



UNIWERSYTET
WARSZAWSKI

CeNT CENTRUM
NOWYCH
TECHNOLOGII

6 lipca, 2023

Prof. dr hab. Joanna Trylska
e-mail: joanna@cent.uw.edu.pl
telefon (22) 55 43 600
<http://bionano.cent.uw.edu.pl>

prof. UAM, dr hab. Beata Messyasz
Dziekan Wydziału Biologii
Uniwersytet in. Adama Mickiewicza w Poznaniu
ul. Uniwersytetu Poznańskiego 6
Collegium Biologicum
61-614 Poznań

Recenzja rozprawy doktorskiej mgra Mateusza Dobrychłopa

Rozprawa doktorska mgra Mateusza Dobrychłopa zatytułowana „Modelowanie struktur przestrzennych kompleksów makromolekularnych” została napisana pod kierunkiem prof. dra hab. Janusza Bujnickiego w Instytucie Biologii Molekularnej i Biotechnologii Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu. Temat badawczy rozprawy dotyczy modelowania hybrydowego struktur trójwymiarowych dużych kompleksów białek na podstawie połączenia metod obliczeniowych z danymi doświadczalnymi, wiedzy o sekwencji komponentów kompleksu i ich ewentualnym podobieństwie do innych molekuł. Znajomość struktury przestrzennej z dokładnością pełnoatomową jest niezwykle istotna dla zrozumienia mechanizmu działania kompleksów makromolekularnych oraz dla zaprojektowania cząsteczek to działanie blokujących, czyli inhibitorów. Znalezienie dobrego celu w dużych makrocząsteczkach jest więc istotne z punktu widzenia projektowania leków.

Przedstawiona mi do recenzji rozprawa doktorska jest oparta o wyniki badań autora opublikowanych w dwóch artykułach. W sumie pan Mateusz Dobrychłop jest współautorem trzech artykułów w bardzo dobrych recenzowanych czasopismach (*FEBS Letters*, *Nucleic Acids Research*, *DNA Repair*). Część wyników badań przedstawionych w rozprawie dopiero oczekuje na publikację.

Rozprawa została napisana w języku polskim, ale zawiera także krótkie streszczenie w języku angielskim. Rozprawa wraz z bibliografią liczy prawie 200 stron. Odnośniki literaturowe to 150 pozycji; są to w większości artykuły anglojęzyczne z listy *Journal Citation Reports*. Rozprawę bardzo dobrze się czyta, gdyż jest napisana poprawnym językiem.

Rozprawa ma standardowy układ. Istotną jej częścią jest Wstęp wprowadzający czytelnika w tematykę wyznaczania struktur przestrzennych biomolekuł, zarówno technikami doświadczalnymi jak i teoretycznymi. Autor opisuje w skrócie kompleksy makromolekularne, typu rybosom czy polimerazy DNA i RNA, których struktury przestrzenne zostały wyznaczone, a poznanie na ich podstawie mechanizmu działania zostało nagrodzone w każdym przypadku nagrodą Nobla. W rozdziale Wstęp

pan Mateusz Dobrychłop omawia też główne metody doświadczalne wykorzystywane do wyznaczania struktur biomolekuł: krystalografię promieni X, spektroskopię magnetycznego rezonansu jądrowego i mikroskopię elektronową. Podoba mi się zwłaszcza Tabela 1, w której autor pokazuje zalety i wady poszczególnych metod doświadczalnych. W kolejnych podrozdziałach autor omawia metody modelowania struktur przestrzennych makromolekuł. Dalej opisuje metody doświadczalne, które są pomocne w przewidywaniu oddziaływań między komponentami makrokompleksów, choć nie służą bezpośrednio do wyznaczania ich struktur pełnoatomowych. Te metody to na przykład sieciowanie chemiczne, FRET, mutageniza ukierunkowana. Autor omawia też metody służące do określania kształtu kompleksów, takie jak tomografia elektronowa, czy niskokątowe rozpraszanie promieniowania rentgenowskiego (SAXS). Co istotne opisuje także metody hybrydowe, które polegają na łączeniu danych pochodzących z doświadczeń z modelowaniem. Są to metody, które pan Mateusz Dobrychłop stosował do stworzenia nieznanych modeli przestrzennych dwóch kompleksów. Metody hybrydowe i łączenie danych z różnych źródeł są ładnie podsumowane na Rycinie 2. W tym rozdziale autor podkreśla przełomowe osiągnięcie firmy DeepMind - oprogramowanie AlphaFold 2 – służące do przewidywania struktur biomolekuł, oparte o sieci neuronowe i głębokie uczenie maszynowe.

