

**RADA NAUKOWA DYSCYPLINY INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA
POLITECHNIKI WARSZAWSKIEJ**

zaprasza na
OBRONĘ ROZPRAWY DOKTORSKIEJ

Mgr. inż. Michała Kadlofa

która odbędzie się w dniu **6.10.2025 roku**, o godzinie **15:30** w trybie stacjonarnym
temat rozprawy:

Komputerowe modelowanie trójwymiarowej struktury chromatyny

promotor: prof. dr hab. Dariusz Plewczyński – Politechnika Warszawska

recenzenci: prof. dr hab. Paweł Mackiewicz – Uniwersytet Wrocławski

prof. dr hab. inż. Andrzej Polański – Politechnika Śląska

prof. dr hab. inż. Marta Szachniuk - Politechnika Poznańska

Obrona odbędzie się w formie stacjonarnej w sali nr 40 w Gmachu Wydziału Matematyki i Nauk Informacyjnych Politechniki Warszawskiej. Z rozprawą doktorską i recenzjami można zapoznać się w Czytelni Biblioteki Głównej Politechniki Warszawskiej, Warszawa, Plac Politechniki 1, oraz na stronie internetowej BIP PW:

www.bip.pw.edu.pl/Postepowania-w-sprawie-nadania-stopnia-naukowego/Doktoraty/Wszczete-po-30-kwietnia-2019-r/Rada-Naukowa-Dyscypliny-Informatyka-Techniczna-i-Telekomunikacja/mgr-inz.-Michal-Kadlof

Przewodniczący Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja
Politechniki Warszawskiej
prof. dr hab. inż. Jarosław Arabas

STRESZCZENIE PRACY DOKTORSKIEJ

Autor: Michał Kadlof

Promotor: prof. Dariusz Plewczyński

Tytuł pracy: *Komputerowe modelowanie trójwymiarowej struktury chromatyny*

Streszczenie:

Chromatyna - skomplikowany układ białek DNA, białek i innych czynników regulacyjnych odgrywa fundamentalną rolę w regulacji ekspresji genów i zachowaniu strukturalnym genomu. Niniejsza praca doktorska stanowi holistyczny przewodnik po komputerowych metodach trójwymiarowego modelowania struktury chromatyny. Autor we wstępie rozpoczyna od omówienia biofizycznych aspektów budowy tej cząsteczki, następnie omawia dostępne typy danych biologicznych o zastosowaniach modelarskich i dokonuje krytycznego przeglądu istniejących metod modelarskich.

W zasadniczej części pracy autor prezentuje własne metody konstruowania modeli, a wśród nich GD-MDS (ang. *Graph Distance Multi Dimensional Scaling*) z metryką opartą o dystans grafowy, spring-model - metodę opartą o mechanikę molekularną - zaimplementowaną i udostępnioną w formie usługi WWW. W dalszej części omawia stworzone przez siebie oprogramowanie md-soft stanowiące uogólnienie metod użytych w spring-model i tworzący złożony ekosystem modelarski umożliwiający integrację komplementarnych typów danych oraz symulację ich ewolucji w czasie. Przedstawione są rozmaite scenariusze przeprowadzania symulacji w różnych skalach czasu i przestrzeni i dla różnych typów danych. Zaproponowana jest systematyczna metoda projektowania strategii eksperymentów modelarskich.

W ostatniej części przedstawione są dwa autorskie algorytmy modelowania na podstawie obrazów mikroskopowych: metoda siłowa - z wykorzystaniem techniki MDFF (ang. *Molecular Dynamics Flexible Fitting*) oraz metoda korelacyjna polegająca na porównywaniu obrazu mikroskopowego z obrazem symulowanym z bieżącej konformacji struktury. Na końcu autor omawia techniczne (w tym infrastrukturalne) aspekty modelowania i prezentuje zasady jakimi kierował się przy budowie i wdrażaniu klastra Eden^N, oraz organizacji ośrodka obliczeniowego MiNI PW - którym kieruje. Ośrodek zapewnia dostęp do największej mocy obliczeniowej w Politechnice Warszawskiej, która między innymi posłużyła do powstania niniejszej pracy, a obecnie służy ponad setce innych badaczy.



Wrocław, 15.05.2025

Recenzja rozprawy doktorskiej Pana mgr inż. Michała Kadlofa pt. "Komputerowe modelowanie trójwymiarowej struktury chromatyny"

Organizacja i ekspresja informacji genetycznej, zwłaszcza u organizmów eukariotycznych, są skomplikowanymi zjawiskami wymagającym interakcji materiału genetycznego i białek. Włókniste kompleksy tych związków zwane chromatyną umożliwiają upakowanie bardzo długiej cząsteczki DNA (u człowieka ok. 2 metry) w niewielkiej objętości jądra komórkowego o średnicy kilku mikrometrów. Poza funkcją strukturalną, chromatyna odgrywa kluczową rolę w regulacji ekspresji genów. Jej organizacja na różnych poziomach umożliwia kontrolę dostępu do konkretnych fragmentów materiału genetycznego. Aby zrozumieć te zjawiska nie wystarcza już znajomość sekwencji genomu oraz białek łączących się z kwasami nukleinowymi, ale potrzebna jest wiedza o trójwymiarowej strukturze chromatyny. Badania nad trójwymiarową organizacją chromatyny mają kluczowe znaczenie dla zrozumienia mechanizmów regulacji genów, różnicowania komórek, odpowiedzi na bodźce środowiskowe oraz powstawania chorób. Chromatyna nie jest statyczna – jej przestrzenna organizacja jest dynamiczna i zależy od typu komórki, fazy cyklu komórkowego oraz sygnałów zewnętrznych. Dlatego bardzo słusznie ambitnym przedmiotem pracy doktorskiej Pana mgr inż. Michała Kadlofa stało się opracowanie metod komputerowego modelowania trójwymiarowej struktury chromatyny.

Rozprawa doktorska została napisana w języku polskim ze streszczeniem w języku angielskim. Zawiera ona: Streszczenie, Abstract, Spis treści, Wykaz użytych skrótów, Część teoretyczną i praktyczną oraz Dodatki zawierające wykaz publikacji autora, warianty strukturalne nukleosomów, Spis rysunków i tabel, Bibliografię i Skorowidz. Nie jest to typowy układ, jednak Część teoretyczną można traktować jako wstęp, a praktyczną jako wyniki i dyskusję. Przydałby się jednak wyróżnić cele pracy w osobnym rozdziale oraz dodać rozdział z wnioskami lub podsumowaniem zawierający najważniejsze informacje uzyskane z opracowanych modeli i przeprowadzonych badań.

W Części teoretycznej we Wstępie w bardzo przystępny i prosty sposób doktorant scharakteryzował chromatynę podając łatwy sposób na jej izolację, którą wykonał, a efekt uwiecznił na zdjęciach. Pod koniec Wstępu podał cel pracy, którym było zaproponowanie nowatorskich metod obliczeniowych do określania przestrzennych modeli chromatyny. Cel pracy doktorskiej został jasno sformułowany i właściwie zrealizowany. Pozostałe rozdziały w przejrzysty i wystarczający sposób zbierają najważniejsze fakty o budowie i roli chromatyny w regulacji ekspresji genów i zachowaniu struktury genomu. Ciekawe są wyliczenia związane z długością, masą i gęstością chromatyny. Zwrócono uwagę, że włókno 30 nm nie jest strukturą powszechnie występującą, jak wcześniej sądzono, a głównymi poziomami organizacji chromatyny są włókno 11 nm, pętle, domeny i kompartmenty. Dobrze byłoby przedstawić schematy pokazujące tę hierarchiczną organizację.

Osobny podrozdział dotyczy technik eksperymentalnych służących do określania struktury chromatyny, genomowych (3C, 4C i 5C, Hi-C, ChIA-PET, ChIA-Drop) i mikroskopowych (FISH, EM-ISH, iPALM). Opisano je w zrozumiały sposób uwzględniając ich wady i zalety, jednak idealnie byłoby do opisu dołączyć schematy tych metod. W dodatkowym rozdziale przedstawione zostały modele chromatyny: łańcuch idealny, model Kratky-Poroda oraz model WLC, model Odijk-Skolnick-Fixman oraz globule fraktalne uwzględniające różne cechy, jak sztywność, efekty odpychania elektrostatycznego oraz bliskiego oddziaływania regionów.

W ostatnim rozdziale części teoretycznej scharakteryzowano metody modelowania struktury chromatyny. W rozdziale zawarto także rozważania natury filozoficznej, a nie tylko suche techniczne fakty. Podkreślono różnice w modelowaniu chromatyny i białek dotyczące natury ich struktur oraz skali i rozdzielczości danych do modelowania. Można byłoby tutaj wspomnieć o słynnym programie AlphaFold. Ciekawe jest porównanie struktury chromatyny do mapy świata. Doktorant przedyskutował ważne parametry związane z modelowaniem, jak: skala, rozdzielczość, typ danych i cechy polimeru oraz przedstawił ważne rozważania dotyczące idei modelowania, co wskazuje, że zdaje sobie sprawę z jego uproszczeń i ograniczeń. Świadczy to o głębokim zrozumieniu zagadnienia przez Pana Michała Kadlofa.

