



Prof. dr hab. Andrzej Koliński
Wydział Chemii
Uniwersytet Warszawski

Warszawa, dnia 25.11.2015 r.

Recenzja rozprawy doktorskiej Pana magistra Filipa Leonarskiego: „Coarse-grained models for molecular dynamics of RNA. Evolutionary algorithm optimization of the potential Energy function parameters.” („Gruboziarniste modele dynamiki molekularnej RNA. Optymalizacja parametrów funkcji energii potencjalnej przy użyciu algorytmu ewolucyjnego.”)

Przedstawiona mi do recenzji rozprawa doktorska pana magistra Filipa Leonarskiego opisuje wyniki jego prac badawczych wykonanych w ramach studium doktoranckiego Wydziału Chemii Uniwersytetu Warszawskiego. Promotorami rozprawy są: pani profesor Joanna Trylska (Laboratorium Maszyn Biomolekularnych, Centrum Nowych Technologii Uniwersytetu Warszawskiego) i pan profesor Andrzej Leś (Pracownia Chemii Kwantowej Wydziału Chemii UW oraz Instytut Farmaceutyczny w Warszawie). W trakcie studiów doktorant był stypendystą Fundacji na Rzecz Nauki Polskiej (w ramach projektu TEAM „Antisense peptide nucleic acids as inhibitors of bacterial translation” kierowanego przez panią profesor Joannę Trylską), uczestnikiem programu „Preludium” Narodowego Centrum Nauki (jako kierownik projektu: „Coarse-grained molecular dynamics of RNA thermometers at elevated temperature”), a także programu Canaletto Ministerstwa Nauki i Szkolnictwa Wyższego (Executive Programme for Scientific and Technological Cooperation between the Italian Republic and the Republic of Poland „Computer modeling and simulations of nucleic acids structure and dynamics”).

Tekst rozprawy zawiera się na blisko 200 stronach maszynopisu. Układ treści jest typowy dla rozpraw doktorskich z nauk ścisłych. Pierwsze 30 (około) stron to wstęp teoretyczny. Następne 30 parę stron to opis metod obliczeniowych będących podstawą wykonanych badań. Na kolejnych trzydziestu stronach opisano przykładowe zastosowania zaproponowanej metody optymalizacji pola siłowego dla gruboziarnistych modeli RNA. Ponad 60 stron (Appendix A) to opis oprogramowania komputerowego RedMDSream, powstałego w trakcie wykonywania pracy doktorskiej i stanowiącego jej główny wynik. Na końcu umieszczony jest spis cytowanej literatury, zawierający 167 pozycji. Do tekstu rozprawy nie zostały dołączone kopie publikacji naukowych doktoranta, ściśle związanych z treścią rozprawy. Rozprawa napisana jest w języku angielskim. Pod względem językowym tekst rozprawy jest czytelny i raczej poprawny, aczkolwiek nie mogę tego dokładnie ocenić, gdyż angielski nie jest moim pierwszym językiem. Mogą trochę razić (nieliczne) wstawki „literackie”. O ile początek wstępu do rozprawy zaczynający się opowiadaniem o tym jak młodszy brat doktoranta biedził się nad szkolnym zadaniem domowym zawierającym wykres powiązań DNA, RNA i białek można uznać za nawiązanie do dalej omawianej roli i funkcji

RNA, to już rozpoczęcie końcowego podsumowania pracy od pytania „How to stop war of robots” („Jak zatrzymać wojnę robotów”) raczej trudno uzasadnić.

