

Jihočeská univerzita v Českých Budějovicích

Přírodovědecká fakulta



Počítačové modelování interakcí molekul s minerálními povrchy

Bakalářská práce

Hana Barvíková

Školitel: RNDr. Milan Předota, Ph.D.

České Budějovice 2012

Barvíková H., 2012: Počítačové modelování interakcí molekul s minerálními povrchy.

[Computer simulation of interactions of molecules with mineral surfaces. Bc. Thesis, in Czech.] – 85 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Abstract

This thesis is focused on molecular dynamics, modelling interactions and their simulation. One of the tasks was to familiarize with GROMACS, a molecular dynamics simulation software. This work also introduces the basics of molecular dynamics and modelling interactions of organic molecules with mineral surfaces. The aim of the thesis was to solve model tasks in GROMACS and analyze the output results.

The thesis describes some of the most important file formats and utilities that are needed for working with GROMACS and the use of both the formats and the utilities. In this program, I built up several systems consisting of combinations of organic molecules such as benzoic acid, phenol, salicylic acid and their conjugate bases with mineral surfaces (quartz crystal) of different surface charge density. Furthermore I analyzed the results of these simulations, behaviour of the structures and adsorption geometries and interaction energies within these systems.

The thesis might also serve as a quick introduction and familiarization with GROMACS with emphasis on simulations and analysis of systems with planar interfaces.

Within the MetaCentrum Project I worked on the Hermes Computer Cluster belonging to the Faculty of Science at the University of South Bohemia.

Prohlašuji, že svoji bakalářskou práci jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

.....
Datum

.....
Hana Barvíková

Ráda bych na tomto místě poděkovala RNDr. Milanu Předotovi, Ph.D., za vedení bakalářské práce, veškerou pomoc a trpělivý přístup. Za dodání vstupních topologií a konfigurací děkuji RNDr. Babaku Minofarovi, MSc. Ph.D., RNDr. Milanu Předotovi, Ph.D, a Ing. Ondřeji Kroutilovi. Za odborné rady a připomínky ke své práci vděčím Ing. Ondřeji Kroutilovi a Mgr. Pavlu Fibichovi. Za poskytnutí výpočetních prostředků a datového úložiště patří dík Ústavu fyziky a biofyziky Přírodovědecké fakulty Jihočeské univerzity v Českých Budějovicích a MetaCentru. V neposlední řadě děkuji své rodině za porozumění a podporu, které se mi dostalo během psaní.

Obsah

Úvod	3
1. Současný stav	4
1.1. Počítačové simulace a molekulové modelování	4
1.2. Teorie molekulární dynamiky	5
1.2.1. Silová pole, potenciálové funkce	5
1.2.2. Parametry silového pole, rodiny silových polí	7
1.2.3. Potenciální energie a jednotlivé složky potenciálové funkce	8
1.2.3.1. Potenciály vazebné	8
1.2.3.2. Potenciály nevazebné	11
1.2.4. Vlastní MD výpočet a integrace pohybových rovnic	13
1.3. Schéma MD simulace a některé technické detaily	15
1.3.1. Solvatace	16
1.3.2. Periodické okrajové podmínky a cutoff	16
1.3.3. Minimalizace energie	17
1.3.4. Ekvilibrace – termostatování a barostatování	18
2. Seznámení s programem GROMACS	20
2.1. Práce s utilitami	20
2.2. Formáty použitých vstupních a výstupních souborů	24
2.2.1. Příprava simulace	24
2.2.1.1. Vstupní souřadnice	24
2.2.1.2. Topologie	25
2.2.1.3. Parametrické soubory	27
2.2.2. Spuštění simulace	32
2.2.3. Analýza výsledků	32
2.2.3.1. Energetický soubor	32

2.2.3.2. Textový soubor	33
2.2.3.3. Indexový soubor	34
2.2.3.4. Trajektorie	35
3. Cíle práce	37
4. Metodika	38
4.1. Postup při přípravě systémů	38
4.2. Fáze a parametry simulací	42
4.3. Přenos souborů	43
5. Výsledky	44
5.1. Seznámení s modelovanými systémy	44
5.1.1. Křemenný povrch	45
5.1.2. Molekuly	47
5.2. Molekula kyseliny benzoové na neutrálním křemenném povrchu	50
5.3. Shrnutí výsledků simulovaných systémů	63
5.3.1. Kyselina benzoová a benzoát	63
5.3.2. Fenol a fenolát	67
5.3.3. Kyselina salicylová	71
5.3.4. Interakční energie molekul s povrchy	75
6. Závěr	77
7. Seznam použité literatury	78
7.1. Internetové a jiné zdroje	78
7.2. Další užitečné odkazy	80
7.2.1. Stránky počítačových klastrů UFY a Hermes	80
8. Přílohy	81

Úvod

Tato bakalářská práce, jak již název napovídá, se zabývá molekulární dynamikou, modelováním interakcí molekul s povrchy a počítačovými simulacemi ve fyzice mnoha částic. Téma počítačových simulací je jistě velmi přitažlivým a obecně se rychle rozvíjícím a perspektivním odvětvím moderní vědy. Tato technika je v dnešní době hojně využívána v mnoha oblastech výzkumu a výroby. Jedná se o interdisciplinární obor, který v sobě spojuje poznatky fyziky, chemie, biologie, matematiky a výpočetní fyziky. Člověku dává *in silico* nahlédnout na chování a vlastnosti systémů na atomární úrovni. Numerické simulace obecně se však netýkají jen malých měřítek; molekulární dynamika je v současné době s úspěchem uplatňována při řešení rozmanitého spektra problémů od studia atomů až po kosmologii a astronomii. V nadneseném slova smyslu lze tedy říci, že tato obsáhlá kapitola současné vědy v mnoha ohledech propojuje mikrokosmos s makrokosmem.

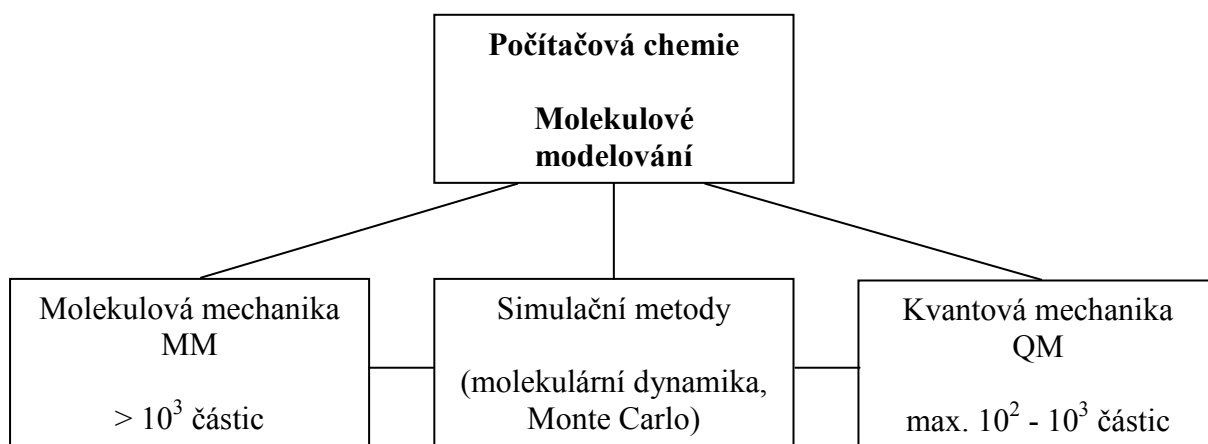
Simulace provedené v praktické části této bakalářské práce jsou součástí projektu nazvaného *Studium struktury a dynamiky minerálních povrchů a biomembrán a jejich interakcí s organickými a anorganickými ligandy pomocí počítačového modelování*. Projekt se mj. zabývá interakcí huminových kyselin, organické hmoty a jejich dílčích částí (včetně organických molekul studovaných v této práci) s minerálními povrchy, konkrétně křemenem.

Práce je strukturována do několika hlavních tematických okruhů. V úvodu se snaží pochopitelným způsobem nastínit základy molekulární dynamiky. Dále seznamuje s prací v programu GROMACS, určeným pro potřeby počítačového modelování a simulací. Nejdůležitější částí je samotné provedení zadaných simulací několika organických molekul (kyseliny benzoové, fenolu, kyseliny salicylové a deprotonovaných forem těchto molekul) s křemenným povrchem o různé hustotě povrchového náboje a zhodnocení výsledků těchto simulací. Snahou autorky bylo pojmout svou práci jako jakési vodítko pro řešitele podobných úloh s ohledem na usnadnění jejich orientace v problematice modelování interakcí molekul s povrchy.

1. Současný stav

1.1. Počítačové simulace a molekulové modelování

Molekulové modelování v současnosti zahrnuje dva základní přístupy (viz Obr. 1) postavené na dvou odlišných teoretických modelech: přístup molekulově mechanický (MM) a kvantově mechanický (QM). Oblast počítačově podporovaného molekulového modelování – simulační metody mohou využívat a kombinovat oba modely, tedy MM i QM. [6] Ve své práci se zabývám pouze MM přístupem. Ten dovoluje simulovat systémy s obecně větším počtem částic (řádově $> 10^3$). Z důvodu velkého množství částic v modelovaném systému však není možné popsat vlastnosti takto komplexního systému analyticky. Molekulární dynamika se s těmito obtížemi vyrovnává využitím numerických metod. U dlouhých simulací však dochází ke kumulaci chyb vznikajících při numerické integraci, které mohou být částečně eliminovány vhodným výběrem algoritmů a parametrů. V případě MM se tedy jedná o výpočetně méně náročnou metodu, která však vyžaduje množství aproximací. QM je přesnější, ale také výpočetně náročnější metoda vhodná pro systémy řádově do 10^2 částic. [6], [7]



Úkolem počítačových simulací je generování konfigurací systému mnoha částic a tato data pak využít ke stanovení a studiu termodynamických či strukturních veličin. Počítačové simulace využívají dvě základní simulační metody: metodu Monte Carlo (MC) a molekulární dynamiku (MD). Metoda MC je stochastická metoda, která generuje konfigurace s ohledem na výpočet středních hodnot na základě generátoru náhodných čísel. [1] Tato práce se vztahuje výhradně ke druhé uvedené metodě.

Molekulární dynamika je technika počítačové simulace vycházející ze zákonů klasické mechaniky – Newtonových pohybových rovnic. Jejím hlavním úkolem je sledovat pohyb a chování částic v čase; její smysl spočívá v generování trajektorie těchto částic. Na rozdíl od MC jde o deterministickou metodu, počáteční konfigurace dané množiny částic tedy zcela určuje časový vývoj celého systému. Jedná se o metodu statistické mechaniky. Trajektorie sledovaných částic je popsána integrací pohybových rovnic atomů a molekul uvnitř systému, kde síly působící mezi částicemi a potenciální energie jsou dány použitým silovým polem. [8]

Metoda MD byla koncipována v rámci teoretické fyziky v padesátých letech dvacátého století. Dnes je využívána především při vývoji nových materiálů a v oblastech biofyziky a biochemie při modelování biomolekul (proteiny, nukleové kyseliny RNA/DNA, membrány). Počítačové simulace obecně často slouží jako alternativa k reálným experimentům, které jsou těžko proveditelné v praxi např. z důvodu vysokých tlaků, nebezpečí exploze, nebo z důvodu finanční náročnosti, a snižují tak náklady za drahé a složitě realizovatelné laboratorní experimenty. Zároveň zajišťují bezpečnost při práci s toxickými nebo radioaktivními látkami a umožňují předpovídat chování a vlastnosti látek nových. Své uplatnění nacházejí ve farmaceutickém průmyslu nebo při vývoji materiálů pro mikroelektroniku. MD se rovněž stala důležitým experimentálním nástrojem při výzkumu v oblasti nanotechnologií, vlastností povrchů, adheze a tření mezi pevnými látkami či zkoumání defektů. V neposlední řadě umožňuje studium tekutin (kapalin i plynů) a s nimi spjatých termodynamických a transportních jevů. [7], [9]

1.2. Teorie molekulární dynamiky

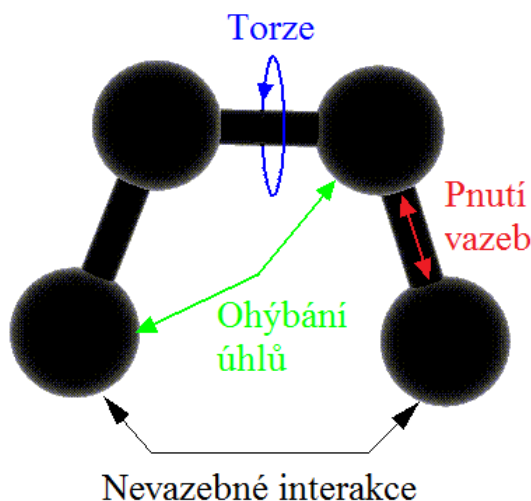
Před samotnou MD simulací je třeba vytvořit model fyzikálního systému, který lze popsat soustavou parametrů. Jedním z takových důležitých měřítek pro popis modelovaného systému je silové pole.

1.2.1. Silová pole, potenciálové funkce

Molekulární dynamika vychází z principů molekulární mechaniky, pro kterou je charakteristická představa atomů jako tuhých kulových těles s van der Waalsovým poloměrem (který aproximuje velikost atomu) a bodovým nábojem ve svém středu. Vzájemné interakce mezi atomy jsou znázorňovány pomocí těles (atomů) spojených

harmonickými pružinami (vazbami), které jsou charakterizované tuhostí k . Komplexním popisem interakcí mezi atomy v systému je silové pole, které je utvářeno jednotlivými potenciálovými funkcemi.

Potenciálové funkce popisují interakce mezi atomy a zahrnují součet všech intermolekulárních a intramolekulárních interakcí daného typu molekuly. Jsou vyjádřeny soustavou rovnic.

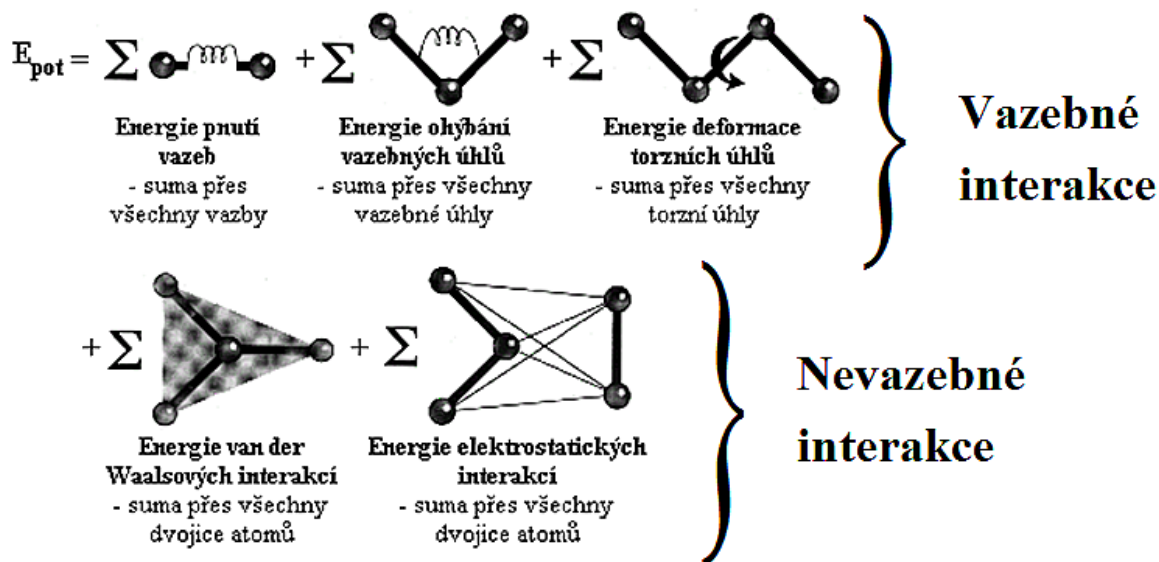


Obr. 2: Grafické znázornění vazebných a nevazebných interakcí [10]

Celková potenciální energie molekuly je dána vnitřní konformací (prostorové uspořádání atomů v molekule) molekuly a její polohou vzhledem k ostatním částicím. Vzájemné působení mezi vázanými atomy vyjadřují tzv. vazebné interakce. Silové pole definuje ideální délky jednotlivých vazeb, vazebných a torzních úhlů (úhly mezi atomy oddělenými třemi vazbami; úhly mezi dvěma rovinami vzniklé otáčením kolem jedné společné vazby). Čím více se skutečné hodnoty liší od hodnot definovaných silovým polem, tím vyšší energii se vyznačuje daná interakce. Nízká potenciální energie vypovídá o stabilitě molekuly. Umožňuje tak stanovit nejstabilnější a tedy nejpravděpodobnější molekulovou konformaci. Druhým typem interakcí jsou nevazebné interakce, které popisují vzájemné působení každých dvou atomů, které není popsáno vazebnými potenciály, tedy jak v rámci jedné molekuly (typicky atomy vzdálené více než 2-3 vazbami), tak mezi atomy různých molekul. Nevazebné interakce se rozdělují na elektrostatické (určují pro každou dvojici atomů, jak na sebe vzájemně působí náboje těchto atomů) a neelektrostatické. [2], [8]

Před samotnou MD simulací se provádí minimalizace energie z důvodu stabilizace a optimalizace systému, odstranění největších repulzí mezi atomy, které by jinak mohly vést

k velkým a prudkým pohybům atomů a narušení stability simulace. Kinetická energie částic související s teplotou je při minimalizaci zanedbána, neboť výpočet probíhá pro teplotu absolutní nuly. [11]



Obr. 3: Výpočet potenciální energie se znázorněním příspěvků jednotlivých interakcí [8]

1.2.2. Parametry silového pole, rodiny silových polí

Jednotlivé atomy jsou rozděleny na atomové typy podle svých fyzikálně-chemických vlastností – chemické značky, vaznosti a geometrického uspořádání a typu dalších vázaných atomů. Silové pole definuje kromě potenciálových funkcí i soustavu parametrů pro každý typ atomu: relativní atomovou hmotnost, van der Waalsův poloměr, náboj, rovnovážnou délku vazeb a velikost vazebných a torzních úhlů a silové konstanty. Sestavení silového pole vychází z empirických metod (na základě experimentálních dat), nebo neempirických technik (z řešení stacionární Schrödingerovy rovnice pomocí kvantově-chemických výpočtů), případně vhodné kombinace obou přístupů. Údaje týkající se malých molekul (např. délky vazeb, velikosti vazebných úhlů, silové konstanty) pocházejí často z výsledků rentgenové krystalografie a spektroskopických měření. [2], [8], [12]

Existují rodiny různých silových polí vyvinutých pro rozdílné účely. Nejznámější z nich jsou uvedeny spolu s nejčastějším použitím v následujícím stručném přehledu¹. [13], [14], [15], [16], [17], [18]

¹ Příklad silového pole použitého v rámci praktické části této bakalářské práce je uložen na přiloženém CD.

GROMOS – biomolekulární systémy; peptidy, sacharidy, nukleotidy

OPLS – vyvinuto pro potřeby simulací tekutin, používá se k simulacím buněčných membrán

AMBER – simulace peptidů a nukleových kyselin (RNA, DNA)

CHARMM – proteiny, lipidy, nukleové kyseliny

MARTINI – proteiny a buněčné membrány

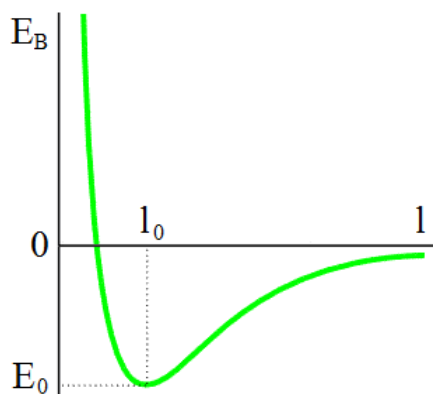
1.2.3. Potenciální energie a jednotlivé složky potenciálové funkce

1.2.3.1. Potenciály vazebné

Popisují interakce mezi atomy jedné molekuly vázanými jednoduchou či násobnou vazbou.

Pnutí vazeb

Energie pnutí vazeb představuje interakci mezi dvěma atomy spojenými kovalentní vazbou² (jednoduchou či násobnou). Závislost potenciální energie chemické vazby na délce této vazby je dána křivkou uvedenou na obrázku (Obr. 4). Uvedený graf mimo jiné vyjadřuje skutečnost, že pokud jsou dva atomy příliš blízko u sebe, vzájemně se odpuzují a hodnota potenciální energie této interakce je vysoká kladná. Na takovéto přiblížení atomů je nutné dodat energii. Naopak pokud jsou atomy od sebe vzdáleny více než na vzdálenost l_0 , vzájemně se přitahují a na další zvětšení jejich vzájemné vzdálenosti je opět potřeba dodat energii. Hodnota l_0 představuje ideální délku vazby, při které je energie pnutí této vazby E_B (bond = vazba) minimální. E_0 odpovídá disociační energii vazby. V tomto ideálním stavu je nutné dodat energii jak na další přiblížení atomů, tak na jejich odtržení.

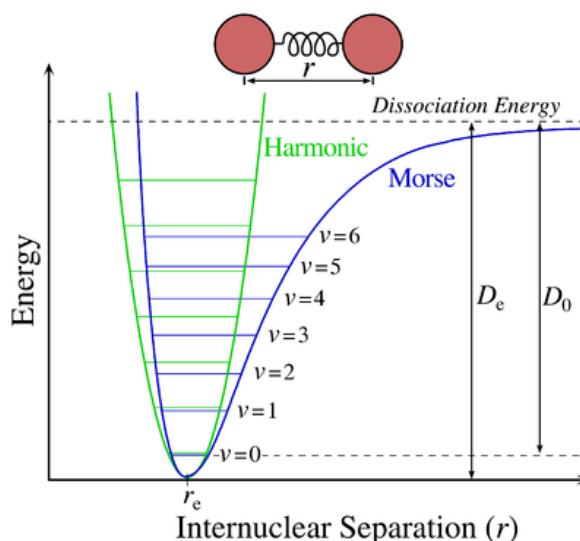


Obr. 4: Závislost potenciální energie na délce vazby [8]

² Chemická vazba charakteristická sdílením elektronových párů mezi atomy [19]

Výše uvedenou závislost lze popsat pomocí několika potenciálů: harmonického, Morseho, kubického či potenciálu FENE [13]. Poslední dva jmenované jsou složité a náročné na výpočet, podobně jako Morseho potenciál, který nejlépe ze všech vystihuje skutečnou závislost. Jeho výpočet však vyžaduje znalost 3 parametrů pro každý typ vazby. Harmonický potenciál je sice méně přesným modelem než Morseho funkce, pro modelování jsou však podstatné hodnoty E_B pro l blízké l_0 a v této oblasti se oba potenciály liší velmi málo (viz Obr. 5). Harmonický potenciál je odvozený z Hookova zákona a vyjadřuje hodnotu potenciální energie vazby jako funkci výchylky od ideální délky této vazby l_0 . Silová konstanta k_B reprezentuje sílu vazby. Oba parametry závisí na typu interagujících atomů a násobnosti vazby. [8]

$$V_B(l) = \frac{1}{2}k_B(l - l_0)^2 \quad [8]$$



Obr. 5: Srovnání harmonického a Morseho potenciálu [19]

Ohýbání úhlů

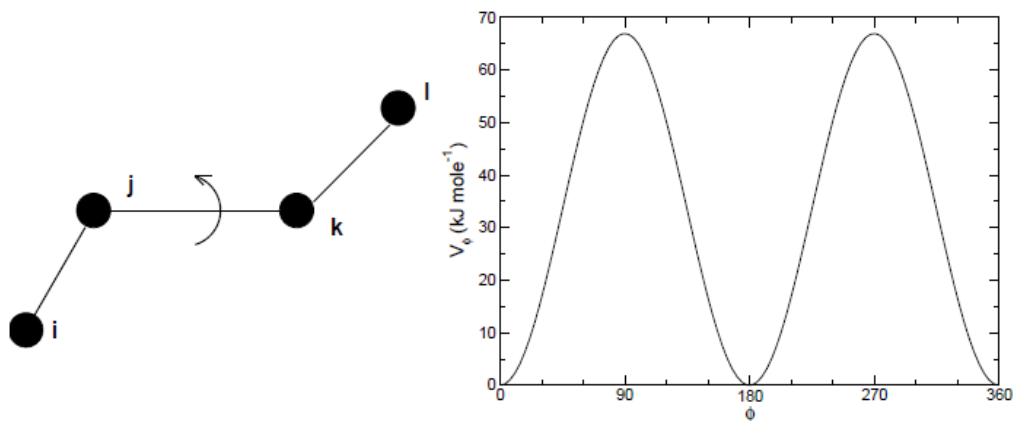
Energie ohýbání úhlů souvisí se změnou velikosti vazebných úhlů. Charakterizuje odchylku od ideální geometrie, která by v optimalizované struktuře měla být blízká nule. Pro reálný vazebný úhel existuje závislost potenciální energie E_A (angle = úhel) této interakce na velikosti úhlu ϕ , kterou lze podobně jako v předchozím případě vyjádřit pomocí několika potenciálů: harmonického, kvadratického, kubického, kosinového nebo Urey-Bradleyho [2], [8]. Často se používá k popisu této závislosti právě harmonický potenciál:

$$V_A(\phi) = \frac{1}{2}k_A(\phi - \phi_0)^2 \quad [13], [8]$$

Hodnoty ϕ_0 (ideální vazebný úhel) a k_A (konstanta příslušející vazebnému úhlu mezi atomy určitých typů) závisí na typu atomů, které tvoří daný úhel. [8]

Torze

Závislost potenciální energie E_T (torsion = rotace, otočení) na velikosti torzního úhlu θ je charakterizována křivkou, jejíž tvar závisí na typu vazby a symetrii ligandů³, např. na Obr. 6 pro dvojnou vazbu:



Obr. 6: Rotace kolem vazby j-k, závislost potenciální energie na torzním úhlu. Uvedená konformace odpovídá úhlu $\theta = k \cdot 180^\circ$, kde $k = 0, 1, 2, \dots$ [13]

Torzní úhel (někdy také nazývaný dihedralní) je úhel mezi atomy oddělenými třemi vazbami (i-j, j-k, k-l), nebo také úhel mezi dvěma rovinami (ijk, jkl) vzniklý otáčením kolem jedné společné vazby (j-k). Potenciální energie je reprezentována periodickou funkcí – rozvojem kosinových funkcí:

$$V_T(\theta) = k_T(1 + \cos(n\theta - \theta_0)) \quad [8]$$

Konstanta k_T přísluší torznímu úhlu θ_0 pro určitou kombinaci atomů, hodnota n vyjadřuje multiplicitu – celé číslo, které udává počet minim funkce $V_T(\theta)$. [8] Dalšími potenciály pro vyjádření energie torzí jsou Fourierův či Ryckaert-Bellemansův. [2]

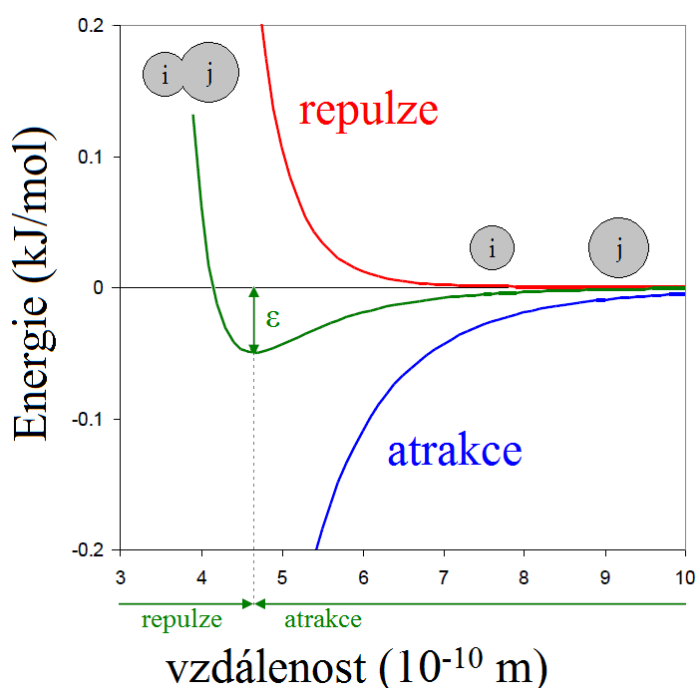
³ ligand = koordinovaná skupina iontů nebo molekul

1.2.3.2. Potenciály nevazebné

Potenciály nevazebné popisují vzájemné působení dvou atomů, které nejsou vázány přímou vazbou. Charakterizují interakce jak mezi atomy v rámci jedné molekuly, tak mezi atomy samostatných molekul.