Mam jednak jeszcze inne uwagi do tej części. W przypadku celów modelowania można byłoby je lepiej sprecyzować, zwłaszcza dotyczące zagadnień badawczych obok wizualizacyjnych. W części dotyczącej wyboru regionów do modelowania prosiłbym dokładniej wyjaśnić kwestię macierzy po normalizacji. Podrozdział dotyczący wyboru skali mógłby zawierać wyszczególnienie tych skal i podanie rzędów wielkości, albo można byłoby go połączyć z podrozdziałem o rozdzielczości. Dobrze byłoby wyjaśnić, co to jest punkt

modelowanego układu. Nie wiem czy wyrażenie "Im lepsza (tzn. mniejsza) rozdzielczość" powinno brzmieć "Im lepsza (tzn. większa) rozdzielczość". Dobrze byłoby wyjaśnić promień bezwładności dla struktur modelowanych oraz potencjał FENE.

Główną częścią doktoratu jest część praktyczna, która jest podzielona na kilka rozdziałów opisujących różne narzędzia do modelowania struktury chromatyny. Doktorant przedstawił dwie metody opracowane przez siebie do wyznaczania użytecznych granic domen. Jedna z metod nazwana CCD Caller służy do identyfikacji regionów o ponadprzeciętnym pokryciu łukami interakcji w porównaniu do sygnału oczekiwanego. Jest on użyteczny, ponieważ w obiektywny sposób wyszukuje takie regiony (domeny typu CCD) i ułatwia automatyzację analiz. Druga metoda MetaTAD Caller analizuje wyniki wielu innych programów wyznaczających domeny typu TAD w oparciu o głosowanie wykorzystując funkcję Gaussa opisującą rozmycie położenia granicy domen. Dobrze byłoby przedstawić jaka jest skuteczność tych metod, np. procent fałszywie i poprawnie przewidzianych regionów w porównaniu do danych rzeczywistych oraz potencjalne ograniczenia. Mam kilka pytań związanych z tą częścią. Prosiłbym o wyjaśnienie zdania, że TAD-y tworzą zbiór stycznych do siebie odcinków dzielących cały chromosom na części. Według mnie styczne odcinki do siebie będą się nakładać i będą równoległe. Mam jeszcze dodatkowe pytanie: Czy liczenie indeksu kierunkowego od tyłu, czyli różnica interakcji późniejszych do wcześniejszych da ten sam wynik, co różnica między wcześniejszymi i późniejszymi interakcjami? Uważam, że Rysunek 6.4 powinien być lepiej opisany.

Jasno jest scharakteryzowana metoda skalowania wielowymiarowego i dystanse grafowe oraz sposoby eliminacji szumu, jednak brakuje wyjaśnienia po co jest ten opis. Czy został on zastosowany w jakimś programie? Wypadałoby pokazać skuteczność działania tej metody na przykładzie.

Pan mgr inż. Michał Kadlof przedstawił także algorytm budowania modeli polimerowych nazwany modelem sprężynowym (ang. SM, Spring Model) polegający na znalezieniu takiej konformacji punktów w przestrzeni trójwymiarowej, która minimalizuje funkcję energii potencjalnej. Model jest atrakcyjny i uwzględniający wiele istotnych parametrów, a jego opis jest jasny i zawiera wiele praktycznych wskazówek. Szczególnie dużo miejsca poświęcono strukturom początkowym. Podano także interesujące przykłady zastosowań oraz implementacje. Mam pewne uwagi również do tej części. We wzorze i opisie na energię niedostępnej objętości dobrze byłoby dać ten sam symbol na średnicę kulki. Chciałbym się także dowiedzieć skąd bierze się potęga 12 w tym wzorze? Następujące stwierdzenie jest dla mnie niejasne: Z tego powodu dobrą praktyką jest wybierać takie struktury

początkowe, które już na początku zostały wybrane w taki sposób by ich energia nie odbiegała znacznie od minimum układu. Jest to dla mnie dziwne, ponieważ jak wybrać te struktury, skoro aby dowiedzieć się jakie jest minimum energetyczne układu trzeba przeprowadzić obliczenia już na jakiś strukturach początkowych? W związku z tym, że struktury początkowe mogą mieć wpływ na końcowy wynik chciałbym uzyskać odpowiedź na pytanie: Jak duża jest powtarzalność wyników dla różnych stanów początkowych? Domyślam się, że może to zależeć od złożoności układu. W przypadku opisu zastosowań tej metody w podrozdziale 8.6.1 dobrze byłoby skomentować uzyskane wyniki na Rys. 8.12. Jakie struktury należałoby oczekiwać w podanym przykładzie? Dobrze byłoby również napisać jaka jest skuteczność modelu sprężynowego w porównaniu z innymi. Chciałbym jeszcze uzyskać odpowiedź na pytanie: Do jakich jeszcze innych polimerów może służyć opracowany model sprężynowy?

Interesujące jest podejście opisane w następnym rozdziale uwzględniające zastosowanie wysokorozdzielczych obrazów mikroskopowych w metodzie siłowej i korelacyjnej służących do modelowania struktur polimerów. Pokazano także wynik działania tego podejścia na oczekiwanej strukturze metodą siłową przy dwóch stanach początkowych oraz na danych jądra komórkowego *Arabidopsis thaliana*. Model generalnie dobrze wypadł, ale pokazano, że wpływ na wynik ma stan początkowy. Ładnie przedyskutowano wyniki symulacji na danych układu testowego. Podobny komentarz przydałby się dla danych rzeczywistych. Mam jeszcze dodatkowe pytanie: Jaki byłby wynik przy zastawianiu metody korelacyjnej oraz innych stanów początkowych dla obu metod?

Ostatni rozdział dotyczy aspektów technicznych zaawansowanych obliczeń. Doktorant przedstawił najważniejsze problemy obliczeniowe i sposoby ich rozwiązania. Szczegółowo opisał architekturę i zarządzanie na klastrze Eden, który służy do wykonywania skomplikowanych obliczeń struktury chromatyny. Klaster ten posłużył do uzyskania wyników przedstawionych w pracy doktorskiej. Należy zaznaczyć, że Pan mgr inż. Michał Kadlof administrował tym klastrem. Dobrze byłoby, aby doktorant wyeksponował jaki miał udział w jego budowie i wdrażaniu.

Warto podkreślić, że doktorant jest współautorem 9 publikacji w renomowanych czasopismach, a dwie metody zostały opublikowane w dwóch pracach.

Generalnie praca jest napisana poprawnym językiem, miejscami występuje żargon i swobodny styl, ale dobrze się ją czyta. Znalazłem następujące błędy z propozycją poprawek: 2m -> 2 m; dwóch typów para zasad -> dwóch typów par zasad; ilością wiązań -> liczbą wiązań; które łatwo tworzą jony H⁺ -> które łatwo przyciągają jony H⁺; 3mm -> 3 mm, dna -> DNA;

odległości ziemi od słońca -> odległości Ziemi od Słońca; Histogram długości chromosomów
 -> Wykres długości chromosomów; parametru GC-content -> zawartości par GC; bilateria ->
 Bilateria; rys. -> Rys; fig. -> Fig., kompartmenty -> kompartmenty, intra i inter
 chromosomalnych -> intra- i interchromosomalnych; nie wielką -> niewielką; trudne do
 sekwencjonowania Np, regiony o wysokiej repetytywności, niemapowalne itp. -> trudne do
 sekwencjonowania, np. regiony o wysokiej repetytywności, niemapowalne itp., odmienne
 interpretacje danych. Np. wartości z macierzy Chia-PET -> odmienne interpretacje danych, np.
 wartości z macierzy Chia-PET, ilości ilości uzyskanych odczytów -> liczby uzyskanych
 odczytów, Prezentuje on ten przypadki -> Prezentuje on przypadki, giętkąść -> giętkość, po
 rzeczywisty -> po rzeczywistym, nie wpływa na pomiary -> nie wpływa on na pomiary, nie
 pożądaný -> niepożądany, płaskie. (rys. 8.4) -> płaskie (Rys. 8.4)., Tek -> Tak, pomiędzy
 każdym trzem sąsiadującymi punktami -> pomiędzy każdym z trzech sąsiadujących punktów,
 problemowi *problemowi* -> *problemowi*, nie zbyt -> niezbyt, nie długich -> niedługich, w w -
 > w, dlatego. Z tego -> dlatego, z tego, przyz -> przy, kolekcję kolekcję -> kolekcję, Na struktura
 -> Ta struktura, Dokładnie identyczne zastosowano -> Dokładnie identyczne założenia
 zastosowano, sać się -> stać się, macierzy. Np. serwerów -> macierzy, np. serwerów, są -> są,
 Obniża bezpieczeństwo -> Obniża to bezpieczeństwo, ilość -> liczba.