Rozdział 1 (Chapter 1) rozprawy, zatytułowany **Introduction**, przeczytałem z dużym zainteresowaniem. Ta część rozprawy jest wprawdzie dość krótka, ale w ciekawy sposób przedstawia podstawowe zagadnienia ważne dla dalszych badań. Na początku autor opisuje rolę cząsteczek RNA w funkcjonowaniu żywych organizmów, w szczególności w przekazywaniu informacji genetycznych. Dobrze omówiona jest złożoność struktury molekuł RNA, od struktury drugorzędowej do struktur wyższego stopnia złożoności oraz obecny stan badań strukturalnych, w szczególności krystalografii rentgenowskiej. Autor zwraca uwagę na konieczność poznania nie tylko statycznych struktur układów makromolekularnych, ale też ich dynamiki. Główna część wstępu teoretycznego przedstawia podstawowe założenia modelowania molekularnego, od mechaniki kwantowej, poprzez klasyczną mechanikę molekularną do modelowania gruboziarnistego. Tę część otwiera trochę dziwne odniesienie do heliocentrycznej teorii kopernikańskiej. Autor rozprawy widzi tę teorię, jako dobry przykład sytuacji nietypowej dla nauki, gdzie teoria wyprzedza eksperyment. Nie wydaje mi się, aby to była sytuacja nietypowa. Istotny postęp w nauce wydaje się bardzo często bazować na nowych teoriach, które wiodą do badań doświadczalnych. Poza tym, to problem do poważniejszej dyskusji i raczej prace Kopernika są bez ścisłego związku z tematem tej rozprawy. Dalsza część wstępu teoretycznego, pokazująca istotne założenia teoretyczne, z których wywodzi się metoda klasycznej mechaniki molekularnej, napisana jest bardzo dobrze. W przekonujący sposób przedstawione są też powody, dla których warto stosować też modele niższej rozdzielczości niż klasyczne modele atomowe. Aby sensownie porównać modele atomowe z modelami gruboziarnistymi autor szczegółowo opisał pole siłowe Amber i odpowiadające mu, a może raczej oparte na podobnym schemacie, pole siłowe gruboziarnistego modelu molekuł kwasów nukleinowych. Model ten, zakładający bardzo grubą reprezentację łańcuchów nukleinowych za pomocą pojedynczego pseudo-atomu na jedną resztę nukleotydową, został zaproponowany przez Trovato i Tozzini dla małych fragmentów DNA i znacznie później rozwinięty przez autora tej rozprawy oraz Trovato, Tozzini, Lesia i Trylską dla bardziej skomplikowanych cząsteczek RNA.

W tej części rozprawy zabrakło mi jakiegokolwiek wzmianki o alternatywnych do dynamiki molekularnej metodach symulacji, przede wszystkim z użyciem reprezentacji gruboziarnistej. Dotyczy to przede wszystkim różnych wariantów dynamiki Monte Carlo, które odpowiednio zaprojektowane mogą być potężnym narzędziem modelowania struktury i (długoczasowej) dynamiki wielkich układów biomakromolekularnych.

Przed sformułowaniem celów swojej rozprawy autor w rzeczowy sposób opisuje problem parametryzacji pól siłowych dla modeli cząsteczkowych różnej rozdzielczości. Jako przykłady wybrano schemat GAFF (General Amber Force Field) dla pełnoatomowego modelu Amber oraz sposób wyprowadzenia pola siłowego dla gruboziarnistego modelu UNRES. UNRES (United Residue) to dość znany model do symulacji struktury i dynamiki białek, rozwijany przez wiele lat przez Liwo i współpracowników. UNRES jest tu dobrym przykładem, gdyż opiera się na wykorzystaniu dobrze rozwiniętych schematów budowania pól siłowych wyprowadzonych dla modeli atomowych do przeniesienia na reprezentację gruboziarnistą. Autor rozprawy wydaje się dobrze rozumieć ograniczenia obu poziomów opisu oddziaływań cząsteczkowych. Po krótko przedstawione jest też alternatywne podejście do modelowania molekularnego, oparte na odpowiednim uogólnieniu wiedzy doświadczalnej. Autor zwraca uwagę na słabą stronę takiego, opartego na wiedzy

(knowledge-based) podejścia, polegającego na ograniczonej możliwości przenoszenia metody na układy słabo poznane doświadczalnie. Moim zdaniem to przynajmniej nie do końca prawda. Problem przenośności (parametryzacji) pola siłowego dotyczy zarówno modeli opartych na prawach fizyki jak i modeli opartych na uogólnionej wiedzy statystycznej. Odpowiednio zaprojektowane statystyczne pola siłowe dla modeli zredukowanych mogą w pewnym zakresie być łatwiej rozszerzane niż pola fizyczne.

Cel rozprawy jest dobrze zdefiniowany na końcu tego rozdziału (Chapter 1 **Introduction**). Autor zwraca uwagę na fakt, że rozwijanych jest bardzo wiele modeli gruboziarnistych i że większość z nich nie jest odpowiednio zdefiniowana i opisana, tak aby umożliwić ich łatwe stosowanie do różnych problemów modelowania molekularnego. Dlatego doktorant postawił sobie zadanie opracowanie w miarę prostego i możliwie uniwersalnego schematu optymalizacji modeli oddziaływań dla gruboziarnistych modeli biomolekuł. Metoda ma być w dużym stopniu zautomatyzowana przez odpowiednie oprogramowanie, umożliwiać optymalizację i rozszerzanie zakresu stosowalności istniejących modeli (reprezentacji) gruboziarnistych, a także wspomagać projektowanie nowych modeli.

