

Recenzja rozprawy doktorskiej mgra **Pawła Krupy**

pt: „*Rozszerzenie pola siłowego UNRES o potencjały lokalne i dane z metod porównawczych w celu lepszego przewidywania struktur białek*”

wykonana w Katedrze Chemii Teoretycznej w Pracowni Symulacji Polimerów na Wydziale Chemii Uniwersytetu Gdańskiego pod kierunkiem dr hab. Cezarego Czaplewskiego prof UG.

Przedstawiona mi do oceny praca doktorska dotyczy zagadnienia przewidywania struktur białkowych z wykorzystaniem pakietu o nazwie UNRES.

Praca składa się z teoretycznego wstępu omawiającego zagadnienie przewidywania struktur białek. W tej części pracy Autor omawia stosowane w chemii teoretycznej metody obliczeniowe w tym metody *ab initio*, półempiryczne i empiryczne. Prezentuje zasady obowiązujące w mechanice molekularnej a także techniki symulacji dynamiki molekularnej. W części końcowej skupia się na sposobach przewidywania struktur białek wyróżniając przewidywanie oparte na analizie porównawczej oraz metody typu *de novo*. Techniki te prezentuje w kontekście eksperymentu o nazwie CASP wymieniając wszystkie edycje konkursu, gdzie rozwijany przez mgra Pawła Krupę model pozwalał na uzyskanie poprawnych wyników przewidywania. Autor wspomina też inicjatywę o nazwie WeFold, w której to inicjatywie Doktorant brał udział wraz z zespołem.

Część „Wyniki” składa się z opisu czterech publikacji, których Doktorant jest współautorem. Publikacje te prezentują zasadniczy dorobek mgra Pawła Krupy. Pace te zostały w całości (w formie kopii) załączone do tekstu pracy doktorskiej.

Łącznie część opisowa pracy doktorskiej zawarta jest na 81 stronach.

Autor dołącza też spis publikacji z rozróżnieniem pozycji z listy Filadelfijskiej (w liczbie 3), recenzowanych publikacji międzynarodowych i książek (rozumiem, że pod tym hasłem znajduje się rozdział w książce) wydane przez wydawców międzynarodowych (w liczbie 3. pozycji) oraz recenzowane publikacje w czasopiśmie polskich (w liczbie 5.).

Lista rysunków zawartych w rozprawie w liczbie 18 poprzedza spis cytowanej literatury fachowej w liczbie 276 pozycji.

W części końcowej dostarczonej mi rozprawy zamieszczone są załączniki w liczbie 4, z czego trzy pierwsze są kopiami publikacji, których Doktorant jest współautorem oraz jeden, który jest manuskryptem pracy przekazanej do publikacji w *Journal of Chemical Information and Modeling*.

Przedstawiona mi do oceny rozprawa doktorska dotyczy rozwijanego w ośrodku Gdańskim modelu o nazwie UNRES (UNited RESidues) bazującego na pierwotnym modelu o nazwie ECEPP. Prace realizowane są w ścisłej współpracy z ośrodkiem w Cornell University kierowanym przez profesora Harolda A. Scheragę.

Zaprezentowany i rozwijany w pracy model wprowadzający pseudo-pole ma na celu uproszczenie zapisu struktury zwłaszcza w początkowym etapie poszukiwania właściwej dla danej sekwencji aminokwasów konformacji. Uproszczenie to wykorzystuje szeroko rozumiane pojęcie atomów „zjednoczonych” uśredniając fizyczne atomy i wprowadzając w ich miejsce elementy reprezentujące wybrane ugrupowania atomów. W ślad za geometrycznym uproszczeniem struktury wprowadzane jest też pseudo-pole siłowe wyrażające oddziaływania wprowadzonych wcześniej „zjednoczonych” atomów.

