

Gdańsk, dnia 3.02.2026

Recenzja rozprawy doktorskiej
Pana mgra Bartosza Ambrożego Grenia
pt. „Topological analysis of biomolecules”

Tworzenie przez niskocząsteczkowe związki organiczne oligomerów o topologii łańcuchów (katenanów) i łańcuchów przewleczonych przez pierścień (rotaksanów) oraz pojawianie się w ich strukturach przestrzennych węzłów jest znane co najmniej od lat 60-tych XX wieku, natomiast dopiero ok. 30 lat temu stwierdzono, że struktury o topologii węzłów występują również w białkach a wcześniej wykryto je w strukturach DNA. W dalszych badaniach stwierdzono występowanie w strukturach białek i DNA topologii pokrewnych: pseudowęzłów, przewleczek, lass, krzywych θ , łańcuchów i „kajdanków”. Obecność tych struktur ma duże znaczenie w funkcjonowaniu białek, w szczególności niektórych enzymów. Do ich identyfikacji i klasyfikacji zastosowano już istniejące i opracowano nowe narzędzia matematyczne. Opracowywaniu matematycznych metod wykrywania i klasyfikacji struktur zawężlonych i ogólnie splotów lub splątków, ich implementacji oraz analizie eksperymentalnych i wymodelowanych struktur białek i RNA jest poświęcona rozprawa doktorska Pana mgra Bartosza Ambrożego Grenia, która została wykonana pod kierunkiem prof. Joanny Sułkowskiej z Centrum Nowych Technologii Uniwersytetu Warszawskiego, która jest jednym z wiodących badaczy skomplikowanych topologii makromolekuł biologicznych w skali światowej.

Przedstawiona mi do recenzji rozprawa doktorska składa się ze spinki jednego preprintu zdeponowanego w arXiv oraz ośmiu publikacji, które ukazały się w międzynarodowych czasopismach naukowych poprzedzonej streszczeniem, słowami kluczowymi oraz częścią wprowadzającą. Część wprowadzająca składa się ze wstępu (rozdział 1), podsumowania zawierającego cele rozprawy i wnioski (rozdział 2), krótkich opisów artykułów naukowych będących rdzeniem rozprawy (rozdział 3) i bibliografii. Cała rozprawa oprócz słów kluczowych, które zostały podane w języku angielskim i polskim, została napisana w języku angielskim.

Wszystkie publikacje składające się na rozprawę są wieloautorskie. Wieloautorskość prac w naukach przyrodniczych jest konsekwencją konieczności zaangażowania i współdziałania wielu badaczy. 8 już opublikowanych artykułów tworzących spinę ukazało się w dobrych i bardzo dobrych czasopismach naukowych: *Nucleic Acids Research* (prace 2 i 4), *Briefings in Bioinformatics* (praca 8), *Journal of Molecular Biology* (praca 6), *PLOS Computational*

Biology (praca 7), *Biophysical Journal* (praca 3), *Protein Science* (praca 5) i *Polymer* (praca 1). Kandydat jest pierwszym autorem preprintu oraz prac 1 i 7. Przy omawianiu poszczególnych publikacji w rozdziale 3 Kandydat deklaruje swoje wkłady do preprintu oraz publikacji 2-8, natomiast w przypadku publikacji 1 wkład ten jest podany w sekcji „Author contributions” tej publikacji, poza tym Kandydat jest pierwszym jej autorem. Ze wspomnianych deklaracji wynika, że wkłady Kandydata do wszystkich artykułów naukowych składających się na rozprawę były wiodące lub bardzo istotne.

