

Gliwice, 04.04.2026

## Recenzja rozprawy doktorskiej mgr-a Sevastianosa Korsaka

### *Multi-scale Computational Modelling of Chromatin*

Ukończonej na Wydziale Matematyki i Nauk Informatycznych  
Politechniki Warszawskiej

Pod opieką promotora Profesora Dariusza Plewczyńskiego

#### **Tematyka i cel pracy, problem badawczy i jego znaczenie**

Przedstawiona mi do recenzji rozprawa doktorska powstała na Wydziale Matematyki i Nauk Informatycznych Politechniki Warszawskiej pod kierunkiem Profesora Dariusza Plewczyńskiego.

Tematyka rozprawy obejmuje wykorzystanie metod komputerowych do wieloskalowego (*multi-scale*) modelowania trójwymiarowej struktury chromatyny. W swoich badaniach Doktorant koncentruje się na stworzeniu zestawu metod, które umożliwią przewidywanie struktury tej cząsteczki ze szczególnym uwzględnieniem interakcji między jej różnymi skalami organizacyjnymi (od poziomu nukleosomów po całe terytoria chromosomowe). Równocześnie część jego badań obejmuje nie tylko samo modelowanie struktury chromatyny, ale także jej dynamiki w czasie całego cyklu komórkowego.

Przestrzenna organizacja chromatyny stanowi obecnie jeden z intensywnie rozwijanych obszarów współczesnej biologii molekularnej. Postęp technologiczny, obejmujący zarówno nowoczesne metody sekwencjonowania, jak i zaawansowane techniki obrazowania, umożliwia coraz dokładniejsze poznawanie trójwymiarowej architektury genomu. Zagadnienie to ma fundamentalne znaczenie dla zrozumienia mechanizmów regulujących funkcjonowanie komórki. Szczególną rolę odgrywają tu pętle chromatynowe, które pośredniczą w kontaktach pomiędzy odległymi regionami DNA, w tym promotorami i enhancerami, wpływając tym samym na kontrolę ekspresji genów. Architektura chromatyny pozostaje również ściśle związana z przebiegiem replikacji materiału genetycznego oraz procesami naprawy DNA, a jej zaburzenia mogą sprzyjać rozwojowi wielu schorzeń, zwłaszcza chorób nowotworowych. Z tego względu badania nad strukturą przestrzenną chromatyny mają istotne znaczenie nie tylko poznawcze, lecz także aplikacyjne, ponieważ mogą przyczynić się do lepszego zrozumienia patogenezы chorób oraz wspierać poszukiwanie nowych strategii terapeutycznych.

W niniejszej rozprawie Doktorant podejmuje tematykę badawczą związaną z wielkoskalowym modelowaniem struktury chromatyny. Autor nie ogranicza swoich analiz do pojedynczego poziomu organizacji, lecz podejmuje próbę ujęcia wzajemnych zależności pomiędzy różnymi skalami jej struktury w ramach spójnego modelu. W tym celu przeprowadzone badania obejmują niemal wszystkie istotne poziomy organizacji chromatyny, w tym symulacje geometryczne nukleosomów, stochastyczne modele ekstruzji pętli, dynamikę kompartmentalizacji oraz oddziaływania chromatyny z laminą jądrową. Celem rozprawy nie jest wyłącznie odtworzenie struktury chromatyny, lecz również poznanie jej dynamiki, przemian fazowych zachodzących pod wpływem kluczowych parametrów, a przede wszystkim związków pomiędzy zmianami strukturalnymi a podstawowym procesami, takimi jak cykl komórkowy oraz związany z nim stres replikacyjny, który jest procesem biologicznym ściśle powiązanego z rozwojem chorób nowotworowych

Oceniając tematykę badawczą podjętą w rozprawie, należy uznać, że dotyczy ona problemu istotnego i aktualnego, wyraźnie osadzonego w aktualnych kierunkach rozwoju bioinformatyki oraz biologii obliczeniowej.

