



## RECENZJA

### Rozprawy doktorskiej mgr Karoliny Mitusińskiej

pt.: *Analiza skutków wyboru modelu zrekonstruowanej pętli na funkcjonalność białka*

wykonanej pod opieką naukową dr hab. Artura Góry

*Podstawę formalną wykonania recenzji stanowi pismo Rady Dyscypliny Nauki Chemicznej z dnia 8 VII 2020 r, zgodnie z Uchwałą Rady Wydziału Chemicznego Politechniki Śląskiej podjętą w dniu 16 I 2019 r.*

Metody obliczeniowe służące do badania struktury i konformacji protein, jak również zajmujące się analizą odpowiednich oddziaływań wewnątrz- i międzycząsteczkowych mają ugruntowaną pozycję w dziedzinie nauk chemicznych oraz biochemicznych. Z uwagi na dużą złożoność badanych układów oraz kompleksowość procesów w nich zachodzących, równoległe z badaniami skupionymi na konkretnych układach podejmowane są liczne starania mające na celu ulepszenie dokładności oraz wydajności dostępnych metodologii oraz towarzyszących im narzędzi obliczeniowych. Oprócz podejść idących w kierunku lepszego odzwierciedlenia właściwości realnych układów (poprzez np. zwiększenie dokładności potencjałów oddziaływań) istotne są również starania sprofilowane na usprawnienie zastosowanych procedur lub zaproponowanie alternatywnych rodzajów analiz danych. Tego typu problematyce poświęcona jest recenzowana rozprawa doktorska, napisana przez mgr Karolinę Mitusińską.

Przedstawiona do oceny rozprawa skonstruowana jest w formie kompilacji artykułów opublikowanych w uznanych czasopismach o zasięgu międzynarodowym, znajdujących się na liście filadelfijskiej. Do kolekcji artykułów dołączony jest wstęp, w którym autorka przedstawiła wybrane zagadnienia bezpośrednio związane z tematyką badań a także skondensowany opis celów oraz rezultatów pracy.

Podstawą recenzowanej pracy jest osiem artykułów naukowych, z czego jeden z nich jest artykułem przeglądowym. Są to artykuły opublikowane w: *International Journal of Molecular Sciences* (2 artykuły, IF = 4,556), *Bioinformatics* (2 artykuły, IF = 5,610), *Biomolecules* (2 artykuły, IF = 4,082), *PloS One* (1 artykuł, IF = 2,740) oraz *Computational and Structural Biotechnology Journal* (1 artykuł, IF =



6,018). Autorka określiła udział w realizacji wszystkich prac w sposób opisowy, definiując wkład nie tylko swój ale często też pozostałych współautorów. Z opisu tego wynika, że jej rola w realizacji wszystkich prac będących podstawą doktoratu była znacząca i dotyczyła m.in. budowania układów molekularnych, przeprowadzania symulacji metodą dynamiki molekularnej, analizowania ich wyników w ramach zarówno istniejących procedur (PCA) lub nowatorskich metod których opracowanie było częścią jej pracy (AQUA-DUCT), analizowania i modelowania struktur białek z pomocą specjalistycznych programów, testowania nowego oprogramowania, pisanie manuskryptów publikacji oraz przeprowadzania analizy literatury. Z tym stwierdzeniem koresponduje miejsce doktorantki na liście autorów w ww. artykułach które jest albo pierwsze (3 artykuły), drugie (z adnotacją o równym wkładzie względem pierwszego autora, 2 artykuły), drugie (bez takiej adnotacji, 2 artykuły) lub czwarte (1 artykuł).

Tematem wiążącym wszystkie prace jest opracowywanie, doskonalenie oraz zastosowanie wybranych metod obliczeniowych w celu badania struktury oraz mechanizmu działania protein. Cel zdefiniowany we wstępie przez Autorkę rozprawy dotyczy dwóch głównych aspektów. Pierwszy, zgodnie z tytułem rozprawy, jest związany z problematyką wyboru modelu zrekonstruowanej pętli białka, drugi natomiast wiąże się z opracowywaniem programów komputerowych służących do analizy funkcjonalności protein. Analizując bardziej szczegółowo charakter uzyskanych wyników można stwierdzić, że trzecim, nie sprecyzowanym wprost celem, jest walidacja oraz testowanie funkcjonalności opracowanego programu AQUA-DUCT połączone z badaniami mechanizmów biomolekularnych leżących u podstaw funkcji czterech wybranych protein.

