

**Univerzita Karlova v Praze, Přírodovědecká fakulta
Katedra fyzikální a makromolekulární chemie**

**Charles University in Prague, Faculty of Science
Department of Physical and Macromolecular Chemistry**

Modelování chemických vlastností nano- a biostruktur
Modelling of Chemical Properties of Nano- and Biostructures

Autoreferát disertační práce
Summary of the Ph.D. Thesis



Modelování mechanických vlastností RNA a DNA
Modeling mechanical properties of RNA and DNA

Tomáš Dršata

Školitel/Supervisor: Ing. Filip Lankaš, Ph.D.

Praha, 2016

Abstrakt

Strukturní a mechanické vlastnosti nukleových kyselin hrají klíčovou roli v řadě biologických procesů i v oblasti nanotechnologií. Práce předkládá výsledky několika studií zaměřených na modelování těchto vlastností. Rozsáhlé simulace atomistické molekulové dynamiky (MD) jsou použity ke studiu strukturní dynamiky nukleových kyselin a k parametrisaci modelů jejich mechaniky. Používáme dva modely, které předpokládají deformační energii ve tvaru obecné kvadratické funkce vhodně zvolených vnitřních souřadnic, a liší se v úrovni rozlišení. První model je založen na popisu konformace na úrovni jednotlivých bází, zatímco druhý, hrubší model, je vhodný k popisu tuhosti vůči globálním deformacím v ohybu a celkové torzi. Tyto modely jsou využity ke studiu mechanických vlastností A-traktů v souvislosti s tvorbou smyček a lokalizací nukleosomů, k charakterizaci spřažených deformací v torzi a prodloužení helikální DNA a RNA, a k predikci změn ve vlastnostech poškozené DNA v souvislosti s procesem rozpoznání a oprav tohoto poškození. Kromě toho je představen návrh obecného modelu alosterických efektů v DNA, který je aplikován k predikci změn ve struktuře DNA způsobených vazbou ligandů do malého žlábku a ke studiu mechanismu alosterické regulace vazby dvou proteinů na DNA. Pečlivé srovnání našich výsledků s experimentálně dostupnými daty demonstreuje všeobecnou spolehlivost použitého postupu a současných silových polí určených pro simulace nukleových kyselin. Získané výsledky by měly pomoci lépe pochopit vliv mechanických vlastností nukleových kyselin na jejich biologickou funkci a najít tak aplikaci např. při návrhu nových léčiv. Celkem tato práce vedla k 11 publikacím v impaktovaných zahraničních časopisech.

Abstract

Structural and mechanical properties of nucleic acids play a key role in a wide range of biological processes, as well as in the field of nucleic acid nanotechnology. The thesis presents results of several studies focused on modelling these properties. Extensive unrestrained atomic-resolution molecular dynamics (MD) simulations are used to investigate structural dynamics of nucleic acids, and to parametrize their mechanical models. The deformation energy is assumed to be a general quadratic function of suitably chosen internal coordinates. Two types of models are employed which differ in the level of coarse-graining. The first one is based on the description of conformation at the level of individual bases and the second, coarser one is used to study global bending and twisting flexibility. The models are applied to explain mechanical properties of A-tracts in the context of DNA looping and nucleosome positioning, to characterize twist-stretch coupled deformations in DNA and RNA, and to predict changes in the properties of damaged DNA that are likely to be relevant for damage recognition and repair. Besides that, we propose a general model of DNA allostery, applied to study the effect of minor groove binding of small ligands and the allosteric coupling between proteins mediated by the DNA. A careful comparison of our results with available experimental data demonstrates the general reliability of contemporary nucleic acid force fields and the adopted modeling approach. Our results should help to better understand the role of mechanical properties of nucleic acids in their biological function with potential application in the development of new drugs. The work presented in the thesis resulted in 11 publications in impacted international journals.

Obsah/Contents

I	Modelování mechanických vlastností RNA a DNA	1
1	Úvod	1
2	Cíle práce	2
3	Materiál a metodika	3
4	Výsledky a diskuze	5
5	Závěry	12
II	Modeling mechanical properties of RNA and DNA	13
1	Introduction	13
2	Aims of the study	14
3	Material and methods	14
4	Results and discussion	17
5	Conclusions	24
III	Přílohy/Appendices	25
	Použitá literatura/References	25
	Seznam publikací/Selected publications	29
	Curriculum vitae	31

Modelování mechanických vlastností RNA a DNA

1 Úvod

DNA a RNA jsou klíčové biomolekuly [1]. DNA uchovává genetickou informaci, která se skládá především z instrukcí pro syntézu proteinů, biomolekul zajišťujících většinu buněčných funkcí. Tato informace je uložena v lineární sekvenci bazí v kódujících částech DNA. Ukazuje se však, že kromě těchto instrukcí jsou v DNA uloženy i další informace. Konkrétní sekvence bazí určuje různé strukturní preference DNA, její mechanické vlastnosti a další fyzikálně-chemické charakteristiky, které mají významný vliv na její biologickou aktivitu [2, 3]. Tyto vlastnosti diktují usporádání DNA v buňce a hrají roli při procesech jako je rozpoznávání DNA proteiny, které jsou dále zásadní pro regulaci exprese genů, transkripcí a replikaci DNA. V případě helikální RNA mají její strukturní a mechanické vlastnosti vliv na interakci s proteiny, určují strukturu virového RNA kapsidu a hrají roli při procesech jako je syntéza proteinů a RNA interference [4].

Existuje mnoho experimentálních technik, které jsou používány ke studiu strukturních a mechanických vlastností DNA. Metody jako je rentgenová (X-ray) krystalografie a nukleární magnetická rezonance (NMR) poskytují informace o struktuře nukleových kyselin v krystalu a v roztoku, a to ve vysokém rozlišení. Jiné metody, jako jsou cyklizační experimenty nebo mikromanipulace pomocí optických či magnetických pinzet, mohou být využity ke studiu tuhosti biopolymeru vůči ohybu, natažení a torzi kolem jeho osy [5, 6]. Metody přímo poskytující informaci o flexibilitě však dosahují pouze nízkého rozlišení a trpí dalšími omezeními. Flexibilitu nukleových kyselin je možné určit také ze znalosti hustoty pravděpodobnosti jejich konformací [7]. Strukturních dat o vysokém rozlišení, dostupných především díky X-ray krystalografii, je ale často příliš málo v porovnání s velikostí prostoru možných konformací.

Další možností je využít ke studiu strukturních a mechanických vlastností nukleových kyselin výpočetních metod. Ukazuje se, že simulace molekulové dynamiky (MD) s explicitní reprezentací vody a iontů jsou schopné úspěšně za-

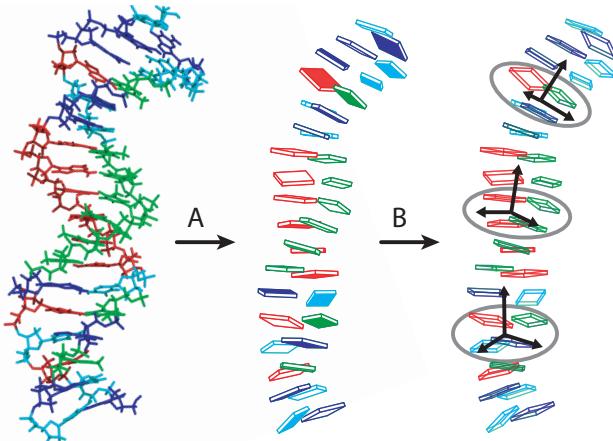
chytit množství důležitých vlastností velkých biologicky relevantních systémů [8, 9]. MD simulaci můžeme chápat jako počítačový experiment, který umožnuje měřit pozice všech atomů v systému v závislosti na čase, při daných vnějších podmínkách jako je teplota a tlak [10, 11]. Časové řady atomových kartézkých souřadnic získané z MD simulace pak mohou být převedeny na časové řady vhodně zvolených vnitřních souřadnic zachycujících konformaci DNA a RNA v hrubším měřítku (coarse-graining). Mezi ně patří např. souřadnice popisující vzájemnou polohu a orientaci jednotlivých bází. Znalost závislosti těchto souřadnic na čase může být dále využita k nalezení parametrů různých modelů mechaniky nukleových kyselin. MD simulace popisují interakce v daném systému s nevyhnutelným zjednodušením. Je tedy vždy nezbytné konfrontovat získané výsledky s experimentálně dostupnými daty, případně z výsledky přesných kvantově-mechanických (QM) výpočtů. Mezi nejdůležitější omezení MD simulací patří dostupné časové škály a výrazné aproximace v popisu interakcí mezi atomy.

2 Cíle práce

Cílem této práce je prostudovat strukturu, dynamiku a mechanické vlastnosti různých oligomerů nukleových kyselin v kontextu biologicky relevantních procesů za pomocí metod počítačového modelování. Jako primární zdroj dat slouží rozsáhlé simulace molekulové dynamiky na úrovni jednotlivých atomů s explicitní reprezentací vody a iontů. Tyto simulace jsou použity ke studiu konformační dynamiky helixů a pro parametrisaci modelů mechaniky nukleových kyselin. Používané modely předpokládají, že deformační energie je nelokální harmonickou funkcí vnitřních souřadnic. Konkrétně je tento postup aplikován na studium mechanických vlastností poly(dA:dT) sekvencí (A-traktů) v souvislosti s tvorbou smyček a lokalizací nukleosomů, dále na studium spřažených deformací v torzi a prodloužení v DNA a RNA a na studium vlastností poškozené DNA v souvislosti s procesem rozpoznání a oprav tohoto poškození. Získané výsledky jsou vždy pečlivě srovnány s dostupnými experimentálními daty a testovány na konvergenci. Dalším cílem této práce je vývoj modelu umožňujícího studium alosterických jevů v DNA. Model umožňuje postihnout změny struktury a mechanických vlastností DNA způsobené omezením některých stupňů volnosti, čímž je napodobován účinek vazebné interakce s ligandem. Tento model je následně použit k predikci změn ve struktuře DNA způsobených vazbou ligandů do malého žlábku a k vysvětlení mechanismu alosterické regulace vazby dvou proteinů na DNA.

