

**Univerzita Palackého v Olomouci**

**Přírodovědecká Fakulta**

**KATEDERA BIOFYZIKY**

**DIPLOMOVÁ PRÁCE**

**Konformace cukrfofátové páteře a deoxyribózy v protein-DNA  
komplexech**



Vypracoval: Filip Černý

Vedoucí práce: doc. RNDr. Petr Jurečka, Ph.D.

Studijní obor: Molekulární biofyzika

Olomouc 2022

## **Bibliografická identifikace**

Jméno a příjmení: Filip Černý

Název práce: Konformace cukrfosfátové páteře a deoxyribózy  
v protein-DNA komplexech

Typ práce: Teoretická

Pracoviště: Katedra biofyziky

Vedoucí práce: doc. RNDr. Petr Jurečka, Ph.D.

Roh obhajoby: 2022

Abstrakt: Pomocí molekulární dynamiky (MD) je možné počítat Gibbsovu energii, predikovat vývoj molekulárních systémů v čase nebo simulovat vazbu ligandu do vazebného místa proteinu. Proto, abychom obdrželi co nejpřesnější výsledky, je nutné mít silové pole, které správně popisuje dané systémy. Tato práce se zaměřuje na ověřování kvality silových polí pro modelování nukleových kyselin v MD simulacích a konkrétně na parametr P cukerného kruhu a torzní úhly  $\delta$  a  $\chi$  při popisu A formy DNA v protein-DNA komplexech silovými poli OL15 a bsc1 v programu AMBER. Analýzou simulací pěti protein-DNA komplexů v obou silových polí bylo zjištěno, že ani jedno silové pole nepopisuje správně stabilitu A formu DNA a bylo by vhodné se při korekcích těchto polí zaměřit na parametr P a jeho popis stability C3'-endo konformace.

Klíčová slova: Molekulární dynamika, molekulová mechanika, silové pole, empirický potenciál, OL15, bsc1, parametr P cukerného kruhu, nukleové kyseliny, A-DNA, glykosidický úhel  $\chi$ , protein-DNA interakce

Počet stran: 69

Počet příloh: 0

Jazyk: Český

**Bibliographical  
identification**

Author's name: Filip Černý

Title of thesis: Conformation of sugar-phosphate backbone and deoxyribose in protein-DNA complexes.

Type of thesis: Theoretical

Department: Department of biophysics

Supervisor: doc. RNDr. Petr Jurečka, Ph.D.

The year of defence: 2022

Abstract: Molecular dynamics (MD) modeling can be used to calculate Gibbs energy, to predict the evolution of a molecular system over time or to simulate the binding of a ligand to the active site of a protein. For accurate molecular dynamics results, it is necessary to have a force field that describes the systems correctly. This work is focused on verification of quality of the force field used to model nucleic acids in MD simulations, specifically on the role of sugar pucker and torsion angles  $\delta$  and  $\chi$  in describing the A form of DNA in protein-DNA complexes simulated by the OL15 and bsc1 force fields in AMBER program. Analysis of the results of five simulations of protein-DNA complexes in both force fields indicates that neither force field correctly describes the A form of DNA. Next force field correction should be aimed at sugar pucker and its description of the C3'-endo conformation.

Key words: Molecular dynamics, molecular mechanics, force field, empiric potential, OL15, bsc1, sugar pucker P, A-DNA, glycosidic angle  $\chi$ , protein-DNA interaction

Number of pages: 69

Number of attachments: 0

Language: Czech

Rád bych poděkoval vedoucímu diplomové práce doc. RNDr. Petru Jurečkovi, Ph.D. za odborné vedení, rady a věcné připomínky, jak při zaučení s danými programy, tak při vypracování textu diplomové práce.

Diplomová práce byla řešena za podpory grantu č. 17-16107S Grantové agentury České republiky.

Prohlašuji, že jsem diplomovou práci vypracoval samostatně pod vedením doc. RNDR. Petra Jurečky, Ph.D. a za použití uvedené literatury.

V Olomouci dne .....

# Obsah

|          |  |    |
|----------|--|----|
| 1.       | Úvod.....                                    | 1  |
| 2.       | Přehled problematiky .....                   | 2  |
| 2.1.     | Molekulární dynamika.....                    | 2  |
| 2.1.1.   | Molekulová mechanika .....                   | 3  |
| 2.2.     | DNA.....                                     | 6  |
| 2.2.1.   | Struktura DNA .....                          | 6  |
| 2.2.2.   | Torzní úhly DNA .....                        | 7  |
| 2.2.3.   | Konformace 2-deoxyribosy v DNA.....          | 9  |
| 2.2.4.   | Helikální parametry nukleových kyselin ..... | 10 |
| 2.2.5.   | Formy DNA .....                              | 11 |
| 2.2.6.   | NtC rodiny (třídy) .....                     | 13 |
| 2.3.     | Proteiny.....                                | 13 |
| 2.3.1.   | Struktura proteinů.....                      | 13 |
| 2.4.     | Protein-DNA komplexy.....                    | 14 |
| 2.4.1.   | Interakce protein-DNA.....                   | 14 |
| 2.4.2.   | Použité protein-DNA komplexy .....           | 16 |
| 2.4.2.1. | 3PVI.....                                    | 16 |
| 2.4.2.2. | 3AAF.....                                    | 17 |
| 2.4.2.3. | 3KXT.....                                    | 17 |
| 2.4.2.4. | 2VJV.....                                    | 18 |
| 2.4.2.5. | 1DP7.....                                    | 19 |
| 3.       | Cíl práce .....                              | 20 |
| 4.       | Materiál a metody .....                      | 20 |
| 5.       | Výsledky .....                               | 24 |
| 5.1.     | Stabilita simulací.....                      | 24 |
| 5.2.     | 3PVI – endonukleasa <i>PvuII</i> .....       | 26 |

|      |                                      |    |
|------|--------------------------------------|----|
| 5.3. | 3AAF – WRN RecQ helikasa .....       | 35 |
| 5.4. | 3KXT – chromatin protein Cren7 ..... | 40 |
| 5.5. | 2VJV – transposasa TnpA.....         | 46 |
| 5.6. | 1DP7 – regulační faktor X1 .....     | 57 |
| 5.7. | Souhrnné výsledky .....              | 64 |
| 6.   | Diskuze .....                        | 69 |
| 7.   | Závěr.....                           | 74 |
| 8.   | Seznam literatury .....              | 76 |

## Seznam zkratek

|                  |   |
|------------------|---|
| 1DP7             | Komplex lidského regulačního faktoru s DNA                    |
| 2VJV             | Komplex dimeru IS608 transposasy TnpA s DNA                   |
| 3AAF             | Komplex WRN RecQ helikasy a DNA                               |
| 3KXT             | Chromatin protein Cren7 s DNA                                 |
| 3PVI             | Komplex mutantního proteinu endonukleasy <i>PvuII</i>         |
| A                | Adenin  |
| Arg              | Arginin   |
| Asp              | Asparagin   |
| C                | Cytosin   |
| -CO <sub>2</sub> | Karboxylová skupina   |
| DFT              | Density functional theory (Teorie funkcionálu hustoty)        |
| DNA              | 2-deoxyribonukleová kyselina                                  |
| FF               | Force field (silové pole)                                     |
| G                | Guanin  |
| Gln              | Glutamin  |
| HF               | Hartree-Focková (metoda)                                      |
| His              | Histidin  |
| HID              | Histidin s atomem H na $\delta$ atomu N imidazolové skupiny   |
| HIE              | Histidin s atomem H na $\epsilon$ atomu N imidazolové skupiny |
| HIP<br>skupiny   | Histidin s atomem H na obou atomech N imidazolové skupiny     |
| HLH              | Helix-loop-helix (helix-smyčka-helix)                         |
| hRFX1            | human regulatory factor X1 (lidský regulační faktor X1)       |
| HTH              | Helix-turn-helix (helix-otočka-helix)                         |
| HUH              | Histidin-hydrofobní aminokyselina-hystidin                    |
| Ile              | Isoleucin   |
| Leu              | Leucin  |
| Lys              | Lysin   |
| MD               | Molekulární dynamika  |

|                  |   |
|------------------|---|
| Mg <sup>2+</sup> | Horečnatý kation  |
| MM               | Molekulární mechanika   |
| Met              | Methionin   |
| NaCl             | Chlorid sodný   |
| NANT             | zkratka, pokud danému nukleotidu nebyla přiřazena žádná Ntc třída                                   |
| -NH <sub>2</sub> | Aminoskupina  |
| NtC              | diNucleosides conformational classes (dinukleotidové konformační třídy)                             |
| Phe              | Fenylalanin   |
| QM               | Kvantová mechanika  |
| R                | Purin   |
| RESP fit         | Restrained electrostatic potencial fit (fit elektrostatického potenciálu s omezením hodnoty náboje) |
| RMSD             | Root mean square deviation (střední kvadratická odchylka)   |
| RQC              | RecQ C-konec  |
| Ser              | Serin   |
| ssDNA            | Single stranded DNA (jednořetězcová DNA)  |
| T                | Thymin  |
| TnpA             | Transposasa A   |
| Tyr              | Tyrosin   |
| Val (V)          | Valin   |
| WFT              | Wave functional theory (Teorie vlnové funkce)   |
| Y                | Pyrimidin   |



# 1. Úvod

Molekulární dynamika je účinný nástroj, pomocí něž je možné lépe interpretovat experimentální data ze spektroskopických měření, počítat volné Gibbsovy energie různých molekulárních systémů, predikovat a simulovat vazby ligandů do aktivního místa enzymů apod. Abychom ale dostali pomocí molekulární dynamiky věrohodné a přesné výsledky, je zapotřebí, aby model správně popisoval energii systému a jeho vývoj v čase.

V molekulové mechanice, která se používá pro simulaci větších systémů jako jsou například celé proteiny nebo řetězce aminokyselin, je modelem tzv. silové pole. Silové pole popisuje systém pomocí jednoduchých základních interakcí, jakými jsou např. Coulombův zákon nebo Lennard-Jonesův potenciál.

V této diplomové práci provádím simulace protein-DNA komplexů pomocí silových polí OL15 a bsc1 a zaměřuji se na stabilitu A formy a přechodných A/B forem nukleové kyseliny v průběhu simulace. Z průběhů simulací plánuji studovat stabilitu daných forem nukleotidů nukleové kyseliny a porovnat použitá silová pole. Výsledky nám pak mohou napovědět, zda je popis A formy DNA vyhovující, nebo zda jsou potřeba nové korekce pro použitá silová pole.

## 2. Přehled problematiky

### 2.1. Molekulární dynamika

Molekulární dynamika (MD) je teoretický nástroj vycházející z výpočetní chemie, který umožňuje modelovat chování molekulárních systémů (Tuckerman a Martyna 2000). Dříve byla MD používána pro lepší interpretaci spekter obdržených z různých spektroskopických metod, jako např. Ramanova spektroskopie, infračervená spektroskopie a další. Dnes se používá nejen pro výpočet volné energie molekul, studování struktur, dynamiky a funkcí biomolekul, ale také pro návrh nových léčiv a predikci vazby proteinu a ligandu do aktivního místa.

V MD může být využíváno dvou přístupů, kvantové mechaniky (QM – z anglického quantum mechanics) a molekulární mechaniky (MM). Kvantová mechanika, jak už z názvu vyplývá, využívá principů kvantové chemie a fyziky a pracuje s elektrony v systému. Oproti tomu MM ignoruje pohyb elektronů a počítá energii systému jako funkci pozici jader (Leach 2001). Proto se QM uplatňuje především pro malé molekuly, protože s nárůstem počtu atomů roste exponenciálně výpočetní náročnost, a pro větší molekuly, jakými jsou např. lipidy, proteiny nebo nukleové kyseliny, se používá MM.

Dva hlavní směry využívané v QM jsou Teorie vlnové funkce (WFT – z anglického Wave Function Theory) a Teorie funkcionálu hustoty (DFT – z anglického Density Functional Theory). Při WFT je stav popsán vlnovou funkcí a pro řešení Schrödingerovy rovnice se využívá Hartree-Fockova (HF) metoda a post-HF metody. Naproti tomu DFT pracuje s elektronovou hustotou systému a energie je výsledkem funkcionálu elektronové hustoty systému. Výhodou QM metod jsou relativně přesně výsledky, v závislosti na použité metodě. Nevýhodou je velká výpočetní náročnost, která roste s velikostí zkoumaného systému.

Protože v této práci jsou použity metody MM, dále je popsána rozsáhleji. Oproti QM metodám dává MM méně přesné výsledky, ale výrazně klesá její výpočetní náročnost, a proto je možné pomocí MM počítat i větší a složitější systémy. Některé biochemické systémy, např. enzymy, jsou příliš velké, aby byly popsány pomocí QM, ale nelze je vhodně popsat pomocí MM, protože ta neumožňuje popsat tvorbu a zánik vazeb.

Proto byla vytvořena hybridní technika QM/MM, která pro podrobný popis malého místa, např. aktivního místa enzymu, využívá QM a pro zbytek systému využívá metod MM. Díky této kombinaci lze získat přesné výsledky za rozumný výpočetní čas (Groenhof 2013).

Výsledkem MD je trajektorie částic systému, která se získá integrací Newtonových pohybových rovnic. Trajektorie dává informaci o tom, jak se měnila pozice a rychlost částic v čase (jednotlivých krocích) v průběhu simulace. Pro provedení MD jsou potřeba tři základní věci (González 2001):

- Počáteční podmínky systému, tj. pozice částic, kterou můžeme získat např. z krystalografické analýzy látek.
- Model pro výpočet energií, ať už metody QM nebo MM.
- Okrajové podmínky, tj. jak se systém bude chovat, pokud se nějaké částice od sebe velmi vzdálí.

### 2.1.1. Molekulová mechanika

Molekulová mechanika, jak již bylo zmíněno, pracuje pouze s pozicemi jader příslušných atomů, díky aplikaci Born-Oppenheimerovy aproximace. Molekulová mechanika využívá jednoduchých principů interakcí, jakými jsou např. natahování vazeb, rotace vazeb apod. I přesto, že jsou zde využity jednoduché fyzikální zákony, např. Hookův zákon, dosahuje MM velmi přijatelných výsledků (Leach 2001).

Modelem pro MM je silové pole (FF – z anglického Force Field), které popisuje interakci částic v systému. Silová pole jsou dána pomocí jednoduchých vztahů popisujících interakce částic. Nejjednodušší popis FF je dán vztahem (1), ve kterém se vyskytuje pět částí. První tři části vztahu popisují intramolekulové síly a zbylé dva členy popisují intermolekulové síly (Leach 2001).

$$U(r^N) = \sum_{vazby} \frac{k_i}{2} (l_i - l_{i0})^2 + \sum_{\text{úhly}} \frac{k_i}{2} (\theta_i - \theta_{i0})^2 + \sum_{torze} \frac{V_i}{2} (1 + \cos(n\omega_i - \gamma)) + \sum_{i=0}^N \sum_{j>i}^N \left[ \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right] + \sum_{i=0}^N \sum_{j>i}^N \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}. \quad (1)$$

První člen popisuje změnu energie při natahování vazeb, kde  $k_i$  je vazebná konstanta,  $l_i$  je velikost natažení  $i$ -té vazby a  $l_{i0}$  je velikost natažení  $i$ -té vazby v rovnovážné poloze. Druhý člen popisuje deformaci vazebného úhlu, kde  $\theta_i$  je velikost  $i$ -tého úhlu a  $\theta_{i0}$  je velikost  $i$ -tého úhlu v minimu energie (rovnovážná velikost úhlu). Třetí člen ve vztahu (1) popisuje rotaci torzních úhlů v molekule.  $V_i$  je výška energetické bariéry, tj. velikost max. energie, která může nastat rotací torzního úhlu o  $360^\circ$ .  $n$  je periodičita úhlu udávající počet minimálních energetických stavů při rotaci torzního úhlu o  $360^\circ$ .  $\omega_i$  je velikost úhlu, který  $i$ -tý torzní úhel nabývá a  $\gamma$  je fázový faktor udávající, ve kterých úhlech se nachází minima energie. Van der Waalsovy síly působící mezi atomy systému jsou popsány čtvrtým členem pomocí Lenard-Jonesova potenciálu. Konstanty  $A_{ij}$  a  $B_{ij}$  lze v tomto vztahu vyjádřit jako  $A_{ij} = 4\epsilon_{ij}\sigma_{ij}^{12}$  a  $B_{ij} = 4\epsilon_{ij}\sigma_{ij}^6$ ,  $\epsilon_{ij}$  je minimální potenciální energie na Lenard-Jonesově křivce  $i$ -tého a  $j$ -tého atomu,  $\sigma_{ij}$  je vzdálenost  $i$ -tého a  $j$ -tého atomu, při které je potenciální energie rovna 0 a  $r_{ij}$  je vzdálenost  $i$ -tého a  $j$ -tého atomu. Poslední člen popisuje Coulombovskou interakci dvou atomů, kde  $q_i$ ,  $q_j$  jsou náboje  $i$ -tého a  $j$ -tého atomu,  $\epsilon_0$  je permitivita vakua a  $r_{ij}$  je vzdálenost  $i$ -tého a  $j$ -tého atomu.

Proto, aby bylo možné simulovat určité molekuly, jsou zapotřebí programy, které budou provádět dané výpočty. První počítačové výpočty vlastností molekul byly provedeny v druhé polovině 40. let dvacátého století. Velký rozmach MM metod pak nastal v 60. letech dvacátého století (Poltev 2015). Od té doby bylo vyvinuto a stále je vyvíjeno mnoho programů pro MM simulace. Mezi nejpoužívanější programy pro simulace biomakromolekul dnes patří AMBER (Assisted Model Building with Energy Refinement), CHARMM (Chemistry at Harvard Macromolecular Dynamics) a GROMACS (GRONingen MACHine for Chemical Simulation). V této diplomové práci je používán program AMBER, a proto zde bude lehce nastíněn vývoj silových polí pro tento program.

Program AMBER byl poprvé popsán v roce 1981 Wienerem a Kollmanem (1981), kteří později dodali podrobnější popis prvního silového pole pro popis nukleových kyselin (Wiener a kol. 1984). O jedenáct let později Cornell a spol. (1995) publikovali tzv. druhou generaci silového pole obvykle označovaného jako ff94, ve kterém při parametrizaci použili větší bázevý set a RESP (restrained electrostatic potencial) fitování pro lepší popis elektrostatického potenciálu atomů. Na konci 90. let proběhly dvě korekce, ff98 (Cheatham 1999) a ff99 (Wang a kol. 2000), která je základem pro dnes používaná silová pole. Ovšem simulace pomocí těchto polí byly stabilní maximálně po dobu 10 ns pro DNA a pro RNA pouze 1 ns. Následně bylo publikováno několik dalších korekcí, pro nukleové kyseliny jsou však důležité následující. V roce 2007 Pérez a kol. upravil popis torzních úhlů  $\alpha$  a  $\gamma$  (úhly páteře nukleových kyselin budou popsány níže). Toto silové pole bylo pak označováno jako parmbsc0 a velmi výrazně se zvýšila stabilita simulací, pro Dickersonův dodecamer až na min. 200 ns a pro RNA až na min. 10 ns. Později byly publikovány korekce pro torzní úhel  $\chi$  (Zgarbová a kol. 2011) a (Krepl a kol. 2012), pole byly označováno jako  $\chi_{OL3}$  a  $\chi_{OL4}$  (OL je zkratkou pro Olomouc, kde byly korekce vyvinuty), které pomohly k vybalancování *syn* a *anti* stavů glykosidického úhlu. Tyto korekce dosahovaly nejlepších výsledků při použití v kombinaci s bcs0. Zgarbová a kol. (2013) vydali další korekci ( $\epsilon\zeta_{OL1}$ ) pro torzní úhly  $\epsilon$  a  $\zeta$ , které napomohly k lepšímu popisu cukrfosfátové páteře B-DNA a guaninovým kvadruplexům. V roce 2015 pak byly vydány další korekce silových polí, a to konkrétně parmbsc1 (Zgarbová a kol. 2013) a  $\beta_{OL1}$  (Ivani a kol. 2016). V případě parmbsc1 bylo použito korekce z parmbsc0 na torzní úhly  $\alpha$  a  $\gamma$ , ale torzní úhly  $\epsilon$ ,  $\zeta$ ,  $\chi$  a parametr P (viz níže) byly nově parametrizovány. V korekci  $\beta_{OL1}$  byl nově parametrizován torzní úhel  $\beta$ . Původně tak bylo zamýšleno pouze pro stabilitu Z-DNA, ale tato korekce napomohla k lepšímu popisu i B-DNA i guaninových kvadruplexů.

Přestože různé FF používají pro popis energie systému stejnou nebo lehce modifikovanou rovnici (1), konečné výsledky jsou odlišné. Je to dáno tím, že používají různé hodnoty konstant pro různé atomové typy. Základem při vytváření každého FF je zvolení tzv. atomového typu. To znamená, že atomy, které jsou si podobné jak po fyzikální, tak chemické stránce, budou brány jako stejné (Wiener a kol. 1984). Například atom uhlíku může mít několik atomových typů. Jedním typem je např.  $sp^3$  hybridizace se dvěma atomy vodíku, dalším může být  $sp^3$  hybridizace se třemi atomy vodíku, dále  $sp^2$  hybridizace v šestičlenném aromatickém kruhu atd. Pro dané typy atomů

jsou pak určeny konstanty, které jsou použity v rovnici (1). Tyto konstanty jsou získávány pomocí kvantově-mechanických výpočtů nebo z krystalografických, spektroskopických a jiných experimentálních dat.

V této práci je použita poslední korekce  $\beta_{OL1}$  z roku 2015 (Zgarbová a kol. 2015). Jelikož se tato korekce používá společně s předešlými korekcemi a celý název všech použitých by byl velmi složitý ( $ff99bsc0\beta_{OL1}\epsilon\zeta_{OL1}\chi_{OL4}$ ), používá se souhrnné označení OL15. Druhým použitým polem v této diplomové práci je parmbsc1 (dále zkracováno na bsc1). V tabulce 1 je pro OL15 i bsc1 přehledně shrnuto, z kterého pole vycházejí a jaké další korekce jsou součástí použitých silových polí.

**Tabulka 1: Použitá silová pole a příslušní korekce torzních úhlů, které jsou použity.**

| Silové pole | Základ | $\alpha/\gamma$ | $\beta$       | $\epsilon/\zeta$      | $\chi$    | P    |
|-------------|--------|-----------------|---------------|-----------------------|-----------|------|
| OL15        | ff99   | bsc0            | $\beta_{OL1}$ | $\epsilon\zeta_{OL1}$ | $X_{OL4}$ | -    |
| bsc1        | ff99   | bsc0            | bsc1          | bsc1                  | bsc1      | bsc1 |

## 2.2. DNA

### 2.2.1. Struktura DNA

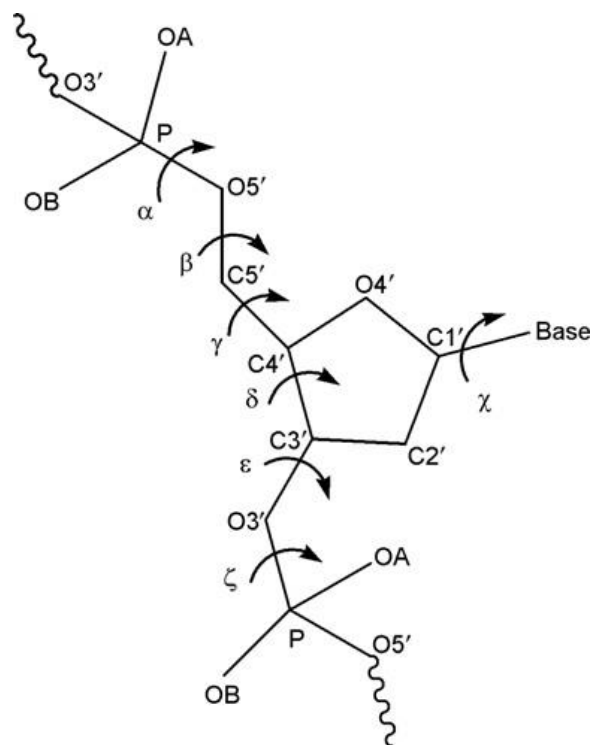
DNA neboli 2-deoxyribonukleová kyselina je nositelka dědičné informace všech živých organismů a nachází se v buněčném jádře a mitochondriích všech buněk. DNA se skládá ze tří částí, cukerné složky, báze a fosfátového zbytku. Jako cukr zde vystupuje 2-deoxyribosa, která nemá hydroxylovou OH skupinu na druhém uhlíku cyklu, a proto je DNA méně reaktivní než RNA. V DNA se vyskytují 2 typy bází, purinové a pyrimidinové. Mezi purinové báze patří adenin (A) a guanin (G), mezi pyrimidinové pak thymin (T) a cytosin (C). Navázáním báze na 1' uhlík 2-deoxyribosy vzniká nukleosid. Zbytek kyseliny fosforečné je navázán na 5' uhlík 2-deoxyribosy a dohromady s bází tvoří nukleotid. Následující nukleotid je vždy navázán přes fosfát na 3' uhlík

2-deoxyribose předchozího nukleotidu. Díky zápornému náboji fosfátového zbytku je celá molekula DNA záporně nabitá (Koolman a Röhm 2012).

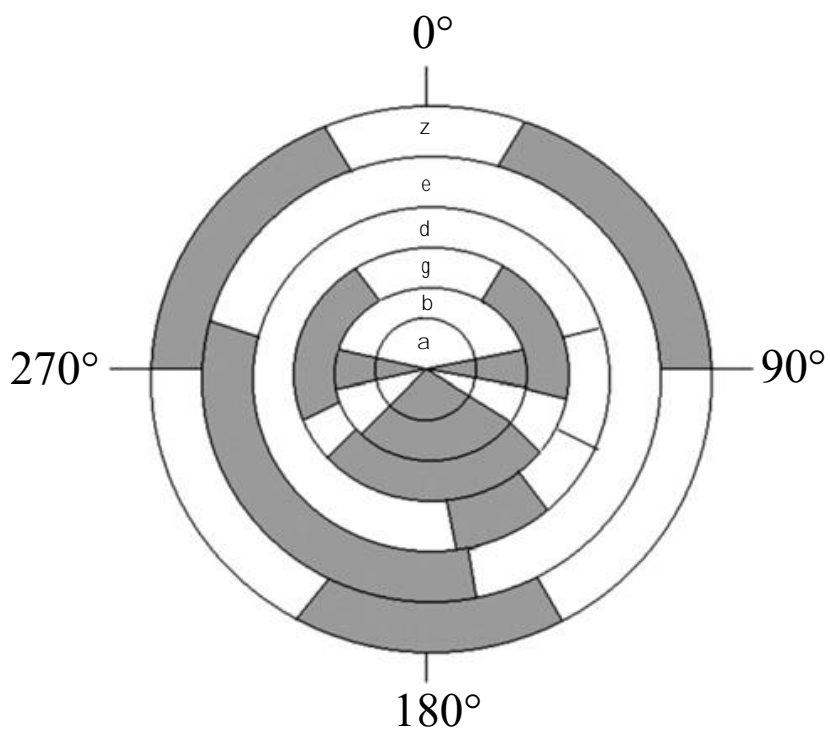
Primární struktura DNA je dána pořadím nukleotidů v řetězci DNA a podle konvence se pořadí zapisuje od 5' konce ke 3' konci řetězce. I když se DNA může vyskytovat jako jedno řetězcová DNA (ssDNA – single stranded DNA), nejčastěji se vyskytuje ve dvouvláknové struktuře, ve které se vlákna obtáčí kolem společné osy. Tato struktura se nazývá dvoušroubovice. Díky takovému obtáčení vzniká na dvoušroubovici malý a velký žlábek, do kterých pak mohou být navázány různé látky, např. proteiny. Druhé vlákno je komplementárně párováno s prvním vláknem pomocí vodíkových vazeb. Komplementární neboli Watson-Crickovské párování bází spočívá v tom, že adenin je párován s thyminem pomocí dvou vodíkových vazeb a cytosin je párován s guaninem za využití tří vodíkových vazeb (Bruce a kol. 2008).

### 2.2.2. Torzní úhly DNA

Fosfodiesterová páteř oligonukleotidů může být popsána šesti torzními úhly  $\alpha - \zeta$  (obr. č. 1). Další torzní úhel, který můžeme v DNA pozorovat, je glykosidický úhel  $\chi$ , který se na vazbě nachází mezi cukernou složkou a bází. Prostorové uspořádání páteře umožňuje daným úhlům nacházet se jen v určitých hodnotách (obr. č. 2) (Niedle 2008).



**Obr. č. 1:** Torzní úhly  $\alpha - \zeta$  na cukrfosfátové páteři na nukleotidu a glykosidický úhel  $\chi$ , převzato z *Niedle 2008*.