Dołączony do zestawu publikacji wstęp jest dość zwięzły lecz pokazuje, że Autorka rozprawy doskonale porusza się w tematyce prowadzonych przez siebie badań. Treść wstępu podzielono kilka sekcji, dotyczących m.in. problematyki modelowania drugorzędowej struktury białek (włączając w to konkretny przypadek pętli) oraz struktur i mechanizmu działania enzymów z rodziny hydrolaz epoksydowych. Całość każdej z tych sekcji stanowi dość spójny ciąg, w którym Autorka wprowadza czytelnika w zagadnienia istotne dla opisanych później badań. Następnie precyzowane są cele badawcze a na końcu prezentowane są (w sposób skondensowany w porównaniu do oryginalnych artykułów naukowych) wyniki uzyskane w toku badań. Prezentację wyników podzielono w analogii do artykułów w których się one oryginalnie ukazały.



Poniżej omawiam badania i ich wyniki, zawarte w pracach będących podstawą rozprawy. Omówienie to jest dość pobieżnie z uwagi m.in. na to, że prace wchodzące w skład rozprawy były już recenzowane. Dodatkowo, z racji faktu, iż niektóre prace dotyczą badań multidyscyplinarnych, skupiam się na elementach dotyczących bezpośrednio recenzowanej rozprawy.

- Praca nr 1 dotyczy tematyki modelowania pętli łączących bardziej ustrukturyzowane fragmenty protein, tj.  $\alpha$ -helisy i  $\beta$ -kartki. Główną idea tej pracy jest wprowadzenie wewnętrznego podziału rodzaju pętli na elastyczne i statyczne, zależnie od skali czasowej reorientacji ich struktury. W pracy dyskutowano alternatywną metodę wyboru optymalnego modelu zrekonstruowanej pętli białka, bazując na założeniu większej elastyczności pętli w porównaniu do tzw. pętli statycznych, jakich struktury są zazwyczaj możliwe do uzyskania na drodze badań krystalograficznych. Opracowano kryteria bazujące na prostych deskryptorach geometrycznych oraz miarze ich fluktuacji, jaka może być uzyskana na drodze symulacji lub aplikacji narzędzi do rekonstrukcji pętli. Przeprowadzono również analizę parametrów kwantyfikujących jakość modelu pętli i uzasadniono używanie kryterium energii potencjalnej w kontekście pętli elastycznych.
- Prace nr 2 i 3 dotyczą prezentacji dwóch wersji programu AQUA-DUCT, dystrybuowanego niezależnie od samych prac. Sam program jest jednak nierozzerwalnie złączony z pracą Doktorantki wykonaną w ramach jej badań naukowych, dlatego też warto oceniać te prace w połączeniu z oprogramowaniem będącym ich istotą. AQUA-DUCT został zaprojektowany jako narzędzie do analizy migracji wybranych części układu (domyślnie są to niewielkie cząsteczki rozpuszczalnika) podczas symulacji układów biomolekularnych. Taka perspektywa pozwala na szerszą interpretację funkcji różnych części biomolekuly jakie mogą być istotne z punktu widzenia pełnego zrozumienia mechanizmu działania danego układu lub też, z bardziej praktycznego punktu widzenia, projektowania celów molekularnych lub modyfikacji biomolekuly zgodnie z pożądanym efektem. W pewnych aspektach zaproponowana w pracy idea uzyskiwania informacji nt. właściwości układu z analizy wyselekcjonowanych fragmentów przestrzeni konfiguracyjnej przypomina metody zorientowane na próbkowanie stanów przejściowych (np. *transition path sampling*). Ta część wyników posiada w opinii Recenzenta duży potencjał naukowy, możliwy do wykorzystania w szerokim zakresie układów biomolekularnych, nie ograniczającym się do przykładów zaprezentowanych w pracy ale obejmującym inne rodzaje białek a nawet układy i zjawiska odmiennego typu, np. penetracja błon lipidowych przez rozpuszczalnik



lub dynamiczne przepływy wody przez hydrofilowe żele. Swobodny dostęp do programu AQUA-DUCT oraz towarzyszące mu obszerne informacje na stronie internetowej świadczą, że autorzy są świadomi potencjalnie istotnej roli jaką może odgrywać ich narzędzie w dziedzinie analizy symulacji biomolekularnych.