3 Materiál a metodika

Mechanické vlastnosti nukleových kyselin jsou v této práci studovány pomocí dvou modelů lišících se úrovní rozlišení. V případě prvního modelu (model tuhých bází) je konformace helixu nukleové kyseliny určena souborem lokálních vnitřních souřadnic, které popisují vzájemnou polohu a orientaci jednotlivých bází, které se považují za tuhá tělesa (viz. obr. 1A). V rámci druhého modelu (model tuhosti v ohybu a torzi) je konformace oligomeru určena pouze trojicí globálních vnitřních souřadnic: velikost ohybu ve směru malého a velkého žlábků, velikost ohybu ve směru cukr-fosfátové páteře a celková torze (viz. obr. 1B). V případě obou modelů se předpokládá, že de-



Obrázek 1: Popis konformace DNA/RNA na dvou různých úrovních rozlišení. (A) Báze v DNA a RNA mohou být v dobrém přiblžení považovány za tuhá tělesa, se kterými je spojen pravotočivý souřadný systém. Lokální konformace molekuly je určena souborem souřadnic, které jsou definovány jako vzájemná posunutí a otočení mezi těmito souřadnými systémy. (B) Podobně může být globální konformace nukleových kyselin, tedy např. globální ohyb, popsána pomocí souřadních systémů reprezentujících vhodně zvolené fragmenty. Tyto body a souřadné systémy jsou získány průměrováním referenční bodů a souřadních systémů vázáných na jednotlivé báze ve zvoleném fragmentu.

formační energie, tj. volná energie spojená s určitou deformací oligomeru, je obecná kvadratická funkce příslušných vnitřních souřadnic,

$$E(\mathbf{w}) = \frac{1}{2} (\mathbf{w} - \hat{\mathbf{w}}) \cdot \mathbf{K} (\mathbf{w} - \hat{\mathbf{w}}), \quad (1)$$

kde \mathbf{w} je vektor vnitřních souřadnic, $\hat{\mathbf{w}}$ jsou jejich rovnovážné hodnoty a K je symetrická, pozitivně-definitní matici tuhostních konstant. Je-li systém je v kontaktu s tepelnou lázní, plyne z teorie termodynamických fluktuací [12], že \mathbf{w} má vícerozměrné Gausovo rozdělení. Snadno pak můžeme nalézt vztahy mezi parametry modelu $\hat{\mathbf{w}}$ a K a momenty pravděpodobnostního rozdělení \mathbf{w} ,

$$\hat{\mathbf{w}} = \langle \mathbf{w} \rangle, \quad K = k_B T C^{-1}, \quad (2)$$

kde $\langle \cdot \rangle$ značí očekávanou hodnotu a C je kovarianční matice \mathbf{w} . Průměrování přes statistický soubor (rov. 2) pak nahrazujeme průměrováním přes časové řady souřadnic, které získáme z atomistické MD simulace studovaného oligomeru. Známe-li parametry $\hat{\mathbf{w}}$ a K pro studovaný systém, může být model použit k odhadu energie spojené s určitou deformací předepsanou souřadnicemi \mathbf{w} s použitím rov. (1), za předpokladu, že deformace nejsou příliš velké a jsou tedy v harmonickém režimu. Bylo ověřeno, že pro systémy a typy deformací studované v této práci je harmonická approximace dostatečně dobře splněna. Marginální rozdělení některých souřadnic vykazují mírné odchylky od harmonicity způsobené konformačními substavy v páteři. Tyto odchylky jsou zanedbány, pozorované rozdělení vnitřních souřadnic je tedy nahrazeno Gausovským rozdělením se stejnými momenty, což vede na efektivní harmonický popis. Mechanické modely nukleových kyselin uvažující harmonickou deformační energii se už dříve ukázaly být užitečné při studiu mnoha procesů [7].

Pro účely modelování alosterických efektů v DNA je model tuhých bází nejdříve rozšířen tak, aby mezi vnitřní souřadnice zahrnul šířky velkého a malého žlábků, což jsou důležité parametry při interakci DNA s ligandy. Souřadnice můžeme rozdělit do dvou podmnožin \mathbf{w}_1 a \mathbf{w}_2 . Abychom napodobili účinek vazby ligandu na DNA, jsou souřadnicím \mathbf{w}_1 předepsány fixní hodnoty $\bar{\mathbf{w}}_1$, diktované vazbou ligandu (např. ligand vázající se do žlásku fixuje jeho šířku). Zbylé souřadnice jsou ponechány, aby volně relaxovaly. V minimu energie pak tyto souřadnice nabývají hodnot $\bar{\mathbf{w}}_2$. Minimalizace kvadratické deformační energie vzhledem k \mathbf{w}_2 , konzistentní s předepsanou vazbou, poskytuje vztah mezi touto vazbou a jejím efektem,

$$\bar{\mathbf{w}}_2 = \hat{\mathbf{w}}_2 - K_{22}^{-1} K_{12}^T (\bar{\mathbf{w}}_1 - \hat{\mathbf{w}}_1), \quad (3)$$

a předpis pro odpovídající deformační energii

$$E = 1/2 (\bar{\mathbf{w}}_1 - \hat{\mathbf{w}}_1) \cdot [K_{11} - K_{12} K_{22}^{-1} K_{12}^T] (\bar{\mathbf{w}}_1 - \hat{\mathbf{w}}_1), \quad (4)$$

kde K_{12} a K_{22} jsou submatice K odpovídající daným podmnožinám vnitřních souřadnic. Pro účely modelování efektu alosterické regulace vazby dvou ligandů na DNA (např. dvou proteinů) je možné model dále rozšířit tak, že

jsou vnitřní souřadnice rozděleny do třech podmnožin. První dvě podmnožiny jsou fixovány, címž je napodobena vazba prvního a druhého ligandu. Zbylé souřadnice jsou znova ponechány, aby volně relaxovaly. Minimalizace kvadratické deformační energie pak poskytuje vztahy analogické k (3) a (4). Efekt alosterické regulace je definovaný rozdílem ΔE mezi deformační energií spojenou s vazbou druhého ligandu v přítomnosti prvního ligandu a při jeho absenci.

MD simulace byly provedeny s použitím moderního silového pole typu AMBER určeného pro nukleové kyseliny (*parm99* s *bsc0* korekcí, a s χ_{OL3} korekcí pro RNA). Simulační protokol používaný pro simulace prováděné v naší laboratoři je založen na dobře zavedeném protokolu ABC konsorcia [13]. Simulovaný systém je uzavřen v boxu oktaedrálního tvaru s periodickými okrajovými podmínkami. Systém obsahuje studovaný oligomer s počáteční konformací B-DNA (příp. A-RNA), který je obklopen molekulami vody (SPC/E model [14]). Záporný náboj nukleové kyseliny je neutralizován přidáním odpovídajícího množství iontů K^+ a další ionty jsou přidány tak, aby bylo dosaženo koncentrace 150 mM KCl. Systém je následně důkladně zekvilibrován. Ve fázi produkce trajektorie je systém simulován za konstantní teploty (300 K) a tlaku (1 atm). Jsou aplikovány oktaedrální periodické okrajové podmínky s využitím metody PME pro výpočet interakcí dalekého dosahu. Vazby obsahující vodíkové atomy jsou fixovány pomocí algoritmu SHAKE a použitý časový krok simulace je 2 fs.

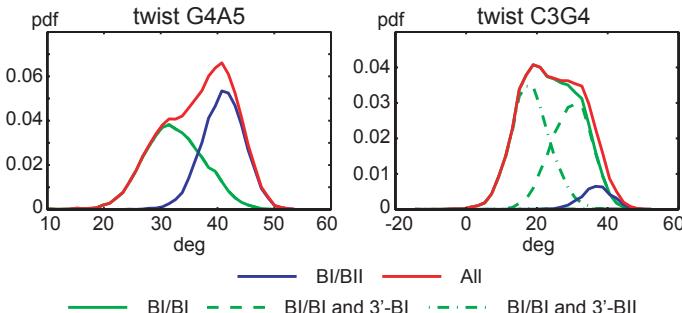
Souřadnice jednotlivých atomů jsou ukládány každých 10 ps simulace. Tímto způsobem je získána množina konformací oligomeru, která odpovídá náhodnému výběru ze statistického souboru NpT, a která je použita pro následující analýzu. Pro každou z konformací jsou vypočteny vnitřní souřadnice modelů mechaniky. Parametry modelů w a K jsou pak jednoduchými funkcemi momentů pravděpodobnostního rozdělení těchto souřadnic (viz. rov. 2).

4 Výsledky a diskuze

4.1 Jemná struktura a substavy v DNA

MD simulace nabízí cenné informace o vlastnostech DNA. Vždy je však třeba ověřovat data získaná z MD simulací při použití určitého silového pole a simulačního protokolu s daty dostupnými z experimentu a také testovat měřené veličiny na konvergenci. Tato analýza umožňuje vytyčit hranice použitelnosti MD simulací a pomáhá při optimalizaci silových polí a simulačních protokolů. My jsme se zaměřili na analýzu jemné struktury a dynamiky B-DNA určené z

rozsáhlých MD simulací a její srovnání s experimentálně dostupnými daty. Pro analýzu byla vybrána konkrétní sekvence, tzv. Dickersonův dodekamer (DD) d-[CGCGAATTCGCG]₂, jehož vlastnosti byly široce studovány množstvím různých experimentálních metod [15].