We wstępnej części rozprawy pan Mateusz Dobrychłop przedstawia też dwa kompleksy makromolekularne, których struktury przestrzenne zostały przez niego wyznaczone w ramach pracy doktorskiej. Jeden kompleks to edytosom, który bierze udział w procesie edycji pre-mRNA u świdrowców (*Trypanosoma brucei*). Kompleks ten zawiera 19 białek i ma masę 0,8 MDa. Drugi badany kompleks to karboksylaza acylo-CoA prątką gruźlicy (*Mycobacterium tuberculosis*), który bierze udział w syntezie kwasu mikołowego u prątków i też zawiera kilkanaście białek.

W kolejnym jednostronicowym rozdziale pan Mateusz Dobrychłop przedstawia cel pracy. Głównym celem badawczym kandydata było stworzenie struktury trójwymiarowej dwóch kompleksów makromolekularnych: edytosomu i karboksylazy acylo-CoA z wykorzystaniem narzędzia PyRy3D oraz danych doświadczalnych. Cel szczegółowy zawierał umożliwienie wizualizacji w programie Chimera danych z narzędzia PyRy3D oraz szczegółowe przetestowanie PyRy3D i porównanie jego działania z innymi podobnymi programami stworzonymi w innych laboratoriach.

Kolejny rozdział to Materiały. W tym rozdziale autor wymienia zasoby obliczeniowe, z których korzystał, kompleksy białek, które wykorzystał do testowania oprogramowania PyRy3D i opisuje formaty danych.

W rozdziale 4 zatytułowanym Metody autor opisuje dane wejściowe do modelowania hybrydowego, przestrzeń symulacji i przeszukiwanie przestrzeni rozwiązań przy nałożonych więzach, funkcję oceny, parametry wejściowe, a także deskryptory wykorzystywane do oceny modeli. Opisuje także program UCSF Chimera, do którego napisał moduł integrujący z narzędziem PyRy3D. W skrócie opisanych jest także sześć innych programów służących do dopasowywania struktur do map gęstości elektronowej, z którymi autor rozprawy porównywał program PyRy3D.

Rozdział 5, Wyniki, dotyczy w pierwszym etapie dokładnego testowania działania oprogramowania PyRy3D, które powstało w laboratorium i pod kierunkiem prof. Bujnickiego. Pan Mateusz Dobrychłop współtworzył to narzędzie i podejścia w nim zawarte. Jego zadaniem było też napisanie modułu tzw.

wtyczki integrującej to narzędzie z programem do wizualizacji Chimera oraz dokładne testy i porównanie działania PyRy3D z innymi dostępnymi narzędziami.

Pierwszy etap obejmował testy PyRy3D na podstawie zestawu testowego 15 kompleksów. Przedstawione porównanie wyników PyRy3D z wynikami uzyskanymi innymi narzędziami jest wnikliwe. Autor wykonał dosyć trudną i żmudną pracę, gdyż musiał instalować i korzystać z oprogramowania innych grup badawczych. Jednak porównanie działania PyRy3D z innymi narzędziami wskazuje, że jest to najlepsze oprogramowanie do stworzenia struktur kompleksów na podstawie danych pochodzących z różnych źródeł. Takie testy powinien przeprowadzać użytkownik nieznający żadnego z testowanych programów, żeby było to sprawiedliwe i najbardziej wiarygodne. Autor rozprawy znał dobrze PyRy3D, bo współtworzył metodę, jednak starał się używać standardowych plików wejściowych, bez wiedzy eksperta. Oprogramowanie PyRy3D dostarczyło najwyższej jakości modele dla największej liczby kompleksów. PyRy3D okazał się więc bezkonkurencyjny, za co należy się pochwała nie tylko autorowi tej rozprawy, ale także wszystkim współautorom tego narzędzia. Myślę, że ta rozprawa jest dobrą i właściwą reklamą narzędzia PyRy3D.

