

Prof. dr hab. Cezary Czaplewski
Wydział Chemii
Uniwersytet Gdański

Gdańsk 14.07.2023

Recenzja rozprawy doktorskiej mgr Mateusza Dobrychłopa
pt. „Modelowanie struktur przestrzennych kompleksów makromolekularnych”

Przedstawiona do recenzji praca doktorska mgr Mateusza Dobrychłopa została wykonana w Instytucie Biologii Molekularnej i Biotechnologii Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu pod kierunkiem prof. dr hab. Janusza Bujnickiego. Tematyka recenzowanej rozprawy doktorskiej jest jednym z wielu wątków badań dotyczących zależności między sekwencją, strukturą i funkcją makrocząsteczek i ich kompleksów jakie są prowadzone w zespole kierowanym przez prof. Janusza Bujnickiego. Wykorzystany w pracy doktorskiej program PyRy3D do modelowania struktur przestrzennych kompleksów makromolekularnych został wcześniej opracowanych w grupie prof. Janusza Bujnickiego, a mgr Mateusz Dobrychłop jest współautorem tego programu, w szczególności jest autorem wtyczki PyRy3D do programu UCSF Chimera. Praca doktorska mgr Mateusza Dobrychłopa poświęcona jest modelowaniu dwóch dużych kompleksów makromolekularnych: edytosomu *Trypanosoma brucei* oraz karboksylazy acylo-CoA z *Mycobacterium tuberculosis*. Dodatkowo doktorant wykonał testy programu PyRy3D i sześciu konkurencyjnych programów dla zestawu 15 kompleksów białkowych o znanych strukturach. Badania przedstawione w pracy doktorskiej mają charakter badań podstawowych, jednak poznanie struktur obu modelowanych kompleksów może mieć w znacznie aplikacyjne. Proces edycji pre-mRNA realizowany przez edytosom *Trypanosoma brucei* jest obiecującym celem dla nowych leków na śpiączkę afrykańską i inne groźne choroby wywoływane przez pasożytnicze pierwotniaki z rodziny świrdowców. Z kolei poznanie struktury karboksylazy acylo-CoA z *Mycobacterium*

tuberculosis może pozwolić na zrozumienie procesu syntezy kwasów mikołowych i zaprojektowanie nowej generacji leków przeciwgruźliczych będących inhibitorami syntezy ściany komórkowej. Wykonane testy najnowszej wersji programu PyRy3D pokazują, że jest to skuteczne narzędzie do modelowania struktury kompleksów makrocząsteczek na podstawie map gęstości elektronowej z uwzględnieniem fragmentów o nieuporządkowanej strukturze przestrzennej.

Recenzowana rozprawa doktorska ma typowy układ prac doktorskich w dziedzinie nauk ścisłych i przyrodniczych. Zawiera siedem głównych rozdziałów poprzedzonych wykazem stosowanych skrótów i streszczeniem, a zakończonych cytowaną literaturą (150 pozycji). Pierwszy rozdział zatytułowany *Wstęp* (33 strony) zgodnie z tytułem wprowadza czytelnika w tematykę rozprawy doktorskiej i spełnia tę rolę bardzo dobrze. Omówione są najważniejsze kompleksy molekularne o znanych strukturach: rybosom, polimerazy DNA i RNA, spliceosom i kompleksy poru jądrowego. Krótko przedstawione są doświadczalne metody określania struktury przestrzennej białek i kwasów nukleinowych: krystalografia rentgenowska, spektroskopia magnetycznego rezonansu jądrowego, kriomikroskopia elektronowa. Metody są omówione zwięźle co nie jest wadą, ale w przedstawieniu niektórych zabrakło istotnych szczegółów. W omówieniu krystalografii rentgenowskiej doktorant nie napisał o konieczności rozwiązania problemu fazowego przy przejściu od rejestrowanych obrazów dyfrakcyjnych do map gęstości elektronowej. Dla makrocząsteczek stosuje się trzy metody rozwiązania problemu fazowego: podstawienia izomorficznego, podstawienia molekularnego oraz metodę dostrojonej dyfrakcji anomalnej (MAD). Problem fazowy jest jednym z ograniczeń krystalografii rentgenowskiej makrocząsteczek. W opisie spektroskopii magnetycznego rezonansu jądrowego powinien być wymieniony jądrowy efekt Overhausera (NOE) jako główne źródło informacji o wzajemnym położeniu jąder atomowych w cząsteczce. W dalszej części wstępu omówiono komputerowe metody modelowania struktur przestrzennych białek i kwasów nukleinowych: metody *de novo*, modelowanie homologiczne oraz metodę AlphaFold wykorzystującą sieci neuronowe i uczenie głębokie. Omówiona jest inicjatywa Critical Assessment of protein Structure Prediction (CASP) pozwalająca na obiektywną ocenę metod przewidywania struktur przestrzennych białek. W tym miejscu spodziewałbym się też wspomnienia o Critical Assessments of Predictions of Interactions (CAPRI). CASP dotyczy zarówno przewidywania struktury przestrzennej białek jak i ich kompleksów jednak to CAPRI jest ukierunkowane na ocenę metod

