



Dr hab. inż. Marek Wojciechowski

Gdańsk 02.07.2019

**Recenzja rozprawy doktorskiej mgr Agnieszki Sylwii Karczyńskiej  
pt. „Przewidywanie struktury białek oraz ich kompleksów przy  
użyciu gruboziarnistego pola siłowego UNRES z wykorzystaniem  
informacji z baz danych”**

Ustalenie pełnej trójwymiarowej struktury cząsteczki białka jest niesłychanie istotne dla większości obszarów działania współczesnej biologii. Jednak pomimo tego iż, dzięki szybkiemu rozwojowi metod sekwencjonowania, ponad 160 milionów sekwencji różnych białek jest dostępnych w bazach takich jak UniProtKB, to tylko nieco ponad 150000 struktur jest dostępnych w bazie PDB. Tymczasem to właśnie znajomość struktury białka ma znaczenie praktyczne np. w projektowaniu leków. Jak łatwo policzyć istnieje tylko ułamek procenta szansy na to, że białko znajdujące się w obszarze zainteresowań ma poznaną strukturę i z roku na rok ta dysproporcja rośnie ze względu na różną dynamikę pojawiania się nowych struktur i nowych sekwencji. Z tego powodu bardzo wzrosło znaczenie teoretycznych metod przewidywania struktur jako potencjalnie jedyne sposoby, który mógłby, do pewnego stopnia, zaspokoić rosnące zapotrzebowanie na dobrej jakości struktury białek, których sekwencje są już znane. Oczywiście metody teoretyczne będą mogły spełnić te oczekiwania tylko wtedy, kiedy używane za ich pomocą modele będą wystarczająco dobre do tego, aby mogły być używane na równi ze strukturami uzyskiwanymi eksperymentalnie. Po części aby to właśnie sprawdzić a z drugiej strony zdopingować grupy naukowe do ciągłego ulepszania swoich metodologii organizowane są konkursy CASP (ang. *Critical Assessment of techniques for protein Structure Prediction*), w których co dwa lata najlepsze zespoły naukowe mogą w obiektywny sposób przetestować opracowane przez siebie algorytmy.

W związku ze wspomnianym zapotrzebowaniem na wysokiej jakości struktury i jednocześnie upowszechnieniem się niskorozdzielczych metod eksperymentalnych dodatkowym obszarem zastosowań metod teoretycznych jest udokładnianie struktur i odtwarzanie modeli pełno-atomowych na podstawie danych pochodzących z niskorozdzielczych eksperymentów takich jak kriomikroskopia elektronowa czy nisko kąto-  
we rozpraszanie promieniowania rentgenowskiego (SAXS).

Przedstawiona mi do oceny rozprawa doktorska Pani mgr Agnieszki Sylwii Karczyńskiej poświęcona jest udoskonalaniu istniejących jak i opracowaniu nowych narzędzi służących do przewidywania struktur białek i ich kompleksów. W swojej pracy doktorantka wykorzystywała przede wszystkim pole siłowe UNRES, które zostało opracowane i od wielu lat jest rozwijane przez pracowników Katedry Chemii Teoretycznej Uniwersytetu Gdańskiego.

Recenzowana rozprawa doktorska przygotowana została jako osiągnięcie publikacyjne. Została napisana na 100 stronach i podzielona na 9 rozdziałów obejmujących wykaz skrótów, wykaz używanego oprogramowania i serwerów obliczeniowych oraz część teoretyczną i wynikową. Na część wynikową składa się streszczenie publikacji stanowiących podstawę osiągnięcia i skrótowo opisanych w pięciu kolejnych podrozdziałach. Rozprawę zamykają wykaz rysunków oraz obejmująca 194 pozycje bibliografia, której około 10% stanowią prace z ostatnich 3 lat. Dodatkowo na końcu rozprawy doktorantka zamieściła pełną treść wspomnianych 5 publikacji, informację o swoim pozostałym dorobku naukowym oraz oświadczenia współautorów poszczególnych artykułów.

W osobnym rozdziale zatytułowanym „*Cel pracy*” doktorantka przedstawiła podstawowe cele jakie zamierzała osiągnąć. Podstawowym celem było, jak sama pisze, „*opracowanie i udoskonalanie narzędzi służących do przewidywania struktur białek oraz ich kompleksów z wykorzystaniem informacji z baz danych oraz wyników uzyskanych metodami eksperymentalnymi*”. Zadanie to doktorantka zamierzała zrealizować rozszerzając pole siłowe UNRES o możliwość definiowania więzów pochodzących z różnych źródeł i optymalizując parametry określające siłę tych więzów a następnie sprawdzając jak wpływają one na przebieg symulacji.

Dodatkowym celem jaki doktorantka miała zamiar osiągnąć było opracowanie metody umożliwiającej przewidywania struktur białek z wykorzystaniem informacji po-

chodzących z pomiarów eksperymentu nisko kąтового rozpraszania promieniowania rentgenowskiego (SAXS).

