

# 1. Cíle práce

Předkládaná diplomová práce je postavena na zvládnutí metody parametrizace silových polí a implementaci těchto parametrů do programu Gromacs. Získání parametrů je nezbytné pro následné molekulárně-dynamické výpočty. Ve své práci jsem se zaměřila na dva systémy: fotosyntetickou membránu jako zástupce „velkých“ molekul a na tryptofanový represor, jehož ligandy lze zařadit mezi malé molekuly. V prvním případě fotosyntetické membrány šlo konstrukci speciálního modelu s následným porovnáním vhodnosti různých silových polí. Dále jsem se zabývala i vytvoření modelu jádra fotosystému II z hrachu (*Pisum sativum*), který bude použit pro studium separace nábojů na pigmentech v reakčním centru fotosystému II.

# 2. Úvod

## 2.1 Biologické membrány

Biologická membrána je stavební a funkční prvek všech buněk a velkého počtu organel, který se sestává především z komplexu lipidové dvojvrstvy a v ní zanořených proteinů (tzv. membránové proteiny). Membrána odděluje buňku od vnějšího prostoru, slouží k výměně hmoty a informací mezi buňkou a okolím a v neposlední řadě může na membránách vznikat energie. Vzhledem k charakteru práce se v následujícím textu zaměřím pouze na lipidovou část membrány.

Biologické membrány jsou tvořeny z lipidů především proto, že lipidy jsou nerozpustné ve vodě, ale dají se rozpouštět v organických rozpouštědlech (např. aceton, ethanol, chloroform...), protože hlavní kapalinou v živých systémech je voda, plní membrány funkci oddělovačů buňky a okolí a také jednotlivých kompartmentů v buňce.

Podle chemického složení můžeme lipidy v membránách rozdělit na tři skupiny: na fosfolipidy, glykolipidy a cholesterol (Brown, 1996).

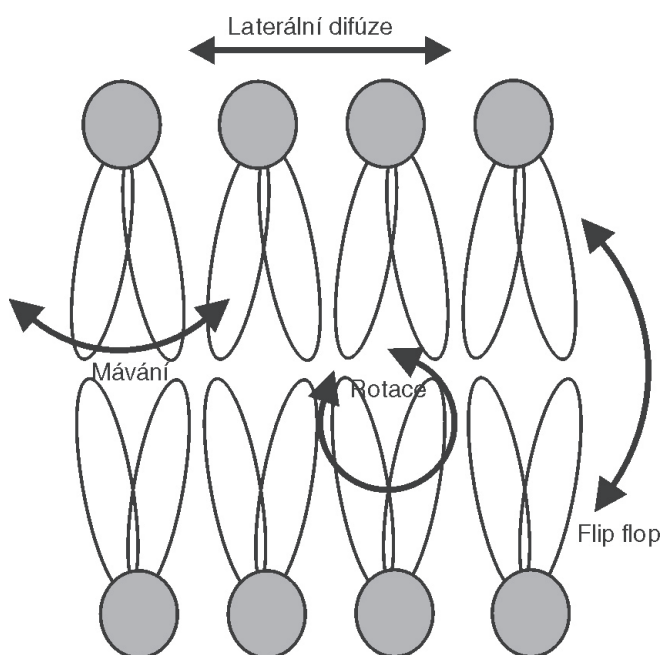
Hlavní součástí biologických membrán jsou fosfolipidy, tj. tuky, u nichž byla jedna mastná kyselina nahrazená fosfátovou skupinou a dusíkatou molekulou. Glycerol je nejčastějším terciárním alkoholem ve fosfolipidech, mastné kyseliny mají 14 až 24 uhlíků. Jejich jeden řetězec je obvykle nenasycený s jednou až čtyřmi cis dvojnými vazbami. Jako dusíkaté látky v nich nejčastěji najdeme cholin, serin nebo ethanolamin. Kromě glycerolu se ve fosfolipidech poměrně často vyskytuje sfingosin.

Další důležitou skupinou lipidů v biologických membránách jsou glykolipidy, tj. lipidy s navázaným cukrem. I zde se jako terciární alkohol nachází u živočichů především sfingosin a u rostlin a bakterií glycerol.

Poslední skupina lipidů má pouze jednoho zástupce – cholesterol. To je lipid obsahující tetracyklický steroid a krátký uhlíkatý řetězec s navázanou hydroxy skupinou. Najdeme ho hlavně v membránách savců a mykoplazem.

Membránové lipidy jsou amfipatické. To znamená, že mají hydrofilní (polární) a hydrofobní (nepolární) část. Aby se omezil kontakt hydrofobní části s vodou tvoří membránové lipidy v polárních rozpouštědlech dvojvrstvy, kde jsou orientované hydrofobními částmi k sobě. V případě, že se dvojvrstva působením vody uzavře do kulovitého útvaru mluvíme o liposomech. Je-li liposom tvořen pouze monovrstvou, nazývá se micela.

Biologické membrány nejsou statické útvary, ale jednotlivé lipidy (stejně jako membránové proteiny) se pohybují. Zde se zaměřím pouze na pohyb molekul lipidů. Ty rotují kolem své nejdelší osy, přičemž „mávají“ mastnými kyselinami do prostoru. Nesmíme zapomínat ani na tzv. laterální difuzi, kdy se lipidy pohybují posuvným pohybem v rámci monovrstvy. Lipidy jsou schopny i přecházet z jedné monovrstvy do druhé. Mluvíme o „flip-flop“ pohybu nebo o tranzverzní difuzi, která, protože hydrofilní část lipidu musí projít hydrofobní oblastí, je mnohem méně častá než předchozí druhy pohybu. Výše popsaný model biologické membrány se nazývá model tekuté mozaiky a byl předpovězen v roce 1972 Singerem a Nicolsonem (Singer, 1972). Pro větší názornost pohyby lipidů v membráně zachycuje obrázek 1 na další straně.



Obr. 1: Pohyby, které vykonávají lipidy v membráně

Důležitou vlastností biologických membrán je tekutost (fluidity), která souvisí s uspořádaností lipidů v rámci membrány. Čím je membrána uspořádanější, tím má vyšší tekutost. Tekutost je ovlivňuje především teplota a tlak, složení membrány a v živočišných membránách i obsah cholesterolu.

Nejméně uspořádanou formou dvojvrstvené membrány je kapalná fáze, kde řetězce mastných kyselin jsou v rámci dvojvrstvy neuspořádané. Tato fáze se nazývá tekutým krystalem. S klesající teplotou nebo rostoucím tlakem se dvojvrstva začíná uspořádat. Tímto způsobem vzniká gel. Pokud se vlivem klesající teploty a nebo rostoucího tlaku lipidy ještě více uspořádají vzniká krystalická fáze.

Kromě těchto víceméně běžných dvouvrstevných fází existují ještě lipidy v jednovrstvené fázi tekutých krystalů. Sem patří především hexagonální a kubická fáze. Pro mou práci je důležitější hexagonální fáze, protože se nachází ve fotosyntetické membráně. Hexagonální fázi tvoří dlouhé tyčovité útvary, ve kterých mastné kyseliny míří buď do prostoru nebo do středu lamely. Nakonec zdůrazním, že ne všechny membrány vykazují všechny fáze, naopak, membrány, které můžeme ve všech fázích najít jsou výjimečné.

## 2.2 Fotosyntéza a fotosyntetická membrána

Fotosyntetická membrána je nejvíce zastoupenou membránou v přírodě. Od většiny ostatních membrán se liší tím, že je tvořena především glykolipidy. Její složení se mění v závislosti na druhu organismu a podmínkách, ve kterých se tento nachází. Obecně lze říct, že asi 50% membrány tvoří monogalaktosyldiacylglycerol (MGDG), dalších asi 30% připadá na digalaktosyldiglycerol (DGDG), asi 10% na sulfoquinovosyldiacylglycerol (SQDG). Zbývajících asi 10% tvoří fosfolipidy, z nichž převažuje fosfatidylglycerol (PG), dále jsou zastoupeny fosfatidylcholin (PC), fosfatidylethanolamin (PE) a fosfatidylinositol (PI) (Lee, 2000; Dörmann, 2002). Při fyziologickém pH jsou neutrální pouze MGDG a DGDG, ostatní mají náboj -1.

I když je MGDG nejhojněji zastoupeným lipidem ve fotosyntetické membráně, ukazuje se, že netvoří typickou dvojvrstvu, ale že vytváří jednovrstvenou hexagonální fázi. Při pohledu na průřez „normálního“ membránového lipidu zjistíme, že tvoří obdelník. Průřez MGDG tvoří lichoběžník, ze kterého není možné vytvořit obdelníkovou membránu (chceme-li nemíchat hydrofilní a hydrofobní části), proto tyto lipidy tvoří válce – lamely. Protože je základna lichoběžníku cukerné části menší než základna lipidové části, shlukují se cukry (spolu s molekulami vody) na vnitřní části válce a hydrofobní lipidy směřují do vodného prostředí. Toto se nazývá hexagonální  $H_{II}$  fáze. V případě, že jsou hydrofobní části lipidů na okraji válce a hydrofilní na okraji, mluvíme o hexagonální  $H_I$  fázi. Ostatní lipidy ve fotosyntetické membráně mají obdelníkový průřez a proto vytváří dvojvrstvy.

Fotosyntetická membrána je zajímavá i z metodologického hlediska. Jedná se totiž o membránu složenou především z glykolipidů, přičemž k dnešnímu datu nebyl publikován žádný model glykolipidové membrány. Samotné téma modelování biologických membrán je v podstatě záležitostí posledních asi 10 let.

První snahy modelovat lipidové dvojvrstvy jsou z počátku 80. let minulého století. Ve většině případů se jednalo o zjednodušené lipidy a navíc se nebrala v úvahu přítomnost rozpouštědla. Modelování biologických membrán umožnil hlavně růst výpočetní kapacity počítačů a především pak vznik silových polí, které jsou schopny popsat pohyb makromolekul a současně vody. S časem a tedy s rostoucí silou počítačů se nejen zesložitovaly modelované systémy, ale také se přidávala voda později i další ionty (Tieleman, 1997). První „reálné“ systémy pochází až z konce 90. let minulého století. Jedná se o krátké ( $\leq 1$  ns) simulace obvykle mající 64 lipidů. Samozřejmě se jednalo o elektricky neutrální lipidy, membrány byly homogenní. Většina prvních membrán se nacházela ve fázi tekutých krystalů, které je v přírodě nejběžnější.

V současné době dosahují nejdelší simulace biologických membrán řádově stovek nanosekund a skládají se z obvykle 64 nebo 128 lipidů. Simulované membrány se obvykle skládají z jednoho lipidu, ikdyž se podařilo simulovat membrány složené ze dvou lipidů (u prvních prací byl obvykle jedním z lipidů smíšené membrány cholesterol) a také se podařilo simulovat nabitě membrány. Samozřejmě jsou v současné době simulace různých fází membrán, převažují simulace fáze tekutých krystalů (nachází se ve většině buněk) a gelové fáze (je pro ni nejvíce experimentálních dat).

Na některých simulacích z přelomu tisíciletí najdeme místo membrány oktan nebo dekan, především proto, že jsou také hydrofobní a jejich hustota přibližně odpovídá hustotě biologických membrán. Tyto aproximace byly vhodné především pro minimalizaci membránových proteinů, ale nedaly se použít např. ke zkoumání interakcí proteinů s membránou.

V dnešních dnech najdeme na internetu 12 rozdílných volně stažitelných membrán v různých fázích. Zajímavostí je, že naprostá většina modelů membrán nebo jejich skupenství byla tématem disertačních prací.