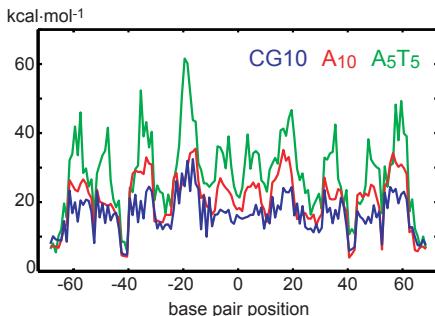
Obrázek 2: Anharmonické rozdělení twistu v G4A5 stepu může být rozloženo na téměř harmonické příspěvky odpovídající BI/BI a BI/BII stavům v daném stepu (vlevo). Rozdělení twistu v C3G4 stepu, odpovídající BI/BI stavu, je stále anharmonické, ale může být dále rozloženo na harmonické příspěvky v závislosti na BI nebo BII stavu G4-A5 fragmentu páteře v 3' sousedním stepu (vpravo). BII/BII stavů se téměř nevyskytuje a přispívají tedy zanedbatelně do uvedeného rozkladu.

Bylo zjištěno, že v MD simulacích mohou báze na obou koncích oligomeru tvorit alternativní *trans* Watson-Crick/sugar edge páry [16], které mají významný vliv na strukturu oligomeru, a to až po třetí step od konce. V současnosti neexistuje experimentální důkaz o existenci těchto páru v DD, což naznačuje, že jsou simulačním artefaktem. Z tohoto důvodu byly dále analyzovány pouze ty části trajektorie, které odpovídají kanonickým páru na obou koncích. Srovnání průměrné konformace DD s experimentem ukazuje na dobrou kvalitativní shodu se strukturami určenými pomocí rentgenové krystalografie a NMR. Odhaleny však byly i menší kvantitativní rozdíly, jako např. mírně podhodnocený twist a nebo obecně širší žlábky v MD simulacích. Za povšimnutí stojí i relativně výrazné rozdíly mezi jednotlivými experimentálními strukturami. Analýza jemné struktury DD dále přinesla nové poznatky o vlivu BI/BII substavů v páteři na helikální konformaci příslušného stepu [17], které poukazují na zajímavý strukturní polymorfismus helixu. Bylo zjištěno, že mírně anharmonické rozdělení některých souřadnic popisujících geometrii určitého stepu může být rozloženo na harmonické příspěvky odpovídající BI/BII substavům ve fragmentu páteře tohoto stepu a 3' sousedního

stepu (viz. obr. 2). Domníváme se, že vliv sousedního stepu byl popsán v této práci poprvé. Výsledky této práce byly publikovány v *J. Chem. Theory Comput.*

4.2 Mechanické vlastnosti A-traktů

A-trakty jsou funkčně významné sekvence DNA známé unikátními strukturními a mechanickými vlastnostmi. Ovlivňují např. lokalizaci nukleosomů vůči vláknu DNA, která je kritická pro regulaci exprese genů [18]. Zatímco strukturní vlastnosti A-traktů jsou již experimentálně dobře charakterizovány [19], zůstávají poznatky o jejich mechanických vlastnostech spíše kontroverzní. Na jednu stranu se A-trakty jeví strukturně rigidní a odolávají tvorbě nukleosomu [18], na druhou stranu se ukazují být flexibilní v případě tvorby smyček za účasti transkripčních faktorů [20]. V této práci jsme analyzovali mechanické vlastnosti symetrických $A_n T_n$ a asymetrických A_n ($n=3,4,5$) A-traktů pomocí modelu tuhých bází a modelu tuhosti v ohybu a twistu, abychom pomohli vysvětlit zdánlivě protichůdná experimentální data a prostudovat rozdíly mezi různými druhy A-traktů.



Obrázek 3: Energie spojená s deformací A_{10} a $A_5 T_5$ traktů a kontrolní sekvence v nukleosomu (pdb kód: 1kx5), zobrazená vůči pozici tohoto fragmentu.

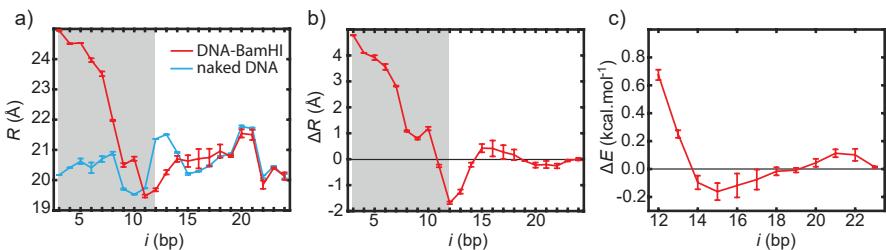
Získané výsledky naznačují, že relativní tuhost A-traktů vzhledem ke kontrolní sekvenci bohaté na C:G páry závisí na typu předepsané deformace. Asymetrické trakty jsou rigidnější než kontrolní sekvence vzhledem k lokálním deformacím, ale flexibilnější vzhledem k globálnímu ohybu a twistu. Lokalizovaná konformační omezení hraje roli při umisťování nukleosomů (viz. obr. 3), zatímco při tvorbě smyček jsou na konformaci helixu aplikovány globální omezení, které dovolují systému nalézt flexibilnější deformační mód. Výsledky tak

nabízejí vysvětlení experimentálně pozorovaných mechanických vlastností A-traktů. Dále bylo zjištěno, že symetrické trakty jsou rigidnější než asymetrické trakty i kontrolní sekvence vzhledem k oběma typům deformací. Symetrické trakty by tedy měly také odolávat tvorbě nukleosomu více než asymetrické, což může otevřít nové možnosti cílené manipulace exprese genů. Výsledky této práce byly publikovány v *Nucleic Acids Res.*

4.3 DNA alosterie

V nedávné době bylo pozorováno, že DNA může sloužit jako zprostředkovatel alosterické regulace. Navázání malého ligandu nebo proteinu způsobí deformaci struktury DNA a změnu její flexibility, což vede k ovlivnění afinity pro navázání dalšího ligandu [21]. Mechanismus těchto efektů však stále není zcela pochopen.

Součástí této práce je návrh modelu umožňujícího studium alosterických efektů v DNA. Tento model je nejdříve aplikován na studium vlivu vazby malých ligandů, diamidinů a pyrol-imidazolových polyamidů, do malého žlábku. Tyto látky jsou považovány za potencionální protinádorová léčiva [22] a byly také např. použity ke kvantifikaci příspěvku tzv. nepřímého čtení při rozpoznávání DNA proteiny [23]. Model alosterie předpovídá ohyb DNA způsobený vazbou těchto ligandů, který je v kvantitativní shodě s experimentem. Kromě toho model naznačuje, že v případě alternující TA sekvence způsobí vazba ligandu posun lokální i globální konformace směrem ke struktuře volného A-traktu, což může mít významné funkční důsledky.



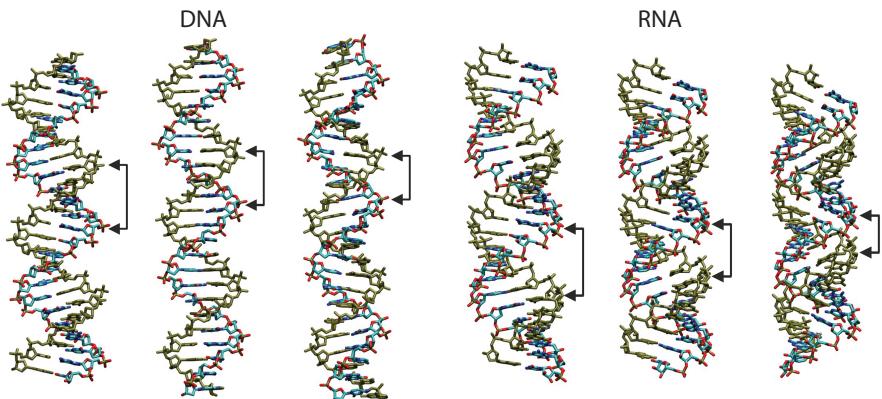
Obrázek 4: Alosterická regulace mezi dvěma proteiny zprostředkovaná helikální DNA. (a) Šířky velkého žlábku volné DNA (modrá) a DNA v komplexu s *BamHI* (červená), a (b) jejich rozdíl. Zastíněná oblast znázorňuje páry bází definující šířky žlábku, které byly fixovány pro účely modelování vazby proteinu. (c) Rozdíl mezi deformační energií navázání GRDBD proteinu v přítomnosti a při absenci *BamHI* proteinu, zobrazený jako funkce vzdálenosti mezi těmito proteinami podél helixu.

V nedávné experimentální studii [24] byl popsán efekt alosterické regulace, při kterém je afinita k navázání proteinu na DNA ovlivněna vazbou dalšího proteinu v jeho blízkosti, ale bez přímého kontaktu mezi těmito proteiny. Tento efekt vykazuje periodicitu ~ 10 párů bází a klesá s rostoucí vzdáleností mezi proteiny. K vysvětlení mechanismu efektu jsme využili model alosterie, který byl parametrizován s pomocí rozsáhlých MD simulací dlouhého oligomeru DNA (33 párů bází, $\sim 123,000$ atomů včetně vody a iontu) o stejně sekvenci jako byla použita v uvedené experimentální studii. Za předpokladu, že jsou vnitřní souřadnice modelu vhodně fixovány tak, aby byla co nejlépe napodobena vazba studovaných proteinů pozorovaná v krystalech jejich komplexu s DNA, vykazuje odhadovaná regulace afinity (vyjádřená rozdílem deformačních energií ΔE) skutečně periodicitu ~ 10 páru bází a klesající amplitudu. Stejnou funkční závislost pozorujeme u rozdílu mezi šírkou velkého žlábků v případě, kdy první protein je již navázán, a šírkou v případě volné DNA, což naznačuje, že právě změny v šířce žlábků zprostředkují daný alosterický efekt (viz. obr. 4). Odhadované maximální a minimální hodnoty ΔE jsou v dobré kvantitativní shodě s experimentem, ačkoliv nás model předpovídá rychlejší pokles se vzdáleností mezi proteiny. Výsledkem této práce je jedna publikace v *J. Phys. Chem. Lett.* a jedna v *Biophys. J.*

4.4 Spřažené deformace v torzi a prodloužení helikální DNA a RNA

Globální konformační flexibilita helikální DNA a RNA má důležitý vliv na biologickou aktivitu těchto molekul. Nedávné mikromanipulační experimenty odhalily pouze malé kvantitativní rozdíly mezi DNA a RNA v tuhosti vůči ohybu, prodloužení (stretch) a celkové torzi (total twist), ale také překvapivě velké kvalitativní rozdíly v tuhosti vůči spřažené deformaci v torzi a prodloužení, vyjádřené nediagonálním elementem matice tuhostních konstant [25]. Při vnuzeném zvýšení celkové torze dochází k prodloužení helikální DNA, zatímco helikální RNA se zkrátí. V této práci jsme využili MD simulace oligomerů DNA a RNA obsahujících všechny existující dvojice po sobě následujících párů bází (basepair steps) a modely mechaniky parametrizované pomocí těchto simulací, abychom odhadli koeficienty helikálního prodloužení vůči změně torze a vyšetřili rozdíly v konformačních změnách mezi DNA a RNA, ke kterým dochází při vnuzeném snížení a zvýšení celkové torze oligomerů.

