



Warszawa, dnia 29.05.2013

Recenzja rozprawy doktorskiej pana magistra Adama K. Sieradzana pt. „Zwiększenie rozdzielczości i zdolności predykcyjnych pola siłowego UNRES do symulacji białek poprzez udoskonalenie potencjałów opisujących oddziaływania z udziałem łańcucha głównego”

Rozprawa doktorska pana magistra Adama Sieradzana, której promotorem jest profesor Adam Liwo, kierownik Zakładu Modelowania Molekularnego Wydziału Chemii Uniwersytetu Gdańskiego, została przedstawiona Radzie Naukowej Wydziału Chemii UG.

Przedstawiona do recenzji rozprawa składa się z około 110 stron tekstu w języku polskim, do którego dołączone są kopie trzech prac opublikowanych w czasopismach naukowych. Dołączony jest też dysk CD zawierający teksty dwóch dodatkowych prac (wysłanej do druku i przygotowywanej do druku) oraz zbiory danych i programów komputerowych potrzebnych do zastosowania zaproponowanych modyfikacji metody UNRES. Układ treści rozprawy nie odbiega od zwykle stosowanego w rozprawach doktorskich z zakresu chemii teoretycznej, czy pokrewnych dziedzin ścisłych. Cytowana literatura jest dobrze wybrana (157 pozycji).

Cel prezentowanej pracy doktorskiej jest krótko przedstawiony na początku rozprawy i dotyczy rozszerzenia pola siłowego UNRES (United Residues). Pole siłowe UNRES rozwijane jest od kilkunastu lat przez profesora Adam Liwo i jego współpracowników jako metoda wydajnego komputerowego modelowania struktury, oddziaływań i dynamiki białek. UNRES zakłada gruboziarnistą reprezentację łańcuchów białek i oparty jest na możliwie ścisłym fizycznie tłumaczeniu klasycznych definicji oddziaływań atomowych na model oddziaływań pomiędzy zjednoczonymi atomami zdefiniowanymi w założonej gruboziarnistej reprezentacji białek. Odróżnia to zasadniczo model UNRES od wielu innych modeli gruboziarnistych wykorzystujących pseudo-potencjały wyprowadzone na podstawie statystycznych analiz regularności obserwowanych w poznanych doświadczalnie strukturach białek. Oba podejścia mają określone zalety. Z jednej strony potencjały typu UNRES oparte są na ścisłej wiedzy fizycznej co powinno umożliwiać dokładne i łatwe do interpretacji wyniki modelowania. Z drugiej strony, przynajmniej do tej pory, niektóre modele oparte na potencjałach statystycznych wydają się bardziej dokładnie odtwarzać struktury przestrzenne prostych białek. Szkoda, że w rozprawie nie przedyskutowano bardziej szczegółowo tych istotnie różniących się podejść do modelowania molekularnego złożonych układów.

Opracowane w ramach przedstawionej pracy doktorskiej modyfikacje pola siłowego UNRES są dwójakiej natury. Po pierwsze, wyprowadzono dodatkowe potencjały mające na celu rozszerzenie zakresu stosowalności pola siłowego dla mniej typowych struktur, uwzględniając izomeryzację cis-trans w białkach zawierających prolinę oraz białka zawierające D-aminokwasy. Istotnie innym celem pracy było „zwiększenie zdolności predykcyjnych” pola siłowego poprzez wprowadzenie potencjałów korelacyjnych łączących wewnętrzne konformacje łańcuchów bocznych. Potencjały te wyprowadzono na podstawie analizy statystycznej znanych struktur. Moim zdaniem to dość zasadnicza modyfikacja całego dotychczasowego podejścia do metody UNRES. Rezultaty pierwszych testów (które też były celem tej rozprawy) wyglądają na bardzo obiecujące.



Część teoretyczna rozprawy zawiera się na 25 stronach. Ta część jest dość lakoniczna, ale prezentuje najważniejsze podstawy teoretyczne potrzebne do opisu prac programistycznych i obliczeniowych prezentowanych w rozprawie. Wyjątkowo krótka (trochę ponad jedną stronę) i raczej słaba jak na tematykę rozprawy jest pierwsza część podrozdziału „Zwijanie białek”. W zasadzie tylko wymienione lub pobieżnie zdefiniowane są niektóre teorie oraz techniki doświadczalne. Jako pierwszy wymienia się model, który (cytuje) „zakłada, że struktura drugorzędowa powstaje znacznie szybciej niż struktura trzeciorzędowa”. To zdanie jest nieco mylące, gdyż może sugerować, że struktura trzeciorzędowa powstaje z gotowej struktury drugorzędowej. Rzeczywiście niektóre, ale zdecydowanie nie wszystkie, elementy struktury drugorzędowej mogą powstawać i zanikać wcześniej, ale kompletna struktura drugorzędowa istnieje dopiero po zwinięciu się białka do struktury trzeciorzędowej, a czasem dopiero po utworzeniu struktury czwartorzędowej. W tym sensie struktura drugorzędowa powstaje razem ze strukturą wyższego rzędu. Są jednak wyjątki od tej reguły. Żadna z istniejących teorii dotyczących mechanizmów zwijania się białek nie została omówiono w zadawalający sposób.