Glykolipidové membrány jsou zajímavé ještě z jiného hlediska. Sacharidy totiž díky svým silným hydrofobním vlastnostem vytváří mezi sebou i mezi přilehlými molekulami vody vodíkové vazby, čímž se mění vlastnosti membrány v porovnání s membránou fosfolipidovou. Jedná se především o posun fázových přechodů (nabalená voda zvyšuje hmotnost molekul) a vliv na tvorbu nelamelárních fází (Corti, 2007).

## 2.3 Tryptofanový represor

Tryptofanový represor je protein, který reguluje transkripci tryptofanu v bakteriích. Vyskytuje se jako homodimer, který má hmotnost 25 kDa. Je-li v buňce nedostatek tryptofanu, represor nasedne na DNA, čímž umožní syntézu aminokyseliny. Je-li v buňce tryptofanu dostatek, dvě jeho molekuly se naváží na represor, čímž změní jeho strukturu a represor se oddělí od DNA a syntéza ustane.

Tryptofanový represor je dobře prostudovaným systémem a proto je možné ho použít jako model na ověřování kvality silových polí. Existují krystalové struktury represoru s DNA i navázanými tryptofany nebo jinými koligandy.

Ramesh a spol. (1996) zkoumali výše uvedený systém pomocí jaderného Overhauserova jevu a dospěli k závěru, že se různé ligandy ve vazebném místě různě orientují. Tento tým používal kromě

tryptofanu jako koligandy 3-indolpropionovou kyselinu, 5-methyltryptofan a trans 3-indolakrylovou kyselinu. Prokázali, že ligandy lze podle orientace ve vazebném místě rozdělit na dvě skupiny. Do první skupiny patří L-tryptofan a 5-methyltryptofan, jejichž orientace ve vazebném místě odpovídá orientaci L-tryptofanu a zbývající dva ligandy jsou o přibližně 120° otočeny ve směru hodinových ručiček. Ramesh ve své práci tento jev nevysvětluje.

Ke stejnému závěru jako Ramesh došli dříve i Lawson a Sigler (1988), kteří jako ligandy používali L-tryptofan a indolpropionát. Strukturu řešili pomocí krystalografie a následné rentgenové difrakce (struktura nebyla v ProteinDataBank nalezena).

Bylo zjištěno (Marmorstein, 1987), že ligandy první skupiny aktivitu represoru téměř neovlivňují, kdežto ligandy druhé skupiny působí jako inhibitory na represor, mají také menší disociační konstantu než ligandy první skupiny.

## 2.4 Fotosystém II

Proces, při kterém dochází k přeměně světelné energie na energii biochemickou, nazýváme oxygení fotosyntézou. Během tohoto procesu dochází k sérii přenosů elektronů z molekul vody na NADP<sup>+</sup>, tyto elektrony se následně použijí na syntézu ATP (Wollman a spol., 1999). Oxygení fotosyntéza se odehrává v thylakoidních membránách, v nichž najdeme 5 typů proteinů: cytochrom *b<sub>6</sub>f* komplex, jádro fotosystému I, jádro fotosystému II, anténní proteiny a proton-ATP syntázu. Všechno jsou multimerní jednotky, všechny kromě cytochromu prostupují membránou a všechny obsahují různé kofaktory – pigmenty, hemy, železno-sírné klastry a kovové ionty.

Fotosystém II si můžeme rozdělit na vlastní reakční centrum, kde dochází ke štěpení náboje, jeho separaci a k přenosu elektronu a na anténní systém, jehož hlavní funkcí je zachyt světelné energie a její přenos do centra fotosystému.

Spolu se svými anténními proteiny se fotosystém II skládá z více než 25 podjednotek a alespoň z devíti různých redoxních komponentů (chlorofyl, feofitin, plastochinon, tyrosin, mangan, železo, cytochrom *b<sub>559</sub>*, katotenoidy a histidin), které se podílejí na přenosu elektronů. Pět z nich se podílí na transportu elektronů z vody na tzv. plastochinon pool. Jsou to manganový klastr (Mn)<sub>4</sub>, tyrosin (Y<sub>z</sub>), reakční centrum chlorofylu (P<sub>680</sub>), feofitin a dvě molekuly plastochinonu (Q<sub>A</sub> a Q<sub>B</sub>) (Whitmarsh, Govindjee, 2002).

Fotosystém II přenáší elektron z vody na plastochinon. Elektron vzniká díky fotoindukované separaci nábojů mezi primárním donorem P<sub>680</sub> a akceptorovou molekulou feofitinu. Odebráním elektronů z vody vzniká kyslík, který se uvolňuje do atmosféry. Je to jediný známý proteinový komplex, který oxiduje vodu na kyslík. PSII se podílí na dvou základních reakcích: oxidaci vody a redukci plastochinonu. Tyto reakce mají za následek štěpení náboje v reakčním centru. Na jeden absorbovaný foton připadá jeden uvolněný elektron a tedy na jednu uvolněnou molekulu vody je třeba absorpce čtyř fotonů.

Fotosystém II se u eukaryotních organismů nachází v thylakoidní membráně, převážně v tzv. granech (seskupení thylakoidů – výběžků vnitřní membrány), zatímco fotosystém I najdeme převážně ve stromální membráně. U prokaryot, která neformují grana, pak najdeme oba fotosystémy v thylakoidní membráně.

Fotosystém II je první z řetězce fotosyntetických přenašečů. Dochází na něm k zachycení světla, které je pomocí fotosběrných antén odváděno do tzv. reakčního centra. Na lumenární straně fotosyn-

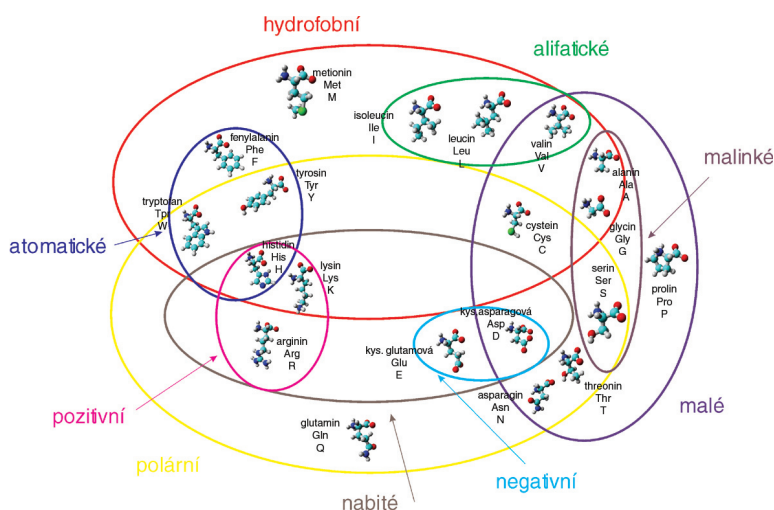
tetické membrány štěpí foton molekulu vody na atom kyslíku, dva vodíky a dva elektrony. Na tomto jevu, jehož princip není ještě zcela objasněn, se podílí vnější proteiny fotosystému II mající u vyšších rostlin a zelených řas označení PsbO, PsbP a PsbQ (v poslední době se spekuluje i o významu proteinů PsbR a PsbT) u ostatních fotosyntetických organismů se označují PsbO, PsbU a PsbV a tzv. kyslík vyvíjející komplex, který sestává ze 4 manganových atomů a atomu sodíku. Kyslík a protony zůstávají v lumenu, kdežto elektrony putují přes systém pigmentů umístěných v jádru fotosystému II na přenašeč plastochinon. Přitom dochází k přenosu nábojů především na systému 6 chlorofylů a 2 feofitinů, což je předmětem našeho výzkumu. Tyto pigmenty jsou vázány čtyřmi proteiny reakčního centra, které mají označení D1 (PsbA), D2 (PsbD), CP47 (PsbB) a CP43 (PsbC). Palenčář (2007) ve své disertační práci ukázal, že aby byl systém stabilní v oktanové membráně nestačí modelovat pouze toto jádro, ale je třeba přidat i další podjednotky fotosystému II.

Do současné doby se nepodařilo získat strukturu fotosystému II z vyšších rostlin v postačujícím rozlišení pro molekulárně-dynamické simulace (existují struktury s rozlišením  $\approx 8 \text{ \AA}$ , což je nedostatečné) a proto bylo třeba proteiny modelovat.

## 2.5 Homologní modelování

Homologní modelování je metoda, která na základě známe primární sekvence bílkoviny vytváří modely jejich terciární struktury. Jsou založeny na předpokladu, že proteiny se stejnou nebo podobnou primární strukturou mají stejnou nebo podobnou terciární strukturu (což ovšem není vždy pravda; Kabsch, 1985).

Jako vstup pro modelování se používá známá 3D struktura bílkoviny a tzv. alignment, což je přiřazení známé a neznámé primární sekvence tak, aby homologní aminokyseliny byly pod sebou. Pokud nejsou aminokyseliny v daném místě homologní je snaha přiřadit k sobě alespoň aminokyseliny s podobnými vlastnostmi, viz. obrázek 2. Alignment se samozřejmě nevytváří ručně, ale na jeho tvorbu existují programy, které ze všech možných alignmentů vybírají ten s nejnižší hodnotou skóre. To je počítáno na základě tzv. vzdálenostních matic, které byly určeny experimentálně a které zachycují pravděpodobnost změny dané aminokyseliny a jinou aminokyselinu. Nejběžněji používané matice jsou PAM80 (Dayhoff, 1978), která se osvědčila pro krátké sekvence s poměrně vysokou homologií a BLOSUM45 (Henikoff, 1992), která se používá pro delší sekvence s nižší homologií, což je při modelování častější případ.



Obr. 2.: Aminokyseliny s podobnými vlastnostmi

Výhodou a současně nevýhodou homologního modelování je, že pokaždé obdržíme model. Proto je třeba správnost modelu uvěřovat experimentálně. Obecně lze říct, že pokud mají primární sekvence homologii větší než 30%, tak modely vychází poměrně přesně. Naopak u sekvencí s homologií menší než asi 25% je přesnost modelu diskutabilní. Přesněji také vychází v modelu oblasti s definovanou sekundární strukturou (šroubovice, skládané listy) než kličky.

Modely vytváří program Modeller (Sali, Blundell, 1993), který využívá metody tzv. srovnávacího modelování (comparative modeling). Jak bylo uvedeno výše, jako vstup do Modelleru používáme sekvenční alignment. Jako výstup dostaneme model bez vodíkových atomů. V prvním kroku program spočítá omezení na vzdálenosti a dihedrální úhly z alignmentu. Z těchto omezení spočítá vztah mezi jednotlivými aminokyselinami. Tyto vztahy se porovnají s databází známých 416 proteinů a podle toho jsou spočítány opravy pro jednotlivé parametry. Na základě těchto hodnot se spočítají pravděpodobnostní hustotní funkce pro každý parametr. Správná poloha daného atomu vznikne kombinací těchto pravděpodobnostních funkcí. V následujícím kroku se z těchto okrajových podmínek a z vypočtené CHARMM energie (Brooks a spol., 1983) vypočítá objektivní funkce, která se následně optimalizuje metodou spojených gradientů (conjugate gradients) a molekulární dynamiky se simulovaným žíháním.