Ponadto w wielu miejscach brakuje przecinków przed: aby, zaś, gdy, gdyby, co, czyli,
 w której, którą, które, w których, ponieważ, jeśli. Brak jest czasownika w kilku zdaniach. Jedno
 zdanie było urwane. Niejasne jest zdanie: W efekcie pokrycie średnie ma bardzo, krótkie
 obszary początkowe i końcowe, a większość chromosomu. Niepoprawnie jest sformatowana
 Tabela 2.2. Niektóre rysunki powinny być większe. Raz jest używane fig. a raz rys. Jako
 separator dziesiętny raz jest stosowany przecinek, a raz kropka.

Reasumując mogę stwierdzić, że powyższe uwagi nie umniejszają znaczenia pracy
 doktorskiej. Doktorant włożył dużo trudu w opracowanie nowatorskich metod i
 przeprowadzone analizy, a przedstawione opisy wyników świadczą o jego dużej dojrzałości
 naukowej, wnikliwości i głębokim zrozumieniu zagadnienia. Nie mam zastrzeżeń do metodyki
 przeprowadzonych analiz.

Tematyka pracy doktorskiej jest bardzo zasadna, ponieważ istnieje potrzeba tworzenia
 nowych metod i narzędzi modelujących struktury przestrzenne chromatyny, co stało się
 przedmiotem rozprawy. Warto zaznaczyć, że zastosowane w pracy modele są nowe i
 zaawansowane, a uzyskane wyniki dowodzą, że wykorzystanie takich metod w biologii
 molekularnej prowadzi do wartościowych i interesujących odkryć. Uważam, że recenzowana

rozprawa stanowi istotne i kompleksowe opracowanie komputerowych metod trójwymiarowego modelowania struktury chromatyny.

Uważam, więc, że przedstawiona do recenzji rozprawa doktorska spełnia wszystkie wymogi Ustawy o Stopniach Naukowych. Zgłaszam, zatem wniosek do Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej o uznanie rozprawy Pana mgr inż. Michał Kadłof za odpowiadającą wymogom stawianym rozprawom doktorskim i o dopuszczenie doktoranta do dalszych etapów przewodu doktorskiego.

Prof. dr hab. Paweł Mackiewicz



Katedra Grafiki, Wizji Komputerowej i Systemów Cyfrowych

Politechniki Śląskiej

Recenzja rozprawy doktorskiej

Autor: mgr inż. Michał Kadlof

Tytuł: Komputerowe modelowanie trójwymiarowej struktury chromatyny

Promotor: prof. dr hab. Dariusz Plewczyński

Ogólna charakterystyka rozprawy

Przedłożona do recenzji rozprawa doktorska liczy 170 stron tekstu, składa się z dziesięciu rozdziałów, dodatków A, B, spisu rysunków, tabel, bibliografii oraz skorowidza. Dziesięć rozdziałów tekstu pracy jest podzielonych na trzy części, teoretyczną, praktyczną i dodatki. Praca przedstawia bardzo szeroki obraz zagadnień związanych z problematyką modelowania struktury chromatyny. Obraz ten obejmuje zarówno aspekty biologiczne jak też eksperymentalne, modelowania matematycznego, bioinformatyki, budowy oprogramowania, technik obliczeniowych łącznie z aspektami konstrukcyjnymi klastra obliczeniowego. Oryginalne, naukowe elementy rozprawy leżą głównie w obszarach modelowania matematycznego oraz bioinformatyki, są udokumentowane recenzowanymi publikacjami naukowymi, których współautorem jest Doktorant.

Praca ma charakter monografii. Uzyskane wyniki naukowe, opublikowane w wieloautorskich, recenzowanych artykułach naukowych są w niezależny sposób opisane, z dodatkowymi wyprowadzeniami i wyjaśnieniami.

Część pierwsza pracy (rozdziały 1-5) przedstawia motywację do podjęcia badań, cele naukowe, przedstawienie i pewna formalizacji pojęć i metod wykorzystywanych w pracy.

Rozdział 1, wstęp stanowi bardzo popularne i poglądowe omówienie obecnej wiedzy biologicznej dotyczącej struktury i własności chromatyny. Doktorant przytacza wiele faktów, które uzasadniają złożoność problematyki badania struktury i funkcjonalności chromatyny. Długość łańcucha DNA upakowanego w chromosomach, w jądrze komórkowym, aspekty

funkcji/zadań biologicznych realizowanych przez DNA, które muszą być precyzyjnie realizowane dla funkcjonowania komórek. Na końcu rozdziału wstępnego autor charakteryzuje obszar badań prowadzonych w ramach pracy. Jest nim powiązanie metod modelowania matematycznego, biologii obliczeniowej i strukturalnej oraz bioinformatyki dla budowy i analizy przestrzennych modeli chromatyny. Autor słusznie podkreśla tu interdyscyplinarny charakter problematyki poruszanej w ramach pracy.

W rozdziale 2, na wstępie przytacza się fakty dotyczące budowy fizycznej i chemicznej łańcucha DNA. Następnie Doktorant przechodzi do wyższych poziomów organizacji chromatyny. Przedstawia koncepcje pętli, domen oraz kompartmentów. Wspomina o mechanizmach i czynnikach formowania pętli. Przykładową strukturę kontaktów i kompartmentów ilustruje na (zapożyczonym z literatury) rysunku 2.9. Cytuje za publikacjami z literatury metodę wykorzystania pierwszej składowej głównej macierzy kontaktów do oceny położenia/kształtu kompartmentów chromatyny.

Rozdział 3 poświęcony jest przedstawieniu szeregu technik eksperymentalnych biologii molekularnej użytecznych do badania struktury chromatyny, 3C, 4C, 5C, Hi-C, ChIA-PET. Spośród omówionych metod dwie ostatnie mają obecnie bardzo duże znaczenie poznawcze, pozwalają na tworzenie map kontaktów w łańcuchu DNA, co pozwala na budowę i rozwój algorytmów do badania trójwymiarowej geometrii chromatyny. Doktorant w rozdziale 3 omawia także jeszcze inne metody eksperymentalne możliwe do zastosowania w badaniu chromatyny, ChIA-Drop, FISH, EM-ISH, iPALM. Ostatnie z wymienionych metod są technikami związanymi z mikroskopią. Doktorant dyskutuje charakterystyczne cechy, zalety i wady dostępnych technik eksperymentalnych.

Rozdział 4 poświęcony jest modelowaniu chromatyny na poziomie mezoskopowym. Doktorant przedstawia przegląd mezoskopowych modeli mechanicznych budowy łańcucha DNA, poczynawszy od łańcucha idealnego, jednowymiarowego, o doskonałej wiotkości, poprzez kolejne modele proponowane w literaturze, Kratky-Poroda, WLC, model OSF oraz model fraktalny. Dyskutuje zasady ich budowania oraz własności powstałych modeli.

Wreszcie ostatni w części pierwszej, rozdział 5 poświęcony jest zasadom modelowania struktur wewnętrzządrowych oraz możliwości ich weryfikacji. Dla obrazowego sformułowania swojej oceny problematyki modelowania chromatyny, Doktorant przeciwstawia problematykę modelowania chromatyny oraz problematykę modelowania struktur białek. Dyskutując różnice w tych zasadach modelowania formułuje słuszną tezę, że zarówno sam proces modelowania chromatyny jak też możliwości weryfikacji modeli struktury i funkcji chromatyny stanowią większe wyzwanie naukowe niż procesy modelowania struktur białkowych. W aspekcie faktu, że przeprowadzenie projektu modelowania struktury chromatyny wymaga podjęcia decyzji o wielu parametrach algorytmu modelowania oraz o jego krokach i kolejności ich wykonywania, Doktorant w ostatnich podpunktach rozdziału 5 stara się sformułować zasady określania strategii modelarskiej dla chromatyny. Decyzje, które należy podjąć dotyczą takich parametrów jak

skali modelowania, własności symulowanego łańcucha DNA, własności otoczenia badanego polimeru, lub jego części.

Część druga pracy obejmująca rozdziały od 6 do 10 nazwana jest przez Doktoranta częścią praktyczną. Zawiera ona opisy algorytmów/narzędzi modelowania opracowanych w ramach pracy oraz opisy ich zastosowania.

Rozdział 6 stanowi pewnego rodzaju wstęp do części drugiej. Doktorant wymienia tradycyjnie definiowane/wykorzystywane w genomice struktury wewnątrzchromosomowe, geny, elementy regulacyjne, centromery, telomery oraz przedstawia ich odpowiedniki w algorytmach modelowania kształtu i struktury chromatyny, domeny i ich definicje TAD, CCD. W dalszej części oraz na ilustracjach przedstawia struktury wykorzystywane w modelowaniu geometrii trójwymiarowej łańcucha DNA, kontakty, mapy kontaktów, pętle, diagramy łuków. Opisuje możliwe metody detekcji zdefiniowanych struktur wewnątrzchromosomowych. W podpunktach 6.1 oraz 6.2 doktoratu autor przedstawia dwa opracowane przez siebie narzędzia do modelowania struktury chromatyny, CCD Caller oraz MetaTAD Caller. Dwa opisane narzędzia mogą być dopasowane do rodzaju dostępnych danych eksperymentalnych stanowiących informacje wejściowe dla modelowania, Hi-C lub ChIA-PET. Dla obu narzędzi dostępne są otwarte kody źródłowe oprogramowania.

