

dr hab. CEZARY CZAPLEWSKI, prof. UG

Gdańsk, 26.05.2017

Recenzja rozprawy doktorskiej mgr Barbary Kalinowskiej

pt.: „The mechanism of protein folding - *in silico* model”

Przedłożona mi do recenzji rozprawa doktorska mgr Barbary Kalinowskiej poświęcona jest badaniom teoretycznym mechanizmu zwijania białek oraz podejmuje próbę zaprojektowania nowej hybrydowej metody przewidywania struktur białek. Wiedza o oddziaływaniach kluczowych dla zwijania białek do struktur natywnych i poznanie mechanizmu zwijania jest niezbędna do pełnego zrozumienia funkcji białek. Błędy w funkcjonowaniu mechanizmu zwijania białek (prowadzące do złego zwinięcia lub agregacji) mogą drastycznie zaburzyć funkcjonowanie całego organizmu, prowadząc do poważnych schorzeń takich, jak choroby prionowe czy amyloidozy (np. choroba Alzheimera). Obecnie jedynie ok. 7% struktur białek o znanej sekwencji jest rozwiązanych metodami eksperymentalnymi. Z uwagi na ogromny przyrost liczby nowo poznawanych sekwencji białek różnica między liczbą sekwencji a liczbą znanych struktur pogłębia się. Powszechnie wykorzystywane eksperymentalne metody określania struktury białek, przede wszystkim krystalografia rentgenowska i magnetyczny rezonans jądrowy wymagają długiego czasu oraz ogromnego nakładu pracy i kosztów. Ponadto, metody eksperymentalne są w stanie dostarczyć jedynie fragmentarycznych informacji o mechanizmie zwijania białek. Dlatego metody teoretyczne stały się powszechnym narzędziem przewidywania struktury i symulacji dynamiki białek. Podjęta przez Doktorantkę tematyka jest aktualna i ważna zarówno pod względem poznawczym, jak i praktycznym.

Praca doktorska mgr Barbary Kalinowskiej powstała pod opieką prof. dr hab. Ireny Roterman-Koniecznej na Wydziale Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego. Rozprawa doktorska ma formę spójnego tematycznie zbioru sześciu artykułów opublikowanych w czasopiśmie naukowych oraz jednego rozdziału w książce:

- A1 I. Roterman, L. Konieczny, M. Banach, D. Marchewska, **B. Kalinowska**, Z. Baster, M. Tomanek, M. Piwowar, Simulation of the Protein Folding Process, in *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes*, pp. 599-638, editor A. Liwo, Heidelberg, Springer (2014). Wkład B.K. 10%
- A2 **B. Kalinowska**, P. Alejster, K. Sałapa, Z. Baster, I. Roterman, Hypothetical in silico model of the early-stage intermediate in protein folding, *J. Mol. Model.* **19**, 4259-4269 (2013). IF=1.867, wkład B.K. 80%
- A3 **B. Kalinowska**, P. Fabian, K. Stapor, I. Roterman, Statistical dictionaries for hypothetical in silico model of the early-stage intermediate in protein folding, *J. Comput. Aided Mol. Des.* **29**, 609-618 (2015). IF=3.199, wkład B.K. 70%
- A4 **B. Kalinowska**, A. Krzykalski, I. Roterman, Contingency Table Browser - prediction of early stage protein structure, *Bioinformation* **11**, 486-488 (2015). Wkład B.K. 70%
- A5 **B. Kalinowska**, M. Banach, L. Konieczny, D. Marchewka, I. Roterman, Intrinsically disordered proteins-relation to general model expressing the active role of the water environment, *Adv. Protein Chem. Struct. Biol.* **94**, 315-346 (2014). IF=3.036, wkład B.K. 70%
- A6 **B. Kalinowska**, M. Banach, L. Konieczny, I. Roterman, Application of divergence entropy to characterize the structure of the hydrophobic core in DNA interacting proteins, *Entropy* **17**, 1477-1507 (2015). IF=1.743, wkład B.K. 70%
- A7 M. Banach, **B. Kalinowska**, L. Konieczny, I. Roterman, Role of Disulfide Bonds in Stabilizing the Conformation of Selected Enzymes-An Approach Based on Divergence Entropy Applied to the Structure of Hydrophobic Core in Proteins, *Entropy* **18**, 1-21 (2016). IF=1.743, wkład B.K. 20%

