta/\gamma = g^+/t$ a další konformací s $\beta/\gamma = t/t$. Je ale také možné, že konformace v tomto místě páteře není dostatečně určena daty z NMR experimentu. Je nutno poznamenat, že úhly β a γ nebyly v NMR experimentu přímo měřeny.

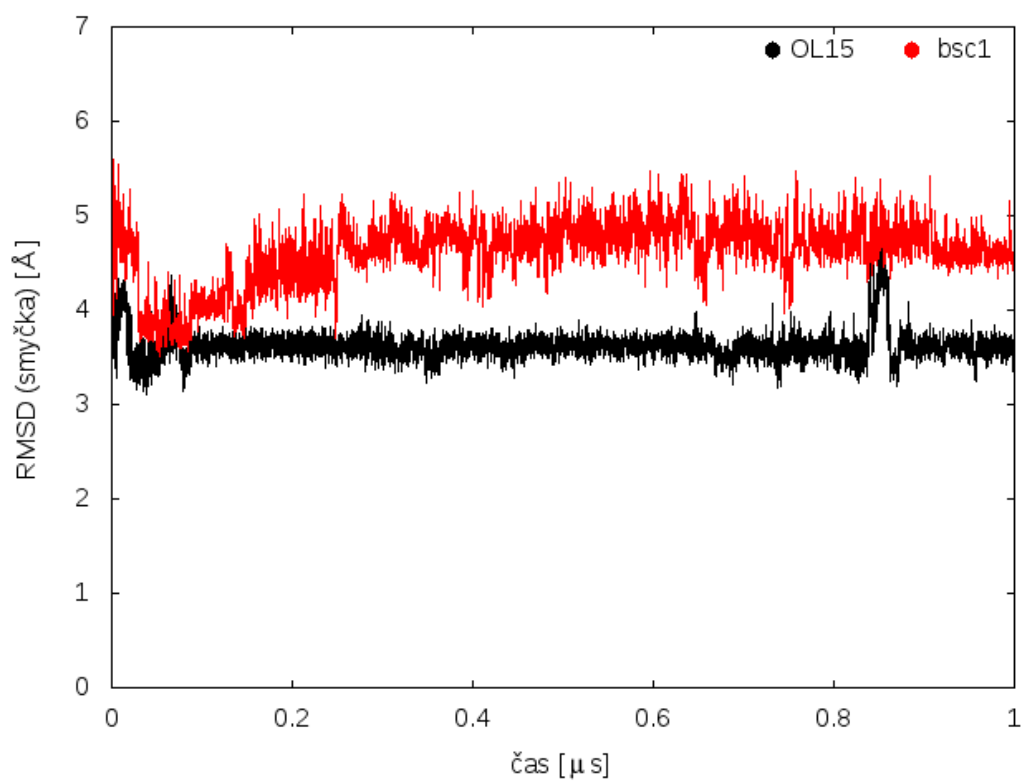
V Grafu 1 a Grafu 2 je zobrazeno RMSD v MD simulaci pro stem a smyčky kvadruplexu zvlášť. Je vidět, že stem kvadruplexu je relativně stabilní s malým RMSD, zatímco RMSD smyček je poměrně velké. Z trajektorie je vidět, že smyčky mění v simulaci svoje uspořádání a střídají různé konformace, např. thymín je vázaný do žlábků kvadruplexu nebo naopak exponovaný do roztoku. Toto je důvodem zvýšeného RMSD smyčky. Pro smyčky kvadruplexu je toto chování relativně časté. [33] Malé RMSD stemu naznačuje, že stem kvadruplexu je stabilní a nedochází v něm k velkým odchylkám od výchozí struktury. Za zmínku stojí relativně vyšší RMSD smyček v poli bsc1, které naznačuje, že toto pole může popisovat smyčky hůře než pole OL15.

Vzhledem k tomu, že v původní struktuře nebyly udány pozice iontů sodíku v kanálu kvadruplexu, byly dodány do struktury tři kationty do prostoru mezi roviny tetrad a dále po jednom nad a pod krajní roviny tetrad. V simulaci dva krajní kationty hned na začátku simulace opustily svá vazebná místa v kvadruplexu a oddifundovaly do roztoku. Zbylé tři kationty zůstaly vázány v prostoru kanálu kvadruplexu. Toto chování je u kvadruplexů obvyklé. [33] Mezi OL15 a bsc1 nebyly v tomto směru významné rozdíly.

Graf 1. RMSD stemu struktury kvadruplexu pro silová pole OL15 a bsc1



Graf 2. RMSD smyčky struktury kvadruplexu pro silová pole OL15 a bsc1

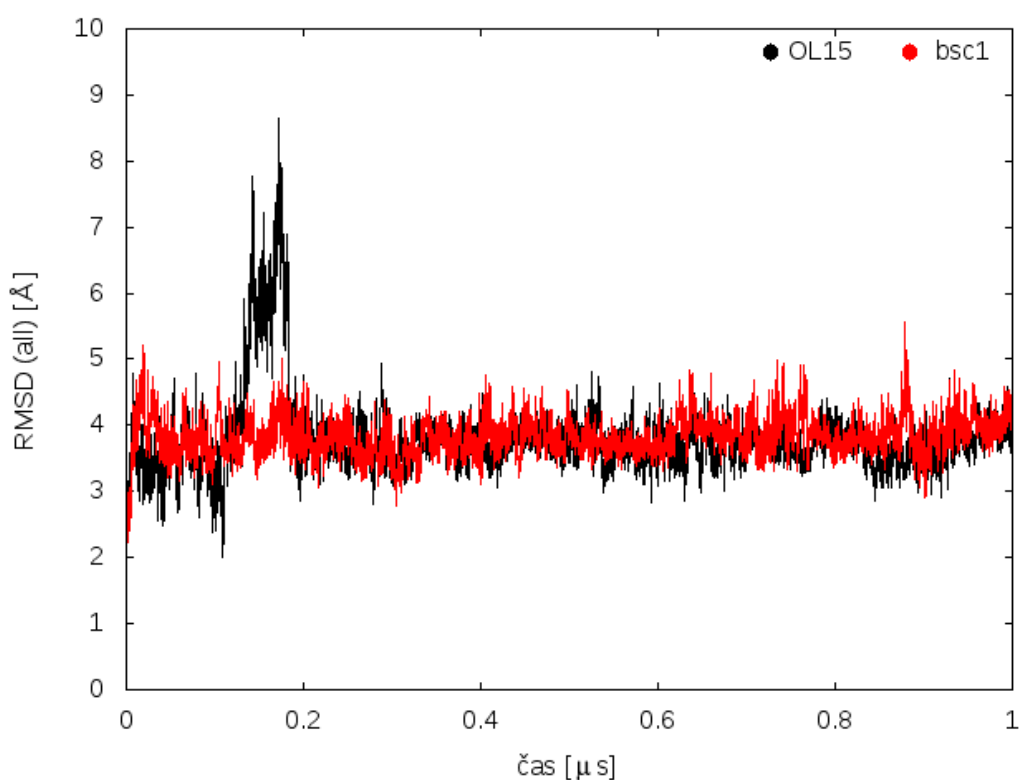


3.3.1.2 HOLLIDAYOVA STRUKTURA

Výchozí Hollidayova krystalová struktura s PDB ID 1NVY byla vybrána díky velmi dobrému rozlišení 1,50 Å. Studovaná konformace β/γ 117 je v řetězcích stejná s $\beta/\gamma = g^+/t = \sim 42^\circ, 155^\circ$. Tato struktura obsahuje také několik kationtů stroncia.

V Grafu 3 je zobrazeno RMSD celé Hollidayovy struktury. RMSD v případě obou polí dosahuje relativně vysokých hodnot. V silovém poli bsc1 je RMSD Hollidayovy struktury téměř po celou dobu simulace stabilní, oproti tomu RMSD Hollidayovy struktury v silovém poli OL15 se na začátku simulace příliš stabilně nejeví, ale po 0,2 μs se ustálí tak jako je tomu v případě silového pole bsc1. Toto vychýlení bylo způsobeno vzdálením dvou koncových dvoušroubovic od sebe, které bylo ale pouze přechodné. Relativně vyšší celková hodnota RMSD je způsobena také vznikem nenativních vodíkových vazeb, které budou popsány níže. Jinak je simulace v čase stabilní a obě silová pole poskytují podobný popis.

Graf 3. RMSD Hollidayové struktury pro silová pole OL15 a bsc1

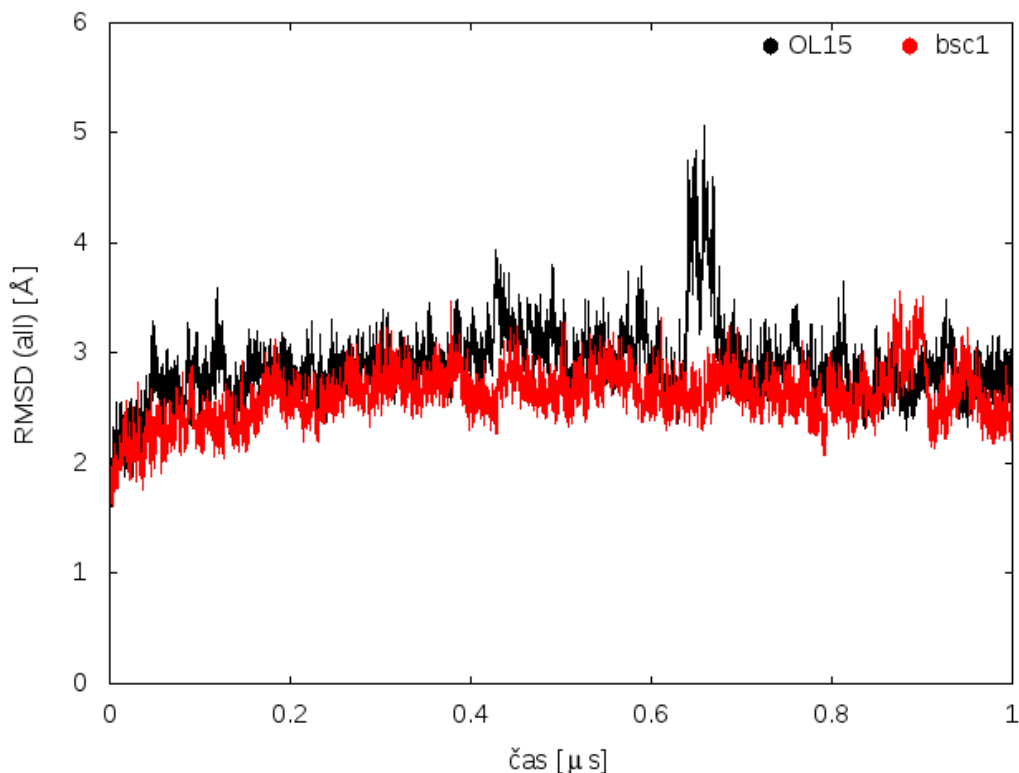


3.3.1.3 RAP1 PROTEIN-DNA KOMPLEX

Výchozí krystalová struktura s PDB ID 1IGN měla rozlišení 2,25 Å. Výchozí struktura obsahuje dva stejné protein-DNA komplexy v jedné krystalové buňce, ze které byl k simulaci vzat pouze první komplex. Protein se k nukleové kyselině váže ve dvou částech dvoušroubovice. Struktura obsahuje nekanonický stav č. 35 mezi rezidui guanin 31 a guanin 32 a tento stav je přítomen v obou komplexech přítomných v krystalové buňce. Výchozí krystalová struktura každého proteinu je na třech místech přerušena a neobsahuje tedy všechna rezidua RAP1 protein DNA komplexu.

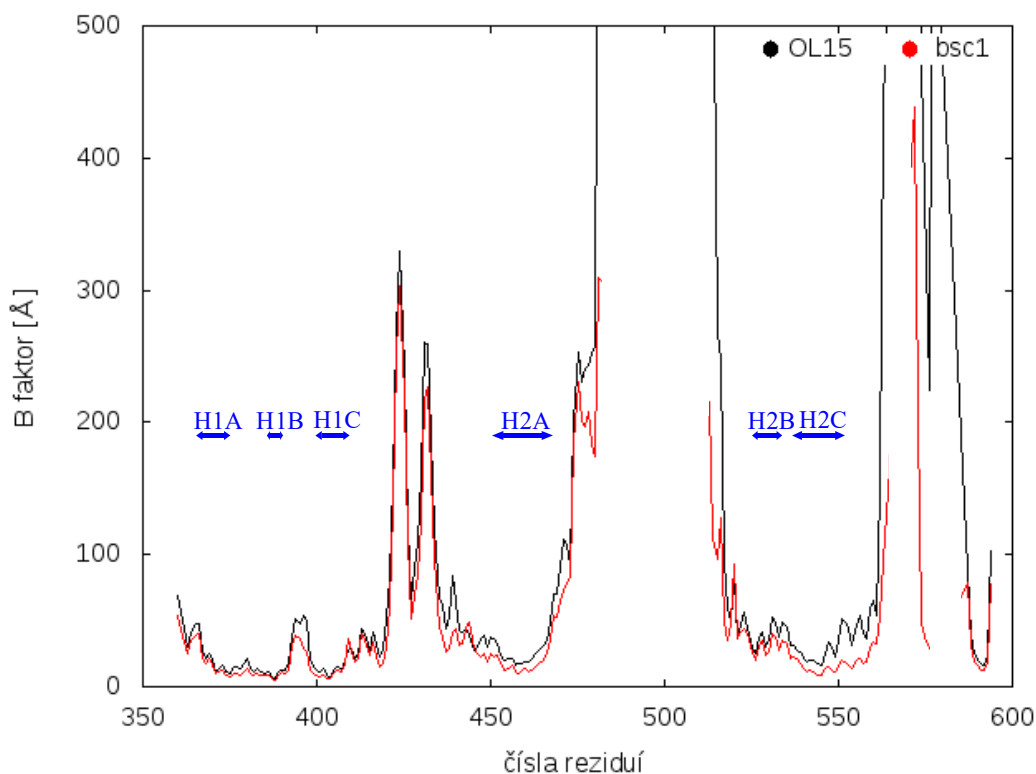
Graf 4 zobrazuje RMSD celé struktury RAP1 protein-DNA komplexu pro silová pole OL15 a bsc1. Až na jeden skok při zhruba 0,6 μ s by se dalo říci, že jsou RMSD pro obě pole stejná a dosahují hodnoty kolem 3 Å. RMSD je relativně velké, je ale potřeba vzít v úvahu, že tento komplex je relativně velký. Simulace se zdá být stabilní v obou silových polích s jen malými rozdíly mezi nimi.

Graf 4. RMSD struktury RAP1 protein-DNA komplexu pro silová pole OL15 a bsc1



U tohoto proteinu byly studovány také B faktory (viz Graf 5). V Grafu 5 jsou vyznačeny jednotlivé helixy proteinu. V blízkosti nekanonického stavu v reziduu guanin 32 RAP1 protein-DNA komplexu se nacházejí následující oblasti proteinu: smyčka mezi helixem H1A a H1B (rezidua 377-385), helix H1B (rezidua 386-391), helix H1C (rezidua 401-409), část regionu nepevně sbalené struktury (rezidua 444-445), N terminální rameno (rezidua 446-450) a C terminální smyčka (rezidua 578-594).

Graf 5. B faktor RAP1 protein-DNA komplexu pro silová pole OL15 a bsc1 (upraveno pomocí programu Inkscape)

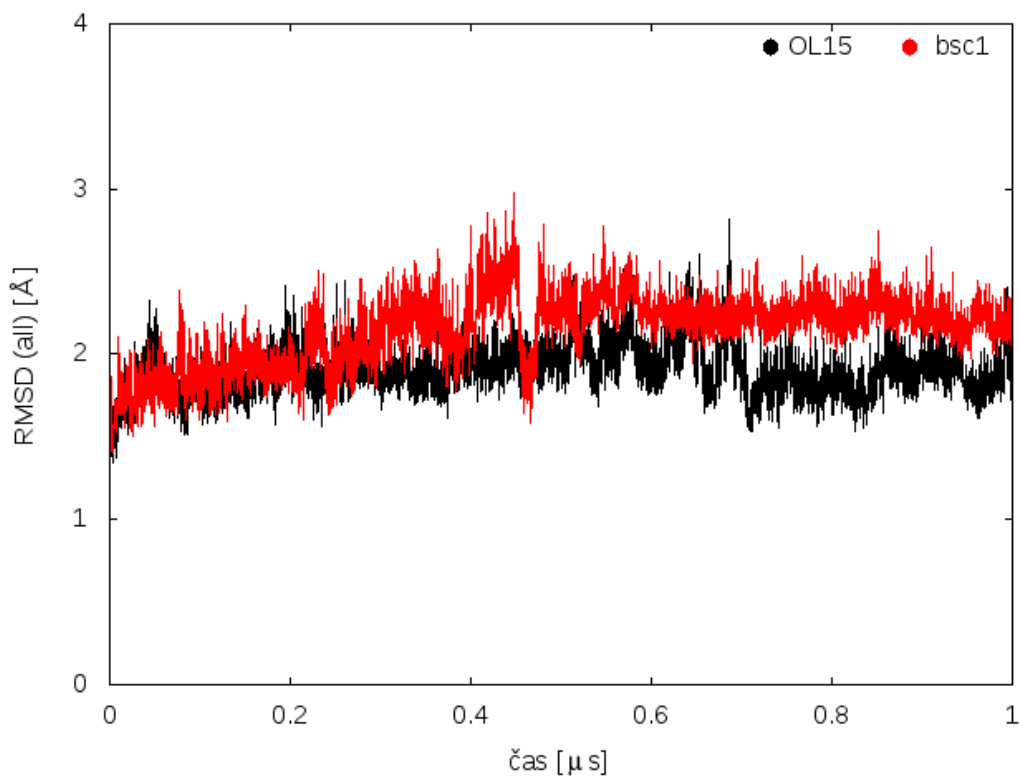


Z grafu je vidět, že fluktuace helixů jsou spíše malé a sekundární struktura proteinu je tedy v simulaci relativně stabilní. To platí i pro výše zmíněné oblasti, které jsou v kontaktu se studovaným reziduem guanin 31 duplexu. Výjimkou jsou části regionu nepevně sbalené struktury, kde jsou fluktuace značné, což lze ale očekávat. Fluktuace v silovém poli OL15 a bsc1 jsou velmi podobné.