W rozdziale drugim rozprawy (Chapter 2, **Methods**) opisano opracowane przez doktoranta oprogramowanie RedMDStream, służące do półautomatycznego optymalizowania pól siłowych dla gruboziarnistych modeli symulacyjnych biomakromolekuł. Celem tej optymalizacji jest wybranie najlepszych możliwych parametrów pól sił dla modeli gruboziarnistych. Typowy schemat metody RedMDStream polega na wykonaniu serii symulacji (dla zadanego układu modelowego) metodą dynamiki molekularnej. Symulacje MD wykonywane są dla klasycznego modelu pełnoatomowego i dla równoważnego modelu gruboziarnistego, z odpowiednio zaprojektowanym schematem oddziaływań pomiędzy pseudoatomami. Współczynniki skalujące różne komponenty pola sił modelu CG traktowane są jako zmienne, które wymagają optymalizacji, tak aby rezultaty MD dla modeli CG były jak najbardziej podobne do obserwowanych w MD dla klasycznego układu pełno-atomowego. Wymaga to zdefiniowania odpowiednich kryteriów podobieństwa oraz sposobów maksymalizacji tego podobieństwa. Autor proponuje zastosowania odpowiedniej kombinacji kilku miar podobieństwa, takich jak średnie kwadratowe odchylenie obu modeli po najlepszym nałożeniu czy odpowiednie średnie oscylacji atomów/pseudo-atomów obserwowane w symulacjach. Zadanie optymalizacji takiego podobieństwa jest bardzo trudne do uzyskania zadawalającego rozwiązania. Wynika to ze złożoności geometrii modeli i skomplikowanego schematu oddziaływań. W tej pracy użyto z powodzeniem metaheurystyki wykorzystujące różne schematy optymalizacji kombinatorycznej, takie jak algorytm ewolucyjny, optymalizacji roju cząsteczek, simplex i proste przeszukiwanie losowe. Doktorant opracował odpowiednie oprogramowanie dla schematu RedMDStream, którego opis i instrukcja użycia dołączone są do treści rozprawy (**Appendix A**). Rozdział **Methods** jest dobrze napisany, chociaż wydaje mi się, że dobrze by było w schematach takich jak ten dla algorytmu genetycznego (strona 55) zastąpić pojęcia z genetyki odpowiednimi pojęciami z biofizyki makromolekuł. Prezentowana metoda jest interesująca i wysoce oryginalna.

W rozdziale 3 rozprawy (Chapter 3, **Method Applied: Coarse-Grained Model of RNA Hairpin Dynamics**) pokazano praktyczne zastosowanie stworzonego przez doktoranta oprogramowania RedMDStream do symulacji małego fragmentu RNA (microROSE RNA thermometer). Wybrano bardzo prosty schemat gruboziarnistej reprezentacji łańcucha RNA, gdzie jeden pseudo-atom zastępuje całą resztę nukleotydową. Optymalizowano szereg schematów oddziaływań tych pseudo-atomów, mapujące (w różnych układach) oddziaływania bliskiego i dalekiego zasięgu, używane w polach siłowych pełnoatomowej

dynamiki molekularnej. Pokazano, że możliwe jest znalezienie kilku schematów (modeli) oddziaływań, które pozwalają na jakościowo poprawną symulację dynamiki modeli gruboziarnistych w otoczeniu założonej struktury (natywnej?). To doskonały wynik, pokazujący, że proponowana metoda pozwala na wybranie stosownej definicji modelu (nie dla wszystkich definicji oddziaływań uzyskano właściwe profile dynamiki) z zestawu różnych możliwości. Ten wynik ma jednak słabe strony. Główna to przypuszczalny brak możliwości przenoszenia generowanych modeli na inne, nawet podobne układy. Autor dobrze zdaje sobie sprawę z tego ograniczenia metody. Pokazały to zresztą obliczenia dla innych układów, krótko omówione na końcu trzeciego rozdziału rozprawy.

W podsumowaniu rozprawy (**Future Outlook**) autor wskazuje na główne zalety metody RedMDStream, zwracając też uwagę na fakt, że wprawdzie metoda ta jest w dużym stopniu zautomatyzowana, ale jej skuteczne zastosowanie wymaga jednak pewnej wiedzy eksperckiej. Dla bardziej złożonych układów odpowiedni dobór optymalizowanych schematów pola sił może być sprawą kluczową.