Pro koeficienty helikálního prodloužení vůči změně torze byly nalezeny hodnoty $0.031 \text{ \AA deg}^{-1}$ v případě DNA a $-0.036 \text{ \AA deg}^{-1}$ v případě RNA, což je v dobré kvalitativní shodě s hodnotami zjištěnými experimentálně [25] ($0.014 \text{ \AA deg}^{-1}$ pro DNA a $-0.024 \text{ \AA deg}^{-1}$ pro RNA). Simulace nicméně mo-



Obrázek 5: Struktury DNA a RNA získané minimalizací helikálních souřadnic a úhlů v páteři v případě vnuzeného snížení helikální torze (struktura vlevo), kanonické struktury (uprostřed) a v případě vnuzeného zvýšení torze (vpravo).

hou poskytnout podrobnější obraz příslušných konformačních změn. Například vnucené zvyšování celkové torze vede u DNA i RNA ke snížení sklonu páru bází vůči helikální ose (inklinace), což má za následek snížení průměru helixu. Zatímco v případě DNA je s tím spojeno zvýšení inklinace, což vede k prodloužení DNA podél helikální osy a k redukci šířky malého žlábku, v případě RNA inklinace klesá, což vede ke zkrácení helixu a výrazné redukci šířky velkého žlábku (viz. obr. 5).

Tyto výsledky byly získány ve spolupráci se skupinou prof. Martina Zachariase na TU v Mnichově a skupinou prof. Jana Lipferta na LMU v Mnichově. Naším příspěvkem byl především odhad koeficientů helikálního prodloužení vůči změně torze pomocí modelů mechaniky. Výsledky byly publikovány v *Nucleic Acids Res.*

4.5 Strukturní a mechanické vlastnosti poškozené DNA

Buněčná DNA neustále podléhá různým druhům poškození. Pokud jsou vzniklé léze ponechány neopravené, může to vést k mutacím, genové nestabilitě, předčasnemu stárnutí nebo rakovině [26, 27]. Buňky si vyvinuly množství mechanismů, jak tyto léze detektovat, signalizovat jejich přítomnost a zajistit opravu [26]. Jedna z otázek, kterou je třeba zodpovědět, je jak proteiny zajišťující opravu rozpoznají danou léze mezi velkým množstvím nepoškozených páru bází. Bylo navrženo, že v mnoha případech jsou léze specificky rozpoznány

na základě změny struktury a flexibility DNA v jejich blízkém okolí (např. [28, 29, 30]).

Celkově jsme se podíleli na 4 studiích zaměřených na vlastnosti poškozené DNA. První dvě se zabývají vlivem 8-oxoguaninu [31] and O⁶-methylguaninu [32] na strukturní a mechanické vlastnosti DNA. Tyto studie byly provedeny ve spolupráci se skupinou prof. M. Zachariase na TU v Mnichově. Naším příspěvkem byla především analýza změn konformace a mechanických vlastností DNA na úrovni jednotlivých bází způsobených poškozením. Výsledky analýzy ukazují, že přítomnost 8-oxoguaninu způsobuje posun konformace DNA blíže ke konformaci pozorované v krystalu komplexu DNA s opravným proteinem, zachyceném v úvodní fázi procesu rozpoznání [33], což naznačuje, že samotná struktura poškozené DNA může usnadnit vznik komplexu s proteinem zodpovědným za rozpoznání léze. Na druhou stranu však v tomto případě nebyly pozorovány žádné změny ve flexibilitě mezi poškozenou a nepoškozenou DNA. V kontrastu s tímto pozorováním ukázala analýza vlastností DNA obsahující O⁶-methylguanin výrazné změny ve flexibilitě způsobené přítomností léze. Tuhost vůči některým deformacím je daleko nižší u páru obsahujícího poškozenou bázi O6-MeG, než u kanonického G:C páru. Tyto deformace zvyšují přístupnost dané báze, což pravděpodobně hraje rozhodující roli při rozpoznání léze opravnými proteiny, které jsou známé tím, že v průběhu procesu opravy vytočí O6-MeG bázi zevnitř helixu do aktivního místa enzymu [34].

Další dvě studie byly provedeny ve spolupráci se skupinou E. Dumontové na ENS v Lyonu. V případě těchto studií jsme přispěli především monitrováním a analýzou lokálních a globálních strukturních změn poškozených oligomerů v MD simulacích. Tyto simulace byly doplněny QM výpočty sloužícími k nalezení optimálních geometrií poškozených míst (tyto výpočty byly provedeny spolupracující skupinou). První z těchto prací je zaměřena na studium struktury, dynamiky a vnitřních interakcí oligomeru obsahujícím C4'-oxidované abazické místo. Tato místa mohou podléhat reakcím s dalšími bázemi, které vedou ke vzniku tzv. cross-linků mezi dvěma řetězci DNA, jež patří mezi nejvíce zhoubné typy lézí [35, 36]. Naše výsledky nabízejí vysvětlení experimentálně pozorované vysoké affinity oxidovaných abazických míst k tvorbě cross-linku s protilehlým cytosinem a dále ukazují, že cross-linky mezi řetězci mohou vznikat také mezi oxidovaným abazickým místem a bází sousedící s protilehlou bází. Druhá práce je zaměřena na dynamiku DNA obsahující cross-link v rámci jednoho řetězce [37], který vzniká kovalentním spojením sousedící purinové a pyrimidinové báze. V současné době neexistují experimentální struktury oligomerů obsahujících tyto léze, které by pomohly vysvětlit jejich vysokou mutagenitu související s nedostatečnou

opravou těchto lézí. V práci jsme využili MD simulace oligomerů obsahujících guanin-thymin cross-link (G[8-5m]T) a guanin-cytosin cross-link (G[8-5]C), doplněné o QM výpočty, k charakterizaci a racionalizaci strukturní dynamiky těchto lézí. Odhaleny byly strukturní rozdíly mezi G[8-5m]T a G[8-5]C a diskutovány důsledky vzhledem k procesu rozpoznání těchto lézí proteiny [38]. Dále bylo zjištěno, že lokální struktura G[8-5m]T a G[8-5]C lézí se výrazně odlišuje od komplexních foto-indukovaných lézí podobné konfigurace, což podporuje hypotézu o odlišném chování těchto dvou typů lézí, pravděpodobně i vzhledem k efektivitě jejich opravy.

Celkově vedly studie zaměřené na poškození DNA, vzniklé ve spolupráci s dalšími skupinami, ke čtyřem publikacím v impaktovaných časopisech (*J. Phys. Chem. B*, *Biopolymers* a dvě práce v *Biochemistry*).

5 Závěry

Tato práce představuje výsledky několika studií zaměřených na strukturní a mechanické vlastnosti helikální DNA a RNA v kontextu biologicky relevantních procesů pomocí metod počítačového modelování. Představen je model alosterických efektů v DNA a výsledky aplikace tohoto modelu na studium alosterické regulace vazby ligandů a proteinů na DNA. Atomistické MD simulace slouží jako primární zdroj dat o konformační dynamice nukleových kyselin a pro parametrisaci modelů, které jsou použity ke studiu jejich mechanických vlastností. Získané výsledky jsou vždy konfrontovány s daty dostupnými z experimentu. Tuhost helixů nukleových kyselin vůči globálním deformacím, jako je ohyb a celková torze, je v téma kvantitativní shodě s experimentem. Je pozoruhodné, že relativně jednoduché modely mechaniky uvažující harmonickou deformační energii jsou schopné reprodukovat i natolik jemné efekty, jako jsou spřažené deformace v torzi a prodloužení DNA a RNA a alosterické regulace zprostředkované DNA. Tato srovnání demonstруjí všeobecnou spolehlivost současných silových polí a modelů mechaniky a činí tak predikce, které tyto modely umožňují, přesvědčivými. Naše výsledky nabízí vysvětlení mechanických vlastností A-traktů v souvislosti s tvorbou smyček DNA a lokalizací nukleosomů, předpovídají změny v konformaci DNA zprostředkující alosterickou regulaci vazby proteinů a ligandů na DNA a předpovídají změny ve vlastnostech DNA způsobené poškozením, které jsou relevantní pro rozpoznání poškození proteiny a jejich opravu. Tyto výsledky mohou pomoci lépe pochopit vliv sekvenčně závislých strukturních a mechanických vlastností nukleových kyselin na jejich biologickou funkci a pomoci tak např. při návrhu nových léčiv.

Modeling mechanical properties of RNA and DNA

1 Introduction

DNA and RNA are key biological molecules [1]. DNA serves as a carrier of genetic information which consists primarily of instructions for making proteins that perform most of the cellular functions. The genetic information is stored in a linear sequence of bases in the coding regions of DNA. It is becoming clear that there is additional information laid over this literal code. A particular sequence of bases implies distinct structural preferences, mechanical properties and other physicochemical characteristics which are of central importance for the biological activity of DNA [2, 3]. These properties govern the organization of DNA in a cell and play a role in processes such as protein/DNA recognition that are, in turn, critical for gene regulation, transcription and DNA replication. In the case of double-stranded RNA, its structural and mechanical properties influence its interactions with proteins, determine the viral RNA capsid compaction and play a role in the processes of protein synthesis and RNA interference [4].