## Charakterystyka rozprawy

Przedstawiona praca doktorska składa się z trzech rozdziałów.

Pierwszy rozdział (*Introduction*) stanowi obszernie, interdyscyplinarne wprowadzenie do niezbędnych koncepcji z dziedziny informatyki, biologii i fizyki związanych z tematyką rozprawy. Autor wyjaśnia, jak matematycznie opisać chromatynę oraz szczegółowo omawia podstawowe terminy, takie jak właściwości wieloskalowe, fraktale, stochastyczność czy przejścia fazowe. Dodatkowo opisane są metody pozyskiwania danych eksperymentalnych oraz wyzwania i powszechnie stosowane podejścia w modelowaniu chromatyny. Na końcu tego rozdziału sformułowano motywację badawczą oraz główne cele rozprawy.

Rozdział drugi (*Overview of Published Research*) to przegląd opublikowanych badań. Jest to rozdział, w którym w zwięzły sposób Doktorant zaprezentował trzy kluczowe modele będące wynikiem badań prowadzonych w ramach doktoratu: LoopSage, RepliSage oraz MultiMM. Rozdział ten ma na celu wyjaśnienie oraz opis zastosowanych rozwiązań metodologicznych.

Pierwszą z opisanych metod jest **LoopSage** – metoda modelowania dynamiki procesu ekstruzji pętli chromatynowych. Zaproponowana metoda LoopSage jest dwuetapowa: w pierwszym etapie symulacji Monte Carlo algorytm bada przestrzeń prawdopodobieństwa, symulując trajektorie LEFs (*Loop Extrusion Factors*). Drugi etap metody to faza dynamiki molekularnej z wykorzystaniem narzędzia OpenMM, gdzie wykorzystuje się otrzymane z poprzedniego etapu parametry (pozycje LEFs). Tym, co w głównej mierze charakteryzuje tę metodę, jest odejście od tradycyjnego, sztywnego kodowania reguł biofizycznych w kodzie programu na rzecz stochastycznego modelu opartego na energii układu.

Druga opisana w tym rozdziale metoda, **RepliSage**, stanowi rozszerzenie metody LoopSage i służy do symulacji dynamiki chromatyny w trakcie całego cyklu komórkowego. Metoda ta działa w trzech etapach. Pierwszy moduł systemu RepliSage zbudowany jest na bazie opisanej wcześniej metody LoopSage i odpowiada za symulację replikacji DNA, skupiając się na losowej aktywacji miejsc inicjacji oraz postępującym

ruchu widełek replikacyjnych. Drugi etap obejmuje symulację stochastyczną łączącą ekstruzję pętli ze zmianami epigenetycznymi (z wykorzystaniem modelu Potts'a). Ostatni etap analizy, to moduł dynamiki molekularnej, który na podstawie wygenerowanych stanów tworzy końcowe, trójwymiarowe struktury chromatyny. Metoda RepliSage została zaprojektowana tak, aby móc skutecznie modelować warunki stresu replikacyjnego, który jest zjawiskiem ściśle powiązaniem z niestabilnością genomową i powstawaniem nowotworów.

Ostatnie opisane w rozdziale trzecim narzędzie to **MultiMM**. Framework MultiMM to deterministyczny, wieloskalowy model obliczeniowy służący do rekonstrukcji trójwymiarowej architektury przestrzennej chromatyny w skali genomowej. Stanowi on kompleksowe narzędzie, którego główną zaletą jest zdolność do integracji organizacji chromatyny na wszystkich istotnych biologicznie poziomach – od rozdzielczości pojedynczych nukleosomów, poprzez pętle chromatynowe i domeny TAD, po kompartmenty, terytoria chromosomowe oraz pełną przestrzeń jądra komórkowego. Model ten charakteryzuje się modułowością. Funkcja całkowitej energii układu może być elastycznie modyfikowana poprzez włączanie i wyłączenie poszczególnych komponentów sił, w zależności od zadanego problemu oraz dostępnych danych wejściowych. Składowe modelu obejmują m.in.: reprezentację pętli chromatynowych, modelowanie zjawiska kompartmentalizacji chromatyny przy użyciu potencjałów kopolimerów blokowych oraz wydajną, geometryczną reprezentację nukleosomów (w przeciwieństwie do kosztownych obliczeniowo symulacji atomowych). Dodatkowo, system MultiMM oferuje szereg opcji pozwalających na personalizację dynamiki, modelowanie wybranych fragmentów genomu oraz generowanie statystycznych zespołów struktur w celu zbadania zmienności i stabilności architektury chromatyny.