W załączniku do rozprawy (**Appendix A, RedMDStream Software Manual**) opisano szczegółowo oprogramowanie przygotowane przez autora rozprawy. Dostępne jest ono na stronie <http://bionano.cent.uw.edu.pl/RedMD>. Opis ten i instrukcje obsługi są dobrze napisane i zapewne pozwalają na stosunkowo łatwe użycie opisanych w pracy metod optymalizacji modeli gruboziarnistych. Nie przeprowadziłem takich testów, gdyż nie jest to prosty serwer internetowy. To, że autor nie próbował stworzyć serwera internetowego jest w pełni uzasadnione charakterem całego schematu obliczeniowego, który raczej nie może być w pełni sposób zautomatyzowany. Z drugiej strony, cała metodyka jest dobrze opisana. Osoby, które potrafią świadomie używać metod modelowania gruboziarnistego powinny też potrafić użyć opisanej w rozprawie metody, nawet bez udostępnionego oprogramowania. Oczywiście gotowe oprogramowanie prezentowane w rozprawie jest olbrzymim udogodnieniem dla projektowania niektórych dobrych schematów modelowania gruboziarnistego.

Oceniając całą rozprawę warto podkreślić, że wyniki uzyskane przez pana Filipa Leonarskiego w trakcie realizacji jego pracy doktorskiej były już przedmiotem oceny przynajmniej paru niezależnych ekspertów, jako że główne wyniki pracy zostały też opublikowane w postaci dwóch artykułów w renomowanych czasopismach (*Biophysical Journal* i *J. Chemical Theory and Computations*) oraz rozdziałów w anglojęzycznych monografiach. Doktorant jest pierwszym autorem tych czterech publikacji. O dobrym przygotowaniu do pracy naukowej świadczą też jego inne publikacje naukowe nie wchodzące w zakres rozprawy oraz prezentacje na konferencjach międzynarodowych.

W podsumowaniu pragnę stwierdzić, że jestem przekonany, iż przedstawiona mi do recenzji rozprawa doktorska pana magistra Filipa Leonarskiego spełnia wszystkie wymagania stawiane pracom doktorskim. Wnoszę zatem do Rady Wydziału Chemii Uniwersytetu Warszawskiego o dopuszczenie pana Leonarskiego do dalszych etapów przewodu doktorskiego.

Wniosuję też o nagrodzenie tej rozprawy stosownym wyróżnieniem. Szereg przesłanek skłania mnie do przedstawienia tego wniosku. Po pierwsze, zakres tematyczny badań będących treścią rozprawy jest doskonale wybrany i bardzo dobrze zrealizowany przez doktoranta. Modelowanie gruboziarniste układów biomolekularnych jest dziś intensywnie rozwijającym się kierunkiem badań o wielkim znaczeniu poznawczym i praktycznym. Realizacja postawionych celów badawczych wymagała szerokiej wiedzy z wielu dziedzin (chemia teoretyczna, biologia molekularna, bioinformatyka). Doktorant wykazał się taką

interdyscyplinarną wiedzą. Po drugie, uważam, że rozprawa napisana jest bardzo dobrze, uzyskane wyniki poprawnie zanalizowane, a także, co ważne, w większości już opublikowane w dobrych czasopismach naukowych. Po trzecie, jestem też przekonany, że badania prezentowane w rozprawie prezentują nowatorskie podejście do istotnych problemów modelowania molekularnego i co bardzo ważne, pokazują dalszy kierunek bardzo obiecujących badań.



Prof. dr hab. Andrzej Koliński



ul. Wita Stwosza 63, 80-308 Gdańsk
tel. 58 523 5124, fax 58 523 5012, email: adam@sun1.chem.univ.gda.pl

Gdańsk, 29 grudnia 2015

**Recenzja rozprawy doktorskiej
pana magistra Filipa Leonarskiego
pt. „Coarse-grained models for molecular dynamics of RNA. Evolutionary
algorithm optimization of the potential energy function parameters”**

Przewidywanie struktury oraz symulacje długoczasowej dynamiki cząsteczek kwasów rybonukleinowych, z których wiele ma charakter globularny, nabiera coraz większego znaczenia w biochemii i biofizyce z uwagi na różnorodne funkcje tych molekuł które, podobnie jak w przypadku białek, są związane z ich strukturą i właściwościami dynamicznymi. Podobnie jak w przypadku białek, przewidywanie struktury można zrealizować metodami bioinformatycznymi, głównie w oparciu o informacje ze strukturalnych baz danych, natomiast symulacje dynamiki należy przeprowadzać metodami używającymi potencjałów zdolnych do zlokalizowania struktury natywnej bez posługiwania się explicite informacjami z baz danych nawet, jeżeli są to potencjały statystyczne lub arbitralne. Ten wymóg pociąga za sobą również konieczność optymalizacji pola siłowego tak, aby było ono zdolne do lokalizowania struktur natiwnych. Ten proces zwany jest również trenowaniem pola siłowego. Na ogół używa się w tym celu układów modelowych.