W rozdziale 7 opisany jest problem skalowania wielowymiarowego (MDS, multidimensional scaling) oraz proponowane dla jego rozwiązania algorytmy jako narzędzia modelowania trójwymiarowej struktury chromatyny. W podrozdziale 7.1 Doktorant proponuje schemat wykorzystania podejścia do konstrukcji trójwymiarowego modelu struktury chromatyny z wykorzystaniem algorytmów dystansu grafowego.

W rozdziale 8 dokładnie opisany jest model oraz związane z nim algorytmy obliczeniowe oznaczane w pracy doktorskiej najczęściej skrótem SM, (spring model, model sprężynkowy, model kulkowo - sprężynkowy). Model ten jest związany z kilkoma publikacjami Doktoranta. Główna idea konstrukcji modelu SM przedstawiona jest w punkcie 8.1. W algorytmie oceny struktury chromatyny dokonuje się iteracyjnej modyfikacji jej konformacji tak aby minimalizować energię polimeru DNA. W punkcie 8.2 opisana jest funkcja energii polimeru DNA składająca się z energii wiązania pomiędzy kolejnymi kulami, energii zgięcia związanej z kątem płaskim pomiędzy trójkami kolejnymi kul, energii wyłączonej objętości oraz energii interakcji (odległych wiązań) pomiędzy kulami tworzącymi oddziałującą parę. Dla energii wiązania kolejnych kul oraz dla energii zgięcia przyjęte są modele harmoniczne. Energia wyłączonej objętości wyliczana jest na podstawie modelu Lennarda – Jonesa. Energia wiązań odległych jest wyznaczana na bazie modelu harmonicznego z obcięciem. Kilka podpunktów rozdziału 8 poświęconych jest możliwym sposobom ustalania początkowej struktury polimeru chromatyny do dalszych obliczeń metodą SM. Struktury polimerów DNA analizowane w trakcie obliczeń, a także struktury wejściowe i wyjściowe są przechowywane/zapisywane w formacie PDB. Podawane są przykłady zastosowania algorytmu i oprogramowania SM, badanie wpływu delekcji regionu

CTCF na przykładzie locusa genu TAL1 (za pracami w literaturze), konstrukcja globuli fraktalnej zgodnie ze scenariuszem literaturowym polegającym na ściągnięciu fragmentów globuli do środka sfery. Rozdział 8 kończy opis struktury i funkcjonalności opracowanego oprogramowania oraz aplikacji internetowej. Przedstawione są możliwości obliczeniowe i możliwe warianty obliczeń.

W rozdziale 9 przedstawione jest jeszcze jedno podejście do modelowania struktury chromatyny, związane z dążeniem do dopasowania wyników modelowania do eksperymentów. Polega ono na konstrukcji algorytmów modyfikacji modelowanej struktury polimeru DNA sterowanych obrazami. Obrazy uzyskiwane technikami mikroskopowymi (EM-ISH, IPALM, STED) służą do zdefiniowania potencjałów lub pól sił powodujących odpowiednią modyfikację struktury modelowanej chromatyny. Wspomina się dwie metody implementacji eksperymentalnych danych obrazowych w algorytmach modelowania chromatyny, metodę siłową i metodę korelacyjną. Obie metody prowadzą do modyfikacji pola sił oddziałujących na polimer DNA. Kolejno obie te metody są dokładniej omówione. Następnie przedstawione są przykłady obliczeniowe. Pierwszy z nich dotyczy obrazu wygenerowanego sztucznie, na podstawie modelu fragmentu polimeru DNA, a drugi bazuje na rzeczywistym obrazie mikroskopowym jądra komórkowego rzodkiewnika.

Rozdział 10 jest opisem infrastruktury obliczeniowej używanej w pracach związanych z pracą, której Doktorant jest autorem i projektantem.

Ocena rozprawy

Recenzowaną rozprawę oceniam bardzo pozytywnie. Bardzo zwracając uwagę cecha pracy, to szeroki zakres tematyczny obejmujący podstawy biologiczne zasady modelowania matematycznego, obszerne bioinformatyki polimerów DNA, wyczerpujące omówienie technik eksperymentalnych orientowanych na badanie budowy polimerów DNA, budowy oprogramowania, technik obliczeniowych łącznie z aspektami konstrukcyjnymi klastra obliczeniowego. Doktorant wykazuje się tym samym szeroką wiedzą i opanowaniem aspektów modelowania, obliczeń, konstrukcji oprogramowania na bardzo szczegółowym i zaawansowanym poziomie. Uzyskiwane wyniki naukowe mają walory interdyscyplinarne. Motorem wielu przeprowadzonych w ramach doktoratu prac jest zrozumienie podstaw biologicznych zjawisk. Opanowany przez Doktoranta warsztat naukowy ma potencjalnie bardzo wiele zastosowań w medycynie, farmakologii, klinicyście.

Warte podkreślenia jest, że praca jest napisana bardzo ładnym, żywym i jasnym językiem. Zawiera formalizmy matematyczne sformułowane poprawnie i odpowiednio charakteryzujące przeprowadzone badania z zakresu modelowania matematycznego.

Za najważniejsze, oryginalne osiągnięcie naukowe pracy uważam opisany w rozdziale 8 algorytm modelowania bazujący na strukturze koralikowo – sprężynkowej, a także w

publikacjach współautorskich Doktoranta. Jest to system o złożonej strukturze, skomplikowanej postaci funkcji energii oraz o potencjalnie złożonej dynamice molekularnej. Uzyskanie poprawnych wyników modelowania, weryfikowanych porównaniem uzyskanych struktur z wynikami eksperymentów jest dużym osiągnięciem naukowym.

Także inne opisywane wyniki naukowe, jak metody wyszukiwania domen chromatynowych, dopasowanie struktur sterowane obrazem są bardzo wartościowe.

Publikacje naukowe, których współautorem jest Doktorant (dodatek A), tematycznie związane z modelowaniem genomu ukazały się w bardzo renomowanych, wysoko punktowanych czasopismach naukowych.

Uwagi krytyczne i dyskusyjne, pytania do doktoratu

Szerokość tematyki badawczej podjętej w pracy powoduje, że wiele zagadnień przedstawionych jest pobieżnie. Przydałoby się więcej przykładów użycia opisywanych w pracy algorytmów do rzeczywistych danych, a także szersze porównania opracowanych metod z wynikami algorytmów z literatury.

Wyniki rozdziału 5 o zasadach modelowania chromatyny mają zbyt mało odniesień w innych częściach pracy.

W pracy brak jest rozdziału, który stanowiłby podsumowanie przeprowadzonych prac oraz uzyskanych wyników.

W pracy jest dość dużo usterek edycyjnych. Przykłady są przedstawione poniżej

Str 104:

Okazuje się, taka częstość w globuli fraktalnej jest zbliżona do częstości obserwowanej w kontaktach chromatyny dlatego. Z tego powodu może być wykorzystywana jako model jądra komórkowego. (urwane zdanie).

str 92:

... kąty pomiędzy każdym trzem sąsiadującymi punktami (błąd gramatyczny)

str 91, 96, 97, 98:

Okrąg (powinno być okrąg)

Str 111:

Minimalizacją energii z opcjonalną relaksacją (błąd gramatyczny)

Str 111:

Pytania:

1. Na rysunku 8.14 przedstawiono schemat generowania oraz rysunek wygenerowanego modelu fraktalnego jądra komórkowego. Skąd wiadomo, że wygenerowana struktura spełnia prawo potęgowe. Ile wynosi wartość alfa? Czy ta wartość zależy od jakichś parametrów modelowania?
2. W pracy w rozdziale 8 wspomina się o termostacie Andersena. Użycie tej idei wymagałoby dodania członu energii kinetycznej do bilansu energii. Czy taka modyfikacja jest dokonywana?
3. Czy bada się wpływ wartości temperatury na konformację chromatyny?
4. Wspominana praca Andersena zawiera nie tylko koncepcję termostatu ale także barostatu. Czy w eksperymentach kontroluje się ciśnienie? Czy ma ono jakiś wpływ na wyniki?
5. W pracy cytuje się artykuł [41] o modelowaniu polimerów. W artykule tym stosuje się inną koncepcję kontroli temperatury, przez odpowiednie zdefiniowanie współczynnika tłumienia w równaniach Langevina oraz dodanie do każdego bilansu sił dodatkowej losowej siły. Czy ta idea jest stosowana w opracowanych algorytmach? Czy porównywano ją z koncepcją termostatu?

Konkluzja

Praca stanowi podsumowanie oryginalnych, interdyscyplinarnych badań naukowych, w których sformułowano i weryfikowano oryginalne hipotezy badawcze. Dokumentuje wiedzę i kompetencje Doktoranta. Osiągnięcia i oryginalne elementy pracy są na pewno wystarczające do jej ogólnej pozytywnej oceny. Stwierdzam, że rozprawa spełnia warunki stawiane pracom doktorskim i wnioskuję o jej dopuszczenie do publicznej obrony.