Energie van der Waalsovských interakcí

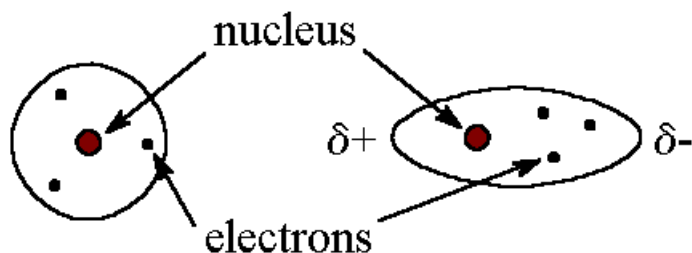
Van der Waalsovské (vdW) interakce jsou výsledkem odpuzivých a přitažlivých sil mezi elektroneutralními molekulami či atomy. Graf závislosti potenciální energie vdW interakcí E_{LJ} na vzdálenosti dvou atomů i a j je zelenou křivkou znázorněn na Obr. 7.



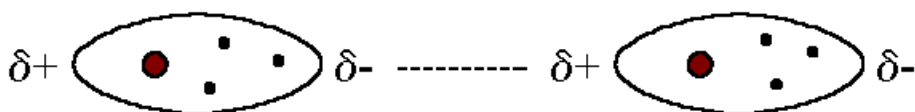
Obr. 7: Závislost energie vdW interakcí na vzdálenosti atomů. Přitažlivé (atrakce) a odpuzivé (repulze) síly [20] – editováno dle [8] a [21]. Schématicky.

K odpuzování atomů (repulzi) dochází v případě, že se atomy přiblíží na takovou vzdálenost, kdy na sebe odpuzivě působí elektrony v obalech těchto atomů. Přitažlivá síla (atrakce) vzniká jako důsledek fluktuační rozložení elektronů v atomu, kdy díky náhodnému nahromadění více elektronů vzniká okamžitý dipól, který indukuje dipól v dalším blízkém atomu či molekule (Londonův efekt, Londonova disperze). Tento jev se uplatňuje u nepolárních atomů či molekul. V důsledku působení polární molekuly na nepolární částici dochází k podobnému jevu, zvanému Debyeův. Síla mezi dvěma permanentními dipóly se nazývá Keesomova. [2], [8], [11], [22], [23]

1. Fluktuační rozložení elektronů v obalu a vznik okamžitého dipólu



2. Indukování dipólu a následná atrakce atomů



Obr. 8 a 9: Londonova disperze [11]

Van der Waalsovy interakce bývají často modelovány pomocí Lennard-Jonesovy 12-6 funkce:

$$V_{LJ}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad [13]$$

Hladina minimální energie E_{LJ} a optimální vzdálenost atomů r_{ij} závisí na parametrech silového pole těchto atomů. Geometrický parametr σ vyjadřuje vzdálenost, ve které je potenciál nulový. Parametr ϵ určuje hloubku potenciálového minima. První člen rozdílu umocněný číslem 12 zastupuje repulzi, druhý člen s faktorem 6 reprezentuje fyzikálně realistickou atrakci atomů (vdW síly ubývají se šestou mocninou vzdálenosti). Lépe než výše uvedená funkce vystihuje vdW interakci méně používaný Buckinghamův potenciál. [8], [12]

Elektrostatické interakce

U nabitých částic dochází vlivem působení elektrostatických sil k odpuzování souhlasně nabitých a přitahování navzájem opačně nabitých částic. V porovnání s vdW jde o síly s dlouhým dosahem. [8], [21] Potenciální energie E_C těchto interakcí vychází z Coulombova zákona:

$$V_C(r_{ij}) = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{q_i q_j}{r_{ij}} \quad [13]$$

Konstanta ε_0 ve vzorci označuje permitivitu vakua, ε_r je relativní permitivita prostředí. Vzájemnou vzdálenost interagujících atomů popisuje parametr r_{ij} a jejich náboje nesou označení q_i a q_j .

Celková potenciální energie systému je dána součtem všech potenciálních energií. Poslední člen potenciálové funkce (v příkladu potenciálové funkce níže) reprezentuje ne vazebné interakce (vdW a elektrostatické) mezi všemi atomovými páry v systému. Uvedená funkce je alternativním zápisem potenciálů popsanych výše. Příklady často používaných silových polí jsou uvedeny v kapitole 1.2.2.

$$V(r^N) = \sum_{bonds} \frac{1}{2} k_b (l - l_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{torsions} \frac{1}{2} V_n [1 + \cos(n\omega - \gamma)] +$$

$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^N \left\{ \varepsilon_{i,j} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$

Příklad potenciálové funkce (silové pole AMBER). [14]

1.2.4. Vlastní MD výpočet a integrace pohybových rovnic

V systému o určitém počtu atomů a molekul, který je popsán pomocí silového pole (potenciálová funkce s parametry), lze předpokládat, že každý atom se bude pohybovat jistou rychlostí. Pro studované atomy je nutno vyřešit Newtonovy pohybové rovnice. Z poloh, rychlostí atomů a sil působících v systému v čase t lze určit následující konfiguraci v čase $t+\delta t$, kde δt je tzv. integrační krok, dostatečně krátký časový interval, zpravidla omezený vibrační vazeb (obvykle se volí 1 až 2 fs).

Celková energie každé molekuly je tvořena součtem její kinetické (dána pohybovým stavem molekuly) a potenciální energie (závisí na konformaci a poloze). Tyto dvě složky se v důsledku pohybu atomů přeměňují jedna v druhou. Vlivem vzájemného působení dochází ke změně souřadnic a tím i změně potenciální energie. Popsaný přístup předpokládá práci s mnoha částicemi a nedovoluje řešit pohybové rovnice analyticky. Pohybové rovnice částic jsou integrovány s využitím metody konečných diferencí. Integrace je rozdělena do mnoha malých kroků, oddělených časovým úsekem δt . Z výsledné síly působící na každou částici v čase t (dané vektorovým součtem všech interakčních sil částice) je vypočítáno zrychlení. Na základě tohoto zrychlení, rychlosti a polohy částice v čase t je vypočtena rychlost a poloha v čase $t+\delta t$. [8]

Rovnice:

$$F = \frac{dp}{dt}$$

je obecným vyjádřením 2. Newtonova zákona (zákonu síly). Síla působící na těleso – částici je dána změnou její hybnosti v čase. Uvedenou funkci lze rovněž přepsat na:

$$F = \frac{d(mv)}{dt} = m \frac{d^2r}{dt^2}$$

kde m je hmotnost částice a v její rychlost. Zrychlení je dáno druhou derivací polohy (nebo první derivací rychlosti) podle času. Pohybové rovnice mohou být aproximovány pomocí Taylorova rozvoje a následně zpracovány pomocí dostupných algoritmů. Nejrozšířenější integrační metodou je Verletův algoritmus, který k výpočtu polohy částice v čase $t+\delta t$ využívá analogii s rovnoměrně zrychleným přímočarým pohybem:

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) \quad \text{poloha částice v následujícím kroku } (t+\delta t)$$

$$r(t - \delta t) = r(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t) \quad \text{poloha částice v předchozím kroku } (t-\delta t)$$

Po sečtení obou rovnic:

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \quad \text{Verletův algoritmus}$$

Poslední kvadratický člen funkce je nejnáročnější na výpočet. Bývá omezen sférickým oříznutím potenciálu (cutoffem, vysvětleno dále). Výhodou Verletova algoritmu je jeho snadná implementace a malé požadavky na paměť. Díky uvažování předchozí konfigurace je také tzv. reversibilní v čase. K výpočtu následující polohy částice je však nutné znát její předchozí pozici, což se u prvního kroku výpočtu, pro který není dostupná informace z času $t-\delta t$, řeší zkráceným Taylorovým polynomem:

$$r(-\delta t) = r(0) - \delta t v(0) + \frac{1}{2} \delta t^2 a(0)$$

Rovněž chybí explicitní vztah pro výpočet rychlosti. Ty jsou odhadovány až po výpočtu pozice částice v dalším kroku:

$$v(t) = \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \quad \text{nebo} \quad v(t + \frac{1}{2}\delta t) = \frac{r(t + \delta t) - r(t)}{\delta t}$$

Dalšími modifikacemi Verletova algoritmu jsou přeskokový leap-frog algoritmus nebo rychlostní Verletův algoritmus. Kromě těchto existuje složitější, výpočetně náročnější, zato však přesnější Beemanův algoritmus s přesnějším odhadem rychlostí a kinetické energie. Další metoda, Gearova, zatížená některými nedostatky se zachováním energie, spadá podle rozdělení integračního kroku na dvě části mezi metody typu prediktor-korektor. [1], [4], [8] Více o metodách integrace pohybových rovnic v literatuře [1] a [5].

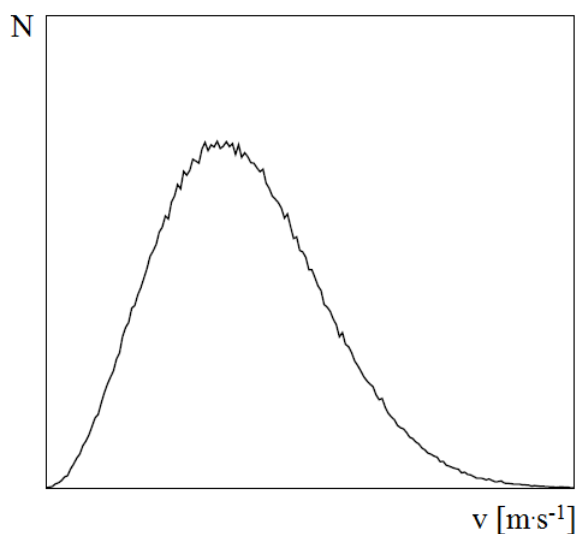
1.3. Schéma MD simulace a některé technické detaily

Průběh celé simulace lze shrnout do následující posloupnosti několika kroků:

generace počátečního stavu → **minimalizace** → **ekvilibrace** → **vlastní MD** → **analýza**

Před spuštěním vlastního MD výpočtu je nutné přiřadit atomům počáteční rychlosti. Ty se obvykle volí s ohledem na Maxwell-Boltzmannovské rozložení odpovídající dané teplotě:

$$p(v_i) = \sqrt{\frac{m_i}{2\pi kT}} \exp\left(-\frac{m_i v_i^2}{2kT}\right) \quad [13]$$



Obr. 10: Maxwell-Boltzmannovské rozložení rychlostí [13]

Uvedený graf popisuje rozložení rychlostí částic v daném systému, kde N vyjadřuje počet částic. Jedná se o pravděpodobnostní rozložení, největší počet částic se pohybuje tzv. nejpravděpodobnější rychlostí (v maximu funkce).

1.3.1. Solvatace

V praktické části této bakalářské práce byly simulovány systémy skládající se z organické molekuly, křemenného povrchu a vodního solventu. Při přípravě systémů bylo provedeno před každou simulací zkombinování modelů povrchu a organické molekuly, které odpovídají zvolenému pH a odpovídajícímu protonovanému/deprotonovanému stavu. Volný prostor byl následně vyplněn (solvatován) molekulami vody. Více o tématu solvatace v literatuře [2] nebo v uživatelském manuálu programu GROMACS [13].

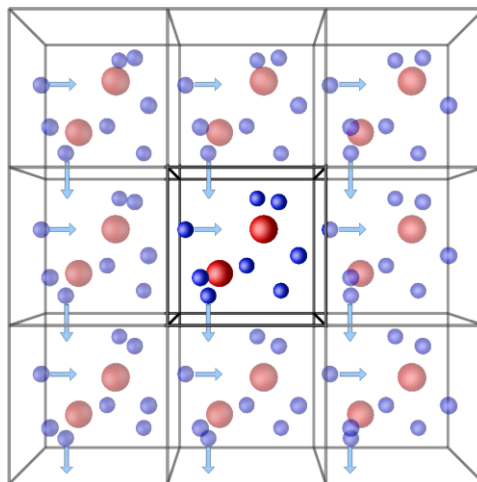
Následně byly přidány ionty za účelem vyrovnání náboje v systému a dosažení požadované koncentrace iontů (popsáno dále).

1.3.2. Periodické okrajové podmínky a cutoff

Molekulární dynamika umožňuje vzhledem ke své časové a výpočetní náročnosti pracovat pouze s omezeným počtem částic. Při jejich umístění do simulačního prostoru (boxu), nejčastěji tvaru krychle nebo kvádrů o určitém objemu, se budou některé částice vyskytovat v nehomogenním prostředí u hranic tohoto prostoru a reagovat s jeho stěnami. Při dostatečné velikosti pracovního boxu bude se stěnami interagovat relativně malá část molekul. Takový systém však předpokládá vysoký počet částic. Proto se zavádějí periodické okrajové podmínky (Periodic Boundary Conditions, PBC). Simulační box je v tomto případě obklopen svými kopiemi (ve dvourozměrném modelu 8, ve 3D podmínkách celkem 26 kopiemi; Obr. 11) a pokud některá částice na jedné straně box opustí, na protější straně se znovu objeví. Tak je zachován konstantní počet částic uvnitř boxu včetně celkové energie. Tato aplikace předpokládá zcela homogenní prostředí (např. nesmí být přítomen gradient teploty).

Z uvedených skutečností vyplývá kritérium pro velikost boxu, jehož nejkratší rozměr musí být minimálně dvakrát větší než radius cutoffu (cut off = oříznout, přerušit). Cutoff neboli oříznutí interakčního potenciálu představuje explicitní počítání nevazebných dlouhodosahových sil (reprezentovaných van der Waalovým a Coulombovým potenciálem) pouze na dvojice molekul (přesněji reziduí či nábojových skupin) vzdálených méně než je

poloměr oříznutí R_C^4 . Zároveň by měla platit podmínka, že každý z rozměrů boxu přesáhne součet nejdelšího rozměru největší makromolekuly a dvojnásobku cutoffu ($2R_C$). Tím je zajištěno, že molekula nebude interagovat se svým periodickým obrazem. [5], [13], [24]

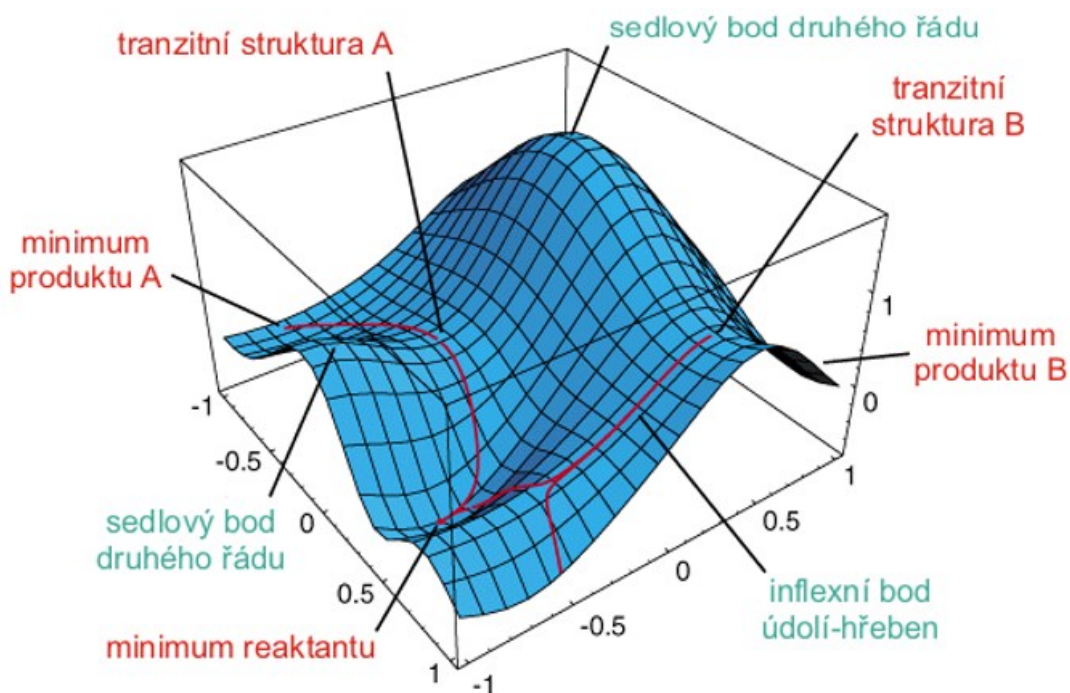


Obr. 11: Periodické okrajové podmínky [25]

1.3.3. Minimalizace energie

Minimalizace se provádí před samotnou MD simulací (např. z důvodu předešlé solvatace systému, která obvykle způsobí prostorové střety a nevhodnou geometrii molekul) k odstranění silných vdW repulzí vedoucích k deformacím struktury. Cílem procesu minimalizace je ustálení a optimalizace systému z energetického hlediska, které souvisí s uspořádáním a geometrií částic uvnitř systému. Při tomto procesu je prohledávána tzv. hyperplocha potenciální energie (Potential Energy Surface; PES), která vyjadřuje závislost energie částic na jejich souřadnicích. „Názorné zobrazení je možné jen pro jednoduché případy v 3D, nicméně matematickým postupům na více rozměrech vadí 'jen' drsně rostoucí nároky na výkon počítače“. [26]

⁴ Interakce vzdálenějších molekul je započtena pro vdW interakce implicitně dlouhodobou korekcí, pro Coulombické interakce metodou reakčního pole nebo složitější Ewaldovou sumací. [1], [12]



Obr. 12: Hyperplocha potenciální energie.

„Užitečnou analogií PES je pohoří. Molekuly pak mění svoji energii jako ti nejlínější turisté: nejvíce se jim líbí v dolíku, a když už musí někam jít, vyrazí do jiného dolíku, cestou nejmenších energetických nároků, přes nejnižší sedlo. Tato analogie v leččems pokulhává. Hlavně v tom, že PES je „pohoří“ v mnohorozměrném prostoru. Také v tom, že přibližování jader pod obvyklé „vazebné“ vzdálenosti vede k velmi špičatým a vysokým vrcholům (na které nikdo nelez, protože na to nemá) a naopak v místech, kde jádra jsou hodně vzdálená, přechází „pohoří“ ve fádni, rozlehlé pláně netknutých atomů.“ [26]

Pro minimalizaci energie je velmi často kromě metody konjugovaných gradientů a dalších technik využívána metoda největšího spádu (steepest descent). Při tomto postupu není uplatněn klasický molekulárně-dynamický výpočet. K nalezení lokálního minima potenciální energie poblíž počáteční struktury je použit algoritmus největšího spádu, což je metoda derivace prvního řádu. Minimalizace je ukončena, je-li splněno konvergenční kritérium – rozdíl energií předchozího a následujícího kroku je menší než stanovená hodnota. Je nutné provést dostatečný počet iterací, aby metoda konvergovala k minimu. Tehdy je iterování (postupné přibližování) ukončeno. [6], [8]

1.3.4. Ekvilibrace – termostatování a barostatování

Dalším krokem ke stabilizaci systému je ekvilibrace (ustálení, uvedení do rovnovážného stavu), která spočívá v ustálení teploty a tlaku v systému okolo zvolených hodnot. Ke stabilizaci teploty, která vede k tzv. kanonickému souboru (NVT) o konstantním látkovém

množství, objemu a teplotě, slouží termostatování. To je realizováno pomocí přeškálování rychlostí nebo za použití některého sofistikovanějšího termostatu – Nosé-Hooverova nebo Berendsenova. [13], [7]

Izotermicko-izobarický soubor (NPT), charakterizovaný konstantním látkovým množstvím, tlakem a teplotou, stabilizuje kromě teploty i tlak – v mém případě na hodnotu atmosférického tlaku (1 bar). Barostatování je prováděno nejčastěji pomocí Berendsenova nebo Parrinello-Rahmanova barostatu. [13], [7]

Po minimalizaci energie a ekvilibraci systému je možné spustit dlouhou MD simulaci (NVT dlouhý běh). Následuje analýza výsledků z kanonického souboru.

2. Seznámení s programem GROMACS

GROMACS⁵ je zkratka pro GRONingen MACHine for Chemical Simulations – nástroj pro chemické simulace vyvinutý na Univerzitě Groningen. Jedná se o unixový programový balík určený pro molekulární dynamiku, šířený pod licenci GPL. [16]

2.1. Práce s utilitami

Tato kapitola je věnována stručnému popisu utilit GROMACSu, které jsem využila k přípravě a provedení simulací a následné analýze výsledků. Nápovědu jakékoli utility lze vyvolat zadáním názvu příslušné utility s volbou *-h*, základní informace o vstupních a výstupních souborech spolu s možnými volbami lze zobrazit pouhým zadáním názvu utility. Více informací o funkcích a použití jednotlivých programů lze nalézt na [6] nebo v uživatelském manuálu [13]. Této problematice se rovněž věnuje K. Šilhavá: Software pro molekulární dynamiku, bakalářská práce, ZSF JU, 2011 [2] v kapitolách 4.1. Základní schéma MD a práce se soubory a 4.3. Práce s utilitami.

Přípony použitých formátů uvádím pro přehlednost za název souboru zvýrazněné kurzívou. Jednotlivé typy topologií (*.tpr*, *.top*, *.itp*) odlišuji v textu jejich příponami vztahujícími se k formátu daného souboru. Všechny tři formáty jsou v odborné literatuře nazývány topologiemi, každý má však odlišnou funkci, obsah i použití.

Utilita	Funkce
editconf	edituje souřadnicové soubory konfigurací, upravuje tvar, velikost a úhly boxu
genbox	generuje box s rozpouštědlem, solvuje systém a vkládá externí molekuly
genion	zaměňuje molekuly rozpouštědla za ionty
make_ndx	utilita sloužící ke generování a editaci indexových souborů
grompp	preprocesor GROMACSu sloužící k přípravě simulace
mdrun	hlavní simulační program balíku GROMACS
G_energy	zobrazuje průměrné hodnoty zvolených veličin, jejich časový vývoj a odchylky
G_density	počítá hustotu systému a hustotní profily zadaných částic napříč systémem
G_rdf	počítá radiální distribuční funkci
G_dist	analyzuje časový vývoj vzdáleností zadaných skupin atomů
G_sgangle	analyzuje úhly a vzdálenosti mezi skupinami atomů

Tab. 1: Stručný přehled využitých utilit a jejich funkcí

⁵ GROMACS implementuje všechna silová pole uvedená v kapitole 1.2.2 [14], [16], [41]

editconf

Tuto utilitu jsem využila při přípravě modelovaných systémů k modifikacím boxů, jmenovitě k úpravě rozměrů boxu s molekulou na rozměry boxu s křemenným povrchem (z důvodu následného sloučení obou boxů) a k posunu molekuly na vhodné místo uvnitř systému (nad pomyslný křemenný povrch).

genbox

Slouží ke sloučení boxu s křemenným povrchem a boxu s molekulou a následné solvataci vzniklého systému. Předchozí úprava rozměrů boxu s molekulou zajistí, že křemenný povrch je „solvatován“ nejprve boxem s molekulou (volbou *-cs*), který je díky vhodné velikosti přidán právě jednou. Tím je zajištěno sloučení obou boxů. V dalším kroku je z databáze do výsledného boxu podobným způsobem přidána voda. Genbox odstraní nadbytečné molekuly vody, které vzniknou překryvem s existujícími molekulami v boxu.

genion

Zamění určitý počet molekul rozpouštědla za jednoatomové ionty (většinou se přidávají Na^+ a Cl^-) za účelem vyrovnaní celkového náboje systému a dosažení požadované koncentrace iontů.

make_ndx

Program k úpravě nebo vytvoření indexového souboru. Prostřednictvím tohoto nástroje lze vytvořit vlastní indexové skupiny podle označení atomů či residuí a sdružovat tak několik částic i celé řetězce, nebo naopak rozdělovat větší skupiny na jednotlivé atomy. Umožňuje editaci skupin, jejich přejmenování či mazání. Indexový soubor je nezbytný při práci se specifickými skupinami atomů při analýze vývoje systému.

grompp

Preprocesor GROMACSu, který ze vstupních souborů vytvoří jediný binární soubor s příponou *.tpr*. Ten slouží jako vstup programu mdrun. Obsahuje veškerá data potřebná ke spuštění a běhu simulace. Po jakékoli změně vstupních parametrů (např. délky doby simulace) je potřeba znovu provést preprocessing (předzpracování) a vytvořit nový spustitelný *.tpr* soubor, aby se tak projevíly požadované změny.

mdrun

Hlavní výpočetní nástroj GROMACSu. Vykonává MD výpočty, stochastickou dynamiku i minimalizaci energie. Během výpočtu vzniká energetický soubor, trajektorie, souřadnice konečné struktury a textový soubor informující o průběhu simulace s časovými údaji.

g_energy

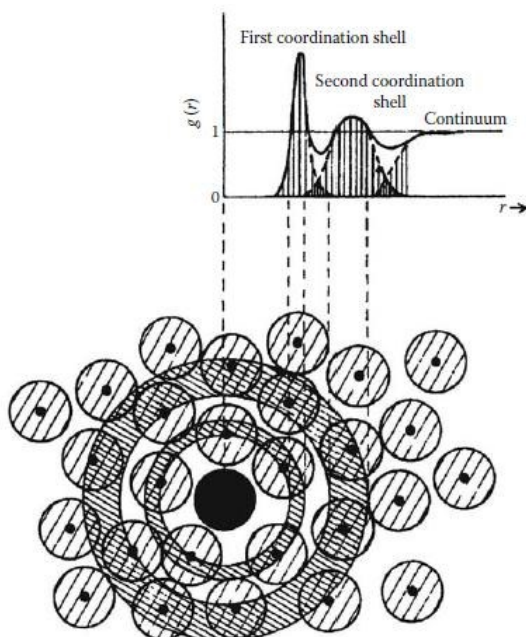
Hlavním vstupním souborem této utility je binární energetický soubor s příponou *.edr*, který vzniká během výpočtu jako výstup programu *mdrun*. Obsahuje data týkající se energií částic, tlaku, velikosti boxu apod. Utilita *g_energy* extrahuje uživatelem zvolená data z tohoto souboru do formátu *.xvg*. Za účelem analýzy jednotlivých energetických příspěvků (např. z interakcí mezi molekulou a povrchem) je potřeba ještě před spuštěním simulace požadované skupiny definovat v parametrickém souboru (viz dále) prostřednictvím parametru *energygrps*.

g_density

Utilita sloužící k výpočtu hustotních profilů částic napříč boxem. Jinými slovy vypočítá rozložení zvolených skupin částic s ohledem na jejich z-ovou souřadnici jako funkci vzdálenosti od povrchu. Po zadání vstupních souborů (indexový soubor, trajektorie a topologie *.tpr*) a volbě skupin/y částic zformuje výstupní data, defaultně uložená do souboru *density.xvg*. Tento formát je určen ke zpracování samostatným unixovým programem *Xmgrace*, WYSIWYG 2D vykreslovacím nástrojem, sloužícím k tvorbě grafů a vizualizaci tabelovaných dat. K úpravě grafů z *g_density* jsem využila programy *g_shift* a *g_sym* RNDr. Milana Předoty, Ph.D. Program *g_shift* posouvá výsledný profil o zvolenou hodnotu, *g_sym* slouží k symetrizaci výsledků pro neutrální/shodně nabitě povrchy (je nutno zadat referenční výšku horního i dolního povrchu).

g_rdf

Radiální distribuční funkce (RDF, Obr. 13) je párová korelační funkce, která udává, s jakou pravděpodobností se dvě částice nacházejí ve vzájemné vzdálenosti r . RDF lze počítat mezi (dvěma nebo více) atomy dvou skupin, které mohou být identické (atomy v rámci jedné skupiny), nebo kolem těžiště skupiny částic, případně pouze kolem jedné částice. Výsledná data, opět uložená ve formátu *.xvg*, lze díky *Xmgrace* vykreslit do grafu. Při své práci jsem utilitu *g_rdf* využívala ke studiu interakcí mezi molekulou a povrchem.



Obr. 13: Schématický význam radiální distribuční funkce a její typický průběh [27]

g_dist

Propočítává vzdálenost mezi těžišti dvou skupin atomů. Výstupní soubor *dist.xvg* obsahuje časový vývoj absolutní vzdálenosti dvou těžišť včetně jednotlivých složek x , y , z . Pro zjištění vzdálenosti molekuly od povrchu jsem využívala právě z -ovou složku prostorově orientovaného vektoru \vec{r} , která udává, v jaké výšce nad povrchem se molekula v daném okamžiku nachází. Ke značnému zjednodušení při zobrazování výsledných grafů této utility napomohl program RNDr. Milana Předoty, Ph.D., nazvaný *g_distmp*, který transformuje výstup utility *g_dist* s oborem hodnot $\langle -L_z/2, L_z/2 \rangle$ do přehledné podoby s oborem hodnot $\langle 0, L_z \rangle$, který byl pro danou geometrii mnohem vhodnější. Hodnota parametru L_z představuje rozměr boxu v ose z (výšku boxu).

g_sgangle

Analyzuje časový vývoj úhlů a vzdáleností mezi zadanými skupinami atomů. Na rozdíl od *g_dist* počítá pouze vzdálenost vyjádřenou prostorově orientovaným vektorem \vec{r} bez jeho jednotlivých složek. Dokáže pracovat pouze se skupinami dvou až tří atomů, z nichž v závislosti na počtu a pořadí v indexovém souboru utvoří orientované vektory či roviny, které jsou definovány normálou danou vektorovým součinem těchto tří atomů. Nástroj *g_sgangle* jsem využívala k analýze natočení molekuly vzhledem k povrchu v době adsorpce.