Celem rozprawy doktorskiej p. mgra Filipa Leonarskiego było opracowanie metody optymalizacji gruboziarnistego pola siłowego do symulacji RNA i przetestowanie jej na przykładzie małej spinki zwanej „termometrem RNA”. Metoda była trenowana tylko na tym RNA i nie było celem Kandydata opracowanie pola przenaszalnego na inne cząsteczki RNA. Tym niemniej, postawione zadanie było bardzo ambitne a zaproponowane algorytmy są bardzo nowatorskie.

Rozprawa doktorska p. mgra Filipa Leonarskiego jest napisana w języku angielskim. O ile mogę to stwierdzić jako osoba, dla której ten język nie jest językiem ojczystym, tekst jest napisany poprawnie. Rozprawa składa się ze streszczenia (2 strony; streszczenie jest również podane w wersji polskiej, słowniczka skrótów (2 strony), organizacji rozprawy (1 strona), wstępu (32 strony), ogólnej części metodologicznej (33 strony), badań własnych (47 stron), która jest moim zdaniem niezbyt szczęśliwie zatytułowana „Methods applied: Coarse-Grained Model of RNA Hairpin Dynamics”, podsumowania i „spojrzenia w przyszłość” (2 strony), podręcznika programu RedMDStream (62 strony) i bibliografii (167 pozycji). W preambule rozprawy podano również wykaz publikacji i wystąpień konferencyjnych Kandydata. Ta lista obejmuje 3 publikacje w czasopismach z listy filadelfijskiej ściśle związane z rozprawą, których Kandydat jest pierwszym autorem, jeden rozdział w książce ściśle związany z rozprawą, którego pierwszym autorem jest Kandydat, 6 innych publikacji w czasopismach z listy filadelfijskiej 3 wystąpienia ustne i 10 prezentacji posterowych. W przypadku kandydata do stopnia doktora nauk chemicznych jest to zdecydowanie wyróżniający dorobek naukowy.

We wstępie Kandydat opisuje strukturę kwasów rybonukleinowych oraz ich rolę w biochemii komórki, nawiązując przy tym do koncepcji „świata RNA”. Ten podrozdział stanowi dobre wprowadzenie do tematu nawet dla czytelnika słabo zorientowanego w problematyce kwasów rybonukleinowych. Następne podrozdziały są poświęcone opisowi hiperpowierzchni energii potencjalnej molekuł i układów molekuł, począwszy od mechaniki kwantowej w wydaniu rozwiązań stacjonarnego równania Schrödingera poprzez pola siłowe w reprezentacjach pełnoatomowych do pól gruboziarnistych takich, jak używane przez Kandydata w recenzowanej rozprawie doktorskiej. W tej części rozprawy jest zawarty jednocentrowy model gruboziarnisty kwasów nukleinowych i sprzężone z nim pole siłowe które, w wersji zmodyfikowanej, Kandydat używa w swojej pracy. Autor omawia również sposoby wyprowadzenia i parametryzacji pól siłowych dla innych biopolimerów. Na przykładzie opartego na fizyce modelu UNRES białek podaje definicję efektywnej funkcji energii jako potencjału średniej siły. Następnie Kandydat omawia model CABS białek opracowany przez grupę prof. Kolińskiego, który jest oparty o potencjały statystyczne. W przedostatnim rozdziale wstępu Kandydat wspomina o parametryzacji potencjałów gruboziarnistym metodą iteracyjnego odwracania funkcji korelacji a w ostatnim definiuje cel pracy. W ostatnim podrozdziale podaje i omawia cel pracy. Ta część rozprawy jest napisana dobrze, jakkolwiek mam do niej parę uwag krytycznych podanych poniżej.

Na str. 12 i 13 poświęconych mechanice kwantowej należałoby wspomnieć o metodach używanych do opisu układów otwartopowłokowych (np. CI czy MCSCF) oraz o metodzie sprzężonych klastrów która jest, podobnie jak MP2-MP5, używana do obliczania energii korelacji elektronowej dla układów zamkniętopowłokowych.