A number of experimental techniques are available that can be used to probe the structural and mechanical properties of nucleic acids. Methods such as X-ray crystallography and nuclear magnetic resonance (NMR) can provide high-resolution information about the structure of nucleic acids in crystal and in solution. Other methods, such as cyclization experiments or force spectroscopies, can be used to probe overall flexibility of the studied biopolymer with respect to twisting, stretching and bending deformations [5, 6]. However, the methods directly measuring flexibility suffer from low resolution and other limitations. Flexibility can be assessed also from the knowledge of the accessible conformational space of nucleic acids [7]. However, the high-resolution structural data, available primarily from X-ray crystallography, may be sparse in comparison to the manifold of possible conformations.

Another viable approach to investigate structural and mechanical properties of nucleic acids is to utilize computational methods. In particular,

atomistic molecular dynamics (MD) simulations with explicit representation of water and ions have proven to be successful in describing many important properties of large systems of biological relevance [8, 9]. An MD simulation is a computer experiment that enables one to measure positions of all atoms in the system as a function of time under the specified external conditions such as temperature and pressure [10, 11]. The time series of atomic coordinates obtained from MD can be transformed to time series of various conformational descriptors such as relative positions and orientations of the bases within the DNA and RNA (coarse-graining), which can then be used to find parameters of various models of nucleic acid mechanics. However, MD simulations necessarily treat the system in a simplified way, and thus should always be validated against experimental data or high-accuracy quantum mechanics calculations. The main limitations include the available time scales and sometimes severe approximations to the potential energy function.

2 Aims of the study

The aim of this study is to investigate structure, dynamics and mechanical properties of various nucleic acid molecules in the context of biologically relevant processes using methods of computer modelling. The unrestrained atomic-resolution MD simulations with explicit representation of water and ions serve as the primary source of data. These simulations are used to study the conformational dynamics of the nucleic acid helices, and to parametrize non-local harmonic energy models that are then employed to study their mechanics. This approach is applied to investigate mechanical properties of DNA poly(dA:dT) sequences (A-tracts) in the context of DNA looping and nucleosome positioning, differences in twist-stretch coupled deformations between DNA and RNA, and properties of lesion-containing DNA related to damage recognition. The results are tested for convergence and always carefully compared to available experiment. Besides that, we propose a model of DNA allosteric coupling based on constrained minimization of the effective harmonic deformation energy. The model is parameterized from MD simulations and then applied to study the effect of minor groove binding of small ligands and the mechanisms of allosteric coupling between proteins mediated by the DNA.

3 Material and methods

To investigate the mechanical properties of nucleic acids, we employ two harmonic energy models which differ in the level of coarse-graining. Within the

scope of the first model (model of rigid bases), the conformation of a nucleic acid helix is specified by a set of internal coordinates describing the relative displacement and orientation of the individual bases which are treated as rigid objects (see Fig. 6A). In the second, coarser one (model of bending and twisting rigidity) the conformation of the whole nucleic acid duplex is specified by just three internal coordinates: bending in the direction of the grooves, bending in the direction of the backbone, and total twist (see Fig. 6B). Both models assume the internal deformation energy, or the free energy

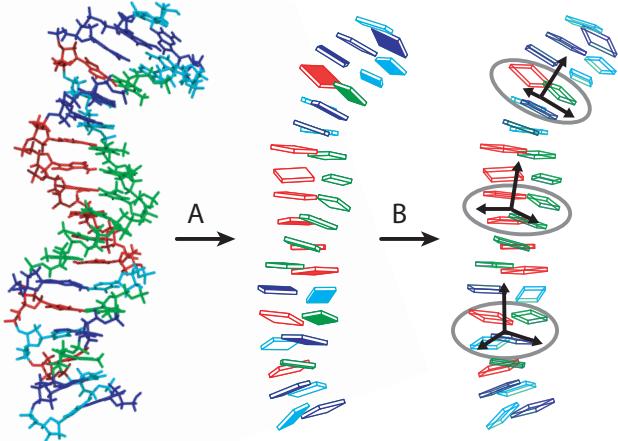


Figure 6: Description of DNA/RNA conformation at different levels of resolution. (A) Bases in the DNA and RNA can be, to a good approximation, treated as rigid objects with reference points and right-handed orthonormal frames attached to them. The conformation of the molecule is then described by relative position and orientation of these points and frames with respect to each other. (B) Similarly, global conformation of nucleic acids, like global bending, can be described by using points and frames representing whole fragments of the nucleic acid molecule. They are determined by averaging base-fixed points and frames in selected groups of bases.

necessary to distort the oligomer, to be a non-local quadratic function of the respective internal coordinates,

$$E(\mathbf{w}) = \frac{1}{2} (\mathbf{w} - \hat{\mathbf{w}}) \cdot \mathbf{K} (\mathbf{w} - \hat{\mathbf{w}}), \quad (5)$$

where \mathbf{w} is the vector of internal coordinates, $\hat{\mathbf{w}}$ are their equilibrium values and \mathbf{K} is a symmetric, positive definite matrix of stiffness constants. Considering the system to be in contact with a thermal bath, it follows from the theory

of thermodynamic fluctuations [12] that the probability distribution of \mathbf{w} is a multivariate Gaussian. It is then straightforward to show that the parameters of the model, $\hat{\mathbf{w}}$ and K , are related to the moments of the distribution as

$$\hat{\mathbf{w}} = \langle \mathbf{w} \rangle, \quad K = k_B T C^{-1}, \quad (6)$$

where $\langle \cdot \rangle$ denotes the expected value and C is the covariance matrix of \mathbf{w} . The moments of the canonical ensemble (eq. 6) are replaced here by the moments of the coordinate time series, obtained from an unrestrained atomic-resolution MD trajectory of a naked DNA/RNA molecule. When the parameters $\hat{\mathbf{w}}$ and K are known, the model can be used to estimate energy changes associated with various deformations prescribed by \mathbf{w} using eq. (5), provided the deformations are in the harmonic regime. It is found that the harmonic approximation is satisfied rather well for the studied systems. The moderate anharmonicity due to conformational substates is coarse-grained by replacing the observed distribution by Gaussian distribution with the same moments, resulting in an effective harmonic description. These models have already proven to be useful in understanding many processes [7].

To model the allosteric effects in DNA, the rigid base model is extended to include minor and major groove widths, which represent important parameters in interactions of DNA with small ligands and proteins. The coordinates are divided into two subsets \mathbf{w}_1 and \mathbf{w}_2 . To mimic binding of a ligand to DNA, the coordinates \mathbf{w}_1 are constrained to have prescribed values $\bar{\mathbf{w}}_1$, dictated by the ligand (e.g. minor groove binder fixes the width of the groove). The remaining coordinates are left free to relax to their energetically optimal values $\bar{\mathbf{w}}_2$. The minimization of the quadratic deformation energy consistent with the constraints then yields relations between the imposed constraints $\bar{\mathbf{w}}_1$ and the response,

$$\bar{\mathbf{w}}_2 = \hat{\mathbf{w}}_2 - K_{22}^{-1} K_{12}^T (\bar{\mathbf{w}}_1 - \hat{\mathbf{w}}_1), \quad (7)$$

and an expression for the corresponding deformation energy

$$E = 1/2 (\bar{\mathbf{w}}_1 - \hat{\mathbf{w}}_1) \cdot [K_{11} - K_{12} K_{22}^{-1} K_{12}^T] (\bar{\mathbf{w}}_1 - \hat{\mathbf{w}}_1), \quad (8)$$

where K_{12} and K_{22} are submatrices of K corresponding to the two subsets. To model allosteric coupling between two ligands binding to the DNA (e.g. two proteins), the model is further extended to divide the coordinates into three subsets. The first and second subsets are constrained to mimic binding of the first and second ligand, respectively. The remaining coordinates are left free to relax. The energy minimization then yields relations analogous to (7) and (8). The allosteric effect is defined by the difference ΔE between the

deformation energy of binding the second ligand in the presence of the first one, and in its absence.

The MD simulations were performed using the modern AMBER force fields for nucleic acids (*parm99* with *bsc0* correction, and with χ_{OL3} correction for RNA). The simulation protocol adopted for the simulations done in our laboratory is based on the well-established protocol of the ABC consortium [13]. The modeled oligomer is built as canonical B-DNA (or A-RNA) and solvated in an octahedral periodic box with SPC/E water model [14]. Cations are added to the system to neutralize the negative charge of the oligomer and additional ions to mimic the physiological concentration of 150 mM of the KCl salt. The systems are properly equilibrated prior to the production run. In the trajectory production phase, the system is simulated at constant temperature (300 K) and pressure (1 atm). The octahedral periodic boundary conditions are applied together with the PME method for calculating long-range interactions. The bonds containing hydrogens are treated with SHAKE and the time step used is 2 fs.

Trajectory snapshots are saved at 10 ps intervals to be used in the analysis. In this way, a set of conformations sampling the NpT statistical ensemble is obtained. For each of the conformations, internal coordinates of the models can be calculated. Ensemble moments of these coordinates are then related to model parameters according to eq. (6).

4 Results and discussion

4.1 DNA fine structure and substates

MD simulations yield valuable information about the properties of DNA. It remains of primary importance, however, to always test how well the MD data compare to experiment, given the force field and simulation setup, and to test the convergence of quantities of interest. This delineates the limits of use of MD and helps in designing new force fields and simulation protocols. To address these issues, we have analyzed fine structure and dynamics of B-DNA determined from extensive MD simulations. We have focused our attention to the so-called Dickerson-Drew dodecamer (DD) d-[CGCGAATTGCG]2, whose properties have been extensively studied by many experiments [15].