2.2. Formáty použitých vstupních a výstupních souborů

V následující kapitole uvádím charakteristiku pouze vybraných formátů vstupních a výstupních souborů využitých utilit programového balíku GROMACS, se kterými jsem pracovala při přípravě jednotlivých systémů, spuštění simulací a následné analýze výsledků. GROMACS v mnoha případech nabízí pro podobné aplikace i jiné analogické formáty, avšak jejich popis by přesáhl rámec této práce. Případné zájemce o tuto problematiku odkazuji do uživatelského manuálu [13] nebo na stránky [6], ze kterých jsou informace převzaty. Ukázky souborů jsou převzaty z [6], pokud není uvedeno jinak.

2.2.1. Příprava simulace

K přípravě simulace jsou nutné 3 soubory: vstupní souřadnice počáteční struktury, topologie *.top* a parametrický soubor *.mdp*. Dohromady obsahují veškerá data potřebná ke spuštění a běhu simulace.

2.2.1.1. Vstupní souřadnice

Formát *.gro*

Souřadnicový soubor vygenerovaný prostřednictvím nástroje editconf nebo vzniklý jako výsledek simulace. Informuje o počtu, souřadnicích a jménech atomů a residuí⁶. Může obsahovat rychlosti jednotlivých atomů.

```
MD of 2 waters, t= 0.0
 6
 1WATER OW1 1 0.126 1.624 1.679 0.1227 -0.0580 0.0434
 1WATER HW2 2 0.190 1.661 1.747 0.8085 0.3191 -0.7791
 1WATER HW3 3 0.177 1.568 1.613 -0.9045 -2.6469 1.3180
 2WATER OW1 4 1.275 0.053 0.622 0.2519 0.3140 -0.1734
 2WATER HW2 5 1.337 0.002 0.680 -1.0641 -1.1349 0.0257
 2WATER HW3 6 1.326 0.120 0.568 1.9427 -0.8216 -0.0244
1.82060 1.82060 1.82060
```

Příklad *.gro* souboru – 2 molekuly vody

První řádek obvykle předkládá stručnou charakteristiku systému s jeho názvem. Následuje číslo na druhém řádku, které odpovídá celkovému počtu atomů v systému a zároveň koresponduje s pořadovým číslem posledního atomu, uvedeným ve třetím sloupci. První sloupec souboru nese číslo a název residua, do kterého náleží atom na odpovídajícím řádku

⁶ Residuum = dílčí část struktury tvořená skupinami atomů nebo celými molekulami

(např. tři atomy tvořící jednu molekulu rozpouštědla – v tomto případě vody – tvoří jedno residuum). Přesné označení jednotlivých atomů v rámci residua, tj. atomových typů (pro které jsou v silovém poli definovány interakční parametry) definuje druhý sloupec. Již zmíněný třetí sloupec obsahuje pořadová čísla všech atomů. Následují 3 sloupce, které udávají kartézské souřadnice jednotlivých atomů v pořadí [x; y; z]. Zbývající 3 sloupce obsahují rychlosti atomů ve směru kartézských souřadnicových os. Poslední řádek *.gro* souboru poskytuje informaci o velikosti celého boxu, opět ve formátu [x; y; z]. Souřadnice atomů jsou uváděny v [nm], jednotkou rychlosti je [nm·ps⁻¹] nebo ekvivalentní [km·s⁻¹].

2.2.1.2. Topologie

Formát *.top*

Systémová topologie. Zahrnuje informaci o zvoleném silovém poli, typech a označení atomů a residuí, definici vazeb a úhlů mezi atomy a jejich náboje. Soubor obsahuje jedinou sekci [system] definující název systému a sekci [molecules], která udává počty obsažených molekul jednotlivých typů. Pořadí uvedených molekul v této sekci se musí shodovat s jejich pořadím v souřadnicovém souboru. Systémová topologie může obsahovat vložené topologie specifických typů molekul, které jsou zahrnuty prostřednictvím instrukce *#include* jako samostatné topologie (vložená topologie, viz kapitola Formát *.itp*), nebo charakterizovány přímo v sekcích [moleculetype] a jejich podsekcích. V sekci [atoms] jsou charakterizovány jednotlivé atomy prostřednictvím čísla a typu atomu, náboje a nábojové skupiny. Čísla atomů majících mezi sebou přímou vazbu jsou vypsána v sekci [bonds]. Sekce [angles] obsahuje trojice atomů, které vytváří vazebné úhly. Čísla krajních atomů, které jsou odděleny třemi vazbami (tzv. 1-4 páry), jsou uvedena v sekci [pairs], která souvisí se sekcí [dihedrals]. Ta definuje torzní úhly.

Topologický soubor lze vytvořit ručně v textovém editoru (tímto způsobem připravil Ing. Ondřej Kroutil topologii křemenných povrchů modifikací silového pole [44] nebo vygenerovat např. pomocí nástrojů *pdb2gmx* či *g_x2top*, které jsou součástí GROMACSu. Topologie organických molekul připravil RNDr. Milan Předota, Ph.D., pomocí skriptu *MKTOP*, který není součástí distribuce GROMACSu a byl vyvinut nezávislou skupinou, ale funkčnost a obliba tohoto skriptu mu již zajistila místo mezi odkazy webových stránek programu GROMACS [28], [43]. Parciální náboje atomů molekul a jejich energeticky

minimalizované konfigurace dodal RNDr. Babak Minofar, MSc. Ph.D., z RESP kvantových výpočtů.

Spolu se souřadnicovým a parametrickým souborem slouží topologie *.top* jako vstupní soubor preprocesoru *grompp*, který zpracuje vstupní data a vytvoří jediný binární soubor s příponou *.tpr*. Ten potom slouží jako vstupní soubor ke spuštění simulace pomocí *mdrun*.

```
;      Example topology file
;
[ defaults ]
; nbfunc      comb-rule      gen-pairs      fudgeLJ fudgeQQ
  1            1             no             1.0     1.0

[ moleculetype ]
; name nrexcl
Urea      3

[ atoms ]
;  nr   type  resnr  residu   atom   cgnr  charge
  1     C     1     UREA    C1     1     0.683
  2     O     1     UREA    O2     1     -0.683
  3     NT    1     UREA    N3     2     -0.622
  4     H     1     UREA    H4     2     0.346
  5     H     1     UREA    H5     2     0.276
  6     NT    1     UREA    N6     3     -0.622
  7     H     1     UREA    H7     3     0.346
  8     H     1     UREA    H8     3     0.276

[ bonds ]
;  ai   aj  funct      c0      c1
  3     4   1 1.000000e-01 3.744680e+05
  3     5   1 1.000000e-01 3.744680e+05
  6     7   1 1.000000e-01 3.744680e+05
  6     8   1 1.000000e-01 3.744680e+05
  1     2   1 1.230000e-01 5.020800e+05
  1     3   1 1.330000e-01 3.765600e+05
  1     6   1 1.330000e-01 3.765600e+05

[ pairs ]
;  ai   aj  funct      c0      c1
  2     4   1 0.000000e+00 0.000000e+00
  2     5   1 0.000000e+00 0.000000e+00
  2     7   1 0.000000e+00 0.000000e+00
  2     8   1 0.000000e+00 0.000000e+00
  3     7   1 0.000000e+00 0.000000e+00
  3     8   1 0.000000e+00 0.000000e+00
  4     6   1 0.000000e+00 0.000000e+00
  5     6   1 0.000000e+00 0.000000e+00

[ angles ]
;  ai   aj   ak  funct      c0      c1
  1     3     4   1 1.200000e+02 2.928800e+02
  1     3     5   1 1.200000e+02 2.928800e+02
  4     3     5   1 1.200000e+02 3.347200e+02
  1     6     7   1 1.200000e+02 2.928800e+02
  1     6     8   1 1.200000e+02 2.928800e+02
  7     6     8   1 1.200000e+02 3.347200e+02
  2     1     3   1 1.215000e+02 5.020800e+02
  2     1     6   1 1.215000e+02 5.020800e+02
  3     1     6   1 1.170000e+02 5.020800e+02

[ dihedrals ]
;  ai   aj   ak   al  funct      c0      c1      c2
```

```

2      1      3      4      1 1.800000e+02 3.347200e+01 2.000000e+00
6      1      3      4      1 1.800000e+02 3.347200e+01 2.000000e+00
2      1      3      5      1 1.800000e+02 3.347200e+01 2.000000e+00
6      1      3      5      1 1.800000e+02 3.347200e+01 2.000000e+00
2      1      6      7      1 1.800000e+02 3.347200e+01 2.000000e+00
3      1      6      7      1 1.800000e+02 3.347200e+01 2.000000e+00
2      1      6      8      1 1.800000e+02 3.347200e+01 2.000000e+00
3      1      6      8      1 1.800000e+02 3.347200e+01 2.000000e+00

[ dihedrals ]
; ai    aj    ak    al funct          c0          c1
  3     4     5     1     2 0.000000e+00 1.673600e+02
  6     7     8     1     2 0.000000e+00 1.673600e+02
  1     3     6     2     2 0.000000e+00 1.673600e+02

; Include SPC water topology
#include "spc.itp"

[ system ]
Urea in Water

[ molecules ]
Urea      1
SOL      1000

```

Příklad topologického souboru – molekula močoviny ve vodě

Formát *.itp*

Vložená topologie je zahrnuta v rámci systémové jako samostatný soubor prostřednictvím instrukce *#include*. Definuje jen část topologie celého systému, např. topologii specifického typu molekuly, rozpouštědla nebo iontů. Vymezuje vazebné a nevazebné interakce pro atomy. Na rozdíl od systémové topologie neobsahuje definici silového pole ani sekce [system] a [molecules]. Pro potřeby simulací lze využít již připravené soubory z gromacsovské databáze nebo ručně vytvořené.

2.2.1.3. Parametrické soubory

Dalším souborem nutným ke spuštění simulace je parametrický soubor s příponou *.mdp*. Používám jsem 4 základní typy parametrických souborů: *sd.mdp* pro minimalizaci energie, *nvt.mdp* pro NVT fázi ekvibrace, *npt.mdp* pro NPT fázi ekvibrace a *md.mdp* pro produkční běh (uvedené označení je jen orientační). Jednotlivé soubory mají různou funkci a liší se v některých parametrech podstatných pro každou konkrétní aplikaci. Obvykle bývají uplatněny v uvedeném pořadí. Jejich použití a vlastnosti uvádím spolu s krátkým popisem a ukázkami.

Minimalizace energie

Cílem procesu minimalizace je snížení potenciální energie systému. Pro simulace za dostatečně vysoké teploty je klíčové (a prakticky postačující) odstranění silných repulzí

atomů, takže je přijatelné, že výsledná konfigurace po minimalizaci reprezentuje pouze jedno z lokálních minim energetické hyperplochy. Při použití tohoto typu *.mdp* souboru není uplatněn klasický molekulárně-dynamický výpočet. K nalezení lokálního minima potenciální energie poblíž počáteční struktury je díky volbě *integrator = steep* použit algoritmus metody největšího spádu (steepest descent) – popsáno v kapitole 1.3.3. Minimalizace je ukončena v případě, že je proveden stanovený počet kroků, častěji však, je-li maximální dosažená síla menší než hodnota parametru *emtol* [$\text{kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-1}$]. (Více o volbě hodnot v uživatelském manuálu [13]). Velikost a počet kroků minimalizace udávají parametry *emstep* [nm] a *nsteps*. Volba *nstlist* vyjadřuje počet kroků, po jehož uplynutí se aktualizuje seznam sousedních částic. Parametr *rlist* [nm] určuje, do jaké vzdálenosti se budou počítat interakce jednotlivých částic (cutoff); v uvedeném příkladu nebudou započítávány interakce ve vzdálenosti větší než 1,0 [nm]. Způsob výpočtu elektrostatických interakcí zajišťuje *coulombtype* s tím, že parametr *rcoulomb* [nm] vymezuje vzdálenost pro Coulombický cutoff. Velikost Lennard-Jonesova nebo Buckinghamova cutoffu stanovuje parametr *rvdw* [nm]. Periodické okrajové podmínky *pb*c (periodic boundary conditions) jsou nastaveny ve všech třech dimenzích (*xyz*).

```

; minim.mdp - used as input into grompp to generate em.tpr
; Parameters describing what to do, when to stop and what to save
integrator      = steep          ; Algorithm (steep = steepest descent
                                minimization)
emtol           = 1000.0        ; Stop minimization when the maximum
                                force < 1000.0 kJ/mol/nm
emstep         = 0.01          ; Energy step size
nsteps         = 50000         ; Maximum number of (minimization) steps to
                                perform

; Parameters describing how to find the neighbors of each atom and how to
calculate the interactions
nstlist        = 1             ; Frequency to update the neighbor list and long
                                range forces
ns_type        = grid          ; Method to determine neighbor list (simple,
grid)
rlist          = 1.0           ; Cut-off for making neighbor list (short range
                                forces)
coulombtype    = PME           ; Treatment of long range electrostatic
                                interactions
rcoulomb       = 1.0           ; Short-range electrostatic cut-off
rvdw           = 1.0           ; Short-range Van der Waals cut-off
pb
```

Příklad *.mdp* souboru pro minimalizaci energie [29]

NVT ekvilibrace

Aplikací tohoto typu *.mdp* souboru lze dosáhnout stabilizace teploty systému na požadovanou hodnotu. Soubor předepisuje podobně jako v předchozím případě použitý typ algoritmu pro výpočet (volba *integrator*; tentokrát *md*) a počet kroků simulace *nsteps*.

Z posledně jmenovaného údaje a velikosti časového kroku dt v [ps] lze lehce odvodit, jak dlouhý časový úsek simulace ve skutečnosti postihuje. Tento soubor rovněž předepisuje, jak často (po kolika krocích) se ukládají informace o souřadnicích (*nstxout*), rychlostech (*nstvout*) a energii (*nstenergy*). Parametr *nstxout* nastavený na nenulovou hodnotu vede k zápisu *.trr* trajektorie, obdobný parametr *nstxcout* zajistí vytvoření *.xtc* trajektorie. Oba výstupní soubory se však z důvodu komprese *.xtc* trajektorie liší ve velikosti (viz dále). Obvykle se proto používá pouze jeden z těchto uvedených parametrů (jedna z trajektorií). V případě absence obou parametrů je implicitně ukládána *.trr* trajektorie po 100 krocích. Pro zapsání trajektorie musí být oba implementované parametry řádově menší než počet kroků *nsteps*. Důležitou sekci v tomto parametrickém souboru jsou instrukce vztahující se k termostatování, kde je stanoven druh termostatu prostřednictvím parametru *tcoupl* a jeho relaxační doba *tau_t* v [ps]. Dále termostatované skupiny *tc_grps* a hodnota referenční teploty *ref_t* v [K], na které se má systém ustálit. Rychlosti jednotlivých částic jsou stanoveny z Maxwellova rozdělení (*gen_vel*) při teplotě 300 K (*gen_temp*). Stabilizace tlaku je v tomto případě vypnuta (*pcoupl = no*), velikost boxu zůstává tedy v průběhu celé simulace konstantní.

```

title           = OPLS Lysozyme NVT equilibration
define          = -DPOSRES           ; position restrain the protein
; Run parameters
integrator      = md                 ; leap-frog integrator
nsteps         = 50000               ; 2 * 50000 = 100 ps
dt             = 0.002               ; 2 fs
; Output control
nstxout        = 100                 ; save coordinates every 0.2 ps
nstvout        = 100                 ; save velocities every 0.2 ps
nstenergy      = 100                 ; save energies every 0.2 ps
nstlog         = 100                 ; update log file every 0.2 ps
; Bond parameters
continuation   = no                 ; first dynamics run
constraint_algorithm = lincs         ; holonomic constraints
constraints    = all-bonds           ; all bonds (even heavy atom-H bonds) constrained
lincs_iter     = 1                   ; accuracy of LINCS
lincs_order    = 4                   ; also related to accuracy
; Neighborsearching
ns_type        = grid                ; search neighboring grid cells
nstlist       = 5                    ; 10 fs
rlist         = 1.0                  ; short-range neighborlist cutoff (in nm)
rcoulomb      = 1.0                  ; short-range electrostatic cutoff (in nm)
rvdw         = 1.0                   ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype    = PME                 ; Particle Mesh Ewald for long-range
                                         electrostatics
pme_order     = 4                    ; cubic interpolation
fourierspacing = 0.16                ; grid spacing for FFT
; Temperature coupling is on
tcoupl        = V-rescale             ; modified Berendsen thermostat
tc-grps       = Protein Non-Protein   ; two coupling groups - more accurate
tau_t         = 0.1 0.1              ; time constant, in ps
ref_t         = 300 300              ; reference temperature, one for each group, in K
; Pressure coupling is off
pcoupl        = no                   ; no pressure coupling in NVT

```

```

; Periodic boundary conditions
pbc          = xyz          ; 3-D PBC
; Dispersion correction
DispCorr= EnerPres        ; account for cut-off vdW scheme
; Velocity generation
gen_vel      = yes         ; assign velocities from Maxwell distribution
gen_temp     = 300         ; temperature for Maxwell distribution
gen_seed     = -1         ; generate a random seed

```

Příklad *.mdp* souboru pro NVT fázi ekvilibrace [30]

NPT ekvilibrace

Použitím NPT souboru je docíleno stabilizace tlaku a s ním související hustoty v systému. Z tohoto důvodu obsahuje na rozdíl od předchozího navíc parametry pro barostatování. Volba *continuation* nastavená na hodnotu *yes* znamená, že tato fáze navazuje na NVT fázi ekvilibrace. Referenční hodnota tlaku, na které se má systém ustálit, je označena *ref_p* [bar], typ použitého barostatu určuje *pcoupl*. *Pcoupltype* (druh barostatování) je v tomto případě nastaven na izotropický – škálování rozměrů boxu probíhá ve všech třech souřadnicových směrech se stejnou kompresibilitou *compressibility* [bar^{-1}]. Z důvodu zachování nastaveného tlaku se box bude rozpínat a smršťovat s časovou konstantou *tau_p* [ps]. Stlačitelnost použitého rozpouštědla je dána veličinou *compressibility* [bar^{-1}]. Oproti NVT souboru nejsou v uvedeném příkladu generovány rychlosti částic (*gen_vel* = *no*).

```

title          = OPLS Lysozyme NPT equilibration
define         = -DPOSRES          ; position restrain the protein
; Run parameters
integrator     = md                ; leap-frog integrator
nsteps        = 50000             ; 2 * 50000 = 100 ps
dt            = 0.002             ; 2 fs
; Output control
nstxout       = 100               ; save coordinates every 0.2 ps
nstvout       = 100               ; save velocities every 0.2 ps
nstenergy     = 100               ; save energies every 0.2 ps
nstlog        = 100               ; update log file every 0.2 ps
; Bond parameters
continuation   = yes              ; Restarting after NVT
constraint_algorithm = lincs      ; holonomic constraints
constraints    = all-bonds        ; all bonds (even heavy atom-H bonds) constrained
lincs_iter    = 1                 ; accuracy of LINCS
lincs_order   = 4                 ; also related to accuracy
; Neighborsearching
ns_type       = grid              ; search neighboring grid cells
nstlist       = 5                 ; 10 fs
rlist         = 1.0               ; short-range neighborlist cutoff (in nm)
rcoulomb      = 1.0               ; short-range electrostatic cutoff (in nm)
rvdw         = 1.0               ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype   = PME               ; Particle Mesh Ewald for long-range
electrostatics
pme_order     = 4                 ; cubic interpolation
fourierspacing = 0.16            ; grid spacing for FFT
; Temperature coupling is on
tcoupl        = V-rescale         ; modified Berendsen thermostat
tc-grps       = Protein Non-Protein ; two coupling groups - more accurate
tau_t         = 0.1 0.1           ; time constant, in ps
ref_t         = 300 300           ; reference temperature, one for each group, in K

```

```

; Pressure coupling is on
pcoupl      = Parrinello-Rahman      ; Pressure coupling on in NPT
pcoupltype  = isotropic              ; uniform scaling of box vectors
tau_p       = 2.0                    ; time constant, in ps
ref_p       = 1.0                    ; reference pressure, in bar
compressibility = 4.5e-5              ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc         = xyz                    ; 3-D PBC
; Dispersion correction
DispCorr    = EnerPres               ; account for cut-off vdW scheme
; Velocity generation
gen_vel     = no                     ; Velocity generation is off

```

Příklad *.mdp* souboru pro NPT fázi ekvibrace [31]

Produkční fáze

Použití tohoto typu souboru závisí na konkrétní aplikaci, je však víceméně volitelné, neboť hodnoty parametrů pro stabilizaci teploty a tlaku systému jsou obdobné jako v předchozích ekvilibračních souborech. Jsou však prodlouženy intervaly ukládání konfigurací a mezivýsledků a celková doba běhu simulace. Pro účely dlouhé simulace je tedy možné využít NVT parametrický soubor s prodlouženou dobou simulace; obvykle se však po předchozím ustálení tlaku dále nepokračuje s barostatováním (na rozdíl od uvedeného příkladu).

```

title      = OPLS Lysozyme MD
; Run parameters
integrator = md                ; leap-frog integrator
nsteps     = 500000            ; 2 * 500000 = 1000 ps, 1 ns
dt         = 0.002             ; 2 fs
; Output control
nstxout    = 1000              ; save coordinates every 2 ps
nstvout    = 1000              ; save velocities every 2 ps
nstxtcout  = 1000              ; xtc compressed trajectory output every 2 ps
nstenergy  = 1000              ; save energies every 2 ps
nstlog     = 1000              ; update log file every 2 ps
; Bond parameters
continuation = yes            ; Restarting after NPT
constraint_algorithm = lincs   ; holonomic constraints
constraints  = all-bonds      ; all bonds (even heavy atom-H bonds) constrained
lincs_iter   = 1              ; accuracy of LINCS
lincs_order  = 4              ; also related to accuracy
; Neighborsearching
ns_type      = grid           ; search neighboring grid cells
nstlist      = 5              ; 10 fs
rlist        = 1.0            ; short-range neighborlist cutoff (in nm)
rcoulomb     = 1.0            ; short-range electrostatic cutoff (in nm)
rvdw         = 1.0            ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype  = PME            ; Particle Mesh Ewald for long-range
electrostatics
pme_order    = 4              ; cubic interpolation
fourierspacing = 0.16        ; grid spacing for FFT
; Temperature coupling is on
tcoupl       = V-rescale      ; modified Berendsen thermostat
tc-grps      = Protein Non-Protein ; two coupling groups - more accurate
tau_t        = 0.1 0.1       ; time constant, in ps
ref_t        = 300 300       ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl      = Parrinello-Rahman ; Pressure coupling on in NPT

```

```

pcoupltype      = isotropic      ; uniform scaling of box vectors
tau_p           = 2.0            ; time constant, in ps
ref_p           = 1.0            ; reference pressure, in bar
compressibility = 4.5e-5         ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc             = xyz            ; 3-D PBC
; Dispersion correction
DispCorr        = EnerPres       ; account for cut-off vdW scheme
; Velocity generation
gen_vel         = no             ; Velocity generation is off

```

Příklad *.mdp* souboru pro produkční fázi MD [32]

2.2.2. Spuštění simulace

Ke spuštění simulace je zapotřebí jediný soubor, který vzniká jako výsledek preprocessingu. Obsahuje veškerá vstupní data potřebná k běhu simulace.

Formát *.tpr*

Binární topologický soubor, který slouží ke spuštění simulace pomocí *mdrun*. Obsahuje počáteční konfiguraci celé simulace, souřadnice a rychlosti všech částic, topologie a veškeré potřebné parametry pro běh simulace. Vzniká jako výstup při preprocessingu a následně je zpracován programem *mdrun* při výpočtu. Z důvodu velkého množství dat i pro malý počet částic v modelovaném systému neuvádím ukázkou tohoto souboru, neboť by si vyžádala mnoho stránek. Soubor je možné si prohlédnout zadáním následujícího příkazu:

```
gmxdump -s topol.tpr | more
```

Případně přesměrováním do souboru:

```
gmxdump -s topol.tpr > soubor.txt
```

2.2.3. Analýza výsledků

Během simulace vznikají 4 výstupní soubory výpočetního programu *mdrun*: energetický soubor, trajektorie, výstupní souřadnice konečné struktury a textový soubor informující o průběhu simulace.

2.2.3.1. Energetický soubor

Formát *.edr*

Binární energetický soubor obsahující veškerá data týkající se energií, teploty a tlaku, uložená během simulace v intervalech specifikovaných v *.mdp* souboru. K získání konkrétních hodnot za účelem analýzy výsledků simulace je třeba použít nástroj *g_energy*, který extrahuje požadovaná data do textového souboru *.xvg*. Tento nástroj umožňuje

uživateli interaktivně specifikovat všechny veličiny, které jsou obsaženy v energetickém souboru. Obsah binárního souboru *.edr* je možné si prohlédnout zadáním příkazu:

```
gmxdump -e ener.edr | more
```

Nebo přesměrováním do souboru:

```
gmxdump -e ener.edr > soubor.txt
```

2.2.3.2. Textový soubor

Formát *.xvg*

Tento soubor je výstupem utility *g_energy*, jež umožňuje analyzovat časovou závislost zvolených veličin, a nebo programu *g_rdf*, který slouží k výpočtu RDF. Jedná se o textový o několika sloupcích, jejichž počet je závislý na množství navolených veličin. První sloupec představuje časové údaje (v případě výstupů z *g_dist* a *g_energy*), případně vzdálenosti (výstupy *g_rdf* a *g_density*) a následující sloupce obsahují časový vývoj hodnot vybraných veličin. Úvodní metapříkazy slouží k vytvoření legendy a popisu grafu. Tento typ souboru obsahuje data ve formátu určeném primárně ke zpracování již zmíněným programem *Xmgrace*. Data je možné po drobné úpravě importovat i do jiných volně dostupných programů jako je např. *Gnuplot*, nebo převést do tabulkové podoby pro práci v oblíbeném tabulkovém procesoru MS Office Excel.

```
# This file was created Sat Nov 19 22:51:34 2011
# by the following command:
# g_energy -f benzoicw.edr -o pokus.xvg
#
# g_energy is part of G R O M A C S:
#
# Giant Rising Ordinary Mutants for A Clerical Setup
#
@ title "Gromacs Energies"
@ xaxis label "Time (ps)"
@ yaxis label "(kJ/mol)"
@TYPE xy
@ view 0.15, 0.15, 0.75, 0.85
@ legend on
@ legend box on
@ legend loctype view
@ legend 0.78, 0.8
@ legend length 2
@ s0 legend "Potential"
@ s1 legend "Kinetic En."
@ s2 legend "Total Energy"
  0.000000 -35024.203125 7766.743652 -27257.458984
  0.100000 -34800.757812 7553.682617 -27247.074219
  0.200000 -35087.812500 7578.834961 -27508.976562
  0.300000 -34834.953125 7533.489258 -27301.464844
  ...
 29.700000 -34893.492188 7663.025879 -27230.466797
```

```

29.800000 -34922.011719 7704.689453 -27217.322266
29.900000 -35139.945312 7708.958984 -27430.986328
30.000000 -34751.621094 7656.285156 -27095.335938

```

Příklad .xvg souboru – vývoj potenciální, kinetické a celkové energie systému molekuly kyseliny benzoové s vodou. Tři tečky nahrazují vynechané hodnoty.

2.2.3.3. Indexový soubor

Formát .ndx

Indexový soubor, který umožňuje dále pracovat s vybranými skupinami atomů. Při své práci jsem tento typ souboru využívala především během analýzy výsledků, z tohoto důvodu jej uvádím až na tomto místě. Pokud není vytvořen, většina utilit umí i bez něj, přímo s použitím konfiguračního souboru, generovat automaticky některé základní indexové skupiny – pro celý systém, proteiny, residua, rozpouštědlo a další. V případě potřeby specifických skupin je nutné indexový soubor založit či upravit ručně nebo za pomoci utility `make_ndx`. Pro mnoho utilit slouží indexový soubor jako volitelný vstupní soubor (`grompp`, `g_rdf`, `g_density`) pro práci se zvláštními skupinami atomů a následnou analýzu vývoje systému.

```

[ Oxygen ]
1         4         7
[ Hydrogen ]
2         3         5         6
8         9

```

Ukázka .ndx souboru o dvou skupinách atomů – sk. Oxygen se 3 atomy a sk. Hydrogen o 6 atomech

```

Reading structure file
Going to read 0 old index file(s)
Analysing residue names:
There are: 2753      Other residues
There are: 4589      Water residues
There are: 16       Ion residues
Analysing residues not classified as Protein/DNA/RNA/Water and splitting into
groups...
Analysing residues not classified as Protein/DNA/RNA/Water and splitting into
groups...