W trakcie swojej pracy badawczej doktoranta miała okazję zweryfikować skuteczność opracowywanych przez siebie modyfikacji pola siłowego UNRES biorąc udział w II i 12 edycji eksperymentu CASP, a także pomagała w implementacji swoich osiągnięć na serwerze internetowym umożliwiającym zdalne przeprowadzanie symulacji we wspomnianym polu siłowym.

Teoretyczne wprowadzenie do zagadnień poruszanych w recenzowanej pracy znalazło się w rozdziale 5 zatytułowanym „Część teoretyczna” i obejmującym 54 strony. Rozdział ten podzielony został na szereg podrozdziałów omawiających dosyć selektywnie takie zagadnienia jak struktury i budowa białek, metody symulacyjne takie jak mechanika i dynamika molekularna, metody poprawy próbkowania w symulacjach dynamiki molekularnej, metody analizy danych uzyskiwanych jako wyniki symulacji MD oraz teoretyczne metody przewidywania struktur białek jak i metody eksperymentalne wspomagające te pierwsze. Wszystkie rozdziały są dosyć skrótowe a bardziej szczegółowo omówione zostały tylko te wybrane techniki obliczeniowe, które były bezpośrednio wykorzystywane w pracy doktorantki. Trochę pogarsza to czytelność tej części rozprawy, zwłaszcza, że podpisy pod niektórymi rysunkami są minimalistyczne a układ podrozdziałów wydaje się chwilami nieco chaotyczny i niekonsekwentny.

Zapewne ze względu na tą skondensowaną formę wkradło się też do treści kilkanaście literówek i drobnych błędów stylistycznych, które niestety gdzieś wpływają też na czytelność i jednoznaczność tekstu. Na stronie 14 jest „analizie badań” powinno być „analizie wyników badań”. W drugim akapicie na stronie 20, badane są raczej zjawiska a efekty są uzyskiwane. W tekście na stronie 64 zostały pomyłone numery rozdziałów. Zamiast (7.1.–7.5) powinno być (6.1.–6.5). Rysunek 5 na stronie 21 przedstawiający różne poziomy reprezentacji łańcucha polipeptydowego byłby czytelniejszy, gdyby przedstawione na nim poszczególne modele były zgodne z opisem (czyli np. gdyby model zjednoczonych atomów był pozbawiony atomów wodorów, które w tym modelu nie są bezpośrednio obecne itd.). Również nie pomaga tu bardzo lakoniczny podpis pod rysunkiem, który powinien zawierać pełniejszą informację, np. że poszczególne elipsy na rysunku oznaczają grupy atomów, które w danym modelu są zastąpione pojedynczym punktem oddziaływań. W podpisie pod rysunkiem 6 określenie „*k*ąt

walencyjny pomiędzy trzema sąsiednimi atomami  $C^{\alpha}$  jest nieuprawnione, ponieważ z definicji kąt walencyjny jest kątem płaskim pomiędzy trzema atomami połączonymi wiązaniami, co nie ma miejsca w przypadku trzech atomów  $C^{\alpha}$ . Z resztą w odnośnej publikacji, której częścią jest opisywany tu fragment, kąt ten określony jest po prostu jako „kąt pomiędzy wirtualnymi wiązaniami” i takie określenie nie budzi już powyższych wątpliwości.

Pewien brak precyzji występuje też w opisie periodycznych warunków brzegowych na stronie 32. Sformułowanie „Symulowana cząsteczka oddziałuje nie tylko na cząsteczki znajdujące się w jej pobliżu, ale także na inne znajdujące się w sąsiednich pudełkach symulacyjnych” jest niepoprawne. W praktyce to cząsteczki z sąsiednich „wirtualnych” pudełek oddziałują na cząsteczki obecne we właściwym pudełku centralnym a pozycje tych pierwszych są po prostu ustalane przez operacje symetrii, więc nie muszą w tym wypadku być liczone działające na nie siły. Inaczej trik z zastosowaniem periodycznych warunków brzegowych byłby bezcelowy i mielibyśmy zwyczajnie do czynienia ze zdecydowanie większym pudełkiem symulacyjnym.

W rozdziale omawiającym teoretyczne metody przewidywania struktur białek znalazło się sformułowanie: „Do najszerzej znanych metod teoretycznych do przewidywania struktur białek należy modelowanie homologiczne oraz symulacje molekularne”. Jest to dosyć nietypowy podział, zwłaszcza, że „symulacje molekularne” to pojęcie bardzo szerokie. Jednocześnie widzę tu pewien brak konsekwencji w stosowanej terminologii gdyż w późniejszych rozdziałach 5.6.1 i 5.6.3 jest mowa odpowiednio o „modelowaniu porównawczym” i „metodach de novo” co jest już zgodne z powszechnie przyjętym podziałem i terminologią.