Procentowy wkład doktorantki do przygotowania opublikowanych artykułów, podany powyżej, jest wykazany w załączonych do rozprawy doktorskiej oświadczeniach o współautorstwie. Dla pięciu artykułów w których mgr Barbara Kalinowska jest pierwszym autorem jest on dominujący. Dla artykułów opublikowanych w *Entropy* wkład poszczególnych autorów jest również podany w samych artykułach: w publikacji A6 mgr Barbara Kalinowska wybrała i wyszukała białka o bezstrukturalnych fragmentach, których dotyczyła publikacja, przeprowadziła analizę ich sekwencji w programie ClustalW oraz przygotowała rysunki trójwymiarowych struktur białek. Natomiast w publikacji A7 jej rola ograniczyła się do

określenia zbioru badanych białek i przygotowania rysunków ilustrujących ich struktury. Nie budzi wątpliwości, że Doktorantka miała udział przy opracowywaniu koncepcji, wykonywaniu części obliczeniowej, opracowaniu i interpretacji wyników badań opublikowanych w zbiorze artykułów przedstawionych do oceny. Z sześciu publikacji przedstawionych do recenzji pięć jest opublikowanych w czasopismach z tzw. Listy Filadelfijskiej a dwie posiadają wysoki współczynnik oddziaływania ($IF > 3$). Mgr Barbara Kalinowska jest także współautorką czterech kolejnych publikacji w czasopismach z Listy Filadelfijskiej, które nie są częścią rozprawy doktorskiej.

Rozprawa doktorska jest przedstawiona w języku angielskim i zgodnie z ustawą jest opatrzona streszczeniem w języku polskim. Jej układ nie budzi zastrzeżeń i odpowiada ogólnie przyjętemu schematowi prezentacji prac naukowych. Oprócz wymienionego wyżej zbioru artykułów oraz oświadczeń współautorów rozprawa zawiera wstęp (6 stron), przedstawienie modelu mechanizmu zwijania białek (11 stron), streszczenie i komentarz do artykułów (12 stron), podsumowanie (4 strony) oraz spis literatury (99 pozycji). Jedyna uwaga do układu pracy to brak osobnego rozdziału definiującego cel pracy. Sam cel: stworzenie i przetestowanie metody przewidywania początkowych struktur na ścieżce zwijania białek (Early Stage) oraz zastosowanie analizy rozkładu hydrofobowości (Fuzzy Oil Drop) dla znanych struktur białek jest przedstawiony w pierwszym akapicie rozdziału 3. *Included papers: summary and comments.*

Rozprawę doktorską otwiera wykaz publikacji, będących podstawą pracy. W pierwsze zdanie wkradł się błąd gdyż lista publikacji liczy siedem a nie osiem pozycji. We wstępie Doktorantka wprowadza czytelnika w tematykę rozprawy doktorskiej, poczynając od znaczenia prowadzonych badań, krótkiego przedstawienia problemu mechanizmu zwijania białek oraz współzależności pomiędzy funkcją i strukturą natywną białek. Następnie opisuje w jakich warunkach zachodzi zwijanie białek *in vivo*. Na kolejnych stronach zostały zwięźle przedstawione oddziaływania decydujące o stabilności struktury natywnej białek. Doktorantka wyróżnia efekt hydrofobowy, wiązania wodorowe, oddziaływania elektrostatyczne i efekty entropowe. Doktorantka zwięźle opisała zaproponowane w literaturze modele mechanizmu zwijania białek, metody symulacji zwijania białek *in silico* oraz metody przewidywania struktury przestrzennej białek.