Cele rozprawy są podane w pierwszym akapicie podrozdziału 2.1 rozdziału 2 zatytułowanego „Summary”. Pierwszym z nich było pogłębienie wiedzy na temat złożonych motywów topologicznych w białkach i RNA, które powtarzają się w wielu rodzinach tych makromolekuł biologicznych i mają znaczenie dla ich funkcjonowania. Drugim celem było opracowanie narzędzi matematycznych i obliczeniowych do znajdowania oraz analizy topologii splotów w makromolekułach biologicznych. W dalszej części sekcji 2.1 Kandydat podaje, który z celów lub jego elementów i w jakim stopniu został zrealizowany w poszczególnych publikacjach. Te bardziej szczegółowe opisy celów rozprawy zawierają również krótkie podsumowania osiągniętych wyników. Podrozdział 2.1 powinien być zatem zatytułowany „The aim of the thesis and summary of results”. Najlepiej by było zastąpić ten rozdział dwoma oddzielnymi rozdziałami poświęconymi odpowiednio celom oraz podsumowaniu i wnioskowi.

Rozdział 1 składa się z dwóch podrozdziałów, które są dalej podzielone na odpowiednio 6 i 8 sekcji. Sekcja 1.1.1 umieszcza węzły i ogólnie sploty w kontekście topologii, definiując relację równoważności między nimi jako izotopię otoczenia, podaje początki teorii węzłów w pracach Maxwella i Thompsona, zastosowanej wtedy odpowiednio do zagadnień elektrodynamiki i budowy atomu, jak również uogólnienie węzłów jako rozmaitości $n-2$ wymiarowej w przestrzeni n -wymiarowej. W sekcji 1.1.2 zdefiniowane są ruchy Rademeistera i wielomianowe niezmienniki węzłów. Jest tam również wprowadzone pojęcie diagramu węzła. Sekcje 1.1.3 – 1.1.5 zawierają kolejno opisy węzłów otwartych oraz innych splotów: „przewleczek” (slipknots), lass, krzywych θ i „kajdanek”. W sekcji 1.1.6 są zdefiniowane sploty podstawowe, które są elementami struktur wymienionych powyżej.

Podrozdział 1.1 jest napisany klarownie, jakkolwiek jest raczej adresowany do czytelników przynajmniej trochę zorientowanych w teorii węzłów oraz posiadających przynajmniej podstawową wiedzę na temat topologii algebraicznej. Tym niemniej, Kandydat mógłby podać krótkie wyjaśnienie diagramów występujących w równaniach 1.1 i 1.2, jak również tych wylistowanych na początku podrozdziału 1.6. Powinny być zdefiniowane wielkości γ_1 , γ_2 , $d\gamma_1$ i $d\gamma_2$, które występują w nienumerowanym równaniu na liczbę połączeń gaussowskich (Gaussian Linking Number). Z definicji zamieszczonej na stronie CENT poświęconej węzłom, pod adresem [https://knotgenom.cent.uw.edu.pl/gln_calc#:~:text=The%20Gaussian%20Linking%20Number%20\(GLN\),winds%20around%20the%20other%20one](https://knotgenom.cent.uw.edu.pl/gln_calc#:~:text=The%20Gaussian%20Linking%20Number%20(GLN),winds%20around%20the%20other%20one) wynika, że te wielkości odpowiadają odpowiednio współrzędnym kartezyjańskim wzdłuż dwóch analizowanych krzywych oraz ich

infinitesimalnym przyrostom, zatem są to wielkości wektorowe i że w równaniu powinien występować iloczyn wektorowy dy_1 i dy_2 w iloczynie skalarnym z (wektorową) funkcją podcałkową.

Podrozdział 1.2 jest poświęcony węzłom i innym skomplikowanym topologicznie strukturom występującym w makromolekułach biologicznych. W sekcji 1.2.1 podana jest krótka historia odkrycia węzłów w cyklicznych strukturach bakteryjnych DNA, natomiast historia badań nad strukturami białek zawierającymi węzły i pokrewne struktury jest opisana w sekcji 1.2.2. Kandydat nawiązuje również do analizy struktur przewidzianych przez AlphaFold. Pozostałe sekcje podrozdziału 2.1 opisują typy węzłów, lass, krzywych θ i „kajdanek” znalezionych w strukturach białek oraz węzłów w strukturach RNA, do odkrycia pierwszej z których przyczynił się Kandydat.