**Obr. č. 2:** Konformační kruh s povolenými oblastmi torzních úhlu (šedé části), převzato a upraveno z *Niedle 2000*.



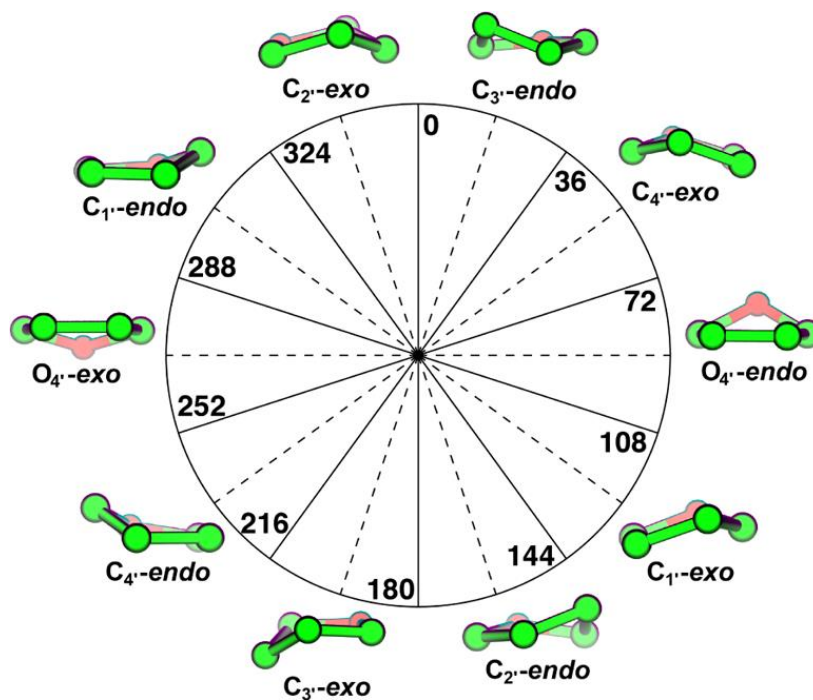
Glykosidický úhel  $\chi$  je definován mezi čtyřmi atomy a to O4'-C1'-N9-C4 pro puriny a O4'-C1'-N1-C2 pro pyrimidiny. O4' a C1' značí atomy 2-deoxyribosy, N1 a C2 pak atomy báze. Glykosidický úhel  $\chi$  se rozděluje do dvou kategorií, *syn*, která nabývá hodnot od 0 – 90°, a *anti*, která je v rozmezí 180 – 240° (Niedle 2008). Konformace anti se často v literatuře ještě dělí na *anti*, která se nachází okolo 200°, a tzv. *high-anti*, která je okolo 260°.

### 2.2.3. Konformace 2-deoxyribosy v DNA

Cyklická 2-deoxyribsa není v řetězci DNA planární, ale vždy je jeden z atomů více vysunut nad a sousedící atom mírně pod rovinu kruhu (endo konformace) nebo naopak, jeden atom více vysunut pod a sousedící atom mírně nad (exo konformace) rovinou kruhu. Například v C3'-endo konformaci je třetí uhlík nad rovinou kruhu a druhý uhlík je vysunut lehce pod rovinu. V jaké konformaci se cukerný kruh nachází, udává parametr P tzv. fázový úhel pseudorotace, který je dán vztahem

$$\tan P = \frac{(\tau_4 + \tau_1) - (\tau_3 + \tau_0)}{2 \cdot \tau_2 \cdot (\sin 36^\circ + \sin 72^\circ)}$$

kde  $\tau_0 - \tau_4$  jsou torzní úhly v ribosovém kruhu (Niedle 2008). Porovnáním parametru P s pseudorotačním kruhem furanosy se určí konformace 2-deoxyribosy v DNA (obr č. 3).

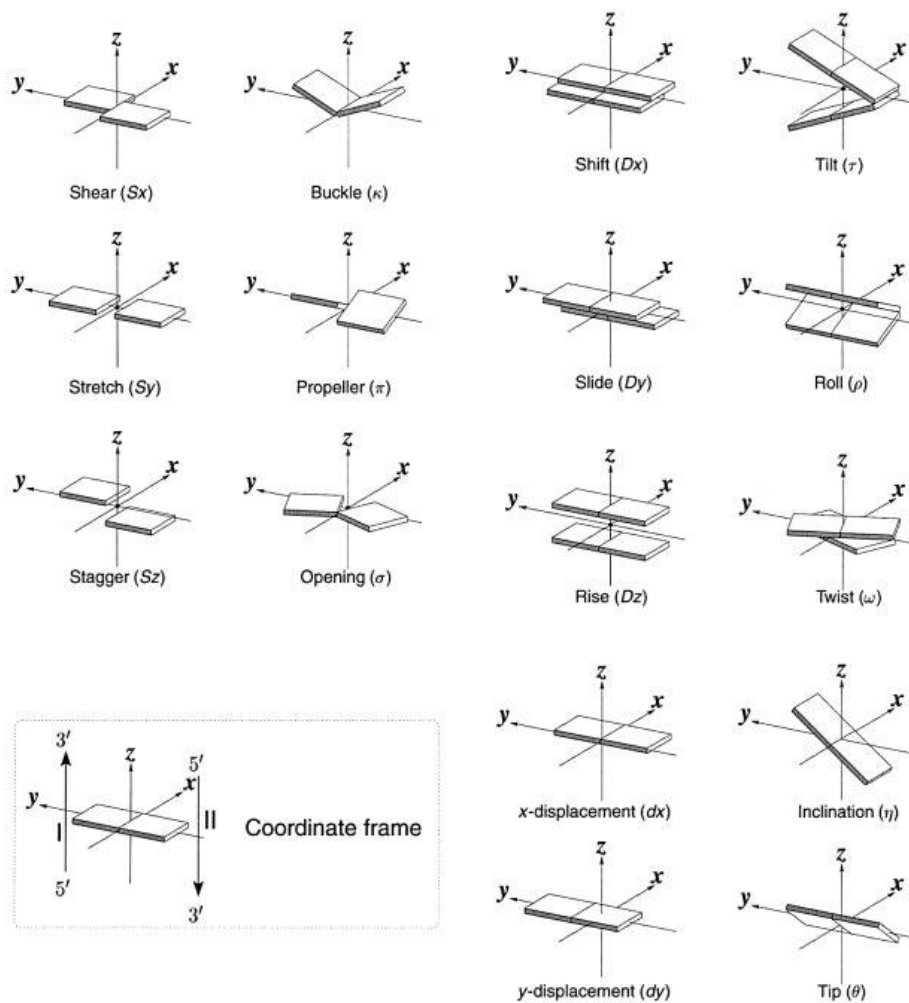


**Obr. č. 3: Pseudorotační kruh furanosy, převzato z Li a Szostak 2014.**

Nejčastěji se v DNA vyskytuje 2-deoxyribosa ve dvou konformacích, a to C3'-endo, která je často nazývána jako severní konformace, a C2'-endo, která bývá nazývána jako jižní konformace. Označení severní a jižní je dáno polohou na pseudorotačním kruhu furanosy (Niedle 2008).

#### 2.2.4. Helikální parametry nukleových kyselin

V nukleových kyselinách lze kromě výše uvedených torzních úhlů určovat tzv. helikální parametry (Dickerson 1989). Jedná se o translace nebo rotace dvou nad sebou nacházejících se párů bází nebo samostatného páru bází vůči osám x, y, z nebo vůči sobě (viz obr. 4). Helikální parametry je možné rozdělit do dvou skupin podle toho, zda se jedná o pohyb dvou nad sebou nacházejících se párů bází, označující se jako parametry bázových stepů, nebo o pohyb samotného páru bází, tzv. parametry bázových párů. Mezi parametry bázových stepů se řadí twist, roll, tilt, rise, slide a shift a mezi parametry bázových párů patří tip, stagger, stretch, shear, opening propeler twist, buckel, inclination (naklonění) a displacement (posunutí). Některé z těchto parametrů lze použít i pro rozlišení různých forem DNA (viz tabulka 2).



Obr. č. 4: Helikální parametry nukleových kyselin, převzato z Lu a Olson (2003).

### 2.2.5. Formy DNA

DNA se nejčastěji vyskytuje v B konformaci tzv. B-DNA. Dvoušroubovice B-DNA je pravotočivá, úzká, dlouhá a obsahuje 10,5 párů bází na otočku. Osa dvoušroubovice vede přes páry bází, které jsou skoro kolmé na tuto osu. Díky tomuto uspořádání vzniká úzký malý žlábek a široký velký žlábek. 2-deoxyribosa je v konformaci C3'-endo a glykosidický úhel  $\chi$  je v konformaci *high-anti*. Druhou nečastější formou, ve které se DNA může nacházet je A-DNA. Ta je oproti B-DNA kratší a širší. Je pravotočivá, obsahuje 11 párů bází na otočku a báze jsou vůči ose otáčení nakloněny. Vlivem toho vzniká mělký malý žlábek. A forma DNA má glykosidický úhel  $\chi$  v konformaci *anti* a 2-deoxyribosa je v konformaci C2'-endo (Dickerson 1992). Další formou, ve které se DNA může nacházet je Z-DNA. Jako jediná je levotočivá a oproti B-DNA je užší

a delší. Obsahuje 12 párů bází na otočku a osa otáčení vede přes malý žlábek (Dickerson a kol. 1982). Strukturální vlastnosti DNA forem jsou zaznamenány v tabulce č. 2, ve které jsou použity hodnoty pro velikost malého a velkého žlábků z práce Hassan a Calladine (1998), stejný postup pro výpočet velikosti žlábků byl zvolen i v programu AMBER.

Dalším rozdílem mezi formami A a B je X posunutí (x displacement), které je definováno jako posunutí párů bází po ose x od osy procházející dvoušroubovicí. V případě A-DNA je toto posunutí okolo 5 Å a pro B-DNA 0 Å (Zgarbová a kol. 2018), to znamená, že páry bází se nacházejí na ose dvoušroubovice. X posunutí je dobře vidět při pohledu na modely DNA shora. A-DNA a B-DNA se také liší v náklonu bází vůči ose dvoušroubovice. B-DNA je kolmá vůči této ose a naklonění bází je tedy 0° podle zavedené konvence, kdežto u A-DNA jsou báze lehce nakloněny asi o 20° (Zgarbová a kol. 2018). Tyto rozdíly mezi A a B formou DNA jsou zaznamenány v tabulce č. 2.

**Tabulka 2 Strukturální vlastnosti forem DNA.** Informace převzaty z Dickerson a kol. (1982), El Hassan a Calladine (1998), Sinden (1994) a Zgarbová a kol. (2018).

| Vlastnosti               | A-DNA                  | B-DNA                       | Z-DNA                           |
|--------------------------|------------------------|-----------------------------|---------------------------------|
| PB/ot.                   | 11                     | 10.5                        | 12                              |
| Otáčivost                | Pravotočivá            | Pravotočivá                 | Levotočivá                      |
| Průměr [Å]               | 23                     | 20                          | 18                              |
| Konformace 2-deoxyribosy | C3'-endo               | C2'-endo                    | C: C2'-endo<br>G: C2'-exo       |
| Glykosidický úhel $\chi$ | <i>Anti</i><br>(~200°) | <i>High-anti</i><br>(~260°) | C: <i>anti</i><br>G: <i>syn</i> |
| Malý žlábek [Å]          | 16                     | 11                          | -                               |
| Velký žlábek [Å]         | 14                     | 17                          | -                               |
| X posunutí [Å]           | ~ 5                    | ~ 0                         | -                               |
| Naklonění [Å]            | ~ 20                   | ~ 0                         | -                               |

## 2.2.6. NtC rodiny (třídy)

Pro lepší klasifikaci nukleových kyselin, jak DNA, tak RNA, z hlediska jejich strukturních parametrů, vytvořili J. Černý a kol. (2020) 97 dinukleotidových konformačních tříd (angl. diNucleotides Conformational classes – zkráceně NtC třídy). Po sobě jdoucí dvojice nukleotidů v řetězci, tzv. stepy, jsou roztrženy do těchto tříd podle sedmi torzních úhlů mezi dvěma ribózovými kruhy, dvou glykosidických úhlů a tří parametrů popisujících vzájemnou orientaci dvou dusíkových bází (Černý a kol. 2020). NtC třídy jsou označovány čtyřmi znaky, většinou dvěma písmeny a dvěma číslicemi, ale objevují se zde i třídy, které mají tři písmena a jednu číslici. Označení je ve formátu AAxx, kde první písmeno označuje formu prvního nukleotidu (A, B nebo Z) ze stepu a druhé písmeno formu druhého nukleotidu ve stepu, x jsou potom číslice nebo písmena. Jsou zde i speciálně značené třídy jako ICxx, které značí paralelní uspořádání bází vůči sobě, OPxx, které je pro báze, které nejsou ani za sebou ani paralelně uspořádány. Poslední třídou je třída NANT, do které patří všechny nezařazené DNA a RNA struktury. Například třída AA00 značí čistě kanonickou A formu, obdobně BB00 je kanonická B forma, ale třída AA02 je označení pro nukleové kyseliny v A formě, které mají  $\chi$  úhel typický pro B formu. Všechny NtC třídy a tabulky, které popisují dané strukturní parametry, lze najít na adrese: <https://dnatco.datmos.org/>. Na stejné adrese lze také najít databázi, ve které jsou DNA a RNA struktury z PDB databáze rozděleny podle NtC tříd.

## 2.3. Proteiny

Proteiny jsou jednou ze základních biomakromolekul a jsou nedílnou součástí každého organismu. Bílkoviny mají v buňkách několik významných funkcí, podle kterých se i dělí do skupin. Těmito skupinami jsou strukturní (např. kolagen), transportní (hemoglobin), obranné (imunoglobuliny), regulační (inzulín), katalytické (alkoholdehydrogenasa), motorové bílkoviny (aktin), histony a zásobní bílkoviny, které se nacházejí převážně u rostlin (např. gluten) (Koolman a Röhm 2012).

### 2.3.1. Struktura proteinů

Primární struktura proteinů je dána pořadím proteogenních aminokyselin v řetězci tzv. peptidickou vazbou. Pro zápis aminokyselin v řetězci se používají třípísmenné nebo

jednopísmenné zkratky (např. Valin, Val, V) a pořadí se zapisuje od N-konce, který má volnou aminoskupinu, k C-konci, který má volnou karboxylovou skupinu poslední aminokyseliny. Primární struktura je zodpovědná za vlastnosti proteinů díky postranním řetězcům aminokyselin, které mohou být nepolární, polární, kyselé nebo zásadité.

Sekundární strukturou je geometrické uspořádání řetězce aminokyselin, které jsou stabilizovány vodíkovými můstky. Existují tři základní uspořádání:  $\alpha$ -helix,  $\beta$ -skládaný list a  $\beta$ -otočka. Nejčastější sekundární strukturou je  $\alpha$ -helix, při kterém je řetězec aminokyselin šroubovitě svinut a vodíkové můstky jsou mezi -NH a -CO skupinami aminokyselin.  $\beta$ -skládaný list je skoro natažená konformace řetězce a vodíkové můstky se nacházejí mezi sousedními řetězci. V závislosti na tom, zda aminokyselinové řetězce běží ve stejném nebo opačném směru, rozdělujeme  $\beta$ -skládaný list na paralelní a antiparalelní.  $\beta$ -otočka se nachází na místech, kde řetězec mění svůj směr.  $\beta$ -otočka je tvořena čtyřmi aminokyselinovými zbytky tak, že se řetězec obrátí do protisměru. Vodíkové můstky mezi první a čtvrtou aminokyselinou stabilizují  $\beta$ -otočku (Bruce a kol. 2008).

Terciální struktura je dána prostorovým uspořádáním sekundárních elementů a neuspořádaných úseků řetězce. Kvarterní struktura je pak celková struktura proteinu, který je tvořen více než jedním polypeptidovým řetězcem. Kvarterní struktura popisuje, jak je protein tvořen podjednotkami, což jsou samostatné polypeptidové řetězce proteinu. Podjednotky drží pospolu díky nekovalentním interakcím mezi jejich vazebnými místy. Protein tvořený dvěma podjednotkami se nazývá dimer, je-li tvořen třemi podjednotkami, nazývá se trimer a podobně (Bruce a kol. 2008).

## 2.4. Protein-DNA komplexy

### 2.4.1. Interakce protein-DNA

DNA jako molekula je velmi konzervativní a málo reaktivní. Aby mohlo docházet k replikaci, transkripci, nebo i shlukování DNA v podobně nukleosomů a následně další kondenzaci do chromozomů, je zapotřebí proteinů. Interakce protein-DNA mohou probíhat specificky, podle určité sekvence DNA, nebo nespecificky. Vazby proteinů na DNA mohou indukovat určité změny ve struktuře DNA, například ohýbat řetězec DNA nebo vytvářet smyčky.

Nespecifické interakce proteinů s DNA probíhají díky solným můstkům. K zápornému náboji fosfátové páteře DNA jsou přitahovány kladně nabitě části postranních řetězců aminokyselin, nejčastěji se jedná o aminokyseliny lysin a arginin. Takto nespecificky se k DNA váží například histony, kolem kterých se DNA dvakrát ovine a vytváří nukleozomy.

Specifické interakce využívají dané sekvence nukleotidů v řetězci DNA. Proteiny se svými doménami nejčastěji váží do velkého žlábků, ale ne vždy tomu tak musí být. Vazba proteinových domén probíhá přes vodíkové vazby, přímé i zprostředkované přes molekuly vody, a van der Waalsovy interakce (Harrison 1991). Velký žlábek slouží také pro rozpoznání sekvence na DNA, a to díky tomu, že daný pár má určité pořadí akceptorů vodíkových vazeb, donorů vodíkových vazeb, vodíků a metylové skupiny. Pomocí těchto čtyř znaků dokáží proteiny rozpoznat, o jaký nukleotidový pár se jedná. Například GC pár obsahuje dva akceptory vodíkové vazby, jeden donor vodíkové vazby a atom vodíku, přesně v tomto pořadí. Nejčastěji je pro vazbu do velkého žlábků využíván  $\alpha$ -helix. Mezi motivy DNA vazebných domén patří helix-otočka-helix, zinkový prst, leucinový zip, helix-smyčka-helix a motivy  $\beta$ -listy. Pomocí specifických interakcí se na DNA vážou proteiny jako transkripční faktory, aktivátory, represory apod.

Helix-otočka-helix (HTH – z angl. helix-turn-helix) je tvořen dvěma  $\alpha$ -helixy, které jsou spojeny  $\beta$ -otočkou a jsou na sebe kolmé. Jeden z helixů, tzv. rozpoznávací, slouží k navázání do velkého žlábků a váže do něj pomocí argininu, serinu nebo asparaginu. Často může být několik těchto domén spojeno pomocí dalších helixů nebo  $\beta$ -listů. Tento motiv je typický např. pro rodinu proteinů Cro represorů pro bakteriofág lambda nebo tryptofanový represor (Harrison 1991). Pod HTH motivy lze zařadit také tzv. winged-helix (okřídlená šroubovice), jedná se HTH motiv, který navíc obsahuje  $\beta$ -listy. Tyto  $\beta$ -listy, případně  $\beta$ -otočky mezi  $\alpha$ -helixy a  $\beta$ -listy, pak napomáhají k dalším kontaktům s DNA.

Zinkový prst je tvořen dvěma krátkými paralelními  $\beta$ -listy, které následuje  $\alpha$ -helix. Smyčka mezi  $\alpha$ -helixem a  $\beta$ -listy je stabilizována zinečnatým kationtem, který interaguje se čtyřmi cysteiny nebo dvěma cysteiny a dvěma histidiny.  $\alpha$ -helix se pak váže do velkého žlábků DNA na guaniny, nejčastěji přes argininy nebo histidiny. Kvůli slabší vazbě na DNA se většinou vyskytuje několik domén za sebou. Pomocí zinkového prstu se např. váže transkripční faktor TFIIIA (Harrison 1991).

Leucinový zip je dimer tvořen dvěma  $\alpha$ -helixy. Tyto  $\alpha$ -helixy mají na polovině směřující k C-konci bazický region tvořen leuciny, které se opakují každou sedmou aminokyselinu, což má za následek, že leuciny jsou pouze na jedné straně helixu. Leuciny sousedících  $\alpha$ -helixů slabě interagují, a tak vytvářejí dimer. N-konce helixů jsou pak pomocí argininu nebo lysinu navázány na velký žlábek DNA. Tento motiv využívá např. transkripční faktor AP-1 (Harrison 1991).

Helix-smyčka-helix (HLH – z angl. helix-loop-helix) je motiv velmi podobný leucinovému zipu. Opět se jedná o dimer, ve kterém je jedna jednotka tvořena dvěma  $\alpha$ -helixy, menším a větším, spojenými smyčkou. Menší  $\alpha$ -helixy, které jsou na N-konci podjednotky, jsou bazické a interagují mezi sebou za vzniku dímeru.  $\alpha$ -helixy na C-konci se pak váží do velkého žlábků dvoušroubovice pomocí argininu a glutamátu.

Vazba proteinu na DNA skrze  $\beta$ -listy není tak obvyklá jako použití  $\alpha$ -helixu, ale existuje několik proteinů s tímto motivem. Jedná se o dimer a každá jeho monomerní podjednotka přispívá jedním  $\beta$ -listem do velkého žlábků DNA a  $\beta$ -listy jsou vůči sobě paralelní. Monomerní jednotka se skládá z  $\beta$ -listu a dvou  $\alpha$ -helixů. Vazbu na DNA pomocí  $\beta$ -listu využívá například MetJ represor. Proteiny, které se vážou do TATA boxu, tzv. TBP, využívají strukturu  $\beta$ -listů pro navázání do malého žlábků, čímž dochází k pokrivení DNA v daném místě (Garvie a Wolberger 2001).

## 2.4.2. Použité protein-DNA komplexy

Protein-DNA komplexy, které jsou použity v této diplomové práci, nesou označení z Protein Data Bank (PDB) databáze a konkrétně se jedná o 3PVI, 3AAF, 3KXT, 2VJV a 1DP7.

### 2.4.2.1. 3PVI

3PVI je komplex mutantního proteinu endonukleasy *PvuII* pocházejícího z bakterie *Proteus vulgaris* se sekvencí DNA. Mutace nastala v pozici 34, kde byl původní arginin nahrazen guaninem. Původní *PvuII* (v PDB databázi označena jako 1PVI) je restriční endonukleasa, která štěpí sekvenci DNA 5'-CAGCTG-3'. *PvuII* je homodimer a každá podjednotka se skládá ze 3 strukturních částí – katalytické, dimerizační a rozpoznávací. Přestože dimerizační a rozpoznávací funkce daných částí je daná, tak His85, který



se nachází v rozpoznávací části proteinu a váže se přes velký žlábek přímou vodíkovou vazbou na atom O6 guaninu a nepřímou (vazba přes molekulu vody) vodíkovou vazbou na atom N4 cytosinu, napomáhá k interakci podjednotek proteinu. Naopak Asp34, který slouží k dimerizaci, napomáhá k rozpoznávání sekvence DNA. Tento Asp34 se váže přes malý žlábek na guanin a to přímou vodíkovou vazbou, s N2 dusíkem guaninu, a nepřímou vodíkovou vazbou, s N3 dusíkem guaninu. K rozpoznání G·C v dané sekvenci DNA napomáhá ještě Gln33 nepřímou vodíkovou vazbou na atom O cytosinu v malém žlábků. Asp34 napomáhá kromě rozpoznání G·C páru také v katalytické aktivitě proteinu. V mutantním proteinu 3PVI je vlivem mutace glycinu za Asp34 snížena afinita proteinu k řetězci DNA a také je snížena jeho katalytická schopnost, což vede k tomu, že nedochází k úplnému rozstřížení DNA (Horton a kol. 1998).

#### 2.4.2.2. 3AAF

Protein-DNA komplex 3AAF obsahuje protein, který patří do rodiny WRN (Werner Syndrome Protein) RecQ helikas DNA. Proteiny z RecQ rodiny jsou aktivní především k různým typům rozpletené DNA, jako jsou bubliny, D smyčky (ang. – D loop) a Hollidayovo spojení, kde poté rozdělují dvoušroubovice do jednotlivých vláken. WRN proteiny nedokáží rozvinout dvoušroubovici a jsou závislé na RQC (RecQ C-koncové) doméně. RQC doména je složena z pěti  $\alpha$ -helixů a čtyřech  $\beta$ -listů. Smyčka mezi helixy  $\alpha 2$  a  $\alpha 3$  ( $\alpha$ -2- $\alpha$ 3 smyčka) se váže do velkého žlábků pomocí sedmi vodíkových vazeb a solným můstkem na fosfátovou páteř. Oba konce DNA byly rozděleny a navázány na  $\beta$ -vlásečku mezi  $\beta 2$  a  $\beta 3$  listy. Tato vlásečka je označována jako tzv.  $\beta$ -křídlo motivu winged-helix (okřídlené šroubovice).  $\beta$ -křídlo se vklíní mezi báze A<sub>1</sub>-C<sub>2</sub> a mezi G<sub>13</sub>-T<sub>14</sub>, což má za následek přerušování patrových (stacking) interakcí mezi těmito bázemi. Vklínění je způsobeno hydrofobními aminokyselinami z  $\beta$ -křídla, přesněji Tyr1034, Phe1037 a Met138. Postranní řetězce aminokyselin z  $\alpha$ 2- $\alpha$ 3 smyčky Arg987 a Arg993 tvoří solné můstky s fosfáty nukleotidů A<sub>7</sub> a C<sub>2</sub>. WRN RQC preferují vazbu na 5'-konec řetězce DNA (Kaitano a kol. 2010).

#### 2.4.2.3. 3KXT

Komplex 3KXT je chromatinový protein Cren7 z Crenarchaeot (jeden z kmenů archeí) s dvoušroubovicí DNA se sekvencí 5'-GCGATCGC-3'. Prokaryota nemají

histony, nedochází u nich tedy k tvorbě nukleozomů a následně chromozomů. Místo histonů mají proteiny vázající DNA, tzv. chromatinové proteiny, které sbalují DNA a tvoří nukleoid. V případě Crenarchaeot je jedním z těchto proteinů Cren7. Cren7 je malý protein tvořen z  $\beta$ -listů a  $\beta$ -otoček, který nespecificky váže DNA, stabilizuje ji a vytváří negativní superhelix.  $\beta$ -listy  $\beta_3$ ,  $\beta_4$  a  $\beta_5$  a smyčka  $\beta_3$ - $\beta_4$  se váží do malého žlábků DNA. Vazba probíhá pomocí kladně nabitých aminokyselin Lys24, Lys31, Arg33, Lys34, Lys42, Arg51, Lys53 a hydrofobních aminokyselin Trp26, Leu28, Pro30, Val36, Ile38, Leu40 a His52. Kvůli vazbě do malého žlábků dochází k jeho rozšíření na  $\sim 20 \text{ \AA}$  (Freng a kol. 2010).

#### 2.4.2.4. 2VJV

V případě 2VJV se jedná o komplex dimeru proteinu IS608 transposasy TnpA s 26 nukleotidovou sekvencí DNA levého konce a 6 nukleotidovou sekvencí DNA. Transposasy jsou enzymy, které katalyzují štěpení, přenášení a vkládání úseků DNA tzv. transpozony na jinou část DNA. U bakterií může být tímto způsobem přenesena část plazmidové DNA, např. s informací rezistence na antibiotika, na genomovou DNA. Transposasy se naváží na donorovou DNA, ze které vyštípou transpozon. Tento transpozon má dva konce, levý a pravý, které se od sebe liší posledními čtyřmi nukleotidy. Před každým koncem je část, která se nazývá nedokonalý palindrom (imperfekt palindrom), ve které jednovláknová DNA tvoří vlásenku (hairpin). Při přenášení transpozonu dochází ke spojení levého a pravého konce a vzniká tzv. transpozonové spojení (transposon junction) a následně je pak transpozon vložen do cílové DNA. Štěpení DNA probíhá na atomu fosforu na cukrfosfátové páteři nukleofilní substitucí. Jako nukleofilní činidla mohou být použity atomy vody, které jsou aktivovány kovovými ionty navázanými na enzym, hydroxylovou skupinou na 5' nebo 3' konci DNA nebo hydroxylovou skupinou aminokyseliny v aktivním místě enzymu (tyrosine nebo serine) (Barabas a kol. 2008). Transposasy obsahují oblast HUH (His-hydrofobní aminokyselina-His), která nasměruje  $\text{Mg}^{2+}$ , a ten napomáhá katalyzovat štěpení. TnpA katalyzuje tyto děje pouze u jednořetězcové DNA. TnpA je dimér a každá podjednotka se skládá ze dvou subdomén. První subdoména obsahuje čtyři antiparalelní  $\beta$ -listy a tři  $\alpha$ -helixy. Druhá subdoména tvořena 17 rezidui na N-konci a 34 rezidui na C-konci obsahuje dva antiparalelní  $\beta$ -listy a dva  $\alpha$ -helixy. Obě subdomény jsou pak spojeny dlouhým  $\beta$ -listem. TnpA obsahuje HUH oblast na  $\beta$ -listu  $\beta_4$  a je tvořen His64, Ile65 a His66. Další

potřebnou aminokyselinou pro katalýzu reakcí je Tyr127, který funguje jako nukleofilní činidlo. TnpA vytváří několik interakcí s cukrfosfátovou kostrou DNA, ale pouze dva z nich jsou bázově specifické. Tyto bázově specifické kontakty jsou zprostředkovány pomocí His11, který se váže na guainin, a Lys 82, který interaguje s cytosinem (Ronning a kol. 2005).

#### 2.4.2.5. 1DP7

1DP7 je komplex lidského regulačního faktoru X1 (hRFX1 – z angl. human regulatory factor X1), který má vysoce konzervovanou DNA vazebnou doménu, a jeho rozpoznávací DNA sekvencí je tzv. X-box, který je 5'-GTNRCC(0-3N)RGYAAC-3', kde N je jakýkoli nukleotid, R je purin a Y je pyrimidin. Tento regulační faktor využívá winged-helix pro interakci s DNA a na jednu sekvenci DNA jsou navázány dva hRFX1. Vazebná doména hRFX1 se skládá ze tří  $\alpha$ -helixu ( $\alpha$ ), tří  $\beta$ -listů ( $\beta$ ) a tří  $\beta$ -smyček (S) v tomto pořadí  $\alpha 1$ - $\beta 1$ - $\alpha 2$ -S1- $\alpha 3$ -S2- $\beta 2$ -S3- $\beta 3$ . Smyčka mezi  $\beta 2$ - $\beta 3$  tvoří tzv. křídlo.  $\beta$ -listy  $\beta 2$  a  $\beta 3$  a smyčka S3 (křídlo), které je spojuje, interagují s velkým žlábkem jedné poloviny symetrického X-boxu. Jedna strana  $\alpha$ -helixu  $\alpha 3$  interaguje s malým žlábkem s druhou polovinou. Vazba hRFX1 k X-boxu je zprostředkována především polárními aminokyselinami přes přímé nebo nepřímé vodíkové vazby. Arg58 (z  $\beta$ -smyčky S3) se váže dvěma vodíkovými vazbami na guanin G11. Tyr67 ( $\beta 3$ ) interaguje s G10 a Arg62 (S3) interaguje vodíkovými vazbami s guaninem druhého řetězce DNA G15'. V případě, že jsou tyto tři báze nahrazeny thyminem, nedochází k tvorbě vodíkových vazeb mezi danými aminokyselinami. Lys45 ( $\alpha 3$ ) se váže na C6 přes malý žlábek, což vede k rozšíření malého žlábků o min. 3 Å. Jelikož se na sekvenci DNA vážou dva stejné proteiny, dochází ke kooperativnímu vázání, kdy se na jednu polovinu vazebného místa váže Ser42 (z jednoho proteinu) pomocí nepřímé vodíkové vazby přes velký žlábek na G6', zatímco Lys45 (z druhého proteinu) se váže přímou vodíkovou vazbou přes malý žlábek na C6. K podobné vazebné kooperaci dochází i na G11 a C11' (Gajiwala a kol. 2000).