W kolejnym rozdziale, poświęconym eksperymentom rozpraszania promieni rentgenowskich pod małymi kątami (SAXS), na rysunku 15 oraz na stronie 62 brak jest wyraźnego wyjaśnienia symboli, przez co pojawiające się w tekście sformułowania „zależność  $I(q)$  od  $q$ ” czy „wzrost  $q$ ” mogą być w pierwszej chwili niezrozumiałe.

Kolejny rozdział, rozdział 6 zatytułowany „Część wynikowa” zawiera bardzo skondensowany (na 12 stronach) opis przeprowadzonych obliczeń i ich rezultatów stanowiących osiągnięcie naukowe, które to wyniki zostały opublikowane w formie pięciu artykułów, a ich pełna treść załączona została na końcu rozprawy. Ponieważ artykuły te zostały opublikowane w dobrych czasopismach o wysokich współczynnikach oddzia-

tywania i przeszły już przez rygorystyczny proces oceny merytorycznej, nie widzę uzasadnienia do ponownej ich oceny, mam jednak do doktorantki kilka pytań.

W artykule opisanym w rozdziale 6.1 znalazło się sformułowanie, że „*przewidywania struktury przez PSIPRED są dokładne więc szansa na pogorszenie struktury pod wpływem nakładanych więzów jest niewielka*”. Chciałbym poznać opinię doktorantki na temat dokładności tego typu oszacowań, zwłaszcza w przypadku struktury  $\beta$ , kiedy to w formowaniu  $\beta$ -kartki istotny jest nie tylko lokalny skład łańcucha polipeptydowego ale jeszcze obecność drugiej wstęgi  $\beta$ . Poza tym czy rezultaty uzyskiwane za pomocą pola UNRES nie byłyby lepsze, (być może takie próby były prowadzone?) gdyby pole siłowe zostało sparametryzowane w oparciu o większą liczbę struktur danej klasy (obecnie używane są dwie wersje parametryzacji, pierwsza opracowana w oparciu o 1 a druga o 2 struktury PDB.)

W publikacji opisanej w rozdziale 6.2, w celu uwzględnienia informacji strukturalnej z konsensusowych modeli danego białka funkcja pseudo-energii pola UNRES została uzupełniona o dodatkowe człony opisujące 4 typy więzów, których wagi wyglądają jednak na dobrane dosyć arbitralnie na 0,5 dla więzów odległości i 1 dla pozostałych 3 członów. We wcześniejszych publikacjach (np. opisanej w rozdziale 6.3 interesującej pracy dotyczącej ergodyczności w symulacjach) doktorantka testowała, co prawda różne ich wielkości w dosyć szerokim zakresie, ale zawsze w zestawach jednorodnych, kiedy to wszystkie 4 typy więzów miały taką samą wagę. Chciałbym poznać uzasadnienie doboru, w tym przypadku, akurat takich wartości.

Ta i wszystkie wcześniejsze uwagi mają jednak charakter jedynie dyskusyjny czy wręcz kosmetyczny. Dotyczą głównie redakcji części wstępnej i nie umniejszają w żaden sposób wartości naukowej prezentowanej rozprawy, a dorobek naukowy doktorantki można uznać za imponujący na tym etapie kariery.

Pani mgr Agnieszka Sylwia Karczyńska realizowała 5 własnych projektów naukowych i brała udział w 3 innych. Odbiła dwa zagraniczne staże naukowe jeden krótkoterminowy i jeden roczny. Opublikowała 8 artykułów w dobrych i bardzo dobrych czasopismach. W 4 z nich jest pierwszym autorem. Prezentowała swoje osiągnięcia na licznych międzynarodowych konferencjach naukowych w tym 18 razy w formie wystąpień ustnych i 12 razy w formie posterów. Sumaryczna liczba cytowań artykułów wchodzących w skład osiągnięcia naukowego wynosi na chwilę obecną 27 (22 bez au-



tocytozań) co daje średnio 5,4 cytowania na artykuł. Indeks Hirscha doktorantki na podstawie Web of Science wynosi 4.

W świetle powyższego, mogę stwierdzić, że przedstawiona mi do oceny rozprawa doktorska spełnia wszystkie wymagania ustawowe oraz zwyczajowe stawiane tego typu pracom i rekomenduję Radzie Wydziału Chemicznego Uniwersytetu Gdańskiego dopuszczenie Pani mgr Agnieszki Sylwii Karczyńskiej do dalszych etapów przewodu doktorskiego. Jednocześnie ze względu na ponadprzeciętny dorobek doktorantki i dużą wartość naukową przedstawionego osiągnięcia wnioskuję do Rady Wydziału Chemicznego Uniwersytetu Gdańskiego o wyróżnienie recenzowanej rozprawy.

  
dr hab. inż. Marek Wojciechowski