Ponadto, pomimo uwag krytycznych, biorąc pod uwagę poziom naukowy pracy, opanowania przez Doktoranta bardzo szerokiego zakresu tematycznego oraz dorobek publikacyjny Doktoranta wnioskuję o wyróżnienie rozprawy.



Recenzja Rozprawy Doktorskiej

Tytuł: Komputerowe modelowanie trójwymiarowej struktury chromatyny
Autor: mgr. inż. Michał Kadlof
Promotor: prof. dr hab. Dariusz Plewczyński

Tematyka badawcza

Rozprawa mgr inż. Michała Kadlofa dotyczy komputerowego modelowania trójwymiarowej struktury chromatyny — zagadnienia leżącego na styku bioinformatyki, biologii strukturalnej oraz nauk obliczeniowych. Autor podejmuje się analizy oraz rozwoju algorytmów służących do rekonstrukcji przestrzennej organizacji genomu, integrując różnorodne typy danych eksperymentalnych — genomowych i mikroskopowych — oraz własne rozwiązania programistyczne. Tematyka pracy jest aktualna i oryginalna, dotyczy istotnego problemu bioinformatyki strukturalnej i wpisuje się w światowe trendy badań nad strukturą genomu.

Układ rozprawy doktorskiej (w tym informacje o jej poszczególnych częściach składowych)

Monografia została napisana w języku polskim, liczy 170 stron i jest podzielona na trzy główne części: teoretyczną (Rozdziały 2-5), praktyczną (Rozdziały 6-10) oraz dodatki (wykaz publikacji autora, zestawienia wariantów histonów). Zawiera także streszczenie w języku polskim i angielskim, wykaz użytych skrótów, szczegółowy spis treści, spisy rysunków i tabel oraz bibliografię. Struktura pracy jest logiczna i przejrzysta. Każdy rozdział pełni jasno określoną funkcję, a układ tekstu sprzyja czytelności i systematycznemu przedstawieniu omawianych zagadnień. Mocną stroną pracy jest wyraźne oddzielenie części przeglądowej od autorskich rozwiązań oraz szczegółowe omówienie aspektów technicznych i infrastrukturalnych.

Rozdział 1 stanowi wprowadzenie do tematyki badawczej, skupiając się na trójwymiarowej strukturze chromatyny oraz jej modelowaniu komputerowym. Autor podkreśla wyjątkową złożoność i piękno chromatyny, która jest nośnikiem informacji genetycznej w każdej komórce eukariotycznej. W rozdziale opisano znaczenie chromatyny zarówno jako nośnika informacji, jak i jako złożonego systemu technicznego, wyposażonego w mechanizmy replikacji, automatyzacji, detekcji i naprawy

błędów. Przedstawiono skalę problemu upakowania DNA, scharakteryzowano wielopoziomą organizację struktury – od podwójnej helisy DNA, przez upakowanie na białkach histonowych, po wyższe poziomy organizacji, które umożliwiają sprawne funkcjonowanie genomu. Opisano wyzwania badawcze związane z analizą konformacji chromatyny oraz podkreślono rosnącą rolę biologii obliczeniowej w podejmowaniu tych wyzwań. Rozdział kończy się określeniem celu pracy, którym było opracowanie nowych metod komputerowego modelowania przestrzennych struktur chromatyny, integrujących różne typy danych biologicznych i umożliwiających lepsze zrozumienie organizacji genomu.

W Rozdziale 2 pracy doktorskiej autor przedstawił wprowadzenie do budowy i organizacji chromatyny w komórkach eukariotycznych. Opisał, w jaki sposób materiał genetyczny człowieka jest upakowany w jądrze komórkowym w postaci chromatyny i podkreślił wielopoziomą strukturę tej cząsteczki. Na początku rozdziału zaprezentowano podstawowe informacje o DNA: jego budowę, długość i ilość w komórce ludzkiej. Następnie omówione zostały poziomy organizacji chromatyny, poczynając od podwójnej helisy DNA, poprzez strukturę nukleosomów (DNA nawiniętego na białka histonowe), aż po wyższe poziomy organizacji, takie jak włókno 11 nm i potencjalne włókno 30 nm. W dalszej części rozdziału przedstawiono składniki chromatyny, w tym różne typy histonów oraz ich masy, a także oszacowano masę i gęstość całej chromatyny w jądrze komórkowym. Autor wyjaśnił, w jaki sposób upakowanie DNA na poziomie nukleosomów oraz obecność tzw. linkerów (łącników) wpływa na fizyczne właściwości chromatyny. Kolejną część rozdziału poświęcono pętłom, domenom (TAD, CCD), kompartmentom oraz mechanizmom ich powstawania – z udziałem białek takich jak CTCF i kohezyzna – wyjaśniając, jak te struktury wpływają na regulację ekspresji genów i funkcjonowanie genomu.

Rozdział 3 prezentuje przegląd metod eksperymentalnych wykorzystywanych do badania struktury chromatyny w jądrze komórkowym. Autor podkreśla, że nie istnieje pojedyncza technika pozwalająca na dokładne określenie struktury i pełne zrozumienie organizacji chromatyny. Opisane techniki dzieli na mikroskopowe, umożliwiające obserwację struktur w skali mikrometrów (np. FISH, EM-ISH, iPALM), oraz molekularne, pozwalające na analizę struktur w skali atomowej (krystalografia rentgenowska czy NMR). Na początku rozdziału Doktorant przedstawił metody genomowe, takie jak: 3C, 4C, 5C, Hi-C, ChIA-PET oraz ChIA-Drop. Następnie opisał metody mikroskopowe, m.in. konfokalną mikroskopię FISH, mikroskopię elektronową EM-ISH oraz technikę iPALM. Autor słusznie zaznacza, że zaawansowane modelowanie struktury chromatyny wymaga integracji danych pochodzących z różnych technik, gdyż tylko takie podejście pozwala na uzyskanie pełniejszego obrazu trójwymiarowej organizacji genomu i zrozumienie jej funkcjonalnych konsekwencji.

Rozdział 4 pracy poświęcono koncepcjom i modelom teoretycznym opisującym strukturę chromatyny w skali mezoskopowej. Przedstawiono tu główne podejścia teoretyczne, takie jak: łańcuch doskonały, model Kratky'ego-Poroda, model WLC, model Odijka-Skolnicka-Fixmana oraz koncepcję globuli

fraktalnej. Rozdział ukazuje ewolucję modeli polimerowych wykorzystywanych do opisu organizacji chromatyny i stanowi teoretyczne podstawy dla jej komputerowego modelowania.

Rozdział 5 stanowi praktyczny przewodnik po procesie modelowania struktur chromatyny, integrując aspekty biologiczne, informatyczne i techniczne. Autor przedstawia kluczowe wyzwania, strategie oraz reguły, które powinny pomóc w tworzeniu użytecznych i wiarygodnych komputerowych modeli chromatyny. Należą do nich: (i) specyfika modelowania chromatyny, (ii) projektowanie strategii modelarskiej z uwzględnieniem określenia celu modelowania, wyboru regionu genomu i skali modelowania, doboru rozdzielczości modelu oraz wyboru i integracji danych eksperymentalnych, (iii) wybór cech polimeru, (iv) maszynowa reprezentacja polimeru.

Rozdział 6, będący pierwszym w części praktycznej, poświęcony jest metodom wykrywania i analizy dużych struktur wewnątrzchromosomowych, takich jak domeny kontaktowe chromatyny (CCD) oraz domeny topologicznie asocjowane (TAD). Autor koncentruje się na praktycznych narzędziach i algorytmach umożliwiających automatyczną identyfikację tych struktur na podstawie danych eksperymentalnych. Przedstawia dwie metody opracowane w ramach pracy doktorskiej – CCD Caller oraz MetaTAD Caller. Pierwsza z nich służy do automatycznej detekcji regionów o ponadprzeciętnym pokryciu łukami interakcji (na podstawie diagramów łuków). Metoda polega na porównaniu rzeczywistego sygnału z sygnałem oczekiwanym – dodatnie odchylenie wskazuje na obecność domen CCD. Podejście to eliminuje subiektywność analizy wzrokowej i umożliwia automatyzację procesu dla całych chromosomów lub genomów. Z kolei MetaTAD Caller to narzędzie oparte na podejściu kompromisowym, które integruje wyniki kilku zewnętrznych algorytmów wyznaczających granice TAD-ów oraz wykorzystuje mechanizm głosowania z rozmyciem granic. Pozwala to uzyskać bardziej wiarygodny podział i zminimalizować błędy wynikające z działania pojedynczych, niedoskonałych metod lub parametrów.

Rozdział 7 pracy doktorskiej skupia się metodzie skalowania wielowymiarowego (MDS) służącej do budowy trójwymiarowych modeli struktury chromatyny na podstawie danych eksperymentalnych. Autor wyjaśnia na czym polega MDS i opisuje dystans grafowy jako jedną z miar odległościowych stosowanych do ustalania podobieństwa między fragmentami chromatyny w tej metodzie. Przedstawił również praktyczne aspekty przetwarzania danych, takie jak przekształcanie macierzy kontaktów w macierze dystansów oraz wyzwania związane z interpretacją sygnałów pochodzących z eksperymentów genomowych.