3.3.1.4 λ REPRESOR PROTEIN-DNA KOMPLEX

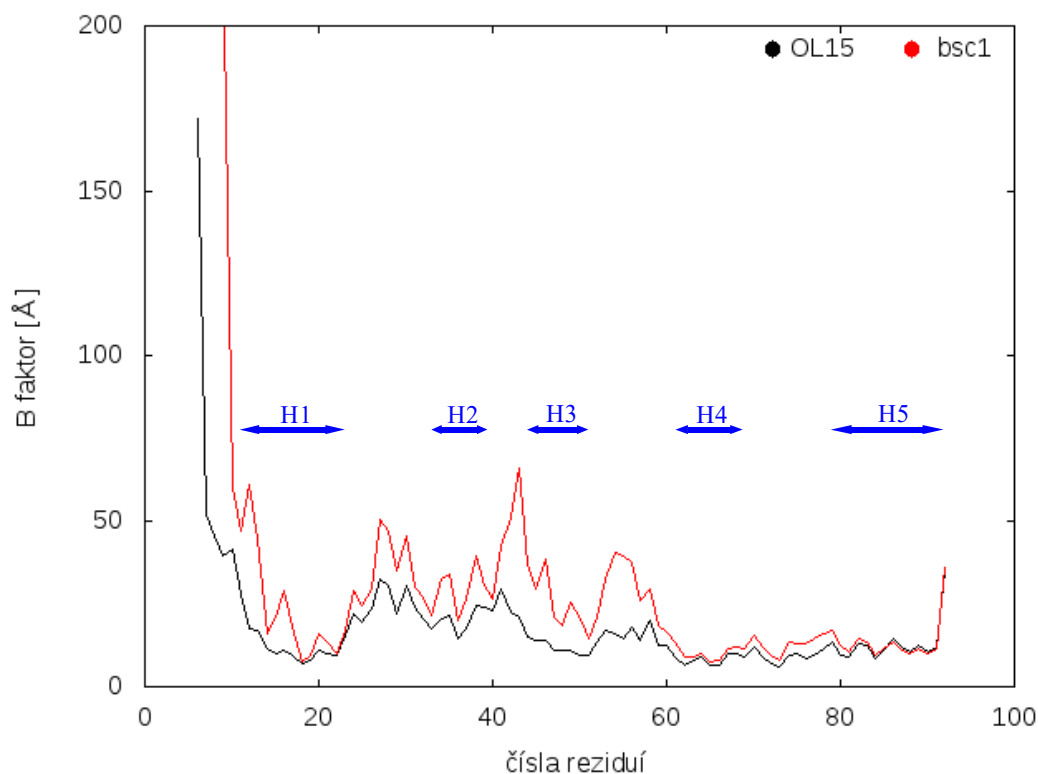
Výchozí krystalová struktura s PDB ID 1LMB má rozlišení 1,80 Å. Ve výchozí krystalové struktuře se protein dotýká nukleové kyseliny ve dvou místech – žlábcích. V Grafu 6 je zobrazeno RMSD této struktury pro silová pole OL15 a bsc1. V případě obou silových polí dosahuje RMSD hodnot okolo 2 Å, které jsou pro takto velký komplex přijatelné. RMSD je pro silové pole bsc1 o něco vyšší než pro pole OL15.

Graf 6. RMSD struktury λ represor protein-DNA komplexu pro silová pole OL15 a bsc1



U tohoto proteinu byly studovány také B faktory (viz Graf 7). V Grafu 7 jsou vyznačeny helixy struktury proteinu. V blízkosti nekanonického stavu v reziduu guanin 13 λ represor protein-DNA komplexu se nacházejí oblasti proteinu helix 3 (rezidua 44-51), smyčka mezi helixem 3 a helixem 4 (rezidua 52-60) a N koncová smyčka (rezidua 1-8) a smyčka mezi helixem 2 a helixem 3 (rezidua 40-43).

Graf 7. B faktor λ represor protein-DNA komplexu pro silová pole OL15 a bsc1 (upraveno pomocí programu Inkscape)



Z Grafu 7 je vidět, že oblasti helixů příliš neoscilují, naopak oblasti mezi helixy nebo výše zmíněná N koncová smyčka či smyčka mezi helixy 2 a 3 fluktuují více. Vysoká hodnota B faktoru zvláště pro pole bsc1 značí, že dochází k oscilaci jednotlivých reziduí a ve zvýšené míře.

3.3.2 STABILITA NEKANONICKÝCH STAVŮ PÁTEŘE β a γ

Jak jsem zmínila v úvodu, ve své práci jsem se zaměřila na dva různé stavy cukrfosfátové páteře, které obsahují nekanonickou konformaci sousedících úhlů β a γ : konformaci č. 117 a konformaci č. 35. [5] Stabilita těchto nekanonických stavů v MD simulacích bude popsána v následujícím textu.

3.3.2.1 KVADRUPLEX

V kvadruplexu jsou dvě symetricky položená místa, kde se nachází nekanonická konformace páteře β , γ (konformer 117), konkrétně mezi cytosinem 2 a guaninem 3, které jsou součástí sousedních vnitřních kvartetů kvadruplexu. Tato pozice byla vybrána, protože se nachází ve střední části stabilního stému kvadruplexu. Nenachází se v blízkosti flexibilních smyček a není tedy ovlivněna změnami jejich konformací. Časový vývoj těchto dvou stavů v reziduu guanin 3 a guanin 14 je vyobrazen v Grafech 8 a 9 pro silové pole OL15 a v Grafech 10 a 11 pro silové pole bsc1.

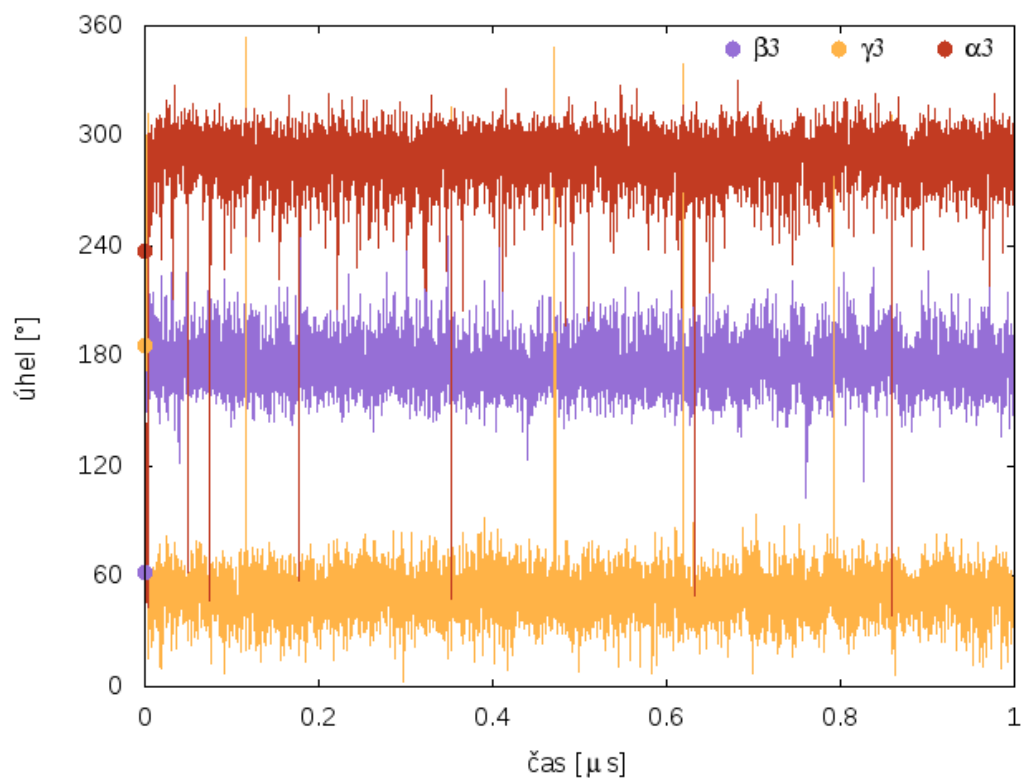
V silovém poli OL15 jsou oba nekanonické stavy nestabilní a hned na začátku simulace změni konformaci z β , $\gamma = g^+/t$ na kanonickou konformaci β , $\gamma = t/g^+$. Sousední úhel α je touto změnou konformace také ovlivněn. Ve výchozí geometrii má úhel α hodnotu přibližně 240° , při rekonformaci se ale změni na kanonickou hodnotu kolem 300° . Chování je v obou symetricky položených místech velmi podobné. V druhém řetězci navíc v průběhu simulace došlo k rekonformaci úhlu α z kanonické hodnoty $\alpha = g^-$ do hodnoty $\alpha = t$ a to naznačuje omezenou stabilitu kanonické konformace v této části molekuly.

V silovém poli bsc1 se první řetězec chová stejně, jako v poli OL15, ale v druhém řetězci je stabilita nekanonického stavu č. 117 vyšší a udrží se téměř do třetiny simulace. Pak ale dojde ke stejné rekonformaci do kanonického stavu, jako v prvním řetězci.

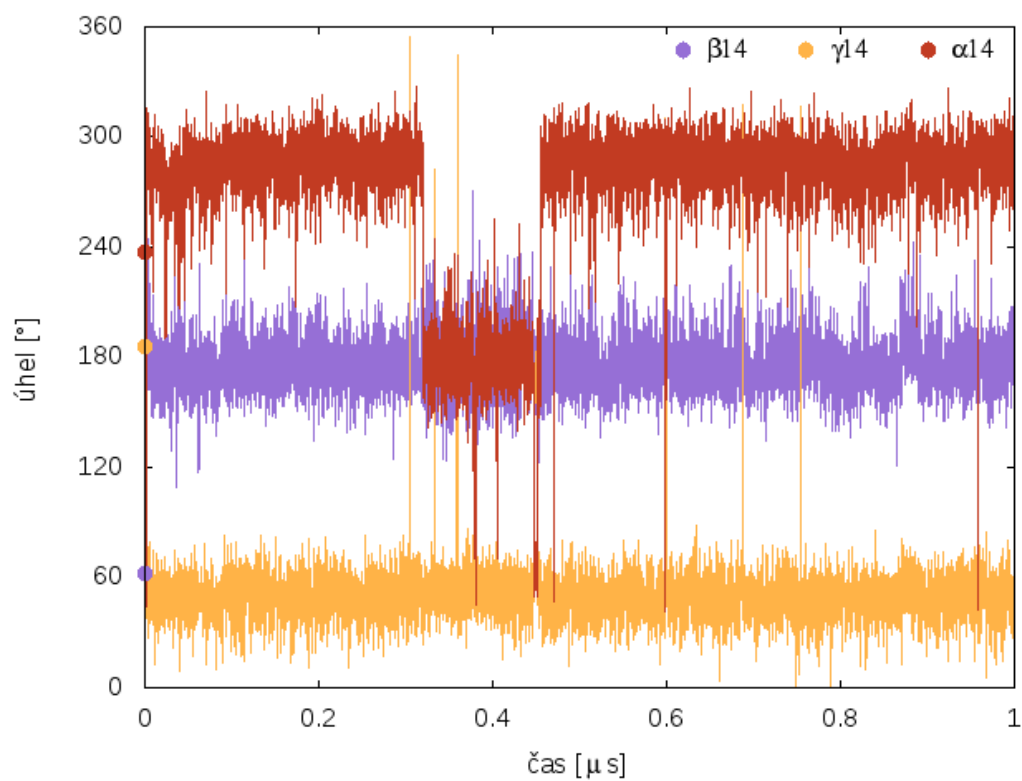
Z MD simulací vyplývá, že silová pole upřednostňují kanonickou konformaci, i když i nekanonický stav č. 117 je polem bsc1 na určitý čas udržen. Porovnáme-li toto chování s experimentálními údaji z Tabulky 4, je vidět, že i v experimentu jsou možné dva stavy, ale žádný z nich není kanonický, protože oba obsahují torzi γ v nekanonické hodnotě $\gamma = t$. Přestože experimentální konformace není zcela jednoznačně určena, zdá se, že oba

empirické potenciály by mohly popisovat tuto část páteře nesprávně a destabilizovat stav $\gamma = t$, v souladu s predikovanými předpoklady.

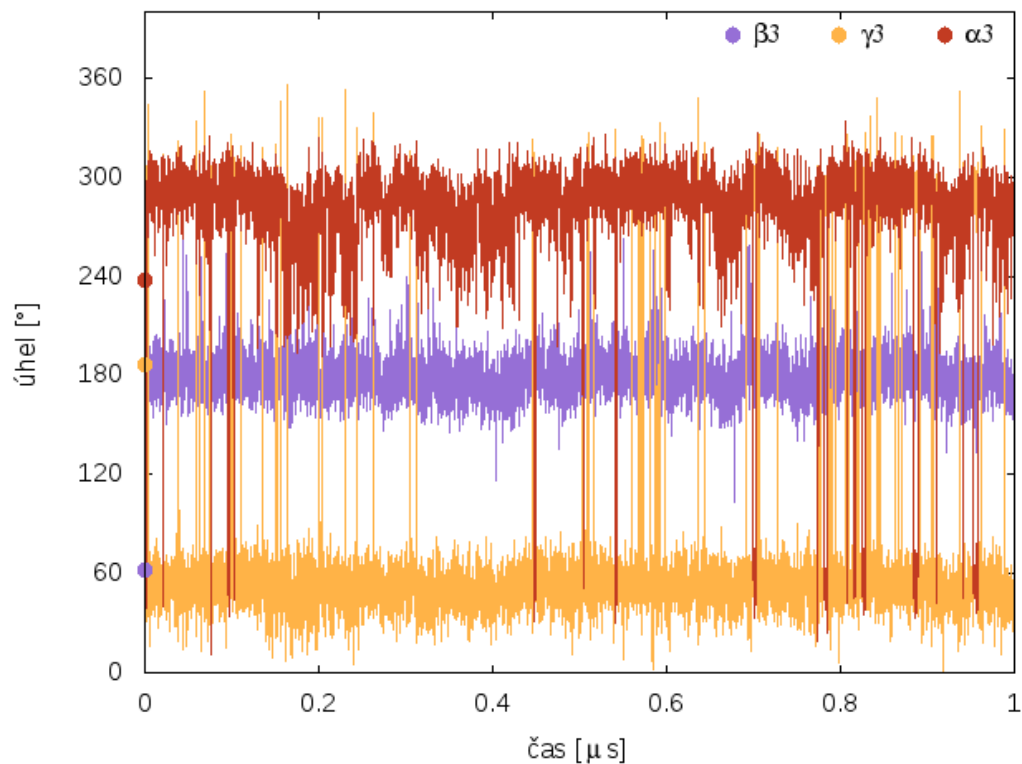
Graf 8. Časový vývoj úhlů β , γ , α v reziduu guanin 3 pro strukturu kvadruplexu v silovém poli OL15. Výchozí experimentální hodnoty jsou označeny tečkami na ose y.