W podrozdziale „Coarse-grained model developed by averaging over secondary degrees of freedom” należałoby wspomnieć o sposobach wyprowadzenia wyrażenia na wkłady do efektywnej energii (której prototypem jest potencjał średniej siły zdefiniowany równaniem 1.20) poprzez kumulanty klastrów jak w modelu UNRES [A. Liwo, C. Czaplewski, J. Pillardy, H.A. Scheraga, J. Chem. Phys., 115, 2323 (2001)] lub metodę dopasowania sił opracowaną przez grupę G. Votha [zob. np. G.S. Ayton, W.G. Noid, G.A. Voth, 32, 929 (2007)]. Te dwa podejścia mają istotną zaletę w stosunku do podejść „neoklasycznych”, gdzie wyrażenia „importuje się” z pól pełnoatomowych, lub potencjałów statystycznych. Tymczasem lektura tego podrozdziału stwarza wrażenie, że formy funkcyjne wkładów do energii w polach gruboziarnistych są jakościowo podobne do tych w polach pełnoatomowych. Tymczasem różnica jest zasadnicza; skutek obecności „ukrytych” stopni swobody pojawiają się tutaj wkłady wielociałowe, których formy funkcyjne można znaleźć dopiero po wnikliwej analizie rozwinięcia potencjałów średniej siły. Ponadto z odsyłaczy [89-91] podanych jako odsyłacze do modelu UNRES odsyłacz 90 nie opisuje modelu UNRES lecz kilkucentrowy gruboziarnisty model kwasów nukleinowych opracowany w grupie Scheragi, odsyłacz 89 nie jest odsyłaczem do modelu UNRES ale do metody optymalizacji pola UNRES a tylko odsyłacz 91 można uznać za odsyłacz do modelu UNRES, jakkolwiek opisuje on jego uogólnienie na białka, kwasy nukleinowe i polisacharydy.

Omawiając gruboziarniste modele kwasów nukleinowych Kandydat ogranicza się do modelu jednocentrowego. Jak sam stwierdza na stronie 19, symulacje przy użyciu tego modelu wymagają podania struktury drugorzędowej (rozumianej jako lista par zasad nukleinowych). Należałoby wspomnieć o modelach kwasów nukleinowych, które potrafią symulować hybrydyzację DNA i zwijanie RNA w trybie ab initio czyli bez podawania informacji o strukturze drugorzędowej. Przykładami są model opracowany w grupie Ouldridge’a [T.E. Ouldridge, A.A. Louis, J.P.K. Doye, Phys. Rev. Lett. 104, 178101 (2010)], model opracowany przez grupę Derreumaux [T. Cragolini, P. Derreumaux, S. Pasquali, J. Phys. Chem. B, 117, 8047 (2013)], dwucentrowy model NARES-2P [rozszerzenie modelu

UNRES na kwasy nukleinowe; Y. He, M. Maciejczyk, S. Ołdziej, H.A. Scheraga, A. Liwo, Phys. Rev. Lett. 110, 098101 (2013)] czy bardziej złożony model Maciejczyka [M. Maciejczyk, A. Spasic, A. Liwo, H.A. Scheraga, J. Chem. Theory Comput., 10, 5020 (2014)].

W rozdziale 2, Kandydat opisuje metodę dynamiki molekularnej, w tym algorytmy całkowania, obliczanie wielkości statystyczno-mechanicznych z symulacji, termostat Berendsena oraz dynamikę Langevina i dynamikę brownowską. Następnie omawia stosowane przez siebie metody optymalizacji pól siłowych, najpierw podając definicję funkcji celu (które są oparte na średnim odchyleniu obliczonych struktur od struktury eksperymentalnej, zgodności wariancji odchylenia oraz zgodności rozkładów odległości i innych wielkości geometrycznych z tymi wyznaczonymi eksperymentalnie lub z symulacji pełnoatomowych), a później algorytmy optymalizacji: algorytm genetyczny, metodę „roju cząstek” i metodę sympleksów. Słusznie zauważa, że ponieważ obliczenie funkcji celu wymaga przeprowadzenia symulacji z danym zestawem parametrów, optymalizacja gradientowa jest tu trudna do zastosowania. W tej części brakuje odniesienia do istniejących algorytmów optymalizacji funkcji energii, z których pierwsze, oparte na maksymalizacji przerwy energetycznej między strukturą natywną a strukturą nienatywną lub optymalizacji „Z-score”, zostały opracowane odpowiednio przez grupę Crippena [G.M. Crippen, M.E. Snow, Biopolymers 29, 1479 (1990)] i grupę Wolynesa [R.A. Goldstein, Z.A. Luthey-Schulten, P.G. Wolynes, Proc. Natl. Acad. Sci. U.S.A., 89, 9029 (1992)] już we wczesnych latach 90-tych ubiegłego wieku tym bardziej, że jak wspomniałem wcześniej, Kandydat, przy okazji omawiania pola UNRES, cytuje odsyłacz 89, gdzie jest omówiona metoda optymalizacji hierarchicznej, będąca rozszerzeniem podejść Crippena i Wolynesa na pośrednie stany na ścieżce zwijania białka. Wymienione metody polegają na cyklach optymalizacji funkcji celu ze zbiorem konformacji uzyskanych z symulacji ze starymi parametrami i symulacji z nowymi parametrami, co umożliwia zastosowanie optymalizacji gradientowej w każdym cyklu.