Rozdział 8 pracy – wg mnie najciekawszy – opisuje autorski model sprężynkowy (Spring Model, SM) opracowany dla komputerowego modelowania struktur polimerowych, takich jak chromatyna. Model ten opiera się na minimalizacji funkcji energii potencjalnej, uwzględniającej wiązania między segmentami, sztywność łańcucha, niedostępną objętość oraz interakcje dalekozasięgowe wynikające z danych eksperymentalnych. Autor przedstawia matematyczne podstawy modelu, różne warianty

struktur początkowych, omawia ich wady i zalety oraz prezentuje przykłady praktycznego zastosowania, takie jak modelowanie domen chromatynowych, analiza wpływu delecji motywów CTCF czy modelowanie jądra fraktalnego. W rozdziale opisano także implementację modelu jako aplikacji internetowej Spring-model oraz programu md-soft, podkreślając elastyczność i użyteczność tych narzędzi w integracji różnych typów danych biologicznych.

Rozdział 9 przedstawia koncepcję modelowania struktur chromatyny sterowanego obrazem, które integrują dane mikroskopowe z komputerową symulacją zwijania polimeru. Autor przedstawia dwa podejścia do problemu: siłowe oraz korelacyjne. W metodzie siłowej obraz mikroskopowy traktowany jest jako pole potencjału, które oddziałuje na elementy modelu, „przyciągając” je do obszarów o większej gęstości sygnału. W metodzie korelacyjnej, na każdym etapie symulacji generowany jest obraz symulowany, który porównuje się z rzeczywistym obrazem mikroskopowym, a funkcja energii minimalizuje różnice między nimi. Rozdział opisuje także techniczne aspekty przetwarzania obrazów mikroskopowych, sposoby ich konwersji do formatu akceptowanego przez narzędzia modelarskie oraz testy obu metod na sztucznych i rzeczywistych danych. Przedstawiono przykłady zastosowania, w tym rekonstrukcję struktur na podstawie zdjęć jądra komórkowego.

Rozdział 10 opisuje zagadnienia związane z infrastrukturą obliczeniową, którą współtworzył oraz administrował Autor pracy. W rozdziale przedstawiono motywację budowy wydajnych klastrów obliczeniowych oraz praktyczne aspekty ich projektowania i zarządzania. Następnie szczegółowo omówiono architekturę i funkcjonowanie klastra Eden. Rozdział zawiera opis typów węzłów obliczeniowych, ich konfiguracji sprzętowej, zastosowanych magazynów danych, a także sieci komunikacyjnych. Przedstawiono system kolejkowy Slurm wykorzystywany do zarządzania zadaniami użytkowników, zasady priorytetyzacji i przydziału mocy obliczeniowej oraz ewolucję polityki dostępu do zasobów. Autor omówił także narzędzia monitoringu klastra, sposoby testowania wydajności magazynów danych oraz praktyczne wyzwania związane z eksploatacją dużych infrastruktur HPC. Rozdział kończy się zestawieniem najważniejszych parametrów klastra Eden oraz statystykami jego wykorzystania przez społeczność naukową.

Zastosowane piśmiennictwo

Autor korzysta z bogatego piśmiennictwa obejmującego zarówno klasyczne prace z zakresu biologii molekularnej i biofizyki, jak i najnowsze publikacje dotyczące modelowania genomu oraz algorytmiki. Cytowane artykuły — w liczbie 61 — pochodzą z lat 1949–2024 i zostały opublikowane w renomowanych czasopismach z dziedziny biologii, bioinformatyki i genomiki, w tym *Journal of Molecular Biology*, *Biochemistry*, *Nature*, *Genome Research*, *Nucleic Acids Research*, *Nature Communications*, *Bioinformatics*, *Scientific Reports*, *PLoS Computational Biology*, *Chromosome Research*, *BMC Bioinformatics*, *Nature Genetics*, *Journal of Structural Biology*, *Nature Methods*, *Nature*

Biotechnology, Nature Structural & Molecular Biology, Proceedings of the National Academy of Sciences of the USA, Genome Biology, Current Opinion in Structural Biology, Journal of Computational Chemistry. Publikacje te dotyczą zagadnień takich jak struktura i organizacja chromatyny, regulacja ekspresji genów i epigenetyka, metody badawcze i technologie eksperymentalne, dynamika i funkcje chromatyny, choroby i nieprawidłowości genomowe oraz ewolucja i różnorodność molekularna. Piśmiennictwo odzwierciedla interdyscyplinarność badań nad chromatyną, łącząc biologię molekularną, biofizykę, genetykę, informatykę oraz nowoczesne technologie badawcze. Bibliografia jest właściwie dobrana i obejmuje zarówno oryginalne prace badawcze, jak i przeglądowe. W pracy nie stwierdzono poważnych braków literaturowych ani istotnych błędów formalnych w cytowaniach. W spisie literatury znaleziono jedynie nieliczne drobne usterki, takie jak nazwy czasopism pisane z małej litery.

Cel pracy oraz zastosowane metody badawcze

Głównym celem pracy doktorskiej Michała Kadlofa było opracowanie i wdrożenie nowych metod komputerowego modelowania trójwymiarowej struktury chromatyny w jądrze komórkowym. Kluczowym założeniem była możliwość integracji różnych typów danych eksperymentalnych w celu uzyskania wiarygodnych modeli przestrzennej organizacji genomu. Autor dążył do rozwoju i implementacji autorskich algorytmów modelarskich, wykorzystania danych pochodzących z różnych technik eksperymentalnych, stworzenia infrastruktury obliczeniowej umożliwiającej zaawansowane symulacje oraz praktycznego zastosowania i walidacji opracowanych narzędzi na rzeczywistych danych biologicznych.

W pracy zastosowano szeroki zakres metod badawczych, obejmujących metody analizy danych, eksploracji wizualnej, modele i techniki matematyki ciągłej oraz dyskretnej, metody optymalizacyjne i algorytmiczne, jak również zróżnicowane techniki programistyczne. Autor szczegółowo przeanalizował dostępne techniki genomowe i mikroskopowe wykorzystywane do badania organizacji chromatyny. Przeprowadził eksplorację danych pochodzących z różnych źródeł oraz zintegrował je w opracowanych przez siebie narzędziach obliczeniowych.

Dbając o transparentność i możliwość replikacji wyników, Autor udostępnił opracowane narzędzia programistyczne w sposób otwarty – kody źródłowe zostały opublikowane za pośrednictwem platformy Bitbucket, powstała również aplikacja sieciowa działająca w domenie mini.pw.edu.pl.

Wyniki badań oraz ich praktyczne zastosowanie

W pracy doktorskiej mgr inż. Michał Kadlof zajął się problematyką komputerowego modelowania trzeciorzędowej struktury chromatyny, integrując szeroki zakres danych eksperymentalnych oraz rozwijając własne narzędzia bioinformatyczne. Autor dogłębnie zapoznał się z problematyką analizy danych genomowych i mikroskopowych, a następnie – na podstawie zdobytego doświadczenia –

opracował autorskie metody i narzędzia obliczeniowe, które zostały przetestowane na różnorodnych zestawach danych biologicznych. Za najważniejsze osiągnięcia badawcze autora uznaje:

- Opracowanie modelu sprężynkowego, który umożliwia modelowanie polimerów na podstawie danych eksperymentalnych. Metoda symulacji łańcuchów chromatyny wykorzystująca ten model została zaimplementowana w aplikacji sieciowej Spring-model, dostępnej pod adresem <https://spring-model.mini.pw.edu.pl/>. Silnik został opublikowany na platformie Bitbucket: https://bitbucket.org/4dnucleome/md_soft/. Spring-model umożliwia szerokiemu gronu użytkowników szybkie modelowanie i wizualizację struktur chromatyny na podstawie własnych danych. Z kolei md-soft, stanowiący silnik aplikacji Spring-model i będący uogólnieniem użytych w niej metod pozwala na integrację różnych typów danych, symulację ewolucji struktur w czasie oraz realizację rozmaitych scenariuszy modelowania w szerszym zakresie niż Spring-model. Md-soft wykorzystuje zewnętrzny silnik OpenMM służący do symulacji dynamiki molekularnej i umożliwia prowadzenie symulacji zarówno pojedynczych, jak i masowych na CPU i GPU. Model sprężynkowy został opisany w artykule naukowym Kadlof *i in.*, *Methods*, 2019 (5Y IF=4,2; 100 punktów ministerialnych).
- Opracowanie i implementację dwóch algorytmów do identyfikacji struktur wewnątrzchromosomowych. Pierwszy algorytm stanowi podstawę narzędzia CCD Caller, które automatycznie identyfikuje domeny kontaktowe chromatyny na podstawie danych ChIA-PET, eliminując subiektywność ręcznej analizy. Algorytm został zaimplementowany w języku Python i udostępniony na Bitbuckecie: <https://bitbucket.org/4dnucleome/ccd-caller>. Drugi algorytm (MetaTAD Caller) jest oparty na metodzie konsensusowej wyznaczania granic domen topologicznych – integruje wyniki wielu zewnętrznych algorytmów i stosuje mechanizm głosowania do wygenerowania wyniku. Również ten algorytm został zaimplementowany w Pythonie i udostępniony przez serwis hostingowy Bitbucket: <https://bitbucket.org/4dnucleome/meta-tad-caller>.
- Praktyczne zastosowanie i walidację opracowanych metod na rzeczywistych danych eksperymentalnych. Opracowane narzędzia zostały wykorzystane m.in. do modelowania rzeczywistych fragmentów genomu, analizy wpływu delecji motywów CTCF, modelowania jądra fraktalnego oraz testowania algorytmów na danych mikroskopowych. Przeprowadzono również integrację i krytyczną analizę danych eksperymentalnych, prowadzącą do konstrukcji wiarygodnych modeli struktur 3D chromatyny.