Modelů obvykle počítáme 10 a nejlepší model z nich vybíráme na základě tzv. Ramachandranova diagramu, což je graf závislosti torzního úhlu  $\psi$  na úhlu  $\phi$ . Experimentálně bylo dokázáno, že aminokyseliny v určitých typech sekundární struktury zaujímají určité oblasti tohoto diagramu a že existují oblasti, kde se aminokyseliny nevyskytují. Dalším ukazatelem kvality modelů je g-faktor, který zohledňuje stereochemické parametry (např. délka vazby, hodnoty torzních úhlů...) a objektivní hodnotící funkce modelleru, kterou program počítá. Obecně lze říct, že v prvním kroku kontrolujeme, jestli se objektivní funkce modelleru neliší vzájemně o více než o 10 %. Pokud se liší, modely s vysokou hodnotou funkce vyřazujeme. Hodnotu ostatních dvou ukazatelů zohledňujeme 60:40 ve prospěch Ramachandranova diagramu. Pokud máme pro celý hledaný protein více templátů, obvykle modelujeme podle každého templátu zvlášť i podle kombinace templátů a navzájem porovnáváme i nejlepší modely jednotlivých skupin mezi sebou.

V případě, že některé části templátu chybí, využíváme serveru ModLoop (Fiser a spol., 2000), který modeluje klíčky na základě metody popsané v citovaném článku. Program modeluje až 14 aminokyselin dlouhou sekvencí. V případě, že je tato sekvence na konci proteinu, lze modelovat libovolně dlouhou sekvencí tak, že počátek proteinu odstraníme v prohlížeči tak, aby zbylo pouze 14 aminokyselin, jejichž poloha není ve vzorové struktuře a modelujeme tento úsek. Následně v témže programu sloučíme výsledek proteinu po modelování klíčků s počáteční sekvencí a tento protein uřízneme tak, aby na jeho konci bylo pouze 14 neznámých aminokyselin, modelujeme klíčky, atd. Nutno zdůraznit, že výsledek tohoto programu velmi závisí na tom, jakou část proteinu modelujeme.

## 2.6 Molekulární dynamika a silová pole

Molekulární dynamika je teoretická metoda, která se snaží reprodukovat pohyb molekul. Stojí na Newtonově pohybovém zákonu

$$F = m \times \frac{d^2r}{dt^2},$$

kde  $F$  je síla působící na částici,  $m$  je její hmotnost,  $r$  polohový vektor a  $t$  čas.

A na vztahu mezi silou a potenciálem

$$F = -\frac{\partial V}{\partial r},$$

kde  $F$  je síla,  $V$  je potenciál a  $r$  je polohový vektor dané částice.

Protože se rovnají levé strany výše uvedených rovnic, rovnají se i jejich pravé strany. Můžeme tedy psát:

$$m \times \partial^2 r / \partial t^2 = -\partial V / \partial r .$$

Tato rovnice se následně řeší pro každý okamžik a každou částici systému.

V této rovnici jsou dvě neznámé: polohový vektor, který počítáme a potenciál. Potenciál, kterému v případě molekulárně-dynamických simulací říkáme silové pole, musí být do systému zanesen z vnějšku. Při zadávání silového pole musíme brát v úvahu pokud možno všechny síly, které působí na danou částici, takže zápis silového pole může vypadat např. takto:

$$U(r_i, \dots) = \frac{1}{2} \sum_{\text{vaz}} K_{\text{vaz}}^i (b^j - b_0^j)^2 + \frac{1}{2} \sum_{\text{uhl}} K_{\text{uhl}}^i (\theta^j - \theta_0^j)^2 + \frac{1}{2} \sum_{\text{rov}} K_{\text{rov}}^i d^{i2} + \sum_{\text{trz}} K_{\text{trz}}^i [1 + \cos(n\phi^i - \delta^i)] + \sum_{\text{nev}} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon r_{ij}} \right)$$

kde  $K_{\text{vaz}}^i$  je vazebná silová konstanta,  $b_0^j$  referenční délka vazby,  $K_{\text{uhl}}^i$  silová konstanta valenčního úhlu,  $\theta_0^j$  referenční valenční úhel,  $K_{\text{rov}}^i$  silová konstanta deformace planárního systému,  $d^i$  je odchylka od planarity systému,  $K_{\text{trz}}^i$  silová konstanta dihedrálního úhlu,  $n$  periodičita torzního potenciálu,  $\Phi$  je torzní úhel a  $\delta^i$  je fáze torzního úhlu, poslední člen rozvoje popisuje nevazebné interakce.

Důležitým parametrem je  $\Delta t$ , kterému se říká časový krok. Jeho délka se nastavuje podle typu interakce, která nás zajímá. Obecně lze říct, že časový krok je limitován nejkratšími vibracemi a rotacemi ve struktuře na 2 fs, na atomistické úrovni jsou to hlavně vibrace, na kterých se podílí vodíky. Použitím aproximace tzv. dummy hydrogens, při které nepočítáme pro vodíky všechny interakce při každém kroku, ale určujeme jejich pozici ze souřadnic okolních těžkých atomů, můžeme docílit delších časových kroků až 5 fs. Toto sice výrazně zkrátí dobu výpočtu, ale na druhou stranu nepostihneme všechny interakce v systému. Dalšími vstupními parametry pro molekulární dynamiku jsou teplota a tlak. Gromacs umožňuje použití periodických okrajových podmínek (periodic boundary conditions), což si lze představit jako, že kolem simulačního boxu umístíme tak obrazy a pro částice na okraji počítáme, jako by interagovaly s obrazy boxu.

Poslední důležitým parametrem, který nastavuje uživatel je délka simulace. I ta závisí na jevu, který se simulací snažíme postihnout. Lze říct, že doba simulace odpovídá době, než se protein dostane do rovnováhy. To poznáme sledováním různých parametrů. Obvykle sledujeme, zda jsou v rovnováze energie systému (kinetická, potenciální i celková energie, která je dána součtem obou předchozích), střední kvadratickou odchylku a poloměr setrvačnosti (radius of gyration), obojí pro  $C_\alpha$  uhlíky a závislost počtu vodíkových vazeb na čase. Poslední parametr je důležitý především pro sledování interakcí dvou a více proteinů.

Před začátkem simulace je třeba jednotlivé složky systému energeticky minimalizovat a protein neutralizovat ionty. Program Gromacs, který používáme, má v sobě funkci, která počítá rozložení náboje na proteinu a následně ionty automaticky umístí do energeticky nejvhodnějších míst. Tento podprogram lze současně použít pro umístění iontů jako ligandů na proteiny.

Jednotlivé hodnoty v silovém poli je třeba určit experimentálně nebo pomocí výpočtů (tento postup se obecně považuje za méně přesný, ale o to rychlejší). O tvorbu co nejvěrohodnějšího silového pole se snažily různé skupiny vědců a proto nás nepřekvapí, že existuje několik více či méně užívaných silových polí. V následujícím textu zmíním několik nejznámějších.



### 2.6.1 Silové pole Gromos

Toto pole bylo vyvinuto jako součást programu pro výpočet molekulární dynamiky Gromos na Univerzitě v Groningenu a na ETH v Curychu. V současné době se nejčastěji používá verze 43A1 (Scott, 1999), která byla optimalizována pro simulace proteinů, původní verze silového pole Gromos z roku 1986 byla optimalizovaná pro alkány v kondenzované fázi. S časem byly přidány parametry i pro jiné skupiny látek.

### 2.6.2 Silové pole Gromacs

Silové pole Gromacs (van der Spoel, 2005) bylo vyvinuto jako součást programového balíku Gromacs a je odvozeno ze silového pole Gromos. Je optimalizované především pro výpočty ve stejnojmenném balíku. Stejně jako předchozí i toto pole se dále vyvíjí.

### 2.6.3 Silové pole OPLS

OPLS je zkratka anglických slov Optimized Potential for Liquid Simulations, tj. optimalizovaný potenciál pro simulace tekutin. Jde o silové pole (Jorgensen, 1996), které bylo vyvinuto W. J. Jorgensenem a jak název napovídá, pole je vhodné pro simulace tekutin včetně membrán. Jako u jednoho z prvních byly k tomuto poli přidány parametry pro karbohydráty.

### 2.6.4 Silové pole Amber

Stejně jako první dvě, i toto pole (Cornell, 1995) je odvozeno od stejnojmenného balíku pro výpočet molekulární dynamiky. V současnosti se toto pole, které bylo původně vyvinuto skupinou kolem Petera Kollemana pro simulace DNA, rozrostlo na celou rodinu silových polí. Nyní je základem silové pole Amber99, které postihuje nukleové kyseliny a proteiny. Poté bylo vytvořeno silové pole GLYCAM, které obohacuje Amber o parametry pro cukry. A konečně GAFF (General Amber Force Field) snažící se postihnout všechny chemické sloučeniny.

## 2.7. Parametrizace silových polí

Parametrizace silového pole je postup, při kterém se experimentálně (měření infračervených spekter) nebo pomocí kvantově-mechanických výpočtů určují rovnovážné hodnoty (tj. vazby, úhly a dihedrální úhly) a silové konstanty pro n-tice atomů v molekule a parciální náboje na atomech molekuly. Většina používaných silových polí má parametry pro uhlovodíky, proteiny a nukleové kyseliny. Postupně jsou přidávány další parametry. Přesto může nastat situace, kdy potřebujeme počítat molekulu s neznámými parametry. V tom případě přichází ke slovu parametrizace, kdy určujeme rovnovážné veličiny molekul, jejich silové konstanty a parciální náboje na jednotlivých atomech molekuly.

Prvním krokem parametrizace je určení rovnovážných vzdáleností, úhlů, dihedrálů atd. molekule. K tomuto se používá kvantově-chemických výpočtů, kdy molekulu minimizujeme. Tato minimizovaná struktura pak slouží jako vstup pro další kroky. Musíme také zkontrolovat, jestli jsou všechny typy atomů v molekule definovány v silovém poli. Pokud ne, je třeba je dodefinovat a samozřejmě pro ně vypočítat všechny potřebné parametry.

Poté následuje výpočet parciálních nábojů. Existuje několik způsobů na výpočet parciálních nábojů, z nichž zmíním tři. Nejjednodušší metodou je výpočet tzv. Mullikenových nábojů (Mulliken, 1955), která se používají pro vyhodnocování reaktivity sloučenin. Většina silových polí používá přesnější

náboje CHELPG (Breneman, 1990). Posledním použitým typem nábojů jsou náboje RESP (Bayly, 1993), které používá silové pole Amber.

Posledním a současně nejobtížnějším typem parametrů, které je třeba určit, jsou silové konstanty, které si lze představit jako tuhost pružiny. Existuje několik způsobů, jak je určit. Nejjednodušší způsob je výpočet energie pro různé hodnoty parametru a jejich následné proložení křivkou potenciálu. Sofistikovanější metoda stojí na tom, že vibrační spektra molekuly (lze vypočítat pomocí kvantové chemie) odpovídají tzv. normálním modům kmitání molekuly. Silové konstanta ovlivňuje velikost normálních modů. Před nedávnem byla publikována tzv. AutoSmilies metoda (Krieger, 2004), ze které jsem ve své práci používala pouze tu část týkající se silových konstant. Tato metoda předpokládá, že podobné typy atomů mají stejné silové konstanty, což ve většině případů platí.

Zde bych chtěla ještě zdůraznit, že každé silové pole bylo vyvinuto na základě jiných předpokladů a tedy nelze parametry z jednoho pole kopírovat do druhého. Kromě toho i každý výpočetní balík požaduje jiný zápis silového pole nehledě na jiné jednotky veličin, proto ne každé silové pole lze bez konverze používat v libovolném výpočetním balíku pro molekulární dynamiku.

## 2.8 Kvantová chemie

Kvantová chemie je odvětví teoretické chemie, která aplikuje postupy kvantové mechaniky a kvantové teorie pole na chemické problémy. Vzhledem ke složitosti a matematické náročnosti kvantové mechaniky můžeme kvantovou chemií přesně popsat pouze velmi jednoduché problémy (např. atom vodíku), pro složitější problémy je třeba používat různá zjednodušení a aproximace. V současné době se kvantová chemie používá především pro výpočty energie molekul (jejíž minimizací můžeme optimalizovat molekuly), na výpočty rozložení elektrostatického potenciálu na molekule a na výpočty spektroskopických vlastností molekul jako je např. vibrační energie molekuly.