Ostatni trzeci rozdział (*Conclusions and Future Perspectives*) zawiera wnioski i przyszłe kierunki badawcze, w których można rozwijać zaprezentowane prace. Rozdział ten podsumowuje całość pracy, analizując w jaki sposób opracowane modele zrealizowały założone cele i jak przyczyniają się do rozwoju wiedzy o strukturze 3D genomu i jej powiązaniach z procesami biofizycznymi. Zawiera on również dyskusję na temat ograniczeń i niedoskonałości zaprezentowanych metod. Na koniec Autor zarysowuje perspektywy dla przyszłych kierunków badań, wskazując m.in. na ogromny potencjał wykorzystania uczenia maszynowego do przewidywania struktury chromatyny bezpośrednio z sekwencji DNA, a także podsumowuje swoje osiągnięcia naukowe, publikacje oraz udział w grantach badawczych.

## Opinia o rozprawie

Przedstawione w rozprawie wyniki stanowią spójny i logicznie skonstruowany cykl badań poświęconych modelowaniu struktury chromatyny. Pierwszy rozdział pracy stanowi wprowadzenie do tematu, drugi opisuje opracowane metody, w trzecim Doktorant podejmuje dyskusję wskazując problemy i słabości istniejących metod oraz możliwe ich rozwiązania. Analizując opracowane w ramach Doktoratu metody szczególnie godne podkreślenia jest to, że przyjęta strategia badawcza realizowana jest konsekwentnie według zasady przejścia od szczegółu do ogółu – od modelowania pojedynczych elementów strukturalnych, poprzez metody modelowania uwzględniające cykl komórkowy chromatyny, aż po framework umożliwiając odwzorowanie organizacji chromatyny jednocześnie na wielu skalach przestrzennych.

Podejmowana problematyka wpisuje się w aktualne trendy badawcze z pogranicza biofizyki, bioinformatyki obliczeniowej oraz biologii strukturalnej genomu — dziedzin, które w ostatnich latach przeżywają dynamiczny rozwój, czego dowodem jest rosnąca liczba publikacji poświęconych trójwymiarowej organizacji chromatyny. Warto przy tym podkreślić, że postęp w tej dziedzinie ma charakter ściśle sprzężony: dynamiczny rozwój metod eksperymentalnych, takich jak techniki Hi-C, ChIP-seq, ATAC-seq czy obrazowanie FISH, dostarcza coraz bogatszych i bardziej precyzyjnych danych, które z kolei stymulują i umożliwiają tworzenie coraz bardziej zaawansowanych metod obliczeniowych. Rozprawa doskonale wpisuje się w tę dwukierunkową zależność, czerpiąc z najnowszych danych eksperymentalnych jako podstawy dla proponowanych modeli. Dopelnieniem całości jest krytyczna analiza metod walidacji modeli chromatyny, która ujawnia istotne słabości obecnie stosowanych podejść walidacyjnych i stanowi cenny wkład w metodologię dziedziny.