Ważnym elementem pracy doktorskiej jest także kompleksowy przegląd i klasyfikacja technik eksperymentalnych, który – choć nie stanowi bezpośrednio wyniku naukowego – ma istotną wartość edukacyjną. Praca zawiera systematyczny opis genomowych i mikroskopowych metod pozyskiwania danych o strukturze chromatyny, przedstawia proces modelowania, dostarczając przydatnych

wskazówek dotyczących poszczególnych etapów modelowania, a także podkreśla konieczność integracji danych wieloskalowych w celu uzyskania spójnych i realistycznych modeli.

Ostatnim istotnym osiągnięciem Autora jest budowa i administracja infrastruktury klastra obliczeniowego Eden. Autor miał znaczący wkład w zaprojektowanie i wdrożenie tego klastra, który zapewnia wysoką moc obliczeniową niezbędną do realizacji zaawansowanych symulacji bioinformatycznych.

Nieprawidłowości i braki w ocenianej rozprawie doktorskiej

Praca została napisana prostym i zrozumiałym językiem oraz bogato zilustrowana. Niemniej jednak, w tak obszernym dokumencie nietrudno o błędy językowe i nieścisłości merytoryczne – część z nich udało mi się zidentyfikować i przedstawiam je poniżej.

- a) W rozprawie zauważalny jest brak konsekwencji w stosowaniu zasad typografii obowiązujących w języku polskim. Na przykład zgodnie z normami języka polskiego spójniki (takie jak "i", "a", "o", "u", "w" czy "z") nie powinny pozostawać na końcu linii, lecz być przenoszone na początek następnego wiersza. Błąd ten obniża estetykę tekstu.
- b) Zdarzają się błędy pisarskie, np. "ora metoda" (str. 6), "dna" (str. 16), "składa się z dwóch typów para zasad" (str. 16), "pod mikroskopem" (str. 17), "biologi" (str. 18), "chromatosmu" (str. 23), "Obecnie nie ma" (str. 40), "Technika ta polega doczepieniu" (str. 41), "posiada pewną całkowitość długość konturową" (str. 45), "wyniki algorytmów modelujących dla zwracały" (str. 50), "31 października" (str. 50), "przyjmie dowolną bzdurę i zawsze zwrócić jakiś wynik" (str. 55), "większe szans załapać" (str. 56), "wiązę się decyzja" (str. 57), "korzysta z formatu z popularnego formatu plików" (str. 60), "o niewielkich cząsteczkach" (str. 60), "Wynika to po z dwóch powodów" (str. 61), "zdefiniowanie więzów które fragmenty powinny wchodzić w interakcję z którymi." (str. 62), "związanego z wyłączoną objętości" (str. 66), "wewnątrzchromosomowych" (w tytule rozdziału, str. 73), "domeny posiadają kilka" (str. 73), "dzielimy na w miejscach w których" (str. 75), "uwzględnia tylko na wzajemne dystanse" (str. 81), "topologi połączeń" (str. 83), "równomierne ułożonymi punktami" (str. 87), "objętości" (str. 90), "problemowi problemowi rozmiaru" (str. 95), "W symulacja pojawiają się" (str. 98), "dalekozasięgowych wiązań działają" (str. 98), "w w przypadku" (str. 103), "kótkiej" (str. 107), "rozdzielczość" (str. 107), "obrazu na na pole" (tytuł rozdziału, str. 118), "Na struktura na symulacji" (str. 119), "rozdzielczości" (str. 129), "urządzaniach" (str. 129), "składał się" (str. 131), "najważniejsza jest ostatnie 8 dysków ost" (str. 137), "sięć" (str. 140).
- c) Podział na akapity miejscami wygląda na przypadkowy (np. str. 19, 22).
- d) Występują błędy interpunkcyjne (myślnik zamiast przecinka, przecinek zamiast kropki, brak kropki na końcu zdania, myślniki wstawiane niezgodnie z zasadami interpunkcji, co powoduje, że tekst

wygląda nienaturalnie, dezorientuje, rytm zdań staje się poszarpany i tekst jest trudny do śledzenia). Przykłady: "- którym kieruje" (str. 6), "szpule, Każdy" (str. 16), "Np," (str. 57), "Obecne techniki biologii, nie tylko są drogie, ale pozwalają na obserwację jedynie niewielkich fragmentów, w ograniczonej rozdzielczości, lub tylko niektórych procesów." (str. 17), "Przez dominującą część czasu, w interfazie, chromosomy występują w formie luźnej, czyli pod postacią chromatyny" (str. 19), "nukleosomów Świadomi" (str. 27), "Z zeskanowanych, w ten sposób, warstw" (str. 40), "obszaru, W przypadku" (str. 62), "tzn," (str. 74), "W efekcie pokrycie średnie ma bardzo, krótkie obszary początkowe i końcowe, a większość chromosomu." (str. 76), "Pamiętać jednak, należy, że w modelu SM, wszystkie rodzaje" (str. 85), "płaskie. (rys. 8.4) Dla" (str. 90).

- e) Zdarzają się także błędy ortograficzne, m.in. niepoprawny zapis liczebników i rozdzielny zapis przymiotników złożonych: "12-sto nukleotydowego" (str. 21) to niepoprawna forma zapisu liczebnika złożonego (Prawidłowy zapis to "dwunastonukleotydowego" lub "12-nukleotydowego".), "chromosomu 1-go" (str. 21) - poprawny zapis to "chromosomu 1" lub "pierwszego chromosomu", "składa się z 2-óch" (str. 22), "chromosomu 3-ego" (str. 31), "4-rech polimerów" (str. 63), "4-rech wariantów" (str. 65), "1-szy z 2-gim, 2-gi z trzecim" (str. 82), "12-ej" (str. 94), "wewnątrz jądrowych" (str. 49), "daleko zasięgowy" (str. 84, 91), "co w prawdzie pozwala" (str. 99), "wewnątrz chromatynowych" (str. 101).
- f) Znalazłam nieliczne błędy stylistyczne i składniowe, np. "Świadomi potencjalnych trudności interpretacyjnych to właśnie te obrazy zostaną użyte podczas opracowywania algorytmów modelowania sterowanego obrazem w dalszej części niniejszej pracy." (str. 40), "Zasada działania modelowaniu opiera się" (str. 86), "przy jednoczesnej czytelność dla człowieka" (str. 100), "Okazuje się, taka częstość w globuli fraktalnej jest zbliżona do częstości obserwowanej w kontaktach chromatyny dlatego." (str. 104), "rodzaj wiązań polimeru mogą być sztywne, lub elastyczne" (str. 127), "przyjęcie na siebie administratora" (str. 130), "Ponadto hosty oprócz kanału do obsługi macierzy każdy z hostów DGX jest wyposażony" (str. 141).
- g) W pracy autor używa określenia „figura” w odniesieniu do ilustracji, co stanowi kalkę z języka angielskiego. W polskojęzycznych tekstach naukowych powszechnie stosuje się termin "rysunek" (np. "rys. 1") – jest on bardziej naturalny i zgodny z przyjętymi standardami redakcyjnymi. "Figura" w języku polskim to kształt (figura matematyczna) lub postać. Ponadto zauważyłam niespójność w sposobie odnoszenia się do ilustracji: autor używa formy "rys. <nr>", "rysunek <nr>", "obrazek <nr>" (polska konwencja), jak i "fig. <nr>" (angielska konwencja). Taka niejednorodność redakcyjna jest niepożądana w tekście naukowym – należy konsekwentnie stosować jedną formę w całej pracy. W szczególności autor nie powinien używać słów/skrótów angielskich w pracy pisanej po polsku jeśli istnieją polskie odpowiedniki tych słów/skrótów.