```

```

0 System          : 17046 atoms
1 Other           : 3263 atoms
2 SiB             : 576 atoms
3 OB              : 1024 atoms
4 OBS             : 256 atoms
5 OS              : 512 atoms
6 SSB             : 128 atoms
7 SOD             : 384 atoms
8 SOU             : 336 atoms
9 SON             : 32 atoms
10 MOL            : 15 atoms
11 NA             : 16 atoms
12 Water          : 13767 atoms
13 SOL            : 13767 atoms

```

```

14 non-Water      : 3279 atoms
15 Ion            : 16 atoms
16 Water_and_ions : 13783 atoms
17 OW             : 4589 atoms
18 HW*           : 9178 atoms
19 AH             : 5 atoms
20 H6             : 1 atoms
21 O1             : 1 atoms
22 O2             : 1 atoms
23 OS12          : 256 atoms

nr : group      ! 'name' nr name 'splitch' nr Enter: list groups
'a': atom      & 'del' nr      'splitres' nr 'l': list residues
't': atom type | 'keep' nr     'splitat' nr  'h': help
'r': residue   'res' nr      'chain' char
'name': group  'case': case sensitive 'q': save and quit
'ri': residue index

```

Systém s křemenným povrchem a molekulou kyseliny benzoové ve vodě – indexové skupiny následované nápovědou možných příkazů editoru make_ndx.

2.2.3.4. Trajektorie

Tyto soubory jsou mimo jiných výsledkem běhu simulace. Během výpočtu se zapisují polohy jednotlivých atomů a v závislosti na nastavení v parametrickém souboru i jejich rychlosti a působící síly. Typ zapisovaných dat a frekvence jejich zápisu jsou dány prostřednictvím parametrů v *.mdp* souboru. Program mdrun defaultně vytváří trajektorii *.trr* a dle nastavení volby *nstxtcout* i komprimovanou verzi s příponou *.xtc*. GROMACS i VMD (program sloužící k vizualizaci molekulárních struktur) pracují s oběma formáty. GROMACS kromě těchto nabízí ještě další jako jsou *.trj*, *.g87*, *.g96*, *.gro* nebo *.pdb* formát, z nichž každý se vyznačuje jiným stupněm komprese. Trajektorie obvykle zahrnují velké množství dat. Všechny binární soubory včetně *.xtc* a *.trr* trajektorií jsou tak z hlediska velikosti daleko úspornější než textové soubory v ASCII kódu.

Formát *.trr*

Obsah trajektorie závisí na volbách v *.mdp* souboru: parametr *nstxout* udává frekvenci zápisu souřadnic, ukládání rychlostí částic je dáno veličinou *nstvout* a informace o velikostech působících sil nastavením hodnoty *nstfout*. Pokud není stanoveno jinak, zapisují se souřadnice do souboru *.trr* defaultně každých 100 kroků. Při ukládání tohoto typu souboru není použit kompresní algoritmus jako je tomu např. u *.xtc* trajektorie. Z tohoto důvodu zabírá oproti *.xtc* trajektorii přibližně 3krát více diskového prostoru.

Formát *.xtc*

Soubor *.xtc* je komprimovanou verzí *.trr* trajektorie obsahující souřadnice všech atomů v čase a údaje o velikosti boxu. V zásadě je shodná s *.trr* trajektorií. Při analýze systémů, kde

vznikají vyšší nároky na přesnost výsledků z `g_density` (např. při zkoumání hustoty vody v modelovaném systému), je však nutno ještě před spuštěním simulace nastavit hodnotu parametru `xtc_precision` v `.mdp` souboru na hodnotu 10000 (pokud není parametr uveden, činí hodnota 1000). Souřadnice jsou díky tomuto nastavení zapisovány s přesností na 4 desetinná místa (oproti implicitním 3) a výsledné grafy hustoty vody nevykazují tak jako v defaultním případě nežádoucí nepřesnosti způsobené zaokrouhlovací chybou. Pokud se v `.mdp` souboru nerovnájí hodnoty parametrů `nstxout` a `nstxtcout`, budou se obě trajektorie lišit v množství zapsaných souřadnic. Obsah trajektorií je možné zobrazit pomocí příkazu:

```
gmxdump -f traj.xtc | more
```

Nebo přesměrovat do textového souboru:

```
gmxdump -f traj.xtc > soubor.txt
```

V případě formátu `.trr` je nutné u výše uvedených příkazů změnit příponu. Prohlížení trajektorií tímto způsobem nemá obvykle význam. U delších simulací vznikají obrovské soubory a jejich zpracování i ukládání je náročné na výpočetní prostředky a úložný prostor.

```
benzoicw.trr frame 0:
  natoms=      2679  step=          0  time=0.0000000e+00  lambda=          0
  box (3x3):
    box[  0]={ 3.05894e+00,  0.00000e+00,  0.00000e+00}
    box[  1]={ 0.00000e+00,  3.05894e+00,  0.00000e+00}
    box[  2]={ 0.00000e+00,  0.00000e+00,  3.05894e+00}
  x (2679x3):
  x[  0]={ 2.63819e+00,  1.70589e+00,  1.28908e+00}
  x[  1]={ 2.58393e+00,  1.69799e+00,  1.38212e+00}
  x[  2]={ 2.57514e+00,  1.75800e+00,  1.17546e+00}
  x[  3]={ 2.47968e+00,  1.80315e+00,  1.19808e+00}
  ...
  x[ 2675]={ 1.89392e+00,  1.17710e+00,  2.13245e-01}
  x[ 2676]={ 2.83529e+00,  6.12726e-01,  1.72240e+00}
  x[ 2677]={ 2.87498e+00,  5.77521e-01,  1.63764e+00}
  x[ 2678]={ 2.90782e+00,  6.27072e-01,  1.78973e+00}
```

Ukázka trajektorie `.trr`. Tři tečky nahrazují vynechané hodnoty.

3. Cíle práce

1. Seznámení se základy molekulární dynamiky a problematikou modelování interakcí molekul s povrchy.
2. Seznámení s prací v programu pro molekulární dynamiku GROMACS.
3. Provedení simulací interakce několika jednoduchých organických molekul s povrchem křemene (SiO_2) pro různé hustoty náboje povrchu (v závislosti na pH roztoku) a disociované stavy molekul.
4. Zpracování výsledků simulací, analýza adsorpčních geometrií a interakčních energií.

4. Metodika

V rámci řešení své bakalářské práce jsem získala přístup k počítačovému klastru UFY a po registraci na stránkách <https://metavo.metacentrum.cz/osobni/prihlaska.do> i k fakultnímu klastru Hermes, který je zařazený do celorepublikového programu organizace MetaCentrum. K oběma strojům jsem se připojovala prostřednictvím SSH klienta PuTTY 0.60 a SFTP klienta WinSCP 4.1.9. Pro přenos grafického výstupu ze vzdálených aplikací (Xmgrace, VMD) jsem využívala program Xming 6.9.0.31. Zmíněné programy jsou volně dostupné a instalovatelné v prostředí OS Windows. Prostřednictvím příkazové řádky OS Linux jsem na klastru UFY připravovala uvedené systémy skládající se z kombinací organických molekul a různě nabitých povrchů křemene (SiO_2), poté posílala soubory nezbytné k provedení výpočtů (.tpr soubory) na klastr Hermes a vzdáleně pomocí dávkových souborů spouštěla simulace těchto struktur. K přípravě simulací a jejich spouštění jsem využívala výše popsaný software určený pro molekulární dynamiku – GROMACS 4.5.3, resp. 4.5.5. Výstupní soubory simulací jsem přesouvala zpět na klastr UFY a analýzu výsledků prováděla zde. Zejména při analýze chování modelovaných systémů jsem kromě prostředků, které nabízí GROMACS, využívala program VMD 1.9 a Xmgrace 5.1.21.

Ve své práci neuvádím vlastnosti ani bližší charakteristiku výše jmenovaných programů, které přímo nesouvisí s modelováním a simulacemi, neboť jejich popis není předmětem této práce. Využití a širší možnosti programu VMD uvádí Krohová, L.: Software pro zobrazování molekulárních struktur, bakalářská práce, ZSF JU, 2011 v kapitole 3.3 VMD. Základní manuál k VMD lze stáhnout ze stránek [33]. Informace o programu Xmgrace lze nalézt na stránkách [34], stručný návod je dostupný na [35].

4.1. Postup při přípravě systémů

Tato kapitola je věnována popisu přípravy systémů sestávajících z kombinací různě nabitých povrchů s jednotlivými molekulami nebo jejich ionty vzniklými deprotonací OH skupiny, ke které dochází při $\text{pH} > \text{pK}_a$ molekuly. Celý postup je prezentován na příkladu kyseliny benzoové s neutrálním křemenným povrchem. Analogický postup platí i v případě ostatních struktur.

V tabulce níže (Tab. 2) jsou uvedeny nejčastější přepínače pro utility použité k přípravě systémů v následujícím postupu.

Přepínač	Použití
-f	hlavní vstupní soubor
-o	hlavní výstupní soubor
-p	topologie <i>.top</i>
-n	indexový soubor
-c	vstupní souřadnice
-s	topologie <i>.tpr</i>
-cp	vstupní souřadnice solutu
-cs	souřadnicový soubor solventu

Tab. 2: Přehled základních přepínačů a jejich použití

1. Prvním krokem k přípravě modelovaného systému je zjištění rozměrů boxu obsahujícího křemenný povrch. Tyto rozměry slouží v následujícím kroku k úpravě velikosti boxu s molekulou z důvodu pozdějšího sloučení obou boxů. Ke zjištění rozměrů lze použít uvedený příkaz k prohlížení konce obsahu *.gro* souboru, kde jsou údaje o velikosti systému zapsány, nebo příslušný soubor otevřít ve WinSCP.

```
tail quartz0.00.gro
```

2. Následuje úprava rozměrů boxu s molekulou (*benzoic.gro*) na rozměry boxu s křemenným povrchem prostřednictvím parametru *-box*. Současně je vhodné molekulu umístit nad pomyslný povrch do prostoru volného boxu pomocí parametru *-center*. Rozměry i souřadnice jsou zadávány v [nm] a v pořadí [x; y; z].

```
editconf -f benzoic.gro -o benz_uprav.gro -box 5.50000 3.98200 8.00000 -center 2.75 1.99 3
```

3. Dále jsou oba boxy sloučeny pomocí příkazu *genbox*, který díky volbě *-cs* zajistí přidání boxu s molekulou do boxu s povrchem. Díky shodným velikostem obou boxů je molekula do systému přidána právě jednou.

```
genbox -cp quartz0.00.gro -cs benz_uprav.gro -o quartz+benz.gro
```

4. Dalším krokem je solvatace vzniklého systému a automatická aktualizace topologického souboru – vložení řádky s počtem přidávaných molekul vody (defaultně SPC či SPC/E; mají identickou geometrii). Volba *-vdwd* v příkazu níže upravuje vdW vzdálenost mezi molekulami solventu a mění tak jeho hustotu. Zároveň zajišťuje vhodnou vzdálenost mezi molekulami solventu a solutu (vody a povrchu nebo vody a vložené organické molekuly).

```
genbox -cp quartz+benz.gro -cs -o qbenzwat.gro -p topol.top -vdwd 0.15
```

5. Před přidáním iontů do systému v následujícím bodě postupu je potřeba nejprve vytvořit topologii *.tpr*, která je požadovaným vstupním souborem utility *genion*. Toho lze

dosáhnout spuštěním programu `grompp`, který provede preprocessing stávajícího systému a vygeneruje zmíněnou topologii.

```
grompp -f genion.mdp -p topol.top -c qbenzwat.gro -o qbenzwat.tpr
```

6. Dalším krokem je úprava koncentrace solí v roztoku a vyrovnání celkového náboje systému nahrazením molekul vody kladnými a zápornými ionty; kationty Na^+ a anionty Cl^- . Parametr `-conc` udává koncentraci (zvolena fyziologická hodnota $0,13 \text{ [mol}\cdot\text{l}^{-1}]$), ze které je pro daný systém vypočítán potřebný počet iontů. V kombinaci s volbou `-neutral` zajistí vytvoření neutrálního systému s danou koncentrací solí.

```
genion -s qbenzwat.tpr -o qbenzwati.gro -p topol.top -conc 0.13 -neutral  
volba SOL
```

7. Důležitým bodem je vytvoření indexového souboru, který před spuštěním simulace slouží k definování skupin atomů, mezi kterými jsou počítány energetické příspěvky (MOL, SOL, SURF) nebo které mají zůstat během výpočtu nepohyblivé (BULK). Tyto skupiny musí být uvedeny v parametrickém souboru `.mdp` u příslušných parametrů. Pro spuštění simulace je nutné definovat pouze skupiny SURF a BULK (ostatní do této chvíle nezbytné skupiny jsou vytvořeny na základě údajů v souřadnicovém souboru). Pro účely pozdější analýzy jsou definovány další skupiny vybraných atomů (viz kapitola 5.1.2.). Detailní popis vytváření atomových skupin je uveden v kapitole 5.2. v části Úprava indexového souboru.

```
make_ndx -f qbenzwati.gro -o index.ndx
```

8. Následuje preprocessing k minimalizaci energie s příslušným parametrickým souborem `.mdp`, ve kterém je třeba prostřednictvím volby `freezegrps` zajistit znehybnění skupiny BULK.

```
grompp -f sd.mdp -p topol.top -c qbenzwati.gro -o qbenzwati.tpr -n index.ndx
```

9. Spuštění minimalizace energie.

```
mdrun -deffnm qbenzwati
```

10. Po minimalizaci energie je možné přikročit k preprocessingu k NVT ekvilibraci systému a spuštění NVT ekvilibrace.*

```
grompp -f nvt.mdp -p topol.top -c qbenzwati.gro -o qbenzwati.tpr -n index.ndx  
mdrun -deffnm qbenzwati
```

11. Dále je nutné v systému stabilizovat tlak prostřednictvím NPT ekvilibrace.*

```
grompp -f npt.mdp -p topol.top -c qbenzwati.gro -o qbenzwati.tpr -n index.ndx  
mdrun -deffnm qbenzwati
```

12. Pokud jsou hodnoty tlaku i teploty v systému stabilizovány, je proveden preprocessing k produkčnímu běhu simulace (dlouhý NVT běh) a následné spuštění dlouhé simulace.*

```
grompp -f nvt.mdp -p topol.top -c qbenzwati.gro -o qbenzwati.tpr -n index.ndx
mdrun -deffnm qbenzwati)
```

* V případě všech déle trvajících výpočtů je dobrým zvykem místo přímého spuštění příkazem `mdrun` použít dávkový soubor (pojmenovaný např. *open.sh*) s vhodně upravenými volbami. Spuštění úlohy probíhá pomocí příkazu:

```
qsub open.sh
```

Díky plánovacímu systému PBS (Portable Batch System) nebo také systému dávkového spouštění úloh bude úloha zařazena do fronty a zpracována s ohledem na sdílené výpočetní prostředky a zatížení systému. Podrobnější informace o obsahu dávkového souboru a použití PBS lze nalézt na http://meta.cesnet.cz/wiki/Plánovací_systém_PBS.

```
#PBS -N qbenzoic0.00
#PBS -m e
#PBS -l nodes=1:ppn=4:mbatch
#PBS -j oe
cat $PBS_NODEFILE
echo "Hostname:" `hostname`
echo "PBS_NODEFILE $PBS_NODEFILE"
cd /home/hb/bp/quartz-0.00/benzoic/nvt_job
mdrun -deffnm qbenzwati
```

Příklad dávkového souboru *open.sh* pro spuštění úloh na klastru UFY

```
#PBS -N qbenzoic0.00
#PBS -q jcu
#PBS -Wgroup_list=jcu
#PBS -l nodes=1:ppn=4
#PBS -j oe
cat $PBS_NODEFILE
cp -r /home/hb/bp/quartz-0.00/benzoic /scratch/hb
cd /scratch/hb/benzoic
module add gromacs-4.5.5-parallel
mpirun -np 4 mdrun_mpi -deffnm qbenzwati
cp -r /scratch/hb/benzoic /storage/home/hb
```

Příklad dávkového souboru *jcu.sh* pro spuštění úloh na klastru Hermes

Na klastru UFY lze k vytvoření analogického spustitelného souboru použít skript `gsub` od Mgr. Pavla Fibicha. Spuštěním `gsub` v adresáři se simulačními soubory (*.tpr* topologií) je vygenerován soubor *run.qsub*, který obsahuje veškeré potřebné údaje ke spuštění úlohy včetně označení poslední (nejnovější) *.tpr* topologie. V příslušném adresáři stačí tedy k vytvoření zmíněného dávkového souboru a spuštění simulace zadat následující dva příkazy:

```
gsub
qsub run.qsub
```

Při spouštění simulací na klastru Hermes je potřeba příslušné soubory nejprve zkopírovat ze školního stroje:

```
scp qbenzwati.tpr jcu.sh hermes:bp/quartz-0.00/benzoic
```

Poté stačí zadáním

```
qsub jcu.sh
```

na klastru Hermes spustit úlohu, která je dle výpočetních možností zpracována či zařazena do fronty.

4.2. Fáze a parametry simulací

Fáze provedených simulací lze obecně shrnout do čtyř po sobě následujících kroků:

1. Minimalizace
2. NVT ekvilibrace – 10000 kroků (20 ps), 300 K, Nosé-Hoover
3. NPT ekvilibrace – 10000 kroků (20 ps), 1 atm, Parrinello-Rahman
4. NVT dlouhý běh – 65000000 kroků (130 ns), 300 K, Nosé-Hoover

Všechny provedené simulace probíhaly s integračním krokem 2 fs, souřadnice byly zapisovány vždy po 1000 krocích. Počáteční velikost boxu před barostatováním byla $5,500 \times 3,982 \times 8,000$ nm.

V některých případech je po rychlém zhodnocení výsledků potřeba simulaci prodloužit a výsledné soubory z po sobě jdoucích simulací z důvodu snadnější manipulace a další analýzy sloučit. Jedná se zejména o výstupní trajektorie a energetické soubory. K tomu slouží nástroje `trjcat` a `eneconv`, které umožňují sloučení příslušných souborů s interaktivní volbou počátečního času (díky parametru `-settime`) každé z trajektorií, respektive každého z energetických souborů. Požadované soubory je nutno zadávat v témže pořadí, v jakém byly spuštěny příslušné simulace.

Sloučení trajektorií:

```
trjcat -f qbenzwati1.xtc qbenzwati2.xtc qbenzwati3.xtc -o trajektorie.xtc -settime
```

Sloučení energetických souborů:

```
eneconv -f qbenzwati1.edr qbenzwati2.edr qbenzwati3.edr -o energie.edr -settime
```

Některé analyzační utility (`g_dist`) vyžadují jako vstupní soubor `.tpr` topologii. Při analýze výsledků z několika po sobě následujících simulací jednoho konkrétního systému lze použít jediný `.tpr` soubor z libovolné části simulace – topologie systému a silové pole jsou totožné.

4.3. Přenos souborů

Jak jsem již zmínila v úvodu nadřazené kapitoly, analýzu výsledků simulací jsem prováděla na výpočetním klastru UFY. Z tohoto důvodu bylo potřeba výstupní soubory zkopírovat zpět na klastr UFY⁷:

```
scp qbenzwati.xtc ufy:bp/quartz-0.00/benzoic/nvt_job
tar cfz vysledky *.* --exclude=qbenzwati.xtc
scp vysledky ufy:bp/quartz-0.00/benzoic/nvt_job
```

a zde archiv rozbalit:

```
tar xfz vysledky
```

⁷ Z bezpečnostních důvodů neuvádím v příkladu kopírování výsledků skutečné jméno klastru

5. Výsledky

5.1. Seznámení s modelovanými systémy

Úkolem praktické části bakalářské práce bylo provedení simulací interakcí několika jednoduchých organických molekul (nebo jejich disociovaných stavů) s povrchem křemene o různé hustotě povrchového náboje. Na CD přiloženém k této práci jsou uloženy soubory s názvy *pka12.xls* a *pka12.doc*, které objasňují, jaké kombinace molekul s různě nabitým povrchem mají za daného pH smysl s ohledem na disociační konstantu (resp. konstanty v případě více OH skupin) pK_a molekuly. Simulace byly provedeny pro neutrální křemenný povrch a pro povrchy s hustotou náboje $-0,03 \text{ C}\cdot\text{m}^{-2}$ a $-0,12 \text{ C}\cdot\text{m}^{-2}$. S ohledem na výpočty ve výše jmenovaných souborech jsem modelovala a simulovala kombinace uvedené v tabulce Tab. 3.

Molekula	Povrch – hustota náboje [$\text{C}\cdot\text{m}^{-2}$]		
	0,00	-0,03	-0,12
benzoic	✓	✗	✗
benzoate	✓	✓	✓
phenol	✓	✓	✓
phenolate	✗	✗	✓
salicylic_1neg	✓	✓	✓

Tab. 3: Modelované kombinace

V případě povrchové hustoty náboje $0,00 \text{ [C}\cdot\text{m}^{-2}]$ se jedná o neutrální (nenabitý) povrch. Struktura použitého křemenného povrchu včetně ostatních, záporně nabitých povrchů, je popsána v kapitole 5.1.1. Geometrie molekul a označení jednotlivých atomů jsou uvedeny v kapitole 5.1.2. Molekuly označuji jejich pracovními názvy vyplývajícími z anglického názvosloví. Zde uvádím jejich ekvivalentní česká pojmenování:

benzoic = kyselina benzoová

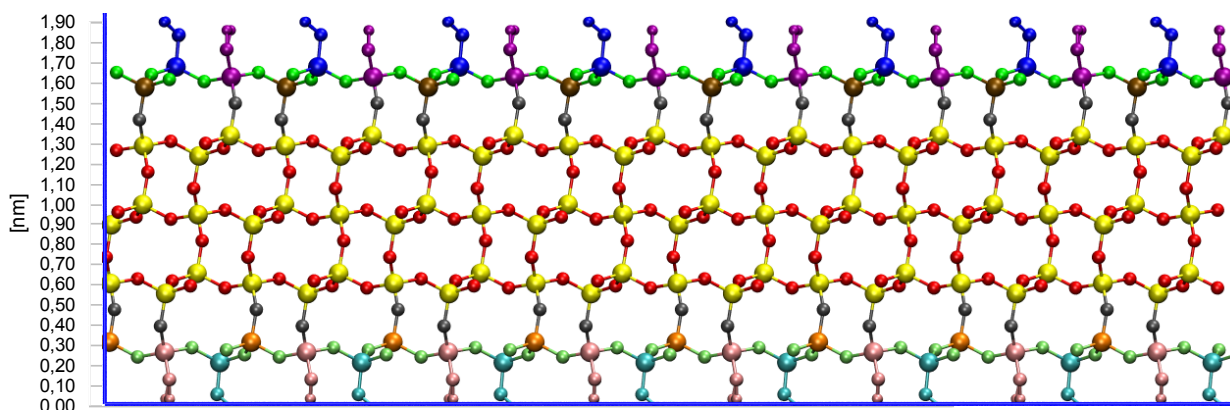
benzoate = benzoát; deprotonovaná forma kyseliny benzoové

phenol = fenol (zastarale kyselina karbolová, hydroxybenzen nebo benzenol)

phenolate = fenolát; deprotonovaná forma fenolu

salicylic_1neg = jednou deprotonovaná kyselina salicylová (deprotonovaná COO^- skupina)

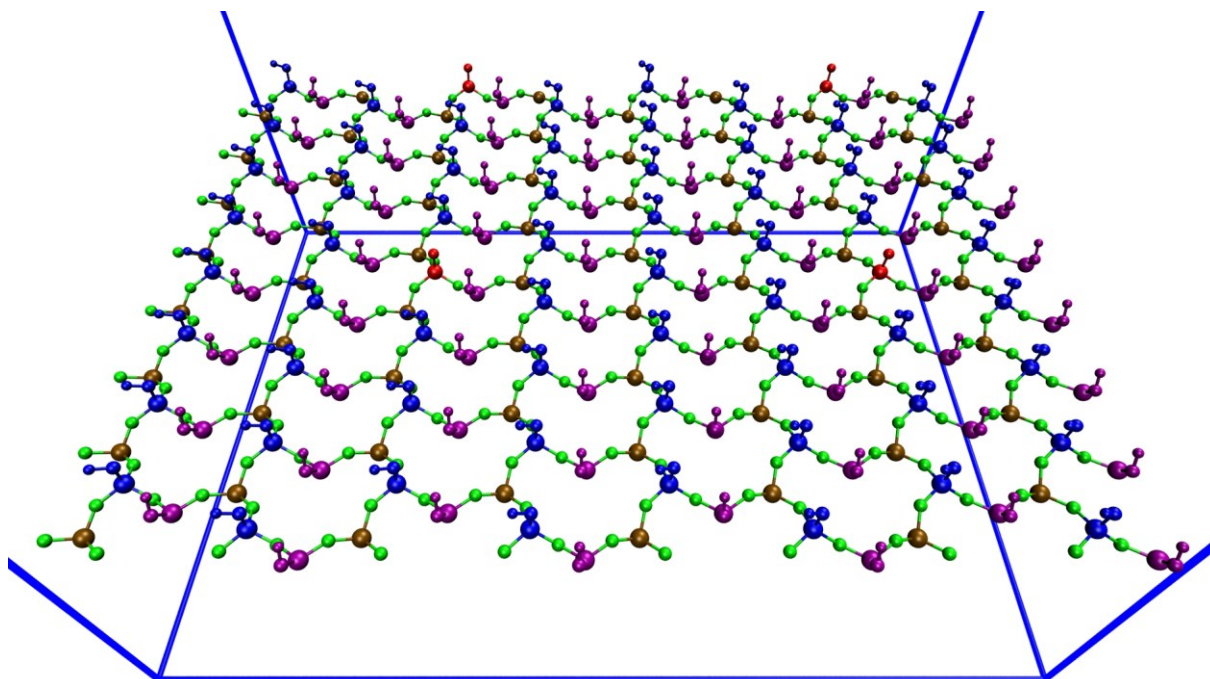
5.1.1. Křemenný povrch



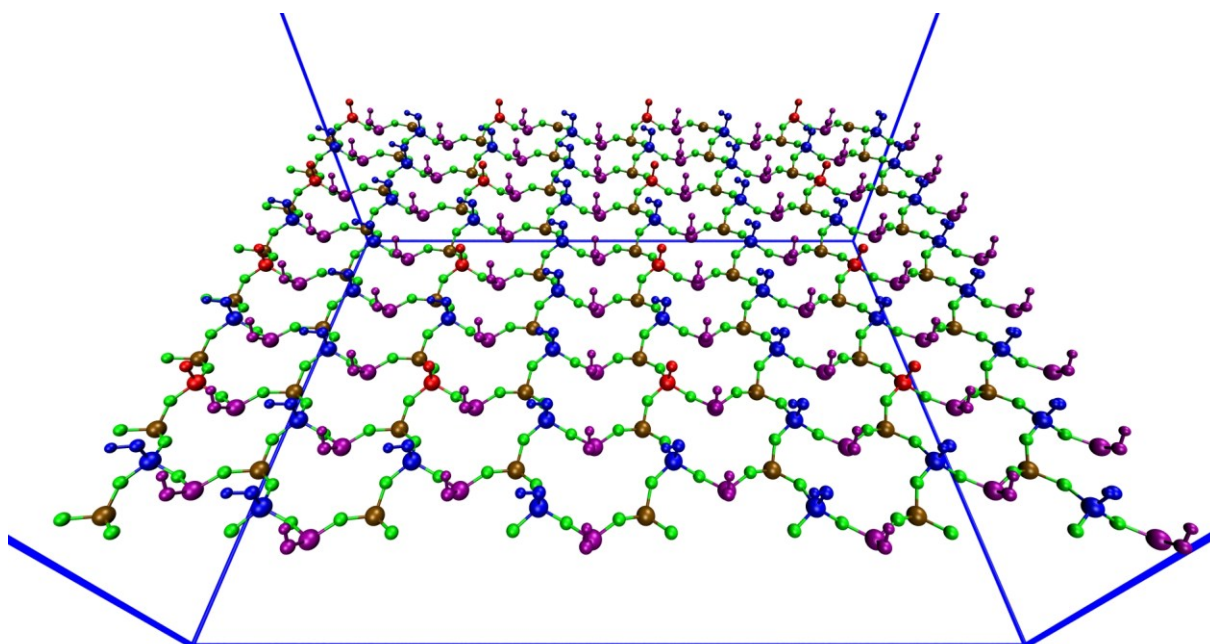
Obr. 14: Neutrální křemenný povrch s rozlišenými skupinami atomů a residuí

●	HU1	vodík navázaný na OU1	}	residuum SU1
●	OU1	kyslík navázaný na atom SU1		
●	SU1	nejvýše položený křemík 1. povrchu		
●	HD1	vodík navázaný na OD1	}	residuum SD1
●	OD1	kyslík navázaný na atom SD1		
●	SD1	níže položený křemík 1. povrchu		
●	OS1	kyslík 1. povrchu; spojuje SD1 s SU1 a SS1		
●	SS1	křemík 1. povrchu bez residua SU1		
●	OBS	kyslík mezi BULKem a povrchovými atomy		
●	SiB	bulkový křemík	}	BULK
●	OB	bulkový kyslík		
●	SS2	křemík 2. povrchu bez residua SU2		
●	OS2	kyslík 2. povrchu; spojuje SD2 s SU2 a SS2		
●	SD2	níže položený křemík 2. povrchu	}	residuum SU2
●	OD2	kyslík navázaný na atom SD2		
●	HD2	vodík navázaný na OD2		
●	SU2	nejvýše položený křemík 2. povrchu	}	residuum SD2
●	OU2	kyslík navázaný na atom SU2		
●	HU2	vodík navázaný na OD2		

Popsaná křemenná struktura obsahuje pět vertikálních vrstev atomů tvořených křemíkovými atomy se čtyřmi mezivrstvami kyslíků. Celý krystal je rozdělen na dva povrchy: první (horní) a druhý (dolní). Atribut níže nebo výše položený je hodnocen z hlediska středové osy – níže položený znamená blíže středové ose a naopak. Residuum SU1 se skládá ze tří atomů: křemíku SU1 označeného shodně s celým residuem, kyslíku OU1 a vodíku HU1. Analogicky jsou tvořena i zbylá residua SU2, SD1 a SD2. Druhé písmeno v názvu residua koresponduje s polohou celého tripletu v povrchu; písmeno U (*up*) značí atomy vzdálenější od středové osy, D (*down*) atomy bližší středové ose. Nejvýše položené skupiny atomů (HU1, HU2) představují atomy nejvzdálenější středové ose celé struktury. Všechny atomy povrchu jsou zahrnuty v jedné skupině pod názvem SURF. BULK je označení pro atomy třech středových vertikálních vrstev křemíků (SiB) se dvěma mezivrstvami kyslíků (OB). Skupina BULK je během všech fází simulace znehybněna pomocí parametru *freeze* v příslušných parametrických souborech z důvodu zachování geometrické struktury celého povrchu a zabránění jeho posunu uvnitř boxu.