Kolejne dwie niezwykle istotne części wyników dotyczą modelowania struktury przestrzennej dwóch kompleksów makromolekularnych, edytosomu i karboksylazy acetylo-CoA, na podstawie ich map o niskiej rozdzielczości pochodzących z kriomikroskopii elektronowej. Opracowanie modeli, które wykorzystywały dane doświadczalne pozwoliło na dostarczenie nowych informacji na temat działania tych kompleksów makromolekularnych. Pan Mateusz Dobrychłop, stworzył modele strukturalne, mimo braku danych dotyczących niektórych fragmentów. Dla nieznanymi czy nieuporządkowanych regionów autor rozprawy modelował ich elastyczną objętość. Następnie przeanalizował stworzone modele pod kątem ich jakości oraz istotności funkcjonalnych regionów nieuporządkowanych. Na podstawie stworzonego modelu struktury kompleksu edytosomu mógł stwierdzić, że regiony nieuporządkowane znajdują się na powierzchni i odpowiadają najprawdopodobniej za aktywność opiekuńczą względem RNA, a uporządkowane domeny występują wewnątrz struktury i odpowiadają za samą katalizę enzymatyczną. W przypadku karboksylazy acetylo-CoA model przestrzenny pozwolił autorowi na zaproponowanie dwuetapowego mechanizmu reakcji karboksylacji cząsteczki acylo-CoA. Mechanizm ten opiera się na ruchliwości łącznika między dwoma domenami kompleksu. W przypadku obydwu kompleksów okazało się, że modelowanie regionów nieuporządkowanych, nawet w sposób gruboziarnisty, jest niezwykle istotne i zaproponowane przez autora podejście do obszarów nieustrukturyzowanych lub nieznanymi bardzo dobrze się sprawdza.

Po rozdziale Wyniki następuje interesująca dyskusja, która świadczy o niezwyklej dojrzałości badawczej autora rozprawy. Autor podkreśla, że oprogramowanie oparte na uczeniu maszynowym nie jest jeszcze w stanie przewidzieć aranżacji białek i RNA w dużych kompleksach molekularnych, które składają się z wielu komponentów. Wiemy, że dla pojedynczych białek AlphaFold 2 radzi sobie świetnie, ale z układanką różnych białek i kompleksów z kwasami nukleinowymi, które wchodzą w skład maszyn biomolekularnych potrzeba więcej danych do trenowania oprogramowania.

Ostatni rozdział zatytułowany Wnioski zawiera krótkie podsumowanie najważniejszych wyników rozprawy.

Rozprawa jest napisana poprawnie językowo, nie zauważyłam prawie żadnych literówek. Widać staranność autora w pisaniu tekstów. Autor obiektywnie ocenia wady i zalety stosowanych metod. Rysunki i schematy w pracy są schludne i dopracowane merytorycznie. Na przykład Rycina 8 ładnie przedstawia działanie programu PyRy3D. Rysunki strukturalne również są klarowne. Czasem autor umieszcza zbyt długie podpisy pod rysunkami, zwłaszcza w rozdziale Wyniki. Z punktu widzenia recenzenta łatwiej by było oznaczyć białka czy ich fragmenty na samym rysunku, ale rozumiem, że te rysunki były przygotowywane z myślą o publikacji. Na stronie 72 pojawiają się jednak zbyt długie informacje dotyczące tego jak możemy przesuwac i obracać struktury za pomocą myszy w środowisku graficznym i jak wyglądają odpowiednie komendy (Tabela 5). Jednak według mojej opinii rozprawa nie powinna być podręcznikiem użytkownika narzędzia Chimera, a tych parę stron na takie wygląda.