First, it was found that the end bases can adopt an alternative *trans* Watson-Crick/sugar edge pairing [16], which has a strong effect on the DD conformation up to the third step from the end. Currently there seems to be no experimental evidence of these pairs in DD, suggesting that they are a

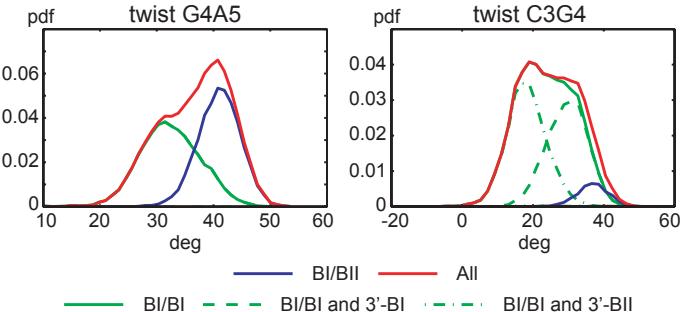


Figure 7: Anharmonic distribution of twist in G4A5 step can be decomposed into nearly harmonic contributions corresponding to the BI/BI and BI/BII states of the step (left). Twist of C3G4 in the BI/BI state is still anharmonic but can be further decomposed into harmonic contributions depending on the BI or BII states of the G4-A5 backbone fragment in the 3' neighboring step (right). The BII/BII states are rare and contribute negligibly to the decomposition.

simulation artifact. Thus, only parts of the trajectory corresponding to the canonical end base pairs were further analyzed. A comparison of mean DNA conformation with X-ray and NMR experimental data reveals good qualitative agreement between simulation and experiment, but also some minor quantitative differences, such as slightly underestimated twist or generally wider minor and major grooves in MD simulations. Interestingly, significant differences between individual experimental structures are also apparent. An inspection of the DD fine structure further revealed rather significant influence of BI/BII backbone substates on the base pairs step conformation, extending previous findings [17] and exposing an interesting structural polymorphism in the helix. It was further found that moderately anharmonic distributions of some of the base pair step coordinates can be decomposed into Gaussian-like components associated with the BI/BII substates in the backbone fragment corresponding to the step itself and to the 3' neighboring step (see Fig. 7). We believe that this was for the first time the influence of the substates of a neighboring step backbone fragment was reported. This work was published in *J. Chem. Theory Comput.*

4.2 Mechanical properties of DNA A-tracts

A-tracts are functionally important DNA sequences that are known to have unique structural and mechanical properties. In particular, A-tracts affect nucleosome positioning, which is, in turn, critical for gene regulation [18]. The structural properties of A-tracts are now well established [19], but their mechanical properties remain rather controversial. A-tracts appear structurally rigid and resist nucleosome formation [18], but seem flexible in transcription factor mediated DNA looping [20]. We have analysed mechanical properties of symmetric A_nT_n and asymmetric A_n ($n=3,4,5$) A-tracts using the two harmonic energy coarse-grained models to investigate the contradictory experimental data and the differences between these two types of A-tracts.

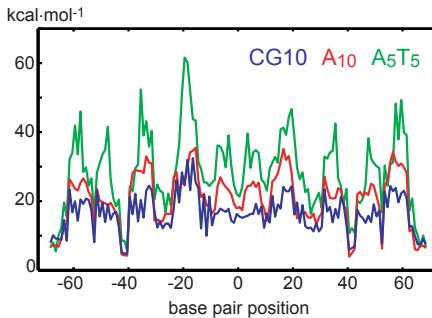


Figure 8: Deformation energy associated with threading of A_{10} and A_5T_5 tracts and the central 10 bp of the control sequence through the 1kx5 nucleosome structure.

We have found that the stiffness of A-tracts depends on the type of deformation imposed. Asymmetric tracts are more rigid than the control C/G-rich sequence with respect to localized deformations, but more flexible with respect to global bending and twisting. The localized constraints play a role in the nucleosome positioning (Fig. 8 shows the energy needed to deform A-tracts and the control to a geometry in the nucleosome as a function of a location in the nucleosome), while global boundary conditions are imposed in the DNA looping, which allow the system to find a more flexible deformation mode. The results thus offer an explanation of the experimental observations. The symmetric A_nT_n tracts were found to be stiffer than the asymmetric tracts and the control with respect to both types of deformation. Thus, they should also resist nucleosome formation more than the asymmetric ones, which would open new possibility of gene expression manipulation. This work was published in *Nucleic Acids Res.*

4.3 DNA allostery

It is becoming increasingly appreciated that DNA can serve as an allosteric mediator. Binding of a small ligand or protein causes deformation of the DNA structure and a change in its flexibility in a way that affects binding affinity of a subsequent ligand [21]. The underlying mechanism of this effect remains poorly understood.

We have proposed a general model of DNA allostery and applied it first to investigate minor groove binding of diamidines and pyrrole-imidazole polyamides, promising antitumor drugs [22] also designed e.g. to quantify protein indirect readout [23]. The model predicts DNA bending caused by the ligand binding that is in quantitative agreement with experiment and further suggests that the binding event brings the conformation of an alternating TA sequence closer to the conformation of a free A-tract both locally and globally, which may have important functional implications.

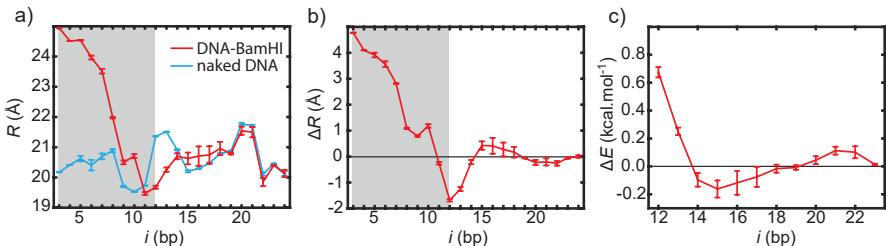


Figure 9: Protein-protein allosteric coupling mediated by the DNA helix. (a) Major groove widths of the unconstrained DNA (blue) and of DNA complexed with *BamHI* (red), and (b) their difference. The shaded area denotes the base pairs defining the groove widths that were constrained to mimic the protein binding. (c) Difference between the DNA deformation energy due to GRDBD binding in the presence and in the absence of *BamHI* as a function of the distance between the two proteins along the helix.

A recent experimental study [24] described an allosteric effect in which the binding affinity of a protein to DNA is influenced by another protein bound nearby, without any direct contact between them. The effect shows ~ 10 -bp periodicity and decays with increasing distance between the proteins. We applied our model to provide a mechanistic explanation. The model of DNA allostery was parametrized from an extensive MD simulation of a long DNA oligomer (33 base pairs, $\sim 123,000$ atoms together with water and ions) of the same sequence as in the experimental study. If the constraints are properly

applied to mimic binding of the two proteins seen in the crystal structures of these proteins complexed with DNA, the estimated energy difference ΔE indeed shows ~ 10 -bp periodicity. The same applies to the difference between the major groove width when the constraints are applied and the width in the naked DNA, offering a mechanistic explanation of the allosteric effect (see Fig. 9). The maximum and minimum values of ΔE are also in a good quantitative agreement with what was measured in the experiment, although our values decay faster. This work resulted in one publication in *J. Phys. Chem. Lett.* and one in *Biophys. J.*

4.4 DNA and RNA twist-stretch coupling

The global conformational flexibility of a double helical DNA and RNA is of central importance for their biological activity. As determined from recent single molecule experiments, DNA and RNA exhibit only a modest quantitative differences in global bending, twisting and stretching flexibility, but show a striking qualitative difference in the twist-stretch coupling [25]. Upon over-twisting, the dsDNA increases its helical extension, while dsRNA behaves in the opposite way. We employed MD simulations of dsDNA and dsRNA containing all possible dinucleotide steps and the models of mechanics to estimate the coupling coefficients and to investigate the differences in the conformational distortions connected with over- and undertwisting deformations.

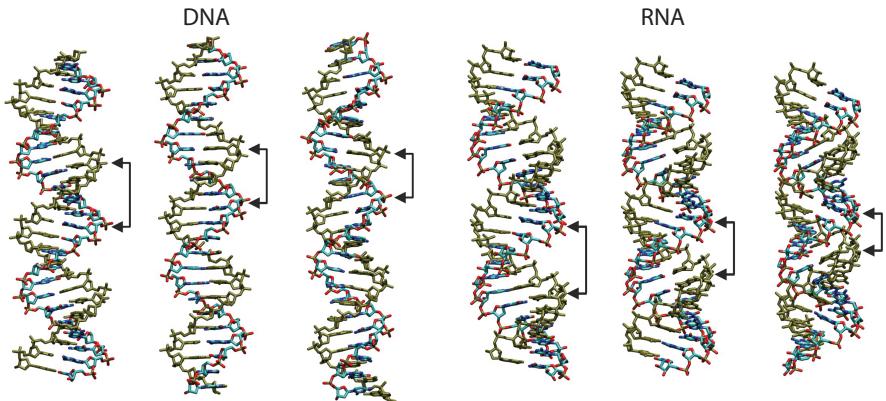


Figure 10: Structures of DNA and RNA obtained by minimizing helical and backbone coordinates keeping the x-displacement and inclination coordinates locked to the values corresponding to undertwisted (left structure), canonical (middle structure) or overtwisted (right structure) oligomer.

The coupling coefficient for the helical extension upon imposed twisting was found to be $0.031 \text{ \AA deg}^{-1}$ for the DNA and $-0.036 \text{ \AA deg}^{-1}$ for the RNA in semi-quantitative agreement with the experiments [25] ($0.014 \text{ \AA deg}^{-1}$ for the DNA and $-0.024 \text{ \AA deg}^{-1}$ for the RNA). The simulations can provide further insight into the corresponding conformational changes. For instance, overtwisting of both DNA and RNA leads to reduction of the absolute value of x-displacement, thus reducing the duplex radius. However, for DNA it leads to more positive values of inclination, which is coupled to an extension of DNA along the helical axis and a reduction of the minor groove width. In the case of RNA, inclination decreases, which is coupled to a reduction of the helical extension and a significant reduction of the major groove width (see Fig. 10).