Obr. 15: Svrchní vrstva křemenného povrchu o povrchové hustotě náboje $-0,03 \text{ C} \cdot \text{m}^{-2}$. Červeně jsou označena residua SUC (C jako „charged“), tedy původní residua SU1 bez atomů HU1 (kyslíky původně označené OU1 jsou nazvány OUC). Celkem 4 deprotonovaná residua jsou rozmístěna s ohledem na rovnoměrné rozložení náboje a periodické okrajové podmínky.

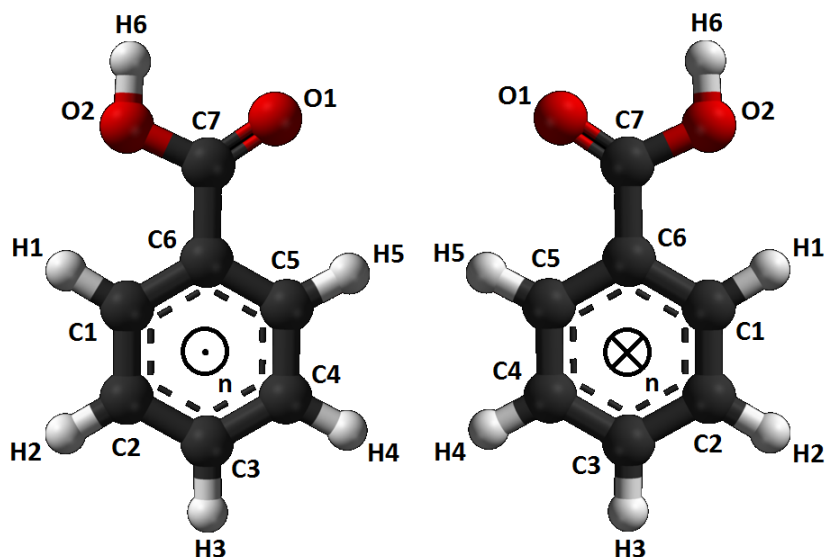


Obr. 16: Svrchní vrstva křemenného povrchu o povrchové hustotě náboje $-0,12 \text{ C} \cdot \text{m}^{-2}$. V tomto případě je deprotonováno 16 původních residuí SU1.

5.1.2. Molekuly

Tato kapitola se věnuje popisu použitých organických molekul (viz Tab. 3). Uvedené obrázky slouží zejména pro představu základní struktury molekul a k popisu jednotlivých atomů, jejichž číslování a pořadí je důležité z hlediska pozdějšího utváření vektorů při analýze adsorpčních okamžiků a definice normály každé z molekul. Stejně označení nesou atomy i v topologiích, souřadnicových a indexových souborech. Orientace normály molekul, definovaná vektorovým součinem dvou vektorů dle vzorce $\overrightarrow{A_1 A_2} \times \overrightarrow{A_1 A_3}$, kde index atomu A označuje pořadí atomu, je dána vzestupným pořadím příslušných atomů v indexovém souboru (při definování skupin jsou atomy automaticky řazeny vzestupně dle svého číselného označení). Podobně určuje pořadí atomů orientaci vektorů v rámci molekul.

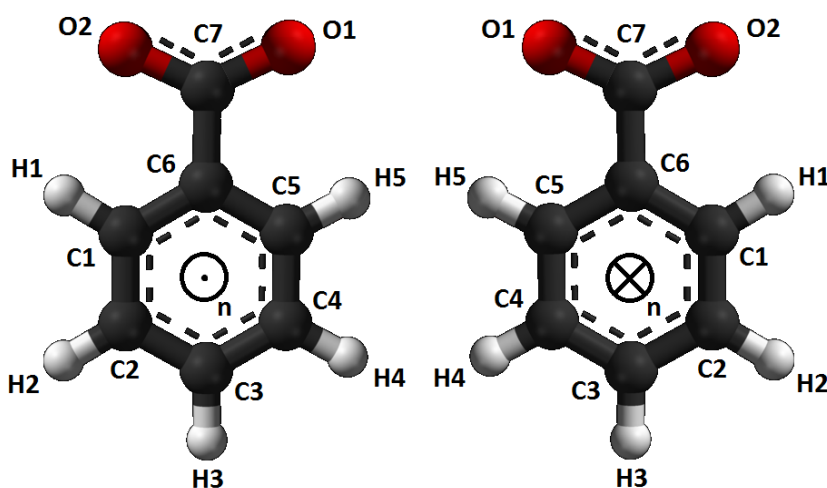
Zavedené skupiny atomů použité k analýze adsorpčních geometrií jsou popsány v tabulkách pod obrázky molekul. Značení atomů odpovídá obrázku příslušné molekuly.



Obr. 17: Kyselina benzoová [36]

Kyselina benzoová	
skupina	použití
O1	kyslík COOH skupiny s dvojnou vazbou
O2	kyslík COOH skupiny s navázaným H6
H6	vodík COOH skupiny
H15	vodíky aromatického kruhu
C1_C3_C5	uhlíky aromatického kruhu definující normálu
C3_C7	vektor natočení COOH k povrchu

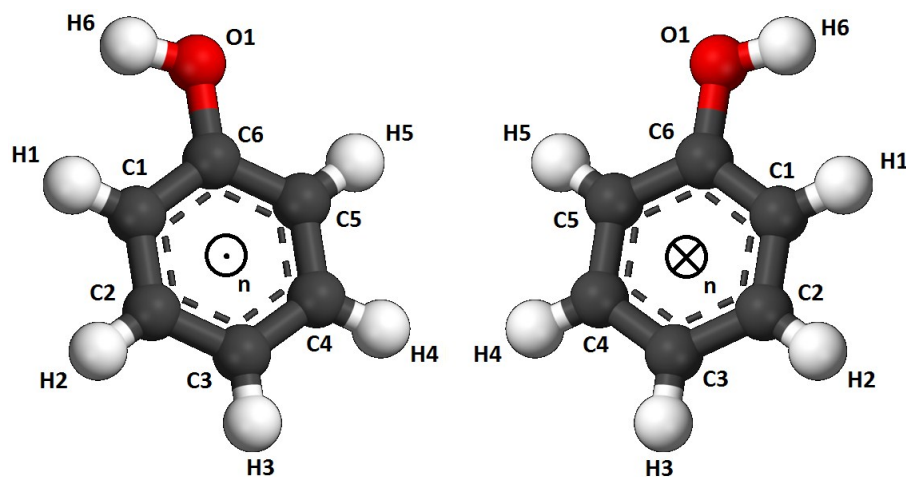
Tab. 4: Indexové skupiny kyseliny benzoové



Obr. 18: Benzoát (deprotonovaná kyselina benzoová) [36]

Benzoát	
skupina	použití
O12	kyslíky COO ⁻ skupiny
H15	vodíky aromatického kruhu
C1_C3_C5	uhlíky aromatického kruhu definující normálu
C3_C7	vektor natočení COO ⁻ k povrchu

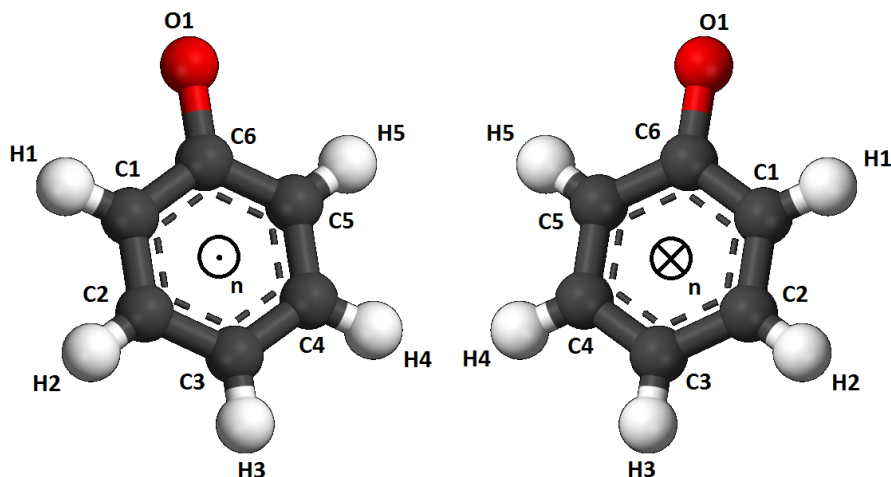
Tab. 5: Indexové skupiny benzoátu



Obr. 19: Fenol [37]

Fenol	
skupina	použití
O1	kyslík OH skupiny
H6	vodík OH skupiny
H15	vodíky aromatického kruhu
C1_C3_C5	uhlíky aromatického kruhu definující normálu
C3_C6	vektor natočení OH k povrchu

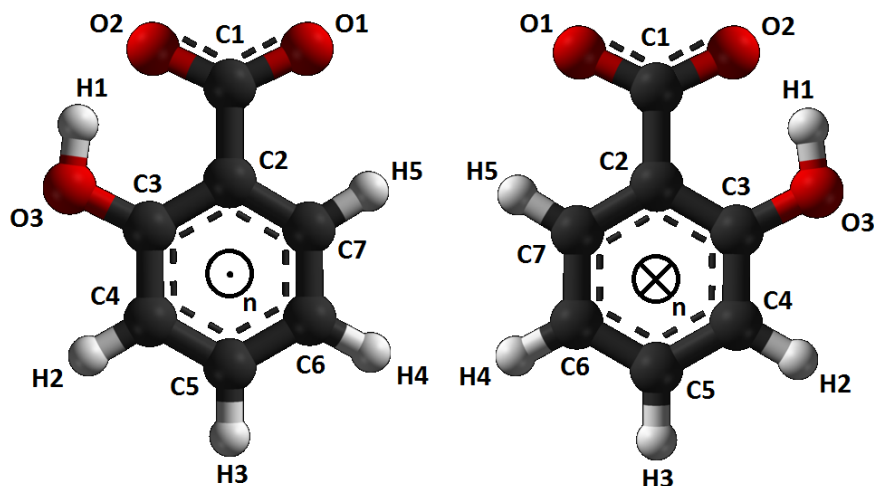
Tab. 6: Indexové skupiny fenolu



Obr. 20: Fenolát [37]

Fenolát	
skupina	použití
O1	kyslík deprotonované OH skupiny
H15	vodíky aromatického kruhu
C1_C3_C5	uhlíky aromatického kruhu definující normálu
C3_C6	vektor natočení deprotonované OH k povrchu

Tab. 7: Indexové skupiny fenolátu



Obr. 21: Deprotonovaná kyselina salicylová [38]

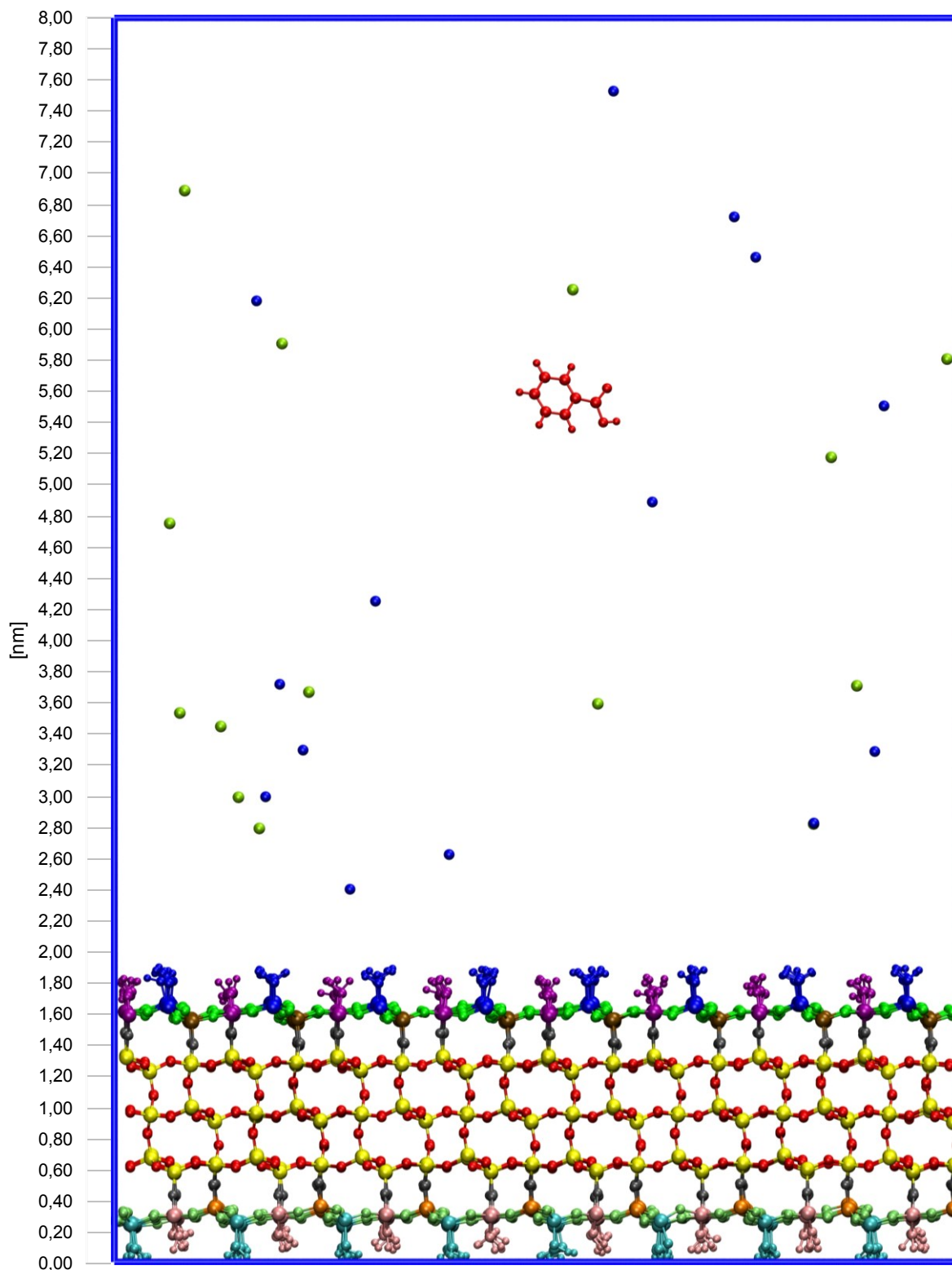
Deprotonovaná kyselina salicylová	
skupina	použití
O12	kyslíky COO ⁻ skupiny
H1	vodík OH skupiny
O3	kyslík OH skupiny
H25	vodíky aromatického kruhu
C3_C5_C7	uhlíky aromatického kruhu definující normálu
C1_C5	vektor natočení COO ⁻ k povrchu
C3_C6	vektor natočení OH k povrchu

Tab. 8: Indexové skupiny deprotonované kyseliny salicylové

5.2. Molekula kyseliny benzoové na neutrálním křemenném povrchu

Na kombinaci molekuly kyseliny benzoové a neutrálního křemenného povrchu popíše schéma analýzy simulovaného systému. Analogický postup by mohl být uplatněn k analýze všech kombinací, výsledky z ostatních simulací však z důvodu prostorových nároků uvádím v redukované podobě, bez použitých příkazů, např. formou srovnání hustotních profilů pro různě nabitě povrchy. Jednotlivé kroky následujícího postupu jsou uvedeny se zřetelem na

jejich pořadí při samotné analýze. Analýza výsledků byla provedena z kanonických NVT souborů.



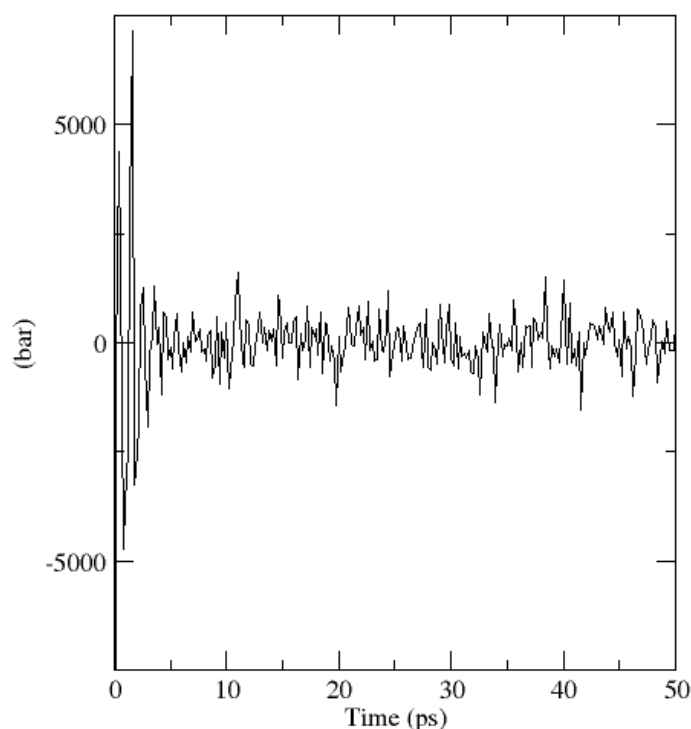
Obr. 22: Ukázka simulačního boxu – neutrální křemenný povrch s molekulou kyseliny benzoové a celkem 28 ionty Na⁺ (modré) a Cl⁻ (zelené) ve vodě (nezobrazena)

Ekvilibrace - ustálení tlaku a teploty

Tlak je obecně tenzorová veličina (matice 3×3), jejíž nediagonální složky by v kapalině v rovnováze měly být nulové a diagonální by se měly rovnat skalárnímu tlaku. Část systému je však tvořena krystalem křemene, navíc obsahujícím skupinu zcela polohově zafixovaných atomů. Z důvodu barostatování pouze v jednom směru (pouze ve směru osy z , aby nedocházelo k posunu a deformaci povrchu) se v ostatních směrech rozměry boxu nemění a odpovídající složky tlakového tenzoru se díky příspěvkům z krystalu mohou lišit. Z tohoto důvodu je nutné kreslit graf pro ověření hodnoty tlaku v systému pouze ve směru osy z (*Pres-ZZ*), nikoliv jako skalární tlak (*Pressure*). Ustálení teploty a tlaku je dosaženo simulací v izotermicko-izobarickém NPT souboru.

```
g_energy -f qbenzwati.edr -s topol.tpr -o press.xvg
          volba Pres-ZZ
xmgrace press.xvg
```

Gromacs Energies

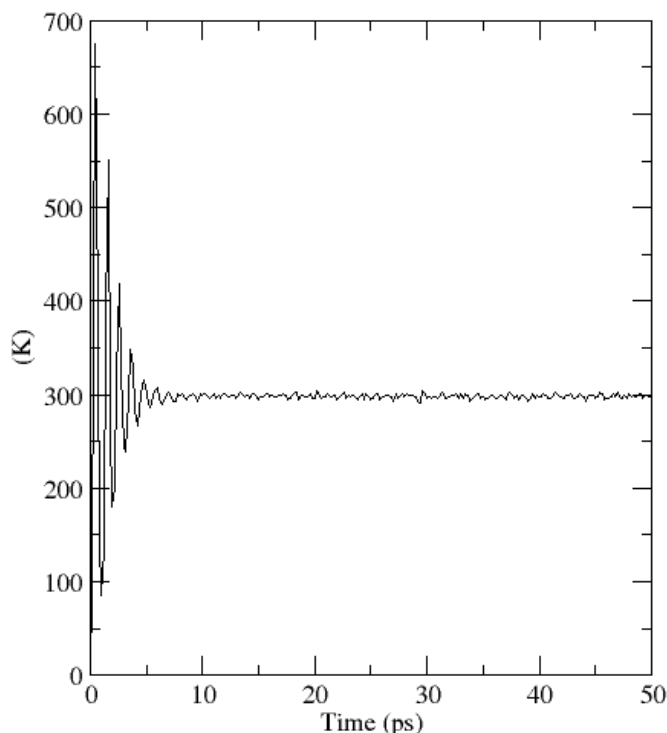


Obr. 23: Ustálení tlaku v systému při ekvilibraci

Ustálení teploty lze prozkoumat pomocí podobných příkazů:

```
g_energy -f qbenzwati.edr -s topol.tpr -o temp.xvg
          volba Temperature
xmgrace temp.xvg
```


Gromacs Energies



Obr. 24: Ustálení teploty v systému při ekvilibraci

Z obou grafů je patrné, že k ustálení tlaku i teploty došlo velmi rychle – během prvních cca 10 ps ekvilibrace. V průběhu simulace však nadále dochází k fluktuacím okamžitých hodnot obou těchto veličin.

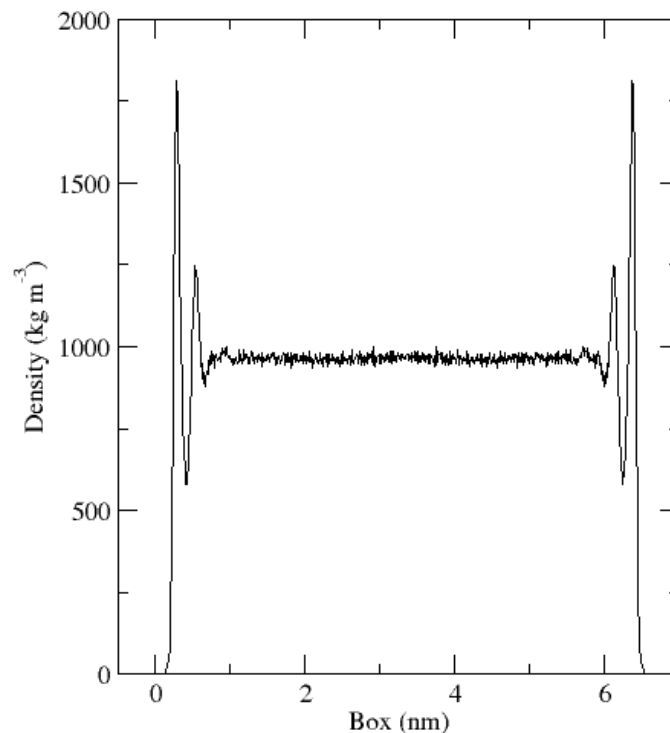
Hustota vody

Hustota vody počítaná ve směru osy z poskytuje histogram hustoty ve směru kolmém k povrchu. Parametr `-sl` udávající počet dílků, na které je box rozdělen (tzv. sliců), je nutno volit s ohledem na celkovou výšku simulačního boxu: box o výšce 8 nm vydělený počtem sliců (v tomto případě 1000) udává tzv. šířku binu – vzdálenost, pro kterou je hustota počítána (0,08 Å). Délka vazeb je typicky 1 až 2 Å, přesnost 0,08 Å je tedy plně dostačující. Příkaz `g_sym` je použit z důvodu zprůměrování hodnot hustoty u obou ekvivalentních neutrálních povrchů (vysvětleno dále).

```
g_density -f qbenzwati.trr8 -s qbenzwati.tpr -n index.ndx -o densH2O.xvg -sl 1000
          volba SOL
g_sym -f densH2O.xvg -o densH2O_sym_both.xvg -walls 1.68 0.24 -all -both
xmgrace densH2O_sym_both.xvg
```

⁸ Z důvodů uvedených v kapitole 2.2.3.4. byla použita trajektorie `.trr`

Partial densities



Obr. 25: Hustota vody jako funkce vzdálenosti od povrchu

Další užitečné volby utility `g_density`:

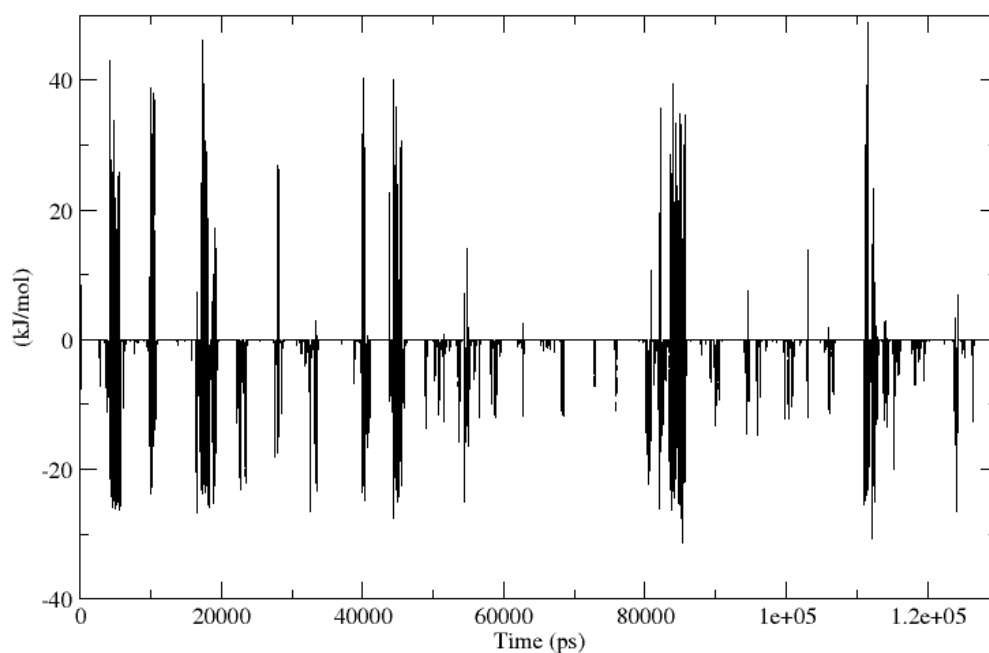
- dt dělí počet snímků daným číslem, určeno pro rychlejší průběh zpracování
- dens hustota: mass = podle rozložení hmotnosti (standardní „hmotnostní“ hustota), number = podle počtu částic („číselná“ hustota)
- sl počet sliců boxu = pomyslných vrstev, na které je box rozdělen. Má vliv na „jemnost“ výsledné funkce. Defaultní hodnota je 50 sliců.
- ng počet skupin, pro které se funkce počítá; jedna křivka pro každou skupinu

Adsorpční energie

Grafy adsorpčních energií mezi molekulou a povrchem jsou užitečným vodítkem při určování adsorpčních okamžiků. Interakce mezi molekulou a povrchem lze kvantifikovat pomocí energie Lennard-Jonesova nebo Coulombova potenciálu.

```
g_energy -f energie.edr -s topol.tpr -o energy_LJ.xvg
        volba LJ-SR:MOL-SURF
xmgrace energy_LJ.xvg
```

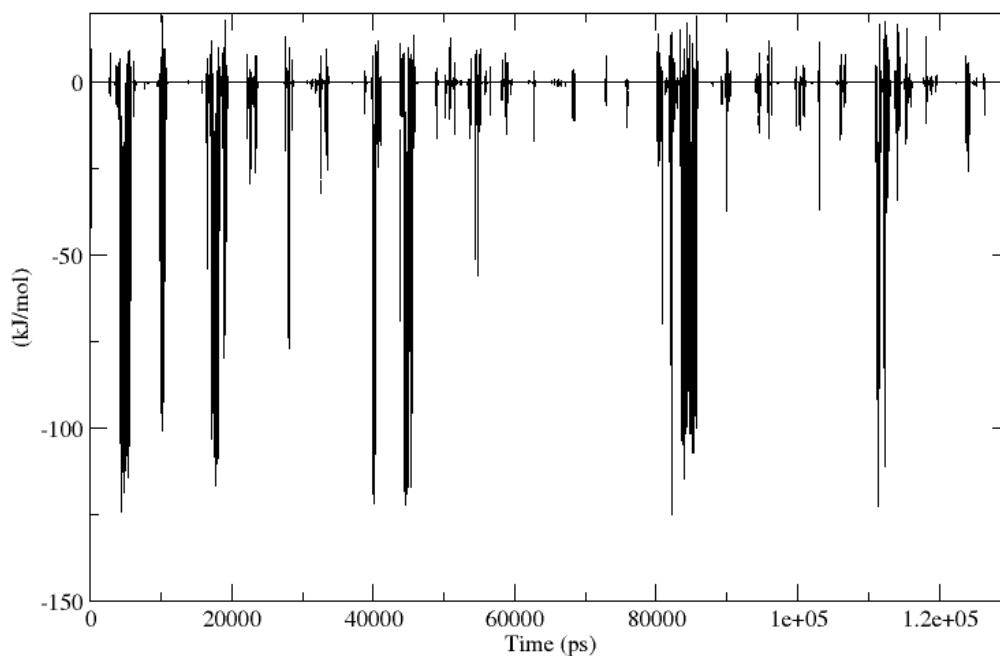
Gromacs Energies



Obr. 26: Van der Waalsovy interakce modelované pomocí Lennard-Jonesova potenciálu

```
g_energy -f energie.edr -s topol.tpr -o energy_C.xvg  
volba Coul-SR:MOL-SURF  
xmgrace energy_C.xvg
```

Gromacs Energies



Obr. 27: Elektrostatické interakce modelované pomocí Coulombova potenciálu

Oba grafy silně korelují a ukazují na stejné adsorpční okamžiky. Během adsorpce je typické, že energie elektrostatických interakcí E_C je výrazně nižší než energie vdW interakcí E_{LJ} . Z energetického hlediska je pro atomy výhodnější lehce odpudivé vdW přiblížení (kladná E_{LJ}) vedoucí k výrazně výhodnější elektrostatické interakci (záporná E_C).