Zakres prac mgra Pawła Krupy związany był z udziałem w obliczeniach w ramach projektu WeFold. Publikacja, której jest współautorem, omawia właśnie przykłady białek stanowiące obiekt analizy w tym eksperymencie (załącznik 1). Nowatorskie podejście zastosowane w ramach rozwijania modelu gruboziarnistego pola siłowego UNRES polega na wprowadzenia więzów w postaci przewidywania kontaktów między resztami aminokwasowymi. W ramach



DZIEKANAT
Wydziału Chemii UG

Wpłynęło dn. 26.05.2015

L.dz. 8010-HCH/JP-866/15

projektu WeFold – elitarnego przedsięwzięcia o charakterze cooptation - grupa o nazwie Cornell-Gdańsk w zespole wFPUNK uzyskała najlepszy wynik dla T0740, natomiast średni wynik uzyskany w oparciu o rozwijany model UNRES plasował się na trzeciej pozycji. Udział Doktoranta w tym przedsięwzięciu – jak sam Autor to określa – polegał na pomocy w tworzeniu schematu przeprowadzania obliczeń w polu siłowym UNRES dla wersji bez więzów oraz z wprowadzonymi więzami. Doktorant pracował także w zespole przeprowadzającym symulacje dynamiki molekularnej. Mgr Paweł Krupa brał także udział w analizie uzyskanych wyników biorąc pod uwagę szczególnie wpływ wprowadzonych więzów na przebieg obliczeń w oparciu o model UNRES. Wykazano, że wprowadzenie więzów w znaczącym stopniu przyspieszyło obliczenia. Wprowadzenie więzów jest o tyle ciekawe, że struktura startowa dla tego typu obliczeń to struktura rozciągniętego łańcucha (forma extended). Wyniki udziału w projekcie WeFold zostały opublikowane w pracy stanowiącej Załącznik 1.

Drugie osiągnięcie Doktoranta to wprowadzenie statystycznych potencjałów korelujących położenie zjednoczonych łańcuchów bocznych i łańcucha głównego w polu siłowym UNRES. Dla standaryzacji parametrów dla takiego modelu wykorzystano strukturę białka 1GAB. Mgr Paweł Krupa przeprowadził analizę statystyczną fragmentów pętli – fragmenty o najwyższym stopniu nieokreśloności w przewidywaniu struktury – w 4585 białkach analizując struktury obecnych w nich pętli oraz fragmentów nieustrukturalizowanych. Ta analiza doprowadziła do konstrukcji alfabetu wyrażającego sekwencję aminokwasową uwzględniającą specyfikę poszczególnych aminokwasów. W efekcie uzyskano pięć kategorii. Charakterystyczne jest wyodrębnienie G i P jako odrębnych klas ze względu na ich istotne zróżnicowanie wobec wszystkich pozostałych aminokwasów. Dla wprowadzonych kategorii – alfabetu – wyznaczono tzw. potencjały średniej siły (PMF) dla odpowiednich kątów torsyjnych dla pseudo-wiązań łączących pozycje atomów C_{α} . Standaryzację poszerzono włączając także analizę struktur białek: 1E0L i 1L2Y. Testy przeprowadzono na ośmiu małych białkach (tak je określa Autor) – 1BDD, 1CLB, 1E0G (dwa ostatnie białka z niewielkim udziałem struktury typu Beta), 1E0L (białko całkowicie beta-strukturalne), 1GAB, 1KOY, 1POU oraz 1PRU. Wynik testów wyrażono za pomocą profilu odległości pozycji atomów C_{α} modeli w stosunku do białka natywnego. Wprowadzenie nowo-zdefiniowanych potencjałów torsyjnych poprawiło wyniki w stosunku do poprzednio stosowanego pola, aczkolwiek najmniejsza zgodność w dalszym ciągu dotyczy głównie tych samych fragmentów łańcucha zlokalizowanych w obrębie pętli między-helikalnych (Rys. 12 str. 65). Efekty wprowadzonych innowacji w postaci zmodyfikowanego potencjału torsyjnego omówiono w publikacji, którą zamieszczono jako Załącznik 2.