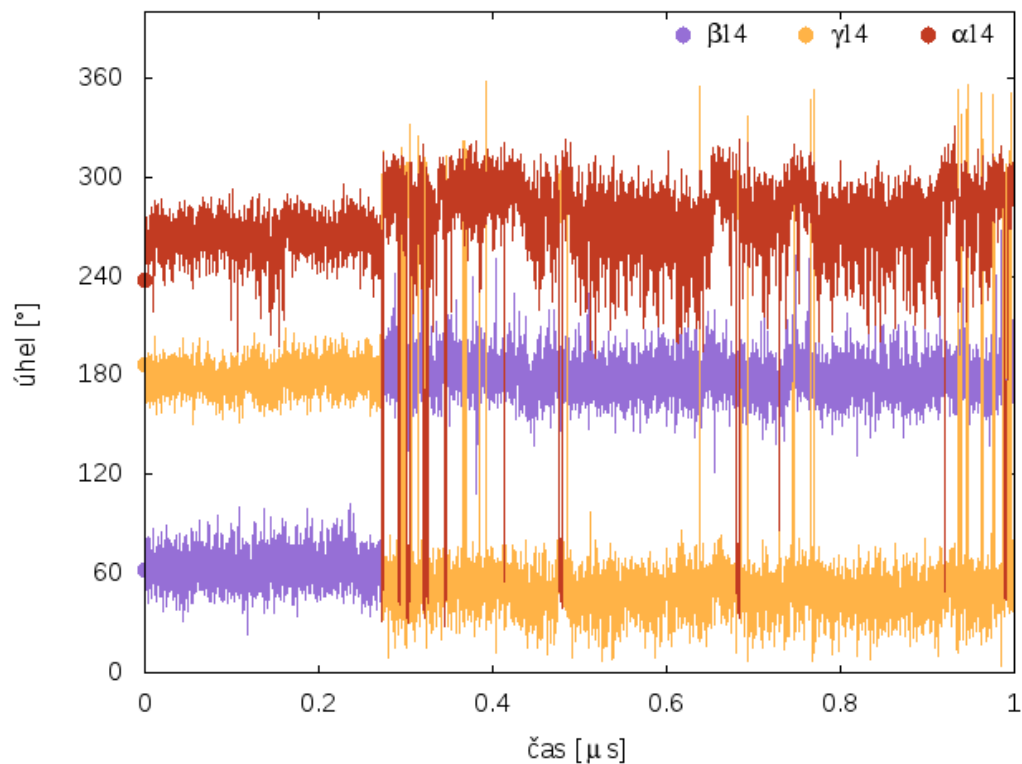
Graf 9. Časový vývoj úhlů β , γ , α v reziduu guanin 14 pro strukturu kvadruplexu v silovém poli OL15



Graf 10. Časový vývoj úhlů β , γ , α v reziduu guanin 3 pro strukturu kvadruplexu v silovém poli bsc1



Graf 11. Časový vývoj úhlů β , γ , α v reziduu guanin 14 pro strukturu kvadruplexu v silovém poli bsc1



3.3.2.2 HOLLIDAYOVA STRUKTURA

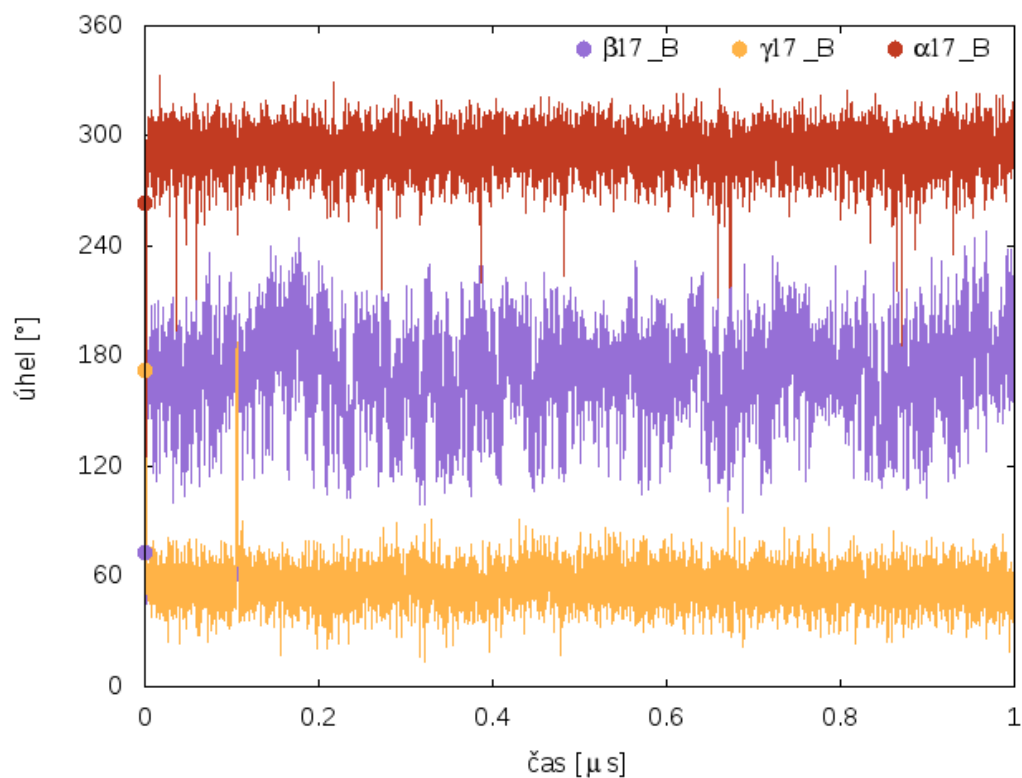
V Hollidayově struktuře se nachází dvě symetricky položená místa, kde se nachází nekanonická konformace páteře $\beta/\gamma = g^+/t$, konformer č. 117. Přesněji mezi rezidui adenin 6 a cytosin 7 v řetězcích B a D.

Začneme popisem simulace v silovém poli OL15. Ve výchozí geometrii má úhel β hodnotu přibližně 72° , úhel $\gamma \sim 172^\circ$ a úhel $\alpha \sim 263^\circ$. Hned v počátcích simulace se β i γ úhly v obou řetězcích neudrží a změní konformaci z $\beta/\gamma = g^+/t$ na kanonickou konformaci $\beta/\gamma = t/g^+$ (Grafy 12 a 13). Úhel α se v obou řetězcích posunuje blíže ke kanonické hodnotě 300° . Nekanonický stav je tedy v silovém poli OL15 nestabilní.

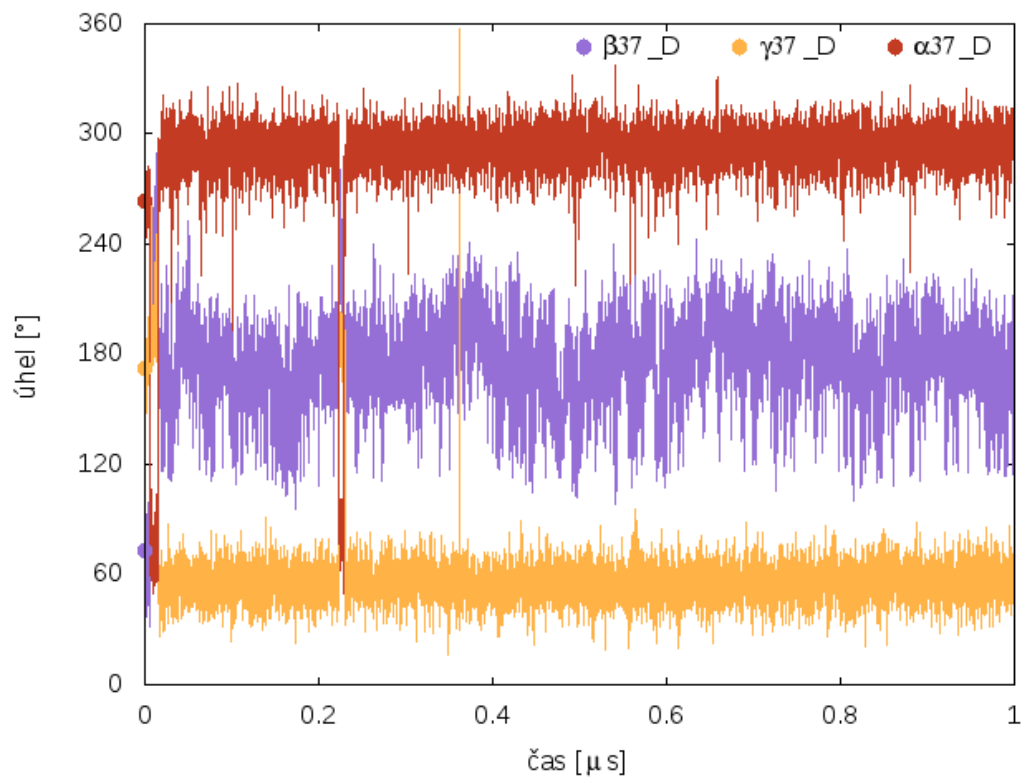
Avšak pro silové pole bsc1 je situace mírně odlišná (Grafy 14 a 15). V obou řetězcích se nekanonická konformace po jistou dobu udrží a až poté přecházejí do kanonických hodnot jako v poli OL15. Důležité ale je, že v obou řetězcích se páteř do nekanonického stavu č. 117 dokáže vrátit a setrvat v něm relativně dlouhou dobu (desítky až stovky ns) To naznačuje, že stav č. 117 může v této struktuře existovat a je lépe popsán silovým polem bsc1.

Souvislost mezi nekanonickým stavem 117 a dalšími okolními úhly ukazují Grafy 16 a 17 pro pole OL15 a 18 a 19 pro pole bsc1. Hlavní informace je obsažena v úhlu χ , jehož hodnota je ve výchozí struktuře $\chi \sim 200^\circ$ (*anti*, typická pro A-DNA) a při rekonformaci β/γ do kanonického stavu se mění na $\chi \sim 250^\circ$ (*high-anti*, typická pro B-DNA). Podobně v bsc1 souvisí hodnota χ s tím, jestli je β/γ stav nekanonický nebo kanonický. Protože parametrizace χ se v OL15 a bsc1 liší, lze soudit, že má vliv i na stabilitu nekanonického stavu 117. V poli bsc1 se zdá být úhel χ popsán lépe.

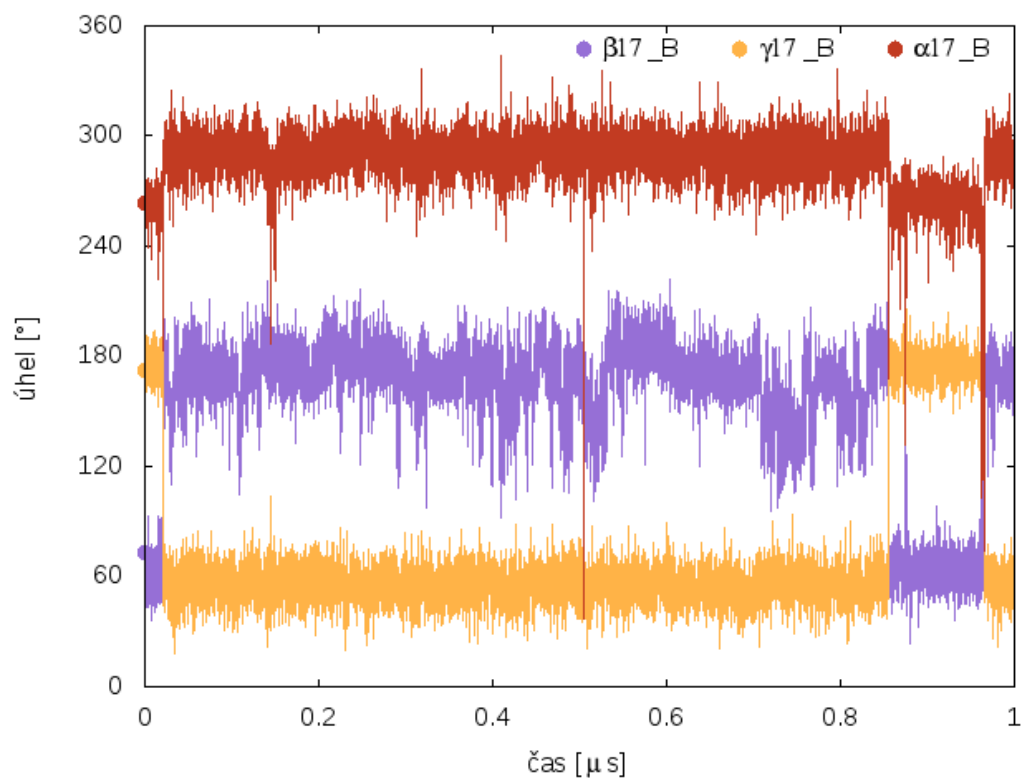
Graf 12. Časový vývoj úhlů β , γ , α v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli OL15



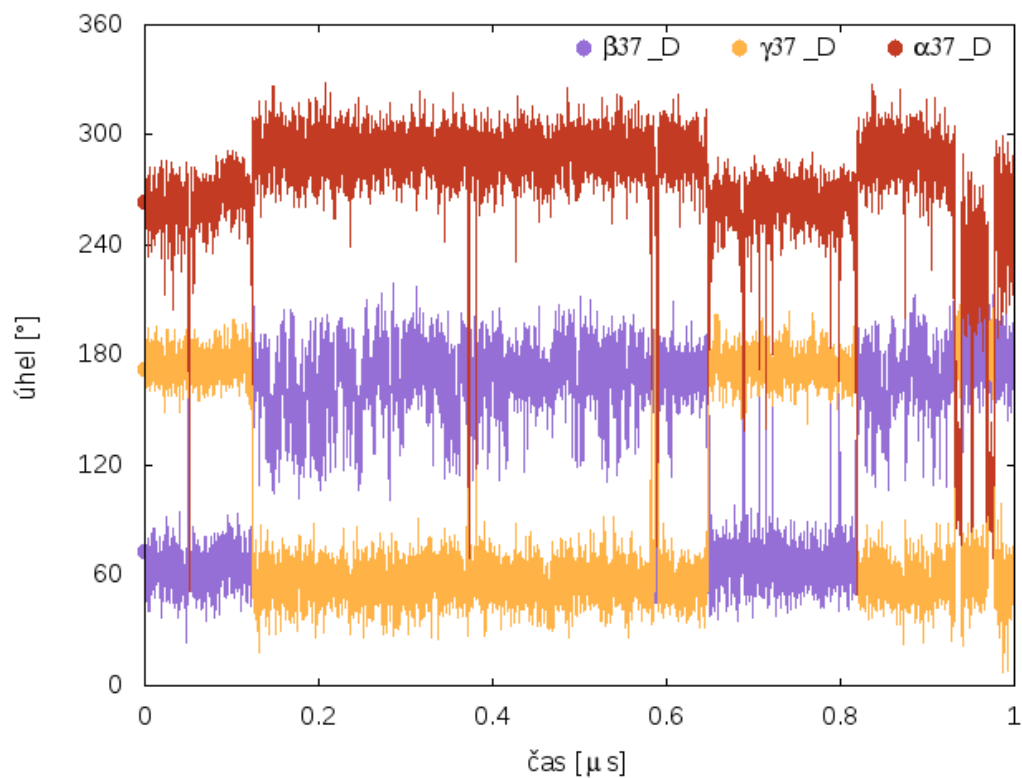
Graf 13. Časový vývoj úhlů β , γ , α v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli OL15



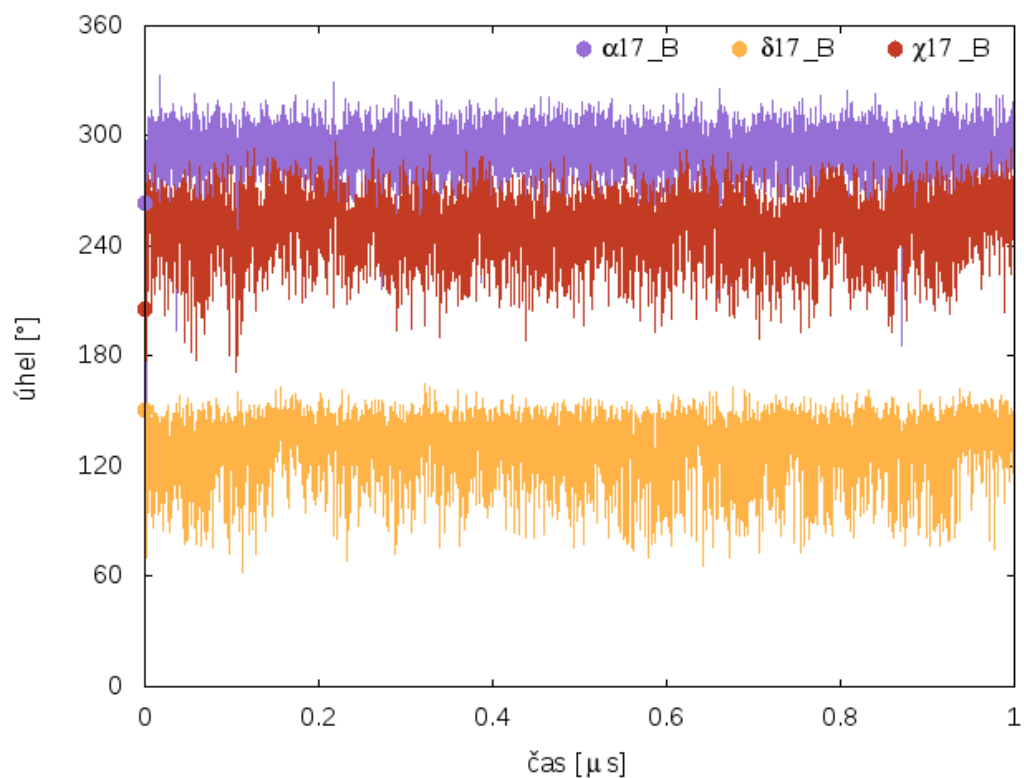
Graf 14. Časový vývoj úhlů α , β , γ v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli bsc1