### 3. Cíl práce

Diplomová práce si klade za cíl studovat stabilitu A formy DNA ve vybraných protein-DNA komplexech v průběhu MD simulací. Dalším cílem je porovnat mezi sebou simulace provedené různými silovými poli a z výsledků posoudit, jak dobře jednotlivá pole popisují stabilitu A formy a případně přechodných A/B forem DNA. Tyto analýzy pak mohou napomoci k úpravám empirických potenciálů a návrhu nových, přesnějších silových polí.

### 4. Materiál a metody

Protein-DNA komplexy byly vybírány z databáze DNATCO, kde byly hledány komplexy, které obsahovaly bazové stepy v AA00, AA02, AB01 nebo BA05 formě. Dalším kritériem pro výběr bylo, aby komplex měl krystalovou strukturu s dobrým rozlišením. Pro MD simulaci jsou nejlepší krystalové struktury s rozlišením do 2 Å. Byly vybrány komplexy s označením 3PVI, 3AAF, 3KXT, 2VJV a 1DP7. Krystalové struktury DNA s proteiny byly staženy z PDB databáze.

Originální PDB soubory byly poté upraveny. Odstranily se atomy krystalizačních činidel a některých iontů, které nebyly v kontaktu s DNA nebo proteinem, a jejich odstranění tedy nemělo vliv na vývoj struktury komplexu během simulace. Byly odstraněny duplicitní atomy v případě dvojitých obsazovacích faktorů. V případě, že byla v databázi uložena jen symetricky nezávislá část molekuly, bylo potřeba strukturu rozšířit o chybějící atomy dle prvků krystalografické symetrie. Podle pH roztoku, ve kterém docházelo ke krystalizaci, byla zvolena protonizační forma aminokyseliny histidinu. Program AMBER rozlišuje mezi třemi formami, HID (atom H na  $\delta$  atomu N imidazolové skupiny), HIE (atom H na  $\epsilon$  atomu N) a HIP (dva atomy H na obou atomech N). Pro určení formy histidinu byl využit on-line program PropKa (<https://server.poissonboltzmann.org/pdb2pqr>). Pokud histidin tvořil zřetelné vodíkové vazby, byl protonizační stav histidinu upraven tak, aby byl konzistentní se sítí předpokládaných vodíkových vazeb ve struktuře.

V případě komplexu 3AAF byly ze souboru odstraněny atomy acetátového iontu a pro komplex 1DP7 byly odstraněny atomy krystalizačních činidel ethylenglykolu

a diethylenglykolu. Pro komplex 1DP7 bylo potřeba udělat více úprav. Nejprve bylo zapotřebí přidat druhý řetězec a k němu i protein, protože originální soubor obsahoval pouze jeden řetězec dvoušroubovice a s ní příslušně interagující protein (symetricky nezávislá část molekuly). Další úpravou bylo nahrazení báze 5-bromouracilu za thymin, která byla provedena proto, že parametry pro thymin jsou lépe ověřené než parametry pro 5-bromouracil. 5-bromouracil byl v původní práci použit kvůli krystalizaci a řešení krystalové struktury a liší se od thyminu tím, že methylová skupina na paté pozici thyminu byla nahrazena brómem, výměna by tak neměla mít vliv na studované chování komplexu.

Všechny simulace byly provedeny v programu AMBER 18 za pomoci příslušných skriptů. Po úpravě pdb souboru se jako první pracuje se skriptem LEAP, který připraví vstupní soubory pro další postup z pdb souboru daného komplexu. V tomto skriptu se vždy musí upravit počet molekul NaCl pro správnou koncentraci roztoku, ve kterém je komplex simulován. V prvním kroku, kdy se skript spouští, se zjistí, kolik je potřeba molekul na neutralizaci náboje komplexu, a v druhém kroku se přidá určitý počet molekul, aby bylo dosaženo 150 mM roztoku NaCl. Při použití tohoto skriptu se také kolem celé molekuly vytvoří box, který je vyplněný rozpouštědlem, ve kterém pak probíhá simulace komplexu. Box je oktaedrální a jeho rozměry jsou nastaveny tak, aby jeho stěny byly vzdáleny 10 Å od krajních atomů molekuly komplexu. Následuje ekvilibrace atomů komplexu a molekul roztoku, během které nejdříve dochází k minimalizaci energie systému.

Poté už proběhla samotná simulace komplexů za použití dvou silových polí, OL15 a bsc1. Simulace, stejně jako ekvilibrace, byla provedena za standardní teploty a tlaku (298 K, 1 atmosféra), délka simulace byla 1 ms a zápis souřadnic atomů byl po 10 ps. Pro ekvilibraci i simulaci byl využit skript PMEMD.CUDA. Pro analýzu trajektorií simulace byl využit modul cpptraj programu AMBER. Kvůli velkému množství dat v jedné trajektorii byl pro analýzu vybrán každý stý snímek z trajektorie (prořezaná trajektorie), a to z důvodu rychlejší práce a interpretace výsledků. Ve výsledné trajektorii byly tedy snímky s 1 ns krokem. Pro určení velikosti žlábků DNA byl použit skript z programu AMBER, který pro výpočet velikosti žlábků použil stejný postup jako byl popsán v práci Hassan a Calladine (1998).

V této práci jsou vyobrazeny grafy, které znázorňují vývoj parametru P a úhlu  $\chi$ , případně velikostí žlábků, v čase v průběhu simulace pro ty nukleotidy daného komplexu, které se ve výchozí struktuře nacházely v A formě. Pro grafické znázornění tohoto vývoje byly použity prořezané trajektorie, jak je popsáno výše. Naopak tabulky znázorňují všechny nukleotidy kromě prvního a posledního nukleotidu daného komplexu a jsou pro ně použita data z nezkrácené původní trajektorie. V tabulkách jsou uvedeny dva druhy rozdělení forem. Prvním z nich je rozdělení pouze na základě pseudorotačního úhlu P, viz tabulka 3.

**Tabulka 3: Rozdělení v tabulkách na základě pseudorotačního úhlu P.**

| Forma    | P [°]     |
|----------|-----------|
| C3'-endo | 0 - 36    |
| C4'-exo  | 36 - 72   |
| O4'-endo | 72 - 108  |
| C1'-exo  | 108 - 144 |
| C2'-endo | 144 - 180 |
| C3'-exo  | 180 - 216 |
| C2'-exo  | 324 - 360 |
| Ostatní  | 216 - 324 |

Druhým používaným rozdělením je rozdělení podle torzního úhlu  $\delta$  a zároveň podle úhlu  $\chi$ . Je zde použito z toho důvodu, že některé nukleotidy se mohou nacházet v tzv. přechodné formě, kdy nukleotid může mít hodnotu úhlu  $\delta$  (respektive parametru P) typickou pro A formu, ale hodnoty úhlu  $\chi$  pro B formu a naopak. Proto je zde použito i označení pro tyto formy jako  $\delta/\chi = A/A, A/B, B/A$  a  $B/B$ , kde první písmeno značí, zda je hodnota úhlu  $\delta$  blíže A nebo B formě, druhé písmeno udává stejnou informaci, avšak o úhlu  $\chi$ . V tabulce 4 je toto rozdělení přiřazeno k příslušným NtC třídám a referenčním hodnotám úhlů  $\delta$  a  $\chi$  pro danou konformaci.

**Tabulka 4: Přiřazení použitého rozdělení parametrů  $\delta/\chi$  k příslušným NtC třídám.**

Referenční hodnoty pro úhel  $\delta$  a  $\chi$  jsou převzaty z databáze dnatco.org

(<https://dnatco.datmos.org/>).

| P/ $\chi$ | NtC  | $\delta$ [°] | $\chi$ [°] |
|-----------|------|--------------|------------|
| A/A       | AA00 | 82           | 199        |
| A/B       | AA02 | 88           | 245        |
| B/A       | BB16 | 138          | 253        |
| B/B       | BB00 | 138          | 204        |

## 5. Výsledky

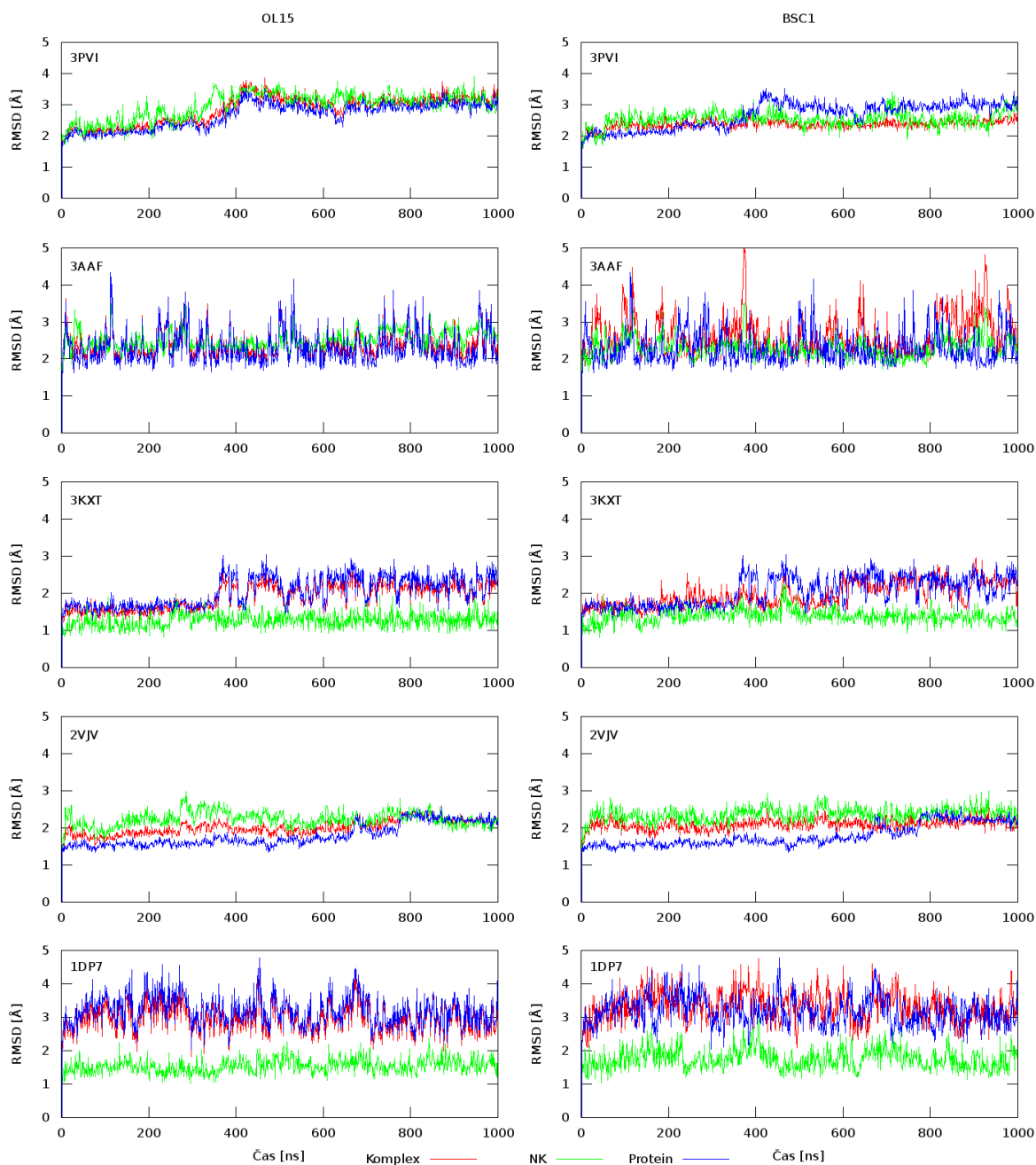
### 5.1. Stabilita simulací

Stabilitu simulací lze posoudit ze střední kvadratické odchylky (RMSD – z angl. root mean square deviation) poloh jednotlivých atomů komplexů v čase vůči původní (krystalografické) pozici. Zde jsme určili RMSD pro celý komplex a zvlášť pro nukleovou kyselinu a protein, abychom mohli posoudit stabilitu, případně nestabilitu částí komplexu. V grafu č. 1 jsou zaznamenány RMSD pro námi simulované komplexy 3PVI, 3AAF, 3KXT, 2VJV a 1DP7. Z grafů lze vyčíst, že nedochází k výrazným a dlouhodobým výchylkám, a komplexy tedy zůstaly v průběhu simulace stabilní.

V případě komplexu 3PVI lze vidět postupný nárůst RMSD v průběhu simulace pro obě silová pole, což může být důsledkem toho, že se nukleová kyselina komplexu, která z větší části začínala v A formě, postupně měnila na B formu. Tento průběh koresponduje s dalšími výsledky komplexu 3PVI (viz dále). V komplexu 3AAF lze vidět fluktuace RMSD kolem hodnoty okolo 2 Å a objevují se zde určité výchylky, které jsou ovšem velmi krátkodobé, proto se jedná o stabilní simulaci. V bsc1 poli jsou odchylky lehce vyšší, ale nenaznačují nestabilitu systému. Takové výchylky v obou polích mohou být způsobeny například pohybem postranních řetězců aminokyselin na vnějším povrchu komplexu. Komplex 3KXT má velmi stabilní nukleovou kyselinu, kdy se její hodnota RMSD pohybuje na hranici 1 Å. Podobnou RMSD má zpočátku i protein a celý komplex, ovšem od cca 360 ns se jejich hodnota RMSD zvyšuje. To může být dáno změnou pozice Arg33, Lys31 a Lys34. Jelikož protein je větší než nukleová kyselina a obsahuje více atomů, dochází k navýšení RMSD proteinu i celého komplexu. V případě pole bsc1 je trend podobný. Komplex 2VJV má velmi malou fluktuaci RMSD kolem hodnoty 2 Å. V čase okolo 800 ns došlo k malému navýšení RMSD nukleové kyseliny. 2VJV obsahuje celkem čtyři jednořetězcové DNA, dvě dlouhé, které interagují samy se sebou a tvoří vlásenky, a dvě krátké, které pomocí několika nukleotidů interagují s konci dlouhých vláken. Navýšení RMSD může být způsobeno pohybem báze nukleotidu, který se nachází ve vlásence a směřuje ven, a také přerušením interakce krátkých vláken s konci dlouhých vláken a tím pádem pohybem konců krátkých řetězců DNA. V poli bsc1 je situace podobná. V posledním komplexu 1DP7 je RMSD nukleové kyseliny nižší než pro protein a celý komplex. To může být zapříčiněno tím, že v tomto komplexu jsou dva stejné



proteiny navázány z obou stran na DNA a jakýkoli menší pohyb proteinů může vyvolat zvýšení RMSD vůči původní pozici. Stejná situace se odehrává i v případě pole bsc1.



**Graf č. 1:** RMSD daných komplexů v závislosti na době trvání simulace. V levém sloupci jsou znázorněny komplexy simulované v silovém poli OL15, v pravém sloupci pak komplexy simulované v poli bsc1. V grafu jsou zaznamenány hodnoty RMSD celého komplexu (červeně), samostatné nukleové kyseliny (zeleně) a samostatného proteinu (modře).

## 5.2. 3PVI – endonukleasa *PvuII*

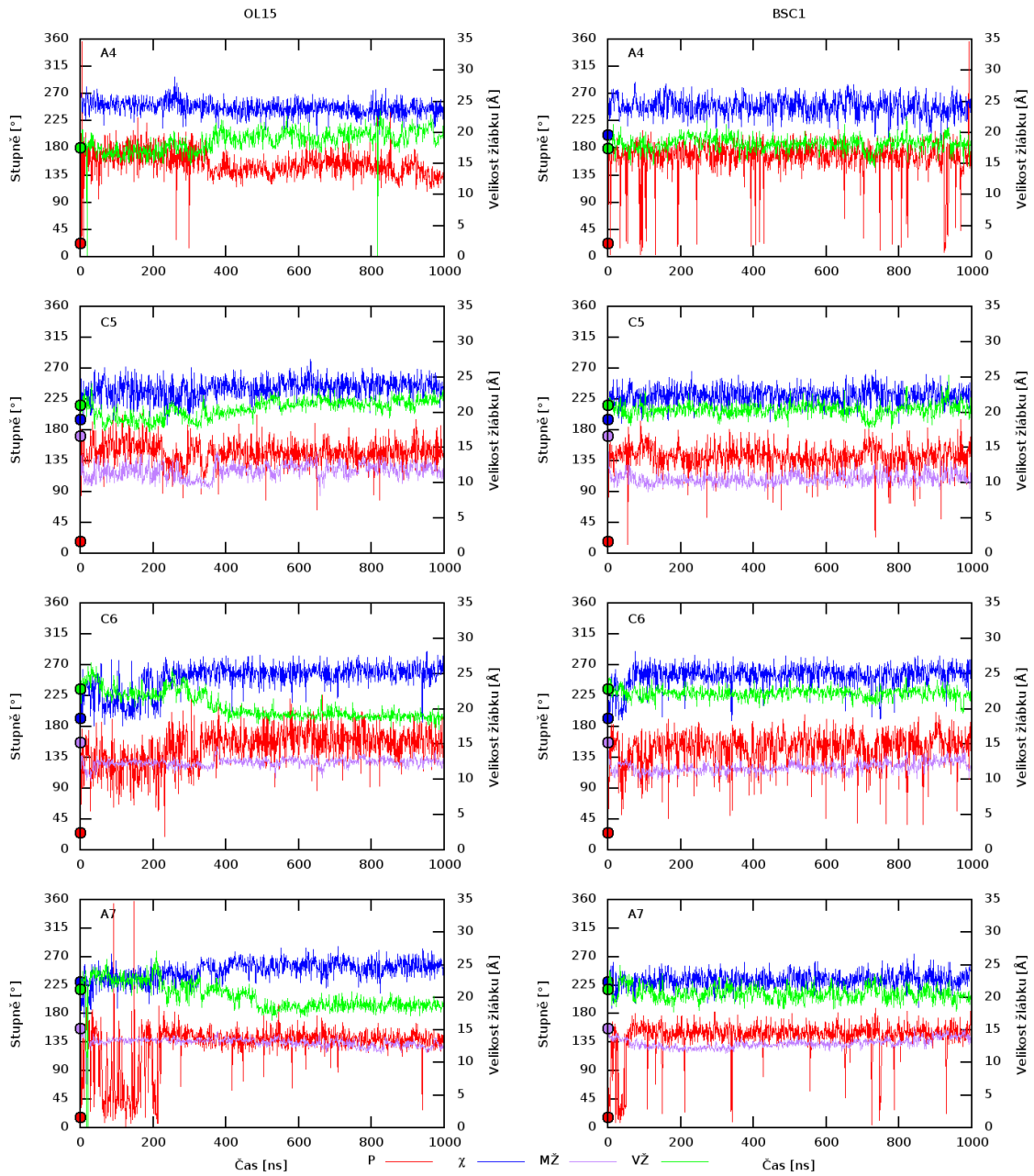
Komplex 3PVI obsahuje dvoušroubovici DNA, která má 13 nukleotidů v jednom řetězci. V případě rozdělení pouze podle parametru P se v C3'-endo konformaci, která je typická pro A-formu DNA, v krystalografické struktuře nacházelo celkem 8 nukleotidů v řetězci C a 9 nukleotidů v řetězci D (viz tabulka 5 a 6, zeleně zvýrazněno). Z těchto 8, respektive 9 nukleotidů se v průběhu simulace v konformaci C3'-endo neudržel ani jeden a většina nakonec přešla do C2'-endo konformace. V tabulce 5 lze vidět, že se v průběhu simulace daný nukleotid nacházel v C3'-endo konformaci jen minimálně, řádově v jednotkách procent. Velké procentuální zastoupení je i pro konformaci C1'-exo. Porovnáním hodnot v tabulce 5 a 6 pro stejné nukleotidy zjistíme, že pole OL15 má vyšší zastoupení konformace C1'-exo na úkor C2'-endo, oproti poli bsc1, kde je to opačné. V grafu č. 1 a č. 2 jsou pak zaznamenány hodnoty parametru P (červená), úhlu  $\chi$  (modrá) a šířky velkého (zelená) a malého žlábků (fialová) v závislosti na čase simulace řetězce C komplexu 3PVI. Z důvodu přehlednosti byly nukleotidy, které začínaly v A/A formě, rozděleny do dvou grafů. V grafech si lze povšimnout, že i když všechny nukleotidy začínaly s hodnotou parametru P okolo 20 – 30° (C3'-endo), což je typické pro A formu, většina nukleotidů přešla během několika desítek nanosekund do C2'-endo konformace, typické pro B formu. Hodnota parametru P se tedy pohybuje okolo 180°. Některé nukleotidy, např. A7, se ještě v průběhu simulace vracely do C3'-endo, ale jen na krátkou dobu. Obdobný trend lze pozorovat pro hodnoty úhlu  $\chi$ , který také začínal v hodnotách typických pro A formu, tedy okolo 200°, ale ihned po začátku simulace došlo k nárůstu jeho hodnoty. Zde nastaly dvě situace. Při první se hodnota úhlu  $\chi$  dostala na hranici 260°, což je typické pro B formu DNA. To se stalo v případě nukleotidů A4, C6, C9 a částečně G11 a G12. V druhé situaci úhel  $\chi$  zůstal někde mezi těmito hodnotami a pohyboval se okolo 220° v mezi-formě mezi A a B DNA. Tendenci nukleotidů přecházet z A formy do B formy naznačuje i šířka malého žlábků, která je ve výchozí struktuře větší (což je typické pro A formu), ihned po začátku simulace se zmenší a poté fluktuuje kolem nižších hodnot typických pro B formu. To lze dobře vidět v grafu nukleotidu C5 a C6.

**Tabulka 5: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem OL15 pro komplex 3PVI.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

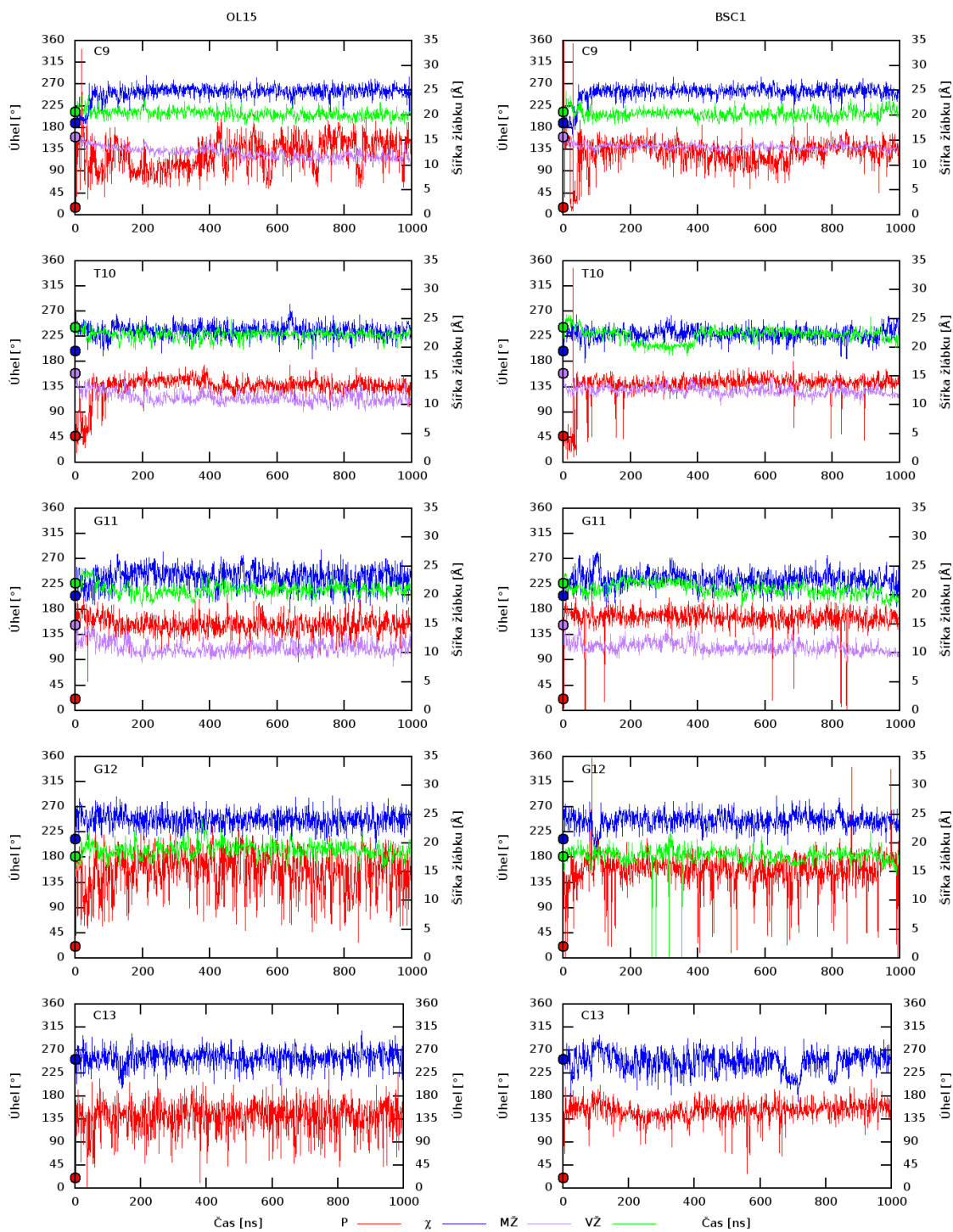
| OL15      |      |                       |                      |                         |              |             |              |             |              |             |         |
|-----------|------|-----------------------|----------------------|-------------------------|--------------|-------------|--------------|-------------|--------------|-------------|---------|
| N         | NtC  | P <sub>0</sub><br>[°] | P <sub>0</sub> forma | Procentuální zastoupení |              |             |              |             |              |             |         |
|           |      |                       |                      | C2'-<br>exo             | C3'-<br>endo | C4'-<br>exo | O4'-<br>endo | C1'-<br>exo | C2'-<br>endo | C3'-<br>exo | Ostatní |
| řetězec C |      |                       |                      |                         |              |             |              |             |              |             |         |
| G3        | BA08 | 149                   | C2'-endo             | 0                       | 0            | 1           | 2            | 25          | 53           | 16          | 2       |
| A4        | AA00 | 23                    | C3'-endo             | 0                       | 1            | 0           | 0            | 33          | 53           | 12          | 0       |
| C5        | AA00 | 18                    | C3'-endo             | 0                       | 0            | 0           | 1            | 40          | 54           | 4           | 0       |
| C6        | AA00 | 26                    | C3'-endo             | 0                       | 0            | 1           | 7            | 27          | 53           | 11          | 0       |
| A7        | AB05 | 16                    | C3'-endo             | 0                       | 5            | 7           | 3            | 49          | 35           | 1           | 0       |
| G8        | BA01 | 162                   | C2'-endo             | 0                       | 1            | 1           | 1            | 5           | 46           | 44          | 3       |
| C9        | AA01 | 16                    | C3'-endo             | 0                       | 0            | 7           | 26           | 32          | 30           | 4           | 0       |
| T10       | AA00 | 48                    | C4'-exo              | 0                       | 1            | 2           | 2            | 63          | 32           | 0           | 0       |
| G11       | AA00 | 21                    | C3'-endo             | 0                       | 0            | 0           | 1            | 25          | 68           | 5           | 0       |
| G12       | AA00 | 21                    | C3'-endo             | 0                       | 0            | 4           | 10           | 18          | 39           | 26          | 2       |
| T13       | AB02 | 21                    | C3'-endo             | 0                       | 0            | 4           | 11           | 33          | 46           | 6           | 0       |
| řetězec D |      |                       |                      |                         |              |             |              |             |              |             |         |
| G3        | BA08 | 162                   | C2'-endo             | 0                       | 0            | 0           | 1            | 29          | 60           | 9           | 1       |
| A4        | AA00 | 160                   | C2'-endo             | 0                       | 0            | 1           | 3            | 16          | 60           | 19          | 1       |
| C5        | AA00 | 27                    | C3'-endo             | 0                       | 0            | 1           | 3            | 63          | 31           | 1           | 0       |
| C6        | AA00 | 21                    | C3'-endo             | 0                       | 0            | 2           | 7            | 27          | 55           | 9           | 0       |
| A7        | AB05 | 11                    | C3'-endo             | 0                       | 0            | 1           | 1            | 33          | 62           | 2           | 0       |
| G8        | BA01 | 162                   | C2'-endo             | 0                       | 1            | 0           | 1            | 13          | 73           | 11          | 0       |
| C9        | AA01 | 20                    | C3'-endo             | 0                       | 2            | 12          | 21           | 18          | 30           | 16          | 1       |
| T10       | AA00 | 32                    | C3'-endo             | 0                       | 0            | 1           | 1            | 31          | 63           | 3           | 0       |
| G11       | AA00 | 24                    | C3'-endo             | 0                       | 0            | 0           | 0            | 15          | 75           | 9           | 0       |
| G12       | AA00 | 14                    | C3'-endo             | 0                       | 1            | 0           | 0            | 4           | 43           | 49          | 2       |
| T13       | AB02 | 17                    | C3'-endo             | 0                       | 0            | 1           | 17           | 55          | 24           | 2           | 0       |

**Tabulka 6: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem bsc1 pro komplex 3PVI.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

| <b>bsc1</b> |      |                       |                      |                         |              |             |              |             |              |             |         |
|-------------|------|-----------------------|----------------------|-------------------------|--------------|-------------|--------------|-------------|--------------|-------------|---------|
| N           | NtC  | P <sub>0</sub><br>[°] | P <sub>0</sub> forma | Procentuální zastoupení |              |             |              |             |              |             |         |
|             |      |                       |                      | C2'-<br>exo             | C3'-<br>endo | C4'-<br>exo | O4'-<br>endo | C1'-<br>exo | C2'-<br>endo | C3'-<br>exo | Ostatní |
| řetězec C   |      |                       |                      |                         |              |             |              |             |              |             |         |
| G3          | BA08 | 149                   | C2'-endo             | 0                       | 0            | 0           | 0            | 16          | 80           | 3           | 0       |
| A4          | AA00 | 22                    | C3'-endo             | 0                       | 3            | 1           | 0            | 5           | 66           | 25          | 0       |
| C5          | AA00 | 18                    | C3'-endo             | 0                       | 0            | 0           | 3            | 52          | 43           | 1           | 0       |
| C6          | AA00 | 26                    | C3'-endo             | 0                       | 0            | 1           | 5            | 27          | 62           | 4           | 0       |
| A7          | AB05 | 16                    | C3'-endo             | 0                       | 3            | 1           | 1            | 25          | 69           | 1           | 0       |
| G8          | BA01 | 162                   | C2'-endo             | 0                       | 1            | 0           | 0            | 1           | 69           | 29          | 0       |
| C9          | AA01 | 16                    | C3'-endo             | 0                       | 2            | 2           | 13           | 53          | 29           | 1           | 0       |
| T10         | AA00 | 47                    | C4'-exo              | 0                       | 2            | 3           | 0            | 52          | 43           | 0           | 0       |
| G11         | AA00 | 21                    | C3'-endo             | 0                       | 1            | 0           | 0            | 3           | 82           | 14          | 0       |
| G12         | AA00 | 21                    | C3'-endo             | 0                       | 1            | 2           | 2            | 17          | 61           | 16          | 0       |
| T13         | AB02 | 21                    | C3'-endo             | 0                       | 0            | 1           | 1            | 27          | 66           | 5           | 0       |
| řetězec D   |      |                       |                      |                         |              |             |              |             |              |             |         |
| G3          | BA08 | 162                   | C2'-endo             | 0                       | 0            | 0           | 0            | 13          | 77           | 9           | 0       |
| A4          | AA00 | 160                   | C2'-endo             | 0                       | 2            | 2           | 1            | 18          | 60           | 17          | 0       |
| C5          | AA00 | 27                    | C3'-endo             | 0                       | 0            | 1           | 3            | 56          | 40           | 1           | 0       |
| C6          | AA00 | 21                    | C3'-endo             | 0                       | 0            | 1           | 2            | 18          | 73           | 6           | 0       |
| A7          | AB05 | 11                    | C3'-endo             | 0                       | 2            | 2           | 1            | 31          | 62           | 1           | 0       |
| G8          | BA01 | 162                   | C2'-endo             | 0                       | 1            | 1           | 0            | 2           | 60           | 36          | 0       |
| C9          | AA01 | 20                    | C3'-endo             | 0                       | 1            | 1           | 9            | 58          | 31           | 0           | 0       |
| T10         | AA00 | 32                    | C3'-endo             | 0                       | 0            | 1           | 0            | 47          | 51           | 0           | 0       |
| G11         | AA00 | 24                    | C3'-endo             | 0                       | 0            | 0           | 0            | 7           | 82           | 11          | 0       |
| G12         | AA00 | 14                    | C3'-endo             | 0                       | 1            | 4           | 3            | 19          | 60           | 12          | 0       |
| T13         | AB02 | 17                    | C3'-endo             | 0                       | 0            | 0           | 1            | 24          | 69           | 5           | 0       |



**Graf č. 2:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů A4, C5, C6 a A7 řetězce C komplexu 3PVI. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole BSC1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .



**Graf č. 3:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů C9, T10, G11, G12 a C13 řetězce C komplexu 3PVI. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .

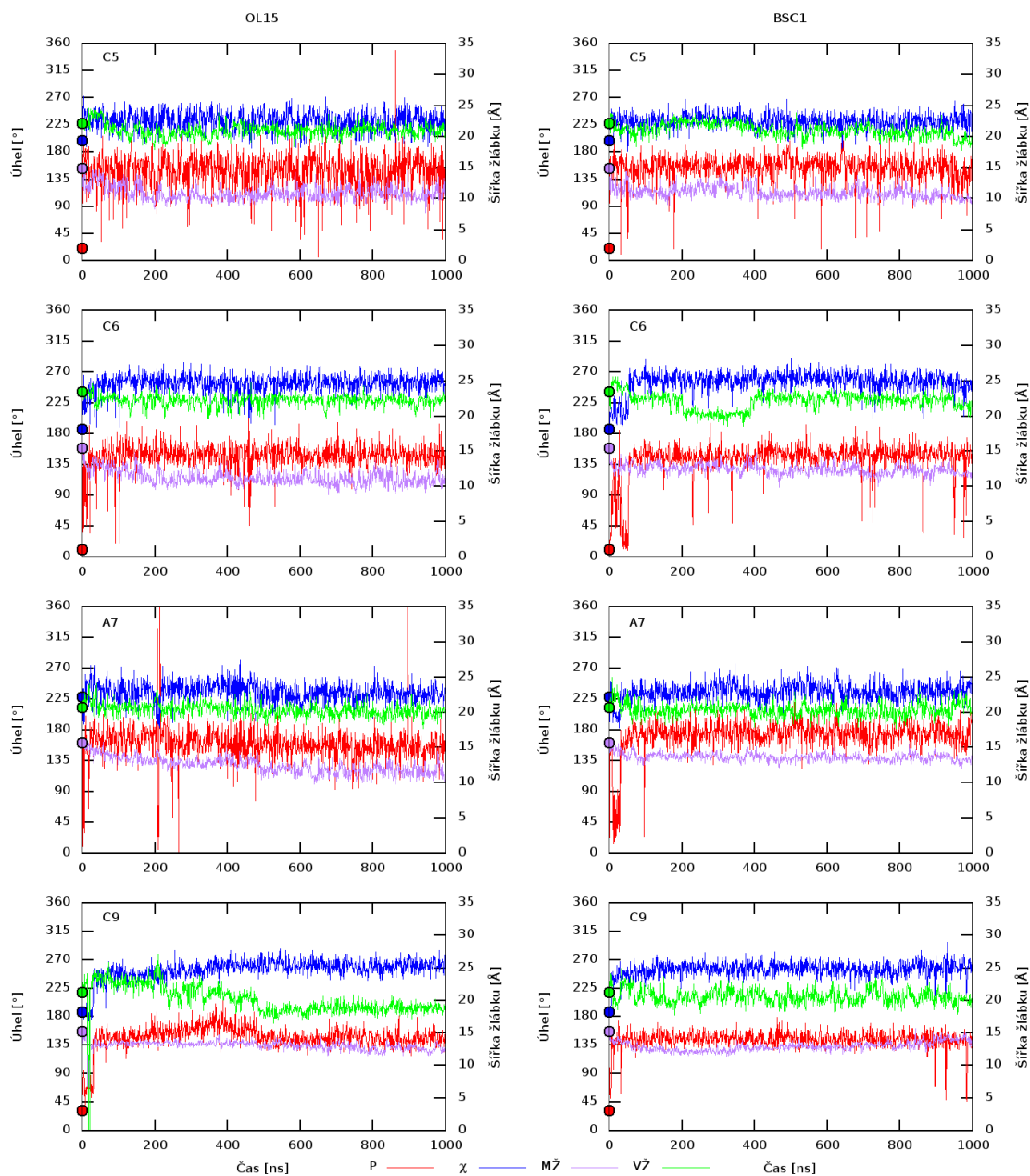
V tabulce 7 je znázorněno procentuální zastoupení konformací v průběhu celé simulace dle rozdělení podle úhlu  $\delta$  a  $\chi$ . I zde lze vidět, že nukleotidy mají tendenci přecházet z A formy do B formy, konformace A/A je tedy velmi málo zastoupena. Výjimku tvoří nukleotid C9, který v A/A konformaci stráví asi třetinu simulace v poli OL15.

V grafu č. 4 a 5 jsou znázorněny průběhy simulací nukleotidu řetězce D komplexu 3PVI. Zde je trend velmi podobný jako v případě řetězce C, většina nukleotidů se z původní A formy přesune během několika desítek nanosekund do B formy.

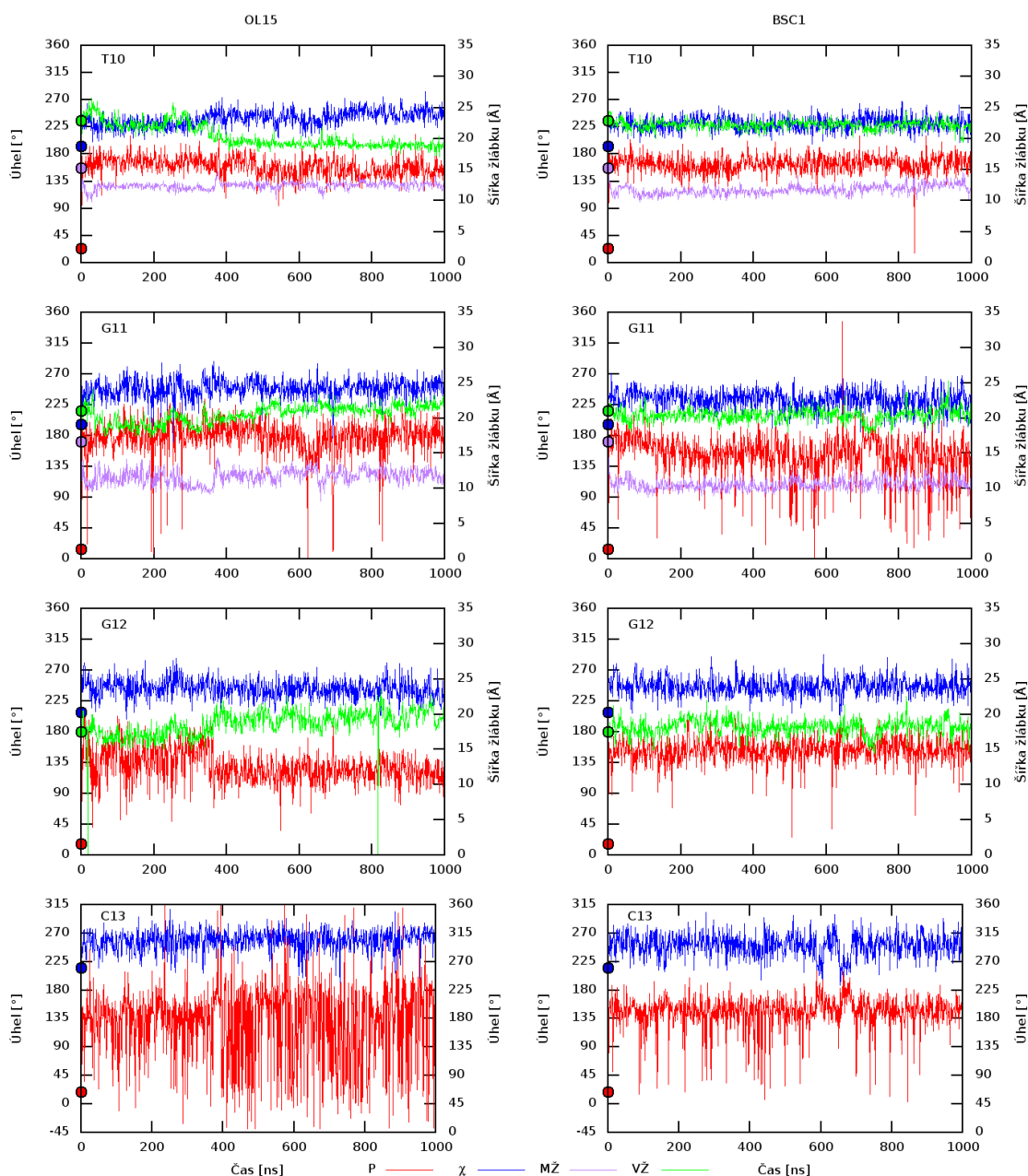
**Tabulka 7: Procentuální zastoupení konformací dle rozdělení podle úhlů  $\delta$  a  $\chi$  v průběhu simulace pro komplex 3PVI. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v A/A konformaci.**

| N         | NtC  | P <sub>0</sub><br>forma | OL15                    |     |     |     | P <sub>0</sub><br>forma | bsc1                    |     |     |     |
|-----------|------|-------------------------|-------------------------|-----|-----|-----|-------------------------|-------------------------|-----|-----|-----|
|           |      |                         | Procentuální zastoupení |     |     |     |                         | Procentuální zastoupení |     |     |     |
|           |      |                         | A/A                     | A/B | B/A | B/B |                         | A/A                     | A/B | B/A | B/B |
| řetězec C |      |                         |                         |     |     |     |                         |                         |     |     |     |
| G3        | BA08 | B/B                     | 0                       | 4   | 0   | 95  | A/A                     | 0                       | 0   | 21  | 78  |
| A4        | AA00 | A/A                     | 1                       | 1   | 0   | 98  | A/A                     | 0                       | 0   | 0   | 100 |
| C5        | AA00 | A/A                     | 1                       | 6   | 3   | 90  | A/A                     | 3                       | 1   | 3   | 93  |
| C6        | AA00 | A/A                     | 7                       | 7   | 10  | 75  | A/A                     | 2                       | 8   | 5   | 84  |
| A7        | AB05 | A/A                     | 12                      | 3   | 0   | 84  | A/A                     | 7                       | 5   | 34  | 54  |
| G8        | BA01 | B/B                     | 2                       | 1   | 8   | 89  | B/B                     | 5                       | 0   | 1   | 93  |
| C9        | AA01 | A/A                     | 35                      | 5   | 16  | 44  | A/A                     | 1                       | 0   | 33  | 66  |
| T10       | AA00 | A/A                     | 5                       | 1   | 1   | 93  | A/A                     | 24                      | 5   | 39  | 32  |
| G11       | AA00 | A/A                     | 1                       | 2   | 27  | 70  | A/A                     | 5                       | 1   | 0   | 94  |
| G12       | AA00 | A/A                     | 12                      | 5   | 13  | 70  | A/A                     | 1                       | 0   | 45  | 54  |
| T13       | AB02 | A/A                     | 3                       | 16  | 3   | 77  | A/A                     | 6                       | 1   | 28  | 65  |
| řetězec D |      |                         |                         |     |     |     |                         |                         |     |     |     |
| G3        | BA08 | B/B                     | 0                       | 2   | 0   | 97  | B/B                     | 0                       | 0   | 0   | 99  |
| A4        | AA00 | B/B                     | 2                       | 4   | 1   | 93  | B/B                     | 3                       | 2   | 1   | 93  |
| C5        | AA00 | A/A                     | 3                       | 9   | 3   | 85  | A/A                     | 3                       | 9   | 5   | 83  |
| C6        | AA00 | A/A                     | 9                       | 3   | 30  | 58  | A/A                     | 3                       | 2   | 36  | 59  |
| A7        | AB05 | A/A                     | 2                       | 1   | 2   | 95  | A/A                     | 5                       | 1   | 0   | 93  |
| G8        | BA01 | B/A                     | 2                       | 0   | 27  | 70  | B/A                     | 2                       | 0   | 25  | 73  |
| C9        | AA01 | A/A                     | 34                      | 5   | 8   | 53  | A/A                     | 16                      | 3   | 51  | 29  |
| T10       | AA00 | A/A                     | 3                       | 0   | 2   | 95  | A/A                     | 1                       | 0   | 1   | 97  |
| G11       | AA00 | A/A                     | 0                       | 1   | 21  | 77  | A/A                     | 0                       | 0   | 42  | 58  |
| G12       | AA00 | A/A                     | 1                       | 0   | 5   | 94  | A/A                     | 8                       | 2   | 27  | 64  |
| T13       | AB02 | A/A                     | 3                       | 28  | 4   | 65  | A/A                     | 1                       | 4   | 6   | 90  |





**Graf č. 4:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů C5, C6, A7 a C9 řetězce D komplexu 3PVI. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .



**Graf č. 5:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů T10, G11, G12 a C13 řetězce D komplexu 3PVI. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .

### 5.3. 3AAF – WRN RecQ helikasa

Ze 14 nukleotidů v komplexu 3AAF se v C3'-endo konformaci vyskytoval v krystalografické struktuře pouze jeden, a, to nukleotid C2. Ten se ovšem v této konformaci dlouho neudržel, jak je vidět v grafu č. 6 pro řetězec B a v grafu č. 7 pro řetězec C. V průběhu simulace se nukleotid C2 vyskytoval převážně v C2'-endo a C3'-exo konformaci (viz tabulka 8 a 9). Lze tedy říct, že se nacházel v tzv. jižní konformaci (při pohledu na pseudorotační kruh na obr. 3), typické pro B formu DNA. Podobné chování vykazují i další nukleotidy komplexu 3AAF, které začínaly buď v čisté B formě, nebo v jedné z přechodných B forem.

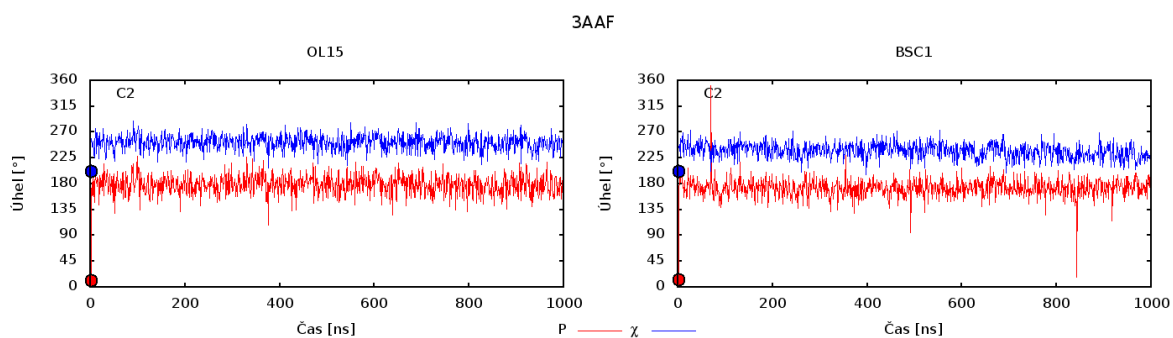
V konformaci A/A se v krystalografické struktuře podle tabulky 10 nacházely nukleotidy C2 v řetězci B a C2 a C3 v řetězci C. Ale ani jeden z těchto nukleotidů se v dané konformaci neudržel a přešel do B/B konformace. Zde je vidět rozdíl mezi silovými poli OL15 a bsc1, kterým je větší procentuální zastoupení přechodné konformace B/A v případě pole bsc1. Pro úhel  $\chi$  je trend velmi podobný jako pro parametr P, tedy, že začíná v hodnotách typických pro A formu, tj. okolo 200°, ale během několika nanosekund dojde k jeho navýšení. Pouze v nukleotidu C3 v řetězci C se v průběhu simulace několikrát vrací k hodnotám 220°. V tabulce 10 můžeme dále vidět, že nukleotidy, které začínaly v A/B formě, konkrétně T8 v obou řetězcích, v této formě setrvávají minimálně třetinu času simulace v případě simulace polem OL15.

**Tabulka 8: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem OL15 pro komplex 3AAF.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci. Zkratka NANT ve sloupci s NtC formami znamená, že danému nukleotidu nebyla přiřazena žádná NtC třída.

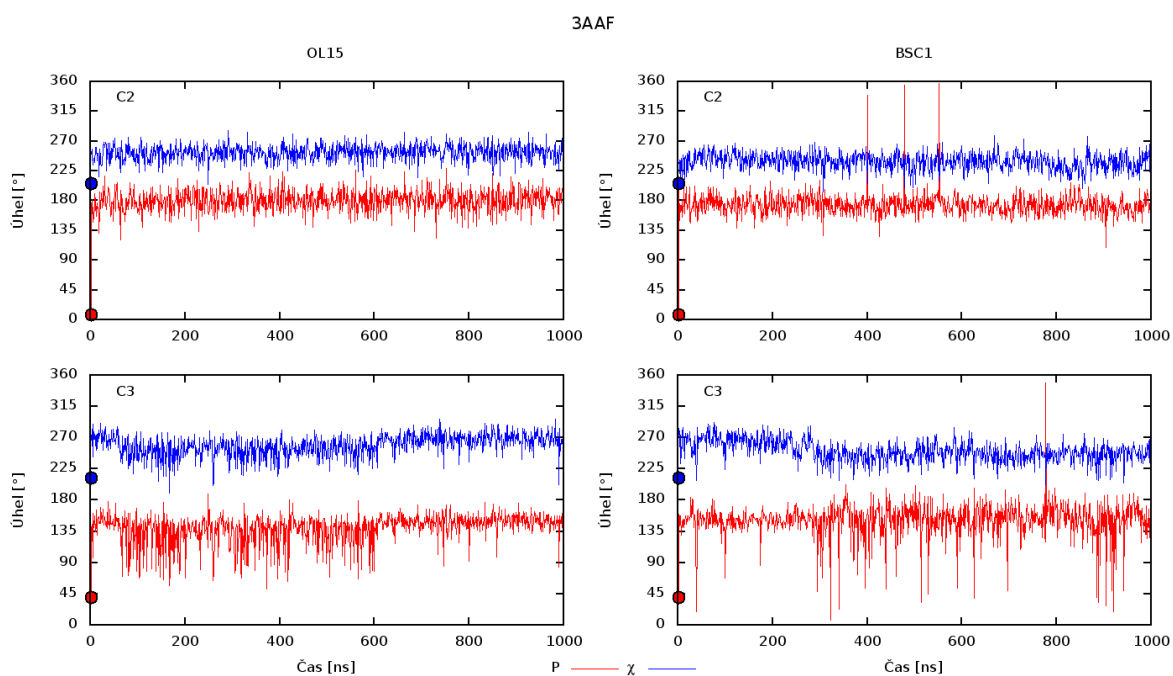
| OL15      |      |                    |                      |                         |          |         |          |         |          |         |         |
|-----------|------|--------------------|----------------------|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| řetězec B |      |                    |                      |                         |          |         |          |         |          |         |         |
| N         | NtC  | P <sub>0</sub> [°] | P <sub>0</sub> forma | Procentuální zastoupení |          |         |          |         |          |         | Ostatní |
|           |      |                    |                      | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo |         |
| C2        | AB01 | 12                 | C3'-endo             | 0                       | 0        | 0       | 0        | 2       | 50       | 47      | 1       |
| C3        | BB01 | 145                | C2'-endo             | 0                       | 0        | 1       | 3        | 30      | 66       | 0       | 0       |
| C4        | BB00 | 116                | C1'-exo              | 0                       | 1        | 2       | 2        | 16      | 61       | 16      | 0       |
| T5        | BB04 | 139                | C1'-exo              | 0                       | 1        | 11      | 14       | 36      | 34       | 3       | 0       |
| A6        | BB01 | 155                | C2'-endo             | 0                       | 0        | 0       | 2        | 25      | 53       | 19      | 0       |
| A7        | BA05 | 139                | C1'-exo              | 0                       | 0        | 0       | 2        | 25      | 57       | 16      | 0       |
| T8        | BB01 | 106                | O4'-endo             | 0                       | 0        | 3       | 25       | 55      | 16       | 0       | 0       |
| T9        | BB00 | 117                | C1'-exo              | 0                       | 2        | 3       | 8        | 30      | 52       | 5       | 0       |
| A10       | BA05 | 185                | Ostatní              | 0                       | 0        | 0       | 1        | 22      | 51       | 23      | 2       |
| G11       | NANT | 59                 | C4'-exo              | 0                       | 0        | 0       | 1        | 22      | 55       | 21      | 1       |
| G12       | NANT | 339                | Ostatní              | 0                       | 0        | 0       | 1        | 39      | 58       | 2       | 0       |
| G13       | NANT | 195                | Ostatní              | 0                       | 0        | 0       | 1        | 12      | 65       | 21      | 0       |
| řetězec C |      |                    |                      |                         |          |         |          |         |          |         |         |
| C2        | AB01 | 7                  | C3'-endo             | 0                       | 0        | 0       | 0        | 1       | 48       | 49      | 1       |
| C3        | BB01 | 40                 | C4'-exo              | 0                       | 0        | 2       | 8        | 36      | 54       | 0       | 0       |
| C4        | BB00 | 147                | C2'-endo             | 0                       | 1        | 2       | 3        | 15      | 61       | 17      | 0       |
| T5        | BB04 | 126                | C1'-exo              | 0                       | 2        | 11      | 17       | 33      | 35       | 1       | 0       |
| A6        | BB01 | 159                | C2'-endo             | 0                       | 0        | 0       | 1        | 21      | 55       | 22      | 1       |
| A7        | BA05 | 123                | C1'-exo              | 0                       | 0        | 0       | 1        | 22      | 56       | 20      | 0       |
| T8        | BB01 | 105                | O4'-endo             | 0                       | 0        | 3       | 28       | 56      | 12       | 0       | 0       |
| T9        | BB00 | 130                | C1'-exo              | 0                       | 3        | 3       | 7        | 29      | 54       | 3       | 0       |
| A10       | BA05 | 137                | C1'-exo              | 0                       | 1        | 1       | 1        | 24      | 52       | 20      | 1       |
| G11       | NANT | 127                | C1'-exo              | 0                       | 0        | 0       | 2        | 20      | 54       | 23      | 1       |
| G12       | NANT | 148                | C2'-endo             | 0                       | 0        | 0       | 1        | 38      | 56       | 4       | 0       |
| G13       | NANT | 174                | C2'-endo             | 0                       | 0        | 0       | 1        | 16      | 63       | 18      | 1       |

**Tabulka 9: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem bsc1 pro komplex 3AAF.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

| bsc1      |      |                       |                      |                         |              |             |              |             |              |             |         |
|-----------|------|-----------------------|----------------------|-------------------------|--------------|-------------|--------------|-------------|--------------|-------------|---------|
| řetězec B |      |                       |                      |                         |              |             |              |             |              |             |         |
| N         | NtC  | P <sub>0</sub><br>[°] | P <sub>0</sub> forma | Procentuální zastoupení |              |             |              |             |              |             |         |
|           |      |                       |                      | C2'-<br>exo             | C3'-<br>endo | C4'-<br>exo | O4'-<br>endo | C1'-<br>exo | C2'-<br>endo | C3'-<br>exo | Ostatní |
| C2        | AB01 | 12                    | C3'-endo             | 0                       | 0            | 0           | 0            | 2           | 69           | 28          | 0       |
| C3        | BB01 | 145                   | C2'-endo             | 0                       | 1            | 1           | 3            | 22          | 65           | 8           | 0       |
| C4        | BB00 | 116                   | C1'-exo              | 0                       | 0            | 4           | 9            | 53          | 33           | 1           | 0       |
| T5        | BB04 | 139                   | C1'-exo              | 0                       | 2            | 4           | 1            | 25          | 66           | 1           | 0       |
| A6        | BB01 | 155                   | C2'-endo             | 0                       | 0            | 0           | 1            | 10          | 58           | 31          | 0       |
| A7        | BA05 | 139                   | C1'-exo              | 0                       | 0            | 0           | 1            | 25          | 65           | 9           | 0       |
| T8        | BB01 | 106                   | O4'-endo             | 0                       | 1            | 4           | 8            | 57          | 30           | 0           | 0       |
| T9        | BB00 | 117                   | C1'-exo              | 0                       | 2            | 1           | 2            | 23          | 69           | 3           | 0       |
| A10       | BA05 | 185                   | Ostatní              | 0                       | 3            | 2           | 1            | 23          | 59           | 12          | 0       |
| G11       | NANT | 59                    | C4'-exo              | 0                       | 1            | 0           | 0            | 8           | 60           | 31          | 0       |
| G12       | NANT | 339                   | Ostatní              | 0                       | 1            | 0           | 0            | 24          | 67           | 8           | 0       |
| G13       | NANT | 195                   | Ostatní              | 0                       | 0            | 0           | 0            | 5           | 71           | 24          | 0       |
| řetězec D |      |                       |                      |                         |              |             |              |             |              |             |         |
| C2        | AB01 | 7                     | C3'-endo             | 0                       | 0            | 0           | 0            | 1           | 69           | 30          | 0       |
| C3        | BB01 | 40                    | C4'-exo              | 0                       | 1            | 1           | 2            | 23          | 68           | 5           | 0       |
| C4        | BB00 | 147                   | C2'-endo             | 0                       | 1            | 2           | 6            | 42          | 44           | 4           | 0       |
| T5        | BB04 | 126                   | C1'-exo              | 0                       | 2            | 4           | 2            | 25          | 65           | 2           | 0       |
| A6        | BB01 | 159                   | C2'-endo             | 0                       | 0            | 0           | 0            | 13          | 60           | 26          | 0       |
| A7        | BA05 | 123                   | C1'-exo              | 0                       | 0            | 0           | 1            | 27          | 62           | 9           | 0       |
| T8        | BB01 | 105                   | O4'-endo             | 0                       | 1            | 3           | 5            | 57          | 33           | 0           | 0       |
| T9        | BB00 | 130                   | C1'-exo              | 0                       | 2            | 2           | 2            | 21          | 70           | 2           | 0       |
| A10       | BA05 | 137                   | C1'-exo              | 0                       | 2            | 1           | 1            | 24          | 60           | 11          | 0       |
| G11       | NANT | 127                   | C1'-exo              | 0                       | 0            | 0           | 0            | 8           | 58           | 34          | 0       |
| G12       | NANT | 148                   | C2'-endo             | 1                       | 2            | 0           | 0            | 21          | 69           | 7           | 0       |
| G13       | NANT | 174                   | C2'-endo             | 0                       | 0            | 0           | 0            | 5           | 76           | 19          | 0       |



**Graf č. 6: Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá) nukleotidu  $C2$  řetězce  $B$  komplexu  $3AAF$ . Vlevo jsou zaznamenány grafy pro pole  $OL15$ , vpravo poté pro pole  $bsc1$ .**



