



Jerzy Haber Institute
of Catalysis and Surface Chemistry
Polish Academy of Sciences



Leading National
Research Centers

Prof. dr hab. Ewa Broclawik

Kraków, 04-05-2017

RECENZJA ROZPRAWY DOKTORSKIEJ

MGR BARBARY KALINOWSKIEJ

ZATYTUŁOWANEJ: “THE MECHANISM OF PROTEIN FOLDING – IN SILICO MODEL”

Przedstawiona do recenzji rozprawa doktorska wpisuje się w jeden z głównych nurtów aktualnej tematyki badań modelowych nad strukturą i właściwościami białek, jaką jest badanie *in silico* procesu fałdowania prowadzącego do utworzenia ostatecznej struktury trzeciorzędowej, odpowiedzialnej za właściwości strukturalne ale i funkcjonalne białka. Praca pani mgr Barbary Kalinowskiej wykonana została w ramach studiów doktoranckich przy Wydziale Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego pod kierunkiem profesor Ireny Roterman-Koniecznej (Zakład Technologii Informatycznych Wydział Fizyki, Astronomii i Informatyki Stosowanej UJ). Ma ona charakter badań podstawowych, w szczególności dotyczy algorytmicznego aspektu tworzenia teoretycznych narzędzi do przewidywania struktury białka na podstawie sekwencji tworzących to białko aminokwasów a będących istotnym elementem cyklu siedmiu artykułów, przedstawionych jako przedmiot rozprawy. Wobec tego zarówno wybrana forma konstrukcji rozprawy jak i wybór języka angielskiego jako języka prezentacji są w pełni uzasadnione.

Podstawę rozprawy stanowią teksty siedmiu artykułów opublikowanych w czasopismach międzynarodowych lub jako rozdziały w książkach monograficznych; całość opatrzona jest wyczerpującym omówieniem zarówno natury i mechanizmu procesu fałdowania białek jak i istniejących dotychczas oraz nowych, współtworzonych w zespole promotora, modeli służących do przewidywania *in silico* trzeciorzędowej struktury proteiny na podstawie jej sekwencji. Wszystkie prace są wieloautorskie ale załączone oświadczenia nie pozostawiają wątpliwości co do istotności oryginalnego wkładu mgr Kalinowskiej (wynoszącego w pięciu przypadkach 70% lub powyżej).

Praca mgr Barbary Kalinowskiej została wykonywana i jest kontynuowana pod kierunkiem prof. Ireny Roterman-Koniecznej a więc jest w sposób oczywisty związana z tematyką promotora, w szczególności z szeroko zakrojonym projektem budowy efektywnego, wielowymiarowego narzędzia (opartego na numerycznych metodach rachunku prawdopodobieństwa i fizyki statystycznej) służącego do przewidywania struktury białek *in silico*. Zespół prof. Roterman specjalizuje się w tym zagadnieniu od wielu lat i uzyskuje rezultaty wysoko cenione w świecie co gwarantuje jej wysoką pozycję w tej dyscyplinie. Przewidywanie trójwymiarowej struktury białka jest zadaniem niezwykle skomplikowanym gdyż proces powstawania tej struktury jest niezmiernie złożony i nie da się opisać za pomocą jednolitego mechanizmu na wszystkich etapach zaawansowania. Przełom w modelowaniu tego procesu przyniosło dopiero niezależne potraktowanie wczesnego etapu fałdowania, prowadzącego do struktury pośredniej, zwanej strukturą Early Stage (ES), oraz drugiego etapu, prowadzącego do struktury Late Stage (LS), bliskiej strukturze rzeczywistej białka natywnego. We wczesnych etapach strukturyzowania łańcucha polipeptydowego, dobre efekty przynosi przewidywanie procesu na podstawie informacji wyekstrahowanej z dostępnych baz danych dla znanych struktur natywnych (o znanych zarówno sekwencji jak i strukturze). Długoletnie badania, w które ogromny wkład wniosła grupa profesor Roterman-Koniecznej wykazały, że dla drugiego etapu fałdowania (decydującego o ostatecznej strukturze białka), podstawowe znaczenie ma dopasowanie powstającej struktury do optymalnego, teoretycznego rozkładu hydrofobowości, opisanego modelem *Fuzzy Oil Drop* (FOD). Dodatkowo, procedury zbudowane na podstawie tego założenia jednoznacznie wykazały, że powodzenie symulacji struktury końcowej zależy w ogromnym stopniu od dobrego punktu wyjściowego dla drugiego etapu, a więc trafnie przewidzianej struktury wczesnego pośrednika (ES).