Dwa następne podrozdziały omawiają pokrótce izomeryzacje cis-trans kontrolujące proces zwijania niektórych białek oraz wpływ D-aminokwasów na struktury drugorzędowe i trzeciorzędowe.

Bardziej obszernie i starannie napisane są podrozdziały opisujące różne poziomy opisu układów molekularnych i metody symulacji komputerowej. Dobrze opisane są teoretyczne podstawy mechaniki molekularnej i ich wykorzystanie w symulacjach na poziomie atomowym oraz w gruboziarnistych modelach zjednoczonych atomów. Starannie opisany jest model UNRES i dotychczas stosowane pole siłowe tego modelu. W części dotyczącej obliczeniowego przeszukiwania przestrzeni konformacyjnej omówiono metodę Monte Carlo, dynamikę molekularną (klasyczną i z wymianą replik). Dobrze opisane są sposoby symulacji i analizy wyników uzyskiwanych dla modeli gruboziarnistych. Zauważyłem tylko parę nieścisłości. Oto dwa przykłady. W podrozdziale „Metody Monte Carlo” opisany jest schemat Metropolisia, nie wspominając, że omawiany jest tylko schemat asymetryczny. Podrozdział „Dynamika molekularna” autor zaczyna stwierdzeniem „W metodzie dynamiki molekularnej, w przeciwieństwie do metody Monte Carlo, ścieżka zwijania jest wyznaczona w sposób deterministyczny”. To nie do końca prawda. Po pierwsze, różne wersje dynamiki molekularnej generują różne (czasem bardzo różne) ścieżki zwijania pomimo identycznych konformacji startowych i identycznej liczby iteracji. Po drugie, czasem wystarczy zmienić typ komputera (jeżeli zmieni to sposób przybliżania liczb) aby ten sam program dynamiki molekularnej produkował inne trajektorie. W opisie metody replik zabrakło mi paru zdań na temat różnych sposobów wyboru temperatury dla poszczególnych replik. Wydajność obliczeniowa tej metody bardzo zależy od wyboru rozkładu temperatur.

Zabrakło mi dyskusji na temat jednego (być może najciekawszego) z wprowadzanych w ramach tej pracy rozszerzeń/modyfikacji pola siłowego UNRES. Zastosowanie w nowej wersji UNRES potencjałów statystycznych uwzględniających korelacje kątów torsyjnych łańcuchów bocznych z kątami torsyjnymi łańcucha głównego. Moim zdaniem to zasadnicza modyfikacja modelu, gdyż dodane składowe oparte są na parametrach statystycznych, wyprowadzonych ze znanych struktur eksperymentalnych. Dlaczego dokonano takiej modyfikacji pola UNRES? Takie zmodyfikowane pole nie jest już oparte na podstawowych prawach fizyki, ale uwzględnia ich końcowy efekt, bez ścisłej analizy oddziaływań atomowych. To zbliża pole UNRES do modeli statystycznych. Jakie są oczekiwania związane z nową wersją? Jak oryginalne pole siłowe UNRES zmodyfikowane potencjałem



statystycznym wpływa na zależności temperaturowe modelowanych układów? Czy planowane są dalsze tego typu modyfikacje pola UNRES? A może warto by sprawdzić jak reprezentacja gruboziarnista UNRES mogłaby wykorzystywać inne pola siłowe, oparte wyłącznie na potencjałach statystycznych?

Wyniki pracy przedstawione są na 75 stronach rozprawy. To główna i najlepiej napisana część rozprawy.

Jako pierwsze opisane jest wyprowadzenie map energii potencjalnej dla izomeryzacji trans-cis w białkach zawierających prolinę. Metodą ab initio w schemacie MP2 z bazą funkcyjną 6-13G(d,p) wyznaczono mapy energii potencjalnej i minima odpowiadające konformacjom cis i trans dla N-metyloacetamidu i N-acetylopirolidyny. Pozwoliło to wyznaczyć potencjały średniej siły izomeryzacji trans-cis grupy peptydowej, następnie dopasowanych do funkcji analitycznych potencjału wykorzystywanego przez model UNRES. Symulacje testujące nową wersję pola siłowego przeprowadzona dla dwupeptydów Gly-Gly i Gly-Pro.