Ve většině případů řešíme pro náš systém bezčasovou Schrödingerovu rovnici:

$$\hat{H}\Psi = E\Psi$$

kde operátor  $H$  (tzv. Hamiltonián) je součet operátorů kinetické a potenciální energie,  $E$  je energie a  $\Psi$  je vlnová funkce.

Protože je mimo dosah současné výpočetní techniky řešit tuto rovnici pro libovolný systém musíme používat různá zjednodušení a aproximace. Podle rozsahu těchto můžeme kvantově-mechanické výpočty rozdělit na několik skupin.

Pokud nepoužíváme žádná empirická data, mluvíme o tzv. *ab initio* výpočtech. Nejjednodušším a současně i výpočetně nejméně náročným a tedy nejpoužívanějším postupem pro *ab initio* výpočty je tzv. Hatree-Fockova (HF) metoda, která nebere v úvahu vzájemné okamžité odpuzování elektronů, ale zohledňuje pouze součet vzájemného odpuzování všech elektronů. Variačním řešením Schrödingerovy rovnice se pak dostává k hodnotám energií. I složitější výpočetní metody začínají touto zjednodušenou metodou a postupně zavádějí korekce pro repulzi elektronů. Jedná se o metodu numeric-kou. Tato metoda je obvykle dostatečná pro jednodušší výpočty kvantové chemie jako je optimalizace molekuly nebo výpočty orbitalů.

Metody, které do řešení systému zanášejí část empirických dat se nazývají semiempirické nebo někdy empirické a i tyto stojí na HF teorii. Používají se především pro velké systémy jako optimalizace systému pro *ab initio* výpočty.

S časem samozřejmě vzniklo mnoho pokročilejších metod na kvantové výpočty, ale jejich výsledky se ve většině případů zásadně neliší od HF metody a jejich nepoměrně větší nároky na výpočetní výkon zastíní jejich přínos pro přesnost výpočtů na systému.

Zajímavou metodou, která hraničí mezi klasickou molekulovou dynamikou a kvantovou chemií je metoda ONIOM, která na část systému aplikuje molekulární dynamiku a na část kvantovou mechaniku. Tato metoda je implementována do programů na kvantovou chemii stejně jako do programů na molekulární mechaniku a je jí předvídána velká budoucnost. Přesto je k dnešnímu datu tato metoda ještě v počátcích a tomu odpovídají jednak její výsledky a pak problémy při jejím spouštění. I když jsem ve své práci této metodě věnovala poměrně hodně času, nepodařilo se mi ji zprovoznit, což přičítám na vrub nepřehledné struktuře komerčního software Gaussian (Frisch, 1998).

# 3. Metody

## 3.1 Parametrizace molekul

DGDG byl vybrán jako stavební prvek homogenní fotosyntetické membrány, protože je to nejčtenější (glyko)lipid v membráně, který tvoří dvojvrstvé struktury (MGDG tvoří lamely, proto je nevhodný). Porovnáním s fázovým diagramem (Shipley, 1973), se zjistilo, že při fyziologických podmínkách se DGDG nachází ve fázi tekutých krystalů.

Jako silové pole pro modelování membrány jsem zvolila pole OPLS, což je zdůvodněno v diskuzi. Parametry pro cukry nejsou součástí tohoto pole, ale byly do něho doplněny Dammem (1997). Rovnovážné vzdálenosti a CHELPG náboje používané v silovém poli OPLS byly vypočítány RHF/6-31-G\*v programem Gaussian98, jemuž kvůli značné velikosti systému předcházela minimizace nejprve semiempirickou AM1 metodou (Dewar, 1985). K určení silových konstant pak byla použita metoda AutoSmilies. Správnost parametrů byla dokázána pomocí molekulárně dynamické simulace po dobu 5 ns ve vakuu, při teplotě 300 K. Vlastnímu běhu předcházela energetická minimizace algoritmem steepest descent.

Stejný postup (kromě semiempirických výpočtů, které zde nebyly třeba) byl použit pro parametrizaci L-tryptofanu, 3-indolpropionové kyseliny, 5-methyltryptofanu a trans akrylové kyseliny pro projekt s tryptofanovým represorem. V tomto případě byla parametrizována silová pole Gromos, Gromacs, OPLS a Amber.

Parametry pro molekuly chlorofylu a feofitinu byly vyvinuty Palenčárem (2007) a byly implementovány do programového balíku Yasara. Mnou používaný program Gromacs vyžaduje jiný způsob zadávání parametrů i topologií molekul a proto byly parametry převedeny i do tohoto softwaru.

## 3.2 Stavba DGDG membrány

Struktura DGDG optimalizovaná kvantově-chemickými výpočty byla načtena do programu Yasara a zde byl zobrazen její molekulový povrch. Poté byla v tomto programu sestavena dvojvrstva o rozměru  $4 \times 4$  molekuly DGDG, která byla následně relaxována v programu Gromacs. Během tohoto procesu odpovídala simulační celá velikosti membrány, teplota byla udržována na 300 K připojováním systému k externí teplotní lázni ( $t = 0,1$  ps). Tlak byl udržován na hodnotě 1 bar připojováním na tlakovou lázeň ( $t = 0,1$  ps). Algoritmus LINCS byl použit pro omezení kovalentní vazbové délky a elektrostatické interakce dlouhého dosahu byly počítány metodou Particle-Mesh Ewald. Pro analýzu na konci běhu byla použita upravená trajektorie, kdy byl zohledněn pouze každý 50. krok.

Tato membrána o velikosti  $4 \times 4$  glykolipidy byla po optimalizaci nahrána do programu Yasara, kde byla rozšířena na velikost monovrstvy  $8 \times 8$  glykolipidů – složením 4 bloků po 32 glykolipidech. Následně byl i tento blok optimalizován v programu Gromacs za stejných podmínek jako blok předchozí.

U relaxované membrány byla počítána plocha na lipid, která se ukázala být více než dvakrát větší než u experimentální hodnoty (Durfane 1998). Proto byly lipidy v programu Gromacs postupně stlačovány o 10% původní velikosti membrány do hodnoty 70% původní velikosti a po každém stlačení byla membrána relaxována pomocí algoritmu „steepest descent“ a následné dynamiky s výše popsaným nastavením a s krokem 0,1 fs po dobu 1 ps. Od velikosti 65% původní membrány byla velikost

stlačování zmenšena na 5% původní velikosti a místo energetické minimalizace pomocí algoritmu „steepest descent“ se k minimalizaci používalo simulovaného žihání na teploty 310 K, 350 K, 330 K a 360 K v časech 0,1 ps, 0,5 ps, 1 ps a 2 ps. Následovala minimalizace s krokem 0,1 fs po dobu 2 ps.

Protože experimentu neodpovídala ani tloušťka membrány, byla membrána stlačena ve směru osy z tak, aby si hodnoty odpovídaly. I po tomto kroku byla membrána minimalizována pomocí simulovaného žihání a následně relaxována se stejnými parametry jaké byly použity pro zmenšování v rovině xy. Následovala vlastní produkce s nezměněnými parametry a krokem 2 fs po dobu 5 ns.

Poté, co se prokázalo, že se membrána nachází ve stavu rovnováhy, byla do systému přidána voda (model tip3p, který je konzistentní a nejčastěji používaný se silovým polem OPLS). K vymazání vody z prostoru membrány, kam ji program chybně umístil, byl použit skript killwater. Následně byla soustava minimalizována pomocí simulovaného žihání (stejně nastavení jako dříve) a voda byla optimalizována po dobu 10 ps s krokem 1 fs a výše uvedeným nastavením. Konečně byla spuštěna vlastní produkce s krokem 2 fs, která zachycuje časové období 5 ns. Simulace měla stejné nastavení jako předchozí produkce.

### 3.3 Vazebná energie tryptofanového represoru

Projekt s tryptofanovým represorem je především metodickým projektem, jehož cílem je na známém systému porovnat různá silová pole.

Jako vstup byla použita krystalová struktura 1TRO (Otwinowski, 1988), která byla získána z ProteinDataBank (Berman, 2000). Ta byla nahrána do programu Yasara, kde byly odstraněny řetězce DNA a jeden homodimer proteinu s navázaným tryptofanem. Zůstal tedy homodimer tryptofanového represoru se dvěma navázanými molekulami tryptofanu (řetězce A až D).

Tato molekula byla nahrána do programu Gromacs, kde byla relaxována po dobu 10 ns s časovým krokem 2 fs při teplotě 300 K a ve vodném prostředí (SCP voda), simulační buňka byla před vlastní relaxací neutralizována sodnými ionty. Simulace probíhala paralelně v silových polích Gromacs, Gromos a OPLS (simulaci v silovém poli Amber se rozběhnout nepodařilo). Po relaxaci systému byla spočítána vazebná energie mezi tryptofanem a proteinem.

Pro porovnání druhé – energeticky nevýhodné – polohy tryptofanu byla aminokyselina otočena manuálně v programu Yasara a celý výše popsaný proces byl zopakován.

Molekula 3-indolpropionové kyseliny byla postavena v programu Gabedit a následně optimalizována RHF/6-31G\*programem Gaussian98. Tato molekula byla následně otevřena v programu Yasara, kde byla manuálně umístěna na vazebné místo v tryptofanovém represoru orientovaná dle předpovědi citace. A poté tak, aby svou orientací odpovídala molekule tryptofanu. Molekula byla opět relaxována v programu Gromacs za stejných podmínek jako tryptofan a následně i zde byla počítána vazebná energie.

### 3.4 Reakční centrum fotosystému II

Primární sekvence monomerů reakčního centra z hrachu (*Pisum sativum*) byly získány z databáze Swissprot (Gasteiger, 2003). V případě potřeby byl z těchto sekvencí odstraněn tranzitní peptid. Dvě rostliny byly používány proto, že u hrachu, na kterém jsou prováděny experimenty, nejsou známy všechny potřebné primární sekvence. Palenčář ve své práci ukázal, že pro modelování reakčního cen-

podjednotka	1S5L	2AXT
PsbH	13 – 65	2 – 65
PsbJ	3 – 40	7 – 40
PsbK	1 – 37	10 – 46
PsbL	1 – 37	1 – 37
PsbM	1 – 30	1 – 36
PsbT	1 – 31	1 – 30
PsbZ	1 – 58	1 – 62

Tabulka 1: Šedě jsou vyznačeny templáty pro jednotlivé anténní podjednotky fotosystému II. Jádro bylo modelováno podle obou templátů současně

tra sinic nestačí použít pouze podjednotky tvořící vlastní reakční centrum (tj. řetězce PsbA až PsbD), ale je třeba modelovat i další řetězce, které drží molekuly chlorofylu a feofitinu ve svých vazebných místech. Jako templáty pro modelování byly použity krystalové struktury s pdb kódem 1S5L (Ferreira, 2004) a 2AXT (Loll, 2005), obě z *Thermosynechococcus elongatus*. Pro podjednotky jádra fotosystému vždy obě, pro anténní podjednotky se obvykle používala pouze jedna z těchto dvou sekvencí, konkrétně ta, která pro danou podjednotku měla více

aminokyselin, viz. tabulka 1, použitý templát je vyznačen šedě.

Pro každou rostlinu a každou podjednotku byly primární sekvence templátu a modelovaného proteinu nahrány do programu ClustalX (Thompson, 1997), kde byl udělán mnohonásobný alignment (matice Gonnet – Gonnet, 1992) a výsledek byl uložen.