- o Prace nr 4-7 demonstrują możliwości programu AQUA-DUCT, zastosowanego do analizy symulacji konkretnych układów biomolekularnych. W szczególności, unaoczniono potencjał aplikacji innych niż czysta woda rozpuszczalników w celu lokacji obszarów białka zdolnych do wykazywania specyficznych oddziaływań międzycząsteczkowych. Nie są to oczywiście jedyne cele ww. prac, a samo zastosowanie metodologii analizy zaproponowanej przez autorów jest jedynie częścią szerszej zaprojonowanych badań, mających na celu wyjaśnienie konkretnych problemów naukowych. W przypadku pracy 7 warto również zauważyć godną docenienia tendencję podążania za aktualnymi potrzebami działań wobec obecnie panującej pandemii COVID-19.
- o Praca nr 8 jest pracą przeglądową, mającą na celu prezentację metodologii oraz towarzyszącym im narzędziom obliczeniowym wykorzystującym obecność wody do analizy właściwości makromolekul o znaczeniu biologicznym. Praca ta stanowi dowód na bardzo dobre rozeznanie Doktorantki w temacie podejść obliczeniowych jakie mogą być zastosowane do badania biomolekul rozpatrywanych w obecności rozpuszczalnika.

Wyniki zawarte w ww. pracach oceniam jako wartościowe i wnoszące istotny wkład w dziedzinę modelowania biomolekularnego a także mające potencjał do wyznaczenia nowych trendów w analizie wyników symulacji. Ponadto, cele badawcze deklarowane przez Doktorantkę zostały w pełni zrealizowane.

Niezależnie od faktu, iż prace stanowiące rozprawę doktorską oraz wyniki w nich zawarte były już uprzednio recenzowane, szczegółowa analiza rozprawy, wliczając w to dołączony wstęp ujawniła pewne niedociągnięcia, które z obowiązku recenzenta przedstawiam w formie uwag krytycznych:

- o Interesującymi wynikami są te zawarte w pracy nr 1, sugerujące przewagę alternatywnej metodologii rozwijania liczniejszych lecz krótszych trajektorii nad przedłużaniem czasu symulacji. Wprawdzie istnieją metody bazujące na uzyskiwaniu informacji o procesach o dużej skali czasowej w oparciu o krótkie odcinki trajektorii (np. *milestoning* lub *transition interface sampling*) lecz w każdym z tych przypadków niezbędny jest czynnik popychający ewolucję układu w stronę konwergentnych wyników. Takim czynnikiem jest generowanie



dużej liczby nierównowagowych trajektorii nie inicjowanych z tego samego punktu przestrzeni fazowej (TIS) lub dążenie do używania równowagowej trajektorii przy generowaniu „cząstkowych” trajektorii (*milestoning*). Omówione przez autorkę wyniki są inicjowane z tej samej konfiguracji, a obserwowana dywergencja może nie odpowiadać stanowi równowagi termodynamicznej z uwagi na ograniczony czas pojedynczej symulacji. Nie twierdzą tu, że wyniki uzyskane w pracy są błędne, jednak użyte stwierdzenie (cyt.: „We therefore recommend for MD simulations using more repetitions of sufficient length, rather than extending the simulations toward microsecond lengths”) jest zbyt mocne w kontekście bardziej ogólnym, obejmującym inne proteiny oraz inne modelowane pętle. Z pewnością zagadnienie to zasługuje na kompleksowe, systematyczne przebadanie przy użyciu szerszego zakresu układów modelowych.

- Wybór parametru *loop-prot-sh*, użytego w pracy nr 1 i zdefiniowanego jako najmniejsza odległość pomiędzy atomami w pętli a tymi poza pętlą jest niejasny. W przypadku braku grupy atomów nie przynależących do żadnej z tych grup (a o takim dodatkowym warunku w definicji nie wspomniano), wartość tego parametru będzie w przybliżeniu stała i równa średniej długości wiązania kowalencyjnego przypadającego na granicę pętla/nie-pętla.
- Tytuł rozprawy nie odzwierciedla w pełni jej zawartości. Problematyka wyboru modelu pętli jest kluczowa dla pracy nr 1 lecz niekoniecznie w pozostałych pracach, gdzie modelowanie jest przeprowadzone bez konieczności rekonstrukcji pętli.
- Najobszerniejszą część wprowadzenia w tematykę pracy (10 stron) stanowi zarys problematyki struktury, funkcji, podziału i mechanizmu działania enzymów z rodziny hydrolaz epoksydowych. Opis ten jest na tyle obszerny, że sugeruje czytelnikowi, iż to właśnie ta rodzina białek będzie stanowiła główny cel badań, tudzież układ modelowy w którego kontekście rozpatrywane będą wszystkie zagadnienia deklarowane w celach pracy. W wyniku dalszej lektury okazuje się, że jedynie dwie na osiem prac (artykuły nr 1 i 6) dotyczą bezpośrednio tej rodziny enzymów
- Z kolei tematyka oddziaływań woda-białko lub rozpuszczalnik białko (zarówno w bardziej ogólnym kontekście, jak i takim ukierunkowanym na bliższy pracy aspekt mobilności cząsteczek rozpuszczalnika w sąsiedztwie makromolekuł) jest we wstępie zupełnie pominięta.