Oprócz preprintu wszystkie załączone artykuły naukowe zostały ocenione pozytywnie przez recenzentów przed dopuszczeniem do publikacji, zatem można przyjąć, że są poprawne merytorycznie. Dlatego poniżej omawiam szerzej jedynie preprint, natomiast w przypadku pozostałych publikacji ograniczam się do podsumowania opisanych w nich najważniejszych wyników, do których uzyskania Kandydat wniósł istotny wkład. Ponieważ nie jestem z wykształcenia matematykiem, moje uwagi odnośnie preprintu dotyczą głównie klarowności tekstu.

Preprint jest poświęcony opracowaniu nowych metod klasyfikacji splotów algebraicznych. Wkład Kandydata do opisanych tam badań była dominujący. Oprócz Promotorki współautorem jest jedynie dr Gabrošek z Uniwersytetu w Ljubljanie.

Sekcja 2 preprintu definiuje pojęcia podstawowe. W tej części szczegółowo opisana jest sfera Conwaya, typy podstawowych splotów, ich reprezentacja w postaci diagramów oraz algebraiczna w postaci binarnego drzewa, która stanowi wkład Kandydata do teorii splotów, operacje symetrii na splotach, ruchy Rademeistera oraz dodawanie i mnożenie splotów. Tekst sekcji 2 jest w większości jasny, jednak zaważyłem drobne niezręczności, które wyszczególniam poniżej:

- (a) Diagramy przedstawiające podstawowe sploty, do których odnosi się definicja 4 określa dopiero definicja 8. Uważam, że ten porządek powinien być odwrotny tym bardziej, że definicja 8 łącznie z rysunkiem 1c określa również operacje dodawania i mnożenia splotów.
- (b) Wydaje się, że według definicji 8 diagramy czterech podstawowych splotów są przedstawione na rysunku 1b. Natomiast operacje ich dodawania i mnożenia (rysunek 1c) oraz operacje symetrii na splotach (rysunek 3) są zilustrowane przy użyciu innych diagramów, z których każdy jest kwadratem o dwóch stronach („dodatniej” i „ujemnej”) z czterema liniami odchodzącymi od wierzchołków wzdłuż przekątnych. Wydaje się, że przedstawienie splotu oparte na kwadracie nie umożliwia odróżnienia splotów achiralnych od chiralnych. Być może określają to symbole R i F w środku kwadratu ale

nie są one wyjaśnione. Na pewno byłoby pomocne zilustrowanie definicji 8 odpowiednim rysunkiem oraz jednoznaczne objaśnienie diagramów opartych na kwadracie.

- (c) Bardziej logiczne byłoby omówienie operacji symetrii na pojedynczym splocie przed omówieniem dodawania i mnożenia splotów.
- (d) W opisie rysunku 3 powinno być zaznaczone, że zmiana znaku współrzędnej z jest przedstawiona jako zacienienie kwadratu.
- (e) Rysunek 1c sugeruje, że dodawanie splotów polega na połączeniu prawych końców splotu po lewej (R) z lewymi końcami splotu po prawej (F) a mnożenie splotów na połączeniu prawych końców splotu po lewej (R) z lewymi końcami splotu po prawej (F), po uprzednim odbiciu R w płaszczyźnie zp (efekt działania tej drugiej operacji jest pokazany na rysunku 3b). Jeżeli tak to $RF = \eta(R) + F$ i w takim przypadku warto by było przytoczyć tę zależność po omówieniu operacji symetrii tym bardziej, że w sekcji 3 jest przedstawiona bardzo interesująca dyskusja kombinacji operacji symetrii i dodawania oraz mnożenia. Zgodnie z definicją 4, rysunek 1c jest jedynym miejscem, gdzie jest określone mnożenie splotów. Zatem jeżeli nie jest ono złożeniem podanych powyżej operacji należałoby opisać je dokładniej.
- (f) W definicji 5 (splotu algebraicznego) stwierdza się, że każdy splot można wyrazić jako iloczyn dwóch splotów, co jest pokazane dopiero w sekcji 3. Poza tym, jaki jest związek symboli L i R pojawiających się w tej definicji z symbolami R i F z rysunku 1c?