Tento soubor byl společně s pdb souborem použit jako vstup do programu Modeller 9v2 (Sali, 1993), kde bylo pro každou rostlinu a každou sekvenci počítáno 10 modelů. Ty byly následně vyhodnoceny v programu PROCHECK (Laskowski, 1993), v úvahu se brala i hodnotící funkce počítaná Modellerem (Modeller objektive function).

V případě potřeby byly počítány struktury neznámých oblastí podjednotek (obvykle N- nebo C-koncové oblasti) pomocí internetového serveru ModLoop (Fiser, 2000). Program předpokládá, že neznámá oblast je tvořena neuspořádanou sekundární strukturou, proto bylo třeba tento předpoklad potvrdit pomocí serveru PredictProtein (Rost, 2003), který předpovídá sekundární strukturu proteinů na základě vlastností aminokyselin.

Jádro fotosystému II vyšších rostlin bylo sestaveno nařítováním vymodelovaných podjednotek hrachu a špenátu na krystalovou strukturu 2AXT. Poloha 6 chlorofylů a 2 feofytinů byla ponechána stejná jako u sinice.

# 4. Výsledky

## 4.1 Fotosyntetická membrána

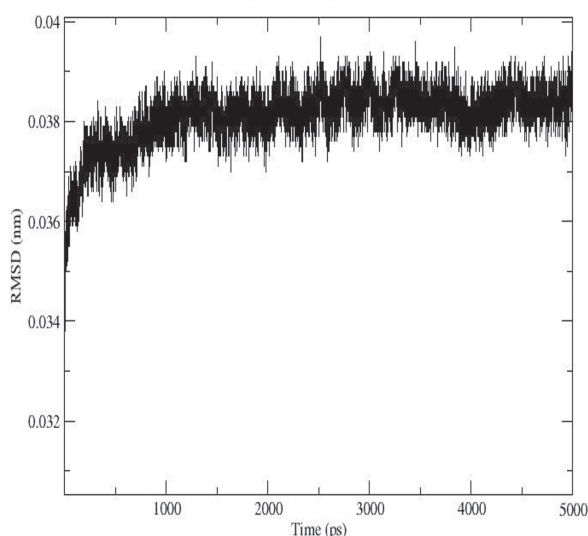
Prvním krokem při parametrizaci DGDG bylo určení atomových typů molekuly a dodefinování typů používaných Dammem (1997) pro cukry.

Následně byly metodou AutoSmilies určeny parametry pro vazby a úhly, které tvoří část atomů určená podle Damma a část atomů podle původních hodnot v silovém poli OPLS, ostatní hodnoty silových konstant nebylo třeba dopočítávat. Aby byly parametrizace kompletní, byly vypočítány CHELPG náboje pro molekulu DGDG, jejichž hodnoty jsou uvedeny v tabulce 2 na následující straně. Obrázek 3 znázorňuje molekulu DGDG, kde jsou atomy očíslovány stejně jako v předchozí tabulce.

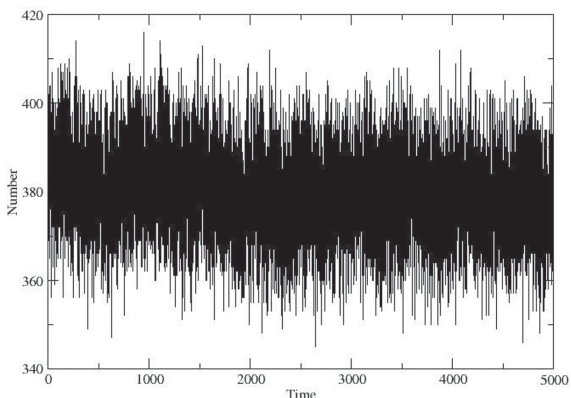
Podářilo se relaxovat membránu o velikosti  $4 \times 4$  glykolipidy a následně i o velikosti  $8 \times 8$  glykolipidů. Ta byla stlačena tak, aby „area per lipid“ a tloušťka dvojvrstvy odpovídala experimentu (Durfene, 1998) a následně relaxována a uvedena do rovnováhy (grafy nejsou ukázány).

Podářilo se relaxovat soustavu membrána-voda (viz. graf 1) a tím vytvořit model DGDG membrány. Správnost modelu potvrzuje i přítomnost vodíkových vazeb mezi cukrovými částmi glykolipidy mezi sebou a mezi cukrovými částmi glykolipidu navzájem, viz. grafy 2 a 3. Výslednou

membránu můžeme vidět na obrázku 4, obrázek 5 pak ukazuje výsek membrány, kde jsou znázorněny vodíkové vazby, oba obrázky jsou na další straně.

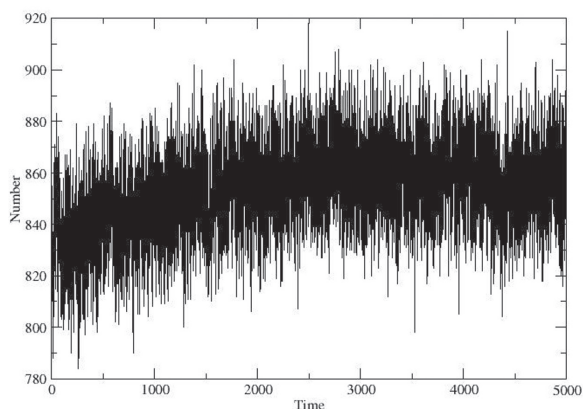


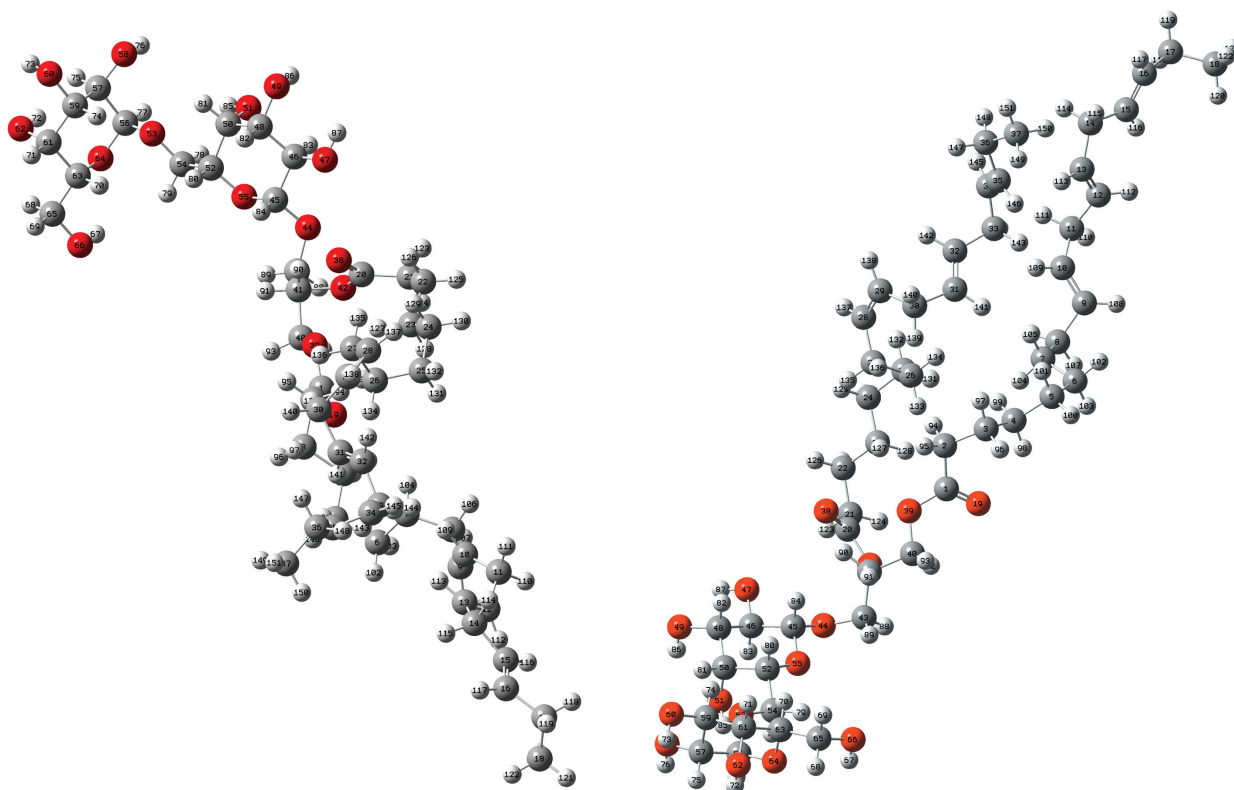
Graf 1 (vlevo nahoře): vývoj střední kvadratické odchylky rozměrů systému v čase pro membránu DGDG s vodou



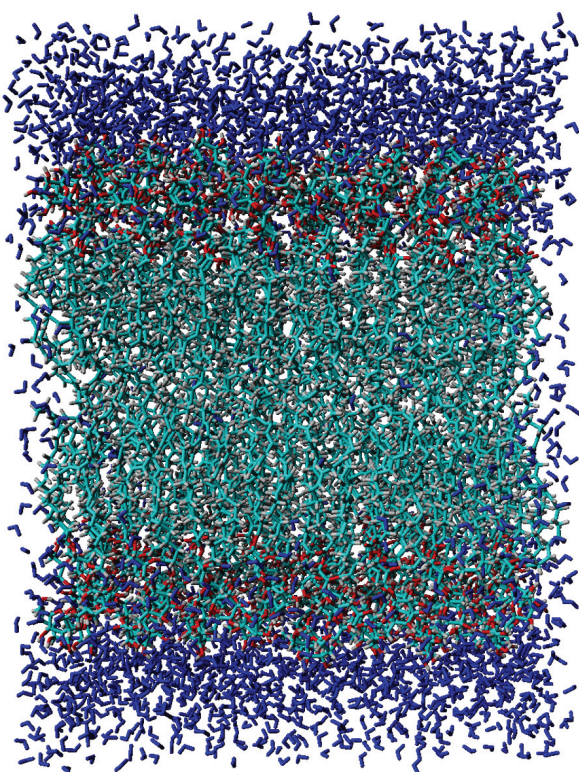
Graf 2 (vlevo dole): závislost počtu vodíkových vazeb mezi glykolipidy navzájem na čase

Graf 3 (vpravo dole): závislost počtu vodíkových vazeb mezi glykolipidy a vodou na čase