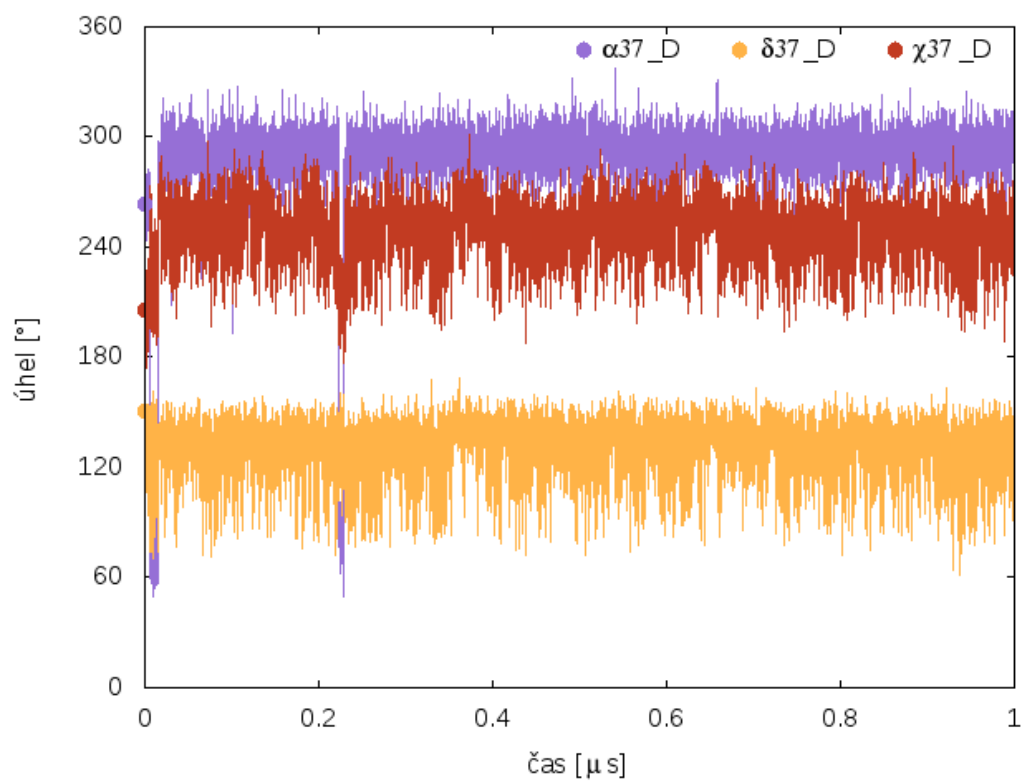
Graf 15. Časový vývoj úhlů α , β , γ v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli bsc1



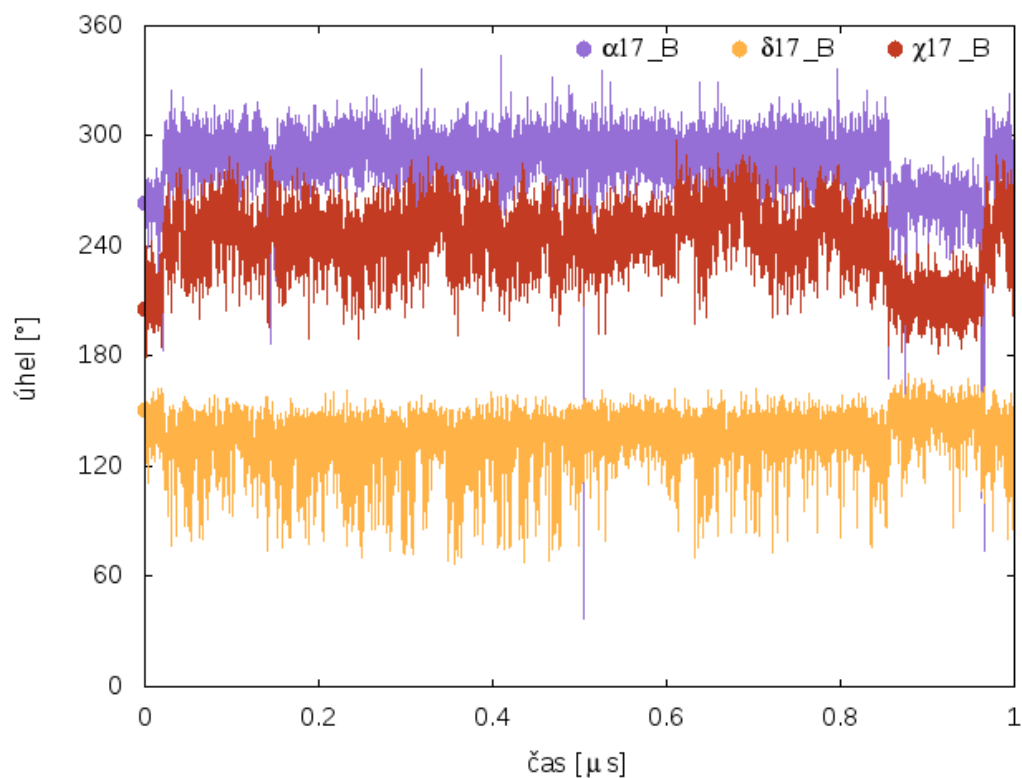
Graf 16. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli OL15



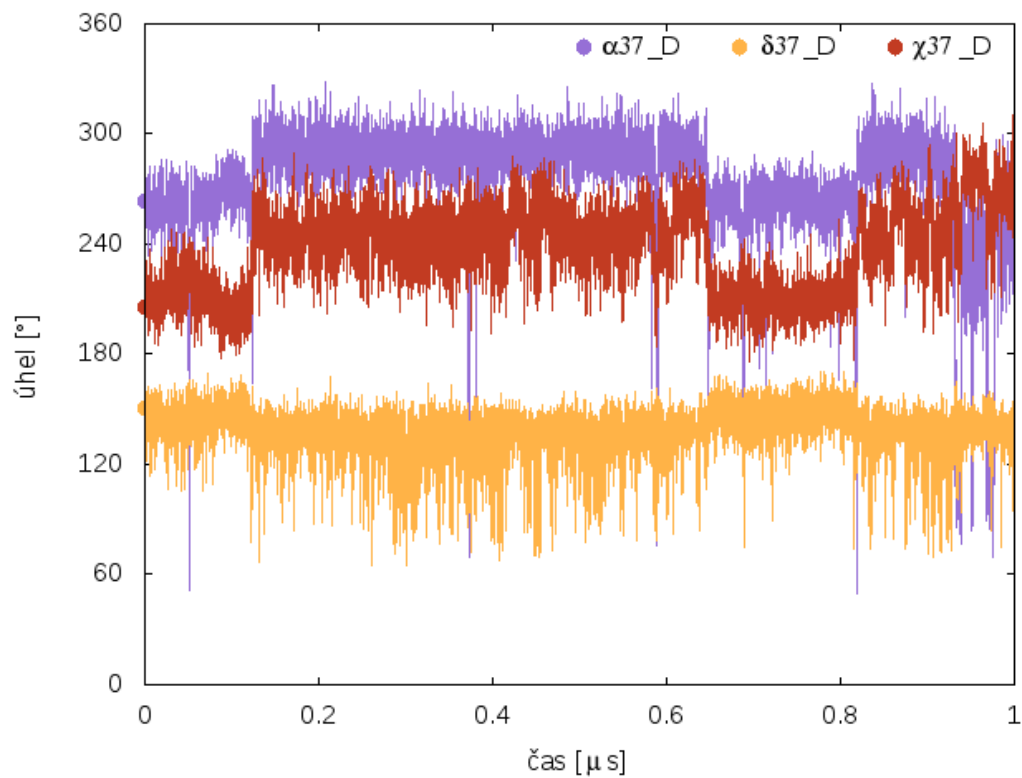
Graf 17. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli OL15



Graf 18. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli bsc1



Graf 19. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli bsc1

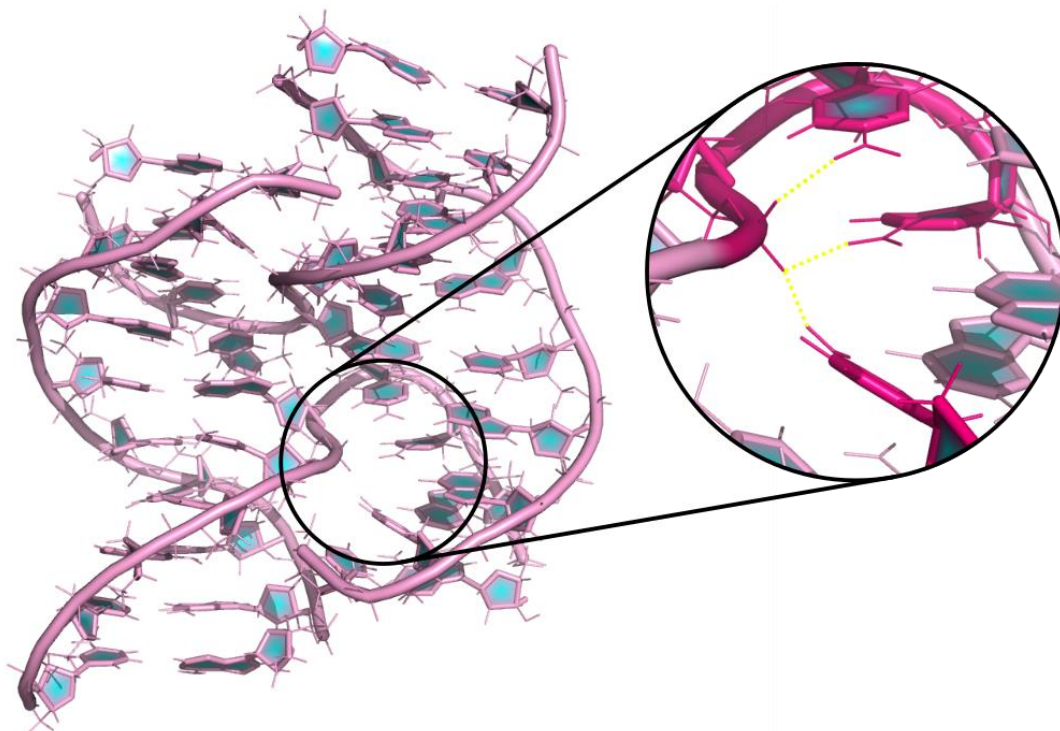


3.3.2.2.1 NENATIVNÍ VODÍKOVÉ VAZBY

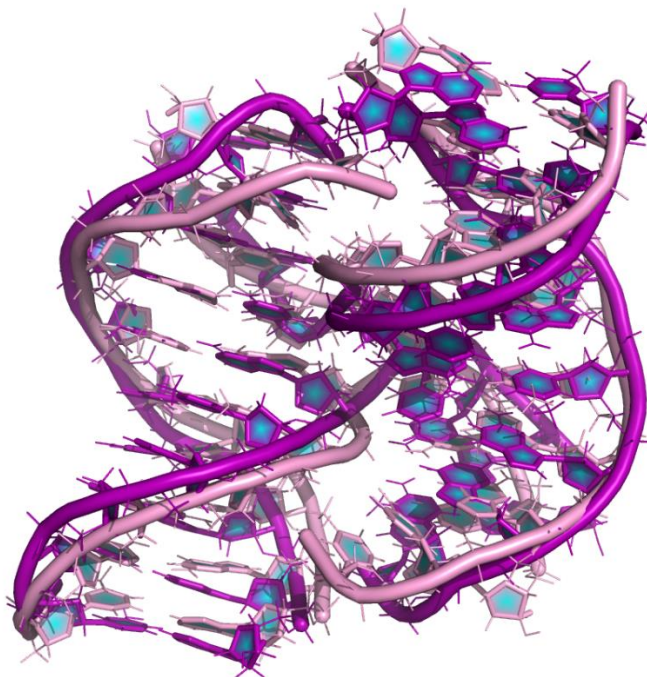
V průběhu simulace Hollidayovy struktury se tvořily vodíkové vazby, které nejsou v původní experimentální struktuře pozorovány (nenativní vodíkové vazby, viz Obrázek 12). Jedná se o vazby mezi fosfátem rezidua adeninu 6 v řetězci A k aminoskupinám reziduí cytosinů 2 v řetězci B, dále cytosinu 7 v řetězci A a cytosinu 8 v řetězci A. Analogické vazby vynikají i v symetricky položeném fosfátu rezidua adeninu 6 v řetězci C. Vzdálenosti vodíkově vázaných reziduí na začátku a v průběhu simulace pole OL15 v prvním a druhém řetězci jsou zobrazeny v Grafech 20 a 21 pro silové pole OL15 a v Grafech 22 a 23 pro silové pole bsc1. Je vidět, že stejné vodíkové vazby se tvoří v obou silových polích, OL15 i bsc1 relativně brzy na začátku simulace a jsou relativně stabilní. Tvorba těchto nenativních vodíkových vazeb vede k mírné deformaci struktury, jak je ukázáno na Obrázku 13.

Podobné vodíkové vazby na fosfátovou skupinu byly pozorovány už dříve např. v simulacích balení RNA smyček [34]. V těchto pracích bylo ukázáno, že vodíkové vazby s fosfátovou skupinou jsou v silovém poli přeceněny, tj. jsou silnější, než by v realitě měly být. Protože fosfátová skupina má v DNA stejnou parametrizaci jako v RNA, dá se předpokládat, že tvorba pozorovaných vazeb je nejspíše i v našich simulacích artefaktem silového pole a MD simulace tedy nepopisuje chování systému správně. Jde o další příklad, který ukazuje, že současné silová pole vyžadují další korekce. Ze srovnání Grafů 18 a 19 s Grafy 22 a 23 je vidět, že tvorba nekanonických vodíkových vazeb nesouvisí s rekonformací nekanonických stavů β/γ (nastávají v čase nezávisle).

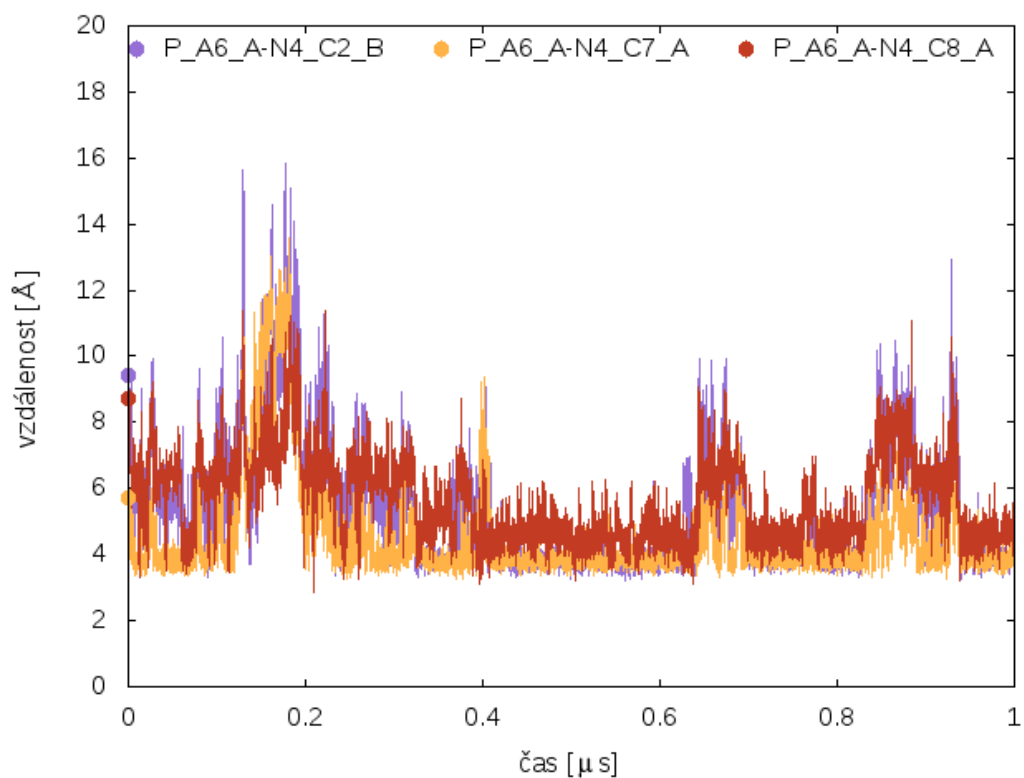
Obrázek 12. Výřez z posledního snímku simulace Hollidayovy struktury – nenativní vodíkové vazby mezi fosfátovou skupinou a aminoskupinami cytosinů (vytvoreno pomocí programu PyMOL, Inkscape)