Poniżej wymieniam, jedynie dla porządku, kilka drobnych błędów, które zauważyłam w rozprawie:

- str. 25, środkowy paragraf, powinno być „uzyskała”
- str. 26, ostatnie zdanie, brakuje przecinka przed „które”
- str. 29, sformułowanie „... stosuje się kulki, tym razem najeżone cząsteczkami kalmoduliny.” Brzmi nieco dziwnie, choć pewnie trudno było dobrać lepszy odpowiednik w języku polskim
- str. 49, Rycina 5, powinno być „współczynnik obsadzenia” a nie „osadzenia”
- str. 59, powinno być „jest reprezentowana przez jedną, dwie lub trzy kule”
- str. 155, powinno być „wygenerować model”
- str. 165, powinno być „wprowadzono kolejnych 150 niezależnych symulacji”

Ciekawa też jestem zdania autora dotyczącego kilku zagadnień, które wymieniam poniżej.

Str. 48 Nie jest dla mnie jasne po co robiono testy programu PyRy3D dla kompleksów, dla których mapy gęstości elektronowej były symulowane. Dlaczego nie badano jedynie kompleksów, dla których znane są eksperymentalne mapy gęstości elektronowej? Czy takie porównanie z idealnie wygenerowanymi mapami miało sens?

Str. 62 Jakie było prawdopodobieństwo akceptacji niżej ocenionego modelu? Czy było ono określane metodą prób i błędów?

Str. 87 Czy wybrane do testów porównawczych kompleksy zawierały rejony nieuporządkowane?

Czy można łączyć w jednym łańcuchu model pełnoatomowy z gruboziarnistym w przypadku fragmentów dłuższych niż 30 aminokwasów?

Jak ustalane są wagi dla składowych funkcji oceny?

Czy testowano i porównywano narzędzie PyRy3D dla kompleksów makromolekularnych zawierających RNA lub DNA?

Str. 130 Mapa gęstości elektronowej edytosomu pochodzi z 2009 roku, czy nie było nowszej mapy?

Jakie autor ma pomysły na zablokowanie działania edytosomu na podstawie wyznaczonej przez siebie struktury tego kompleksu?

Program AlphaFold 2 dobrze przewiduje struktury białek i pewnie jego odpowiednik dla kwasów nukleinowych pojawi się niedługo, gdy tylko będzie dostępnych więcej struktur RNA pozwalających na wyuczenie tego oprogramowania. Jednak wyznaczenie struktur przestrzennych kompleksów kilkudziesięciu białek i kwasów nukleinowych nie jest łatwym procesem i wymaga integracji wielu metod doświadczalnych i obliczeniowych. Taką integrację wielu metod przeprowadził mgr Mateusz Dobrychłop w ramach rozprawy doktorskiej. Wymagało to dopasowywania struktur pojedynczych komponentów układu do różnych map gęstości elektronowej, wykorzystania więzów NMR, czy wiedzy o regionach nieuporządkowanych. O ile metody wyznaczania struktur przestrzennych pojedynczych białek są już w zasadzie ogólnodostępne i problem można uznać za rozwiązany, to przewidywanie struktur ich kompleksów wymaga jeszcze sporo pracy. Istotną pracę badawczą w tym kierunku wykonał właśnie pan Mateusz Dobrychłop, zarówno jeśli chodzi o narzędzia jak i stworzenie modeli przestrzennych dwóch istotnych makrokompleksów. Pracę tę doceniam i uważam, że badania pana Mateusza Dobrychłopa znacząco przyczyniły się do rozwoju metod hybrydowych przewidywania struktur kompleksów makromolekularnych.

Podsumowując stwierdzam, że rozprawa doktorska mgra Mateusza Dobrychłopa stanowi oryginalne rozwiązanie problemu naukowego, potwierdza jego wiedzę teoretyczną w dyscyplinie nauk biologicznych oraz umiejętność samodzielnego prowadzenia pracy naukowej. Przedstawiona mi do oceny rozprawa doktorska spełnia warunki określone w art. 13 ust. 1 Ustawy z dnia 14 marca 2003 roku o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki. W związku z tym wnoszę o dopuszczenie mgra Mateusza Dobrychłopa do dalszych etapów przewodu doktorskiego.

J. Trzebska