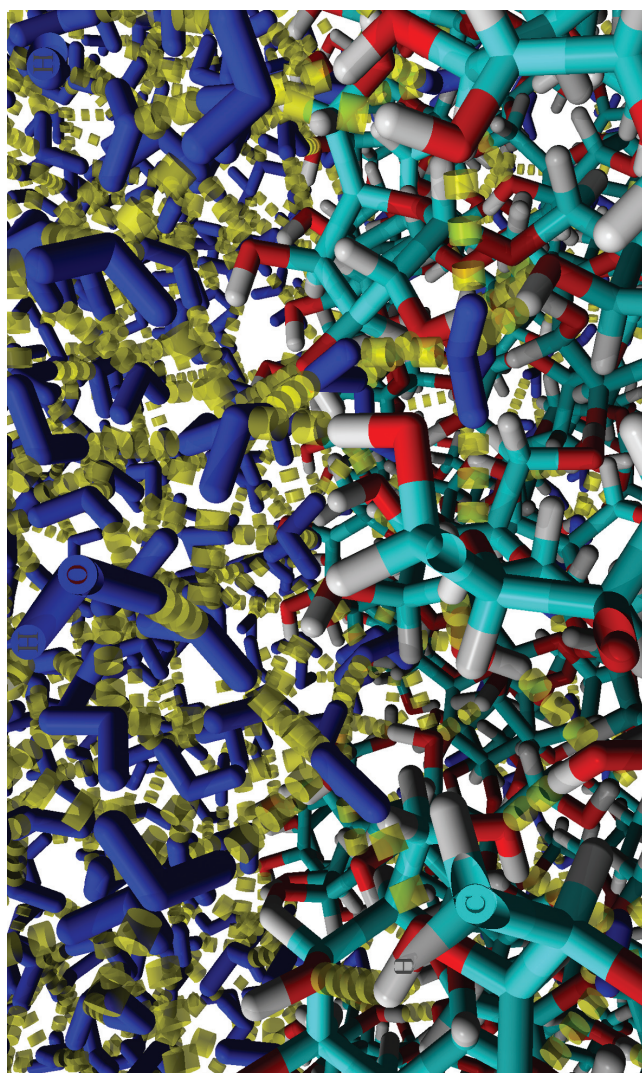


Obr. 3.: Čísla atomů DGDG k tabulce 2 na následující straně



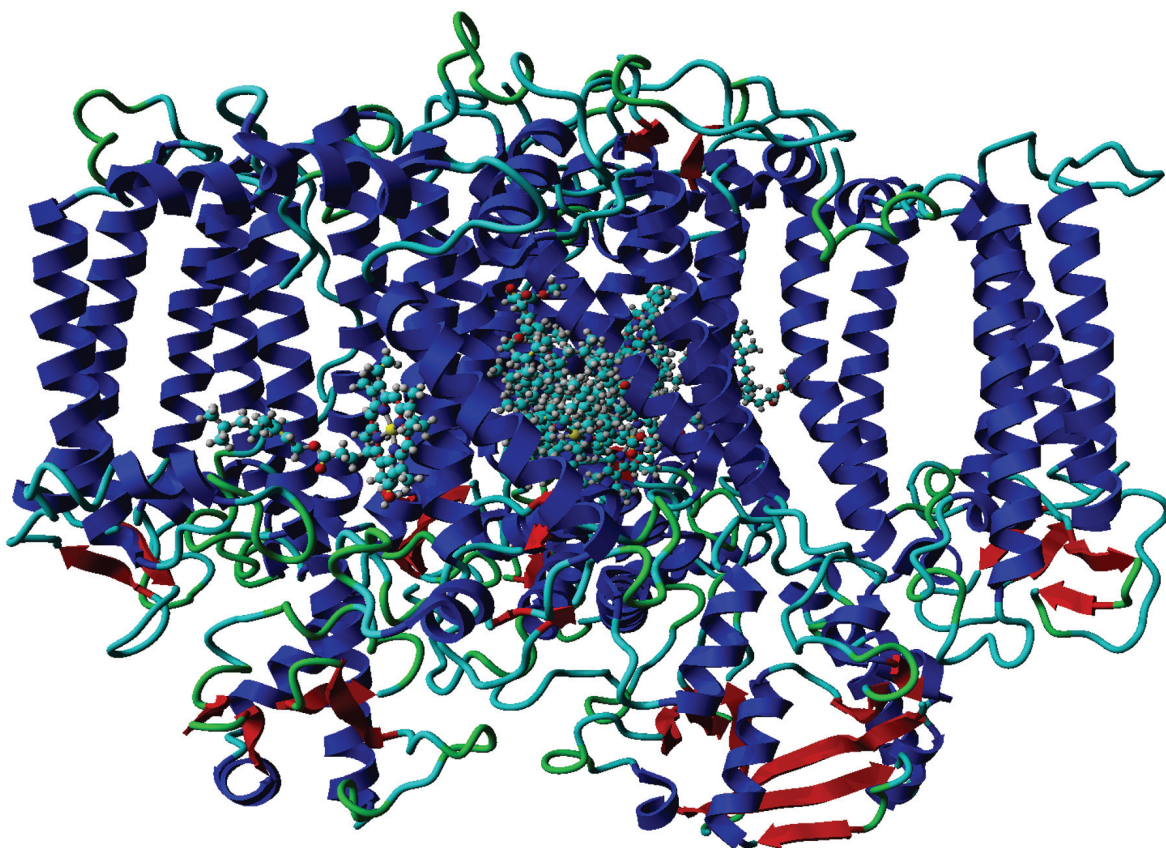
Obr. 4. (nahore): DGDG membrána po 5 ns relaxace. Voda se na okraje membrány dostala během simulace, díky změnám objemu simulačního boxu vyvolaných změnami tlaku.

Obr 5. (vpravo): Oblast sacharidů v DGDG membráně a na ni navazující vodné prostředí (modře). Zřetelně jsou vidět vodíkové vazby v systému (žlutě).





1	C	0.956944	51	O	-0.633289	101	H	0.009742
2	C	-0.371999	52	C	-0.143299	102	H	0.004231
3	C	0.051341	53	O	-0.458490	103	H	-0.008514
4	C	-0.037405	54	C	0.309370	104	H	0.013372
5	C	-0.075513	55	O	-0.439217	105	H	0.011461
6	C	-0.093734	56	C	0.278654	106	H	0.036565
7	C	0.146205	57	C	0.290071	107	H	0.006597
8	C	-0.103841	58	O	-0.703203	108	H	0.018705
9	C	-0.183724	59	C	0.257752	109	H	0.003461
10	C	-0.216017	60	O	-0.693580	110	H	0.125123
11	C	-0.012865	61	C	0.005649	111	H	0.134514
12	C	-0.195462	62	O	-0.648064	112	H	0.057868
13	C	-0.259767	63	C	-0.089048	113	H	0.043701
14	C	0.095921	64	O	-0.465116	114	H	0.147951
15	C	-0.224011	65	C	0.453878	115	H	0.122768
16	C	-0.328131	66	O	-0.722054	116	H	0.056702
17	C	0.253508	67	H	0.111355	117	H	0.054090
18	C	-0.219280	68	H	0.104349	118	H	0.147557
19	O	-0.606807	69	H	0.028604	119	H	0.173051
20	C	0.754266	70	H	0.000826	120	H	-0.004384
21	C	-0.231762	71	H	0.010179	121	H	0.000503
22	C	-0.012027	72	H	0.037304	122	H	0.045133
23	C	0.083449	73	H	0.013981	123	H	0.060855
24	C	-0.044227	74	H	0.041574	124	H	0.052012
25	C	-0.009546	75	H	0.027176	125	H	0.077717
26	C	-0.103463	76	H	0.020025	126	H	0.053431
27	C	0.141284	77	H	-0.009958	127	H	0.006320
28	C	-0.241873	78	H	-0.001601	128	H	0.103370
29	C	-0.237477	79	H	0.045077	129	H	0.053203
30	C	0.111122	80	H	0.050867	130	H	0.036183
31	C	-0.296663	81	H	0.137911	131	H	0.067059
32	C	-0.118995	82	H	0.142811	132	H	0.455916
33	C	-0.022492	83	H	0.057366	133	H	0.075078
34	C	-0.178815	84	H	0.056224	134	H	0.444833
35	C	-0.305710	85	H	0.154333	135	H	0.192632
36	C	0.178190	86	H	0.166673	136	H	0.446039
37	C	-0.218368	87	H	0.036036	137	H	0.126965
38	O	-0.571663	88	H	0.031293	138	H	0.034768
39	O	-0.533323	89	H	0.160411	139	H	0.059168
40	C	0.018631	90	H	0.171526	140	H	0.124508
41	C	0.532486	91	H	-0.024121	141	H	-0.003427
42	O	-0.529795	92	H	-0.020840	142	H	0.451107
43	C	-0.085262	93	H	0.041284	143	H	0.073889
44	O	-0.380269	94	H	0.058953	144	H	0.442601
45	C	0.529414	95	H	0.049676	145	H	0.139018
46	C	0.040466	96	H	0.092303	146	H	0.450719
47	O	-0.649756	97	H	0.076155	147	H	0.115108
48	C	0.253285	98	H	0.020501	148	H	-0.034739
49	O	-0.683461	99	H	0.023412	149	H	0.017414
50	C	-0.022213	100	H	-0.012600	150	H	0.438176



Obr. 6.: Model reakčního centra fotosystému II z hrachu s navázanými kofaktory (feofitiny a 6 chlorofylů).

## 4.2. Jádru fotosystému II pro hrách a špenát

Podářilo se sestavit homologní modely podjednotek reakčních center fotosystému II a jejich přilehlé podjednotky z hrachu (*Pisum sativum*). Byl sestaven multimer reakčních center fotosystému II a do něj byly implementovány fotosyntetické pigmenty tak, aby model co nejvíc odpovídal skutečnosti. Alignment pro jednotlivé modelované podjednotky neobsahuje téměř žádné inserce ani delece (není uveden). Na obrázku 6 lze vidět výsledný multimer jádra fotosystému II včetně pigmentových molekul.

Tento model bude v dalším kroku bude použit pro vložení do membrány a zkoumání distribuce nábojů na pigmentech fotosystému II.

## 4.3 Tryptofanový represor

Podářilo se minimizovat, spočítat CHLPG náboje a určit atomové typy silových polí Gromos, Gromacs a OPLS pro molekuly L-tryptofanu a 3-indolpropionové kyseliny. Vazebné konstanty v případě potřeby byly určeny metodou AutoSmilies (podařilo se parametrizovat i molekuly 5-methyltryptofanu a trans 3-indolakrylové kyseliny, ale protože tyto nejsou v diplomové práci používány, parametry nejsou uvedeny).

Vazebné energie pro oba ligandy vypočítané po 10ns molekulárně-dynamickém běhu jsou uvedeny v tabulce 3. Na jejich základě můžeme prohlásit, že použitá silová pole lépe popisují 3-indolpropionovou kyselinu než L-tryptofan, v jehož případě se výsledky pohybují na hranici chyby nebo vykazují oproti předpovědi opačný trend (především Gromos). Pro 3-indolpropionovou kyselinu poskytuje

nejlepší výsledky silové pole Gromos, naopak silové pole OPLS vychází na hranici chyby.

Celkově mohu prohlásit, že simulace s tryptofanovým represorem (dle očekávání doc. Carey) dosud nepřinesly žádné průkazné výsledky a bude třeba ho zopakovat s dalšími vstupními polohami našich ligandů příp. přidat další ligandy.

<b>L-tryptofan</b>				
<b>pole</b>	<b>experiment</b>	<b>otočeno</b>	<b>rozdíl</b>	<b>procenta</b>
gromacs	-1824	-1804	-20	98,9
gromos	-1093	-1319	226	120,7
OPLS	-7894	-7997	103	101,3

<b>3-indolpropionová kyselina</b>				
<b>pole</b>	<b>experiment</b>	<b>otočeno</b>	<b>rozdíl</b>	<b>procenta</b>
gromacs	-1516	-1452	-64	95,8
gromos	-1157	-854	-303	73,8
OPLS	-7798	-7730	-68	99,1

Tabulka 3: Hodnoty vazebné energie pro L-tryptofan a 3-indolpropionovou kyselinu umístěné v poloze pozorované experimentálně a v otočené poloze v kJ/mol, jejich rozdíly a procentuelní poměr prvního vůči druhému.

# 5. Diskuse

## 5.1 Glykolipidová membrána

Glykolipidy jsou z molekulárně-dynamického hlediska zajímavé především proto, že mají polární a nepolární část. Toto se ukazuje jako problém především při výběru silového pole pro jejich výpočty. Především bylo třeba vybrat silové pole, pro které existují parametry jak pro lipidy, tak pro sacharidy. Parametry pro druhý typ sloučenin do nich byly dodány dodatečně. Můžeme diskutovat, zda volba silového pole OPLS je nejvhodnější.