Sekcja 3 artykułu opisuje operacje na podstawowych splotach. Kandydat udowadnia, że dodawanie splotów można wyrazić przez mnożenie tych splotów z włączeniem elementarnego splotu neutralnego („0”) do operacji. Podaje również działanie kombinacji operacji symetrii z operacjami dodawania i mnożenia. Zdefiniowane są również n -sploty otrzymane przez dodanie do siebie n razy elementarnych splotów typu 1 lub o przeciwnej chiralności, jak również tworzenie bardziej skomplikowanych splotów w wyniku kombinacji operacji dodawania/mnożenia i symetrii. Ta sekcja jest napisana jasno a jej materiał jest przedstawiony w sposób wzbudzający zainteresowanie. Występują jedynie drobne niekonsekwencje w oznaczeniach w stosunku do sekcji 2, np. obrót względem osi x o 180° (ρ_x) w równaniu 2 jest oznaczany przez podanie x jako indeksu dolnego splotu bez jawnego określenia, co taka symbolika oznacza. Sekcja 4 jest poświęcona nowej kanonicznej reprezentacji splotów w oparciu o drzewa binarne, która jest pomysłem Kandydata i bardzo ułatwia zapis transformacji oraz klasyfikację splotów. W tej sekcji podany jest również algorytm transformacji splotów do reprezentacji kanonicznej. Sekcja 5 jest poświęcona klasyfikacji splotów, która była celem badań opisanych w artykule a sekcja 6 grupom symetrii splotów.

Praca 1 dotyczy analizy lass (w których tworzeniu biorą udział mostki dwusiarczkowe) w białkach. Kandydat stwierdził, że dominuje jeden typ lassa z ujemnym przewleczeniem łańcucha przez pętlę zamkniętą wiązaniem dwusiarczkowym. Analiza konserwacji reszt

cysteinowych w obrębie analizowanych rodzin białek wykazała, że obecność lassa ma znaczenie dla ich funkcji.

W pracy 2 jest opisany serwer AlphaLasso, który umożliwia analizę nie tylko struktur eksperymentalnych ale i modelowanych ponieważ sam znajduje możliwe topologie wiązań dwusiarczkowych, które zwykle nie są określone w modelach zwracanych przez serwery. Ponadto zaimplementowany algorytm bierze pod uwagę poziom wiarygodności przewidzianej struktury dla poszczególnych reszt aminokwasowych (pLDDT), aby uniknąć zaklasyfikowania artefaktów jako rzeczywistych lass. Serwer zawiera również bazę danych lass. Kandydat miał znaczący udział w opracowaniu tego serwera. Opracował algorytm detekcji lass oraz określania ich zapętlenia na podstawie maksymalnej liczby połączeń gaussowskich (maxGLN) jak również wprowadził wykorzystanie pLDDT. Wielkość pLDDT powinna być zdefiniowana w skróconym opisie publikacji na stronie 31. Z opisu wynika jedynie, że jest to miara jakości przewidywania natomiast to, że jest ona określona dla każdej reszty aminokwasowej.

Praca 3 dotyczy badań nad zachowaniem topologii węzłów i przewleczek w białkowych transporterach transmembranowych. Kandydat przeprowadził analizę węzłów oraz przewleczek w eksperymentalnych oraz modelowanych przez AlphaFold białkach tej klasy, jak również zidentyfikował ich rdzenie.