komputerowych przewidywania struktury kompleksów i oddziaływań pomiędzy makrocząsteczkami. We wstępie przedstawiono też doświadczalne metody badania oddziaływań między komponentami kompleksów molekularnych: sieciowanie chemiczne, FRET, mutagenezę ukierunkowaną, drożdżowy system dwyhybrydowy, tandemową chromatografię powinowactwa a także doświadczalne metody badania kształtu kompleksów makromolekularnych pozwalające na uzyskanie informacji o niższej rozdzielczości: tomografię elektronową, niskokątowe rozpraszanie promieniowania rentgenowskiego i neutronów. Krótko przedstawiono metody komputerowe dopasowania struktur atomowych do map gęstości elektronowej. Wstęp kończy omówienie budowy i funkcji obu modelowanych kompleksów: edytosomu *Trypanosoma brucei* oraz karboksylazy acylo-CoA z *Mycobacterium tuberculosis*.

Cel pracy jest jednoznacznie postawiony w rozdziale drugim: podstawowym celem było stworzenie modeli przestrzennych dwóch kompleksów makromolekularnych: edytosomu *Trypanosoma brucei* oraz karboksylazy acylo-CoA z *Mycobacterium tuberculosis*, a dodatkowym analiza działania programu PyRy3D i porównanie go z konkurencyjnymi metodami komputerowymi. Rozdział trzeci *Materiały* (7 stron) przedstawia zasoby obliczeniowe, dane wykorzystane w testowaniu programu PyRy3D (zestaw 15 kompleksów białkowych o znanych strukturach przestrzennych) oraz formaty plików (PDB, FASTA, CCP4, Filter3D) wykorzystane w obliczeniach. Rozdział czwarty *Metody* (33 strony) omawiający metody obliczeniowe stosowane w pracy doktorskiej przedstawia tylko najważniejsze zagadnienia: program PyRy3D (dane wejściowe, ich reprezentacja w programie, przeszukiwanie przestrzeni rozwiązań, funkcję oceny, parametry), programy UCSF Chimera i UCSF Chimera X wykorzystane do modelowania i wizualizacji. W opisie symulacji objętości fragmentów elastycznych w programie PyRy3D są wymienione trzy metody wyznaczania pozycji pseudoreszt: krótki (<30 reszt) łańcuch na końcu lub początku białka, łańcuch zakotwiczony wewnątrz białka oraz reszty w formie grona dla dłuższych elastycznych fragmentów. Nie opisano, jak jest realizowana struktura grona, czy ta forma nie uwzględnia łańcucha polimeru? Jako miary wykorzystywane do oceny modeli opisano średnie odchylenie kwadratowe (RMSD) i średnie odchylenie kwadratowe dla odległości międzyatomowych (dRMSD) oraz współczynnik korelacji krzyżowej (CCC). Nie zdefiniowano precyzyjnie miary dRMSD mimo, że jest to główna miara wykorzystywana w pracy. Zabrakło też uzasadnienia wyboru tej miary do oceny podobieństwa

struktur dużych kompleksów makromolekularnych. Jakie inne miary są wykorzystywane do porównania podobieństwa struktur kompleksów? W pracy przy omawianiu eksperymentów CASP przedstawiona miarę GDT_TS, czy mogłaby ona zastąpić dRMSD? W dalszej części rozdziału są przedstawione konkurencyjne względem PyRy3D programy do dopasowania struktur atomowych do map gęstości elektronowej: Multifit, gamma-TEMPy, Situs, ADP_EM, Powerfit, gEMFitter. Pomocne w porównaniu tych narzędzi byłoby zestawienie pokazujące metodę przeszukiwania i funkcję celu wykorzystywaną przez każdy z programów.