V současné době se pro simulace membrán nejčastěji používají silová pole CHARMM (Brooks, 1983) a Bergerovo silové pole (Berger, 1997), které stojí na starších verzích silových polí Amber a OPLS. Bergerovo silové pole nemá parametry pro sacharidy, silová pole CHARMM a Amber nejsou standardní součástí programového balíku Gromacs, proto bylo použito silové pole OPLS, se kterým jsou navíc zde na oddělení zkušenosti pro molekulární dynamiku membrán. (Existují balíky, které rozšiřují program Gromacs o silová pole Amber a CHARMM, avšak jejich implementace do Gromacsu není zcela bezproblémová, proto – i přes počáteční snahu o jejich použití – bylo nakonec jako silové pole pro membránu zvoleno silové pole OPLS.) Není ovšem vyloučeno následné převedení membrány i do jiného pole (uvažuje se hlavně o silovém poli Amber).

Diskuze, do jaké míry parametry pro cukry v použitém silovém poli OPLS vystihují realitu je mimo dosah předkládané práce, proto fakt, že byly publikovány v impaktovaném časopise (Journal of Computational Chemistry, impakt faktor v roce publikace 2,8) beru za dostatečné ověření jejich správnosti.

DGDG se v rostlinných membránách nevyskytuje jako jedna přesně definovaná sloučenina, ale liší se počtem a polohou dvojných vazeb v lipidové části. Toto je třeba brát v úvahu při porovnávání modelu s experimentálně naměřenými hodnotami, pokud nebyly měřeny na čistém, přesně definovaném DGDG, ale pokud byly měřeny na DGDG z rostlinného extraktu. Jeho rozdílné isomerie může ovlivnit vlastnosti DGDG v řádech procent.

Nahrazení fotosyntetické membrány membránou z DGDG, tedy důvod, proč byla DGDG membrána konstruována, je také předmětem diskuze. Především se ukazuje, že vlivem negativně nabitých lipidů má membrána náboj, který membrána složená pouze z neutrálního DGDG mít nemůže. Bez povšimnutí nemůže zůstat ani rozdílná hustota a vlastnosti související s fázovými přechody mezi fotosyntetickou a DGDG membránou. Bylo dokázáno, že přítomnost MGDG je především kvůli tvaru jeho průřezu nezbytná pro správné zanoření fotosystému II. Toto je třeba brát v úvahu při tvorbě modelu fotosystému II v membráně a případným odstraněním anténních podjednotek se pokusit dostat průřez proteinového multimeru aby byl pokud možno čtvercový a tím se zbavit nutnosti přítomnosti MGDG.

## 5.2 Molekulární modelování

Molekulární modelování je metoda, která stojí na předpokladu, že proteiny se stejnou nebo podobnou (do cca 20%) primární strukturou mají i stejnou nebo podobnou terciární a potažmo kvartérní strukturu. Ukázalo se (Kabsch, 1985), že tato podmínka není vždy splněna. V našem případě však modelujeme proteiny s vysokou homologií kolem 65% s modelovou strukturou, která má navíc stejnou funkci jako náš protein, proto zde lze předpokládat, že výše uvedený předpoklad platí.

Vzhledem k vzájemné vysoké homologii modelovaných proteinů a vzorových struktur nebylo třeba dělat manuální úpravy alignmentu, jehož správnost potvrzují i vysoké hodnoty stereochemického g-faktoru a nízké počty aminokyselin v zakázaných oblastech Ramachandranova diagramu. Podobnost primárních sekvencí je natolik vysoká, že bylo možno použít zvolenou matici programu ClustalX, tj. matici Gonnet 250, která i když nepatří mezi nejobvyklejší matice se v praxi osvědčila pro alignment sekvencí s homologií do cca 30 %.

Pro homologní modelování byl z dostupných programů vybrán program MODELLER proto, že tento program po vlastním vytvoření modelu ještě model optimalizuje pomocí metody simulovaného žhánání, tj. že se během optimalizace několikrát výrazně zvýší teplota, která se následně nechá klesnout na původní hodnotu. Tím se předchází tomu, že model zůstane v lokálním energetickém minimu.

Modelování klíčků pro chybějící vzorové struktury je obecně problém. Ve své práci jsem klíčky modelovala pomocí serveru ModLoop, který pro danou neznámou oblast hledá homologní oblast/oblasti, podle kterých klíčku vymodeluje. Je zajímavé sledovat, jak se se změnou velikosti zadávané oblasti na modelování mění i struktura klíčky. V těchto případech se správnost modelu klíčky dá odhadnout na základě běhu molekulární dynamiky, příp. i intuici modeláře.

Ve výše uvedeném textu je, že se nejlepší model vybírá na základě stereochemického g-faktoru a Ramachandranova diagramu. Tolik teorie, která poskytuje pouze hrubé rysy výběru. V praxi je třeba ještě všechny modely prohlédnout v prohlížeči a porovnat je mezi sebou. Obecně lze říct, že nejvíc modelů se shlukuje kolem jakési „nejpravděpodobnější struktury“ a čím víc se model od této struktury liší, tím je četnost podobných modelů menší. Přesto v některých případech nelze říct, že nejčetnější model je nejlepší a je třeba zapojit intuici a znalosti o funkci a výskytu proteinu.

Zajímavou otázkou je i kolik modelu od daného proteinu počítat. Obecně lze říct, že čím bude víc modelů, tím je větší pravděpodobnost, že se nám vyskytnou všechny energeticky možné struktury a my budeme moci vybrat tu nejpravděpodobnější na základě rozložení podobnosti modelů mezi sebou. Na druhou stranu působí tlak používat co nejmenší výpočetní čas (mít modely co nejdřív) a zde obecně platí, že čím méně modelů tím lépe. Na základě analýzy (Sovová, nepublikované výsledky) jsem dospěla k závěru, že pro většinu proteinů je optimální počet 10 modelů, který lze v případě potřeby zvýšit.

### 5.3 Molekulární dynamika

Molekulární dynamiku lze v současné době prohlásit za standardní prostředek bioinformatiky. Přesto i ona má jistá omezení, která je třeba při jejím použití brát v patrnost. Asi nejzákladnějším omezením je silové pole, tedy potenciál, který popisuje působení částic mezi sebou navzájem. Toto pole je třeba určit experimentálně, případně se do něj dají doplnit parametry na základě kvantově mechanických výpočtů. Všechna experimentální data jsou získávána s určitou nepřesností a proto se tato nepřesnost přenáší i do silového pole. Silová pole jsou obvykle počítána pro sloučeninu s určitými chemickými vlastnostmi a parametry pro ostatní látky jsou do nic dopočítávány *in silico*, přičemž počítačové určování parametrů se bere za méně přesné než experimentální určení parametrů. Mnou použitá metoda na určování parametrů metoda AutoSmilies nepatří mezi nejdůvěryhodnější metody, protože předpoklad, že podobné typy atomů mají stejné silové konstanty nemusí být vždy splněn. Navíc zde

vyvstává otázka co je a co už není podobný typ atomů. Použití metody AutoSmilies jsem ve své práci omezila pouze na určování silových konstant, rovnovážné konstanty byly počítány pomocí kvantové mechaniky, která je považována za spolehlivější a při parametrizaci *in silico* se bere za standardní postup. Jediný způsob, jak určit, že silové konstanty byly určeny správně je vizuální sledování již proběhlé simulace a hledání, jestli se nějaký atom nepohybuje příliš nebo naopak moc málo.

Další výrazný rozdíl mezi skutečností a molekulární dynamikou je existence tzv. cut-offu v molekulárně dynamických simulacích. To znamená, že Lennard-Jonesovy a někdy i Coulombovy interakce se počítají jen do předem nastavené hodnoty i když ve skutečnosti mají nekonečný dosah. Hodnota cut-offu se obvykle nastavuje alespoň na 7,5 Å a za touto vzdáleností se působení výše uvedených sil neuvažuje, ve skutečnosti tam ale tyto síly působí (i když nepatrně).

Nepřirozené jsou i periodické hranice, které používáme proto, že simulovaný systém je malý a u reálného systému stejné velikosti bychom museli zohledňovat okrajové jevy, čemuž se tímto vyhneme.

Molekulárně-dynamické simulace předpokládají i to, že jsou atomy v základním stavu, a proto nejsou schopné modelovat chemické reakce a procesy při kterých dochází k přenosu elektronů.

V neposlední řadě je třeba zdůraznit, že molekulární dynamika používá aparát klasické fyziky. To je pro většinu systémů za normální teploty v pořádku, ale při vyšších teplotách a pro lehké částice je třeba někdy používat relativistickou fyziku.

## 5.4 Reakční centrum fotosystému II

Palenčár ve své disertační práci prováděl výpočetní experimenty na fotosystému II v oktanu, který díky svým fyzikálním vlastnostem (především hustotě) byl hlavně v minulosti používán jako aproximace biologických membrán. Je otázkou, jestli lipidové okolí nemá stabilizující funkci na jádro fotosystému a tedy by kvůli potenciálnímu šetření výpočetního času nebylo od věci spustit krátkou dynamiku pouze na systém sestávající s jádra fotosystému II a pigmentů zanořených do membrány, aby se ukázalo, zda je v tomto případě nezbytná přítomnost ostatních podjednotek. Pokud nebude, budeme mít před sebou zajímavý příklad zobrazující omezení použití oktanu místo lipidové dvojvrstvy.

## 5.5 Výběr silových polí a ligandů do projektu s tryptofanem

První podmínkou pro výběr silových polí bylo, aby dané silové pole bylo možné implementovat do programového balíku Gromacs, ve kterém simulace probíhaly. Ukázalo se, že kromě standardně dodávaných polí Gromacs, Gromos a OPLS lze do Gromacsu ještě přidat silová pole CHARMM a Amber. I když k těmto polím existují poměrně podrobné návody pro jejich implementaci do Gromacsu, nepodařilo se mi je instalovat tak, abych mohla prohlásit, že fungují bezchybně a proto bylo nakonec od jejich použití pro diplomovou práci upuštěno, což neznamená, že se jejich použitím nebudu zabývat v budoucnosti. Velký důraz je přitom kladen na spolehlivé fungování silového pole Amber, ve kterém Palenčár prováděl experimenty s fotosystémem II sinic a ve kterém plánujeme zopakovat stejné experimenty pro vyšší rostliny s výměnou oktánové vrstvy za DGDG membránu.

Ramesh a spol (1996) ve své práci, na které stojí náš experiment, používali 4 ligandy tryptofanového represoru: L-tryptofan, 3-indolpropionovou kyselinu, 5-methyltryptofan a trans 3-indolakrylovou kyselinu. Prokázal, že L-tryptofan a 5-methyltryptofan zaujímají ve vazebném místě stejnou orientaci

a zbývající dva ligandy (3-indolpropionová kyselina a 3-indolakrylová kyselina) mají identickou orientaci vůči sobě, která je rozdílná vůči prvním dvěma ligandům.

Při výběrů ligandů pro naše experimenty jsme vybrali z každé skupiny orientací jeden ligand, přičemž bylo zřejmé, že budu používat L-tryptofan jako přirozený ligand tryptofanového represoru. Při výběru mezi 3-indolpropionovou kyselinou a 3-indolakrylovou kyselinou byl rozhodující fakt, že první jmenovaná zaujímá ve vazebném místě víc poloh než druhá a tedy není tak náročná na přesné umístění ve vazebném místě represoru a tedy je časově méně náročné tento systém relaxovat.