Úprava indexového souboru

Úpravu indexového souboru je nutné provést před další analýzou a prací s utilitami `g_density`, `g_dist`, `g_rdf` aj., kde je potřeba v jednotlivých případech specifikovat požadované skupiny atomů. Editace indexového souboru spočívá v načtení stávajícího souboru s automaticky generovanými skupinami atomů, které byly vytvořeny na základě údajů v souřadnicových souborech při preprocessingu, a následné úpravě a přidání dalších skupin. Typicky byla u všech modelovaných systémů vytvořena skupina SURF, která reprezentuje všechny atomy křemenného povrchu, a skupina BULK představující atomy třech středových vertikálních vrstev křemíků povrchu se dvěma mezivrstvami kyslíků. Dále byly u systémů s neutrálním povrchem zavedeny referenční skupiny atomů pro horní i dolní povrch (nejvýše položené atomy křemíku SU1 a SU2) a skupiny OUD (zahrnuje residua OU1, OU2, OD1 a OD2) a HUD (HU1, HU2, HD1 a HD2). V případě systémů s nabitým povrchem byla provedena analýza pouze pro horní (nabitý) povrch; skupiny OUD a HUD byly nahrazeny skupinami OUD1 (OU1, OD1) a HUD1 (HU1, HD1). Navíc byla rozlišena skupina OUC představující deprotonované kyslíky povrchu. Pro jednotlivé molekuly byly zavedeny indexové skupiny uvedené v kapitole 5.1.2.

Úprava indexového souboru s použitím konfiguračního souboru se provádí pomocí příkazu:

```
make_ndx -f qbenzwati.gro -n index.ndx
```

Editaci skupin lze provést pomocí následujících vzorových příkazů:

<code>a HU1 a HD1</code>	vytvoření a sloučení 2 skupin atomů povrchu HU1 a HD1
<code>22 23</code>	sloučení skupin č. 22 a 23
<code>a SU1</code>	vytvoření referenční skupiny povrchu obsahující atomy SU1
<code>del 39</code>	smazání skupiny č. 39
<code>name 41 C1_C3_C5</code>	přejmenování skupiny č. 41 na C1_C3_C5

U modelového příkladu kyseliny benzoové na neutrálním křemenném povrchu jsem oproti ostatním systémům rozlišila vyšší (U) a nižší (D) skupiny atomů horního a dolního povrchu

a sloučila odpovídající si skupiny obou ekvivalentních povrchů vždy do jedné bez číselného indexu:

OU1 a OU2 zahrnuty do skupiny OU

HU1 a HU2 zahrnuty do skupiny HU

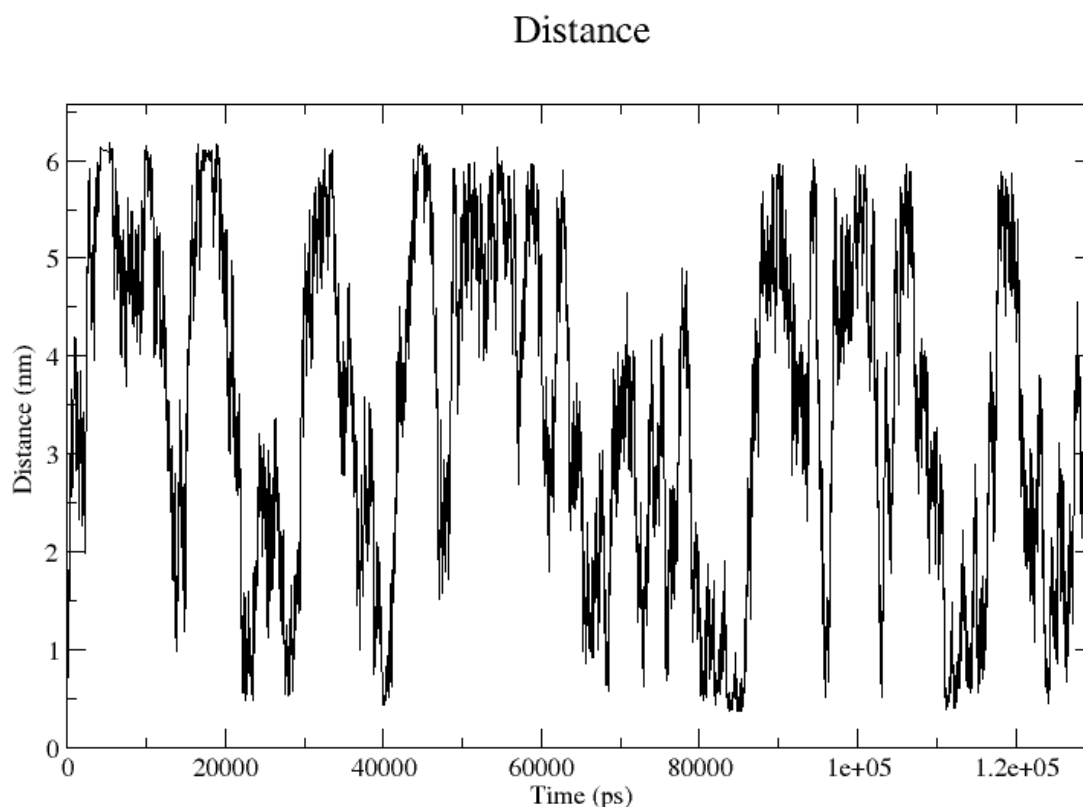
OD1 a OD2 zahrnuty do skupiny OD

HD1 a HD2 zahrnuty do skupiny HD

Trajektorie molekuly a její vzdálenost od povrchu

Chování molekuly lze dobře popsat pomocí časového vývoje její vzdálenosti od povrchu. Utilita `g_dist` počítá vzdálenosti těžišť zvolených skupin atomů jako funkci času, přesněji zaznamenává x , y , z složky vzdálenosti a velikost prostorově orientovaného vektoru \vec{r} . Z těchto údajů je programem `g_distmp` zpracována pouze složka z . Obr. 28 ukazuje, že molekula několikanásobně putovala mezi oběma povrchy, na které po kratší dobu adsorbovala.

```
g_dist -f trajektorie.xtc -s qbenzwati.tpr -n index.ndx
    1. volba MOL
    2. volba SU1
g_distmp dist.svg
xmgrace distmp.svg
```



Obr. 28: Časový vývoj vzdálenosti těžiště molekuly od těžiště povrchové skupiny SU1

Hustotní profil molekuly

Hustotní profil molekuly udává, jak často se molekula vyskytovala v určité výšce – např. kolikrát častěji se nacházela u povrchu oproti střední části boxu. Hustotní profily molekul počítá utilita `g_density`, jejíž výstup je vhodné upravit pomocí programů `g_sym` nebo `g_shift`. Pro neutrální povrch jsem použila program `g_sym` ke zprůměrování a symetrizaci profilů (horní i dolní povrch jsou v případě neutrálního křemene ekvivalentní; zadává se výška skupin SU1 a SU2. Je-li díky periodickým okrajovým podmínkám výška skupiny SU2 nižší než SU1, což nastává (viz Obr. 22), `g_sym` automaticky přičte k zadané výšce druhé referenční skupiny výšku boxu L_z známou z `.xvg` souboru). U nabitých povrchů jsem aplikovala `g_shift`, který posouvá výsledný profil o zvolenou hodnotu (zadává se pouze výška referenční skupiny SU1, takže opět poloha molekul je udávána od roviny skupiny SU1).

Postup pro neutrální povrch:

```
g_density -f trajektorie.xtc -s qbenzwati.tpr -n index.ndx -o densMOL.xvg -sl 1000  
      volba MOL
```

```
g_density -f trajektorie.xtc -s qbenzwati.tpr -n index.ndx -o densSU1.xvg -sl 1000  
      volba SU1
```

```
xmgrace densSU1.xvg      určení výšky referenční skupiny SU1 (první číselný údaj  
                        parametru -walls příkazu g_sym)
```

```
g_density -f trajektorie.xtc -s qbenzwati.tpr -n index.ndx -o densSU2.xvg -sl 1000  
      volba SU2
```

```
xmgrace densSU2.xvg      určení výšky referenční skupiny SU2 (druhý číselný údaj  
                        parametru -walls příkazu g_sym)
```

```
g_sym -f densMOL.xvg -o densMOL_sym_both.xvg -walls 1.68 0.24 -all -both  
xmgrace densMOL_sym_both.xvg
```

Postup pro nabité povrchy:

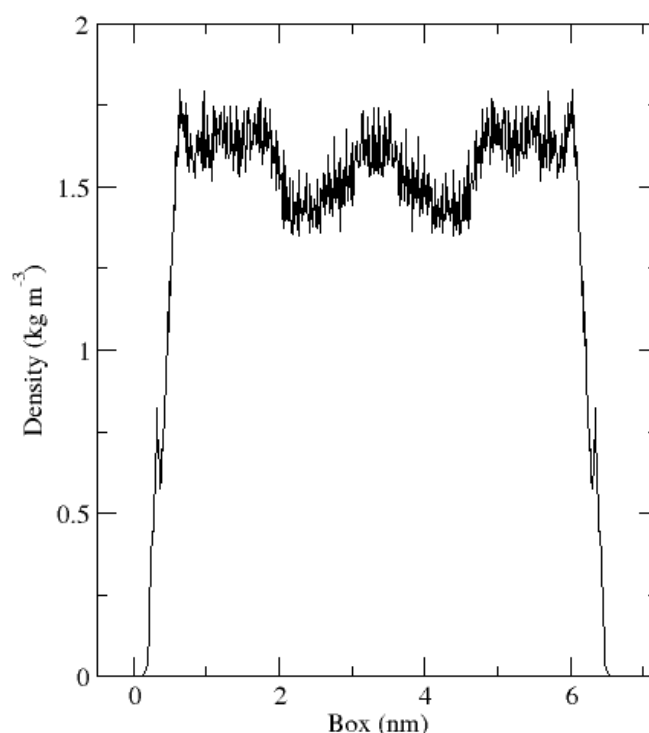
```
g_density -f trajektorie.xtc -s qbenzwati.tpr -n index.ndx -o densMOL.xvg -sl 1000  
      volba MOL
```

```
g_density -f trajektorie.xtc -s qbenzwati.tpr -n index.ndx -o densSU1.xvg -sl 1000  
      volba SU1
```

```
xmgrace densSU1.xvg      určení výšky referenční skupiny SU1 (absolutní hodnota  
                        parametru -shift příkazu g_shift)
```

```
g_shift -f densMOL.xvg -o densMOL_shift.xvg -shift -1.67  
xmgrace densMOL_shift.xvg
```

Partial densities



Obr. 29: Zesymetrizovaný hustotní profil molekuly kyseliny benzoové na neutrálním povrchu

Radiální distribuční funkce

Radiální distribuční funkce (zkráceně RDF) udává, s jakou pravděpodobností se nacházejí dvě částice v definované vzájemné vzdálenosti. Typický průběh RDF je znázorněn na Obr. 13.

```
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o HU_O1.xvg -n index.ndx -bin 0.005
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o HU_O2.xvg -n index.ndx -bin 0.005
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o OU_H6.xvg -n index.ndx -bin 0.005
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o OU_H15.xvg -n index.ndx -bin 0.005
```

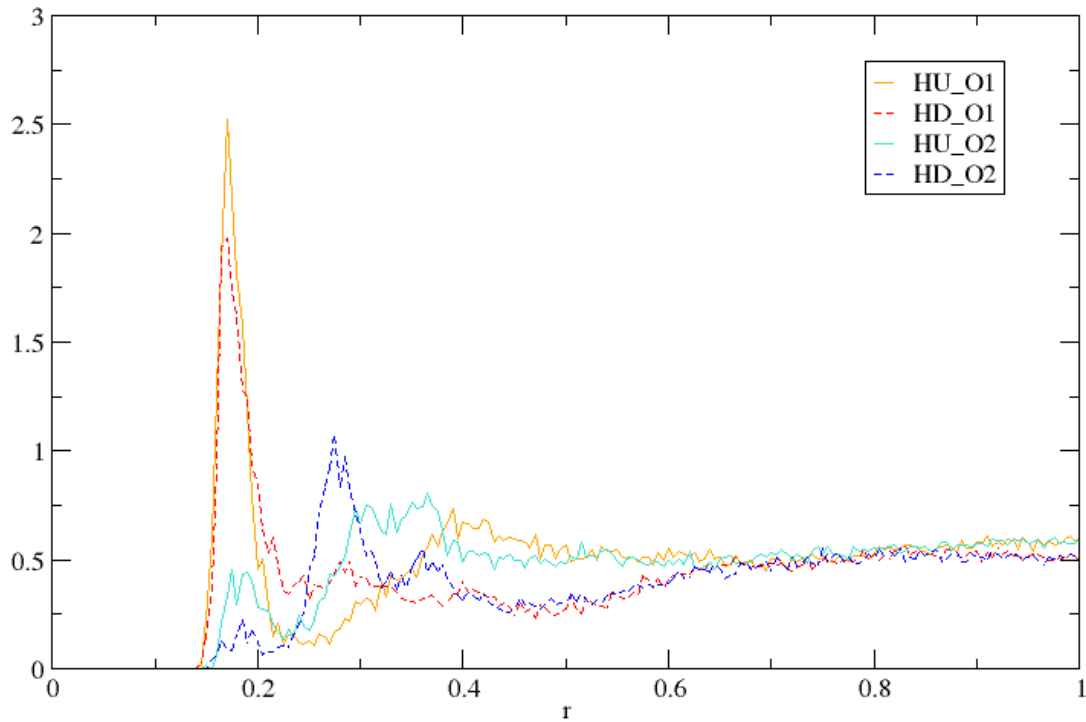
```
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o HD_O1.xvg -n index.ndx -bin 0.005
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o HD_O2.xvg -n index.ndx -bin 0.005
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o OD_H6.xvg -n index.ndx -bin 0.005
g_rdf -f trajektorie.xtc -s qbenzwati.tpr -o OD_H15.xvg -n index.ndx -bin 0.005
```

Další užitečné volby utility `g_rdf`:

<code>-bin</code>	rozlišení, šířka binu; defaultní hodnota je 0,002 [nm]
<code>-ng</code>	počet skupin, pro které se funkce počítá; jedna křivka pro každou skupinu
<code>-dt</code>	dělí počet snímků daným číslem, určeno pro rychlejší průběh zpracování

xmgrace -nxy HU_O1.xvg HD_O1.xvg HU_O2.xvg HD_O2.xvg

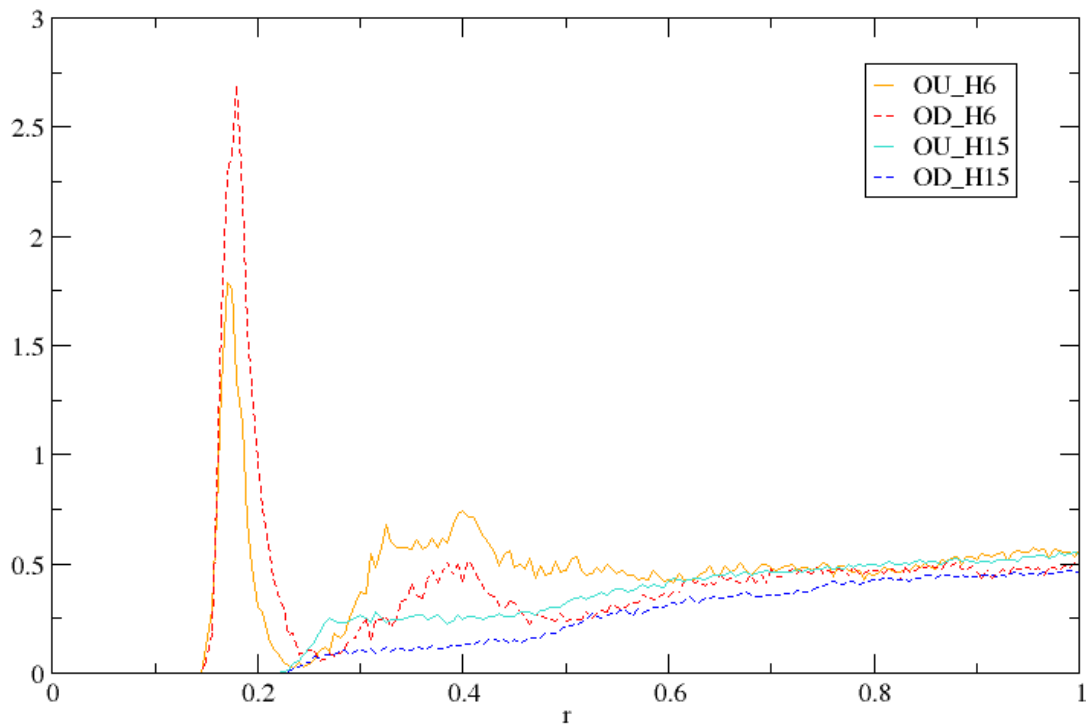
Radial distribution



Obr. 30: RDF pro páry HU_O1 a HD_O1, HU2_O2 a HD2_O2

xmgrace -nxy OU_H6.xvg OD_H6.xvg OU_H15.xvg OD_H15.xvg

Radial distribution



Obr. 31: RDF pro páry OU_H6 a OD_H6, OU_H15 a OD_H15

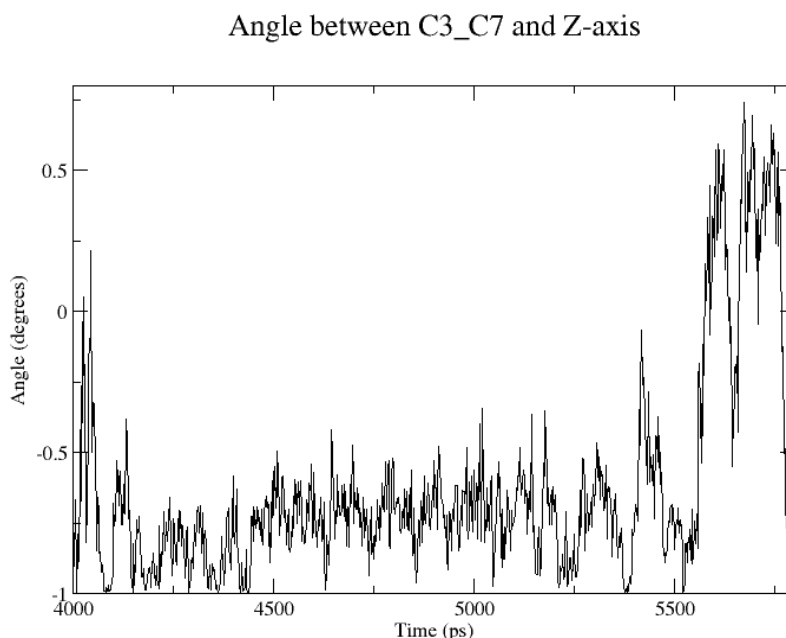
Adsorpční geometrie – natočení molekuly vzhledem k povrchu

Pro zjištění natočení molekuly vzhledem k povrchu během adsorpčních okamžiků jsem využila utilitu `g_sgangle`. Zkoumala jsem jak natočení funkčních skupin jednotlivých molekul, tak orientaci těchto molekul danou normálou aromatického kruhu. Pro úhel natočení normály molekul vzhledem k ose *z*, respektive k povrchu, nebyla zjištěna žádná významnější závislost, tyto výsledky tedy ve své práci vůbec neuvádím. Natočení karboxylové skupiny kyseliny benzoové je dáno orientací vektoru C3_C7 vzhledem k ose *z*, respektive k povrchu, jak ukazuje Tab. 9:

Povrch	Vektor C3_C7	
	cos	orientace
Horní (SU1)	+	COOH k povrchu
	-	COOH od povrchu
Dolní (SU2)	+	COOH od povrchu
	-	COOH k povrchu

Tab. 9: Orientace vektoru C3_C7 kyseliny benzoové vzhledem k povrchu v závislosti na cosinu úhlu⁹

```
g_sgangle -f trajektorie.xtc -s qbenzwati.tpr -oa lz5podel.xvg -n index.ndx -z -b  
4000 -e 5800  
volba C3_C7
```



Obr. 32: Vzájemné natočení vektoru C3_C7 a osy *z*. V době adsorpce (od 4200 do 5600 ps), kdy byla molekula u spodního povrchu, značí záporný cosinus natočení COOH skupiny směrem k povrchu.

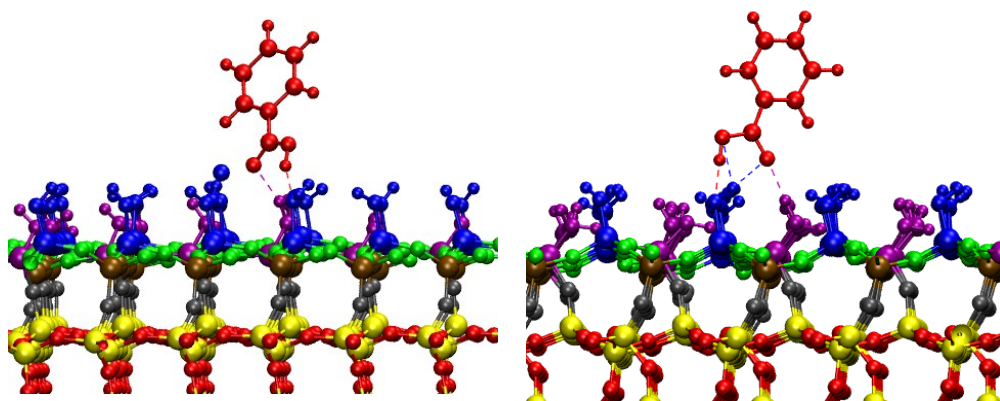
⁹ Díky očíslování atomů lze uvedenou tabulku využít k určování natočení funkčních skupin i u ostatních použitých molekul kromě kyseliny salicylové.

Shrnutí adsorpčních časů kyseliny benzoové na neutrálním povrchu

Povrch	Adsorpční čas v [ps]	Pořadí snímků	Doba adsorpce v [ps]
2	4200 - 5600	2100 - 2800	1400
2	9900 - 10550	4950 - 5275	650
2	16500 - 19200	8250 - 9600	2700
1	22200 - 23500	11100 - 11750	1300
1	27500 - 28600	13750 - 14300	1100
2	32300 - 33600	16150 - 16800	1300
1	39800 - 40400	19900 - 20200	600
2	44500 - 45500	22250 - 22750	1000
1	83300 - 85600	41650 - 42800	2300
1	111000 - 113000	55500 - 56500	2000
1			7300
2			7050
Celkem			14350

Tab. 10: Adsorpční okamžiky molekuly kyseliny benzoové s neutrálním povrchem

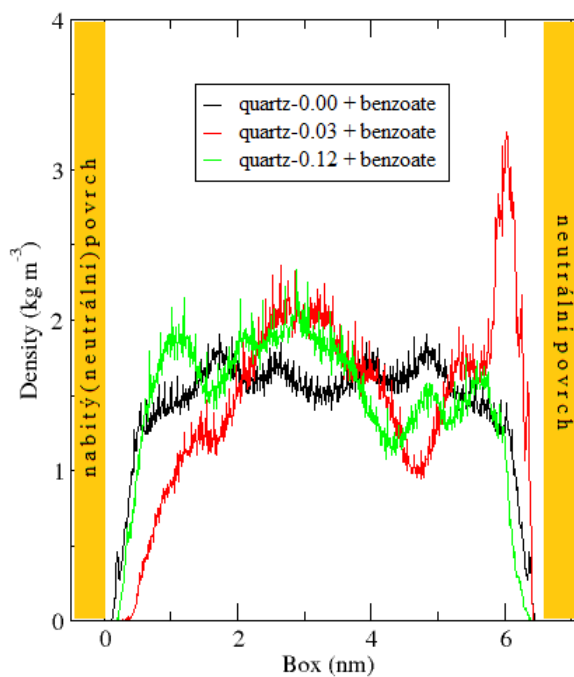
Molekula kyseliny benzoové interagovala s neutrálním křemenným povrchem převážně COOH skupinou (prostřednictvím atomů O1 a H6). Po většinu času byla plně solvatovaná, k výraznější adsorpci docházelo v časech uvedených v tabulce Tab. 10 (celkem 10 – 12% z celkového simulačního času). Z grafů RDF je patrné, že molekula adsorbovala častěji na výše položených skupinách atomů povrchu představující residua OU a HU, než na atomech OD a HD. Došlo však i k silnější interakci atomů H6 s OD a O2 se skupinou HD. Atom O1, který je dvojně vázaný na karboxylový uhlík C, je aktivní v tvorbě vazeb s vodíky povrchu (HU), zatímco atom O2, na který je kovalentně navázaný vodík H6, tvoří další vazby s vodíky povrchu podstatně hůře. Nejdéle adsorbovala molekula na 2 – 2,7 ns. Karboxylová skupina COOH je při bezprostřední adsorpci natočená k povrchu a celá molekula tak zaujímá charakteristickou orientaci patrnou z Obr. 32.