Kolejnym rozszerzeniem pola siłowego UNRES zrealizowanym przez autora rozprawy było uwzględnienie D-aminokwasów. Wymagało to wyprowadzenia wszystkich składowych pola UNRES dla układów zawierających D-aminokwasy połączone z różnymi typami L-aminokwasów. Pierwsze testy nowego potencjału przeprowadzono dla peptydu 1AL4 składającego się naprzemiennie z L i D aminokwasów. Struktura o najniższej energii uzyskana w trakcie symulacji była zbliżona do struktury eksperymentalnej, mającej formę heterodimeru w formie podwójnej helisy. Podobnie udane były symulacje dla gramicydyny D, której sekwencje nieco zmodyfikowano, zastępując D-allo-treoninę bardzo podobną do niej D-treoniną. Również w tych symulacjach uzyskano struktury podobne do doświadczalnych. Najbardziej złożonym testem nowego pola siłowego były symulacje zawierające D-prolinę peptydu BBAT1, który przyjmuje formę tetrameru. Pojedynczy łańcuch w tym tetramerze ma złożoną formę $\beta\beta\alpha$ zwoju. Symulacje pokazały wiarygodny mechanizm tworzenia peptydu o strukturze bardzo zbliżonej do wyznaczonej doświadczalnie.

Ostatnia modyfikacja UNRES przedstawiona w tej rozprawie nie dotyczy już rozszerzenia opisu pola siłowego o nowe typy białek (o różnych formach cis-trans i D-L składowych aminokwasów), ale rozszerza definicję istniejącego wcześniej pola siłowego. W nowej wersji pola siłowego uwzględniono sprzężenie stanów konformacyjnych łańcuchów bocznych z lokalnymi konformacjami łańcucha głównego. Korelacje te wyprowadzono na podstawie analizy statystycznej znanych struktur białek o wysokiej rozdzielczości. Stwierdzono, że taka modyfikacja pola UNRES poprawia rozdzielczość.

Podsumowanie wyników (strony 97-99) jest bardzo dobrze napisane, wymieniając najważniejsze rezultaty osiągnięte w trakcie realizacji pracy, ale pomijając mniej istotne szczegóły.

Ocena uzyskanych wyników i ich prezentacji w treści rozprawy. Poniżej ustosunkuję się do całego projektu rozprawy, opracowanej modyfikacji modelu UNRES, przeprowadzonych obliczeń i ich analizy. Oddzielnie ocenię język rozprawy.

Język rozprawy jest bardzo niestaranny. Zbyt dużo jest niezręcznych sformułowań oraz dziwnej składni. Oto kilka przykładów, a wymieniam tylko niektóre. Strona 13: „Metoda ta wymaga całkowania numerycznego w celu obliczenia uśrednionego hamiltonianu, która jest niezwykle kosztowna.” Strona 14/15: „badanie ścieżki zwijania części szczytowej wiliny”. Co tu jest częścią szczytową? Czy jest to tłumaczenie na język polski terminu „headpiece subdomain”? Strona 15: „większość model zorientowanych”. Strona 19: „Potencjały średniej siły są związane z gruboziarnistą funkcją energii”. Strona 2: „akceptacja



kolejnej konformacji jest zaakceptowana”. Strona 27: „Ilustracja skal czasowych poszczególnych ruchów: wibracji wiązania, obrotowych łańcuchów bocznych,...” czyta się tak jakby chodziło o „obrotowe łańcuchy boczne”. Strona 69: „w każdej temperaturze prowadziłem 4 trajektorie”. Strona 75: „Pierwszym pełnym zastosowaniem zmodyfikowane pola UNRES” – nie wiadomo jak rozumieć to „pełne zastosowanie”. Inne trudne do zrozumienia zdanie na tej samej stronie: „W przypadku tetrameru poszczególne monomery były losowo rozmieszczone w przestrzeni w odległości co najmniej rozciągniętego łańcucha monomeru”. Też na tej stronie pojawia się „C-terminalna reszta łańcucha”. Pewnie chodzi o C-koniec łańcucha? Strona 78: „asocjacja czterech łańcuchów do tetrameru”, a w następnej linii „Dla wysokich temperatur dominuje forma monomeru”. Strona 80: „Następnym etapem jest kolaps struktur do struktury „pęczka””. Strona 81: „Proces ten rozpoczyna się gdy około 90% molekuł jest w stanie tetrameru”. Rozdział „Podsumowanie i wnioski” zaczyna zdanie „W toku mojej pracy doktorskiej zrealizowałem zadania badawcze, które wraz z wyciągniętymi z przeprowadzonych przeze mnie badań wnioskami są podsumowane poniżej”. Zdanie to jest w zasadzie poprawne, ale jednak można to było napisać trochę ładniej i prościej. Pomimo licznych błędów i niezręczności językowych prezentowane cele rozprawy i opis uzyskanych wyników są zrozumiałe.