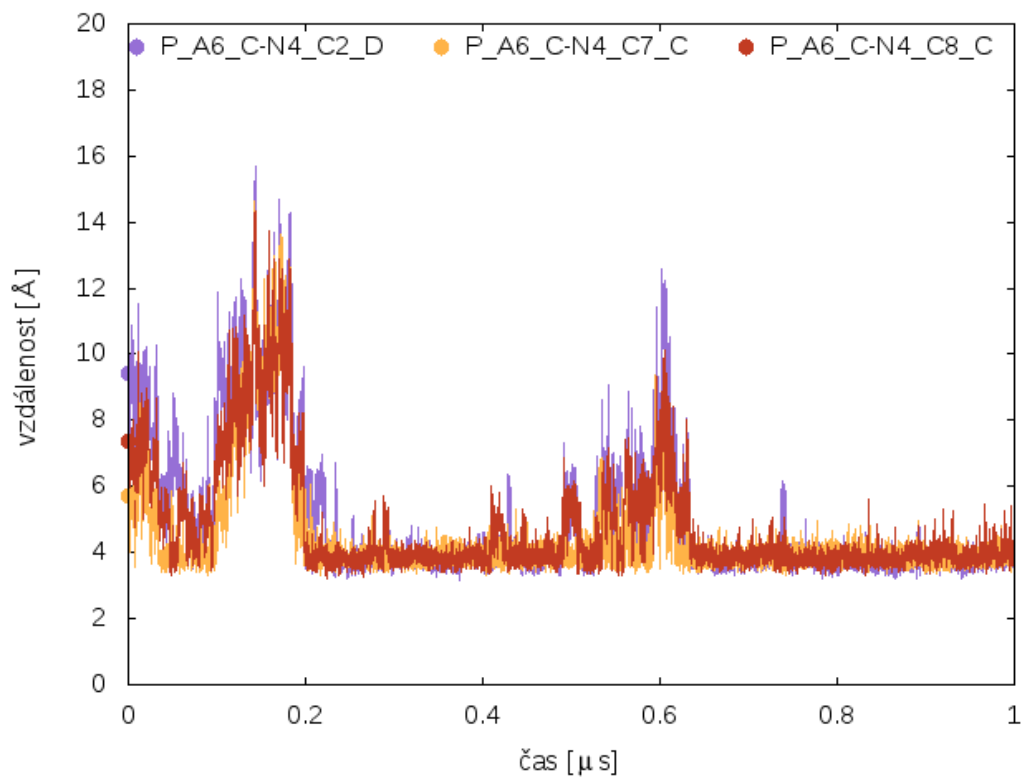
Obrázek 13. Srovnání krystalové geometrie (fialová) a posledního snímku 1 μ s simulace (růžová) Hollidayovy struktury (vytvoreno pomocí programu PyMOL)



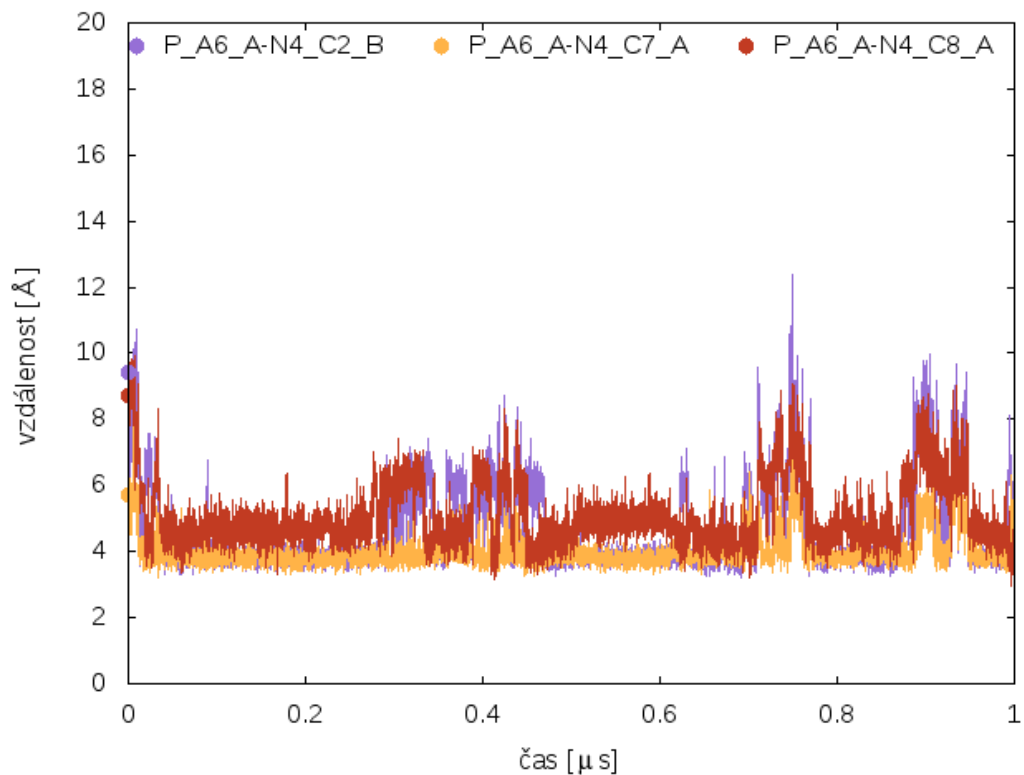
Graf 20. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli OL15 pro Hollidayovu strukturu v prvním řetězci



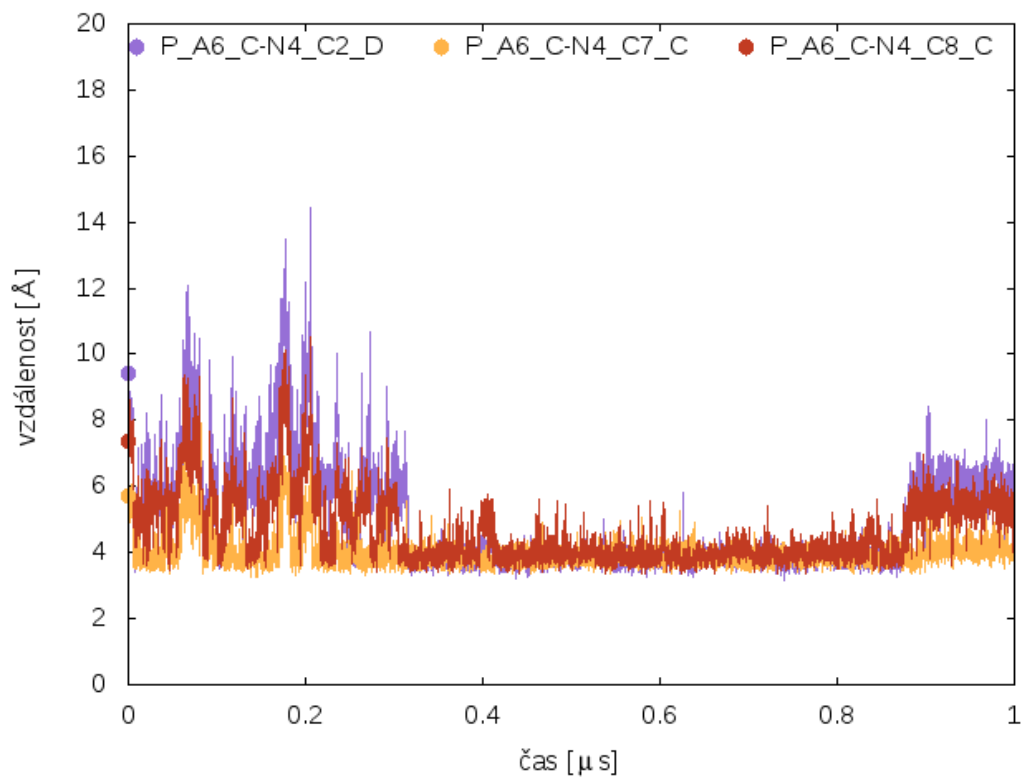
Graf 21. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli OL15 pro Hollidayovu strukturu v druhém řetězci



Graf 22. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli bsc1 pro Hollidayovu strukturu v prvním řetězci



Graf 23. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli bsc1 pro Hollidayovu strukturu v druhém řetězci

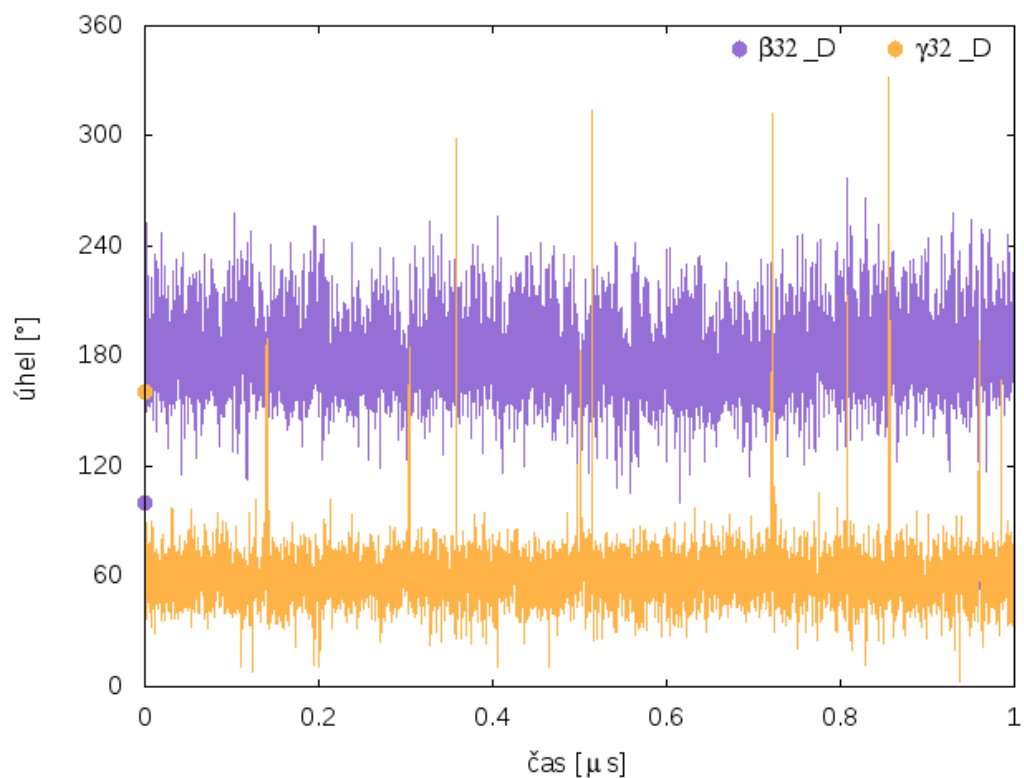


3.3.2.3 RAPI PROTEIN DNA KOMPLEX

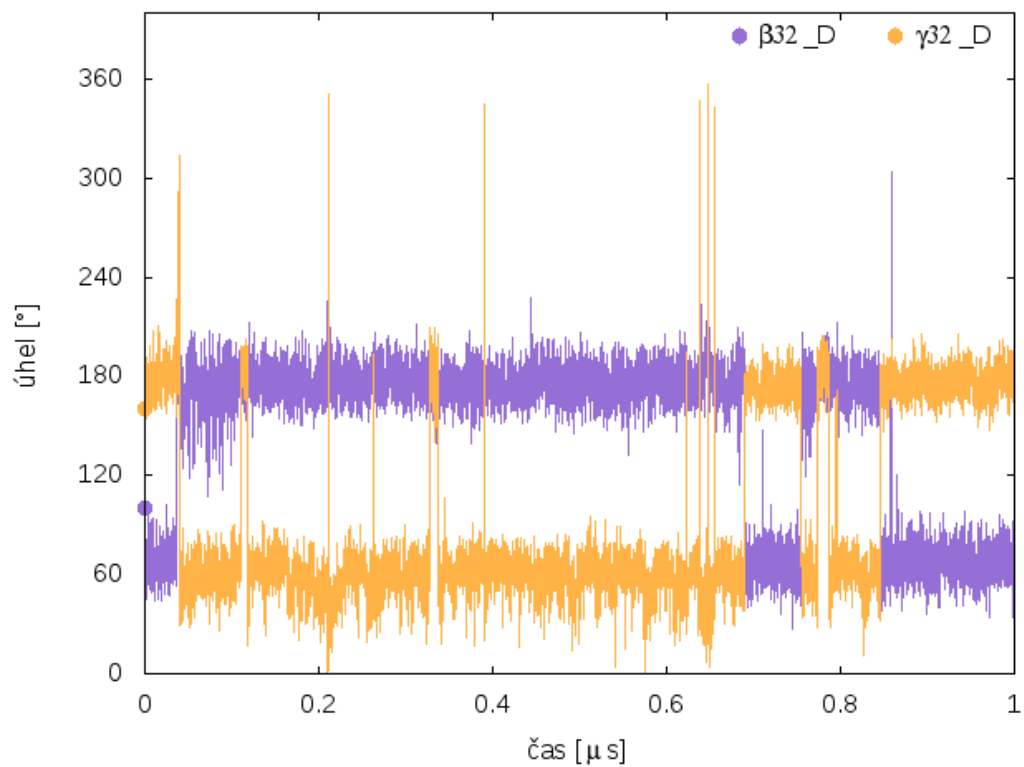
Ve výchozí geometrii vykazují úhly β a γ rezidua guaninu 32 hodnot $\beta \sim 100^\circ$, a $\gamma \sim 160^\circ$ (konformer 35). V konformeru 35 jsou tyto hodnoty spojeny se severním puckerem (N) a glykosidickou torzí $\chi = anti$ v daném reziduu. V silovém poli OL15 se konformace $\beta/\gamma = g^+/t$ na začátku simulace velmi rychle změní na kanonickou $\beta/\gamma = t/g^+$ a zůstane v ní po zbytek simulace (Graf 24). Naproti tomu v silovém poli bsc1 je situace odlišná (Graf 25). Na počátku simulace se konformer 35 po jistou dobu udrží, a přestože pak rekonformuje do kanonického stavu $\beta/\gamma = t/g^+$, je páteř schopna se do konformace 35 v průběhu simulace zase na kratší čas několika desítek ns vrátit.

V Grafech 26 a 27 (pro OL15 a bsc1) jsou zobrazeny časové průběhy úhlů α , δ , χ rezidua guaninu 32. Ve výchozí geometrii vykazují hodnoty $\alpha \sim 245^\circ$ (nekanonická), $\delta \sim -93^\circ$ (severní pucker, N) a $\chi \sim 174^\circ$. V silovém poli OL15 se v průběhu simulace tyto úhly změní do kanonických hodnot. V silovém poli bsc1 ale vykazují jasnou souvislost s nekanonickým stavem β/γ . Když je β i γ nekanonická, drží si své nekanonické hodnoty také zmíněné úhly. Protože se pole OL15 a bsc1 liší v parametrizaci puckeru ribózy a úhlu χ , lze říci, že pole bsc1 popisuje díky těmto odlišnostem konformaci 35 lépe. Dále lze říci, že obě pole pravděpodobně podceňují populaci stavu 35, možná díky nepřesnému popisu stavu $\gamma = t$, v souladu s našimi předpoklady.

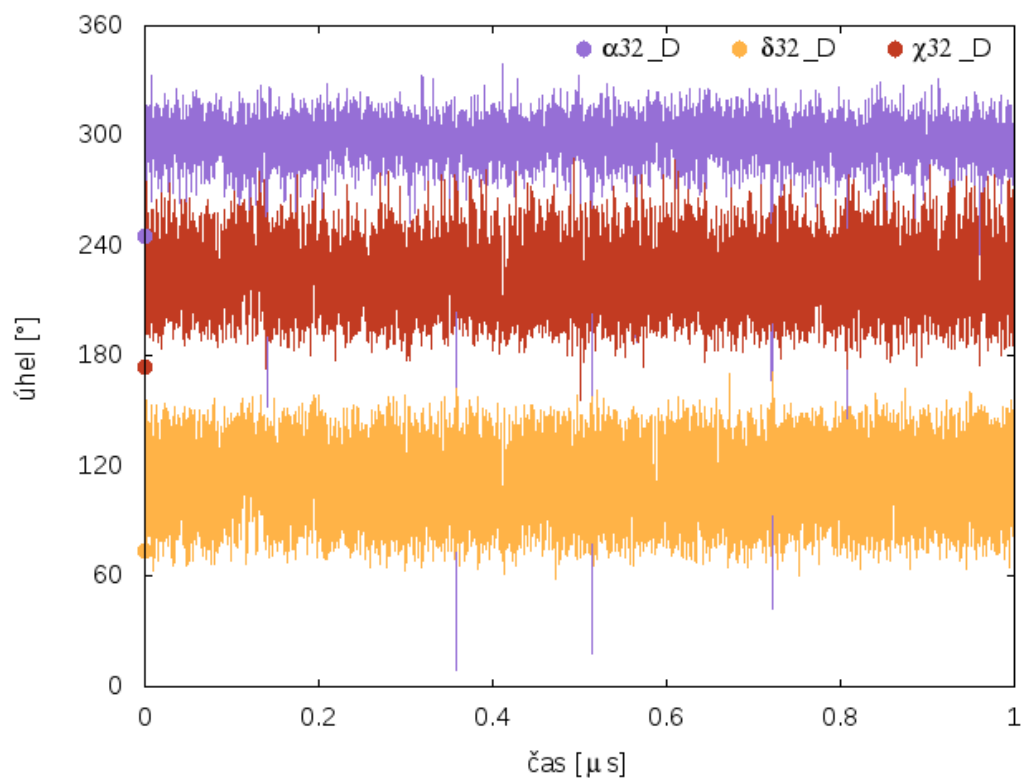
Graf 24. Časový vývoj úhlů β , γ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli OL15