Pozwoliłam sobie na to uproszczone podsumowanie tematyki na wstępie mojej recenzji pomimo tego, że nie jestem specjalistą w dziedzinie badania struktur olbrzymich układów molekularnych, w celu wstępnego uzasadnienia mojej bardzo wysokiej oceny pracy pani Barbary Kalinowskiej. Efektywne zebranie, analiza oraz przedstawienie ogromnej ilości istniejącej informacji o relacjach lokalnej sekwencji do struktury dla znanych białek w takiej formie, która pozwoli w efektywny sposób, jak najtrafniej przewidzieć strukturę wczesnego pośrednika (utworzoną na wczesnym etapie fałdowania) dla polipeptydu o znanej sekwencji ale nieznanej strukturze, to zadanie wymagające niezmiernie zaawansowanych narzędzi statystycznych. Bez dobrze przewidzianej struktury ES szanse na powodzenie całego procesu symulacji struktury końcowej maleją a więc istotna rola badań pani Kalinowskiej w uzyskaniu całościowych wyników prezentowanych w omawianych artykułach nie budzi żadnych wątpliwości. Wniesiona przez nią rzetelna wiedza na temat metod analizy zjawisk współzależnych statystycznie, ale opisywanych przez wielość cech jakościowych, ukrytych w informacji zawartej w dużych bazach danych, pozwoliła autorce zbudować odpowiednie narzędzia numeryczne. Nie byłoby to jednak możliwe bez dogłębnego zrozumienia natury procesów podlegających symulacjom czyli najprawdopodobniej dodatkowych studiów nad budową i strukturą białek. Dlatego też nie waham się zaklasyfikować ocenianej rozprawy jako podstawowego osiągnięcia naukowego a nie wyłącznie pracy pomocniczej w sensie zadań techniczno-informatycznych.

Ponieważ głównym elementem rozprawy doktorskiej jest siedem artykułów, opublikowanych w czasopiśmie z listy filadelfijskiej bądź jako rozdział w książce wydanej przez wydawnictwo Springer, a więc po uzyskaniu pozytywnych recenzji od badaczy o uznanej reputacji międzynarodowej, bliższych tym zagadnieniom od recenzenta rozprawy, nie ma potrzeby przedstawiać mojej recenzji merytorycznej całości badań. Skupię się raczej na wkładzie własnym doktorantki w przedstawioną do recenzji rozprawę doktorską, świadczącym o jej przygotowaniu do podjęcia badań naukowych. Część wstępna rozprawy zawiera (na 22 stronach) krótki ale wyczerpujący opis całego procesu fałdowania białka, wraz z analizą kryterium podziału na dwa etapy wymagające różnych form opisu oraz typu modelowania. Wyjaśnione są kryteria, na podstawie których sformułowane zostały hipotezy do budowy zarysu mechanizmu każdego z etapów, rozumiane jako kryteria, na podstawie których tworzone są metody symulacyjne. Podane jest także uzasadnienie dla spodziewanej poprawy efektywności całościowej symulacji struktury przy założeniu o dwóch, zasadniczo różnych podejściach do symulacji struktury wczesnego pośrednika oraz do optymalizacji