These results were obtained in collaboration with the groups of prof. Zacharias at TU Munich and prof. Jan Lipfert at LMU Munich. We have contributed by estimating the twist-stretch couplings from unrestrained MD simulations using the harmonic stiffness model. The study was published in *Nucleic Acids Res.*

4.5 Structural and mechanical properties of damaged DNA

DNA is constantly being damaged. If left unrepaired, lesions may lead to mutations, genomic instability, premature aging or cancer [26, 27]. To deal with the damage, cells have evolved mechanisms to detect DNA lesions, signal their presence and promote their repair [26]. One of the underlying questions that remains to be fully answered is how repair proteins locate the lesion in such a large surplus of undamaged base pairs. It is believed that in many cases the lesion alters the structural and mechanical properties of DNA helix in the vicinity of the lesion, which facilitates the recognition (e.g. [28, 29, 30]).

We have performed four MD studies concerned with the properties of damaged DNA. The first two of them investigated the effect of 8-Oxoguanine [31] and O⁶-methylguanine [32] lesions. These studies were done in collaboration with the group of M. Zacharias at TU Munich. We contributed mainly by analysing the changes of properties of the DNA induced by the lesions in terms of conformation and flexibility at the level of rigid bases. The results show that 8-Oxoguanine lesion brings DNA closer to its conformation observed in the crystal structure of the initial repair protein-DNA recognition complex [33]. This indicates that the damaged DNA may be partially preformed to facilitate the creation of the complex. On the other hand, the rigid base model suggests that the damaged and undamaged DNA have very similar flexibility.

In contrast to this, the analysis of the properties of O⁶-methylguanine lesion-containing DNA showed significant changes in flexibility due to the presence of the lesion. In particular, the stiffness of shear, stretch and opening deformations of the base pair containing O6-MeG was found to be much lower compared to a G:C pair. These motions enhance the accessibility of the base, which is likely to play a decisive role for the induced fit recognition of the lesion by the repair enzymes, which are known to loop out the O6-MeG base to transfer it into the enzyme active site during the repair [34].

The other two studies were done in collaboration with the group of E. Dumont at ENS Lyon. In case of these studies, we contributed especially by monitoring and analyzing the structural rearrangements in MD simulations of the damaged oligomers in terms of various structural descriptors such as rigid base coordinates and global bending angles. The MD simulations were complemented by QM calculations to quantify optimal geometries of the damaged sites (done by the collaborating group). The first study concerned the structure, dynamics and interactions of oligomers containing C4'-oxidized abasic site. It has been observed experimentally that these sites can undergo a coupling reaction with other nucleobases resulting in interstrand cross-links which are among the most deleterious lesion [35, 36]. Our results provide a mechanistic explanation of the experimentally observed high affinity of oxidized abasic site in forming cross-link with opposing cytosine base and further indicate that cross-links can be formed with strand offset. The second study concerned properties and dynamics of DNA containing intrastrand cross-link lesion [37], featuring a purine and a pyrimidine base that are covalently linked together within the strand. The field suffers from the absence of experimental structural data that would explain the lack of repair of these highly mutagenic lesion. We have employed MD simulations of guanine-thymine (G[8-5m]T) and guanine cytosine (G[8-5]C) intrastrand cross-links combined with QM calculations to characterize and rationalize the structural dynamics of these lesion. We expose structural differences between G[8-5m]T and G[8-5]C and discuss possible implications for the recognition process [38]. We further show that the resulting structures differ locally from complex photoinduced lesions, which reinforces a hypothesis of contrasted behaviour of light-induced versus oxidatively generated lesions, possibly in line with the lack of repair.

Taken together, our collaborative studies on damaged DNA resulted in four journal publications (*J. Phys. Chem. B*, *Biopolymers* and 2 in *Biochemistry*).

5 Conclusions

The thesis presents results of several computational studies focused on the structural and mechanical properties of DNA and RNA helices in the context of biologically relevant problems. It further presents the novel DNA allostery modelling approach and results of its application to mechanism of allosteric coupling between proteins mediated by the DNA. Atomic-resolution MD simulations serve as the primary source of data used to study the conformational dynamics of the nucleic acids and to parametrize coarse-grained models of mechanics. Whenever possible, the results are confronted with available experimental findings. Modern nucleic acid force fields reproduce the average local and global conformation very well. In addition, the stiffness of nucleic acid helices with respect to global deformations, such as bending and twisting, is in almost quantitative agreement with experiment. Remarkably, the relatively simple models of mechanics can reproduce even such subtle effects as DNA and RNA twist-stretch coupling or allosteric modulation in DNA. The basic reliability and robustness of the contemporary force fields and our modelling approach thus have been demonstrated, bringing confidence into the predictions that can be made using these models. Our results have provided an explanation of mechanical properties of A-tracts in the context of DNA looping and nucleosome positioning, predict changes in the conformation responsible for the effect of allosteric coupling between proteins binding to DNA, and predict changes in the properties of DNA upon damage that are likely to be relevant for damage recognition. We believe that these results can help to better understand the role of sequence dependent properties of nucleic acids in these processes that can, in turn, be utilized e.g. in the development of new drugs.

Přílohy / Appendices

Použitá literatura / References

- [1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular biology of the cell*. Garland Science, New York, 2002.
- [2] S. C. J. Parker and T. D. Tullius. DNA shape, genetic codes, and evolution. *Curr. Opin. Struct. Biol.*, 21:342–347, 2011.
- [3] R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, and R. S. Mann. Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.*, 79:233–269, 2010.
- [4] J. Wereszczynski and I. Andricioaei. On structural transitions, thermodynamic equilibrium, and the phase diagram of DNA and RNA duplexes under torque and tension. *P. Natl. Acad. Sci. USA*, 103:16200–16205, 2007.
- [5] J. P. Peters and L. J. Maher III. DNA curvature and flexibility *in vitro* and *in vivo*. *Q. Rev. Biophys.*, 43:23–63, 2010.
- [6] R. Lavery, A. Lebrun, D. cois Allemand, J.-F., and Bensimon, and V. Croquette. Structure and mechanics of single biomolecules: experiment and simulation. *J. Phys.: Condens. Matter*, 14:R383–R414, 2002.
- [7] F. Laskaš. *Innovations in biomolecular modeling and simulations.*, chapter 1. Royal Society of Chemistry, London, 2012.
- [8] J. Šponer, P. Banáš, P. Jurečka, M. Zgarbová, P. Kührová, M. Havrla, M. Krepl, P. Stadlbauer, and M. Otyepka. Molecular dynamics simulations of nucleic acids. from tetranucleotides to the ribosome. *J. Phys. Chem. Lett.*, 5:1771–1782, 2014.
- [9] A. Pérez, F. J. Luque, and M. Orozco. Frontiers in molecular dynamics simulations of DNA. *Accounts Chem. Res.*, 45:196–205, 2012.

- [10] D. Frenkel and B. Smit. *Understanding molecular simulation*. Academic press, 1996.
- [11] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, 1989.
- [12] L. D. Landau and E. M. Lifshits. *Statistical Physics, Part 1*. Elsevier, Amsterdam, 1980.
- [13] R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. Cheatham III, S. Dixit, B. Jayaram, F. Lankaš, Ch. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, Nad'a Špačková, and Jiří Šponer. A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.*, 38:299–313, 2010.
- [14] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91:6269–6271, 1987.
- [15] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. USA*, 78:2179–2183, 1981.
- [16] N. B. Leontis, J. Stombaugh, and E. Westhof. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.*, 30:3497–3531, 2002.
- [17] D. Djuranovic and B. Hartmann. DNA fine structure and dynamics in crystals and in solution: The impact of BI/BII backbone conformations. *Biopolymers*, 73:356–368, 2004.
- [18] Eran Segal and Jonathan Widom. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.*, 19:65–71, 2009.
- [19] T. E. Haran and U. Mohanty. The unique structure of A-tracts and intrinsic DNA bending. *Q. Rev. Biophys.*, 42:41–81, 2009.
- [20] S. Johnson, Y.-J. Chen, and R. Phillips. Poly(dA:dT)-rich DNAs are highly flexible in the context of DNA looping. *PLoS One*, 8:e75799, 2013.
- [21] J. B. Chaires. Allostery: DNA does it too. *ACS Chem. Biol.*, 3:207–209, 2008.

- [22] J. A. Raskatov, N. G. Nickols, A. E. Hargrove, G. K. Marinov, B. Wold, and P. B. Dervan. Gene expression changes in a tumor xenograft by a pyrrole-imidazole polyamide. *Proc. Natl. Acad. Sci. USA*, 109:16041–16045, 2012.
- [23] R. Moretti, L. J. Donato, M. L. Brezinski, R. L. Stafford, H. Hoff, J. S. Thorson, P. B. Dervan, and A. Z. Ansari. Targeted chemical wedges reveal the role of allosteric DNA modulation in protein-DNA assembly. *ACS Chem. Biol.*, 3:220–229, 2008.
- [24] Sangjin Kim, Erik Broströmer, Dong Xing, Jianshi Jin, Shasha Chong, Hao Ge, Siyuan Wang, Chan Gu, Lijiang Yang, Yi Qin Gao, Xiao-dong Su, Yujie Sun, and X. Sunney Xie. Probing allostery through DNA. *Science*, 339:816–819, 2013.
- [25] J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Köber, Z. Yu, S. P. Donkers, F. C. Chou, R. Das, and N. H. Dekker. Double-stranded RNA under force and torque: Similarities to and striking differences from double-stranded DNA. *Proc. Natl. Acad. Sci. USA*, 111:15408–15413, 2014.
- [26] S. P. Jackson and J. Bartek. The DNA-damage response in human biology and disease. *Nature*, 461:1071–1078, 2009.
- [27] M. S. Cooke, M. D. Evans, M. Dizdaroglu, and J. Lunec. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J.*, 17:1195–1214, 2003.
- [28] N. E. Geacintov, S. Broyde, T. Buterin, H. Naegeli, M. Wu, S. Yan, and D. J. Patel. Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers*, 65:202–210, 2002.
- [29] M. Jaciuk, E. Nowak, K. Skowronek, A. Tanska, and M. Nowotny. Structure of UvrA nucleotide excision repair protein in complex with modified DNA. *Nature Struct. Mol. Biol.*, 18:191–U254, 2011.
- [30] Y. Cai, K. Kropachev, R. Xu, Y. Tang, M. Kolbanovskii, A. Kolbanovskii, S. Amin, D. J. Patel, S. Broyde, and N. E. Geacintov. Distant neighbor base sequence context effects in human nucleotide excision repair of a benzo[a]pyrene-derived DNA lesion. *J. Mol. Biol.*, 399:397–409, 2010.