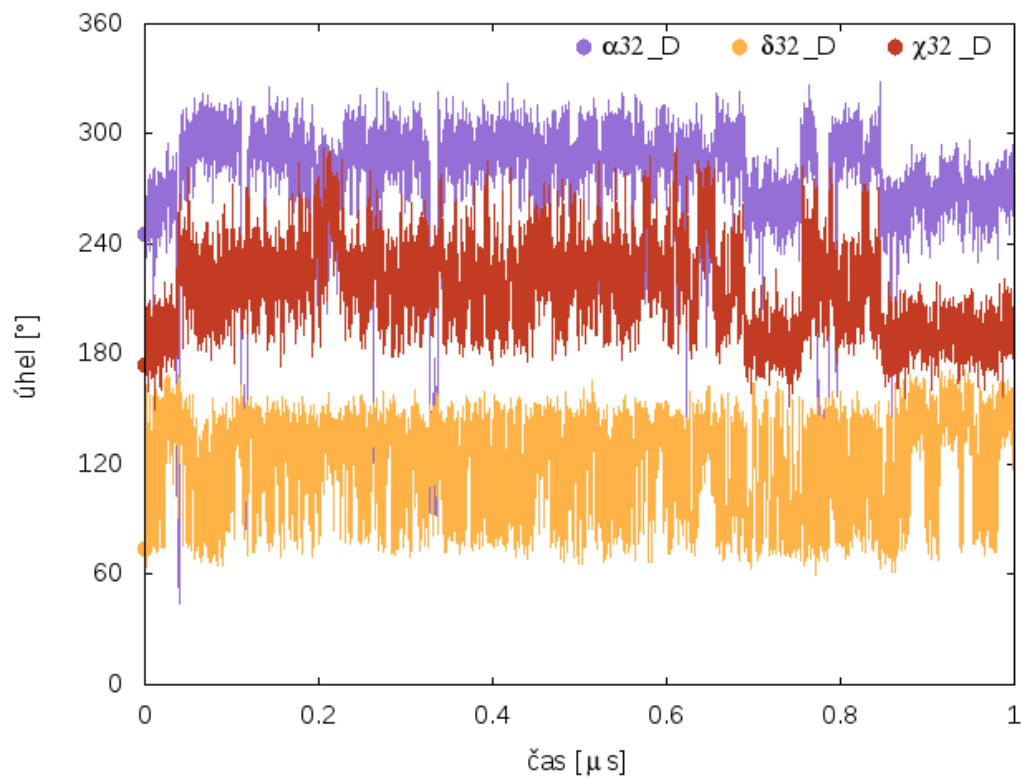
Graf 25. Časový vývoj úhlů β , γ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli bsc1



Graf 26. Časový vývoj úhlů α , δ , χ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli OL15



Graf 27. Časový vývoj úhlů α , δ , χ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli bsc1

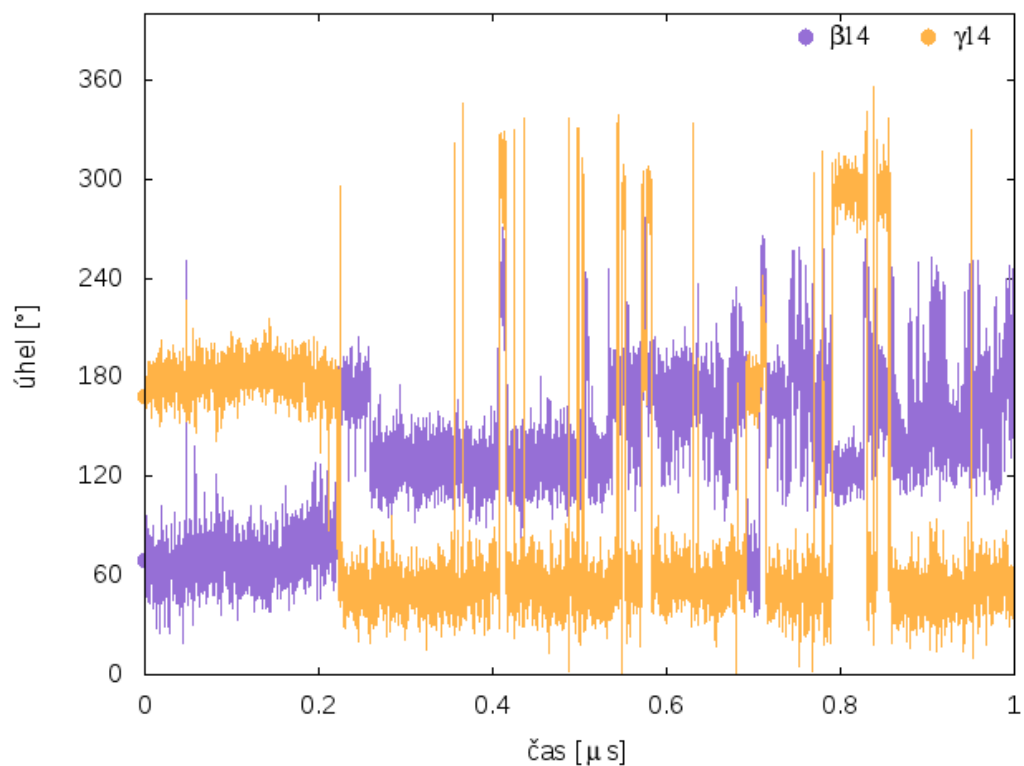


3.3.2.4 λ REPRESOR PROTEIN-DNA KOMPLEX

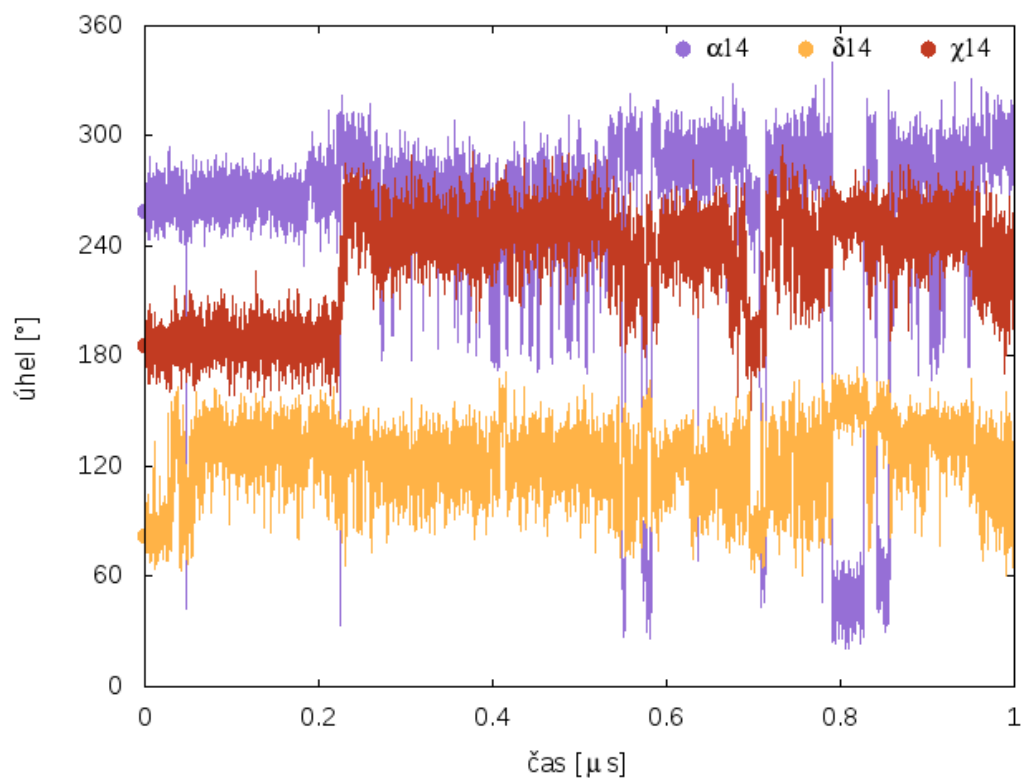
V tomto komplexu se nekanonický stav $\beta/\gamma = g^+/t$ nachází na reziduu guanin 14. Ve výchozí geometrii nabývá úhel β hodnoty $\sim 69^\circ$ a $\gamma \sim 168^\circ$. V průběhu simulace v silovém poli OL15 (Graf 28) se úhly β i γ po dobu prvních 0,2 μ s drží ve své výchozí konformaci, avšak poté svou konformaci změní na hodnoty blízké kanonickým, $\beta = t$, $\gamma = g^+$. Po zbytek simulace ale nejsou příliš stabilní a úhel γ mění svoji konformaci i na g^- . Naproti tomu v silovém poli bsc1 se úhly β i γ drží ve své výchozí konformaci (viz Graf 30).

Úhly α , δ , χ , které souvisejí s nekanonickou konformací 35, odpovídaly v původní geometrii hodnotám 258° , 82° a 185° v tomto pořadí. Časový vývoj těchto úhlů je zaznamenán v Grafu 29 pro silové pole OL15 a Grafu 31 pro silové pole bsc1. Je vidět, že vývoj těchto úhlů kopíruje výskyt nekanonického stavu 35 v simulaci. To bylo pozorováno i v předchozí simulaci protein-DNA komplexu RAP1. Stejně, jako v předešlém případě tedy lze vyvodit, že pole bsc1 popisuje díky odlišnostem v parametrizaci ribózy a glykosidického úhlu konformaci 35 lépe. Pole OL15 pravděpodobně podceňuje populaci stavu 35, což je možná i díky nepřesnému popisu stavu $\gamma = t$, v souladu s předpoklady.

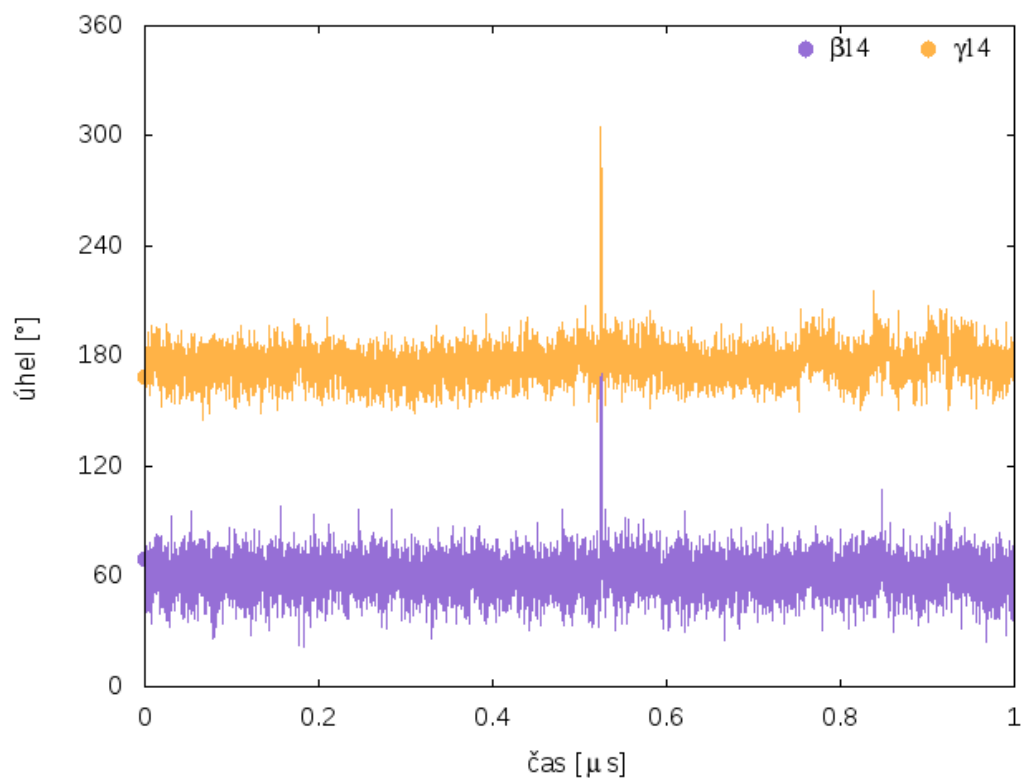
Graf 28. Časový vývoj úhlů β , γ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli OL15



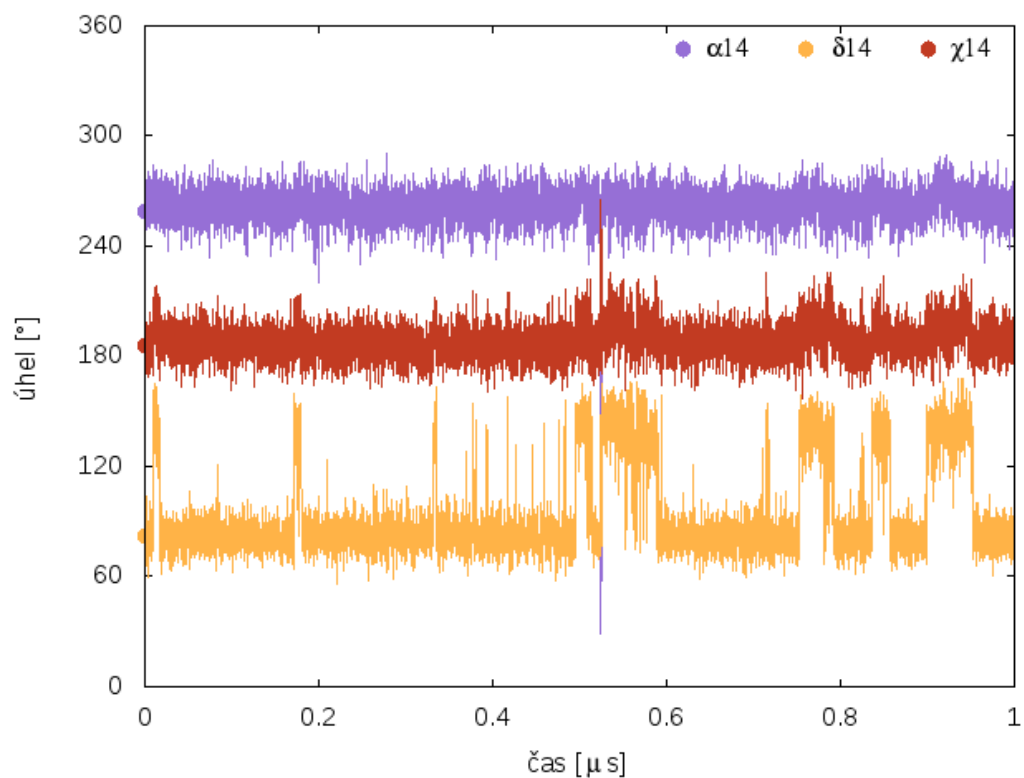
Graf 29. Časový vývoj úhlů α , δ , χ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli OL15



Graf 30. Časový vývoj úhlů β , γ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli bsc1

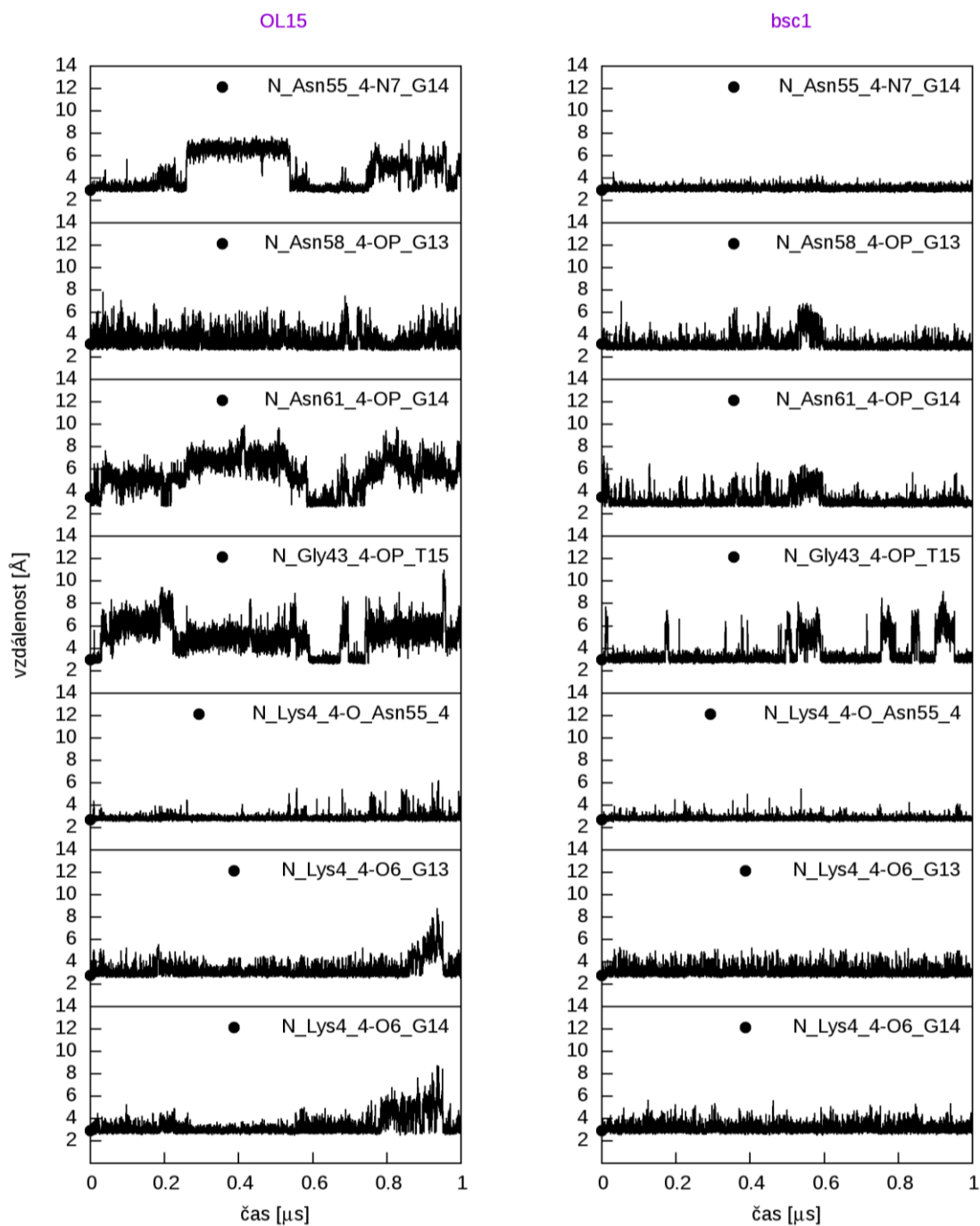


Graf 31. Časový vývoj úhlů α , δ , χ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli bsc1



Protože v tomto komplexu byl studovaný nekanonický stav 35 v poli OL15 málo stabilní a v poli bsc1 stabilní, rozhodla jsem se prozkoumat souvislost jeho stability se sítí vodíkových vazeb mezi proteinem a rezidui guanin 13 a guanin 14, mezi kterými se stav 35 nachází. Časový vývoj délky vodíkových vazeb je pro obě pole ukázán v Grafu 32. Je vidět, že v poli bsc1, ve kterém je stav 35 stabilní, jsou v průběhu simulace většinou zachovány také všechny související vodíkové vazby. U pole OL15, kde je stav 35 stabilní jen na začátku simulace (asi po 0,2 μ s), je méně stabilní i síť vodíkových vazeb. Zatímco některé nativní vodíkové vazby jsou stabilní po celou dobu simulace OL15, jiné jsou nestabilní. Nelze ale nalézt žádnou korelaci mezi nestabilitou některé vodíkové vazby a stavu 35. Z toho lze usuzovat, že stabilita nekanonického stavu 35 je určena spíše kvalitou popisu dihedrálních úhlů než kvalitou popisu vodíkových vazeb v empirickém potenciálu.