Wyniki przedstawione są w rozdziale piątym *Wyniki* na 90 stronach, który podzielony jest na trzy podrozdziały omawiające kolejno: testy programu PyRy3D dla 15 kompleksów białek o znanych strukturze i porównanie z innymi narzędziami, modelowanie struktury przestrzennej edytosomu *Trypanosoma brucei* oraz karboksylazy acylo-CoA z *Mycobacterium tuberculosis*. Program PyRy3D stosuje podejście hybrydowe do budowania modeli strukturalnych kompleksów makrocząsteczek i pozwala na integracje danych o strukturze i oddziaływaniach pomiędzy składowymi kompleksów. Testy programu PyRy3D wykonane przez mgr Mateusza Dobrychłopa dotyczyły modelowania struktury kompleksów na podstawie map gęstości elektronowej zarówno eksperymentalnych (o rozdzielczości od 4,2 do 12,6Å dla 5 testowanych kompleksów) jak i symulowanych na podstawie znanej struktury przestrzennej (dla 10 kompleksów wygenerowano mapy o rozdzielczości 20Å). Dla 13 z 15 symulowanych kompleksów uzyskano modele, w których komponenty zajmują prawidłowe rejony map gęstości elektronowej. Bardzo dobre (o dRMSD poniżej 3Å) i dobre (o dRMSD pomiędzy 3 a 6Å) modele, w których zarówno rejony jak i pozycje zajmowane przez białka budujące kompleks są prawidłowe uzyskano odpowiednio dla 5 i 8 z 15 symulowanych w PyRy3D układów. W porównaniu z konkurencyjnymi narzędziami program PyRy3D wygenerował najwyższej jakości modele dla 6 z 15 testowanych kompleksów, co jest najlepszym wynikiem ze wszystkich testowanych metod. Dla modeli wygenerowanych przez PyRy3D średnie dRMSD jest jedynie o 5,26Å gorsze od najlepszych modeli ze wszystkich testowanych metod. Przeprowadzone testy symulacji objętości regionów elastycznych pokazały, że metoda ta pozwala na uzyskanie lepszych wyników niż pominięcie składowych kompleksów o nieznannej strukturze. We wszystkich opisanych w pracy symulacjach wykorzystywano model pełnoatomowy (fa), czy były prowadzone testy pokazujące jak na wyniki wpływa zastosowanie uproszczonych modeli dostępnych w programie PyRy3D? Dla

modeli symetrycznych struktur kompleksów niższa jakość modeli PyRy3D była w kilku przypadkach związana z nieprawidłową orientacją jednego z komponentów, w jaki sposób można rozszerzyć program PyRy3D o możliwość uwzględniania symetrii?

Model struktury przestrzennej edytosomu *Trypanosoma brucei* zbudowany przez mgr Mateusza Dobryłopa w programie PyRy3D opiera się na danych literaturowych: mapie gęstości elektronowej o rozdzielczości 22Å, modelach komputerowych 15 składowych białek oraz 16 więzach odległościowych między składowymi białkami (7 ustalonych w eksperymentach sieciowania chemicznego, 8 na podstawie eksperymentów drożdżowego systemu dwyhybrydowego, ostatni z więzów wynikał z podziału białka TbMP99 na dwa fragmenty). W modelowaniu uwzględniono regiony nieuporządkowane. Wynikowy model edytosomu zgadza się z wynikami uzyskanymi z transmisyjnej mikroskopii elektronowej zgodnie z którymi miejsce wiązania cząsteczek RNA przez białka zawierające domeny OB-fold zlokalizowane jest w centralnej części kompleksu. Przeprowadzona przez doktoranta analiza rozmieszczenia regionów nieuporządkowanych pokazuje, że prawie połowa wszystkich reszt na powierzchni kompleksu to reszty z regionów nieuporządkowanych, regiony uporządkowane są zlokalizowane wewnątrz edytosomu.

Model struktury przestrzennej karboksylazy acylo-CoA z *Mycobacterium tuberculosis* zbudowany przez mgr Mateusza Dobryłopa w programie PyRy3D wykorzystuje mapę gęstości elektronowej o rozdzielczości 24Å dostarczoną przez grupę prof. Andrzeja Dziembowskiego z IBB PAN, model białka AccA3 uzyskany z programu AlphaFold2 oraz model heteroheksameru AccD4/AccD5 opracowany przez dr Joannę Kasprzak. Białko AccA3 zostało podzielone w programie PyRy3D na trzy fragmenty: N-kończącą domenę BC oraz C-kończącą domenę BCCP połączone elastycznym łącznikiem. Model karboksylazy został opracowany w czterech etapach: wstępne dopasowanie kompleksu AccA3 i AccD4/AccD5 w programie PyRy3D, udokładnienie dopasowania ustrukturyzowanych regionów białka AccA3 w PyRy3D, symulacja objętości łącznika między domenami BC i BCCP białka AccA3 w PyRy3D oraz budowa pełnego modelu karboksylazy acylo-CoA w programie UCSF Chimera.