Bibliografia rozprawy jest obszerna i obejmuje 278 pozycji literaturowych, co świadczy o gruntownym zapoznaniu się Doktoranta z dorobkiem naukowym w omawianej dziedzinie. Pewne zdziwienie budzi jednak fakt, że zaledwie około 1/4 przywołanych źródeł stanowią publikacje z ostatnich pięciu lat (tj. z roku 2021 lub nowsze). Biorąc pod uwagę, że pole badawcze objęte rozprawą cechuje się wyjątkowo szybkim tempem rozwoju, wskazane byłoby szersze uwzględnienie najnowszego piśmiennictwa, które mogłoby wzbogacić dyskusję wyników oraz zapewnić pełniejsze osadzenie pracy w bieżącym stanie wiedzy.

Pozytywnie oceniam również fakt, że kod źródłowy wszystkich opracowanych metod i narzędzi obliczeniowych został udostępniony publicznie w repozytoriach serwisu GitHub. Praktyka ta, zgodna z postulatami otwartej nauki, znacząco podnosi wartość uzyskanych w ramach rozprawy wyników, umożliwiając innym badaczom weryfikację uzyskanych wyników oraz dalszy rozwój zaproponowanych rozwiązań.

Podsumowując, w pracy umiejętnie połączono modelowania fizycznego z podejściami obliczeniowymi — formalizmy mechaniki statystycznej i dynamiki molekularnej zostały z powodzeniem zintegrowane z nowoczesnymi technikami symulacyjnymi i analizą danych genomicznych, tworząc spójne ramy badawcze. Zastosowane metody badawcze są adekwatne do postawionych celów naukowych, a przyjęte w pracy podejścia do przetwarzania i analizy danych wskazują na dobrą znajomość przez Doktoranta nowoczesnych i efektywnych narzędzi stosowanych we współczesnej biologii obliczeniowej.

### **Uwagi krytyczne i dyskusyjne**

Na wstępie niniejszej części recenzji chciałabym wyraźnie zaznaczyć, że problematyka podjęta przez Doktoranta należy do zagadnień o istotnym znaczeniu naukowym i znacznym stopniu trudności. Po szczegółowej analizie przedłożonej rozprawy nie stwierdzam obecności błędów merytorycznych o zasadniczym charakterze. Uwagi przedstawione poniżej formułuję w duchu dyskusji naukowej i nie stanowią one podstawy do obniżenia ogólnej, pozytywnej oceny pracy.

Pewną wątpliwość budzi niejednoznaczna forma redakcyjna rozprawy. Czytając odnosiłam wrażenie, że Doktorant nie podjął do końca jednoznacznej decyzji co do przyjętego modelu prezentacji — czy praca ma stanowić monografię w tradycyjnym ujęciu, czy też cykl powiązanych tematycznie artykułów naukowych. W efekcie rozprawa jest z jednej strony stosunkowo obszerna, z drugiej zaś w odniesieniu do szeregu przedstawionych metod brakuje pełnej informacji dotyczącej ich wydajności i skuteczności na rzeczywistych danych, oraz opisu tych danych. Kwestia walidacji modeli sygnalizowana jest jedynie skrótowo, ograniczając się w wielu miejscach do ogólnych komentarzy na temat trudności z nią związanych, bez przedstawienia konkretnych wyników ilościowych. Co więcej, w pracy nie zamieszczono bezpośrednich odnośników do artykułów lub ich konkretnych sekcji, w których wyniki te zostały opublikowane. Recenzent jest wprawdzie w stanie domyślić się, w których publikacjach należy szukać brakujących informacji, niemniej jednak rozprawa doktorska powinna zawierać tego rodzaju informacje podane wprost, a nie pozostawiać je do samodzielnego ustalenia przez czytelnika. Uważam, że jest to istotny minus pracy. Wskazane byłoby uzupełnienie odpowiednich fragmentów o jednoznaczne odesłania do powiązanych publikacji oraz tam, gdzie to możliwe o związku podsumowanie kluczowych wyników ilościowych.