Graf 32. Měřené vzdálenosti vodíkových vazeb okolo rezidua guanin 14 ve struktuře λ represor protein-DNA komplexu pro obě silová pole



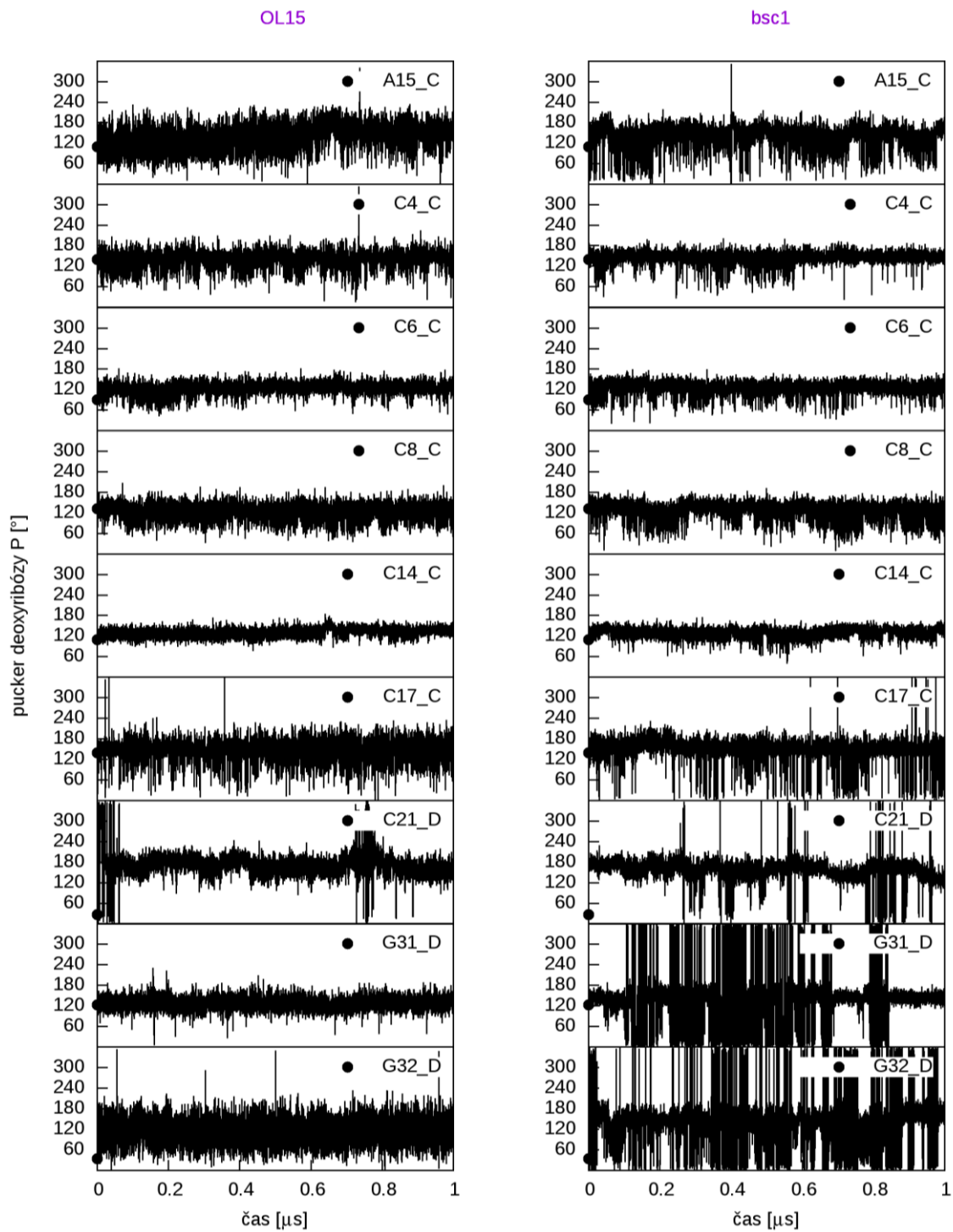
3.3.3 POPIS KONFORMACE DEOXYRIBÓZY

Při tvorbě komplexů proteinů s DNA dochází často k určitému konformačnímu přizpůsobení interagující DNA. Hlavním rysem tohoto přizpůsobení je změna konformace deoxyribózy z kanonické jižní (S) konformace do konformace severní (N), typické pro A-DNA, nebo alespoň konformace bližší k N. Protože kvalita popisu S/N rovnováhy v protein-DNA komplexech není pro empirické potenciály příliš prozkoumána, rozhodla jsem se prozkoumat ji pro RAP1 i λ represor komplex.

3.3.3.1 RAP1 PROTEIN DNA KOMPLEX

Ze všech reziduí molekuly DNA jsem vybrala pouze ty, u kterých byla konformace deoxyribózy výrazně vzdálena od kanonické konformace S pro B-formu. Časový vývoj pseudorotačních úhlů (puckerů) deoxyribózy (P) je ukázán v Grafu 33. Počáteční (experimentální) hodnota P je v grafech znázorněna tečkou na ose y. Je vidět, že pokud je počáteční hodnota P v okolí $\sim 120^\circ$, obě pole (OL15 i bsc1) ji většinou udrží a popisují ji tedy správně. Pokud je ale nativní hodnota v okolí N ($\sim 18^\circ$), mají obě pole tendenci posouvat P směrem k vyšším hodnotám (blíže ke kanonické B-DNA formě). Potvrdilo se tak naše podezření, že obě zmiňovaná pole nepopisují konformaci deoxyribózy dostatečně přesně.

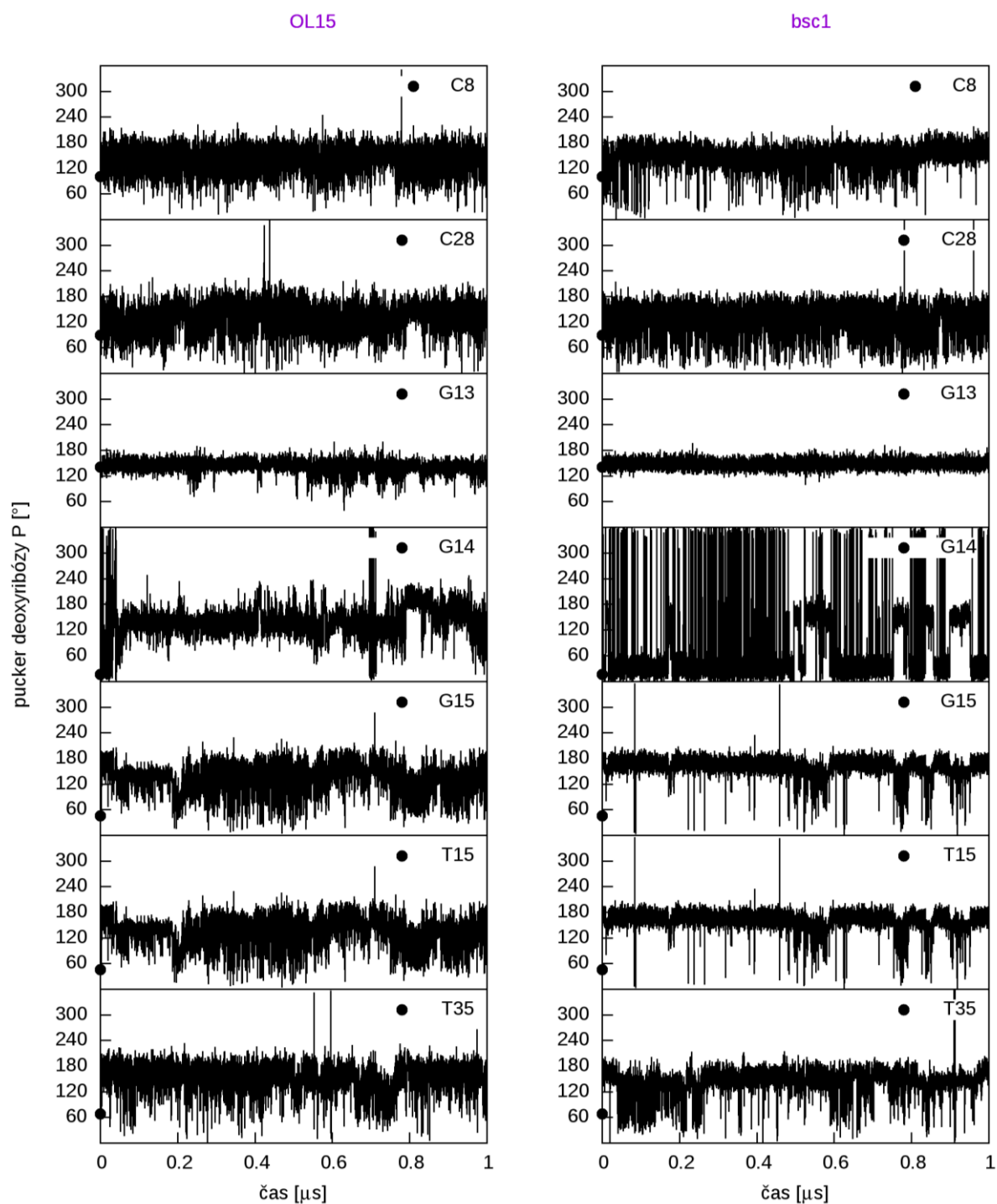
Graf 33. Časový průběh puckeru deoxyribózy, P, u reziduí, která se v krystalové struktuře odchylojí od B formy struktury RAP1 protein-DNA komplexu pro obě silová pole



3.3.3.2 λ REPRESOR PROTEIN DNA KOMPLEX

Také u λ represor protein-DNA komplexu jsem ze všech reziduí molekuly DNA vybrala pouze ty, u kterých byla konformace deoxyribózy výrazně vzdálena od kanonické konformace S pro B-formu. Časový vývoj pseudorotačních úhlů deoxyribózy (P) je ukázán v Grafu 34. Počáteční (experimentální) hodnota P je v grafech znázorněna tečkou na ose y. Jako u předchozího protein-DNA komplexu je vidět, že pokud je počáteční hodnota P v okolí $\sim 120^\circ$, obě pole (OL15 i bsc1) ji většinou udrží a popisují ji více méně správně. Pokud je ale výchozí hodnota v severní oblasti N ($\sim 18^\circ$), mají obě pole tendenci posouvat P směrem k vyšším hodnotám (blíže ke kanonické B-DNA formě). I u tohoto komplexu se tedy znovu potvrdilo, že obě zmiňovaná pole nepopisují konformaci deoxyribózy dostatečně přesně.

Graf 34. Časový průběh puckeru deoxyribózy, P, u reziduí, která se v krystalové struktuře odchylojí od B formy struktury λ represor protein-DNA komplexu pro obě silová pole



4 ZÁVĚR

Nukleové kyseliny jsou jednou z nejdůležitějších složek všech živých organismů. V chemii se k jejich studiu může využívat teoretická chemie, do které spadají molekulově dynamické simulace. Za pomoci těchto simulací se dá zjistit mnoho zajímavých informací. Výsledek simulací je ale ovlivněn tím, jaké silové pole (empirický potenciál) bylo použito. V této práci byla pro studium nekanonických konformací DNA použita dvě silová pole – OL15 a bsc1. Tato pole jsou si svými parametry velmi podobná, a tudíž byly výsledky z obou polí snadněji porovnatelné.

V této bakalářské práci byly molekulově dynamickými simulacemi studovány čtyři DNA struktury, které obsahovaly nekanonické konformery páteřních úhlů β a γ : kvadruplex, Hollidayova struktura a dva protein-DNA komplexy. U všech čtyř systémů byla studována stabilita a dynamika systémů. V případě kvadruplexu se ukázalo, že je stabilní stem kvadruplexu, avšak jeho smyčky v průběhu simulací mění svou konformaci více. Pole OL15 poskytovalo o něco menší RMSD než pole bsc1. U Hollidayovy struktury bylo RMSD v případě obou silových polí relativně velké, i když celková konformace molekuly se oproti výchozí krystalové struktuře příliš neměnila. V případě prvního protein-DNA komplexu (RAP1) byl krom RMSD studován ještě B faktor, který měl ukázat, jak flexibilní je část proteinu vážící se na DNA. RMSD komplexu bylo pro obě pole přijatelně vysoké a mezi poli byly jen malé rozdíly. Při studiu B faktoru se ukázalo, že sekundární struktura proteinu (helixy) je relativně rigidní a také části interagující s DNA mají spíše malé fluktuace. To znamená, že studované komplexy jsou silovými poli popsány spíše dobře. U druhého protein-DNA komplexu bylo studováno RMSD i B faktor jako v předchozím případě. RMSD opět dosahovalo hodnot kolem 2 Å, avšak pro tak velký komplex je toto RMSD přijatelné. B faktor byl pro obě silová pole vysoký, což značí velké oscilace reziduí.

Dále byla pro všechny systémy studována stabilita nekanonických stavů páteře β a γ . U kvadruplexu se nacházely dvě symetricky položená místa, kde se vyskytovala nekanonická konformace páteře β , γ , a to mezi cytosinem 2 a guaninem 3, které byly součástí sousedních vnitřních kvartetů kvadruplexu. Ukázalo se, že se stavy neudrží a touto změnou je ovlivněn i blízký úhel α . V obou silových polích tato situace byla velmi podobná. I v Hollidayově struktuře byla dvě symetricky položená místa, kde se nacházela

nekanonická konformace páteře β/γ . Nekanonické stavy se v průběhu simulací také většinou neudržely a změnily své konformace do kanonických hodnot. Jen v případě silového pole bsc1 docházelo k tomu, že stavy β a γ své konformace sice mění, ale mají tendence se vracet ke svým počátečním konformacím. V průběhu simulace vznikají v Hollidayově struktuře vodíkové vazby, které nebyly v původní experimentální struktuře pozorovány. Tvorba těchto pozorovaných vazeb je ale spíše artefaktem silových polí a MD simulace tedy nepopisují chování systému v tomto směru správně. U RAP1 protein-DNA komplexu dochází ke změnám stavů β/γ rezidua guaninu 32, avšak každém silovém poli jinak. V silovém poli OL15 oba stavy změni své konformace do kanonické a v té se stabilně drží, naopak u silového pole bsc1 se vrací ke svým výchozím (nativním) konformacím. U λ represor komplexu byly studovány stavy β i γ rezidua guanin 14. V silovém poli OL15 stavy své výchozí konformace relativně dlouhou dobu (0,2 μ s) udrží a následně se změni. U silového pole bsc1 se drží ve svých výchozích konformacích. Stabilita této konformace naznačuje, že může hrát důležitou strukturní roli.

Kromě nekanonických β/γ stavů jsem u protein-DNA komplexů zkoumala i to, jak dobře jsou vybranými silovými poli popsány konformace deoxyribózy. Ukázalo se, že obě silová pole destabilizují nativní severní (N) pucker u těch deoxyribóz, které se v krystalových strukturách v tomto nekanonickém stavu nacházely. V obou silových polích docházelo k rekonformaci do jižní (S) kanonické formy typické pro B-DNA.

Závěrem lze říci, že vybrané nekanonické β/γ stavy páteře DNA nejsou v současných silových polích pravděpodobně popsány dostatečně přesně. Původ tohoto problému spočívá z části v nepřesném popisu γ potenciálu, u kterého je příliš destabilizován stav $\gamma = t$. To je v souladu s předešlými studiemi, které to ukázaly u molekul RNA. Dále ale lze najít i nepřesnosti v dalších torzních úhlech, například χ a pseudorotačních úhlech, které popis nekanonických stavů β/γ také ovlivňují. Mé závěry ukazují, že současná empirická pole pro DNA nejsou dostatečně přesná a budou potřebovat další úpravy.