Kolejnym osiągnięciem mgra Pawła Krupy jest Jego udział w opracowaniu fizykochemicznego potencjału pół-empirycznego w oparciu o obliczenia bazujące na technice AM1 dla rotacji na wirtualnym wiązaniu łączącym pozycje atomów C_{α} . Tutaj parametryzację wykonano wykorzystując strukturę 1L2Y i 1LE1. Wyznaczone potencjały torsyjne zastosowane zostały do obliczeń testowych symulacji dynamiki molekularnej (osiem białek) dla 5. różnych temperatur.

W konkluzji analizy porównawczej uzyskanych wyników okazuje się, że pół-empiryczna analiza odtworzyła wyniki uzyskiwane za pomocą potencjałów statystycznych wprowadzonych i analizowanych w poprzednim punkcie omawianych osiągnięć mgra Pawła Krupy. Struktury wynikowe poddane zostały analizie skupień. Analiza porównawcza z poprzednimi wynikami wskazała wartość wagi dla potencjału torsyjnego obniżając ją z 0.5 na 0.25. W konkluzji podano też, że metoda pół-empiryczna dała lepsze wyniki na skutek lepszego rozpoznania przez pole UNRES typów struktury drugorzędowej. Dzięki

zastosowaniu wszystkich poprawek uzyskano znaczącą polepszenie przewidywanych struktur w tym szczególnie dla białka 2CRB. Ważne jest też, że wprowadzenie nowego podejścia nie skutkowało pogorszeniem jakiegokolwiek wyniku.

Rola mgra Pawła Krupu w tej części badań polegała na uzupełnieniu symulacji pół-empirycznych blokowanych reszt aminokwasowych o prolinę, wyznaczeniu hiperpowierzchni energii potencjalnej dla blokowanych dipeptydów, optymalizacji wag potencjałów oraz przeprowadzeniu obliczeń testowych.

Wyniki omawianego eksperymentu zawarte zostały w publikacji, która w całości zamieszczona jest w rozprawie jako Załącznik 3.

Dalsze rozszerzenie pola siłowego UNRES to uwzględnienie informacji pochodzących z technik przewidywania struktur białkowych bazujących na białkach homologicznych. Ta innowacja wprowadza do modelu typu ab initio - jakim jest model UNRES - elementy technik stosowanych w metodach bazujących na podobieństwie sekwencji i struktur białek. Wprowadzona innowacja polega na wprowadzeniu dodatkowego składnika pola siłowego U_{restr} nakładającego więzy ograniczające zmianę odległości między centrami oddziaływującymi (atomy zjednoczone) a także dla kątów torsyjnych. To podejście jest hybrydą kilku modeli białek do obliczeń energii więzów. Model ten został zaimplementowany i przetestowany z wykorzystaniem pięciu białek jedno-domenowych z edycji CASP9 i czterech z edycji CASP10 oraz dla białek dwu-domenowych. Wyniki wysoce pozytywne zostały uzyskane dla konstrukcji domen, natomiast położenie między-domenowe okazało się niezadowolające. Jak zwykle w procedurze testowania przeprowadzono liczne symulacje dynamiki molekularnej w oparciu o zmodyfikowane pole siłowe. W rezultacie uzyskano kolejną korektę wyrażającą się zwiększoną zgodnością przewidzianych struktur z białkiem docelowym. Postęp w rozwijaniu pola UNRES jest widoczny na przykładzie białka T0663 (str 78) gdzie poprawnie przewidziano znaczący fragment białka o strukturze typu Beta.

W „Podsumowaniu” Autor wymienia zalety i wady pola siłowego UNRES wskazując jako pierwsze: poprawne przewidywanie struktury drugo- i trzecio-rzędowej i jako drugie – przeciętną rozdzielczość uzyskiwanych struktur.

Generalnie wprowadzone przez mgra Pawła Krupę modyfikacje przyczyniły się do poprawy skuteczności działania pola siłowego UNRES.