struktury końcowej. Autorka nie zaznacza na tym etapie, która część symulacji jest bezpośrednim przedmiotem jej pracy. Poprawia to jednolitość formy tekstu oraz spójność całego opisu ale utrudnia nieco wyodrębnienie i ocenę własnego wkładu merytorycznego doktorantki. Jedyne miejsce, w którym autorka umieściła bardziej szczegółowy opis jej własnego wkładu w prezentowany materiał to wstępny paragraf do krótkiego omówienia składowych artykułów na stronie 22, gdzie stwierdza ona, że skupiała się na krytycznym testowaniu statystycznych narzędzi typu „knowledge based” używanych do przewidywania struktury wczesnego pośrednika (ES) na podstawie informacji w sposób celowy wyselekcjonowanej z baz danych i przetworzonej dla efektywnego wykorzystania w dalszym procesie symulacji. Jako kolejne zadanie własne przedstawia swój udział w stosowaniu analizy rozkładu hydrofobowości w natywnych strukturach protein oraz interpretacji przyczyn odchylenia tego rozkładu od idealnego modelu teoretycznego, w kontekście dodatkowych oddziaływań jako przyczyny samoistnego braku zgodności struktury natywnej z modelową hydrofobowością. Podczas gdy pierwsze z tych zadań oceniałabym jako kluczowe dla powodzenia całego programu badawczego ale w pewnym procencie techniczne, drugie zadanie jest w pełni naukowym zadaniem badawczym

Kolejna część (strony 22-39) to skrótowe omówienie załączonych artykułów, zakończone krótkim podsumowaniem przedstawionych prac oraz całego procesu konstrukcji wieloetapowego narzędzia do całościowego przewidywania struktury białka na podstawie znanej sekwencji aminokwasów, w który to program autorka wniosła i wnosi bardzo istotny wkład. Rozprawę zamykają kopie siedmiu przedmiotowych artykułów oryginalnych oraz oświadczenia współautorów.

We wstępnej części mojej recenzji podkreśliłam już moją wysoką ocenę recenzowanej rozprawy doktorskiej, jednakże do niewątpliwych obowiązków recenzenta należy także szczegółowe wypunktowanie słabszych stron czy niedopowiedzeń rozprawy. Zdecydowanie oświadczam, że nie znalazłam żadnych znaczących uchybień czy istotnych nieścisłości w analizowanym tekście własnym autorki. Nie będę tu oczywiście wymieniała literówek, drobnych niespójności gramatycznych czy prostych przejęzyczeń, które są absolutnie nieznaczące i jest ich bardzo niewiele, zwłaszcza dla tekstu pisanego w języku angielskim.

- Moje główne zastrzeżenie dotyczy pewnej trudności, jaką napotkałam początkowo próbując zdefiniować własną tematykę doktorantki. Jak już wspomniałam, jedyny paragraf (na str. 22), w którym osiągnięcia własne są wprost opisane przez autorkę jest bardzo krótki i może nieco zbyt skromny. Dopiero po przeczytaniu całości rozprawy i jej dogłębnym przemyśleniu zrozumiałam, jak ważną rolę w całości dużego programu kierowanego przez prof. Irenę Roterman-Konieczną odgrywa praca pani Kalinowskiej. Wysoka jakość i poziom zaufania do narzędzi skonstruowanych przez autorkę w znaczącym stopniu zwiększa szanse powodzenia wieloetapowego procesu przewidywania struktury białka na podstawie znanej sekwencji aminokwasów. Rozumiem jednak, że całość rozdziałów pierwszego i drugiego rozprawy czyta się dobrze przez to, że jest napisany w sposób ciągły, spójnym językiem, bez rozbijania na przyczynki własne oraz innych autorów. Czy to jest główną przyczyną, dla której autorka podejmuje ryzyko nie eksponując swojego udziału w części merytorycznej od początku i pozostawiając to częściowo domyślności recenzentów?