5 SUMMARY

Nucleic acids are one of the most important components of all living organisms. In chemistry, they can be studied by using methods of theoretical chemistry, which include molecular dynamics simulations. With the help of MD simulations, a lot of interesting information can be obtained. However, the results of the simulations are affected by quality of force field (empirical potential) used. In this work, two force fields - OL15 and bsc1 - were used to study of non-canonical DNA conformations. Because these force fields are very similar, they can be compared quite easily.

In this bachelor's thesis, four DNA structures containing non-canonical conformers of the backbone angles β and γ were studied by molecular dynamics simulations: quadruplex, Holliday junction and two protein-DNA complexes. Stability and dynamics of these four systems was studied. In the case of a quadruplex, it has been shown that while its stem is stable, its loops change their conformation considerably in the simulation. The OL15 field provided a slightly smaller RMSD than the bsc1 field. For the Holliday structure, the RMSD was relatively large for both force fields, although the overall conformation of the molecule did not change much compared to the initial crystal structure. In the case of the first protein-DNA complex (RAP1), in addition to RMSD, B factor was also studied to show how flexible the DNA-binding portion of the protein is. The RMSD of the complex was acceptable for both fields and there were only small differences between the force fields. The study of B factor showed that the secondary structure of the protein (helices) is relatively rigid and also the parts interacting with DNA have rather small fluctuations. That means that the studied complexes are described rather well by the force fields. For the second protein-DNA complex, both RMSD and B factor were studied as in the previous case. The RMSD again reached values around 2 Å, but for such a large complex this RMSD is acceptable. The B factor was high for both force fields, indicating large residue oscillations.

Furthermore, the stability of non-canonical states of the β and γ dihedrals of the sugar-phosphate backbone was studied for all systems. In the quadruplex, there were two symmetrically placed sites where the non-canonical conformation of the β , γ dihedrals occurred, between cytosine 2 and guanine 3, which were part of the adjacent inner quartets of the quadruplex. It turned out that the non-canonical states are not maintained in the MD simulation and this fact also affects the neighboring dihedral angle α . In both force fields, the situation was very similar. Also in the Holliday junction there were two

symmetrically placed places where the non-canonical conformation of the dihedrals β/γ was located. The non-canonical states were also unstable during the simulations and changed their conformations to canonical values. Only in the case of the bsc1 force field, did the β and γ dihedrals tend to return to their initial conformations, fluctuating between their native and canonical values. During the simulation, hydrogen bonds were formed in the Holliday junction, which were not observed in the original experimental structure. Formation of these hydrogen bonds is probably a force field artefact, which means that MD simulations do not describe the behavior of the system correctly in this respect. In the RAP1 protein-DNA complex, the β/γ state of guanine residue 32 behaves differently, in each force field. In both force fields it transitions to the canonical conformation in the beginning, but in the bsc1 force field it is able to return to its initial (native) conformation, while in OL15 this does not happen. In the λ repressor protein-DNA complex, the β and γ states of the guanine residue 14 were studied. In the OL15 force field, the states changed their initial conformations only after a relatively long time (0,2 μ s) and in the bsc1 force field they remained stable in their initial conformations. This suggests that non-canonical β/γ state plays an important structural role in this protein-DNA complex.

In addition to non-canonical β/γ states, I also studied how well the deoxyribose conformations are described by the chosen force fields in protein-DNA complexes. Both force fields have been shown to destabilize the native north (N) sugar pucker in those deoxyriboses, for which this non-canonical state was found in the crystal structures. In both force fields I observed reformation into the south (S) canonical form typical of B-DNA.

In conclusion, the selected non-canonical β/γ states of the DNA backbone are probably not described with sufficient accuracy in the current force fields. The origin of this problem lies, at least partially, in the inaccurate description of the γ potential, for which the $\gamma = t$ state is overly destabilized. This is in line with previous findings that have shown this for RNA molecules. However, we can also find inaccuracies in other torsion angles, such as χ and pseudorotation angles, which also affect the description of the non-canonical β/γ states. My conclusions show that the current empirical force fields for DNA are not accurate enough and will need further adjustment.

6 LITERATURA

- [1] K. M. V. Aklank Jain, Guliang Wang, “DNA Triple Helices: biological consequences and therapeutic potential,” pp. 1117–1130, 2009.
- [2] P. Karlson, *Základy biochemie*. Academia, 1971.
- [3] Stephen Neidle, *Principles of Nucleic Acid Structure*. 2008.
- [4] M. Bansal, “Sequence preference for BI / BII conformations in DNA : MD and crystal structure data analysis Sequence Preference for BI / BII Conformations in DNA : MD and Crystal Structure Data Analysis,” no. September 2005, 2014.
- [5] P. Čech, J. Kukul, J. Černý, B. Schneider, and D. Svozil, “Automatic workflow for the classification of local DNA conformations,” *BMC Bioinformatics*, vol. 14, no. 1, pp. 30–32, 2013.
- [6] M. Zgarbova, P. Jurečka, and P. Banáš, “Noncanonical α / γ Backbone Conformations in RNA and the Accuracy of Their Description by the AMBER Force Field,” pp. 2420–2433, 2017.
- [7] N. Špačková, “Tři jsou málo , pět je moc aneb seznamte se s kvadruplexy,” pp. 98–100, 2009.
- [8] P. Hubálková, “G-kvadruplexy v oblasti lidských telomer a jejich terapeutický potenciál,” *Chem. List.*, vol. 109, no. 12, pp. 918–922, 2015.
- [9] B. H. M. M. Brandt F. Eichman, Jeffrey M. Vargason, “The Holliday junction in an inverted repeat DNA sequence: Sequence effects on the structure of four-way junctions,” *Proc. Natl. Acad. Sci.*, vol. 97, no. 8, pp. 3971–3976, 2002.
- [10] F. A. Hays *et al.*, “How sequence defines structure: A crystallographic map of DNA structure and conformation,” *Proc. Natl. Acad. Sci.*, vol. 102, no. 20, pp. 7157–7162, 2005.
- [11] G. D. Stormo, *Introduction to Protein-DNA interactions*. John Inglis, 2013.
- [12] L. J. Beamert and C. Pabo, “Refined 14 % L Crystal Structure of the A Complex,” pp. 177–196, 1992.
- [13] P. König, R. Giraldo, L. Chapman, and D. Rhodes, “The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA,” *Cell*, vol. 85, no. 1, pp. 125–136, 1996.
- [14] A. R. Leach, *Molecular Modelling: Principles and Applications (2nd Edition)*. Pearson Education, 2001.
- [15] W. F. Van Gunsteren and H. J. C. Berendsen, “Molecular dynamics computer simulation. Method, application and perspectives in chemistry.,” *Sect. Title Phys. Org. Chem.*, vol. 102, no. 9, pp. 1020–1055, 1990.
- [16] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, “The missing term in

- effective pair potentials,” *J. Phys. Chem.*, vol. 91, no. 24, pp. 6269–6271, 1987.
- [17] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, “Comparison of simple potential functions for simulating liquid water,” *J. Chem. Phys.*, vol. 79, no. 2, pp. 926–935, 1983.
- [18] H. W. Horn *et al.*, “Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew,” *J. Chem. Phys.*, vol. 120, no. 20, pp. 9665–9678, 2004.
- [19] I. Shvab and R. J. Sadus, “Atomistic water models: Aqueous thermodynamic properties from ambient to supercritical conditions,” *Fluid Phase Equilib.*, vol. 407, pp. 7–30, 2015.
- [20] Y. Liu, E. Haddadian, T. R. Sosnick, K. F. Freed, and H. Gong, “A novel implicit solvent model for simulating the molecular dynamics of RNA,” *Biophys. J.*, vol. 105, no. 5, pp. 1248–1257, 2013.
- [21] R. J. W. D. A. Case, Thomas E. Cheatham, Tom Darden, Holger Gohlke, Ray Luo, Keneth M. Merz, JR., Alexey Onufriev, Carlos Simmerling, Bing Wang, “The Amber Biomolecular Simulation Programs,” *J. Comput. Chem.* 26, pp. 1668–1688, 2005.
- [22] B. R. Brooks, C. L. Brooks III., A. D. Mackerell, Jr., and E. Al., “CHARMM : The Biomolecular Simulation Program,” *J. Comput. Chem.* 30, pp. 1545–1614, 2009.
- [23] C. I. Bayly *et al.*, “A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules,” *J. Am. Chem. Soc.*, vol. 117, no. 19, pp. 5179–5197, 1995.
- [24] T. E. Cheatham, P. Cieplak, and P. A. Kollman, “A modified version of the cornell et al. force field with improved sugar pucker phases and helical repeat,” *J. Biomol. Struct. Dyn.*, vol. 16, no. 4, pp. 845–862, 1999.
- [25] J. Wang, P. Cieplak, and P. A. Kollman, “How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?,” *J. Comput. Chem.*, vol. 21, no. 12, p. 1049, 2000.
- [26] A. Pérez *et al.*, “Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ conformers,” *Biophys. J.*, vol. 92, no. 11, pp. 3817–3829, 2007.
- [27] M. Zgarbová, F. J. Luque, J. Šponer, M. Otyepka, and P. Jurečka, “A novel approach for deriving force field torsion angle parameters accounting for conformation-dependent solvation effects,” *J. Chem. Theory Comput.*, vol. 8, no. 9, pp. 3232–3242, 2012.
- [28] M. Zgarbová *et al.*, “Refinement of the Cornell et al. Nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles,” *J. Chem. Theory Comput.*, vol. 7, no. 9, pp. 2886–2902, 2011.

- [29] M. Krepl *et al.*, “Reference simulations of noncanonical nucleic acids with different χ variants of the AMBER FORCE Field: Quadruplex DNA, quadruplex RNA, and Z-DNA,” *J. Chem. Theory Comput.*, vol. 8, no. 7, pp. 2506–2520, 2012.
- [30] M. Zgarbová, F. J. Luque, J. Šponer, T. E. Cheatham, M. Otyepka, and P. Jurečka, “Toward improved description of DNA backbone: Revisiting epsilon and zeta torsion force field parameters,” *J. Chem. Theory Comput.*, vol. 9, no. 5, pp. 2339–2354, 2013.
- [31] R. Galindo-Murillo *et al.*, “Assessing the Current State of Amber Force Field Modifications for DNA,” *J. Chem. Theory Comput.*, vol. 12, no. 8, pp. 4114–4127, 2016.
- [32] M. H. M. Olsson, C. R. SØndergaard, M. Rostkowski, and J. H. Jensen, “PROPKA3: Consistent treatment of internal and surface residues in empirical pK_a predictions,” *J. Chem. Theory Comput.*, vol. 7, no. 2, pp. 525–537, 2011.
- [33] M. Zgarbova, J. Šponer, M. Otyepka, T. E. Cheatham, R. Galindo-murillo, and P. Jurec, “Refinement of the Sugar – Phosphate Backbone Torsion Beta for,” 2015.
- [34] P. Kührová *et al.*, “Correction to: Improving the performance of the amber rna force field by tuning the hydrogen-bonding interactions(Journal of the American Chemical Society (2019)15:5 (3288–3305) DOI: 10.1021/acs.jctc.8b00955),” *J. Chem. Theory Comput.*, vol. 16, no. 1, pp. 818–819, 2019.

SEZNAM GRAFŮ

Graf 1. RMSD stemu struktury kvadruplexu pro silová pole OL15 a bsc1	31
Graf 2. RMSD smyčky struktury kvadruplexu pro silová pole OL15 a bsc1	31
Graf 3. RMSD Hollidayové struktury pro silová pole OL15 a bsc1	32
Graf 4. RMSD struktury RAP1 protein-DNA komplexu pro silová pole OL15 a bsc1.	33
Graf 5. B faktor RAP1 protein-DNA komplexu pro silová pole OL15 a bsc1 (upraveno pomocí programu Inkscape)	34
Graf 6. RMSD struktury λ represor protein-DNA komplexu pro silová pole OL15 a bsc1	35
Graf 7. B faktor λ represor protein-DNA komplexu pro silová pole OL15 a bsc1 (upraveno pomocí programu Inkscape).....	36
Graf 8. Časový vývoj úhlů β , γ , α v reziduu guanin 3 pro strukturu kvadruplexu v silovém poli OL15. Výchozí experimentální hodnoty jsou označeny tečkami na ose y.....	39
Graf 9. Časový vývoj úhlů β , γ , α v reziduu guanin 14 pro strukturu kvadruplexu v silovém poli OL15.....	39
Graf 10. Časový vývoj úhlů β , γ , α v reziduu guanin 3 pro strukturu kvadruplexu v silovém poli bsc1	40
Graf 11. Časový vývoj úhlů β , γ , α v reziduu guanin 14 pro strukturu kvadruplexu v silovém poli bsc1	40
Graf 12. Časový vývoj úhlů β , γ , α v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli OL15.....	42
Graf 13. Časový vývoj úhlů β , γ , α v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli OL15.....	42
Graf 14. Časový vývoj úhlů α , β , γ v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli bsc1	43
Graf 15. Časový vývoj úhlů α , β , γ v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli bsc1	43
Graf 16. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli OL15.....	44
Graf 17. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli OL15.....	44
Graf 18. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci B pro Hollidayovu strukturu v silovém poli bsc1	45

Graf 19. Časový vývoj úhlů α , δ , χ v reziduu cytosin 7 v řetězci D pro Hollidayovu strukturu v silovém poli bsc1	45
Graf 20. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli OL15 pro Hollidayovu strukturu v prvním řetězci	48
Graf 21. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli OL15 pro Hollidayovu strukturu v druhém řetězci.....	48
Graf 22. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli bsc1 pro Hollidayovu strukturu v prvním řetězci	49
Graf 23. Mění se velikosti vazeb mezi fosfáty v závislosti na čase v silovém poli bsc1 pro Hollidayovu strukturu v druhém řetězci.....	49
Graf 24. Časový vývoj úhlů β , γ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli OL15	51
Graf 25. Časový vývoj úhlů β , γ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli bsc1	51
Graf 26. Časový vývoj úhlů α , δ , χ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli OL15	52
Graf 27. Časový vývoj úhlů α , δ , χ v reziduu guanin 32 pro strukturu RAP1 protein-DNA komplexu v silovém poli bsc1	52
Graf 28. Časový vývoj úhlů β , γ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli OL15	54
Graf 29. Časový vývoj úhlů α , δ , χ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli OL15	54
Graf 30. Časový vývoj úhlů β , γ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli bsc1	55
Graf 31. Časový vývoj úhlů α , δ , χ v reziduu guanin 14 pro strukturu λ represor protein-DNA komplexu v silovém poli bsc1	55
Graf 32. Měření vzdálenosti vodíkových vazeb okolo rezidua guanin 14 ve struktuře λ represor protein-DNA komplexu pro obě silová pole	57
Graf 33. Časový průběh puckeru deoxyribózy, P, u reziduí, která se v krystalové struktuře odchylují od B formy struktury RAP1 protein-DNA komplexu pro obě silová pole	59
Graf 34. Časový průběh puckeru deoxyribózy, P, u reziduí, která se v krystalové struktuře odchylují od B formy struktury λ represor protein-DNA komplexu pro obě silová pole	61

SEZNAM OBRÁZKŮ

Obrázek 1. Struktura polynukleotidového řetězce DNA (vytvořeno pomocí programu ChemSketch).....	9
Obrázek 2. Párování bází – cytosin s guaninem (vlevo), thymin s adeninem (vpravo) (vytvořeno pomocí programu ChemSketch)	10
Obrázek 3. Struktura B-DNA (vytvořeno pomocí programu PyMOL)	11
Obrázek 4. Struktura A-DNA (vytvořeno pomocí programu PyMOL).....	12
Obrázek 5. Struktura Z-DNA (vytvořeno pomocí programu PyMOL)	13
Obrázek 6. Struktura triplexu (vytvořeno pomocí programu PyMOL)	14
Obrázek 7. Nukleotidová jednotka (vytvořeno pomocí programu ChemSketch)	15
Obrázek 8. Guaninová tetráda G·G·G·G (vlevo) a tetráda obsahující cytosin C·G·C·G (vpravo) obrázek převzat a upraven v programu ChemSketch) [9].....	18
Obrázek 9. Hollidayova struktura (vytvořeno pomocí programu PyMOL).....	19
Obrázek 10. Struktura λ represor opetátor protein-DNA komplexu (vytvořeno pomocí programu PyMOL).....	21
Obrázek 11. Struktura RAP1 protein-DNA komplexu (vytvořeno pomocí programu PyMOL).....	22
Obrázek 12. Výřez z posledního snímku simulace Hollidayovy struktury – nenativní vodíkové vazby mezi fosfátovou skupinou a aminoskupinami cytosinů (vytvořeno pomocí programu PyMOL, Inkscape).....	47
Obrázek 13. Srovnání krystalové geometrie (fialová) a posledního snímku 1 μ s simulace (růžová) Hollidayovy struktury (vytvořeno pomocí programu PyMOL)	47

SEZNAM TABULEK

Tabulka 1. Definice torzních úhlů (tabulka převzata a upravena) [4]	15
Tabulka 2. Hodnoty dihedrálních úhlů nekanonických konformací (převzato a upraveno) [5]	16
Tabulka 3. Přehled silových polí pro DNA.....	26
Tabulka 4. Konformace páteře ve 4 modelech NMR struktury 1A6H	30