Oceniając przedstawioną mi do recenzji rozprawę trudno jest polemizować z modelem o tak wysokich światowych notowaniach, jakie posiada model UNRES. Świadczą o tym wysokie pozycje publikacji naukowych omawiających zagadnienie konstruowania struktur białek z wykorzystaniem tego pola. Pole UNRES rozwijane jest przez zespół Cornell-Gdańsk od roku 1999. Sukcesy, jakie daje zastosowanie tego pola, skutkują zaliczeniem zespołu Cornell-Gdańsk do światowej czołówki wiodących jednostek zajmujących się profesjonalnie zagadnieniem przewidywania struktur białek. Z przyjemnością i satysfakcją a wręcz z poczuciem dumy narodowej czyta się prace, gdzie nazwiska specjalistów z zespołu Uniwersytetu Gdańskiego zamieszczone są wraz z nazwiskami laureatów nagrody Nobla (czego przykładem jest publikacja podana w Załączniku 1).

Jedno, co pozostaje Recenzentowi, to życzyć zespołowi dalszych sukcesów oraz pomyślnego rozwoju wygenerowanych narzędzi.

Praca doktorska wymaga jednak oceny i jako uwagę krytyczną poddam pod wątpliwość przyszłość pola siłowego UNRES. Pierwsza opublikowana wersja tego pola zawierała 7

składników (Lee J, Liwo A, Scheraga HA. Energy-based de novo protein folding by conformational space annealing and an off-lattice united-residue force field: application to the 10-55 fragment of staphylococcal protein A and to apo calbindin D9K. Proc Natl Acad Sci U S A. 1999 Mar 2; 96(5):2025-30). Obecnie liczba składników wynosi 15 (strona 31 rozprawy doktorskiej). Z treści pracy można wnioskować, że Autorzy tego modelu skupiają się na ciągłej rozbudowie tego pola poprzez wprowadzanie kolejnych składników mających na celu poprawę skuteczności. Biorąc pod uwagę wirtualność elementów (pseudo-atomów) oraz parametrów ich opisujących, a w tym szczególnie wartości wag dla każdego z komponentów można obawiać się nie tyle dużej liczby zmiennych (którą to liczbę modele gruboziarniste starają się zredukować), ale dużej liczby parametrów, które nie podlegają weryfikacji eksperymentalnej.

Niewątpliwym jednak postępem, jaki jest odnotowany od roku 1999, jest trafne przewidywanie struktury ponad-drugorzędowej z dużą zawartością struktury typu Beta (np. str 78 – T0663).

Przechodząc jednak to samej pracy doktorskiej. Została ona napisana w sposób bardzo staranny. Swoboda, z jaką Autor porusza się po całym spektrum zagadnień związanych ze strukturą białek i technik obliczeniowych do tego celu stosowanych wymaga bardzo wysokiej oceny. Autor biegle posługuje się nowoczesnymi technikami informatycznymi, co czyni Go specjalistą wysokiej klasy.

Mimo wielkiej staranności można jednak doszukać się pewnych niedociągnięć (taka jest rola Recenzenta).

Na str 59 bardzo szczegółowo podano wymagania komputerowe i moce obliczeniowe użytej maszyny. Nie podano jednak, w jakim ośrodku te obliczenia były wykonane.

Przy omawianiu każdej z publikacji Autor podaje swój udział w danym projekcie. Jest to lista czynności wykonanych przez Autora. Nie ma natomiast informacji o Jego kreatywnej roli. Informacje te robią wrażenie wykonania obliczeń wg wcześniej ustalonego scenariusza.

Mimo wspomnianej już wielkiej staranności w opracowaniu tekstu, jaką niewątpliwie reprezentuje praca, sugeruję, aby Autor ustrzegł się zdań tak długich, jak ma to miejsce na str 71 (12 linii).

Tekst na str. 72 opisujący wzór (19) nie jest jasny. Kąt ω nie został zdefiniowany.

Drobny błąd językowy (gramatyczny) w trzeciej linii tekstu str 72 to zastosowanie niewłaściwego przypadku.

W podsumowaniu stwierdzam, że przedstawiona praca doktorska w pełni odpowiada kryteriom stawianym pracom doktorskim. Pracę oceniam jako BARDZO DOBRĄ i wnoszę do Pana Dziekana o dopuszczenie mgra Pawła Krupy do dalszych etapów przewodu doktorskiego.

Kraków, 11 maja 2015


Prof dr hab Irena Rotermań-Konieczna
Recenzent