Rozdział szósty *Dyskusja* (10 stron) zgodnie z tytułem dyskutuje uzyskane wyniki w świetle funkcji biologicznych obu modelowanych kompleksów oraz jakości modeli generowanych przez program PyRy3D. Cenna dla innych użytkowników programu PyRy3D jest dyskusja zależności

zalecanej długości symulacji od liczby niezależnych komponentów kompleksu. Model edytosomu sugeruje jego dwuczęściową budowę: zewnętrzną nieuporządkowaną powierzchnię oraz rdzeń zbudowany z uporządkowanych domen. Nieuporządkowana powierzchnia edytosomu odpowiada za początkowy kontakt z RNA i funkcję opiekuńczą a uporządkowany rdzeń pełni funkcję katalityczną edytosomu. Taki opis jest zgodny z wynikami eksperymentalnymi dla aktywności opiekuńczej względem RNA poszczególnych składowych edytosomu. Model karboksylazy acylo-CoA pozwala na zaproponowanie strukturalnych podstaw mechanizmu reakcji karboksylacji biotyny w miejscu aktywnym domeny BC białka AccA3. Kluczową rolę w mechanizmie odgrywa elastyczny łącznik między domenami BC i BCCP. W pierwszym etapie reakcji, domena BCCP z przyłączoną biotyną przemieszcza się w okolice miejsca aktywnego domeny BC. Po karboksylacji biotyny domena BCCP przemieszcza się ponownie oddziałując z białkami AccD4 i AccD5 co prowadzi do przeniesienia grupy karboksylowej z biotyny na cząsteczkę acylo-CoA.

Ostatni rozdział *Wnioski* (2 strony) stanowi zwięzłe podsumowanie przeprowadzonych badań. Przeprowadzone badania wykazały, że program PyRt3D skutecznie buduje modele kompleksów na podstawie danych doświadczalnych i teoretycznych różnego rodzaju, metoda symulacji objętości regionów elastycznych znacznie poprawia wyniki dla kompleksów o nieokreślonej strukturze niektórych fragmentów. Zbudowane modele edytosomu i karboksylazy acylo-CoA są zgodne z danymi doświadczalnymi i pozwalają na wyciągnięcie wniosków na temat mechanizmów działania tych kompleksów.

Przedstawione w rozprawie wyniki jednoznacznie pokazują, że mgr Mateusz Dobrychłop zrealizował wszystkie postawione sobie cele badawcze. Szczegółowa dyskusja osiągniętych wyników, precyzyjne wnioski znamionują osobę przygotowaną do prowadzenia tego typu badań naukowych. Recenzowana praca doktorska przygotowana została na wysokim poziomie merytorycznym, a zawarte w niej wyniki stanowią oryginalny i istotny wkład do nauki. Przedstawione powyżej uwagi mają charakter dyskusyjny i nie obniżają wartości naukowej rozprawy.

Mgr Mateusz Dobrychłop jest współautorem dwóch publikacji związanych bezpośrednio z tematyką rozprawy doktorskiej: A. Czerwoniec, J.M. Kasprzak, P. Bytner, M. Dobrychłop, J.M. Bujnicki, Structure and intrinsic disorder of the proteins of the Trypanosoma brucei editosome, *FEBS Letters* **589**, 2603-2610 (2015) i C. Voigt, M. Dobrychłop, E. Kruse, A. Czerwoniec, J.M. Kasprzak,

P. Bytner, C. Del Campo, W.-M. Leeder, J.M. Bujnicki, H.U. Göringer, The OB-fold proteins of the *Trypanosoma brucei* editosome execute RNA-chaperone activity, *Nucleic Acids Res* **46**, 10353-10367 (2018) oraz jednego artykułu o tematyce nie związanej z doktoratem: M. Jaciuk, P. Swuec, V. Gaur, J.M. Kasprzak, L. Renault, M. Dobrychłop, S. Nirwal, J.M. Bujnicki, A. Costa, M. Nowotny, A combined structural and biochemical approach reveals translocation and stalling of UvrB on the DNA lesion as a mechanism of damage verification in bacterial nucleotide excision repair, *DNA Repair* **85**, 102746 (2020).

Reasumując stwierdzam, że przedstawiona mi do recenzji praca doktorska spełnia formalne wymagania stawiane przez Ustawę o stopniach i tytule naukowym jak i zwyczajowe warunki stawiane rozprawom doktorskim. Wnoszę do Rady Dyscypliny Nauki Biologiczne Uniwersytetu im. Adama Mickiewicza w Poznaniu o dopuszczenie mgr Mateusza Dobrychłopa do dalszych etapów przewodu doktorskiego.

Cezary Czaplowski