**Graf č. 7: Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá) nukleotidů  $C2$  a  $C3$  řetězce  $C$  komplexu  $3AAF$ . Vlevo jsou zaznamenány grafy pro pole  $OL15$ , vpravo poté pro pole  $bsc1$ .**

**Tabulka 10: Procentuální zastoupení konformací dle rozdělení podle úhlů  $\delta$  a  $\chi$  v průběhu simulace pro komplex 3AAF. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v A/A konformaci.**

| N         | NtC  | P <sub>0</sub> forma | OL15                    |     |     |     | P <sub>0</sub> forma | bsc1                    |     |     |     |
|-----------|------|----------------------|-------------------------|-----|-----|-----|----------------------|-------------------------|-----|-----|-----|
|           |      |                      | Procentuální zastoupení |     |     |     |                      | Procentuální zastoupení |     |     |     |
|           |      |                      | A/A                     | A/B | B/A | B/B |                      | A/A                     | A/B | B/A | B/B |
| řetězec B |      |                      |                         |     |     |     |                      |                         |     |     |     |
| C2        | AB01 | A/A                  | 0                       | 0   | 4   | 96  | A/A                  | 0                       | 0   | 25  | 75  |
| C3        | BB01 | B/B                  | 1                       | 4   | 0   | 95  | B/B                  | 2                       | 7   | 5   | 86  |
| C4        | BB00 | B/B                  | 5                       | 3   | 6   | 86  | B/B                  | 3                       | 21  | 8   | 67  |
| T5        | BB04 | B/B                  | 6                       | 25  | 1   | 67  | B/B                  | 3                       | 6   | 0   | 91  |
| A6        | BB01 | B/B                  | 1                       | 6   | 0   | 93  | B/B                  | 1                       | 1   | 4   | 95  |
| A7        | BA05 | B/B                  | 1                       | 3   | 4   | 91  | B/B                  | 1                       | 3   | 7   | 90  |
| T8        | BB01 | A/B                  | 13                      | 35  | 5   | 46  | A/B                  | 9                       | 19  | 8   | 63  |
| T9        | BB00 | B/B                  | 3                       | 13  | 3   | 81  | B/B                  | 2                       | 4   | 3   | 91  |
| A10       | BA05 | B/B                  | 1                       | 2   | 0   | 97  | B/B                  | 3                       | 3   | 1   | 92  |
| G11       | NANT | A/B                  | 0                       | 3   | 3   | 94  | A/B                  | 1                       | 1   | 4   | 95  |
| G12       | NANT | B/A                  | 1                       | 0   | 0   | 98  | B/A                  | 2                       | 0   | 1   | 97  |
| G13       | NANT | B/A                  | 0                       | 1   | 16  | 83  | B/A                  | 0                       | 0   | 19  | 80  |
| řetězec C |      |                      |                         |     |     |     |                      |                         |     |     |     |
| C2        | AB01 | A/A                  | 0                       | 0   | 2   | 98  | A/A                  | 1                       | 0   | 19  | 80  |
| C3        | BB01 | A/A                  | 2                       | 11  | 0   | 87  | A/A                  | 2                       | 4   | 3   | 91  |
| C4        | BB00 | B/B                  | 3                       | 5   | 10  | 82  | B/B                  | 3                       | 16  | 8   | 73  |
| T5        | BB04 | B/B                  | 7                       | 29  | 1   | 63  | B/B                  | 3                       | 6   | 1   | 90  |
| A6        | BB01 | B/B                  | 0                       | 3   | 1   | 95  | B/B                  | 1                       | 1   | 4   | 95  |
| A7        | BA05 | B/B                  | 1                       | 2   | 4   | 93  | B/B                  | 1                       | 2   | 18  | 78  |
| T8        | BB01 | A/B                  | 10                      | 41  | 4   | 45  | A/B                  | 8                       | 15  | 8   | 69  |
| T9        | BB00 | B/B                  | 3                       | 12  | 2   | 82  | B/B                  | 3                       | 5   | 3   | 89  |
| A10       | BA05 | B/B                  | 1                       | 4   | 1   | 94  | B/B                  | 2                       | 5   | 0   | 93  |
| G11       | NANT | B/B                  | 1                       | 4   | 3   | 93  | B/B                  | 0                       | 0   | 1   | 98  |
| G12       | NANT | B/B                  | 0                       | 1   | 0   | 98  | B/B                  | 3                       | 0   | 2   | 95  |
| G13       | NANT | B/B                  | 0                       | 2   | 7   | 91  | B/B                  | 0                       | 1   | 23  | 76  |

## 5.4. 3KXT – chromatin protein Cren7

Protein-DNA komplex 3KXT neobsahoval ve své krystalografické struktuře ani jeden nukleotid, který by začínal v čisté A-formě, respektive žádná jeho 2-deoxyribosa nebyla v C3'-endo konformaci, viz tabulka 11 a 12. Z tohoto důvodu jsou v grafech č. 8 a 9 zaznamenány průběhy nukleotidů, které byly v krystalografické struktuře v přechodné A/B konformaci. V řetězci B to byly nukleotidy G103 a T105 (graf č. 8) a v řetězci C to byly nukleotidy A112 a T113 (graf č. 9). 2-deoxyribosa nukleotidu G103 se v průběhu simulace silovým polem OL15 nacházela ve všech konformacích a z grafu č. 8 lze vidět, že cukerná složka měla po 400 ns tendenci setrvat spíše v C3'-endo nebo C4'-exo konformaci, zatímco v případě simulace silovým polem bsc1 se deoxyribosa nacházela polovinu času simulace právě v C3'-endo konformaci. V řetězci C nukleotid T113 a jeho cukerná složka měla ve výchozí struktuře hodnotu parametru  $P$  na  $90^\circ$ , ale během několika nanosekund se tato hodnota snížila a poté se pohybovala okolo  $45^\circ$ , což dokládá i procentuální zastoupení C4'-exo konformace v tabulce 8 pro pole OL15. V poli bsc1 je to velmi podobné, zde je ale podobné zastoupení forem C3'-endo a C3'-exo. V grafech vidíme, že šířka malého žlábků je po celou dobu simulace srovnatelná s šířkou velkého žlábků, v případě nukleotidu G103 je i větší. Rozšíření malého žlábků je dáno vazbou proteinu do malého žlábků, jak je popsáno výše. Hodnota úhlu  $\chi$  se ve všech nukleotidech komplexu 3KXT pohybuje okolo  $250^\circ$ , což není typické pro A formu, a proto v tabulce 13 u těchto nukleotidů převažuje přechodná forma A/B. Výjimku tvoří nukleotid T105, u kterého se hodnota úhlu  $\chi$  pohybuje na hodnotě  $225^\circ$ , a proto je u něj v tabulce 13 hojně zastoupena i forma A/A, obzvláště pak v bsc1 poli. Nukleotid A112 v obou polích setrvává polovinu času simulace v konformaci B/B a zbylá část simulace se dělí mezi konformace A/A a A/B, zastoupení těchto dvou konformací se v příslušných polích liší. Při pohledu na graf č. 9 lze vidět, že hodnota parametru  $P$  se u nukleotidu A112 v průběhu celé simulace velmi rychle a významně měnila, z čehož lze usoudit, že ani jedno pole nepopisuje správně tento nukleotid.

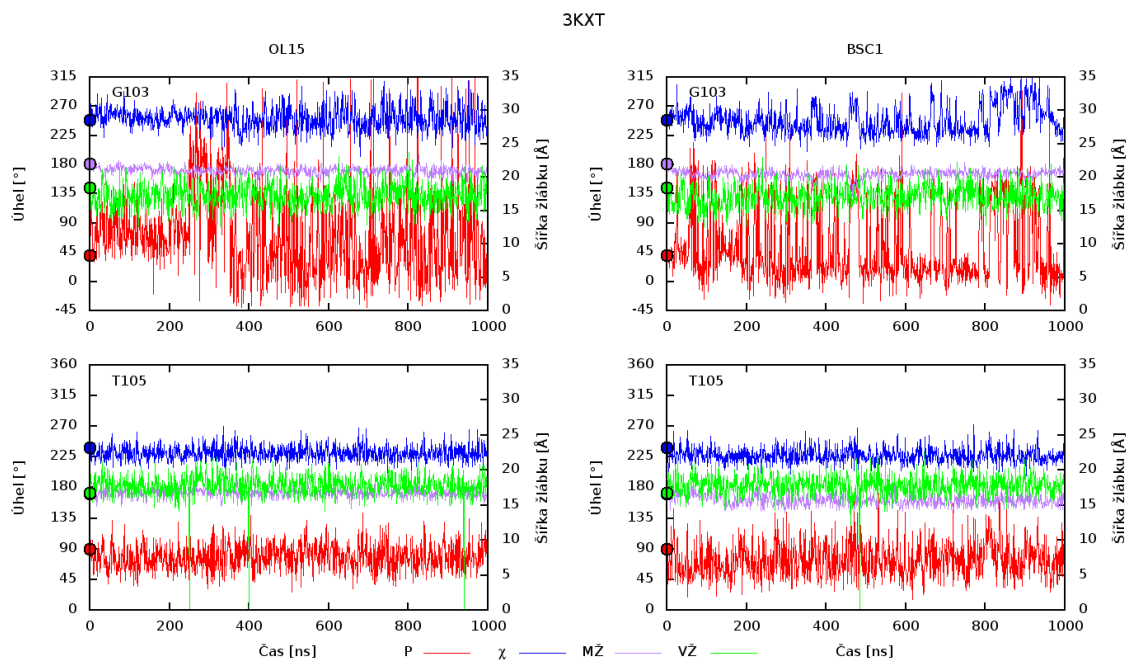


**Tabulka 11: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem OL15 pro komplex 3KXT.** V tabulce zaznamenáno rozdělení pouze dle parametru P.

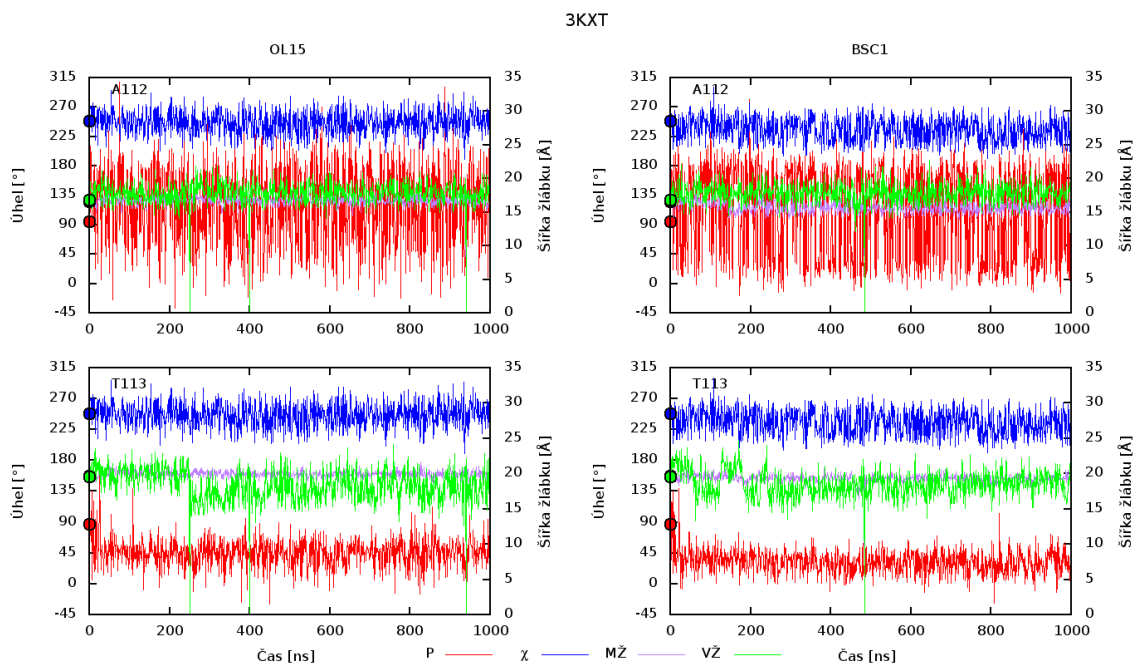
| OL15      |      |                    |                      |                         |          |         |          |         |          |         |         |
|-----------|------|--------------------|----------------------|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| N         | NtC  | P <sub>0</sub> [°] | P <sub>0</sub> forma | Procentuální zastoupení |          |         |          |         |          |         | Ostatní |
|           |      |                    |                      | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo |         |
| řetězec B |      |                    |                      |                         |          |         |          |         |          |         |         |
| C102      | NANT | 145                | C2'-endo             | 0                       | 6        | 17      | 10       | 31      | 33       | 3       | 0       |
| G103      | AB01 | 40                 | C4'-exo              | 12                      | 21       | 23      | 20       | 11      | 5        | 4       | 4       |
| A104      | BA05 | 178                | C2'-endo             | 0                       | 0        | 0       | 0        | 5       | 35       | 52      | 7       |
| T105      | AB01 | 90                 | O4'-endo             | 0                       | 1        | 39      | 54       | 6       | 0        | 0       | 0       |
| C106      | BB07 | 152                | C2'-endo             | 0                       | 0        | 0       | 0        | 14      | 79       | 7       | 0       |
| G107      | BA01 | 162                | C2'-endo             | 0                       | 0        | 0       | 0        | 5       | 58       | 36      | 1       |
| řetězec C |      |                    |                      |                         |          |         |          |         |          |         |         |
| C110      | NANT | 163                | C2'-endo             | 0                       | 0        | 0       | 4        | 27      | 57       | 11      | 0       |
| G111      | AB01 | 163                | C2'-endo             | 0                       | 0        | 2       | 2        | 30      | 51       | 14      | 1       |
| A112      | BA05 | 95                 | O4'-endo             | 2                       | 8        | 17      | 15       | 16      | 26       | 14      | 3       |
| T113      | AB01 | 87                 | O4'-endo             | 2                       | 26       | 65      | 6        | 0       | 0        | 0       | 0       |
| C114      | BB07 | 128                | C1'-exo              | 1                       | 2        | 4       | 4        | 18      | 57       | 12      | 2       |
| G155      | BA01 | 129                | C1'-exo              | 0                       | 0        | 0       | 3        | 43      | 45       | 8       | 1       |

**Tabulka 12: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem bsc1 pro komplex 3KXT.** V tabulce zaznamenáno rozdělení pouze dle parametru P.

| bsc1      |      |           |          |                         |              |             |              |             |              |             |         |
|-----------|------|-----------|----------|-------------------------|--------------|-------------|--------------|-------------|--------------|-------------|---------|
| N         | NtC  | P0<br>[°] | P0 forma | Procentuální zastoupení |              |             |              |             |              |             |         |
|           |      |           |          | C2'-<br>exo             | C3'-<br>endo | C4'-<br>exo | O4'-<br>endo | C1'-<br>exo | C2'-<br>endo | C3'-<br>exo | Ostatní |
| řetězec B |      |           |          |                         |              |             |              |             |              |             |         |
| C102      | NANT | 145       | C2'-endo | 0                       | 19           | 23          | 6            | 13          | 34           | 5           | 0       |
| G103      | AB01 | 40        | C4'-exo  | 10                      | 48           | 15          | 2            | 8           | 13           | 2           | 1       |
| A104      | BA05 | 178       | C2'-endo | 0                       | 1            | 0           | 0            | 5           | 54           | 40          | 1       |
| T105      | AB01 | 90        | O4'-endo | 0                       | 5            | 47          | 38           | 9           | 1            | 0           | 0       |
| C106      | BB07 | 152       | C2'-endo | 0                       | 0            | 0           | 0            | 7           | 88           | 4           | 0       |
| G107      | BA01 | 162       | C2'-endo | 0                       | 0            | 0           | 0            | 0           | 64           | 35          | 0       |
| řetězec C |      |           |          |                         |              |             |              |             |              |             |         |
| C110      | NANT | 163       | C2'-endo | 0                       | 0            | 0           | 0            | 16          | 75           | 9           | 0       |
| G111      | AB01 | 163       | C2'-endo | 0                       | 0            | 2           | 2            | 32          | 59           | 5           | 0       |
| A112      | BA05 | 95        | O4'-endo | 3                       | 28           | 10          | 3            | 4           | 33           | 18          | 1       |
| T113      | AB01 | 87        | O4'-endo | 2                       | 57           | 40          | 1            | 0           | 0            | 0           | 0       |
| C114      | BB07 | 128       | C1'-exo  | 0                       | 2            | 1           | 1            | 14          | 76           | 5           | 0       |
| G155      | BA01 | 129       | C1'-exo  | 0                       | 0            | 0           | 0            | 37          | 58           | 4           | 0       |



**Graf č. 8: Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů G103 a T105 řetězce B komplexu 3KXT. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .**



**Graf č. 9:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů A112 a T113 řetězce C komplexu 3KXT. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole BSC1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .

**Tabulka 13: Procentuální zastoupení konformací dle rozdělení podle úhlů  $\delta$  a  $\chi$  v průběhu simulace pro komplex 3KXT. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v A/B konformaci.**

|           |      | <b>OL15</b>             |                         |     |     |     | <b>bsc1</b> |                         |                         |     |     |     |
|-----------|------|-------------------------|-------------------------|-----|-----|-----|-------------|-------------------------|-------------------------|-----|-----|-----|
| N         | NtC  | P <sub>0</sub><br>forma | Procentuální zastoupení |     |     |     | NtC         | P <sub>0</sub><br>forma | Procentuální zastoupení |     |     |     |
|           |      |                         | A/A                     | A/B | B/A | B/B |             |                         | A/A                     | A/B | B/A | B/B |
| řetězec B |      |                         |                         |     |     |     |             |                         |                         |     |     |     |
| C102      | NANT | B/B                     | 9                       | 25  | 10  | 56  | NANT        | B/B                     | 20                      | 29  | 19  | 32  |
| G103      | AB01 | A/B                     | 4                       | 73  | 0   | 23  | AB01        | A/B                     | 8                       | 68  | 0   | 24  |
| A104      | BA05 | B/B                     | 0                       | 1   | 0   | 98  | BA05        | B/B                     | 1                       | 1   | 4   | 95  |
| T105      | AB01 | A/B                     | 20                      | 78  | 0   | 2   | AB01        | A/B                     | 35                      | 61  | 1   | 3   |
| C106      | BB07 | B/B                     | 0                       | 0   | 0   | 100 | BB07        | B/B                     | 0                       | 0   | 0   | 100 |
| G107      | BA01 | B/B                     | 0                       | 0   | 3   | 97  | BA01        | B/B                     | 0                       | 0   | 9   | 91  |
| řetězec C |      |                         |                         |     |     |     |             |                         |                         |     |     |     |
| C110      | NANT | B/B                     | 3                       | 5   | 7   | 85  | NANT        | B/B                     | 1                       | 1   | 12  | 87  |
| G111      | AB01 | B/B                     | 0                       | 42  | 1   | 56  | AB01        | B/B                     | 2                       | 43  | 1   | 54  |
| A112      | BA05 | A/B                     | 8                       | 36  | 1   | 55  | BA05        | A/B                     | 28                      | 17  | 1   | 54  |
| T113      | AB01 | A/B                     | 2                       | 97  | 0   | 1   | AB01        | A/B                     | 11                      | 88  | 0   | 0   |
| C114      | BB07 | B/B                     | 2                       | 12  | 3   | 83  | BB07        | B/B                     | 3                       | 3   | 10  | 84  |
| G155      | BA01 | B/B                     | 1                       | 3   | 0   | 96  | BA01        | B/B                     | 0                       | 1   | 0   | 99  |

## 5.5. 2VJV – transposasa TnpA

Protein-DNA komplex 2VJV se skládá ze čtyř jednořetězcových DNA vláken, dvou delších a dvou kratších. Simulovány byly všechny řetězce, ale analyzovány byly jen dva delší řetězce. Každý z dlouhých řetězců tvoří smyčku a vytváří Watson-Crikovské vazby a dvoušroubovici se sebou samým. Jedná se o nukleotidy od C20 po G40 s výjimkou T28, 29, T30, T31 a A32 v obou řetězcích C a D. V C3'-endo konformaci se nacházely v krystalografické struktuře nukleotidy C20, C21, C22, C34, G38, G39 a G40 řetězce C, v řetězci D pak navíc nukleotid T31 (viz tabulky 14, 15, 16 a 17). V případě simulací polem OL15 zůstaly v konformaci C3'-endo po většinu času pouze nukleotidy C20 a C34, a to pro oba řetězce. Druhou nejpočetnější konformací, ve které se vyskytovaly C20 a C34, je C2'-exo, nacházející se nachází v severní oblasti pseudorotačního kruhu. Ostatní nukleotidy, které začínaly v C3'-endo konformaci, pak přecházely do C2'-endo nebo C1'-exo konformace. Pro pole bsc1 jsou výsledky podobné, akorát konformace C3'-endo v nukleotidech C20 a C34 má o cca 20 % větší zastoupení.

V těchto dvou nukleotidech se i úhel  $\chi$  pohyboval okolo hodnoty 200°, která je typická pro A formu, v průběhu celé simulace (grafy č. 10, 11, 12 a 13). V ostatních nukleotidech se úhel  $\chi$  pohyboval ve vyšších hodnotách typických spíše pro B formu, výjimkou je jen nukleotid T37 řetězce C v poli bsc1, jehož úhel  $\chi$  se pohyboval v hodnotách okolo 200°. V grafu č. 10, respektive 12, byla hodnota šířky velkého žlábků v nukleotidu C20 v obou řetězcích okolo 15 Å, což je typické pro A formu DNA. Vše tedy nasvědčuje, že C20 po většinu času simulace setrvalo v A/A konformaci, přesněji v řetězci C je to 96 % pro pole OL15 a 100 % pro pole bsc1 (tabulka 18). V řetězci D je to 97 % v poli OL15 a 100 % v poli bsc1 (tabulka 20). Ostatní nukleotidy se z větší části simulace nacházely v konformaci B/B. Výjimkou jsou nukleotidy C23 (který má rovnoměrné zastoupení všech konformací), C34 (A/A), T37 (A/B), G38 (B/A) v obou řetězcích a obou silových polích.

**Tabulka 14: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem OL15 řetězce C komplexu 2VJV.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

| OL15      |      |                    |                      |                         |          |         |          |         |          |         |         |
|-----------|------|--------------------|----------------------|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| N         | NtC  | P <sub>0</sub> [°] | P <sub>0</sub> forma | Procentuální zastoupení |          |         |          |         |          |         | Ostatní |
|           |      |                    |                      | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo |         |
| řetězec C |      |                    |                      |                         |          |         |          |         |          |         |         |
| A17       | NANT | 165                | C2'-endo             | 0                       | 0        | 0       | 0        | 43      | 53       | 3       | 0       |
| A18       | BB00 | 174                | C2'-endo             | 2                       | 2        | 4       | 19       | 24      | 22       | 24      | 2       |
| G19       | NANT | 145                | C2'-endo             | 0                       | 0        | 0       | 0        | 36      | 64       | 0       | 0       |
| C20       | AA00 | 13                 | C3'-endo             | 28                      | 60       | 5       | 2        | 1       | 2        | 1       | 0       |
| C21       | AA00 | 14                 | C3'-endo             | 1                       | 10       | 1       | 1        | 8       | 64       | 15      | 0       |
| C22       | AA00 | 17                 | C3'-endo             | 0                       | 1        | 2       | 16       | 49      | 29       | 2       | 0       |
| C23       | AB05 | 39                 | C4'-exo              | 0                       | 2        | 12      | 37       | 29      | 18       | 3       | 0       |
| T24       | BB00 | 159                | C2'-endo             | 0                       | 0        | 2       | 13       | 44      | 38       | 2       | 0       |
| A25       | BB04 | 159                | C2'-endo             | 0                       | 0        | 0       | 3        | 44      | 48       | 5       | 0       |
| G26       | BB02 | 148                | C2'-endo             | 0                       | 3        | 3       | 8        | 26      | 45       | 14      | 0       |
| C27       | BB00 | 192                | Ostatní              | 0                       | 0        | 1       | 5        | 36      | 49       | 9       | 0       |
| T28       | NANT | 154                | C2'-endo             | 0                       | 5        | 32      | 28       | 18      | 17       | 0       | 0       |
| T29       | NANT | 153                | C2'-endo             | 0                       | 0        | 7       | 15       | 62      | 13       | 1       | 0       |
| T30       | NANT | 136                | C1'-expo             | 0                       | 0        | 0       | 0        | 33      | 65       | 1       | 0       |
| T31       | NANT | 214                | Ostatní              | 0                       | 0        | 0       | 2        | 56      | 38       | 4       | 0       |
| A32       | BBS1 | 198                | Ostatní              | 0                       | 0        | 0       | 0        | 0       | 25       | 68      | 6       |
| G33       | BA13 | 149                | C2'-endo             | 0                       | 0        | 0       | 0        | 57      | 43       | 0       | 0       |
| C34       | NANT | 6                  | C3'-endo             | 54                      | 45       | 1       | 0        | 0       | 0        | 0       | 0       |
| T35       | BB16 | 161                | C2'-endo             | 0                       | 0        | 0       | 0        | 3       | 63       | 34      | 0       |
| A36       | NANT | 157                | C2'-endo             | 0                       | 0        | 0       | 0        | 17      | 74       | 9       | 1       |
| T37       | NANT | 52                 | C4'-exo              | 0                       | 19       | 77      | 4        | 0       | 0        | 0       | 0       |
| G38       | AA01 | 26                 | C3'-endo             | 1                       | 5        | 3       | 5        | 34      | 47       | 4       | 0       |
| G39       | AA00 | 17                 | C3'-endo             | 0                       | 1        | 3       | 11       | 51      | 32       | 2       | 0       |
| G40       | AA01 | 15                 | C3'-endo             | 0                       | 0        | 0       | 5        | 53      | 39       | 2       | 0       |

**Tabulka 15: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem OL15 řetězce D komplexu 2VJV.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

| OL15      |      |                    |                      |                         |          |         |          |         |          |         |         |
|-----------|------|--------------------|----------------------|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| N         | NtC  | P <sub>0</sub> [°] | P <sub>0</sub> forma | Procentuální zastoupení |          |         |          |         |          |         |         |
|           |      |                    |                      | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo | Ostatní |
| řetězec D |      |                    |                      |                         |          |         |          |         |          |         |         |
| A17       | NANT | 158                | C2'-endo             | 0                       | 0        | 0       | 1        | 43      | 53       | 3       | 0       |
| A18       | BB00 | 177                | C2'-endo             | 2                       | 2        | 4       | 16       | 23      | 15       | 32      | 5       |
| G19       | NANT | 142                | C1'-exo              | 0                       | 0        | 0       | 0        | 32      | 68       | 0       | 0       |
| C20       | AA00 | 14                 | C3'-endo             | 28                      | 64       | 4       | 2        | 1       | 1        | 0       | 0       |
| C21       | AA00 | 12                 | C3'-endo             | 2                       | 12       | 1       | 0        | 8       | 64       | 13      | 0       |
| C22       | AA00 | 17                 | C3'-endo             | 0                       | 1        | 2       | 14       | 43      | 37       | 3       | 0       |
| C23       | AB01 | 37                 | C4'-exo              | 0                       | 1        | 9       | 33       | 31      | 22       | 3       | 0       |
| T24       | BB03 | 158                | C2'-endo             | 0                       | 0        | 3       | 14       | 43      | 37       | 2       | 0       |
| A25       | BB02 | 184                | Ostatní              | 0                       | 0        | 0       | 4        | 41      | 46       | 9       | 0       |
| G26       | BA05 | 191                | Ostatní              | 0                       | 1        | 1       | 9        | 32      | 45       | 11      | 0       |
| C27       | NANT | 92                 | O4'-endo             | 0                       | 1        | 0       | 1        | 29      | 64       | 5       | 0       |
| T28       | NANT | 180                | C2'-endo             | 0                       | 0        | 0       | 0        | 28      | 70       | 2       | 0       |
| T29       | NANT | 167                | C2'-endo             | 0                       | 0        | 0       | 0        | 1       | 77       | 22      | 0       |
| T30       | NANT | 167                | C2'-endo             | 5                       | 25       | 2       | 1        | 10      | 53       | 4       | 0       |
| T31       | NANT | 34                 | C3'-endo             | 0                       | 3        | 5       | 5        | 34      | 46       | 5       | 0       |
| A32       | BBS1 | 170                | C2'-endo             | 0                       | 0        | 0       | 3        | 16      | 47       | 34      | 0       |
| G33       | BA13 | 143                | C1'-exo              | 0                       | 0        | 0       | 0        | 53      | 47       | 0       | 0       |
| C34       | NANT | 7                  | C3'-endo             | 55                      | 43       | 1       | 0        | 0       | 0        | 0       | 1       |
| T35       | BB16 | 172                | C2'-endo             | 0                       | 0        | 0       | 0        | 2       | 83       | 15      | 0       |
| A36       | NANT | 154                | C2'-endo             | 0                       | 0        | 0       | 0        | 14      | 63       | 18      | 4       |
| T37       | NANT | 59                 | C4'-exo              | 1                       | 24       | 73      | 2        | 0       | 0        | 0       | 0       |
| G38       | AA01 | 20                 | C3'-endo             | 4                       | 5        | 3       | 4        | 28      | 43       | 11      | 1       |
| G39       | AA00 | 14                 | C3'-endo             | 0                       | 1        | 2       | 10       | 45      | 39       | 3       | 0       |
| G40       | NANT | 21                 | C3'-endo             | 0                       | 0        | 0       | 3        | 57      | 36       | 3       | 0       |