Praca 4 opisuje serwer AlphaKnot do wykrywania i analizy węzłów w modelach struktur białek wygenerowanych przez AlphaFold. Podobnie jak AlphaKnot, serwer wykorzystuje wartości pLDDT do oceny wiarygodności znalezionych węzłów, w szczególności stwierdzenia czy nie są artefaktami. Kandydat wniósł istotny wkład do tworzenia oprogramowania serwera poprzez zaprogramowanie dystrybucji zadań określania typów węzłów oraz obliczania map Alexandra.

Praca 5 jest poświęcona analizie węzłów w modelach struktur białek utworzonych przy użyciu AlphaFold i RoseTTaFold. Kandydat przeprowadził analizę 23391 modeli w celu znalezienia struktur zawężlonych oraz określenia typów węzłów. W wyniku analizy została zidentyfikowana struktura białka o identyfikatorze UniProt O00534, zawierająca węzeł 63, który nie został wcześniej stwierdzony w strukturach białek. Według informacji podanej przez Kandydata w streszczeniu pracy 5 obecność tego węzła została potwierdzona w rozwiązanej później strukturze eksperymentalnej.

W pracy 6 jest przedstawiona przeprowadzona przez Kandydata analiza węzłów w strukturach RNA, w wyniku której znaleziono pierwszą zawężloną strukturę molekuly tego typu (z węzłem trójlistnym). W streszczeniu tej pracy podanym w rozdziale 3 należałoby wyjaśnić, co to jest baza danych BGSU.

Praca 7 dotyczy analizy modeli RNA z 15 edycji ogólnoswiatowego eksperymentu oceny metod przewidywania struktur białek i RNA (CASP15) pod kątem obecności węzłów i wynikania tych węzłów z artefaktów modeli. Badania zostały w dużej części przeprowadzone przez Kandydata. Na podstawie analizy wiarygodności poszczególnych segmentów oraz porównania ze upublicznionymi po zamknięciu eksperymentu rozwiązanymi strukturami stwierdzono, że duża

część węzłów w modelach wysłanych na eksperyment CASP15 stanowi artefakty, przy czym liczba artefaktów była zdecydowanie większa w przypadku modeli uzyskanych w przy użyciu metod opartych na sztucznej inteligencji. Również więcej struktur zawężlonych znaleziono w modelach nienaturalnych makrocząsteczek RNA.

W pracy 8 jest opisany pakiet Topoly, napisany w języku Python, do analizy topologii łańcuchów polimerów. Kandydat zaimplementował algorytm generowania losowych struktur polimerów, ulepszenie algorytmu generowania krzywych θ , jak również włączenie napisanych w języku C++ kodów algorytmów do pakietu.

Najważniejsze wyniki uzyskane przez Kandydata można podsumować następująco:

1. Rozwinięcie matematycznych teorii splotów Conwaya o ich nową kanoniczną reprezentację opartej na drzewach binarnych, opracowanie opartych na niej algorytmów wykrywania i analizy splotów oraz użycie opracowanego formalizmu do uzupełnienia dotychczas znanych typów struktur zawężlonych (preprint).
2. Udział w opracowaniu pakietów Topoly (praca 8), AlphaLasso (praca 2) i AlphaKnot (praca 4) do analizy złożonych struktur polimerów, w szczególności tych zawierających sploty. Pakiety AlphaLasso i AlphaKnot umożliwiają analizę zarówno eksperymentalnych struktur białek i kwasów nukleinowych jak i modeli otrzymanych w wyniku przewidywań tych struktur.
3. Zbadanie, przy użyciu opracowanych narzędzi, topologii eksperymentalnych i modelowanych struktur białek i RNA (prace 1, 3, 5-7). Na szczególną uwagę zasługuje odkrycie pierwszej zawężlonej struktury RNA (praca 6), znalezienie najbardziej dotąd złożonego węzła 6_3 w wymodelowanej przez AlphaFold strukturze białka O00534 (praca 5) oraz pokazanie, że węzły pojawiające się w modelowanych metodami opartymi na sztucznej inteligencji strukturach RNA są w większości artefaktami (praca 7).