W rozdziale 3 Kandydat opisuje wyniki swoich badań nad optymalizacją pola siłowego dla jednocentrowego modelu gruboziarnistego spinki RNA złożonej z 29 nukleotydów. Stosowane pole ma charakter neoklasyczny. Parowanie zasad nukleinowych jest symulowane poprzez potencjały harmoniczne (zamiast potencjałów Morse'a jak w oryginalnym modelu) nałożone na odpowiednie centra oraz centra przyległe (te drugie zapewne po to, żeby zapewnić odpowiednią kierunkowość oddziaływań); występują również jawne wkłady opisujące energetykę pętli. W tym miejscu nasuwa się pytanie, czy potencjały harmoniczne były specyficzne tylko dla par typów zasad nukleinowych, niezależnie od tego czy dane dwie reszty tworzą w strukturze natywnej parę czy nie, czy też potencjały harmoniczne nakłada się tylko na reszty tworzące pary w strukturze natywnej. W pierwszym przypadku należałoby wprowadzić funkcję obcięcia, włączającą dany potencjał harmoniczny tylko jeżeli reszty znajdują się blisko siebie, ponieważ włączenie jego dla wszystkich potencjalnych par powodowałoby „zapadanie się” struktury. W drugim przypadku pole może być co prawda przenaszalne na inne spinki ale może służyć jedynie do symulacji w otoczeniu struktury natywnej. Mam nadzieję, że Kandydat wyjaśni tę wątpliwość w odpowiedzi na recenzję.

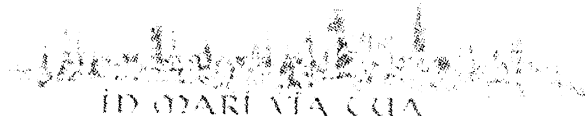
Na str. 83 Kandydat pisze, że jedynym potencjałem niewiążącym był potencjał kulombowski dany równaniem 3.4 i że wprowadzenie innych potencjałów (Lennarda-Jonesa?) nie poprawiało wyników. Z tym stwierdzeniem mogę się zgodzić ponieważ jest to jedyny potencjał z osobliwością dla zerowej odległości. Mam jednak wątpliwości co do ładunków w równaniu 3.4. Ładunek całego nukleotydu wynosi zawsze -1 w jednostkach ładunku elektronu chyba, że uwzględnimy ekranowanie. Czy zatem ładunki były równe -1 czy też były efektywnymi ładunkami ekranowanymi; jeżeli tak, to jak je określono?

Do generowania struktur odniesienia Kandydat używał pola AMBER ff14SB, które dobrze oddaje strukturę badanej spinki, jakkolwiek przejście od struktury zwiniętej do rozwiniętej jest skokowe nie tak, jak obserwowane eksperymentalnie, a temperatura mięknięcia jest znacznie zawyżona. Kandydat przetestował wszystkie 3 metody optymalizacji: algorytm genetyczny, metodę „roju cząstek” oraz metodę sympleksów. Dla porównania prowadził również przypadkową selekcję parametrów. Dwie pierwsze metody okazały się bardzo wydajne w znajdowaniu optymalnego rozwiązania, natomiast zachowanie metody sympleksów niewiele odbiegało od przypadkowej selekcji. Takie wyniki nie jest zaskoczeniem ponieważ metoda sympleksów jest z definicji lokalna a zatem nie jest w stanie szybko odejść od punktu startowego. Natomiast zarówno algorytm genetyczny jak i metoda „roju cząstek” operują na populacjach rozwiązań, co umożliwia wybranie najlepszego z nich. Kandydat przetestował optymalizację 10 wariantów pola siłowego, z czego najlepsze okazały się dwa, w których potencjały harmoniczne są nałożone nie tylko na pary zasad $i:j$ ale również na przyległe $i:j+1$ oraz $i+1:j+1$. Kandydat przeprowadził bardzo rzetelną i wyczerpującą dyskusję wyników optymalizacji oraz jakości otrzymanych pól siłowych.

Na podstawie moich doświadczeń w optymalizacji gruboziarnistych pól siłowych mogę stwierdzić, że Kandydat wykazał ponadprzeciętną inwencję i pomysłowość w projektowaniu metod optymalizacji i wykonał ogromną pracę nad jej przeprowadzeniem, co nie jest umniejszone faktem, że obiektem była tylko jedna spinka RNA. Nad optymalizacją pól siłowych pracuje zwykle latami sztab ludzi, dochodząc do jakichkolwiek wyników dopiero po latach prób i błędów. Kandydat natomiast wykonał takie zadanie samodzielnie. Opracowana przezeń metoda optymalizacji pól siłowych znacznie odbiega od stosowanych dotychczas i będzie z pewnością nie tylko rozwijana przez Kandydata ale zostanie podchwycona przez innych badaczy.