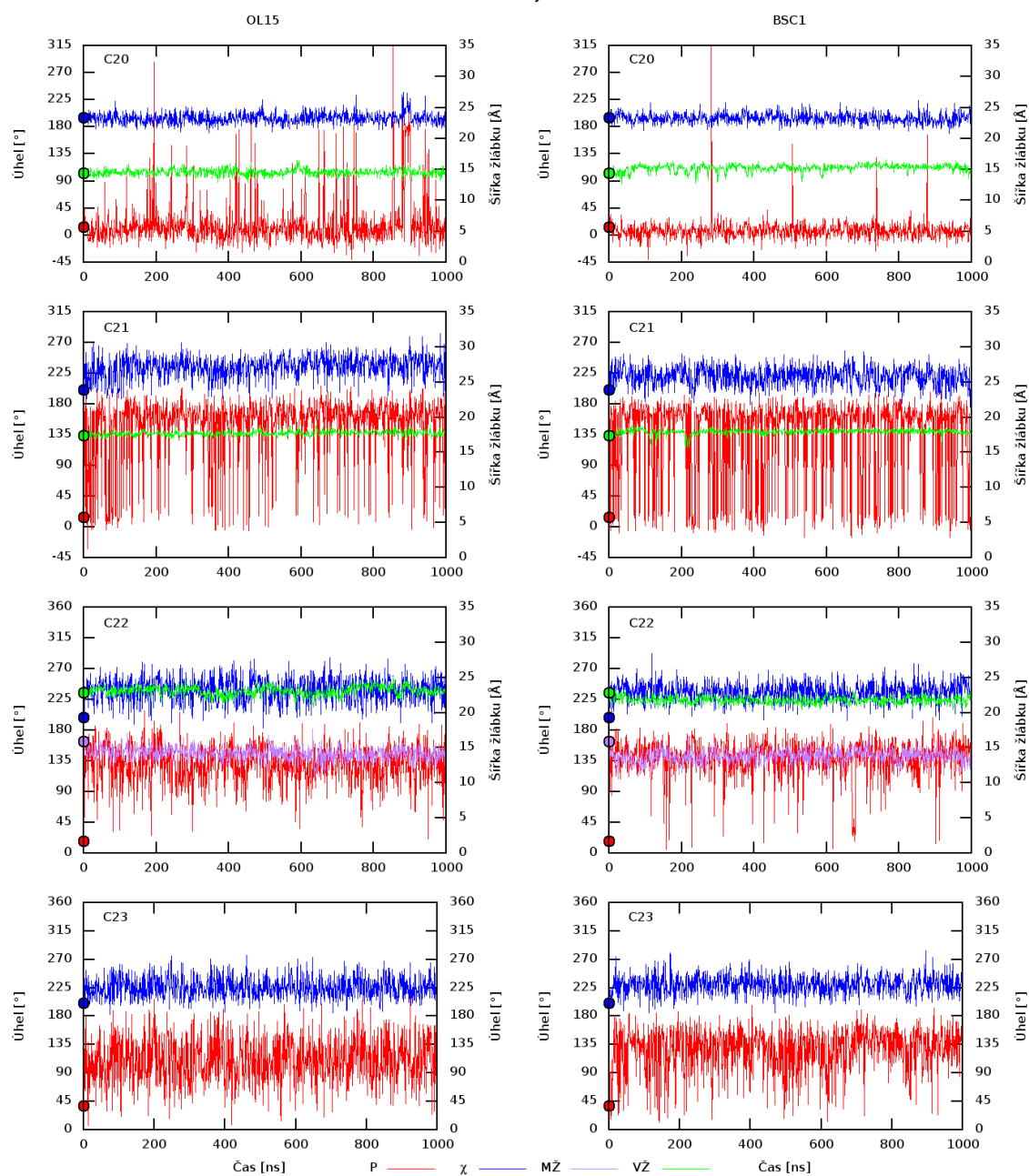


**Tabulka 16: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem bsc1 řetězce C komplexu 2VJV.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

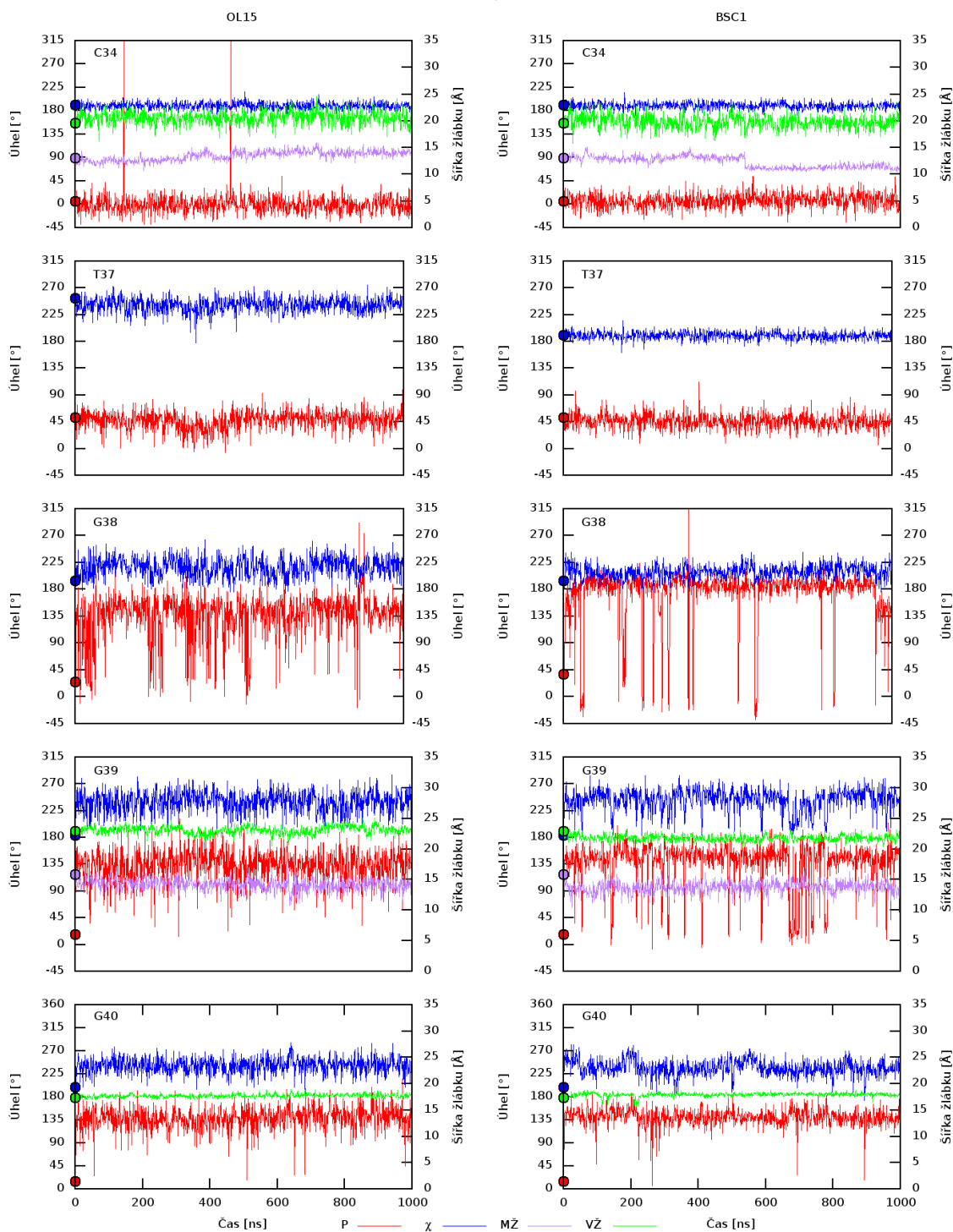
| bsc1      |      |                    |                      |                         |          |         |          |         |          |         |         |
|-----------|------|--------------------|----------------------|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| N         | NtC  | P <sub>0</sub> [°] | P <sub>0</sub> forma | Procentuální zastoupení |          |         |          |         |          |         | Ostatní |
|           |      |                    |                      | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo |         |
| řetězec C |      |                    |                      |                         |          |         |          |         |          |         |         |
| A17       | NANT | 164                | C2'-endo             | 0                       | 0        | 0       | 0        | 35      | 61       | 5       | 0       |
| A18       | BB00 | 174                | C2'-endo             | 0                       | 1        | 1       | 17       | 67      | 13       | 1       | 0       |
| G19       | NANT | 145                | C2'-endo             | 0                       | 0        | 0       | 0        | 12      | 87       | 2       | 0       |
| C20       | AA00 | 13                 | C3'-endo             | 23                      | 76       | 0       | 0        | 0       | 0        | 0       | 0       |
| C21       | AA00 | 14                 | C3'-endo             | 4                       | 19       | 0       | 0        | 3       | 64       | 10      | 0       |
| C22       | AA00 | 17                 | C3'-endo             | 0                       | 2        | 1       | 5        | 46      | 44       | 2       | 0       |
| C23       | AB05 | 39                 | C4'-exo              | 0                       | 3        | 6       | 14       | 39      | 36       | 2       | 0       |
| T24       | BB00 | 160                | C2'-endo             | 0                       | 2        | 4       | 4        | 35      | 54       | 1       | 0       |
| A25       | BB04 | 160                | C2'-endo             | 0                       | 1        | 0       | 1        | 17      | 64       | 16      | 0       |
| G26       | BB02 | 148                | C2'-endo             | 0                       | 2        | 3       | 5        | 38      | 49       | 2       | 0       |
| C27       | BB00 | 192                | Ostatní              | 0                       | 0        | 1       | 0        | 19      | 76       | 4       | 0       |
| T28       | NANT | 145                | C2'-endo             | 0                       | 0        | 0       | 0        | 10      | 89       | 1       | 0       |
| T29       | NANT | 153                | C2'-endo             | 0                       | 0        | 0       | 0        | 15      | 75       | 10      | 0       |
| T30       | NANT | 136                | C1'-exo              | 0                       | 0        | 0       | 0        | 30      | 68       | 1       | 0       |
| T31       | NANT | 214                | Ostatní              | 0                       | 0        | 0       | 0        | 14      | 73       | 13      | 0       |
| A32       | BBS1 | 198                | Ostatní              | 0                       | 0        | 0       | 0        | 0       | 5        | 89      | 6       |
| G33       | BA13 | 149                | C2'-endo             | 0                       | 0        | 0       | 0        | 29      | 71       | 0       | 0       |
| C34       | NANT | 6                  | C3'-endo             | 30                      | 68       | 2       | 0        | 0       | 0        | 0       | 0       |
| T35       | BB16 | 160                | C2'-endo             | 0                       | 0        | 0       | 0        | 2       | 86       | 13      | 0       |
| A36       | NANT | 157                | C2'-endo             | 0                       | 0        | 0       | 0        | 20      | 78       | 1       | 0       |
| T37       | NANT | 52                 | C4'-exo              | 0                       | 18       | 80      | 2        | 0       | 0        | 0       | 0       |
| G38       | AA01 | 26                 | C3'-endo             | 4                       | 2        | 0       | 0        | 4       | 26       | 63      | 0       |
| G39       | AA00 | 19                 | C3'-endo             | 0                       | 7        | 3       | 0        | 31      | 56       | 2       | 0       |
| G40       | AA01 | 15                 | C3'-endo             | 0                       | 0        | 0       | 1        | 59      | 39       | 1       | 0       |

**Tabulka 17: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem bsc1 řetězce D komplexu 2VJV.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

| bsc1      |      |                    |                      |                         |          |         |          |         |          |         |         |
|-----------|------|--------------------|----------------------|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| N         | NtC  | P <sub>0</sub> [°] | P <sub>0</sub> forma | Procentuální zastoupení |          |         |          |         |          |         |         |
|           |      |                    |                      | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo | Ostatní |
| řetězec D |      |                    |                      |                         |          |         |          |         |          |         |         |
| A17       | NANT | 158                | C2'-endo             | 0                       | 0        | 0       | 0        | 36      | 61       | 2       | 0       |
| A18       | BB00 | 177                | C2'-endo             | 0                       | 1        | 1       | 17       | 65      | 14       | 1       | 0       |
| G19       | NANT | 142                | C1'-exo              | 0                       | 0        | 0       | 0        | 11      | 88       | 1       | 0       |
| C20       | AA00 | 14                 | C3'-endo             | 27                      | 72       | 1       | 0        | 0       | 0        | 0       | 0       |
| C21       | AA00 | 13                 | C3'-endo             | 4                       | 21       | 0       | 0        | 3       | 63       | 9       | 0       |
| C22       | AA00 | 17                 | C3'-endo             | 0                       | 2        | 1       | 4        | 35      | 55       | 2       | 0       |
| C23       | AB01 | 37                 | C4'-exo              | 0                       | 6        | 8       | 13       | 35      | 35       | 3       | 0       |
| T24       | BB03 | 158                | C2'-endo             | 0                       | 6        | 5       | 5        | 35      | 47       | 2       | 0       |
| A25       | BB02 | 184                | Ostatní              | 0                       | 2        | 0       | 1        | 24      | 60       | 12      | 0       |
| G26       | BA05 | 191                | Ostatní              | 0                       | 2        | 5       | 10       | 38      | 41       | 4       | 0       |
| C27       | NANT | 92                 | O4'-endo             | 0                       | 1        | 0       | 0        | 18      | 76       | 4       | 0       |
| T28       | NANT | 180                | C2'-endo             | 0                       | 0        | 0       | 0        | 16      | 78       | 6       | 0       |
| T29       | NANT | 167                | C2'-endo             | 0                       | 0        | 0       | 0        | 0       | 67       | 33      | 0       |
| T30       | NANT | 167                | C2'-endo             | 2                       | 20       | 1       | 0        | 5       | 68       | 4       | 0       |
| T31       | NANT | 34                 | C3'-endo             | 0                       | 3        | 2       | 0        | 19      | 67       | 8       | 0       |
| A32       | BBS1 | 170                | C2'-endo             | 0                       | 0        | 0       | 0        | 1       | 40       | 59      | 0       |
| G33       | BA13 | 143                | C1'-exo              | 0                       | 0        | 0       | 0        | 35      | 65       | 0       | 0       |
| C34       | NANT | 7                  | C3'-endo             | 38                      | 61       | 1       | 0        | 0       | 0        | 0       | 0       |
| T35       | BB16 | 172                | C2'-endo             | 0                       | 0        | 0       | 0        | 1       | 87       | 12      | 0       |
| A36       | NANT | 154                | C2'-endo             | 0                       | 0        | 0       | 0        | 21      | 78       | 1       | 0       |
| T37       | NANT | 59                 | C4'-exo              | 0                       | 14       | 83      | 2        | 0       | 0        | 0       | 0       |
| G38       | AA01 | 20                 | C3'-endo             | 20                      | 11       | 0       | 0        | 0       | 20       | 48      | 1       |
| G39       | AA00 | 15                 | C3'-endo             | 5                       | 23       | 2       | 1        | 25      | 42       | 2       | 0       |
| G40       | NANT | 21                 | C3'-endo             | 0                       | 6        | 1       | 1        | 36      | 54       | 2       | 0       |



**Graf č. 10:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů C20, C21, C22 a C23 řetězce C komplexu 2VJV. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v Å.



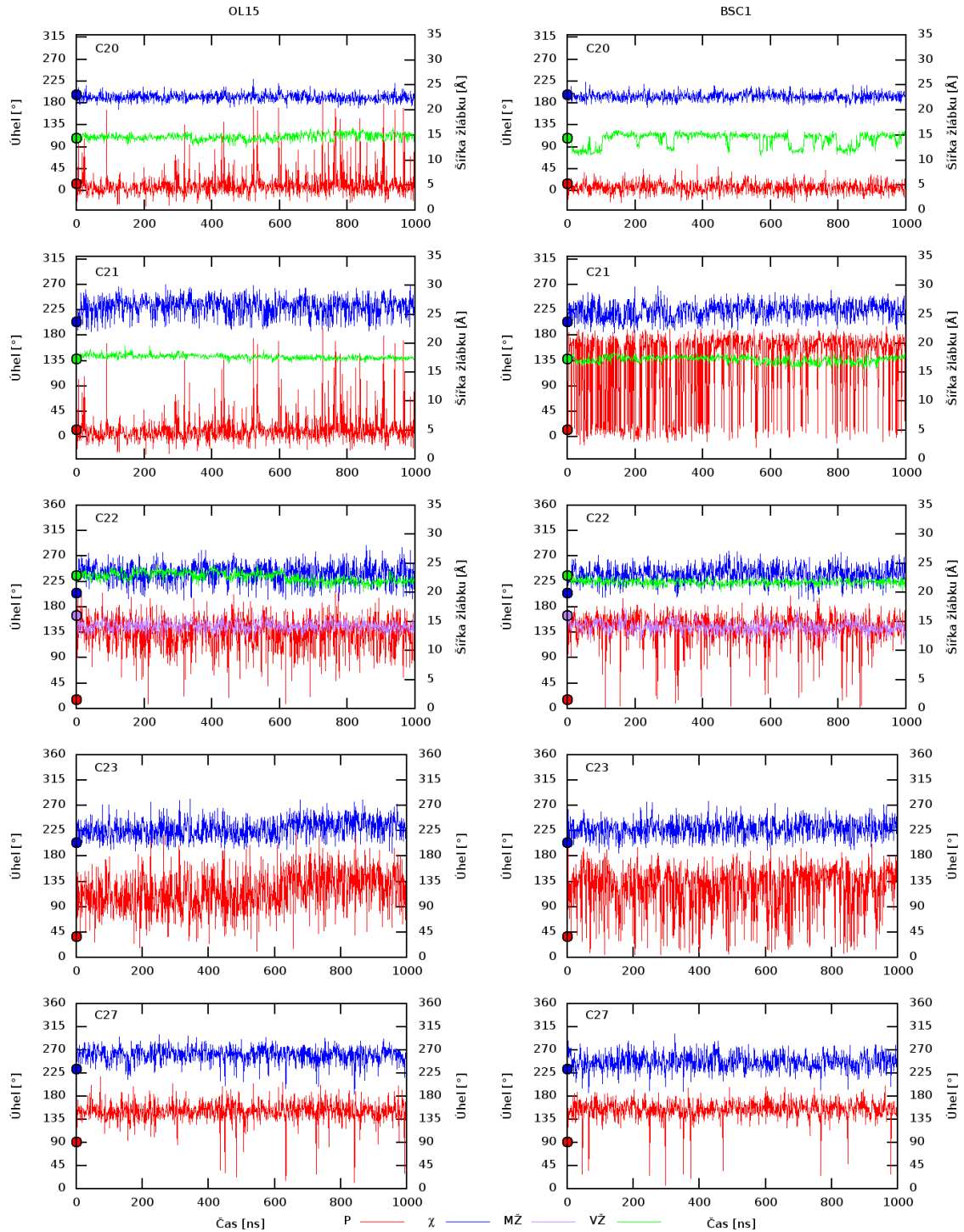
**Graf č. 11:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů C34, T37, G38, G39 a G40 řetězce  $C$  komplexu 2VJV. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^\circ$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v Å.

**Tabulka 18: Procentuální zastoupení konformací dle rozdělení podle úhlů  $\delta$  a  $\chi$  v průběhu simulace řetězce C komplexu 2VJV.** Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v A/A konformaci.

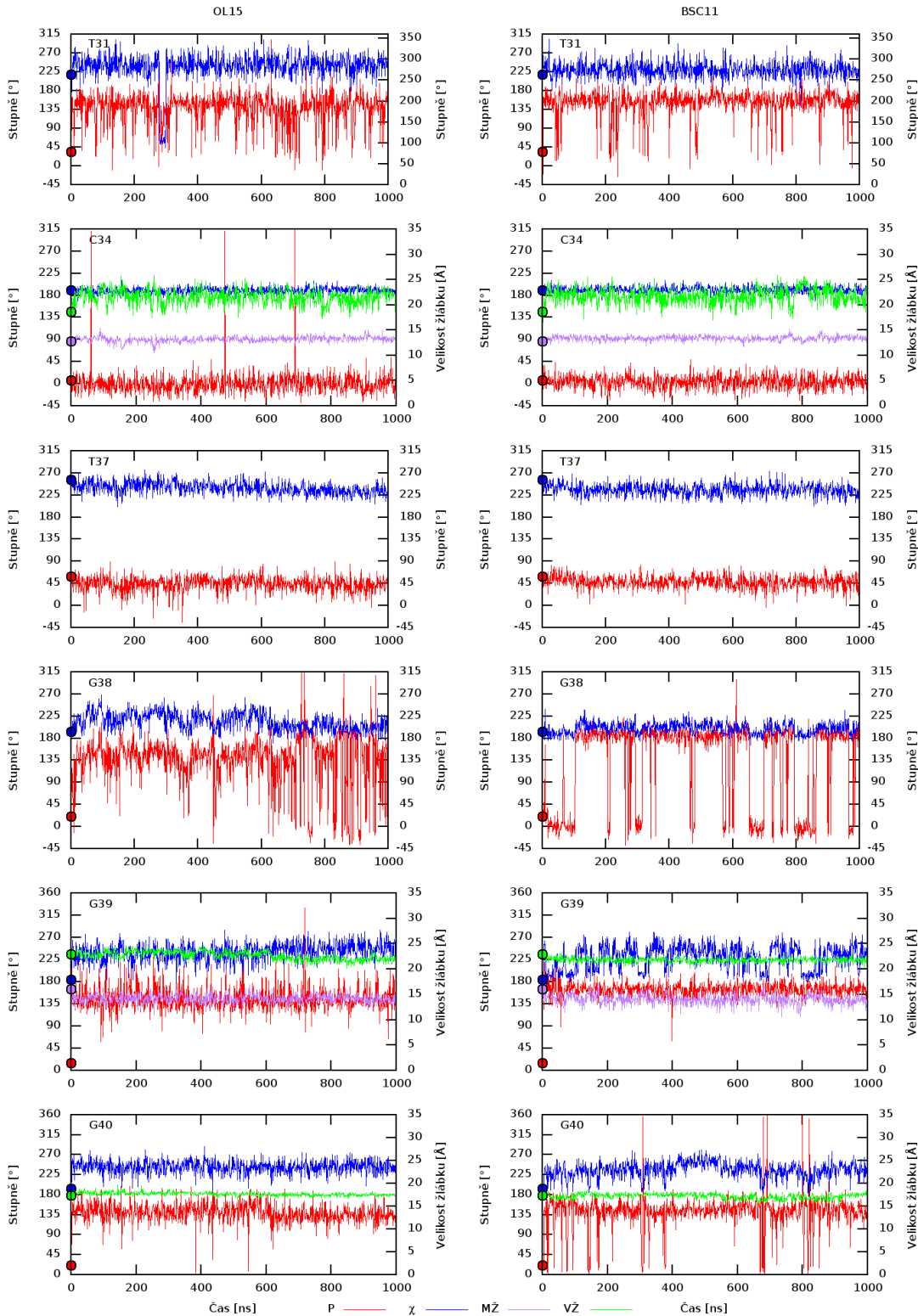
| N         | NtC  | P <sub>0</sub> forma | OL15                    |     |     |     | P <sub>0</sub> forma | BSC1                    |     |     |     |
|-----------|------|----------------------|-------------------------|-----|-----|-----|----------------------|-------------------------|-----|-----|-----|
|           |      |                      | Procentuální zastoupení |     |     |     |                      | Procentuální zastoupení |     |     |     |
|           |      |                      | A/A                     | A/B | B/A | B/B |                      | A/A                     | A/B | B/A | B/B |
| řetězec C |      |                      |                         |     |     |     |                      |                         |     |     |     |
| A17       | NANT | B/A                  | 0                       | 2   | 2   | 96  | B/A                  | 0                       | 0   | 0   | 100 |
| A18       | BB00 | B/A                  | 4                       | 28  | 12  | 56  | B/A                  | 5                       | 26  | 40  | 28  |
| G19       | NANT | B/A                  | 0                       | 0   | 0   | 100 | B/A                  | 0                       | 0   | 0   | 100 |
| C20       | AA00 | A/A                  | 93                      | 1   | 6   | 0   | A/A                  | 100                     | 0   | 0   | 0   |
| C21       | AA00 | A/A                  | 4                       | 10  | 27  | 59  | A/A                  | 14                      | 9   | 51  | 26  |
| C22       | AA00 | A/A                  | 5                       | 27  | 18  | 50  | A/A                  | 2                       | 16  | 24  | 58  |
| C23       | AB05 | A/A                  | 26                      | 34  | 25  | 15  | A/A                  | 10                      | 24  | 29  | 37  |
| T24       | BB00 | B/A                  | 1                       | 22  | 10  | 67  | B/A                  | 1                       | 11  | 8   | 80  |
| A25       | BB04 | B/A                  | 0                       | 8   | 4   | 87  | B/A                  | 1                       | 3   | 12  | 84  |
| G26       | BB02 | B/A                  | 4                       | 15  | 21  | 61  | B/A                  | 3                       | 11  | 10  | 76  |
| C27       | BB00 | B/A                  | 1                       | 10  | 5   | 84  | B/A                  | 1                       | 1   | 53  | 45  |
| T28       | NANT | B/A                  | 13                      | 53  | 9   | 26  | B/A                  | 0                       | 0   | 14  | 86  |
| T29       | NANT | B/A                  | 4                       | 27  | 14  | 55  | B/A                  | 0                       | 0   | 42  | 58  |
| T30       | NANT | B/A                  | 0                       | 2   | 25  | 73  | B/A                  | 0                       | 0   | 80  | 19  |
| T31       | NANT | B/A                  | 1                       | 3   | 75  | 21  | B/A                  | 0                       | 1   | 83  | 17  |
| A32       | BBS1 | B/A                  | 0                       | 0   | 100 | 0   | B/A                  | 0                       | 0   | 100 | 0   |
| G33       | BA13 | B/A                  | 0                       | 0   | 0   | 100 | B/A                  | 0                       | 0   | 0   | 100 |
| C34       | NANT | A/A                  | 100                     | 0   | 0   | 0   | A/A                  | 100                     | 0   | 0   | 0   |
| T35       | BB16 | B/A                  | 0                       | 0   | 81  | 18  | B/A                  | 0                       | 0   | 49  | 51  |
| A36       | NANT | B/A                  | 0                       | 0   | 0   | 100 | B/A                  | 0                       | 0   | 0   | 100 |
| T37       | NANT | A/A                  | 6                       | 94  | 0   | 0   | A/A                  | 21                      | 79  | 0   | 0   |
| G38       | AA01 | A/A                  | 11                      | 6   | 59  | 24  | A/A                  | 6                       | 0   | 90  | 4   |
| G39       | AA00 | A/A                  | 5                       | 20  | 18  | 57  | A/A                  | 7                       | 6   | 10  | 77  |
| G40       | AA01 | A/A                  | 1                       | 17  | 16  | 66  | A/A                  | 1                       | 10  | 25  | 64  |

**Tabulka 19: Procentuální zastoupení konformací dle rozdělení podle úhlů  $\delta$  a  $\chi$  v průběhu simulace řetězce D komplexu 2VJV. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v A/A konformaci.**

| N         | NtC  | P <sub>0</sub> forma | OL15                    |     |     |     | P <sub>0</sub> forma | BSC1                    |     |     |     |
|-----------|------|----------------------|-------------------------|-----|-----|-----|----------------------|-------------------------|-----|-----|-----|
|           |      |                      | Procentuální zastoupení |     |     |     |                      | Procentuální zastoupení |     |     |     |
|           |      |                      | A/A                     | A/B | B/A | B/B |                      | A/A                     | A/B | B/A | B/B |
| řetězec D |      |                      |                         |     |     |     |                      |                         |     |     |     |
| A17       | NANT | B/A                  | 0                       | 3   | 1   | 96  | B/A                  | 0                       | 1   | 1   | 98  |
| A18       | BB00 | B/A                  | 3                       | 27  | 8   | 61  | B/A                  | 5                       | 28  | 32  | 35  |
| G19       | NANT | B/A                  | 0                       | 0   | 0   | 100 | B/A                  | 0                       | 0   | 0   | 100 |
| C20       | AA00 | A/A                  | 97                      | 0   | 3   | 0   | A/A                  | 100                     | 0   | 0   | 0   |
| C21       | AA00 | A/A                  | 5                       | 10  | 31  | 54  | A/A                  | 16                      | 9   | 39  | 35  |
| C22       | AA00 | A/A                  | 4                       | 23  | 18  | 55  | A/A                  | 2                       | 11  | 22  | 65  |
| C23       | AB01 | A/A                  | 19                      | 33  | 24  | 25  | A/A                  | 10                      | 24  | 31  | 34  |
| T24       | BB03 | B/A                  | 8                       | 92  | 0   | 0   | B/A                  | 13                      | 87  | 0   | 0   |
| A25       | BB02 | B/A                  | 0                       | 8   | 4   | 87  | B/A                  | 1                       | 5   | 10  | 83  |
| G26       | BA05 | B/A                  | 2                       | 15  | 18  | 66  | B/A                  | 4                       | 19  | 22  | 56  |
| C27       | NANT | A/A                  | 0                       | 3   | 2   | 94  | A/A                  | 0                       | 2   | 11  | 87  |
| T28       | NANT | B/A                  | 0                       | 0   | 30  | 70  | B/A                  | 0                       | 0   | 28  | 71  |
| T29       | NANT | B/A                  | 0                       | 0   | 97  | 3   | B/A                  | 0                       | 0   | 97  | 3   |
| T30       | NANT | B/A                  | 27                      | 6   | 56  | 11  | B/A                  | 21                      | 2   | 69  | 9   |
| T31       | NANT | A/A                  | 5                       | 12  | 22  | 62  | A/A                  | 3                       | 4   | 45  | 48  |
| A32       | BBS1 | B/A                  | 12                      | 0   | 88  | 0   | B/A                  | 0                       | 0   | 100 | 0   |
| G33       | BA13 | B/A                  | 0                       | 0   | 0   | 100 | B/A                  | 0                       | 0   | 0   | 100 |
| C34       | NANT | A/A                  | 100                     | 0   | 0   | 0   | A/A                  | 100                     | 0   | 0   | 0   |
| T35       | BB16 | B/A                  | 0                       | 0   | 48  | 52  | B/A                  | 0                       | 0   | 46  | 54  |
| A36       | NANT | B/A                  | 0                       | 0   | 0   | 100 | B/A                  | 0                       | 0   | 0   | 100 |
| T37       | NANT | A/A                  | 8                       | 92  | 0   | 0   | A/A                  | 14                      | 86  | 0   | 0   |
| G38       | AA01 | A/A                  | 15                      | 4   | 58  | 23  | A/A                  | 31                      | 0   | 69  | 1   |
| G39       | AA00 | A/A                  | 3                       | 19  | 16  | 62  | A/A                  | 24                      | 9   | 16  | 50  |
| G40       | NANT | A/A                  | 1                       | 17  | 12  | 70  | A/A                  | 5                       | 7   | 24  | 64  |



**Graf č. 12:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů C20, C21, C22, C23 a C27 řetězce D komplexu 2VJV. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^\circ$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .



**Graf č. 13:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů T31, C34, T37, G38, G39 a G40 řetězce C komplexu 2VJV. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .



## 5.6. 1DP7 – regulační faktor X1

V komplexu 1DP7 bylo potřeba zrcadlově rozšířit původní pdb soubor, který obsahoval jen jeden protein a jeden řetězec DNA, o druhý řetězec a druhý protein. Z celkem 16 nukleotidů se v C3'-endo konformaci v krystalografické struktuře nacházely pouze nukleotidy C7 a A8 v obou řetězcích (tabulka 20 a 21). Nukleotid A8 v řetězci C vykazuje podobné chování v obou polích, tedy po většinu času simulace zůstává v C3'-endo konformaci, konkrétně 68 % pro pole OL15 a 77 % pro bsc1. Nukleotid C7 v řetězci C se v poli OL15 nachází převážnou část simulace v C1'-exo konformaci, ale v poli bsc1 zůstává v C3'-endo. Tento vývoj je zaznamenán v grafu č. 14, ve kterém lze vidět, že C7 v poli OL1 se po prvních cca 100 ns přesune do konformace C1'-exo. Stejný trend má i úhel  $\chi$  v C7, kdy se opět po 100 ns zvýší až na hodnotu 260°, která je typická pro B formu. V poli bsc1 parametr P i úhel  $\chi$  nukleotidu C7 po většinu času simulace setrvávají v hodnotách typických pro A formu.

V řetězci D se pak oba nukleotidy v poli OL15 nacházejí ve všech konformacích od C3'-endo až po C2'-endo, u nukleotidu A8 lehce převládá C2'-endo konformace, ve které tráví deoxyribosa 46 % času simulace. V tomto nukleotidu se úhel  $\chi$  na začátku simulace pohybuje v hodnotách okolo 200 – 220°, ale po 400 ns se dostává na hodnotu 260°. V nukleotidu C7 úhel  $\chi$  osciluje po celou dobu simulace mezi hodnotami 200° a 260°. V poli bsc1 pak nukleotidy řetězce D setrvávají spíše v C3'-endo konformaci 46 % pro nukleotid C7 a 77 % pro A8.

V tabulce 22 jsou zvýrazněny nukleotidy, které začínaly v A/A konformaci, jsou jimi T3, C7, A8 a T9 pro oba řetězce. Nukleotid T3 se v této konformaci vůbec neudrží a ihned přechází do B/B konformace v OL15 i bsc1, s malým zastoupením i B/A konformace. Obdobné chování vykazuje i nukleotid T9, který se ale nachází pouze v B/B konformaci. V řetězci D v poli OL15 tráví více času v přechodných konformacích A/B pouze T9, ale pořád zde převažuje konformace B/B, a to 62 % času simulace. Nukleotid C7 se v poli OL15 v řetězci C nachází nejvíce v konformaci B/B (65 %), ale v řetězci D je zastoupení konformací v průběhu simulace rozděleno mezi všechny čtyři. V poli bsc1 ale převažuje v obou řetězcích konformace A/A. Nukleotid A8 se v poli bsc1 chová podobně, opět zde převažuje konformace A/A, přes 80 % v obou řetězcích. V poli OL15 v řetězci C má pak největší zastoupení konformace A/A, v řetězci D pak konformace B/B,

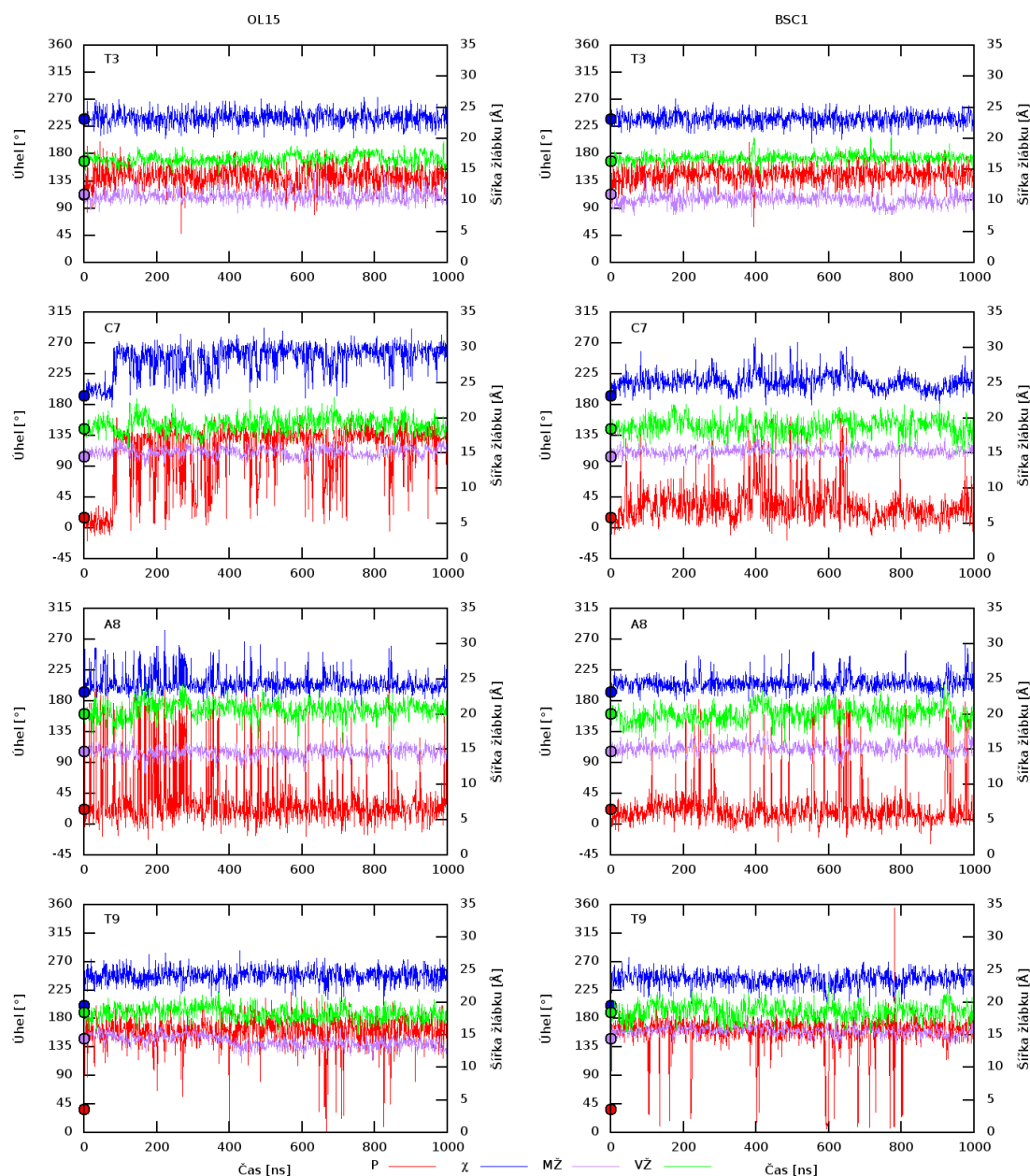
i když v tomto řetězci se pětinu simulace nachází v A/A konformaci. To lze vidět v grafu č. 15, kde v A8 v prvních 400 ns se deoxyribosa vrací k hodnotám okolo  $20^\circ$  a úhel  $\chi$  se pohybuje v oblasti  $200 - 220^\circ$ .

**Tabulka 20: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem OL15 pro komplex 1DP7.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

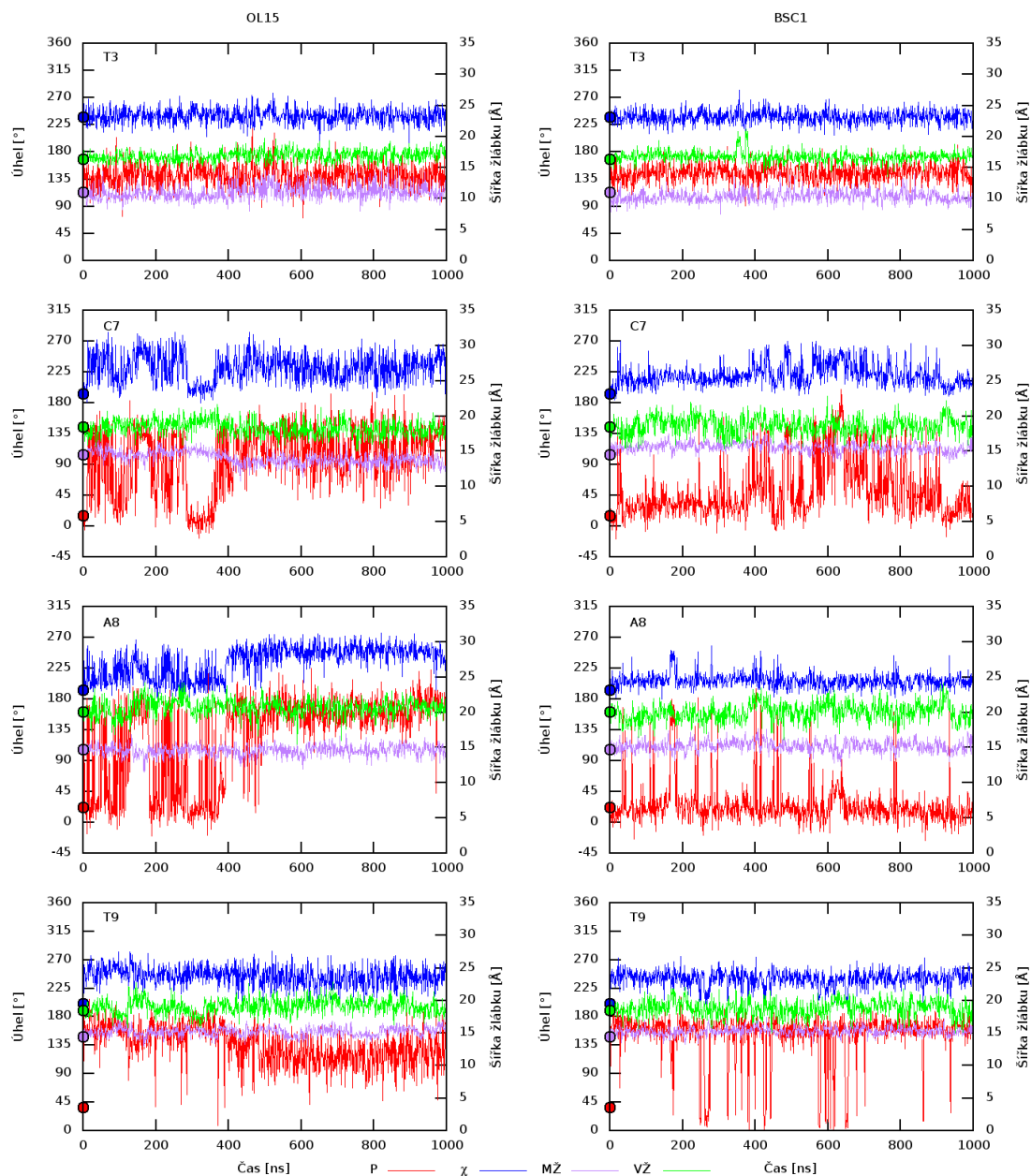
| OL15      |      |                       |                      |                         |              |             |              |             |              |             |         |
|-----------|------|-----------------------|----------------------|-------------------------|--------------|-------------|--------------|-------------|--------------|-------------|---------|
| řetězec C |      |                       |                      |                         |              |             |              |             |              |             |         |
| N         | NtC  | P <sub>0</sub><br>[°] | P <sub>0</sub> forma | Procentuální zastoupení |              |             |              |             |              |             |         |
|           |      |                       |                      | C2'-<br>exo             | C3'-<br>endo | C4'-<br>exo | O4'-<br>endo | C1'-<br>exo | C2'-<br>endo | C3'-<br>exo | Ostatní |
| G2        | BA06 | 170                   | C2'-endo             | 0                       | 0            | 0           | 0            | 2           | 65           | 32          | 0       |
| T3        | BB01 | 113                   | C1'-exo              | 0                       | 0            | 0           | 3            | 50          | 45           | 1           | 0       |
| T4        | BB00 | 152                   | C2'-endo             | 0                       | 0            | 1           | 8            | 48          | 42           | 0           | 0       |
| A5        | BB00 | 171                   | C2'-endo             | 0                       | 1            | 0           | 2            | 10          | 50           | 34          | 2       |
| C6        | BA13 | 152                   | C2'-endo             | 0                       | 0            | 0           | 3            | 28          | 58           | 10          | 0       |
| C7        | AA00 | 15                    | C3'-endo             | 2                       | 12           | 9           | 5            | 60          | 12           | 0           | 0       |
| A8        | AA00 | 22                    | C3'-endo             | 8                       | 68           | 14          | 1            | 1           | 7            | 2           | 0       |
| T9        | AB05 | 37                    | C4'-exo              | 0                       | 1            | 1           | 1            | 13          | 69           | 14          | 0       |
| G10       | BB00 | 164                   | C2'-endo             | 0                       | 0            | 0           | 3            | 47          | 48           | 1           | 0       |
| G11       | BB00 | 181                   | Ostatní              | 0                       | 1            | 2           | 5            | 18          | 53           | 19          | 1       |
| T12       | BB00 | 139                   | C1'-exo              | 0                       | 0            | 1           | 4            | 33          | 56           | 5           | 0       |
| A13       | BB00 | 167                   | C2'-endo             | 0                       | 0            | 0           | 1            | 26          | 50           | 20          | 1       |
| A14       | BB04 | 155                   | C2'-endo             | 0                       | 0            | 0           | 2            | 27          | 57           | 13          | 0       |
| C15       | BB00 | 169                   | C2'-endo             | 0                       | 2            | 2           | 7            | 29          | 48           | 12          | 0       |
| řetězec D |      |                       |                      |                         |              |             |              |             |              |             |         |
| G2        | BA06 | 170                   | C2'-endo             | 0                       | 0            | 0           | 0            | 2           | 69           | 29          | 0       |
| T3        | BB01 | 113                   | C1'-exo              | 0                       | 0            | 0           | 3            | 53          | 42           | 2           | 0       |
| T4        | BB00 | 152                   | C2'-endo             | 0                       | 0            | 2           | 8            | 41          | 48           | 1           | 0       |
| A5        | BB00 | 171                   | C2'-endo             | 0                       | 1            | 1           | 1            | 8           | 50           | 38          | 2       |
| C6        | BA13 | 152                   | C2'-endo             | 0                       | 0            | 0           | 5            | 39          | 51           | 5           | 0       |
| C7        | AA00 | 15                    | C3'-endo             | 2                       | 13           | 13          | 20           | 36          | 14           | 1           | 0       |
| A8        | AA00 | 22                    | C3'-endo             | 1                       | 19           | 5           | 2            | 12          | 46           | 14          | 0       |
| T9        | AB05 | 37                    | C4'-exo              | 0                       | 0            | 2           | 17           | 38          | 37           | 6           | 0       |
| G10       | BB00 | 164                   | C2'-endo             | 0                       | 0            | 0           | 1            | 39          | 58           | 1           | 0       |
| G11       | BB00 | 181                   | Ostatní              | 0                       | 0            | 0           | 1            | 11          | 58           | 28          | 1       |
| T12       | BB00 | 139                   | C1'-exo              | 0                       | 0            | 1           | 9            | 37          | 50           | 1           | 0       |
| A13       | BB00 | 167                   | C2'-endo             | 0                       | 0            | 0           | 0            | 12          | 57           | 28          | 1       |
| A14       | BB04 | 155                   | C2'-endo             | 0                       | 0            | 0           | 2            | 26          | 51           | 19          | 1       |
| C15       | BB00 | 169                   | C2'-endo             | 0                       | 1            | 3           | 11           | 39          | 40           | 6           | 0       |

**Tabulka 21: Procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace polem bsc1 pro komplex 1DP7.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci.

| <b>bsc1</b> |      |                    |                      |                         |          |         |          |         |          |         |         |
|-------------|------|--------------------|----------------------|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| řetězec C   |      |                    |                      |                         |          |         |          |         |          |         |         |
| N           | NtC  | P <sub>0</sub> [°] | P <sub>0</sub> forma | Procentuální zastoupení |          |         |          |         |          |         |         |
|             |      |                    |                      | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo | Ostatní |
| G2          | BA06 | 170                | C2'-endo             | 0                       | 0        | 0       | 0        | 1       | 68       | 30      | 0       |
| T3          | BB01 | 113                | C1'-exo              | 0                       | 0        | 0       | 0        | 45      | 54       | 1       | 0       |
| T4          | BB00 | 151                | C2'-endo             | 0                       | 0        | 0       | 1        | 39      | 59       | 1       | 0       |
| A5          | BB00 | 171                | C2'-endo             | 1                       | 14       | 3       | 1        | 6       | 53       | 22      | 0       |
| C6          | BA13 | 152                | C2'-endo             | 0                       | 0        | 2       | 2        | 28      | 63       | 4       | 0       |
| C7          | AA00 | 15                 | C3'-endo             | 3                       | 61       | 30      | 2        | 3       | 1        | 0       | 0       |
| A8          | AA00 | 22                 | C3'-endo             | 9                       | 77       | 9       | 0        | 1       | 4        | 0       | 0       |
| T9          | AB05 | 37                 | C4'-exo              | 0                       | 3        | 0       | 1        | 6       | 80       | 10      | 0       |
| G10         | BB00 | 164                | C2'-endo             | 0                       | 1        | 0       | 0        | 46      | 52       | 0       | 0       |
| G11         | BB00 | 181                | Ostatní              | 0                       | 1        | 0       | 0        | 4       | 56       | 37      | 1       |
| T12         | BB00 | 139                | C1'-exo              | 0                       | 0        | 1       | 1        | 23      | 72       | 2       | 0       |
| A13         | BB00 | 167                | C2'-endo             | 0                       | 2        | 0       | 0        | 6       | 60       | 31      | 0       |
| A14         | BB04 | 155                | C2'-endo             | 0                       | 0        | 0       | 0        | 23      | 65       | 12      | 0       |
| C15         | BB00 | 169                | C2'-endo             | 0                       | 0        | 1       | 4        | 33      | 58       | 3       | 0       |
| řetězec D   |      |                    |                      |                         |          |         |          |         |          |         |         |
| G2          | BA06 | 170                | C2'-endo             | 0                       | 0        | 0       | 0        | 1       | 66       | 33      | 0       |
| T3          | BB01 | 113                | C1'-exo              | 0                       | 0        | 0       | 0        | 45      | 53       | 1       | 0       |
| T4          | BB00 | 151                | C2'-endo             | 0                       | 0        | 0       | 1        | 40      | 58       | 0       | 0       |
| A5          | BB00 | 171                | C2'-endo             | 1                       | 15       | 2       | 0        | 5       | 53       | 24      | 0       |
| C6          | BA13 | 152                | C2'-endo             | 0                       | 1        | 2       | 2        | 29      | 63       | 3       | 0       |
| C7          | AA00 | 15                 | C3'-endo             | 2                       | 46       | 32      | 5        | 10      | 5        | 0       | 0       |
| A8          | AA00 | 22                 | C3'-endo             | 8                       | 77       | 10      | 0        | 1       | 3        | 0       | 0       |
| T9          | AB05 | 37                 | C4'-exo              | 0                       | 7        | 1       | 0        | 7       | 77       | 8       | 0       |
| G10         | BB00 | 164                | C2'-endo             | 0                       | 0        | 0       | 0        | 38      | 60       | 1       | 0       |
| G11         | BB00 | 181                | Ostatní              | 0                       | 1        | 0       | 0        | 4       | 58       | 36      | 0       |
| T12         | BB00 | 139                | C1'-exo              | 0                       | 0        | 0       | 1        | 29      | 68       | 1       | 0       |
| A13         | BB00 | 167                | C2'-endo             | 0                       | 1        | 0       | 0        | 8       | 63       | 28      | 0       |
| A14         | BB04 | 155                | C2'-endo             | 0                       | 0        | 0       | 1        | 23      | 66       | 10      | 0       |
| C15         | BB00 | 169                | C2'-endo             | 0                       | 1        | 1       | 4        | 31      | 59       | 4       | 0       |



**Graf č. 14:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů T3, C7, A8 a T9 řetězce C komplexu 1DP7. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole bsc1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .



**Graf č. 15:** Časový průběh parametrů  $P$  (červená), úhlu  $\chi$  (modrá), šířky velkého (zelená) a malého (fialová) žlábků nukleotidů T3, C7, A8 a T9 řetězce D komplexu IDP7. Vlevo jsou zaznamenány grafy pro pole OL15, vpravo poté pro pole BSC1. Levá y-osa znázorňuje velikost úhlů ve  $^{\circ}$  pro parametr  $P$  a úhel  $\chi$ , pravá y-osa znázorňuje šířku žlábků v  $\text{Å}$ .

**Tabulka 22: Procentuální zastoupení konformací dle rozdělení podle úhlů  $\delta$  a  $\chi$  v průběhu simulace pro komplex 1DP7. Zeleně jsou vyznačeny nukleotidy, které jsou ve výchozí struktuře v A/A konformaci.**

| OL15      |      |          |                         |     |     |     | bsc1     |                         |     |     |     |
|-----------|------|----------|-------------------------|-----|-----|-----|----------|-------------------------|-----|-----|-----|
| N         | NtC  | P0 forma | Procentuální zastoupení |     |     |     | P0 forma | Procentuální zastoupení |     |     |     |
|           |      |          | A/A                     | A/B | B/A | B/B |          | A/A                     | A/B | B/A | B/B |
| řetězec C |      |          |                         |     |     |     |          |                         |     |     |     |
| G2        | BA06 | B/A      | 0                       | 0   | 18  | 82  | B/A      | 0                       | 0   | 19  | 81  |
| T3        | BB01 | A/A      | 1                       | 11  | 15  | 73  | A/A      | 0                       | 7   | 17  | 75  |
| T4        | BB00 | B/A      | 1                       | 16  | 7   | 76  | B/A      | 0                       | 3   | 3   | 93  |
| A5        | BB00 | B/A      | 0                       | 4   | 5   | 91  | B/A      | 7                       | 13  | 15  | 66  |
| C6        | BA13 | B/A      | 0                       | 6   | 3   | 91  | B/A      | 1                       | 6   | 4   | 89  |
| C7        | AA00 | A/A      | 14                      | 15  | 7   | 64  | A/A      | 79                      | 17  | 2   | 1   |
| A8        | AA00 | A/A      | 82                      | 8   | 8   | 3   | A/A      | 91                      | 4   | 4   | 1   |
| T9        | AB05 | A/A      | 0                       | 3   | 5   | 92  | A/A      | 2                       | 2   | 9   | 87  |
| G10       | BB00 | B/A      | 0                       | 6   | 5   | 90  | B/A      | 1                       | 1   | 2   | 97  |
| G11       | BB00 | B/A      | 3                       | 9   | 17  | 72  | B/A      | 1                       | 2   | 9   | 89  |
| T12       | BB00 | B/A      | 0                       | 8   | 3   | 89  | B/A      | 0                       | 3   | 4   | 92  |
| A13       | BB00 | B/A      | 0                       | 3   | 2   | 95  | B/A      | 0                       | 2   | 4   | 94  |
| A14       | BB04 | B/A      | 0                       | 5   | 5   | 90  | B/A      | 0                       | 2   | 5   | 93  |
| C15       | BB00 | B/A      | 1                       | 13  | 12  | 74  | B/A      | 1                       | 10  | 16  | 74  |
| řetězec D |      |          |                         |     |     |     |          |                         |     |     |     |
| G2        | BA06 | B/A      | 0                       | 0   | 14  | 86  | B/A      | 0                       | 0   | 21  | 79  |
| T3        | BB01 | A/A      | 0                       | 13  | 15  | 72  | A/A      | 0                       | 6   | 16  | 79  |
| T4        | BB00 | B/A      | 1                       | 13  | 7   | 79  | B/A      | 0                       | 3   | 3   | 93  |
| A5        | BB00 | B/A      | 0                       | 3   | 5   | 92  | B/A      | 6                       | 12  | 15  | 67  |
| C6        | BA13 | B/A      | 0                       | 8   | 4   | 87  | B/A      | 1                       | 7   | 5   | 88  |
| C7        | AA00 | A/A      | 25                      | 30  | 15  | 30  | A/A      | 60                      | 26  | 5   | 9   |
| A8        | AA00 | A/A      | 21                      | 7   | 11  | 60  | A/A      | 92                      | 4   | 3   | 1   |
| T9        | AB05 | A/A      | 3                       | 26  | 9   | 62  | A/A      | 4                       | 3   | 10  | 82  |
| G10       | BB00 | B/A      | 1                       | 2   | 2   | 95  | B/A      | 0                       | 1   | 3   | 96  |
| G11       | BB00 | B/A      | 0                       | 3   | 7   | 90  | B/A      | 1                       | 1   | 10  | 88  |
| T12       | BB00 | B/A      | 1                       | 15  | 8   | 77  | B/A      | 0                       | 2   | 2   | 95  |
| A13       | BB00 | B/A      | 0                       | 2   | 3   | 96  | B/A      | 0                       | 1   | 3   | 96  |
| A14       | BB04 | B/A      | 0                       | 8   | 4   | 88  | B/A      | 0                       | 3   | 4   | 93  |
| C15       | BB00 | B/A      | 2                       | 20  | 13  | 65  | B/A      | 1                       | 11  | 24  | 64  |

## 5.7. Souhrnné výsledky

Hodnoty parametru P, zprůměrované přes všechny nukleotidy daného komplexu, jsou zapsány v tabulce 23. Pro porovnání jsou v tabulce výsledky pro obě použitá silová pole. Na první pohled je patrné, že všechny komplexy po většinu dobu simulace přetrvávaly v jižní konformaci, tedy převážně v C2'-endo a poté v C1'-exo. Je zde ovšem rozdíl mezi silovými poli, kdy zastoupení C2'-endo hodnot v poli bsc1 je vyšší, většinou přibližně o 10 % pro každý komplex, než v poli OL15. Pole OL15 tedy modeluje širší distribuci parametru P a ve větší míře umožňuje navštěvovat hodnoty v C1'-exo oblasti.

**Tabulka 23: Průměrné procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace silovými poli OL15 a bsc1 pro všechny nukleotidy daných komplexů.** V tabulce zaznamenáno rozdělení pouze dle parametru P. Sloupec #N značí počet započítaných nukleotidů z obou řetězců daného protein-DNA komplexu.

| OL15           |    |                         |          |         |          |         |          |         |         |
|----------------|----|-------------------------|----------|---------|----------|---------|----------|---------|---------|
| Komplex        | #N | Procentuální zastoupení |          |         |          |         |          |         |         |
|                |    | C2'-exo                 | C3'-endo | C4'-exo | O4'-endo | C1'-exo | C2'-endo | C3'-exo | Ostatní |
| 3PVI           | 22 | 0                       | 1        | 2       | 5        | 30      | 49       | 12      | 1       |
| 3KXT           | 12 | 1                       | 5        | 14      | 10       | 17      | 37       | 13      | 2       |
| 3AAF           | 14 | 0                       | 0        | 2       | 6        | 26      | 51       | 15      | 0       |
| 2VJV           | 48 | 4                       | 7        | 6       | 6        | 28      | 40       | 8       | 1       |
| 1DP7           | 28 | 0                       | 4        | 2       | 5        | 28      | 48       | 12      | 0       |
| Celkový průměr |    | 1                       | 4        | 5       | 6        | 26      | 45       | 12      | 1       |
| bsc1           |    |                         |          |         |          |         |          |         |         |
| 3PVI           | 22 | 0                       | 1        | 1       | 2        | 26      | 61       | 9       | 0       |
| 3KXT           | 12 | 1                       | 13       | 12      | 4        | 12      | 46       | 11      | 0       |
| 3AAF           | 14 | 0                       | 1        | 1       | 2        | 23      | 60       | 13      | 0       |
| 2VJV           | 48 | 3                       | 9        | 4       | 2        | 21      | 50       | 9       | 0       |
| 1DP7           | 28 | 1                       | 11       | 3       | 1        | 19      | 53       | 11      | 0       |
| Celkový průměr |    | 1                       | 7        | 4       | 2        | 20      | 54       | 11      | 0       |



V tabulce 24 jsou zaznamenány hodnoty procentuálního zastoupení konformací 2-deoxyribosy v průběhu simulace zprůměrované přes všechny nukleotidy daného protein-DNA komplexu, jejichž 2-deoxyribosa byla ve výchozí struktuře v C3'-endo konformaci. V případě komplexu 3PVI simulovaného v poli OL15 trávila 2-deoxyribosa v C2'-endo konformaci až polovinu času simulace daných nukleotidů a pouze 1 % v počáteční C3'-endo. Velmi zastoupená je zde i C1'-exo. 2-deoxyribosa má tedy tendenci přecházet spíše do jižních konformací typických pro B formu DNA. V poli bsc1 je tento trend podobný, zde tráví 2-deoxyribosa daných nukleotidů skoro dvě třetiny simulace v C2'-endo konformaci. V komplexu 3KXT se nenacházel žádný nukleotid s 2-deoxyribosou v C3'-endo konformaci v krystalografické struktuře. Jižní konformaci zaujímají i 2-deoxyribosy příslušných nukleotidů v komplexu 3AAF v poli OL15. Konformace C2'-endo a C3'-exo jsou téměř shodně zastoupeny. V poli bsc1 je výrazné navýšení zastoupení konformace C3'-endo na úkor C3'-exo. V komplexu 2VJV je zastoupení forem 2-deoxyribosy rozprostřené přes všechny konformace, ale nejvíce jsou zastoupeny C2'-endo a C1'-exo v poli OL15. V poli bsc1 je zastoupení konformací podobné, je zde akorát vyšší zastoupení C3'-endo oproti poli OL15. Obdobné chování vykazuje i komplex 1DP7, který má v poli OL15 rovnoměrné zastoupení všech konformací, zde ale převažují dvě konformace C3'-endo a C1'-exo. V poli bsc1 má dominantní zastoupení konformace C3'-endo.