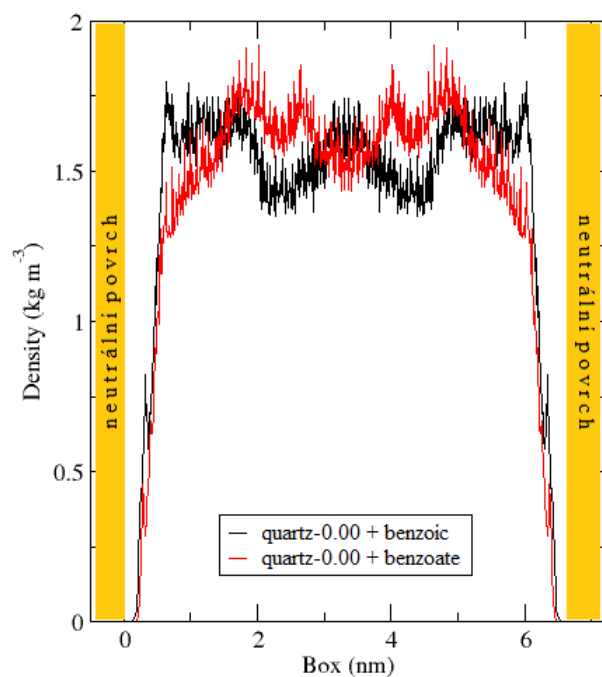
Obr. 33 a 34: Vodíkové vazby molekuly kyseliny benzoové s neutrálním křemenným povrchem při adsorpci

5.3. Shrnutí výsledků simulovaných systémů

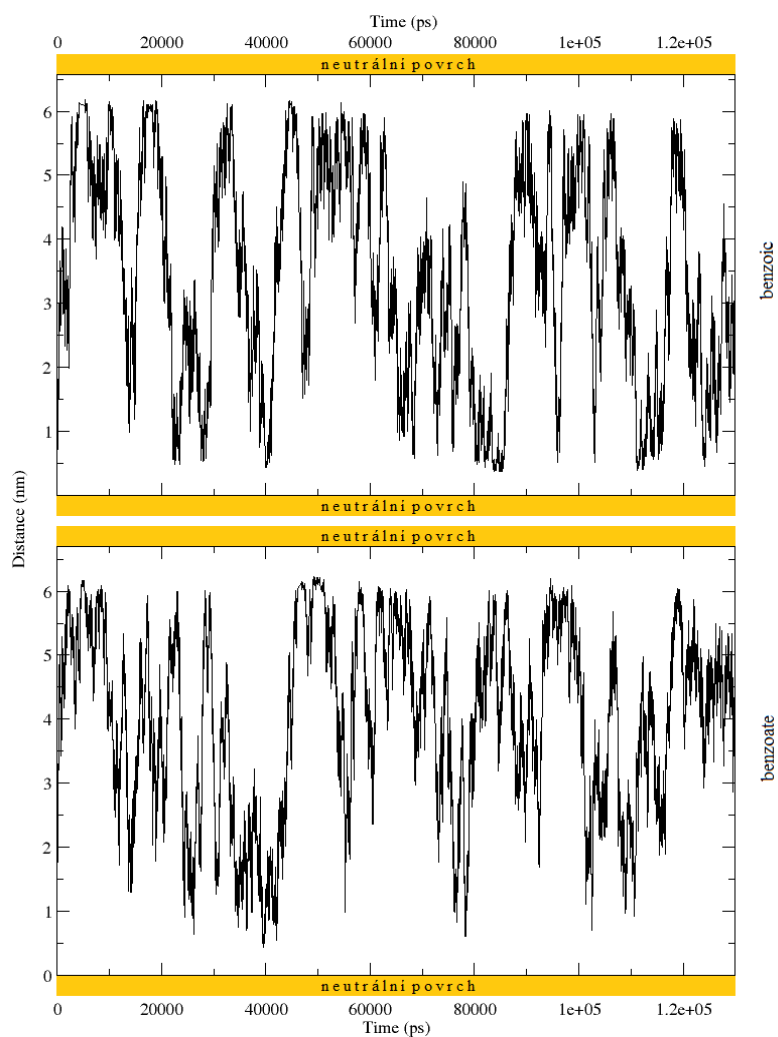
5.3.1. Kyselina benzoová a benzoát



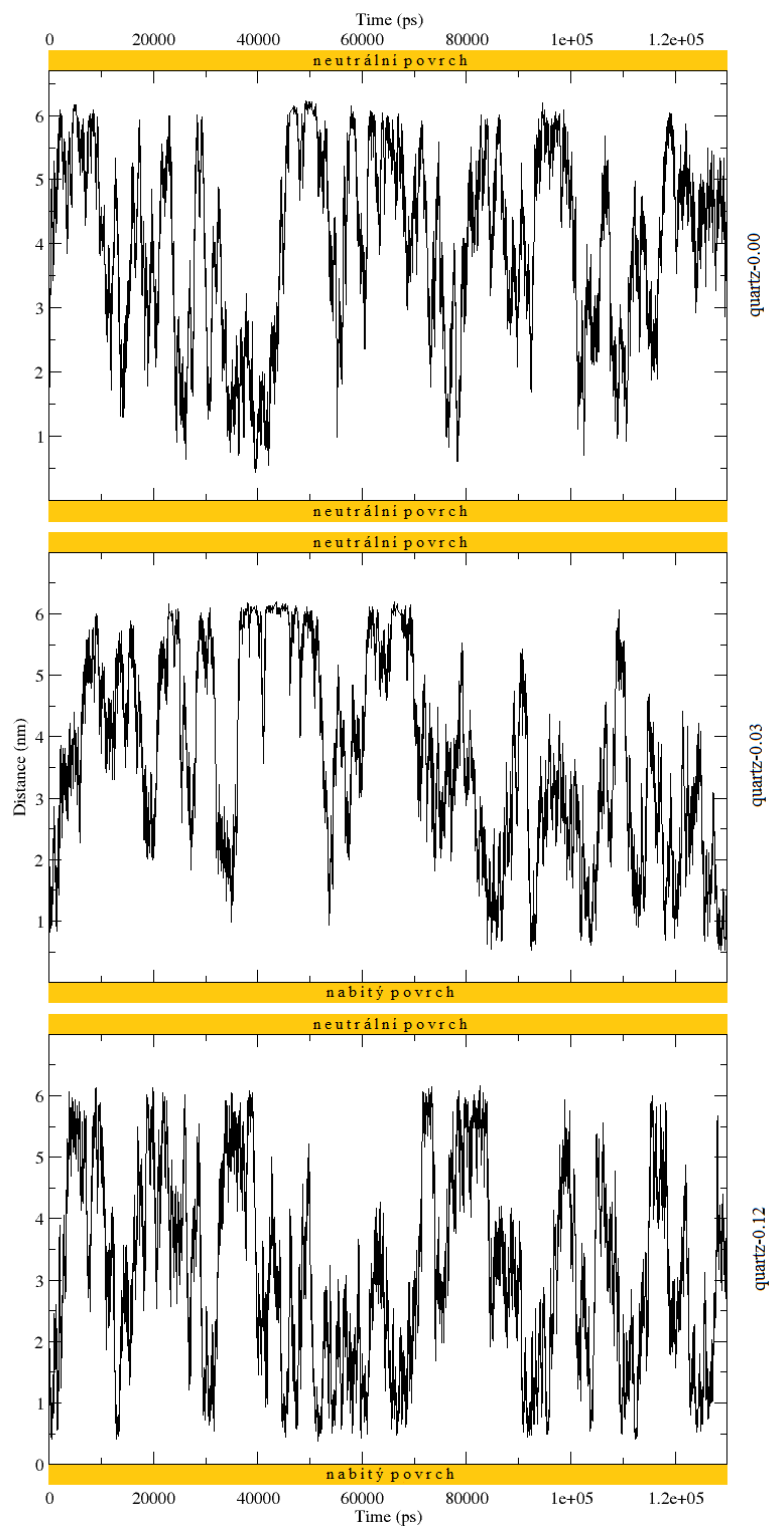
Obr. 35: Srovnání hustotních profilů molekuly benzoátu u různě nabitých povrchů



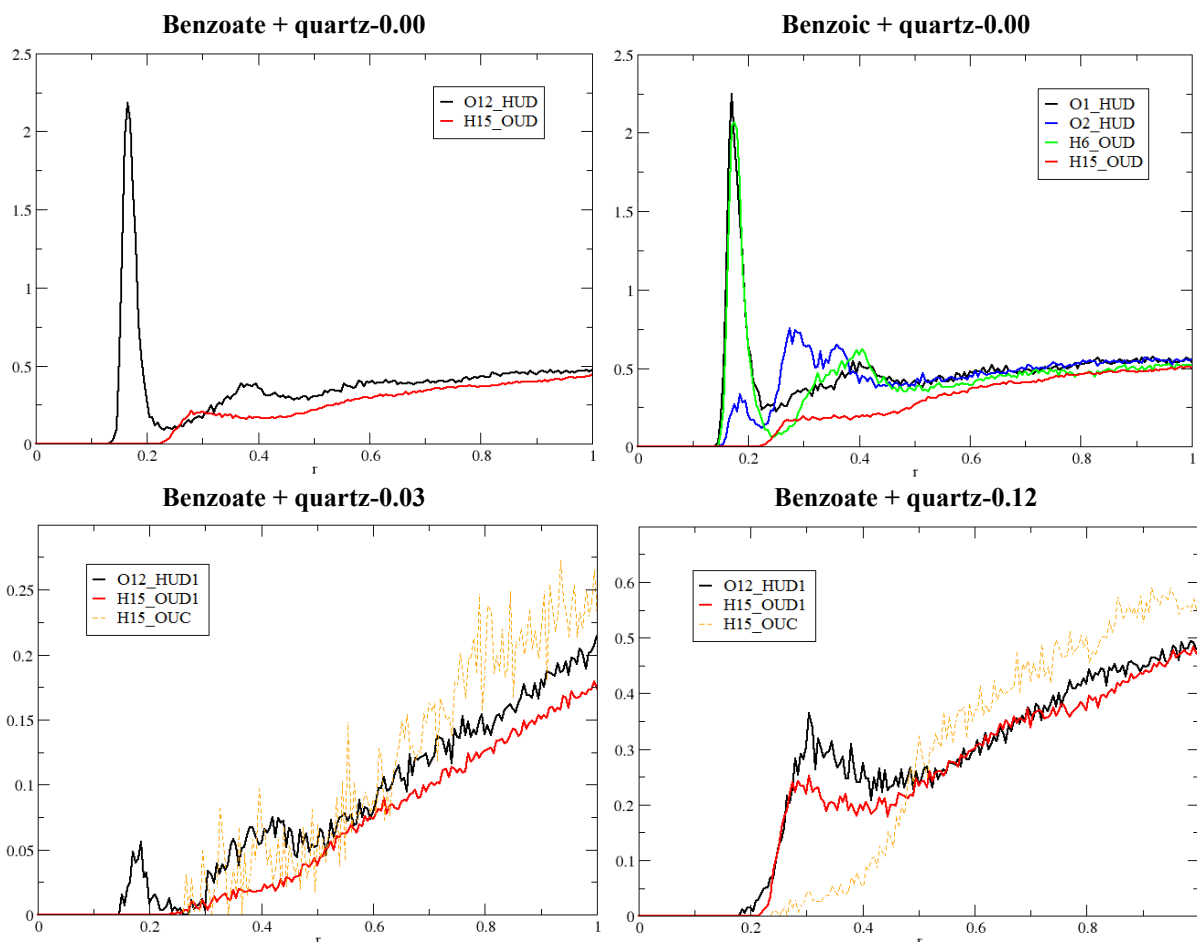
Obr. 36: Srovnání hustotních profilů kyseliny benzoové a benzoátu na neutrálním povrchu



Obr. 37: Srovnání časového vývoje vzdálenosti kyseliny benzoové a benzoátu od neutrálního povrchu

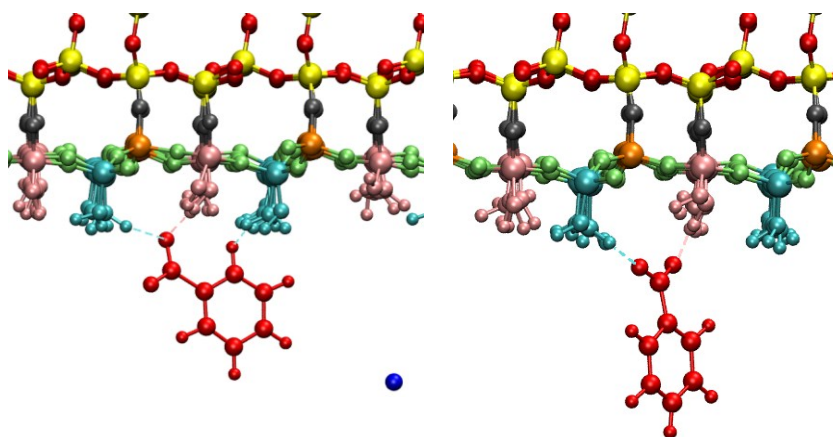


Obr. 38: Časový vývoj vzdálenosti molekuly benzoátu od povrchu pro různé hustoty povrchového náboje



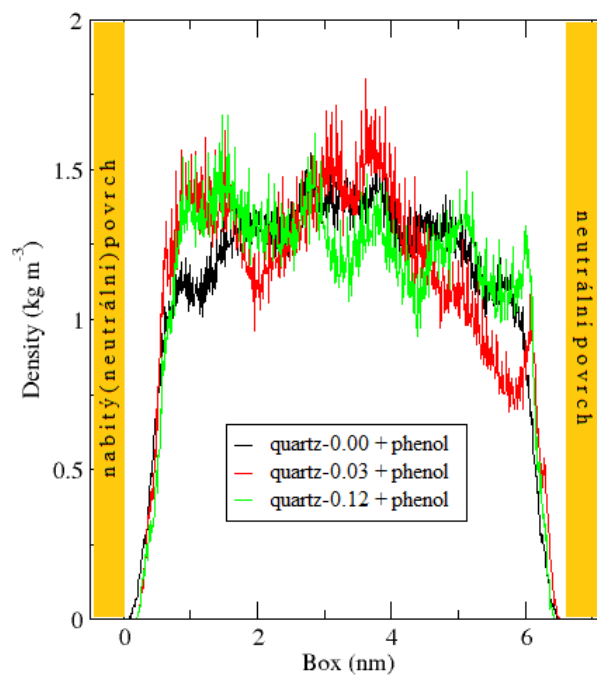
Obr. 39: RDF molekuly benzoátu pro různě nabitě povrchy, srovnání RDF benzoátu a kyseliny benzoové u neutrálního křemenného povrchu

Molekula benzoátu interagovala se všemi povrchy především COO^- skupinou, která byla v době adsorpce natočena směrem k povrchu. V případě interakce benzoátu s nabitými povrchy docházelo k tvorbě vodíkových vazeb v mnohem menší míře než s povrchem neutrálním (Obr. 39). S neutrálním povrchem vytvářel benzoát srovnatelné množství vodíkových vazeb jako kyselina benzoová u téhož povrchu (Obr. 39).

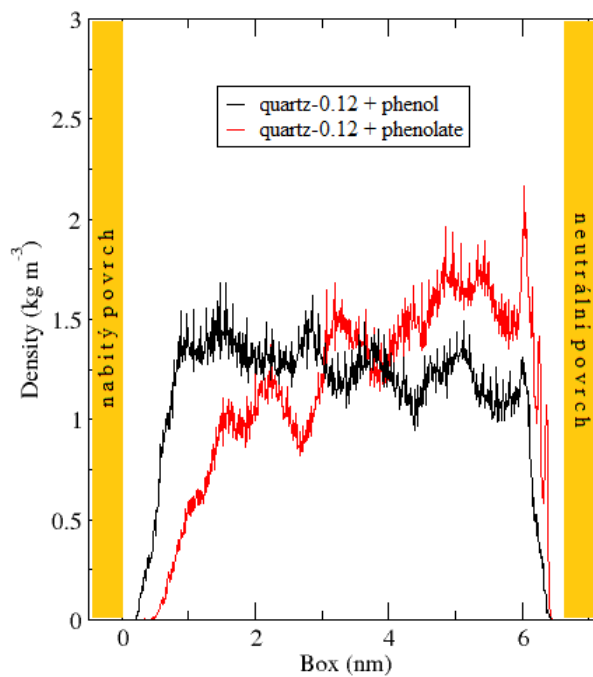


Obr. 40 a 41: Vodíkové vazby molekuly benzoátu s neutrálním křemenným povrchem při adsorpci

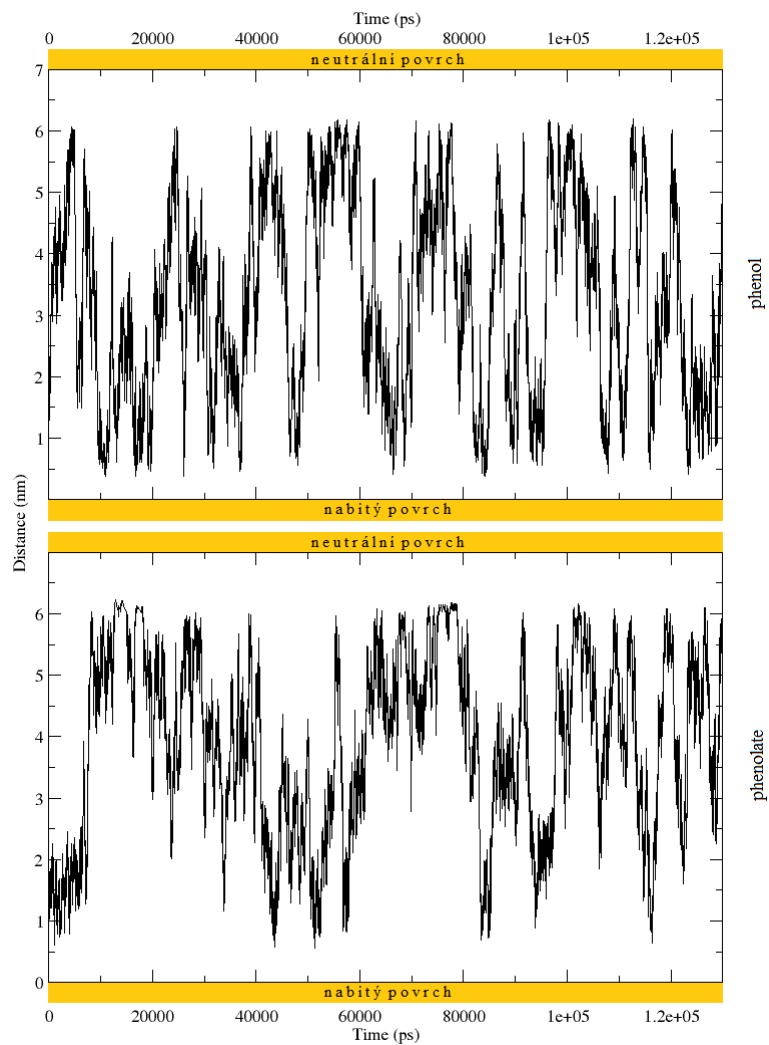
5.3.2. Fenol a fenolát



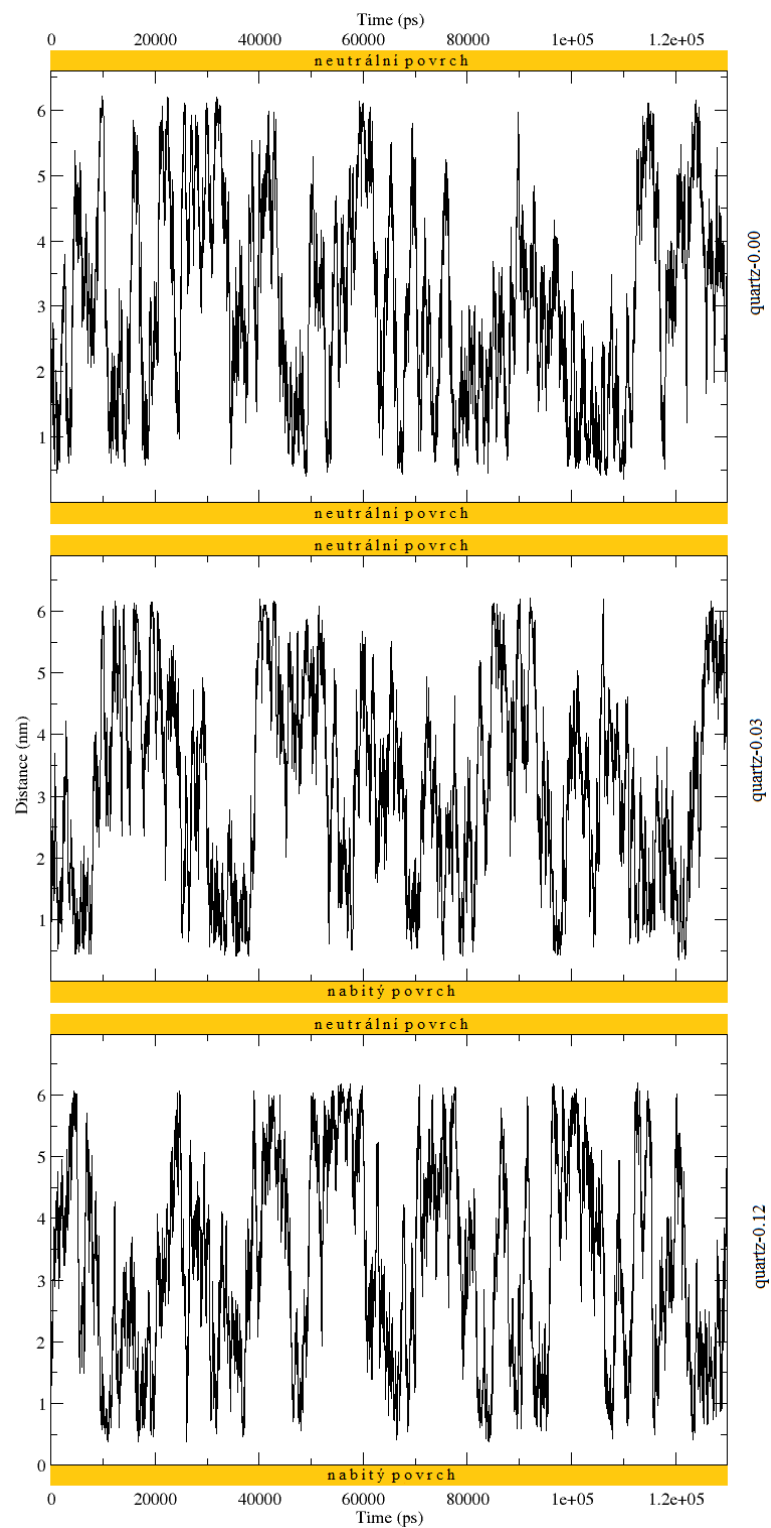
Obr. 42: Srovnání hustotních profilů molekuly fenolu u různě nabitých povrchů



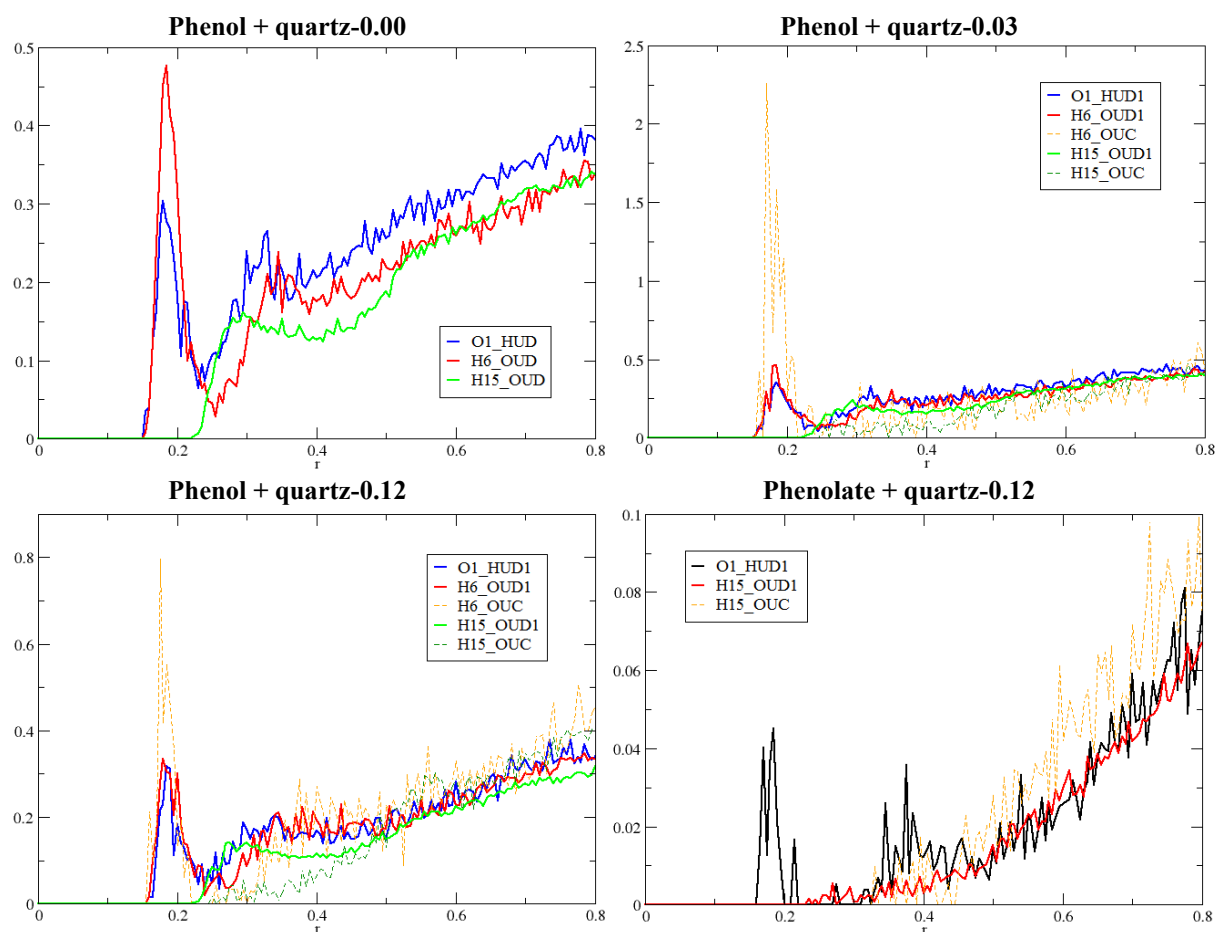
Obr. 43: Srovnání hustotních profilů molekuly fenolu a fenolátu na povrchu o nábojové hustotě $-0,12 \text{ C} \cdot \text{m}^{-2}$



Obr. 44: Srovnání časového vývoje vzdáleností molekul fenolu a fenolátu od povrchu o nábojové hustotě $-0,12 \text{ C} \cdot \text{m}^{-2}$

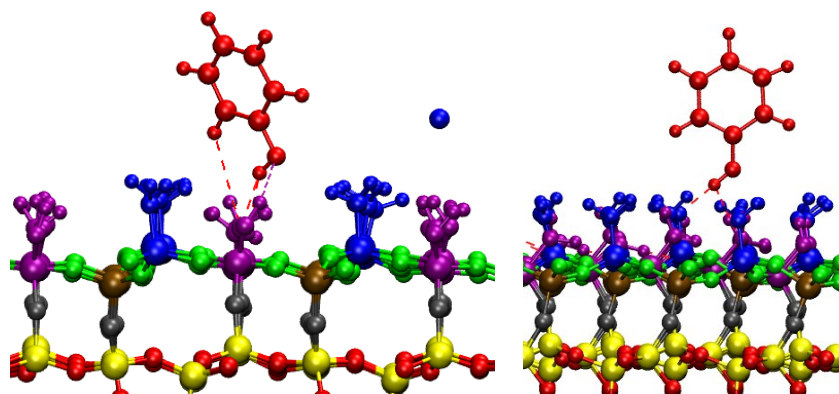


Obr. 45: Časový vývoj vzdálenosti molekuly fenolu od povrchu pro různé hustoty povrchového náboje

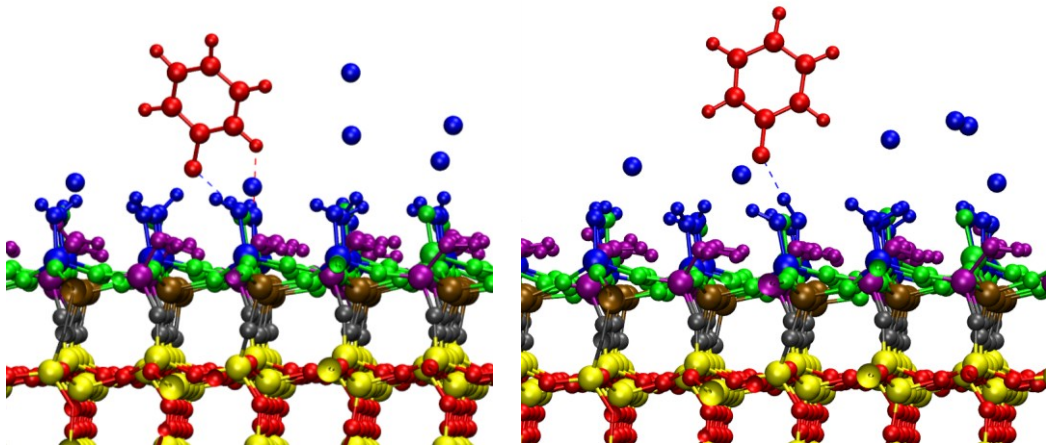


Obr. 46: RDF molekuly fenolu pro různě nabitě povrchy a srovnání RDF fenolu a fenolátu u povrchu s hustotou náboje $-0,12 \text{ C}\cdot\text{m}^{-2}$

Z grafů RDF je patrné, že fenol interagoval se všemi povrchy především prostřednictvím atomu H6, vždy však jen na velmi krátký okamžik (Obr. 45). Z grafu hustotních profilů (Obr. 42) vyplývá, že oproti neutrálnímu povrchu strávila molekula fenolu více času u povrchů nabitých, se kterými však tvořila méně vodíkových vazeb (Obr. 46). Dle Obr. 46 interagoval fenol s povrchem o nábojové hustotě $-0,12 \text{ C}\cdot\text{m}^{-2}$ silněji než fenolát, který byl dle uvedených hustotních profilů přitahován k neutrálnímu povrchu (Obr. 43).