Podsumowując, rozprawę doktorską p. mgra Filipa Leonarskiego uważam za wyróżniającą. Założony cel został w pełni zrealizowany. Wymienione powyżej uwagi krytyczne, z których większość ma charakter dyskusyjny, nie umniejszają w żadnym stopniu jej wartości. Autor opracował i przetestował na przykładzie gruboziarnistego modelu spinki „termometru” RNA oryginalną metodę optymalizacji pól siłowych, która ma bardzo duże szanse na zastosowanie praktyczne. Metodę tę zakodował w postaci programu komputerowego, który może być wykorzystany także do optymalizacji innych pól siłowych. Rozprawa spełnia z należytą należytą wymagania stawiane rozprawom doktorskim przez Ustawę z dnia 14 marca 2003 roku o stopniach naukowych i tytułach naukowych oraz o stopniach i tytułach w zakresie sztuki (Dz. U. nr 65, poz. 595 z późn. zm.), jak również zwyczajowe standardy stawiane rozprawom doktorskim w dziedzinie nauk przyrodniczych i ścisłych. Dlatego z pełnym przekonaniem zwracam się do Wysokiej Rady Wydziału Chemii Uniwersytetu Warszawskiego o dopuszczenie p. mgra Filipa Leonarskiego do dalszych etapów przewodu doktorskiego. Biorąc pod uwagę zaproponowanie nowatorskiej metody optymalizacji pól siłowych, ogrom i wysoki profesjonalizm wykonanej pracy, znakomitą jakość wyników oraz ich znaczenie dla modelowania RNA wnoszę o wyróżnienie rozprawy. Szczegółowe uzasadnienie wniosku o wyróżnienie stanowi załącznik do niniejszej recenzji.


prof. dr hab. Józef Adam Liwo



ul. Wita Stwosza 63, 80-308 Gdańsk
tel. 58 523 5124, fax 58 523 5012, email: adam@sun1.chem.univ.gda.pl

Gdańsk, 29 grudnia 2015

**Uzasadnienie wniosku o wyróżnienie rozprawy doktorskiej
pana magistra Filipa Leonarskiego
pt. „Coarse-grained models for molecular dynamics of RNA. Evolutionary
algorithm optimization of the potential energy function parameters”**

W ramach rozprawy doktorskiej Pan magister Filip Leonarski opracował nowatorską metodę optymalizacji funkcji energii do obliczeń gruboziarnistych kwasów nukleinowych i udowodnił jej poprawność używając do optymalizacji jednocentrowego pola gruboziarnistego do symulacji spinki RNA. Funkcja celu w tej metodzie składa się z dopasowania do struktury eksperymentalnej, jak również dopasowania fluktuacji odchyień od struktury eksperymentalnej oraz rozkładów odległości i innych wielkości geometrycznych do tych otrzymanych z symulacji pełnoatomowych. Do optymalizacji zastosował algorytm genetyczny, metodę „roju cząstek” oraz metodę sympleksów; algorytmy te zostały przezeń dostosowane do rozpatrywanego zagadnienia. W każdym kroku funkcja celu jest obliczana z symulacji z danymi parametrami pola. Podejście stosowane przez Kandydata nie było dotąd stosowane do optymalizacji pól siłowych mimo, że sam problem ma co najmniej 25-letnią historię. Opracowana metoda może nadać nowy kierunek pracom nad optymalizacją pól siłowych.

O wysokiej jakości pracy Kandydata świadczy Jego dotychczasowy dorobek naukowy, w którym zawierają się 3 prace oryginalne w bardzo dobrych czasopismach z listy filadelfijskiej oraz 1 rozdział w książce na temat materiału samej rozprawy (we wszystkich Kandydat jest pierwszym autorem), jak również 6 innych publikacji w czasopismach z listy filadelfijskiej. Wykonał samodzielnie ogromną pracę podczas, gdy optymalizację danego pola siłowego na ogół realizuje sztab ludzi.

W świetle przedstawionych wywodów a także argumentów przytoczonych w mojej recenzji rozprawy doktorskiej p. mgra Filipa Leonarskiego, z pełnym przekonaniem wnoszę do Wysokiej Rady Wydziału Chemii Uniwersytetu Warszawskiego o wyróżnienie jego rozprawy doktorskiej.

prof. dr hab. Józef Adam Liwo