- [31] D. E. Barnes and T. Lindahl. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.*, 38:445–476, 2004.
- [32] D. Fu, J. A. Calvo, and L. D. Samson. Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer*, 12:104–120, 2012.
- [33] R.-J. Sung, M. Zhang, Y. Qi, and G. L. Verdine. Sequence-dependent structural variation in DNA undergoing intrahelical inspection by the DNA glycosylase MutM. *J. Biol. Chem.*, 287:18044–18054, 2012.
- [34] D. S. Daniels, T. T. Woo, K. X. Luu, D. M. Noll, N. D. Clarke, A. E. Pegg, and J. A. Tainer. DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat. Struct. Mol. Biol.*, 11:714–720, 2004.
- [35] B. Demple and M. S. DeMott. Dynamics and diversions in base excision DNA repair of oxidized abasic lesions. *Oncogene*, 21:8926–8934, 2012.
- [36] P. Zhang, U. Herbig, F. Coffman, and M. W. Lambert. Non-erythroid α spectrin prevents telomere dysfunction after DNA interstrand cross-link damage. *Nucleic Acids Res.*, 41:5321–5340, 2013.
- [37] H. C. Box, E. E. Budzinski, J. B. Dawidzik, J. S. Gobey, and H. G. Freund. Free radical-induced tandem base damage in DNA oligomers. *Free Radical Biol. Med.*, 23:1021–1030, 1997.
- [38] M. Fuxreiter, N. Luo, P. Jedlovszky, I. Simon, and R. Osman. Role of base flipping in specific recognition of damaged DNA by repair enzymes. *J. Mol. Biol.*, 323:823–834, 2002.

Seznam publikací/Selected publications

- [1] Tomáš Dršata, Marie Zgarbová, Petr Jurečka, Jiří Šponer, and Filip Lankaš. On the use of molecular dynamics simulations for probing allostery through DNA. *Biophys. J.*, 110:874-876, 2016.
- [2] Tomáš Dršata and Filip Lankaš. Multiscale modelling of DNA mechanics. *J. Phys.: Condens. Matter*, 27:323102, 2015.
- [3] Korbinian Liebl, Tomáš Dršata, Filip Lankaš, Jan Lipfert, and Martin Zacharias. Explaining the striking difference in twist-stretch coupling between DNA and RNA: A comparative molecular dynamics analysis. *Nucleic Acids Res.*, 43:10143-10156, 2015.
- [4] Mahmut Kara, Tomáš Dršata, Filip Lankaš, and Martin Zacharias. Effect of O6-guanine alkylation on DNA flexibility studied by comparative molecular dynamics simulations. *Biopolymers*, 103:23-32, 2015.
- [5] Elise Dumont, Tomáš Dršata, Célia F. Guerra, and Filip Lankaš. Insights into the structure of intrastrand cross-link DNA lesion-containing oligonucleotides: G[8-5m]T and G[8-5]C from molecular dynamics simulations. *Biochemistry*, 54:1259-1267, 2015.
- [6] Tomáš Dršata, Marie Zgarbová, Naďa Špačková, Petr Jurečka, Jiří Šponer, and Filip Lankaš. Mechanical model of DNA allostery. *J. Phys. Chem. Lett.*, 5:3831-3835, 2014.
- [7] Tomáš Dršata, Naďa Špačková, Petr Jurečka, Marie Zgarbová, Jiří Šponer, and Filip Lankaš. Mechanical properties of symmetric and asymmetric DNA A-tracts: implications for looping and nucleosome positioning. *Nucleic Acids Res.*, 42:7383-7394, 2014.
- [8] Tomáš Dršata, Mahmut Kara, Martin Zacharias, and Filip Lankaš. Effect of 8-oxoguanine on DNA structure and deformability. *J. Phys. Chem. B*, 117:11617-11622, 2013.
- [9] Tomáš Dršata and Filip Lankaš. Theoretical models of DNA flexibility. *WIREs Comput. Mol. Sci.*, 3:355-363, 2013.
- [10] Chandan Patel, Tomáš Dršata, Filip Lankaš, and Elise Dumont. Structure, dynamics, and interactions of a C4'-oxidized abasic site in DNA: A concomitant strand scission reverses affinities. *Biochemistry*, 52:8115-8125, 2013.

- [11] Tomáš Dršata, Alberto Pérez, Modesto Orozco, Alexandre V. Morozov, Jiří Šponer, and Filip Lankaš. Structure, stiffness and substates of the Dickerson-Drew dodecamer. *J. Chem. Theory Comput.*, 9:707-721, 2013.

Curriculum vitae

Tomáš Dršata

Born on 5th July 1988 in Jihlava, Czech Republic

Nationality: Czech

Martial status: single

Address: V Občanském domově 3, 140 00, Praha 4

Phone: (+420) 732 820 729 **E-mail:** tomas.drsata@gmail.com

Education

- Since 2012: Ph.D. student, Modelling of Chemical Properties of Nano- and Biostructures, Faculty of Science, Charles University, Prague. Thesis: Modelling mechanical properties of RNA and DNA (supervisor Dr. Filip Lankaš). Anticipated graduation: November 2016.
- Additional optional courses, Faculty of Mathematics and Physics, Charles University, Prague: Thermodynamics and statistical physics I (NTMF043, 2012/2013), Thermodynamics and statistical physics II (NTMF044, 2012/2013), Probabilistic methods in physics (NOFY062, 2013/2014). All exams passed.
- 2010-2012: Master studies, Modelling of Chemical Properties of Nano- and Biostructures, Faculty of Science, Charles University, Prague. Diploma thesis: Design, parametrization and verification of a coarse-grained model of DNA (supervisor Dr. Filip Lankaš). Mgr. title - Euro-master with honours (study average: 1.0, diploma thesis: 1.0, final state examination: 1.0).
- General physics (2 years of bachelor study, study average: 1.33), Faculty of Mathematics and Physics, Charles University, Prague, September 2010 to June 2012.
- 2007-2010: Bachelor studies, Chemistry, Faculty of Chemical Engineering, University of Chemistry and Technology, Prague. Bachelor thesis: Influence of ionic environment on the structure, dynamics and mechanical properties of DNA (supervisor Dr. Filip Lankaš). Bc. title - Chemistry Eurobachelor (study average: 1.28, bachelor thesis: 1.0, final state examination: 1.5).

International experience

- One semester DAAD scholarship, group of prof. M. Zacharias, Department of Physics, Technische Universität München, Germany. From 1st October 2015 to 31st January 2016.
- Visiting research assistant, group of prof. M. Zacharias, Department of Physics, Technische Universität München, Germany. From 23rd September to 15th October 2013 and from 6th October to 17th October 2014.

Scientific interests

Multiscale modelling of structural and mechanical properties of DNA and RNA macromolecules. Using large scale computer simulations of molecular dynamics to construct models of DNA and RNA with help of statistical mechanics. Application to problems in molecular biology and nanotechnology.

Teaching

- Winter semester 2014/2015: Assistant, course Mathematical methods for physical chemistry (N403045, course tutors: prof. J. Kolafa, Dr. F. Lankaš, Dr. A. Malijevský), Department of Physical Chemistry, University of Chemistry and Technology, Prague. Preparation and evaluation of exercises, general assistance.
- Since 2015: Assistance for diploma students (Jiří Reiniš and Bc. Hana Dohnalová, tutor Dr. Filip Lankaš), University of Chemistry and Technology, Prague.

Grants and Awards

- DAAD scholarship (German Academic Exchange Service), Research Grants - Short-Term Grants, 2015 (57130097). Intended for stay at TU Munich, October 2015 - January 2016
- Outstanding student presentation award, Conference: XIV Discussions in Structural Molecular Biology, Nové Hrady, Czech Republic, 17.3 - 19.3.2016.
- 2014-2016: Team member, Nanomechanics of RNA and DNA structural motifs, Grant Agency of the Czech Republic (14-21893S, principal investigator Dr. Filip Lankaš)

- 2013-2015: Principal investigator, Modelling allosteric effects in DNA, Grant Agency of the Charles University (584213)

Presentations

- **Dršata, T.** and Lankaš F. (2016, March). *Computer simulations of nucleosome positioning and allosteric effects in DNA*. Talk given at the XIV Discussions in Structural Molecular Biology, Nové Hrady, Czech Republic.
- **Dršata, T.** and Lankaš F. (2015, October). *Applying a coarse-grained model to investigate structure and dynamics of oxidatively generated DNA lesions*. CECAM workshop "DNA damages: modeling and rationalize structure and reactivity", CECAM-FR-RA, Centre Blaise Pascal, ENS Lyon, France.
- **Dršata, T.** and Lankaš F. (2014, April). *Mechanical Properties of DNA A-tracts*. Poster presented at the CECAM workshop "Intersection of simulative and experimental approaches for the investigation of natural and artificial DNA nanostructures", Lausanne, Switzerland.

Languages

- Cambridge English Level 2 Certificate, FCE, Grade A (Europe Level C1)