Z obowiązku recenzenta poniżej wymieniam znalezione przeze mnie potknięcia semantyczne, stylistyczne oraz literówki. Nie utrudniły one zrozumienia tekstu. Punkty 1-4 dotyczą rozdziału 3, punkt 5 preprintu.

1. Ostatnia linia na stronie 34 i dwie pierwsze linie na str 35, zdanie: "*If a given knot was present in over 80% of homologue structures and in a model generated by RoseTTaFold, then the was treated as knotted (51), otherwise it was treated as potentially knotted (18).*". Albo pomiędzy "the" i "was" powinien być rzeczownik albo "the" należy zastąpić przez „it”.
2. Strone 35, linia 8, początek zdania: „*Its the only crystallized domain forms an unknot, ...*". Albo "its" zdanie ma nawiązywać do poprzedniego albo "the" jeżeli nie ma nawiązywać.
3. Strona 35, linie 11-13, zdanie: "*Von Willebrand factor A domain-containing protein 5A (UniProtKB ID: O00534), contains brand new 6₃ knot, which haven't been observed in*

any experimentally solved structure.” Powinno być “which has not been observed”. W zdaniu występuje niezgodność gramatycznej liczby rzeczownika i czasownika a kontrakcji w języku angielskim używa się raczej w mowie potocznej i tekstach beletrystycznych.

4. Strona 35, 3 ostatnie linie, zdanie: „*Some RNA molecules are composed of few RNA chains – I joined them if their ends were closer than 4 Å, and appended them to the list.*” Czy chodzi tutaj o uzupełnianie “dziur” w strukturach eksperymentalnych czy łańcuchy polimeru były łączone? Jeżeli tak to jaki był cel takiego łączenia?
5. Strona 4, linia 6 preprintu: „abd” zamiast „and”.

Rozprawę doktorską oraz dorobek Pana mgra Bartosza Ambrożego Grenia oceniam bardzo wysoko. Kandydat rozwinął metody matematyczne do analizy węzłów i motywów pokrewnych w makromolekułach biologicznych i zaimplementował je w pakietach obliczeniowych, które są wykorzystywane przez innych naukowców. Przy użyciu opracowanych narzędzi wykrył nowe węzły oraz lassa w białkach i RNA. Uzyskane przezeń wyniki stanowią nowość naukową i wnoszą znaczący wkład do naszej wiedzy o tworzeniu i roli topologicznie skomplikowanych struktur w makromolekułach biologicznych a przez to do biofizyki. Uwagi krytyczne zawarte w mojej recenzji, z których część ma charakter dyskusyjny, nie umniejszają wartości rozprawy. Rozprawa zdecydowanie spełnia wymagania stawiane rozprawom doktorskim przez Ustawę z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (tekst jednolity: DzU z 2022 r. poz. 574 z późn. zm.), jak również zwyczajowe standardy stawiane rozprawom doktorskim w dziedzinie nauk ścisłych i przyrodniczych. Dlatego **konkluzja mojej recenzji jest pozytywna** i stawiam wniosek do Wysockiej Rady Dyscypliny Nauki Fizyczne Uniwersytetu Warszawskiego o dopuszczenie Pana mgra Bartosza Ambrożego Grenia do dalszych etapów przewodu doktorskiego. Z uwagi na wybitny charakter badań i ważność uzyskanych wyników, czego wymiernym efektem jest ich opublikowanie większości materiału rozprawy w prestiżowych czasopismach naukowych oraz dominujący bądź znaczący wkład Kandydata do każdej z ośmiu publikacji oraz preprintu składających się na rozprawę doktorską, wnoszę ponadto o wyróżnienie rozprawy.



Prof. dr hab. Józef Adam Liwo