W rozdziale drugim mgr Barbara Kalinowska opisała zaproponowany wcześniej przez prof. Roterman model procesu zwijania białek będący podstawą badań przedstawionych w recenzowanej rozprawie doktorskiej. Model ten jest szczegółowo omówiony w rozdziale książki A1. Głównym założeniem modelu jest wieloetapowość procesu fałdowania białka,

potwierdzona eksperymentalnie obserwowanymi stanami pośrednimi. Wyróżniono dwa główne etapy: Early Stage i Late Stage. Etap Early Stage wykorzystuje jedynie informacje o lokalnych preferencjach konformacyjnych łańcucha głównego i nie uwzględnia oddziaływań dalekozasięgowych. W celu uproszczenia opisu struktury łańcucha polipeptydowego wprowadzono siedmioliterową klasyfikację konformacji każdej reszty aminokwasowej. Klasyfikacja ta opiera się rzutowaniu mapy Ramachandrana na eliptyczną ścieżkę podzieloną na siedem stref (oznaczonych literami A-G), częściowo związanych z dobrze zdefiniowanymi strukturami drugorzędowymi (kod C odpowiada α -helisie, kod E płaskiej β -kartce, kod F wygiętym końcom β -kartki). W rozdziale książki A1 jest błędnie zacytowana pozycja 22. Publikacja Ivankov, D.N., Finkelstein, A.V.: Prediction of protein folding rates from the amino acid sequence-predicted secondary structure, PNAS 101, 8942-8944 (2004) nie opisuje generowania struktur zgodnie z przewidywanymi strukturami drugorzędowymi ani nie wyjaśnia wprowadzenia eliptycznej ścieżki jako uproszczenia mapy Ramachandrana.

Na podstawie kodów strukturalnych Doktorantka wyznaczyła biblioteki motywów strukturalnych, których prawdopodobieństwa wystąpienia w znanych strukturach wykorzystwała do wyznaczenia najbardziej prawdopodobnych struktur Early Stage. W publikacji A2 przedstawiła metodę przewidywania tych struktury opartą o tabele współzależności prawdopodobieństwa dla tetrapeptydów, natomiast w publikacji A3 zaproponowała bardziej efektywną metodę opartą na słownikach statystycznych nie ograniczonych do czterech reszt aminokwasowych, ale w zależności od dostępnych danych statystycznych wykorzystujących fragmenty od 1 do 13 reszt aminokwasowych. W publikacji A4 przedstawiła narzędzie do wizualizacji relacji pomiędzy strukturą zakodowaną siedmioliterowym kodem a sekwencją badanego białka. Jednym z celów określenia struktury Early Stage jest przewidywanie struktury drugorzędowej (str. 15), jednak jedynie w publikacji A3 porównano wyniki zastosowania metody przewidywania struktury drugorzędowej SPINE-X ze strukturami Early Stage. Struktury Early Stage lepiej przewidują α -helisy i β -kartki, ale metoda SPINE-X jest bardziej wiarygodna przy przewidywaniu struktur kłębka statystycznego (random coil). Dlaczego do porównania wybrano metodę SPINE-X a nie jedną z bardziej popularnych metod jak PSIPRED, JPRED ?

Drugi etap, nazwany Late Stage, zakłada dominujący wpływ oddziaływań hydrofobowych na zwijanie białka. Zgodnie z przyjętym modelem idealny rozkład reszt aminokwasowych w strukturze przestrzennej białka ze względu na ich właściwości hydrofobowe opisuje model Fuzzy Oil Drop (FOD). Model ten zakłada teoretyczny rozkład hydrofobowości w formie trójwymiarowej funkcji Gaussa. Dla znanych struktur białek