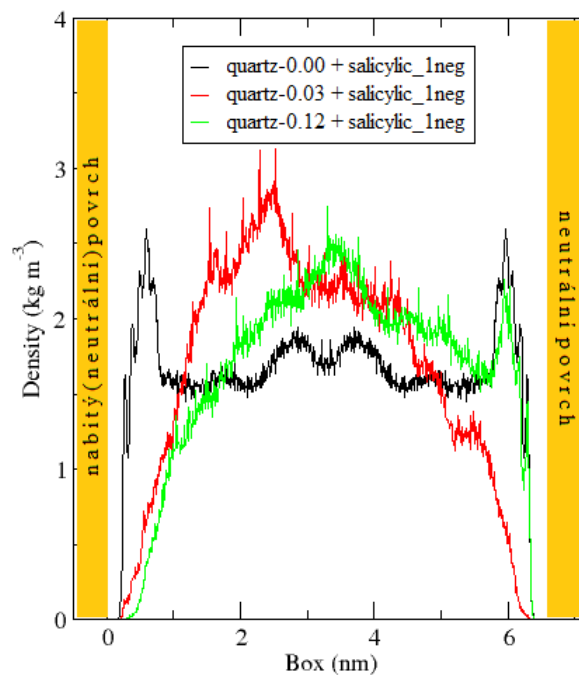


Obr. 47 a 48: Vodíkové vazby molekuly fenolu s neutrálním křemenným povrchem při adsorpci

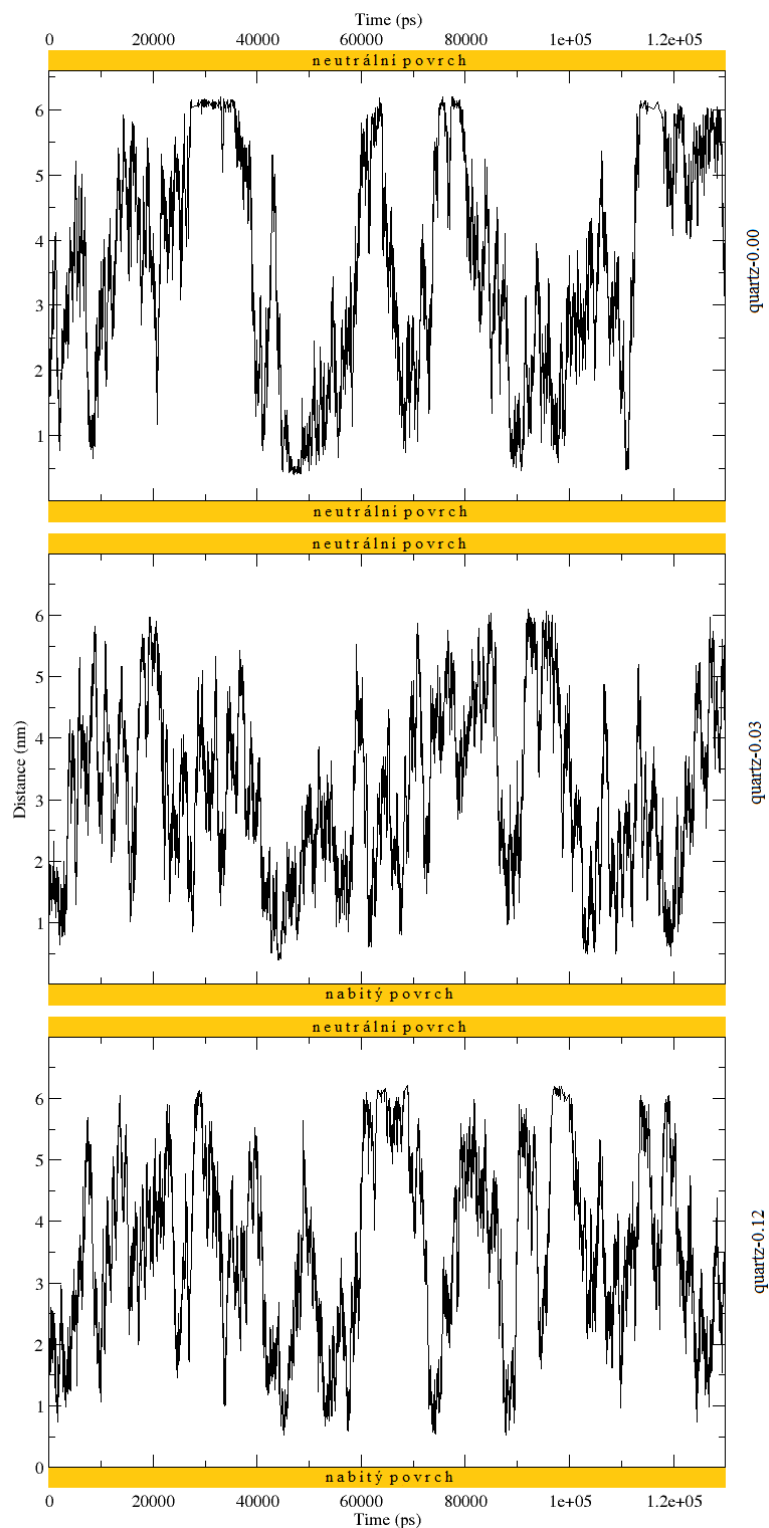


Obr. 49 a 50: Vodíkové vazby molekuly fenolátu s křemenným povrchem o povrchové hustotě náboje $-0,12 \text{ C} \cdot \text{m}^{-2}$ při adsorpci

5.3.3. Kyselina salicylová



Obr. 51: Srovnání hustotních profilů jedné deprotonované kyseliny salicylové u různě nabitých povrchů

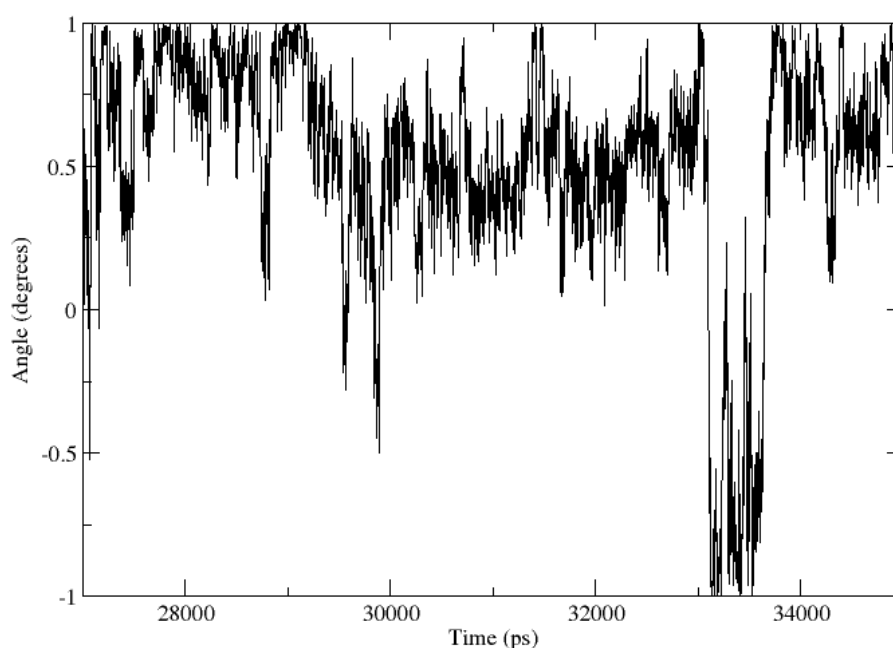


Obr. 52: Časový vývoj vzdálenosti jednou deprotonované kyseliny salicylové od povrchu pro různé hustoty povrchového náboje

Povrch	Vektor C1_C5	
	cos	orientace
Horní (SU1)	+	COO ⁻ od povrchu
	-	COO ⁻ k povrchu
Dolní (SU2)	+	COO ⁻ k povrchu
	-	COO ⁻ od povrchu

Tab. 11: Orientace vektoru C1_C5 jednou deprotonované kyseliny salicylové vzhledem k povrchu v závislosti na cosinu úhlu

Angle between C1_C5 and Z-axis

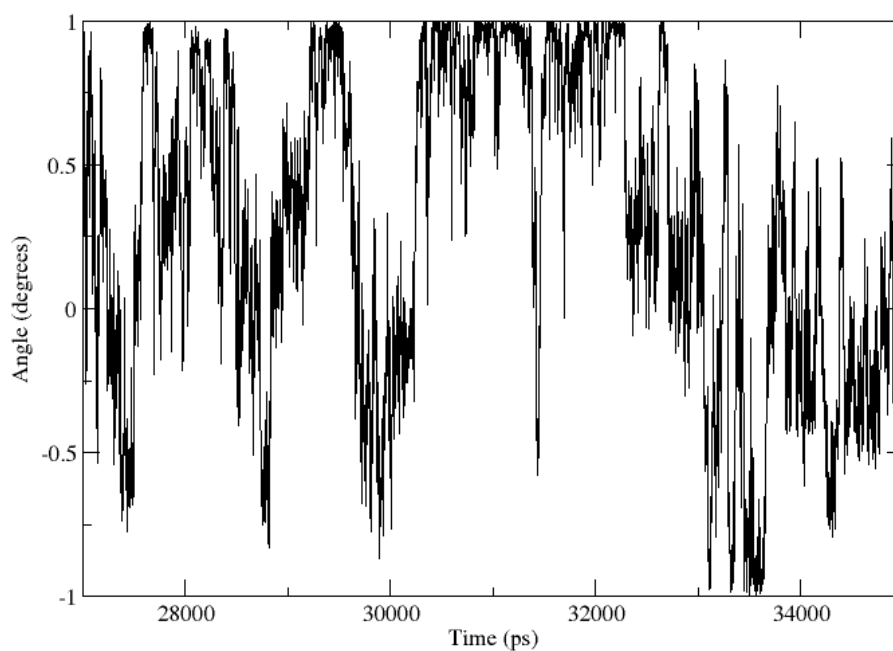


Obr. 53: Vzájemné natočení vektoru C1_C5 a osy z. V době adsorpce u spodního povrchu (SU2), značí kladný cosinus natočení COO⁻ skupiny směrem k povrchu.

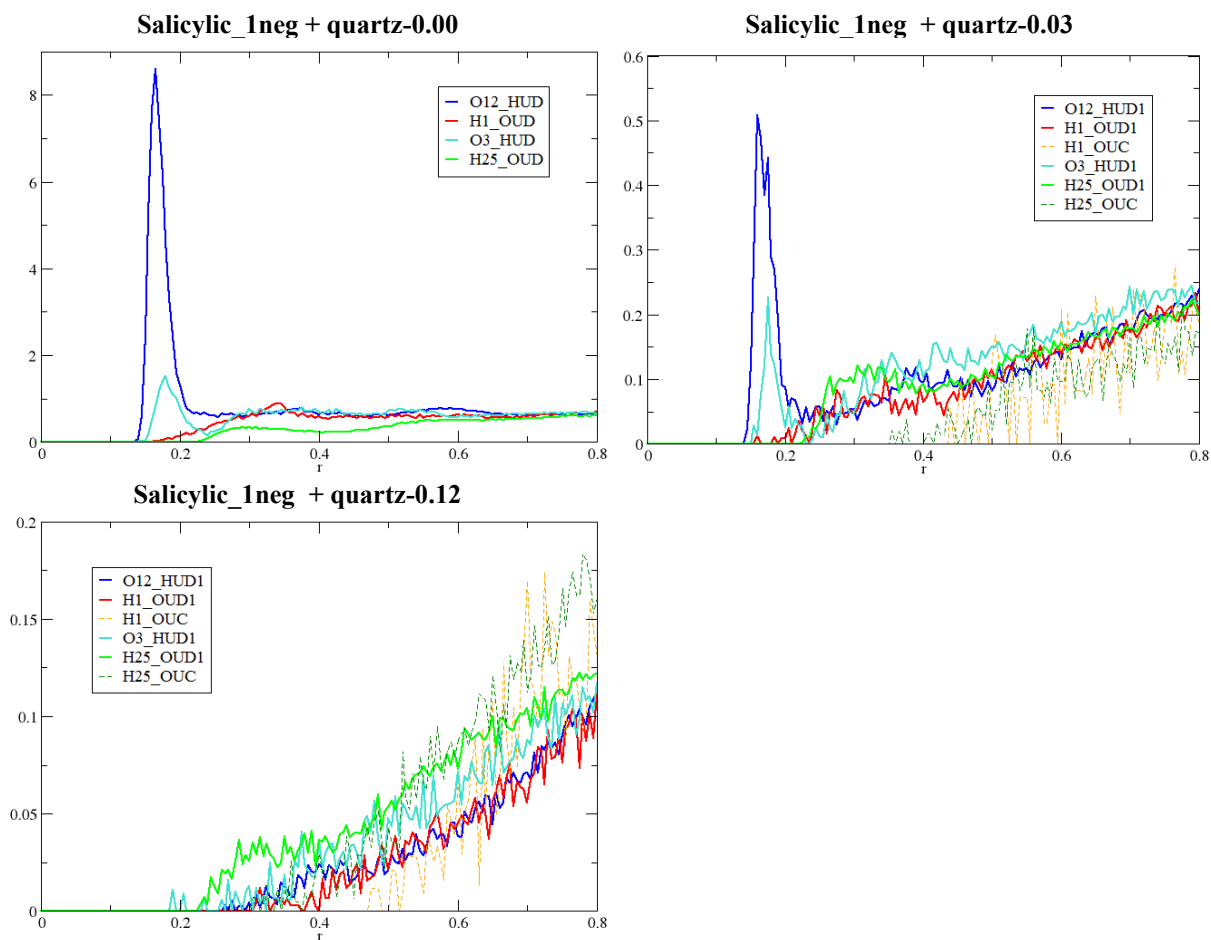
Povrch	Vektor C3_C6	
	cos	orientace
Horní (SU1)	+	OH od povrchu
	-	OH k povrchu
Dolní (SU2)	+	OH k povrchu
	-	OH od povrchu

Tab. 12: Orientace vektoru C3_C6 jednou deprotonované kyseliny salicylové vzhledem k povrchu v závislosti na cosinu úhlu

Angle between C3_C6 and Z-axis

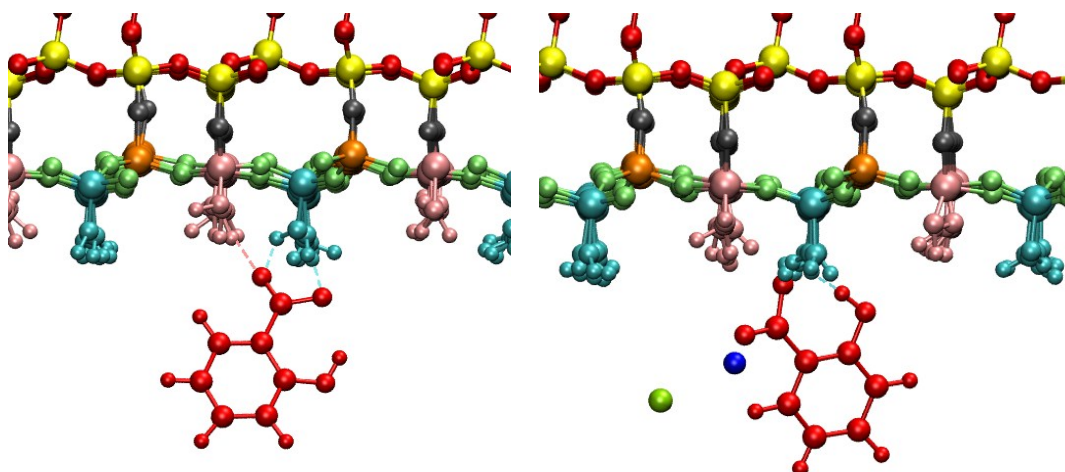


Obr. 54: Vzájemné natočení vektoru C3_C6 a osy z. V době adsorpce u spodního povrchu (SU2), značí záporný cosinus natočení OH skupiny směrem k povrchu.



Obr. 55: RDF jednou deprotonované kyseliny salicylové pro různě nabitě povrchy

V případě kombinace deprotonované kyseliny salicylové a neutrálního povrchu došlo k velmi výrazné a téměř nepřetržité adsorpci, která trvala přes 8 ns (Obr. 52). Molekula interagovala především COO⁻ skupinou (Obr. 55), při silnější interakci docházelo zároveň k natočení OH skupiny směrem k povrchu. Dle Obr. 53 a Obr. 54 je však hlavní skupinou podílející se na vazbách s povrchem právě deprotonovaná karboxylová skupina. Se záporně nabitými povrchy tvořila deprotonovaná kyselina salicylová vodíkové vazby podstatně obtížněji (Obr. 51). Čím byla povrchová hustota náboje povrchu větší, tím byl jev patrnější. U povrchu o nábojové hustotě $-0,12 \text{ C}\cdot\text{m}^{-2}$ docházelo při přiblížení molekuly k povrchu k jejímu odpuzování, které bylo viditelné zejména při pohledu na systém v programu VMD. Povrchová residua SUC nesoucí záporný náboj jsou rozmístěna natolik blízko sebe, že molekula nacházela velmi těžko výhodnou adsorpční polohu.



Obr. 56 a 57: Vodíkové vazby molekuly jednou deprotonované kyseliny salicylové s neutrálním křemenným povrchem při adsorpci

5.3.4. Interakční energie molekul s povrchy

Tab. 14 ukazuje hodnoty interakčních energií všech simulovaných systémů v rámci celé jejich simulační doby (130 ns) i pro vybrané interakční časy, které jsou vypsány v Tab. 13. Vazebná energie mezi molekulou a povrchem vyjádřená pomocí energie vdW (Lennard-Jonesův potenciál) a elektrostatických (Coulombův potenciál) interakcí vypovídá o povaze a síle vazeb mezi molekulou a atomy povrchu a míře přiblížení molekuly k povrchu. Z Tab. 14 je patrné, že hodnoty interakčních energií konkrétních interakčních časů korelují s uvedenými grafy radiálních distribučních funkcí jednotlivých molekul. U jednou deprotonované kyseliny salicylové je z grafů RDF (Obr. 55) v souladu s hodnotami interakčních energií vidět negativní vliv záporně nabitého povrchu na tvorbu vodíkových

vazeb mezi molekulou a atomy povrchu. Podobně je tomu u neutrální molekuly fenolu (Obr. 46). Adsorpce záporně nabitého benzoátu je dle grafů RDF i hodnot adsorpčních energií mírně zeslabena se vzrůstajícím záporným nábojem povrchu s tím, že odpudivý vliv nabitého povrchu je patrnější u křemene o povrchové hustotě náboje $-0,03 \text{ C} \cdot \text{m}^{-2}$.

Povrch	Molekula	Začátek v [ps]	Konec v [ps]
0,00	benzoic	4200	5600
	benzoate	4700	5400
	phenol	22300	22550
	salicylic_1neg	113000	117000
-0,03	benzoate	128200	129000
	phenol	96500	98000
	salicylic_1neg	44000	44600
-0,12	benzoate	12700	13400
	phenol	84100	84200
	phenolate	51250	51300
	salicylic_1neg	44500	45500

Tab. 13: Vybrané interakční časy simulovaných systémů použité k analýze adsorpčních energií

Interakční energie [kJ/mol]		Celá simulace		Vybraný interakční okamžik	
		E_C	E_{LJ}	E_C	E_{LJ}
benzoic	0,00	-3,42	-1,14	-57,01	-9,10
	-0,03	-	-	-	-
	-0,12	-	-	-	-
benzoate	0,00	-3,70	-0,68	-95,80	1,31
	-0,03	-8,42	-0,43	-0,03	-3,48
	-0,12	-0,67	-1,11	-9,07	-11,01
phenol	0,00	-0,79	-0,73	-27,03	-9,72
	-0,03	-0,98	-0,87	-8,64	-6,08
	-0,12	-0,88	-0,77	-3,79	-2,81
phenolate	0,00	-	-	-	-
	-0,03	-	-	-	-
	-0,12	-6,43	-0,03	-9,50	-3,72
salicylic_1neg	0,00	-2,56	-0,27	-103,25	6,21
	-0,03	-0,47	-0,40	-94,06	-3,73
	-0,12	-5,47	-0,45	1,93	-3,37

Tab. 14: Adsorpční energie simulovaných systémů v rámci celé simulační doby a pro vybrané interakční časy z Tab. 13

6. Závěr

Hlavním úkolem mé práce bylo s využitím programu GROMACS připravit a analyzovat několik systémů složených vždy z jedné organické molekuly a křemenného povrchu o různé hustotě povrchového náboje, provést simulace chování těchto systémů a následně zpracovat výsledky simulací.

Doposud byly prováděny simulace pouze na neutrálním křemenném povrchu. [44], [45]. Autoři projektu uvedeného v úvodu této práce, zaměřeného na zkoumání interakcí huminových kyselin a organické hmoty s minerálními povrchy, vycházeli z předpokladu, že záporný náboj křemenného povrchu může ovlivnit jednak orientaci molekul vody, adsorpci iontů, ale i chování biomolekul u těchto povrchů. Předmětem této práce, která vznikala v rámci jmenovaného projektu, je analýza a zhodnocení interakcí neutrálních organických molekul (kyseliny benzoové, fenolu a kyseliny salicylové) a jejich záporně nabitých forem s křemenným povrchem o různé (záporné) hustotě povrchového náboje. Výstupem je zhodnocení vlivu změn povrchového náboje na chování biomolekul. Díky provedeným simulacím tak lze porovnat vliv nabitých povrchů na interakce s molekulami oproti povrchu neutrálnímu a posoudit jeho tendenci se změnou velikosti povrchového náboje.

Na základě analýzy zpracovaných simulací lze konstatovat, že míra adsorpce a tvorba vazeb s povrchem je u deprotonované kyseliny salicylové, nesoucí záporný náboj, viditelně negativně ovlivněna rostoucím záporným nábojem povrchu. Čím je povrch nabitější, tím více je deprotonovaná kyselina salicylová odpuzována. Podobnou tendenci vykazuje i neutrální molekula fenolu. Naopak jeho deprotonovaná forma fenolát je k záporného povrchu přitahována díky vazbě na nenabitá residua povrchu. Adsorpce záporně nabitého benzoátu je mírně zeslabena se vzrůstajícím záporným nábojem povrchu, zatímco kyselina benzoová simulovaná pouze v kombinaci s neutrálním povrchem tvoří srovnatelné množství vazeb s atomy nenabitého povrchu jako benzoát.

Získané výsledky provedených simulací rozšiřují doposud zpracované téma simulací na neutrálních površích o oblast interakcí na záporně nabitém povrchu. Limity rozsahu bakalářské práce neumožnily provést detailnější analýzu včetně dalších simulací (např. dalších organických molekul), které by si na základě tendencí plynoucích z doposud získaných výsledků zadané téma jistě zasloužilo. Část výsledků získaných při řešení této bakalářské práce byla prezentována na konferenci [46].

7. Seznam použité literatury

- [1] Nezbeda, I.; Kolafa, J.; Kotrla, M.: Úvod do počítačových simulací, Metody Monte Carlo a molekulární dynamiky, Karolinum 1998
- [2] Šilhavá, K.: Software pro molekulární dynamiku, bakalářská práce, ZSF JU, 2011
- [3] Krohová, L.: Software pro zobrazování molekulárních struktur, bakalářská práce, ZSF JU, 2011
- [4] Frenkel, D.; Smit, B.: Understanding molecular simulation, Academic Press 2002
- [5] Matěnová, M.: Studium interakce membránových proteinů na molekulární úrovni pomocí silové spektroskopie, optické spektroskopie a metod výpočetní biochemie, diplomová práce, PřF JU, 2011

7.1. Internetové a jiné zdroje

- [6] http://www.chemicke-listy.cz/docs/full/2000_01_05.pdf [2012-02]
- [7] http://en.wikipedia.org/wiki/Molecular_dynamics [2012-01]
- [8] http://www.ncbr.muni.cz/~n19n/vyuka/pocitacova_chemie/ [2012-02]
- [9] http://www.cdm.cas.cz/publications/hora/ph_habl_sld.pp4.pdf [2012-02]
- [10] http://faf.vfu.cz/export/struktura-fakulty/sekce_ustavy/ustav_chemickych_leciv/vyuka/molekularni-zaklady-vyvoje-leciv/Molekulove-modelovani_2011.pdf [2012-02]
- [11] <http://www.xray.cz/ms/bul2000-1/huml.pdf> [2012-02]
- [12] <http://marge.uochb.cas.cz/~jungwirt/kkmd.pdf> [2012-02]
- [13] <http://www.gromacs.org/Documentation/Manual> [2012-02]
- [14] <http://en.wikipedia.org/wiki/AMBER> [2012-02]
- [15] <http://en.wikipedia.org/wiki/CHARMM> [2012-02]
- [16] <http://en.wikipedia.org/wiki/GROMACS> [2012-02]
- [17] <http://en.wikipedia.org/wiki/GROMOS> [2012-02]
- [18] <http://en.wikipedia.org/wiki/OPLS> [2012-02]
- [19] http://en.wikipedia.org/wiki/Covalent_bond [2012-02]
- [20] http://fch.upol.cz/skripta/chs/qm_V.pdf [2012-02]
- [21] <http://cheminfo.chemi.muni.cz/ianua/ZFCh/molekuly/vdW.htm> [2012-02]
- [22] <http://polymer.bu.edu/Wasser/robert/work/> [2012-02]
- [23] http://en.wikipedia.org/wiki/Van_der_Waals_force [2012-02]
- [24] http://physics.ujep.cz/~mmaly/vyuka/poc_fyz_1/pocitacova_fyzika_1.pdf [2012-01]
- [25] <http://isaacs.sourceforge.net/phys/psc.html> [2012-01]
- [26] http://is.muni.cz/do/rect/el/estud/prif/js11/fyz_chem/web/molekuly/PES.htm [2012-02]

- [27] <http://nanoparticlesandquantumdots.blogspot.com/> [2012-02]
- [28] http://www.gromacs.org/index.php?title=Download_%26_Installation/Related_Software/MKTOP [2012-03]
- [29] <http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/Files/minim.mdp> [2011-12]
- [30] <http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/Files/nvt.mdp> [2011-12]
- [31] <http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/Files/npt.mdp> [2011-12]
- [32] <http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/Files/md.mdp> [2011-12]
- [33] http://people.sissa.it/~raugei/lecture_notes/vmd.pdf [2011-12]
- [34] <http://plasma-gate.weizmann.ac.il/Grace/> [2011-12]
- [35] <http://mintaka.sdsu.edu/reu/grace.printout.html> [2011-12]
- [36] <http://en.wikipedia.org/wiki/File:Benzoic-acid-3D-balls.png> [2012-01]
- [37] <http://en.wikipedia.org/wiki/File:Phenol-3D-balls.png> [2012-01]
- [38] <http://en.wikipedia.org/wiki/File:Salicylic-acid-from-xtal-2006-3D-balls.png> [2012-01]
- [39] <http://en.wikipedia.org/wiki/File:Morse-potential.png> [2012-02]
- [40] <http://www.chem.purdue.edu/gchelp/liquids/disperse.html> [2012-02]
- [41] <http://www.gromacs.org> [2012-02]
- [42] <http://www.studiumchemie.cz/NMR/skripta.pdf> [2012-02]
- [43] Ribeiro, A. A. S. T.; Horta, B. A. C.; de Alencastro, R. B. J. *Braz. Chem. Soc.*, Vol. 19, No. 7, 1433-1435, 2008
- [44] Lopes, P. E. M.; Murashov, V.; Tazi, M.; Demchuk, E.; MacKerell, A. D.: Development of an Empirical Force Field for Silica. Application to the Quartz–Water Interface *J. Phys. Chem. B* 110 , 2782-2792 2006
- [45] Skelton, A. A.; Fenter, P.; Kubicki, J. D.; Wesolowski, D. J.; Cummings, P. T.: Simulations of the Quartz(1011)/Water Interface: A Comparison of Classical Force Fields, Ab Initio Molecular Dynamics, and X-ray Reflectivity Experiments *J. Phys. Chem. C* 115, 2076 2011
- [46] X Discussions in Structural Molecular Biology, Nové Hrady, March 22-24, 2012, poster O. Kroutil, Z. Chval, H. Barvíková and M. Předota: "Behavior of Water, Ions and Small Organic Molecules Near Quartz (101) Surfaces"

7.2. Další užitečné odkazy

7.2.1. Stránky počítačových klastrů UFY a Hermes

<http://www.metacentrum.cz/cs/resources/machines/hermes.html>

<http://cluster.prf.jcu.cz/>

<http://ufy.prf.jcu.cz/cluster/index.html>

8. Přílohy

Příložené CD obsahuje:

1. Topologie a konfigurace křemenných povrchů a studovaných organických molekul, použítá silová pole
2. Datové soubory provedených simulací (energetické soubory a trajektorie o velikostech několika GB nejsou přiloženy, ale uloženy na klastru UFY)
3. Zdrojové kódy doplňkových utilit `g_distmp`, `g_shift`, `g_sym` a jejich spustitelné binární soubory pro Linux (x64), skript `gsub`
4. Ukázky videí zobrazujících vývoj systémů