W tym miejscu pozwolę sobie na uwagę tylko pośrednio związaną z tą rozprawą. Według nowego trybu przyznawania stopni doktorskich przedstawiana rozprawa może składać się z paru współautorskich publikacji naukowych doktoranta i krótkiego wstępu. Wydaje mi się, że magister Sieradzan mógłby spełniać wszystkie wymagania nowego trybu, choć być może z koniecznością niewielkiej zmiany tytułu pracy. Wtedy nie byłoby moich krytycznych uwag na temat języka rozprawy. Publikacje pana magistra w czasopismach naukowych napisane są w dobrym języku angielskim.

Moja ocena wyników pracy doktorskiej magistra Sieradzana jest bardzo wysoka. Autor rozprawy dokonał znaczącego rozszerzenia pola siłowego UNRES, umożliwiając jego wydajne stosowanie do modelowania dynamiki i struktur peptydów i białek zawierających nietypowe formy aminokwasów (dwie formy proliny oraz D-aminokwasów). Atrakcyjność pola siłowego UNRES i jego wykorzystanie w różnych technikach symulacyjnych była wielokrotnie demonstrowana w badaniach dynamiki i procesów związania białek. UNRES jest bardzo dobrze rozpoznawany w literaturze światowej. Niewątpliwie rozszerzenie modelu wykonane w ramach omawianej pracy doktorskiej umożliwi badanie nowych zagadnień biofizyki białek.

Modyfikacja pola siłowego UNRES poprzez uwzględnienie korelacji konformacji łańcuchów bocznych z najbliższymi fragmentami łańcucha głównego jest nowatorskim pomysłem w rozwoju metody. Korelacje te wyprowadzone zostały na podstawie statystycznej analizy znanych struktur. Niewątpliwie, statystyczne pola siłowe dobrze sprawdzają się w kilku często używanych gruboziarnistych metodach modelowania struktur białek. Ciekawe jest jak sprawdza się ta modyfikacja UNRES. Przedstawione w rozprawie testy należy raczej uznać za fragmentaryczne. Autor zauważa, że „rozdzielczość UNRES” poprawia się średnio o 0.5Å. Nie bardzo rozumiem co oznacza tu słowo „rozdzielczość”. Dużo bardziej interesująca byłaby ocena dokładności przewidywań struktur trzeciorzędowych (i czwartorzędowych). Autor rozprawy zapowiada wykonanie takich testów przez innego doktoranta. Szkoda, że wyniki takich testów nie pojawiły się w tej rozprawie. Nie może to jednak być zarzutem dla autora recenzowanej rozprawy – i tak zakres wykonanych przez niego prac jest imponujący.



Oceniając rozprawę pana magistra Sieradzana warto zwrócić uwagę na to, że znaczna część wyników pracy została już opublikowana w formie trzech doskonałych artykułów naukowych w renomowanym Journal of Chemical Theory and Computation. IF (Impact Factor) tego czasopisma wynosi 5.215. Oczywiście wymienione publikacje, a więc także wyniki uzyskane przez magistra Sieradzana, zostały pozytywnie ocenione przez wymagający zespół recenzentów. Pan Adam Sieradzan jest pierwszym autorem tych prac. Czwarta praca została wysłana do tego samego czasopisma, a piąta jest w przygotowaniu.

Podsumowując powyższe, uważam, że rozprawa doktorska pana magistra Adama Sieradzana spełnia wszelkie zwyczajowe i ustawowe wymagania stawiane pracom doktorskim. Wnoszę zatem do Rady Naukowej Wydziału Chemii Uniwersytetu Gdańskiego o dopuszczenie doktoranta do dalszych etapów przewodu doktorskiego.

Ponadto, biorąc pod uwagę olbrzymie znaczenie uzyskanych przez doktoranta wyników, polegających na opracowaniu bardzo użytecznych programów obliczeniowych istotnie rozszerzających zakres stosowalności metody UNRES do analizy i przewidywania struktury białek, wnioskuję do Rady Naukowej Wydziału Chemii Uniwersytetu Gdańskiego o wyróżnienie rozprawy doktorskiej pana magistra Adama Sieradzana.

Prof. dr hab. Andrzej Koliński