wyznaczenie odległości pomiędzy obserwowanym rozkładem hydrofobowości a rozkładem teoretycznym poprzez obliczenie względnej entropii pozwala na ilościową ocenę uformowania jądra hydrofobowego danego białka. W publikacji A5 zastosowano tę analizę do białek z nieuporządkowanymi fragmentami łańcucha i wykazano, że fragmenty te mogą zarówno wspólnie z uporządkowanymi fragmentami białka budować jądro hydrofobowe zgodne z rozkładem teoretycznym jak i naruszać ten rozkład. Szczegółowa analiza pozwala na postawienie hipotez co do mechanizmu fałdowania tych białek, szczególnie dla białek złożonych z kilku domen lub kilku łańcuchów. Podobną analizę przeprowadzono dla białek wiążących DNA w artykule A6. Dla białek z rodziny Jumonji obserwowany rozkład hydrofobowości zgadza się z rozkładem teoretycznym. Istotne dla wiązania DNA β -zwoły mają rozkład hydrofobowości zgodny z teoretycznym dla domen ARID, częściowo zgodny dla domen PHD bez związanych ligandów i różny od teoretycznego dla domen PHD ze związanymi ligandami. W publikacji A7 przeanalizowano współdziałanie efektu hydrofobowego i wiązań disulfidowych jako czynników stabilizujących strukturę białek. Wiązania disulfidowe często stabilizują struktury białek niezgodne z teoretycznym rozkładem hydrofobowości. Najważniejszy wniosek to, że oddziaływania hydrofobowe w połączeniu z wiązaniami disulfidowymi pozwalają na powstanie stabilnych, ale zarazem elastycznych konformacji niezbędnych dla prawidłowej funkcji białek.

W podrozdziale 2.3 przedstawiono pełny schemat symulacji łączący modele Early Stage i Late Stage. Jednak w badaniach opisanych w publikacjach składających się na rozprawę doktorską nie zastosowano go w pełnej formie: publikacje A2-A4 dotyczą tylko modeli Early Stage, a publikacje A5-A7 dotyczą analizy natywnych struktur białek z wykorzystaniem modelu Late Stage. Jedynie w wnioskach w publikacji A2 jest uwaga, że połączone symulacje Early i Late Stage prowadzą do poprawnych struktur. Jak jest efekt zastosowania połączonego modelowania Early Stage i Late Stage, czy testowano taką hybrydową metodę przewidywania struktur przestrzennych białek na ich większym zbiorze? Tu nasuwa się pytanie dlaczego ograniczać się do jednej struktury wynikającej z modelowania początkowego etapu fałdowanie białka (Early Stage)? Przekazanie do kolejnego etapu (Late Stage) kilku możliwych struktur mogłoby zwiększyć efektywność zaproponowanej metody przewidywania struktur białek. Kolejne moje pytanie dotyczy zastosowanej procedury optymalizacji, w której osobno optymalizujemy energię wewnętrzną białka (uwzględniając oddziaływania elektrostatyczne, van der Waalsa i torsyjne), i osobno optymalizujemy strukturę ze względu na efekt hydrofobowy. Dlaczego nie połączyć, z uwzględnieniem odpowiednio

dobrej wagi dla efektu hydrofobowego, wszystkich oddziaływań w jednym kroku optymalizacji ?

Przedstawione powyżej uwagi mają głównie charakter dyskusyjny i nie obniżają wartości naukowej rozprawy. Przedstawione w rozprawie wyniki jednoznacznie pokazują, że mgr Barbara Kalinowska zrealizowała wszystkie postawione sobie cele badawcze. Recenzowana praca doktorska przygotowana została na wysokim poziomie merytorycznym, a zawarte w niej wyniki stanowią oryginalny i istotny wkład do nauki. Z pracy widać, że autorka rozprawy, Pani mgr Barbara Kalinowska świetnie porusza się w obszarze badań.

Reasumując stwierdzam, że przedstawiona mi do recenzji praca doktorska z dużym nadatkiem spełnia formalne wymagania stawiane przez Ustawę o stopniach i tytule naukowym, jak i zwyczajowe warunki stawiane rozprawom doktorskim. Wnoszę do Rady Wydziału Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego o dopuszczenie mgr Barbary Kalinowskiej do dalszych etapów przewodu doktorskiego.

KIEROWNIK
Pracowni Symulacji Polimerów
Cezary Czaplewski
dr hab. Cezary Czaplewski, prof. UG