**Tabulka 24: Průměrné procentuální zastoupení konformací 2-deoxyribosy v průběhu simulace silovými poli OL15 a bsc1 pro nukleotidy, které jsou ve výchozí struktuře v C3'-endo konformaci, pro všechny komplexy.** V tabulce zaznamenáno rozdělení pouze dle parametru P. V komplexu 3KXT ve výchozí struktuře nebyl ani jeden nukleotid s 2-deoxytibosou v C3'-endo konformaci. Sloupec #N značí počet započítaných nukleotidů z obou řetězců daného protein-DNA komplexu.

| <b>OL15</b>    |    |                         |              |             |              |             |              |             |         |
|----------------|----|-------------------------|--------------|-------------|--------------|-------------|--------------|-------------|---------|
| Komplex        | #N | Procentuální zastoupení |              |             |              |             |              |             |         |
|                |    | C2'<br>-exo             | C3'<br>-endo | C4'<br>-exo | O4'<br>-endo | C1'<br>-exo | C2'<br>-endo | C3'<br>-exo | Ostatní |
| 3PVI           | 17 | 0                       | 1            | 2           | 7            | 32          | 48           | 10          | 0       |
| 3KXT           | 0  | -                       | -            | -           | -            | -           | -            | -           | -       |
| 3AAF           | 2  | 0                       | 0            | 0           | 0            | 2           | 49           | 48          | 1       |
| 2VJV           | 15 | 10                      | 15           | 2           | 5            | 25          | 37           | 6           | 0       |
| 1DP7           | 4  | 3                       | 28           | 10          | 7            | 27          | 20           | 4           | 0       |
| Celkový průměr |    | 3                       | 11           | 4           | 5            | 22          | 38           | 17          | 0       |
| <b>bsc1</b>    |    |                         |              |             |              |             |              |             |         |
| 3PVI           | 17 | 0                       | 1            | 1           | 3            | 29          | 59           | 6           | 0       |
| 3KXT           | 0  | -                       | -            | -           | -            | -           | -            | -           | -       |
| 3AAF           | 2  | 0                       | 0            | 0           | 0            | 1           | 69           | 29          | 0       |
| 2VJV           | 16 | 10                      | 25           | 1           | 1            | 17          | 35           | 10          | 0       |
| 1DP7           | 4  | 6                       | 65           | 20          | 2            | 4           | 3            | 0           | 0       |
| Celkový průměr |    | 4                       | 23           | 6           | 1            | 13          | 42           | 11          | 0       |

Tabulka 25 znázorňuje průměrné hodnoty procentuálního zastoupení forem všech nukleotidů daných komplexů podle rozdělení dle úhlu  $\delta$  a  $\chi$ . Opět je zde vidět, že nukleotidy všech komplexů nejčastěji zaujímaly B/B konformaci. Při porovnání obou silových polí lze z tabulky pozorovat rozdílné preference k přechodným konformacím u jednotlivých polí. Konformace A/B má vyšší zastoupení v poli OL15, zatímco konformace B/A má větší zastoupení v poli bsc1.

**Tabulka 25: Průměrné procentuální zastoupení konformací nukleotidů v průběhu simulace silovými poli OL15 a bsc1 pro všechny nukleotidy.** V tabulce je zaznamenáno rozdělení podle úhlu  $\delta$  a  $\chi$ . Sloupec #N značí počet započítaných nukleotidů z obou řetězců daného protein-DNA komplexu.

| <b>OL15</b>                    |    |                         |     |     |     |
|--------------------------------|----|-------------------------|-----|-----|-----|
| Komplex                        | #N | Procentuální zastoupení |     |     |     |
|                                |    | A/A                     | A/B | B/A | B/B |
| 3PVI                           | 22 | 6                       | 5   | 8   | 80  |
| 3KXT                           | 12 | 4                       | 31  | 2   | 63  |
| 3AAF                           | 24 | 3                       | 9   | 3   | 85  |
| 2VJV                           | 48 | 12                      | 15  | 22  | 50  |
| 1DP7                           | 28 | 6                       | 9   | 8   | 77  |
| Celkem přes všechny nukleotidy |    | 6                       | 14  | 9   | 71  |
| <b>bsc1</b>                    |    |                         |     |     |     |
| 3PVI                           | 22 | 5                       | 2   | 18  | 75  |
| 3KXT                           | 12 | 9                       | 26  | 5   | 60  |
| 3AAF                           | 24 | 2                       | 5   | 7   | 85  |
| 2VJV                           | 48 | 13                      | 10  | 29  | 48  |
| 1DP7                           | 28 | 13                      | 6   | 8   | 74  |
| Celkem přes všechny nukleotidy |    | 8                       | 10  | 14  | 68  |

V tabulce 26 jsou pak zaznamenány průměrné hodnoty procentuálního zastoupení konformací nukleotidů, ale zde je to z hlediska úhlů  $\delta$  a  $\chi$  zároveň. V tabulce jsou započítány hodnoty nukleotidů, které začínaly v A/A konformaci. V komplexu 3PVI a 3AAF nukleotidy trávily většinu času v konformaci B/B v obou silových polích. Akorát v poli bsc1 je větší zastoupení přechodné B/A konformace. Nukleotidy komplexu 2VJV mají procentuální zastoupení konformací v průběhu simulace rovnoměrně rozložené mezi všechny čtyři konformace. Opět je zde patrný rozdíl mezi poli v přechodných stavech, kdy pole OL15 upřednostňuje konformaci A/B, zatímco pole bsc1 spíše konformaci B/A. V komplexu 1DP7 převažuje zastoupení B/B formy v poli OL15, v poli bsc1 je zastoupení konformací A/A a B/B téměř vyrovnané.

**Tabulka 26: Průměrné procentuální zastoupení konformací nukleotidů v průběhu simulace silovými poli OL15 a bsc1 pro nukleotidy, které jsou ve výchozí struktuře v A/A konformaci.** V tabulce je zaznamenáno rozdělení podle úhlu  $\delta$  a  $\chi$ . V komplexu 3KXT nebyl ve výchozí struktuře ani jeden nukleotid v konformaci A/A. Sloupec #N značí počet započítaných nukleotidů z obou řetězců daného protein-DNA komplexu.

| <b>OL15</b>                    |    |                         |     |     |     |
|--------------------------------|----|-------------------------|-----|-----|-----|
| Komplex                        | #N | Procentuální zastoupení |     |     |     |
|                                |    | A/A                     | A/B | B/A | B/B |
| 3PVI                           | 17 | 8                       | 5   | 9   | 78  |
| 3KXT                           | 0  | -                       | -   | -   | -   |
| 3AAF                           | 3  | 1                       | 4   | 2   | 93  |
| 2VJV                           | 20 | 25                      | 21  | 18  | 36  |
| 1DP7                           | 8  | 18                      | 14  | 11  | 57  |
| Celkem přes všechny nukleotidy |    | 10                      | 9   | 8   | 53  |
| <b>bsc1</b>                    |    |                         |     |     |     |
| 3PVI                           | 18 | 5                       | 2   | 21  | 71  |
| 3KXT                           | 0  | -                       | -   | -   | -   |
| 3AAF                           | 3  | 1                       | 1   | 16  | 82  |
| 2VJV                           | 20 | 28                      | 15  | 24  | 33  |
| 1DP7                           | 8  | 41                      | 9   | 8   | 42  |
| Celkem přes všechny nukleotidy |    | 15                      | 5   | 14  | 46  |

## 6. Diskuze

Aby bylo možné vyhodnotit MD simulace, je zapotřebí, aby simulované systémy zůstaly v průběhu simulace stabilní a nedocházelo v nich k výrazným nežádoucím konformačním změnám. Přestože v reálném systému může obecně k větším změnám docházet, rozsáhlé a rychlé změny konformací v simulacích na takto krátkých časových škálách by mohly ukazovat nestabilitu systému. Z toho by mohl vyplývat nefyzikální vývoj trajektorie v čase a nevhodnost jejích použití pro další analýzy. Pokud jsou MD simulace stabilní a konformační změny, které proběhly, jsou omezené, můžeme usuzovat zda mají tyto změny fyzikální a chemický význam, případně, zda jsou tyto změny v použitém silovém poli správně modelovány.

Stabilita simulací byla posuzována pomocí střední kvadratické odchylky (RMSD) poloh jednotlivých atomů komplexů v čase vůči původní krystalografické struktuře. Většina komplexů vykazuje fluktuaci RMSD kolem hodnoty 2 – 3 Å. V případě několika komplexů dochází k velmi mírnému nárůstu RMSD, což může být způsobeno přeměnou konformace DNA a s tím je spojen i pohyb proteinu. Nicméně v průběhu simulace nedochází k výrazným výchylkám RMSD, které by mohly naznačit velký pohyb částí komplexu případně až roztržení komplexu. Z toho vyplývá, že simulace daných komplexů jsou z globálního pohledu stabilní a jejich výsledky jsou relevantní.

V této diplomové práci se zaměřuji na stabilitu A forem (konformací) DNA v protein-DNA komplexech. Rozdělení DNA forem závisí především na parametru P, jenž popisuje konformaci 2-deoxyribosy, a na glykosidickém úhlu  $\chi$ . Zvolil jsem zde rozdělení podle úhlu  $\delta$ , který popisuje konformaci 2-deoxyribosy podobně jako parametr P, a jeho referenční hodnoty pro dané konformace jsou uvedeny v databázi dnatco.org, a úhlu  $\chi$  pouze na čtyři konformace A/A, A/B, B/A a B/B (viz tabulka 4). Toto rozdělení se může zdát zjednodušující vzhledem k tomu, jakých hodnot může úhel  $\delta$ , respektive parametr P, nabývat, a tedy i v jakých konformacích se může 2-deoxyriobsa nacházet. Nicméně toto rozdělení nám umožňuje zjednodušeně a přehledně sledovat stabilitu A formy (A/A) a její změnu na B (B/B) formu přes přechodné formy A/B a B/A, které mohou být způsobeny jen změnou jednoho parametru, a to jen P nebo jen změnou úhlu  $\chi$ .

Jednou z možností, jak posoudit stabilitu konformací a forem DNA v průběhu simulace, je porovnat je s původní krystalografickou strukturou daného komplexu. Pro obě silová pole platí ve všech mých simulacích podobný trend a z výše uvedených tabulek a grafů vyplývá, že většina nukleotidů, které se ve výchozí krystalografické struktuře nacházely v A/A konformaci, během několika desítek ns tuto konformaci ztratily a přešly do konformace kanonické B-DNA. Nicméně, nukleotidy C9 (řetězec C a D) v komplexu endonukleasy 3PVI, C20, C23 a C34 (řetězce C a D) v komplexu transposasy 2VJV a nukleotidy C7 (pouze řetězec D) a A8 (řetězce C a D) v komplexu regulačního faktoru X1 1DP7 si v poli OL15 uchovaly A/A konformaci po určitou část simulace. Většina nukleotidů komplexů, které měly v krystalografické struktuře přechodnou A/B nebo B/A konformaci, si ji částečně udržely v průběhu simulace, což lze vidět na nukleotidech G103, T105, A112 a T113 komplexu 3KXT.

V poli bsc1 si výše uvedené nukleotidy komplexů také zachovaly A/A konformaci po určitou dobu simulace s výjimkou několika nukleotidů. Jedním z nich je nukleotid C9 v komplexu 3PVI, který v bsc1 setrval polovinu simulace v B/A konformaci. Opačně to bylo v případě nukleotidu T37 komplexu 2VJV, který v poli OL15 většinu simulace zaujímal konformaci A/B (87 % v řetězci C a 82 % v řetězci D) a zbytek v konformaci A/A. V poli bsc1 se podíl konformace A/A zvýšil až na třetinu vůči konformaci A/B. Z těchto výsledků vyplývá, že podíl daných konformací ve stejných nukleotidech se může mezi simulacemi různými poli lišit až o desítky procent.

Při porovnávání polí mezi sebou lze vidět, že pole bsc1 v některých nukleotidech, které obsahují cytosin, upřednostňují severní konformaci 2-deoxyribosy. Příkladem mohou být nukleotidy C20 a C34 komplexu 2VJV, kdy v případě pole OL15 dosahuje parametr P hodnoty C3'-endo konformace v 60 % a 45 % simulace (řetězec C), a v poli bsc1 mají stejné nukleotidy procentuální zastoupení parametru P v C3'-endo konformaci 76 % a 68 %. Podobné je to i pro nukleotid C7 v komplexu 1DP7, kde je v případě pole OL15 hodnota zastoupení konformace C3'-endo 12 % a 13 % (tabulka 20) pro řetězce C a D a v poli bsc1 je to pak 61 % a 46 % (tabulka 21). To má za následek větší procentuální zastoupení i A/A konformace v průběhu celé simulace v komplexu 1DP7 v poli bsc1, 79 % a 60 %, oproti poli OL15, 14% a 25 %. Podobná preference silového pole pro určitou konformaci pro danou bázi nukleotidu již nebyla pro jiné nukleotidy pozorována. Zdá se tedy, že i když se obě silová pole liší parametrizací parametru

P 2-deoxyribosy, poskytují velmi podobné výsledky pro nukleotidy dA, dG, a dT, ve kterých je severní oblast (C3'-endo) málo stabilní, a výrazně se liší jen pro nukleotid dC, kde poskytuje pole bsc1 stabilnější popis severní konformace.

Dalším porovnáním polí zjistíme, že pole OL15 má oproti poli bsc1 pro 2-deoxyribosu vyšší procentuální zastoupení konformace C1'-exo na úkor konformace C2'-endo. V literatuře se zjednodušeně udává, že typická konformace cukerné složky pro B-DNA je C2'-endo, nicméně studiem B-DNA dodekameru (Drew a kol. 1981) se zjistilo, že deoxyribosa B-DNA není pouze v C2'-endo konformaci, ale má tendenci se vyskytovat i v C1'-exo konformaci. Z toho můžeme usoudit, že pole OL15 lépe popisuje B formu DNA, protože má oproti poli bsc1 má více zastoupenou C1'-exo konformaci v B formách DNA.

Nicméně porovnávání simulace vůči výchozí krystalografické struktuře tak, jak ho provádím v předešlém textu, je mírně zjednodušené. V krystalografické struktuře je totiž většinou zobrazena jen jedna, energeticky nejstabilnější konformace DNA. Pokud navíc může docházet ke konformačním změnám i v už vytvořeném krystalu, bude energeticky stabilnější konformace preferována ještě více, protože krystalografická měření se často provádějí za velmi nízkých teplot. Naproti tomu v roztoku mohou konformace vlivem vyšší teploty jednak snadněji přecházet mezi sebou, ale také může být nezanedbatelná i populace energeticky méně výhodných konformačních stavů. Určitý výskyt jiných než krystalograficky určených stavů v MD simulaci v roztoku není v rozporu s experimentem, pokud není příliš vysoký.

Dalším faktorem, na kterém závisí konformace DNA v roztoku, je aktivita vody. Změnou aktivity vody v roztocích a krystalech mohou konformace DNA přecházet mezi sebou. V roztoku tak nastává rovnováha mezi různými konformacemi DNA, která může být v krystalu posunuta. B-DNA se vyskytuje v prostředí s vysokou aktivitou vody, pokud ale dojde ke snížení aktivity vody, DNA změní svou konformaci na A formu (Knee a kol. 2008). Tento jev je známý z krystalů samotné DNA, ale v krystalech protein-DNA komplexů je pozorován méně často. Ty části DNA, které nejsou v kontaktu s proteinem, se vyskytují převážně v B formě, což může být způsobeno větším podílem vody v těchto krystalech v porovnání s krystaly samostatné DNA. Většina nukleotidů, které

se v protein-DNA komplexech nacházejí v A formě, je v kontaktu s proteinem a A forma je tímto kontaktem podmíněná.

Dále, i přestože v PDB souboru z krystalografické analýzy je většinou pouze jedna konformace DNA, může se stát, že v reálném krystalu byly přítomny obě konformace. Poté záleží na rozlišení krystalografické struktury a schopnosti posouzení daných výsledků, zda budou správně interpretovány obě konformace nebo pouze jedna. Výsledky krystalografické analýzy tedy nemusí vždy správně popisovat daný krystal, zvláště, když v něm mohou koexistovat obě formy DNA, podobně jako v roztoku.

V této práci byly pro posouzení stability A konformace nukleotidů DNA během simulace vybrány protein-DNA komplexy, protože protein svou interakcí s DNA indukuje změnu B formy na A formu (Lu a kol. 2000). Předpoklad byl takový, že právě protein-DNA komplexy by si mohly udržet A konformaci po většinu času simulace. Porovnáním výsledků zastoupení konformací 2-deoxyribosy v protein-DNA komplexech (tabulka 23 a 24) s výsledky z práce Zgarbová a kol. (2018) (tabulka č. 3), ve které byla zjišťována stabilita severní konformace 2-deoxyribosy v samotné DNA (bez proteinu), zjistíme, že protein opravdu indukuje A konformaci v DNA. Průměrné hodnoty zastoupení C3'-endo konformace jsou v protein-DNA komplexech 11 % pro pole OL15 a 23 % pro bsc1, ale pro samostatné řetězce DNA, které krystalizují převážně v A formě, jsou průměrné hodnoty severní konformace pouze 2 % pro obě silová pole.

Výsledky této práce ukázaly, že A forma nebyla v průběhu simulace stabilní a přecházela na B formu, případně přechodné formy. Výjimkou jsou určité nukleotidy, které se v A konformaci vyskytovaly většinu času simulace. V roztoku sice dochází k přechodům mezi formami, jak je popsáno výše, ale v simulacích docházelo k těmto změnám během několika desítek nanosekund od začátku simulace a výchozí simulace pak již nebyla populována. to naznačuje, že stabilita A formy je v MD simulacích podhodnocena. Z toho vyplývá, že silová pole, jak OL15, tak bsc1, nepopisují správně A formu DNA a bylo by vhodné navrhnout korekci, která by umožnila větší stabilitu A formy DNA. Otázkou zůstává, pro který z parametrů by měla být udělána korekce, aby simulace dosahovala lepších výsledků pro popis A formy. Dle výsledků úhel  $\chi$  vykazoval větší stabilitu v oblasti *anti* (typická pro A-DNA) v nukleotidech, které v krystalografické struktuře byly v A formě. Naopak parametr P popisující konformaci 2-deoxyribosy nebyl



v příslušných nukleotidech stabilní. Z tohoto důvodu bych navrhoval se v další korekci pro silová pole zaměřit na popis parametrů P.

## 7. Závěr

Tato diplomová práce se zaměřuje na stabilitu A formy nukleové kyseliny v protein-DNA komplexů, které zkoumá MD simulací. Pět protein-DNA komplexů, vybraných tak, aby obsahovaly dobře definované strukturní motivy A-DNA formy, bylo simulováno pomocí programu AMBER ve dvou různých silových polích OL15 a bsc1.

Všechny protein-DNA komplexy byly stabilní po celou dobu simulace a nedocházelo k výrazným konformačním změnám, pro které by nebylo možné najít fyzikální nebo chemický význam. Stabilita simulovaných komplexů byla zjišťována na základě porovnání RMSD komplexu v průběhu simulace vůči výchozí krystalografické struktuře a dále vizuální kontrolou trajektorie simulace.

Z výsledků simulací vyplývá, že A forma DNA je v průběhu simulací nestabilní skoro ve všech nukleotidech, které byly v A formě ve výchozí (krystalografické) struktuře. Přestože v roztoku obecně může docházet ke změnám konformací nukleových kyselin, v našem případě jsme tyto změny pozorovali v prvních desítkách nanosekund od začátku simulace a často i v prvních krocích simulace. To naznačuje nesprávný popis A formy nukleových kyselin silovými poli, kdy je A geometrie pravděpodobně modelována s vyšší energií, a tedy méně stabilní, než jak je tomu ve skutečnosti. Ve výsledcích byly ale některé nukleotidy, které si v simulaci udržely A formu, případně přechodnou konformaci mezi A a B, která se v komplexech také může vyskytovat, po delší dobu. To naznačuje, že protein je v některých případech schopen svou interakcí s DNA A formu stabilizovat téměř tak silně, jak je v silovém poli destabilizována.

Při porovnávání výsledků silových polí OL15 a bsc1 byly pozorovány určité rozdíly. Velké rozdíly nastaly zvláště v případě simulace nukleotidů s bází cytosin. V některých případech pole bsc1 upřednostňovalo právě v těchto nukleotidech konformaci A/A oproti ostatním konformacím a rozdíly vůči výsledkům pole OL15 byly až několik desítek procent. Lze tedy říci, že pole bsc1 lépe popisuje A formu DNA u nukleotidů, které obsahují cytosin. Další rozdíl mezi silovými poli nastal při popisu B formy DNA, při které se 2-deoxyribosa může nacházet nejen v konformaci C2'-endo, ale i v C1'-exo. Tomuto faktu lépe odpovídá popis nukleotidů pomocí silového pole OL15, protože v případě simulace silovým polem OL15 se 2-deoxyribosa nacházela

v konformaci C1'-exo častěji než v případě bsc1. Z toho vyplývá, že pole OL15 popisuje B formu DNA lépe.

Závěrem této diplomové práce je, že v nukleotidech, jež měly počáteční formu v A/A konformaci, se tato forma ve většině případu v MD simulaci neudržela a byla nahrazena kanonickou B-DNA. To vedlo k odchylkám geometrie komplexů od výchozích krystalografických struktur a dá se očekávat, že je to spojeno i s energetickou destabilizací studovaných komplexů v MD simulacích. Zdá se tedy, že A-DNA je pomocí současných silových polí OL15 a bsc1 popsána nepřesně. Tato nepřesnost může mít vliv na věrohodnost a prediktivní hodnotu MD simulací systémů obsahujících komplexy proteinů s DNA. Silová pole OL15 a bsc1 tedy vyžadují korekci (reparametrizaci) parametrů popisujících pseudorotaci P deoxyribosového cyklu, konkrétně zvýšení stabilizace C3'-endo (severní) konformace této molekuly. Toto zjištění může pomoci zlepšit popis nukleových kyselin v příštích verzích silových polí z rodiny AMBER.

## 8. Seznam literatury

- Barabas O., Ronning D. R., Guynet C., Hickman A.B., Ton-Haong B., Chandler M., Dyda F. (2008) Mechanism of IS200/IS605 Family DNA Transposases: Activation and Transposon-Directed Target Site Selection. *Cell* **132**:2, 208-220.
- Bruce A., Johnson A., Lewis J., Raff M.C., Roberts K., Walter P., Wolson J.H., Hunt T. (2008) *Molecular biology of the cell*. 5. vyd., Garland Science, New York, U.S.A, 1392 stran.
- Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz M. Jr., Ferguson D. M., Spellmeyer D. C., For T. Caldwell J.W., Kollman P. A (1995) 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.
- Černý J., Božíková P., Svoboda J., Schneider B. (2020) A unified dinucleotide alphabet describing both RNA and DNA structures. *Nucleic Acids Research* **48**:11, 6367-6381.
- Dickerson R.E., Drew H.R., Conner B.N., Wing R.M., Frantini A. V., Kopka M. L. (1982) The Anatomy of A-, B-, and Z-DNA. *Science* **216**:4545, 475-485.
- Dickerson R.E. (1989) Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Research* **17**:5: 1797-1803.
- Dickerson R.E. (1992) [5] DNA structure from A to Z. In: *Methods in Enzymology*, Vol 211, Academic Press, Massachusetts, U.S.A, 67-111.
- Drew H.R., Wing R.M., Takano T., Broka Ch., Tanaka S., Itakura K., Dickerson R. E. (1981) Structure of a B-DNA dodecamer: conformation and dynamics. *Proceedings of the National Academy of Sciences* **78**:4. 2179-2183.
- El Hassan M., Calladine C. (1998) Two distinct modes of protein-induced bending in DNA. *Journal of Molecular Biology* **282**:2, 331-343.
- Feng Y., Yao H., Wang J. (2010) Crystal structure of the crenarchaeal conserved chromatin protein Cren7 and double-stranded DNA complex. *Protein Science* **19**:6, 1253-1257.
- Gajiwala K.S., Chen H. Cornille F., Roques B.P., Reith W., Mach B., Burley S. K. (2000) Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature* **403**:6772, 916-921.
- Garvie C.W., Wolberger C. (2001) Recognition of Specific DNA Sequences. *Molecular Cell* **8**:5, 937-946.
- González M.A. (2011) Force fields and molecular dynamics simulations. *École thématique de la Société Française de la Neutronique* **12**, 169-200.
- Groenhof G. (2013) Introduction to QM/MM Simulations. In: *Biomolecular Simulations. Methods in Molecular Biology*, vol 924, Humana Press, Totowa, U.S.A., 43-66.
- Harrison S.C. (1991) A structural taxonomy of DNA-binding domains. *Nature* **353**:6346, 715-719.
- Horton J.R., Nastro H.G., Riggs P.D., Cheng X. (1998) Asp34 of Pvu II endonuclease is directly involved in DNA minor groove recognition and indirectly involved in catalysis. *Journal of Molecular Biology* **284**:5. 1491-1504.
- Cheatham T.E., Cieplak P., Kollman P.A. (1999) A Modified Version of the Cornell et al. Force Field with Improved Sugar Pucker Phases and Helical Repeat. *Journal of Biomolecular Structure and Dynamics* **16**:4, 845-862.
- Ivani I., Dans P. D., Noy A., Pérez A., Faustino I., Hospital A., Walther J., Andrio P., Goñi R., Balaceanu A., Portella G., Battistini F., Gelpi J.L., González C., Vendruscolo M., Laughton Ch.A., Harris S.A., Case D.A., Orozco M. (2016) Parmbsc1: a refined force field for DNA simulations. *Nature Methods* **13**:1, 55-58.
- Kitano K., Kim S.-Y., Hakoshima T. (2010) Structural Basis for DNA Strand Separation by the Unconventional Winged-Helix Domain of RecQ Helicase WRN. *Structure* **18**:2, 177-187.
- Knee K.M., Dixit S. B., Aitken C. E. Ponomarev S., Beveridge D. L., Mukerji I. (2008) Spectroscopic and Molecular Dynamics Evidence for a Sequential Mechanism for the A-to-B Transition in DNA. *Biophysical Journal* **95**:1, 257-272.

- Koolman J., Röhm K. -H. (2012) Barevný atlas biochemie. Grada, 4. vydání, Praha, Česká republika, 498 stran.
- Krepl M., Zgarbová M., Stadlbauer P., Otyepka M., Banáš P., Koča J., Cheatham T.E., Jurečka P., Šponer J. (2012) Reference Simulations of Noncanonical Nucleic Acids with Different  $\chi$  Variants of the AMBER Force Field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *Journal of Chemical Theory and Computation* **8**:7, 2506-2520.
- Leach A.R. (2001) Molecular Modelling: Principles and Applications. Pearson Education, London, U.K., 744 stran.
- Li L., Szostak J.W. (2014) The Free Energy Landscape of Pseudorotation in 3'-5' and 2'-5' Linked Nucleic Acids. *Journal of the American Chemical Society* **136**:7, 2858-2865.
- Lu X.-J., Shakked Z., Olson W.K. (2000) A-form Conformational Motifs in Ligand-bound DNA Structures. *Journal of Molecular Biology* **300**:4, 819-840.
- Lu X.-J., Olson W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research* **31**:17, 5108-5121.
- Neidle S. (2008) In Principles of Nucleic Acid Structure, 1. vydání, Elsevier, London, U.K., 289 stran.
- Peréz A., Marchán I., Svozil D., Sponer, J. Cheatham T.E., Laughton Ch. A., Orozco M. (2007) Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of  $\alpha/\gamma$  Conformers. *Biophysical Journal* **92**:11, 3817-3829.
- Poltev V. (2015) Molecular Mechanics: Principles, History, and Current Status. In *Handbook of Computational Chemistry*, Springer, Dordrecht, Netherlands, 1-48.
- Ronning D.R., Guynet C., Ton-Hoang B., Perez Z.N., Ghirlando R., Chandler M., Dyda F. (2005) Active Site Sharing and Subterminal Hairpin Recognition in a New Class of DNA Transposases. *Molecular Cell* **20**:1, 143-154.
- Sinden R.R. (1994) DNA Structure and Function. Elsevier. London, U.K., 398 stran.
- Tuckerman M.E., Martyna G.J. (2000) Understanding Modern Molecular Dynamics: Techniques and Applications. *The Journal of Physical Chemistry B* **104**:2, 159-78.
- Wang J., Cieplak P., Kollman P.A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules. *Journal of Computational Chemistry* **21**:12, 1049-1074.
- Weiner S.J. Kollman A, Case D.A., Singh C., Ghio C., Alagona G., Profeta S., Weiner P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society* **106**:3, 765-784.
- Weiner P.K., Kollman P.A. (1981) AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *Journal of Computational Chemistry* **2**:3, 287-303.
- Zgarbová M. Otyepka M. Šponer J., Mládek A. Banáš P., Cheatham T. E., Jurečka P. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *Journal of Chemical Theory and Computation* **7**:9,2886-2902.
- Zgarbová M., Luque F. J., Šponer J., Cheatham T.E., Otyepka M., Jurečka P (2013) Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *Journal of Chemical Theory and Computation* **9**:5, 2339-2354.
- Zgarbová M., Šponer J., Otyepka M., Cheatham T.E., Galindo-Murillo R., Jurečka P. (2015) Refinement of the Sugar-Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *Journal of Chemical Theory and Computation* **11**:12, 5723-5736.
- Zgarbová M., Jurečka P, Šponer J., Otyepka M. (2018) A- to B-DNA Transition in AMBER Force Fields and Its Coupling to Sugar Pucker. *Journal of Chemical Theory and Computation* **14**:1, 319-328.