Niedosyt budzi sposób przedstawienia celów badawczych we wstępie rozprawy. Wstęp zawiera jedynie bardzo ogólne stwierdzenie, że przedmiotem pracy jest modelowanie trójwymiarowej struktury chromatyny, bez precyzyjnego określenia szczegółowych celów naukowych i pytań badawczych, które Doktorant zamierzał podjąć. Informacje te pojawiają się dopiero w końcowej części rozdziału pierwszego, co utrudnia czytelnikowi — a także recenzentowi — właściwe osadzenie prezentowanych wyników w zamierzonym kontekście badawczym już na etapie lektury wstępnej. Zgodnie z przyjętymi standardami redakcji rozpraw doktorskich, cele pracy powinny być wyraźnie sformułowane we wstępie, tak aby od samego początku wyznaczały perspektywę dla oceny zawartości wszystkich kolejnych rozdziałów.

W sekcji *Limitations and Open Challenges* rozdziału trzeciego, Doktorant słusznie zauważa, że przedstawione w pracy metody wymagają określenia zestawu hiperparametrów, co w praktyce stanowi niebagatelne wyzwanie. Doktorant wspomina o możliwości zastosowania przeszukiwania siatki jako strategii ich doboru, jednocześnie jednak przyznaje, że brak jednoznacznych danych eksperymentalnych opisujących strukturę 3D chromatyny istotnie utrudnia ocenę jakości uzyskanych rozwiązań. Doceniam fakt, że w rozdziale trzecim Doktorant podejmuje dyskusję nad możliwymi podejściami do wyznaczania hiperparametrów i jest świadomy złożoności tego zagadnienia. Niemniej jednak chciałabym w tym miejscu zadać Doktorantowi pytanie: czy na podstawie dotychczasowych doświadczeń z opracowanymi modelami jest w stanie zaproponować potencjalnym użytkownikom konkretny, praktyczny „przepis” na dobór hiperparametrów? Innymi słowy — czy możliwe jest sformułowanie choćby ogólnych wytycznych lub heurystyk, które pozwoliłyby nowemu użytkownikowi na sensowne zainicjowanie procesu optymalizacji, bez konieczności przeprowadzania kosztownych obliczeń metodą prób i błędów lub implementacji nowych metod optymalizacji?

## Podsumowanie

Pan mgr Sevastianos Korsak przedstawił rozprawę doktorską rozwiązującą aktualny problem naukowy, która przyczyni się do rozwoju reprezentowanej dyscypliny naukowej. Nie mam wątpliwości, że rozprawa zawiera oryginalne rozwiązanie problemu naukowego, a Doktorant wykazał, że zarówno posiada ogólną wiedzę teoretyczną w dyscyplinie informatyka techniczna i telekomunikacja jak i umiejętność prowadzenia pracy naukowej.

Doktorant jest współautorem 9 publikacji naukowych, z których jedna jest aktualnie dostępna wyłącznie na platformie *bioRxiv* i pozostaje w trakcie procesu recenzyjnego. Jest on pierwszym autorem w czterech spośród tych prac, trzech opublikowanych oraz jednej recenzowanej. Przy czym należy podkreślić, że w pracach opisujących główne metody obliczeniowe przedstawione w rozprawie, Doktorant figuruje jako pierwszy autor, co świadczy o jego samodzielnym i wiodącym wkładzie w prowadzone badania. W okresie doktoratu brał on ponadto aktywny udział w projektach naukowych, co dodatkowo potwierdza jego zaangażowanie w działalność badawczą.

Biorąc pod uwagę ocenę przedstawioną w niniejszej recenzji, nie mam wątpliwości, że przedstawiona do oceny praca doktorska w pełni odpowiada warunkom określonym w Art. 187 ustawy z dnia 20 lipca 2018 r. Prawo o Szkolnictwie Wyższym i Nauce (tekst jednolity Dz. U. z 2023 r. poz. 742 z późn. zm.) i na tej podstawie wnoszę do Wysokiej Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej o dopuszczenie mgr-a Sevastianosa Korsaka do dalszych etapów przewodu doktorskiego.

dr inż. hab. Aleksandra Gruca, prof. PŚ.