## 6. Závěr

Podařilo se zvládnout metodu parametrizace silových polí a aplikovat ji na reálné biologické systémy. Podařilo se sestavit model jádra fotosystému II s přilehlými proteinovými podjednotkami a implementovat fotosyntetické pigmenty do tohoto modelu pro hrách a špenát. Parametry pro chlorofyl a feofitin byly převedeny ze silového pole v programu Yasara do silového pole v programu Gromacs. Tento postup nebyl nikým ke dnešnímu dni zveřejněn. Byly učiněny pokusy na tryptofanovém represoru, jejichž cílem bylo porovnat kvalitu silových polí dodávaných v programovém balíku Gromacs. Tyto pokusy však nebyly dostatečně průkazné. Za vrchol své práce považuji model glykolipidové membrány složené z DGDG, který odpovídá experimentům, protože k dnešnímu dni nebyl publikován žádný model glykolipidové membrány.



## 7. Použitá literatura

Bayly I. C., Cieplak P., Cornell W. D., Kollman P. A. (1993): A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* 97, 10269 – 10280

Berger O., Edholm O., F. Jähnig F. (1997): Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* 72, 2002 – 2013

Berman H. M., Westbrook J., Feng Z., Gilliland G, Bhat T. N., Weissing H., Shindyalov I. N., Bourne P. E. (2000): The Protein Databank. *Nucleic Acids Res.* 28, 235 – 242

Breneman C. M., Wiberg K. B. (1990): Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* 11, 361 – 373

Brooks B. R., Bruccoleri R. E., Olafson B. D., States D. J., Swaminathan S., Karplus M (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comp Chem* 4: 187–217

Brown B. S. (1996): *Biological Membranes*. BACS 8, 1 – 43

Cornell W. D., Cieplak P., Bayly C. I., Gould I. R., Merz Jr. K. M., Ferguson D. M., Spellmeyer D. C., Fox T., Caldwell J. D., Kollman P. A. (1995): A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 117, 5179 – 5197

Corti M., Cantu L., Brocca P., Del Fábero E. (2007): Self-assembly in glycolipids. *Curr Opin Colloid Interface Sci.* 12, 148 – 154

Damm W., Frontera A., Tidaro-Rives J., Jorgensen W. L. (1997): OPLS All-Atom Force Field for Carbohydrates. *J. Comput. Chem.* 18, 1955 – 1970

Dayhoff M. O., Schwarz R. M., Orcutt B. C. (1978): In *Atlas of protein sequences and structure* 5, 345 – 358

Dewar M. J. S., Zoebisch E. G., Healy E. F., Stewart J. J. P. (1985): Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* 107, 3902 – 3909

Dörmann P., Benning C. (2002): Galactolipids rule in seed plants. *Trends Plant Sci* 7, 112 – 118

Durfene Y. F, Boland T., Schneider J. W., Barger W. R., Lee G. U. (1998): Charakterization oh the physical properties of model biomembranes at the nanometer scale with the atomic force microscope. *Faraday Discuss.* 111, 79 – 94

Gasteiger E., Gattiker A, Hoogland C., Ivany I., Appel R. D., Bairoch A. (2003): ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784 – 3788

Ferreira K. N., Iverson T. M., Maghlaoui K., Barber J., Iwate S. (2004): Architecture of the Photosynthetic Oxygen-Evolving Center. *Science* 303, 1831 – 1838

Fiser A., Do R. K., Sali A. (2000): Modeling of loops in protein structures. *Protein Sci.* 9

Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Zakrzewski V. G., Montgomery, Jr. J. A., Stratmann R. E., Burant J. C., Dapprich S., Millam J. M., Daniels A. D., Kudin K. N., Strain M. C., Farkas O., Tomasi J., Barone V., Cossi M., Cammi R., Mennucci B., Pomelli C., Adamo C., Clifford S., Ochterski J., Petersson G. A., Ayala P. Y., Cui Q., Morokuma K., Malick D. K., Rabuck A. D., Raghavachari K., Foresman J. B., Cioslowski J., Ortiz J. V., Baboul A. G., Stefanov B. B., Liu G., Liashenko A., Piskorz P., Komaromi I., Gomperts R., Martin R. L., Fox D. J., Keith T., Al-Laham M. A., Peng C. Y., Nanayakkara A., Gonzalez C., Challacombe M., Gill P. M. W., Johnson B. G., Chen W., Wong M. W., Andres J. L., Head-Gordon M., Replogle E. S., Pople J. A. (1998): Gaussian, Inc., Pittsburgh PA

Gonnet G. H., Cohen M. A., Banner S. A. (1992): Exhaustive matching of the entire protein sequence database. *Science* 256, 1443 –1445

Henikoff S., Henikoff J. G. (1992): Amino acid substitution matrices for protein blocks. *Proc. Natl. Acad. Sci USA* 89, 10915 – 10919

Jorgensen W. L., Maxwell D. S., Tirado –Rivas J. (1996): Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118, 11225 – 11236

Kabsch W., Sander C. (1985): Identical pentapeptides with different backbones. *Nature* 317, 19 – 25

Krieger E., Darden T., Nabuurs S. B., Filkenstein A., Vriend G. (2004): Making optimal use of empirical energy functions: Force-field parametrization in crystal space. *Proteins* 57, 678 – 683

Laskowski R. A., MacArthur M. W., Moss D. S., Thornton J. M. (1993): PROCHECK: a program to check stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283 – 291

Lawson C. L., Sigler P. B. (1988): The structure of trp pseudorepressor at 1,65 Å shows why indole propionate acts as a trp „inducer“. *Nature* 333, 869 – 871

Lee A. G. (2000): Membrane lipids: It's only a phase. *Curr. Biology* 10, R377 – R380

Loll B., Kern J., Saenger W., Zouni A., Biesiadka J. (2005): Towards complete cofactor arrangement in 3,0 Å resolution structure of photosystem II. *Nature* 438, 1040 – 1044

Marmorstein R. Q., Joachimiak J., Sprinzl M., Singel P. B. (1987): The structural basis for the interaction between L-tryptophan and the Escherichia coli trp Aporepressor. *J. Biol. Chem.* 262, 4922 – 4927

Mulliken R. S. (1955): Electronic population analysis on LCAO – MO molecular wave functions I. *J. Chem Phys.* 23, 1833 – 1840

Otwinowski Z., Schevitz R. W., Zhang R.-G., Lawson C. L., Joachimiak J. L., Marmorstein L., Luisi B. F., Sigler P. B. (1988): Crystal structure of trp repressor/operator complex at atomic resolution. *Nature* 335, 321 –329

Palenčár P. (2007): The theoretical study of photosystem II, Chlorosomes and Flavoprotein Wrba, disertční práce, Ústav fyzikální biologie Jihočeské univerzity, Nové Hradky

Ramesh V., Syed S. E. H., Frederick R. O., Sutcliffe M. J., Barnes M., Roberts G. C. K. (1996): NMR studies of the mode of binding corepressors and inducers to Escherichia coli trp repressor. *Eur. J. Biochem* 235, 804 – 813

- Rost B., Yachdav G., Liu J. (2003): The PredictProtein Server. *Nucleic Acid Res.* 32, W321 – W326
- Sali A., Blundell T. L. (1993): Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779 – 815
- Scott W. R. P., Hünenberger P. H., Tironi I. G., Mark A. E., Billeter S. R., Fennen J., Torda A. E., Huber T., Krüger P., van Gunsteren W. F. (1999): The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* 103, 3596 – 3607
- Shibley G. G., Green J. P., Nichols B. W. (1973): The phase behavior of monogalactosyl, digalactosyl, and sulphoquinovosyl diglycerides. *Biochim. Biophys. Acta* 311, 531 – 544
- Singer S. J., Nicolson G. J. (1972): The Fluid Mosaic Model of the Structure of Cell Membranes. *Science* 175, 720 – 731
- van der Spoel D., Lindahl E., Hess B., Groenhof G., Mark A. E., Berendsen H. C. J. (2005): Gromacs: Fast, flexible, and free. *J. Comput. Chem.* 26, 1701 – 1718
- Thompson J. D., Gibbon T. J., Plewniak F., Jeanmougin F., Higgins D. G. (1997): The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 24, 4876 – 4882
- Tieleman D. P., Marrink S. J., Berendsen H. J. C. (1997): A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochim. Biophys. Acta* 1331, 235 – 270
- Wollman F-A., Minai L., Nachushtai R., (1999): The biogenesis and assembly of photosynthetic proteins in thylakoid membranes. *Biochim Biophys Acta* 1411, 21 – 81
- Whitnarsch J., Govindjee (2002): in *Encyclopedia of life science*

Přírodovědecká fakulta Jihočeské univerzity



# **Aplikace kvantově chemických metod a molekulárně dynamických simulací na reálné biologické systémy**

Diplomová práce

**Žofie Sovová**

**vedoucí práce: Mgr. Michal Kutý, Ph.D.**

Nové Hrady 2008

## Diplomová práce

Sovová Ž. (2008): Aplikace kvantově chemických metod a molekulárně dynamických simulací na reálné biologické systémy [Application of quantum-chemical methods and molecular-dynamics simulations on real biological systems Mgr. Thesis, in Czech] – 27 p., Faculty of Science, The University of South Bohemia, České Budějovice, Czech republic.

This metodological work is focused on parametrization of biological macromolecules in force field implementable to Gromacs software. The second part of the work deal with construction of glycolipid membrane.

Prohlašuji, že svoji diplomovou 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 diplomové práce, a to v nezkrácené podobě – v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou - elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách.

Dne 25.4.2008

.....  
Žofie Sovová

## Poděkování

Děkuji všem, kteří mi jakkoli pomohli s realizací diplomové práce. V první řadě děkuji svému školiteli, Mgr. Michalu Kutému, Ph.D., za množství cenných rad a připomínek, které mi věnoval v průběhu práce a za možnost tuto práci realizovat ve skvělém kolektivu Oddělení struktury a funkce proteinů Ústavu fyzikální biologie na Nových Hradech. Velký dík patří také mým konzultantům Doc. RNDr. Rüdigeru Ettrichovi, Ph.D. z ÚFB na Nových Hradech a Assoc. Prof. Janette Carey z Princeton University za jejich laskavý přístup, čas a trpělivost, kterou mi věnovali. V neposlední řadě chci poděkovat i ostatním kolegům z Oddělení struktury a funkce proteinů, kteří mají více zkušeností především s programem Gromacs, za jejich předání a za cenné rady k mé práci, ale především za kamarádský přístup, oporu a podporu ve všem. Ještě jednou Ti moc děkuji...

Děkuji především všem svým kamarádům a rodině za to, že mi po celou dobu byli oporou a také svým želvičkám Amálce a Pepíkovi za to, že jsou...

*The acces to the METACentrum computing facilities provided under the research intent MSM6383917201 is highly appreciated.*

# Obsah

<b>1. Cíle práce</b>	<b>1</b>
<b>2. Úvod</b>	<b>2</b>
2.1 Biologické membrány	2
2.2 Fotosyntéza a fotosyntetická membrána	3
2.3 Tryptofanový represor	4
2.4 Fotosystém II	5
2.5 Homologní modelování	6
2.6 Molekulární dynamika a silová pole	7
2.7. Parametrizace silových polí	9
2.8 Kvantová chemie	10
<b>3. Metody</b>	<b>12</b>
3.1 Parametrizace molekul	12
3.2 Stavba DGDG membrány	12
3.3 Vazebná energie tryptofanového represoru	13
3.4 Reakční centrum fotosystému II	13
<b>4. Výsledky</b>	<b>15</b>
4.1 Fotosyntetická membrána	15
4.2. Jádru fotosystému II hráchu	18
4.3 Tryptofanový represor	18
<b>5. Diskuse</b>	<b>20</b>
5.1 Glykolipidová membrána	20
5.2 Molekulární modelování	20
5.3 Molekulární dynamika	21
5.4 Reakční centrum fotosystému II	22
5.5 Výběr silových polí a ligandů do projektu s tryptofanem	22
<b>6. Závěr</b>	<b>24</b>
<b>7. Použitá literatura</b>	<b>25</b>