- To samo dotyczy choćby zgrubnej charakterystyki pozostałych układów badanych w pracach nr 4, 5 i 7, a mianowicie, oksydazy D-aminokwasów, izomerazy fosfoglukozy oraz proteazy wirusa SARS-CoV-2. Wątki takie byłyby pomocne w odniesieniu do zawartego później opisu wyników dotyczących tych właśnie układów.
- Choć wstęp do zestawu prac został napisany starannie, Autorka nie ustrzegła się kilku niewielkich błędów natury gramatycznej, typograficznej czy edytorskiej lub niejasnych sformułowań. Przykładowo:
  - str. 13: „Ideally, MD simulations will generate trajectories which follow the free energy landscape for a given system according to the force field used in the simulations” --- czy jest możliwa sytuacja alternatywna, w której trajektoria nie odpowiada takiemu krajobrazowi?;
  - str. 18: „A structure is an  $\alpha/\beta$  hydrolase member ...” --- z kontekstu wynika, że chodziło raczej o hydrolazy epoksydowe;
  - str. 25: w jednym zdaniu określono AQUA-DUCT jako „unique tool” oraz skomentowano pracę skupioną na innych tego typu programach („other software utilizing water molecules...”);
  - str. 23 “several attempt”, str. 39.: “focus the water...” --- błędy gramatyczne;
  - str.: 35: pracę nr 4 błędnie określono jako pracę nr 3;
  - str. 46: „intramolecular correlations” --- nie podano jakie są to korelacje.

Z drugiej strony, liczba takich błędów jest niewielka i nie stanowi żadnej przeszkody w zrozumieniu treści wyводу.

Zaznaczam, że ww. uwagi dotyczą, za wyjątkiem dwóch pierwszych z nich, wyłącznie podsumowania wyników oraz wprowadzenia czytelnika w tematykę pracy, które to zagadnienia zostały przedstawione we wstępie do zestawu publikacji i w dużej mierze są niezależne od wyników oraz dyskusji zawartych w załączonych artykułach.

Oprócz diskutowanej powyżej rozprawy doktorskiej oraz wyników w niej zawartych, należy także docenić dużą aktywność Doktorantki na innych polach związanych z nauką. W szczególności chodzi tu o ponadprzeciętną aktywność konferencyjną (20 prezentacji konferencyjnych, w tym 7 w roli prezentującej wyniki) oraz udział w roli współorganizatorki warsztatów naukowych. Autorka rozprawy wykazuje również duże zaangażowanie w realizowanie projektów badawczych (udział w 3 projektach, z czego w 1 jako główny laureat). Ponadto,



Doktorantka przygotowała obszerny podręcznik użytkownika (*User Guide*) dla programu AQUA-DUCT odpowiadający treści pracy nr 2 (częściowo) oraz 3 (w całości).

Oceniając wysoko poziom badań naukowych będących podstawą recenzowanej rozprawy, w podsumowaniu stwierdzam, że przedstawiona przez Doktorantkę praca doktorska spełnia warunki stawiane rozprawom doktorskim przez *Ustawę o stopniach naukowych i tytule naukowym w zakresie sztuki* z dnia 14 marca 2003 r (wraz z późniejszymi zmianami). Dlatego też wnioskuję do Rady Wydziału Chemicznego Politechniki Śląskiej o dopuszczenie mgr Karoliny Mitusińskiej do dalszych etapów przewodu doktorskiego.

Ponadto, biorąc pod szereg czynników pozytywnie wyróżniających pracę naukową oraz aktywność Doktorantki, wnioskuję o wyróżnienie rozprawy. Uzasadniam to przede wszystkim ponadprzeciętnym dorobkiem naukowym wnoszącym istotny wkład do tematyki modelowania struktury białek oraz opracowywania nowatorskich metod służących do analizy wyników symulacji. Nie bez znaczenia jest też fakt, że wszystkie prace, będące podstawą recenzowanej rozprawy powstały w ciągu zaledwie 3 lat studiów doktoranckich.

Z poważaniem,

dr hab. Wojciech Plaziński, prof. IKiFP PAN