- krótka uwaga techniczna: jest siedem artykułów włączonych w rozprawę a nie osiem (jak podano na stronie 2)

- Pytanie: Jak pisze autorka (m.in. na str. 23), istnieje pewna liczba białek, dla których struktura natywna jest bardzo bliska strukturze wczesnego pośrednika. Czy wobec tego dalsza optymalizacja struktury ES zgodnie z modelem *Fuzzy Oil Drop* pogarsza w tym przypadku zgodność struktury uzyskanej w jej wyniku ze strukturą natywną? Jeśli tak, to czy istnieją jakieś przesłanki aby zaniechać tej procedury i pozostać przy strukturze wczesnego pośrednika dla nieznanego białka?

Przedstawione krótkie uwagi, dotyczące raczej pewnych nieścisłości czy nieścisłości, aczkolwiek istotne ze względu na czystość i jasność sformułowań oraz dobre rozumienie meritum, w niczym nie umniejszają mojej wysokiej oceny recenzowanej rozprawy. Uważam, że praca mgr Barbary Kalinowskiej spełnia z nawiązką wymogi stawiane rozprawom doktorskim przez Ustawę z dnia 14 marca 2003 roku o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki oraz z przyjemnością wnoszę o dopuszczenie pani Barbary Kalinowskiej do dalszych etapów przewodu doktorskiego.





Jerzy Haber Institute
of Catalysis and Surface Chemistry
Polish Academy of Sciences



Leading National
Research Centers

Prof. dr hab. Ewa Broclawik

Kraków, 04-05-2017

Wniosek o wyróżnienie rozprawy doktorskiej mgr Barbary Kalinowskiej zatytułowanej “The mechanism of protein folding – in silico model”

Składam wniosek o wyróżnienie rozprawy doktorskiej mgr Barbary Kalinowskiej zatytułowanej “The mechanism of protein folding – in silico model” gdyż stanowi ona duże osiągnięcie w tematyce badań modelowych nad strukturą i właściwościami białek. Ma ona charakter badań podstawowych, w szczególności dotyczy algorytmicznego aspektu tworzenia teoretycznych narzędzi do przewidywania struktury białka na podstawie sekwencji tworzących to białko aminokwasów. Praca ta jest częścią szeroko zakrojonego projektu budowy efektywnego, wielowymiarowego narzędzia (opartego na numerycznych metodach rachunku prawdopodobieństwa i fizyki statystycznej) służącego do przewidywania struktury białek *in silico*.

Procedury symulacji końcowej struktury białka, zbudowane na podstawie założenia *Fuzzy Oil Drop* jednoznacznie wykazały, że jej powodzenie zależy w ogromnym stopniu od dobrego punktu wyjściowego dla drugiego etapu fałdowania (decydującego o ostatecznej strukturze białka), a więc od trafnie przewidzianej struktury wczesnego pośrednika, której poświęcona jest praca mgr Kalinowskiej. Efektywne zebranie, analiza oraz przedstawienie ogromnej ilości istniejącej informacji o relacjach lokalnej sekwencji do struktury dla znanych białek w takiej formie, która pozwoliłaby w efektywny i trafny sposób przewidzieć strukturę wczesnego pośrednika dla polipeptydu o znanej sekwencji ale nieznannej strukturze, to zadanie wymagające bardzo zaawansowanych narzędzi statystycznych. Bez dobrze przewidzianej struktury wczesnego pośrednika, szanse na powodzenie całego procesu symulacji struktury końcowej maleją a więc istotna rola badań pani Kalinowskiej w uzyskaniu całościowych wyników nie budzi żadnych wątpliwości. Bez jej udziału najprawdopodobniej nie powstały by wyniki, stanowiące podstawę siedmiu artykułów, na których oparta jest rozprawa.

Dlatego też oceniam rozprawę jako wybitne osiągnięcie i wnoszę o jej wyróżnienie.