- h) W streszczeniu autor pracy odnosi się do samego siebie w trzeciej osobie liczby pojedynczej ("autor rozpoczyna", "autor prezentuje", "autor omawia") , co jest nietypowe i niezalecane w rozprawach jednoautorskich. Przyjętą i zalecaną praktyką jest pisanie w formie bezosobowej lub w pierwszej osobie liczby pojedynczej.
- i) Niektóre informacje są podane niedokładnie, przez co przekaz traci klarowność. Przykładowo:
- niefortunny podział treści między zdania: "Po drugie składa się z dwóch typów para zasad A-T i C-G, które różnią się ilością wiązań wodorowych, które je tworzą. Pary A-T cechuje niższa energia wiązań, a to oznacza, że stabilność struktury i łatwość jej rozplecenia jest zależna od sekwencji." (str. 16) - Informacja o niższej energii wiązań A-T powinna znaleźć się już w pierwszym zdaniu – to bezpośrednia konsekwencja różnicy w liczbie wiązań wodorowych. Drugie zdanie powinno odnosić się ogólnie do wpływu sekwencji na stabilność struktury.
 - "By zwizualizować sobie trudność zagadnienia" (str. 16) - nie wiadomo o jakie zagadnienie chodzi (nawijanie chromatyny na białko? rozplątanie dwuniciowego DNA?)
 - "Gdyby chromatyna była grubości włóczki 3mm, jej długość wynosiłaby ponad 2500 km." (str. 16) - nie rozumiem co autor miał na myśli; jaki wpływ ma grubość łańcucha na jego długość? raczej przypuszczałabym, że grubszy łańcuch ma większą pojemność informacyjną, a zatem powinien być krótszy od swojego cienkiego odpowiednika (jeśli ma przechowywać taki sam wolumen danych)
 - "W przeciwnym kierunku białka działa blokująco" (str. 29) - zdanie jest niepoprawne stylistycznie i składniowo. Brzmi niejasno i jest trudne do zrozumienia.
- j) Inne błędy i nieścisłości:
- W podpisie pod rysunkiem 2.2 znajduje się informacja, że przedstawiono widok z boku (u góry) oraz w osi helisy (u dołu), podczas gdy rysunek zawiera dwa ujęcia rozmieszczone obok siebie — po lewej i prawej stronie.
 - Fragment "A zatem z prostego rachunku wynika, że w naszym ciele znajduje się: <wzór>" jest nieprecyzyjny. Autor pracy nie informuje, co dokładnie jest przedmiotem obliczeń - nie wiadomo, jaka wielkość czy liczba jest wyznaczana przedstawionym wzorem.
 - "Dlatego aby obliczyć precyzyjnie średnią masę jednej pary ludzkiego genomu konieczna jest znajomość" - co to jest "para ludzkiego genomu"?
 - Puste komórki w tabeli 2.2. oznaczają - jak mi niekam - brak atomu a nie niekompletną tabelę. Dużo bardziej intuicyjna w odbiorze byłaby wartość 0.
 - W pracy obserwuję niespójności w zapisie liczb zmiennoprzecinkowych - niektóre liczby zapisane są z kropką jako separatorem dziesiętnym, a inne z przecinkiem. W polskim piśmiennictwie naukowym standardowo stosuje się przecinek jako separator dziesiętny.

Stosowanie naprzemiennie kropki i przecinka powoduje, że praca traci czytelność i wartości liczbowe przestają być zrozumiałe. Przykładowo na stronach 24-25 mamy kropki (wzory, tabela 2.1, tabela 2.3) i przecinki (tabela 2.2).

- "Gdyby zastąpić punkty sferami (tzn. łańcuch nabrałby objętości) to jego rozmiar wzrósłby Ree $N3/5$." (str. 44) - niezrozumiałe; rozmiar wzrósłby o Ree, czy do wartości Ree, czy może Ree razy?
- Rozdział "łańcuch idealny" rozpoczyna się od definicji łańcucha doskonałego. Zakładam, że oba terminy oznaczają to samo, jednak nie rozumiem, dlaczego autor wprowadza pojęcie „łańcucha idealnego”, skoro dalej w pracy używa wyłącznie terminu „łańcuch doskonały”. Wprowadzenie dwóch określeń dla tego samego pojęcia jest niepotrzebne i wprowadza terminologiczny chaos. Od początku należy konsekwentnie stosować jedno, spójne nazewnictwo. Taka niespójność obniża czytelność tekstu i sprawia wrażenie braku staranności redakcyjnej.
- str. 55: zupełnie nie rozumiem wtrąconego zdania na temat pól siłowych
- "Rozdzielczością modelowania nazywamy ilość par zasad reprezentowaną przez pojedynczy punkt modelowanego układu." - pary są policzalne, a zatem mówimy o ich liczbie a nie ilości.
- W pracy umieszczono wiele wzorów, ale tylko niektóre z nich (str. 59, str. 114-116) mają nadane numery. Uważam, że wszystkie wzory powinny być ponumerowane.
- "W praktyce modelarskiej wykształciły się dwa pojęcia do określania domen" - chodzi raczej o podejścia do określania domen.
- Ilustracje zawarte w pracy są merytorycznie trafne i wizualnie bardzo dobrze dopracowane, jednak sposób przygotowania podpisów pod rysunkami pozostawia wiele do życzenia. Autor stosuje w nich czcionki w trzech różnych rozmiarach, co wygląda nieestetycznie i sprawia wrażenie niespójności. Szczególnie mylące jest to, że uszczegółowienie podpisu (dodatkowe opisy elementów ilustracji) zapisane jest czcionką większą niż właściwy podpis rysunku, mającą ten sam styl i rozmiar jak tekst główny pracy. Sprawia to wrażenie błędu formatowania. Na początku było też dla mnie bardzo mylące - nie wiedziałam czy jest to podpis pod rysunkiem czy doklejona, nieco inaczej sformatowana treść rozdziału.
- W przypisie na stronie 89 autor deklaruje, że w tekście będzie stosował określenie "niedostępna objętość" jako tłumaczenie angielskiego terminu excluded volume. Niestety, już dwa zdania dalej zaczyna używać sformułowania "wyłączona objętość", które – jak należy domniemywać – odnosi się do tego samego pojęcia. Taka niespójność terminologiczna jest myląca i negatywnie wpływa na klarowność oraz odbiór tekstu. Skoro autor zdecydował się na konkretne tłumaczenie, powinien konsekwentnie się go trzymać w całej pracy.

- "W większości zastosowań mechaniki molekularnej postać struktury wynikowej." (str. 92) - niepełne zdanie
- W rozdziale dotyczącym struktur początkowych autor pisze: "Kolejnym przykładem struktury początkowej jest okrąg" (str. 96). Należy zwrócić uwagę, że w języku polskim termin „okrąg” odnosi się przede wszystkim do jednostki administracyjnej lub organizacyjnej, natomiast figura geometryczna, o którą zapewne autorowi chodziło, to „okrąg”.
- W pracy brak jest konsekwencji w stosowaniu separatorów tysięcy w zapisie dużych liczb. W niektórych miejscach Autor używa znaku apostrofu (np. 2'000'000), w innych – spacji (np. 2 000 000), a czasami nie używa separatora w ogóle. W języku polskim standardem zapisu dużych liczb jest stosowanie spacji jako separatora tysięcy, zgodnie z normą PN-N-01204:1992 oraz zasadami poprawnej polszczyzny.
- Na stronie 112 zamieszczono wzór, jednak brak jest wyjaśnienia występujących w nim zmiennych i parametrów.

Wyżej wymienione nieścisłości i usterki nie wpływają istotnie na wartość naukową rozprawy.

Inne komentarze i pytania dotyczące rozprawy

Poniżej przedstawiam uwagi ogólne oraz pytania, które nasunęły mi się podczas lektury monografii:

- a) Praca momentami jest bardzo obszerna, a w niektórych miejscach pojawiają się powtórzenia, które warto byłoby skrócić w celu zwiększenia zwięzłości i klarowności przekazu.
- b) W wybranych fragmentach rozdziałów praktycznych przydałaby się szersza dyskusja na temat ograniczeń stosowanych metod oraz możliwych artefaktów modelowania.
- c) Opisy infrastruktury obliczeniowej mogłyby zostać skrócone i bardziej skoncentrowane na aspektach istotnych z punktu widzenia prowadzonych badań naukowych.
- d) Styl pracy jest rzeczowy i techniczny; miejscami warto byłoby zadbać o większą spójność językową i redakcyjną całego tekstu.
- e) Jakie hipotezy badawcze postawił autor rozprawy?
- f) Na jakiej podstawie autor stwierdza we Wstępie, że struktura chromatyny nie płacze się? Nie podano przy tym stwierdzeniu odnośnika do żadnej pracy na ten temat. Czy faktycznie przebadano struktury chromatyny pod kątem splątań / węzłów topologicznych? Taka analiza jest dość wymagająca obliczeniowo... Z tego, co znalazłam w literaturze wnioskuję, że w chromatynie powstają splątania. Większość z nich jest jednak nietrwała, gdyż są na bieżąco usuwane. Co więcej, na str. 66 wspomniany jest mechanizm powstawania splątań w strukturze chromatyny, co przeczy stwierdzeniu umieszczonemu na początku pracy.
- g) Podoba mi się pomysł umieszczenia w pracy przepisu na izolację DNA z cebuli, jednak jest on nieścisły - brak informacji o tym ile potrzeba soli, wody i etanolu.

- h) W Rozdziale 4 autor wielokrotnie posługuje się pojęciem segmentu, między innymi w kontekście definicji homopolimeru, przy wprowadzaniu parametrów $R(s)$ i $P(s)$ oraz przy opisywaniu modeli. Niestety, pojęcie segmentu nie zostało wprost zdefiniowane, co utrudnia zrozumienie dalszych wywodów. Dodatkowym źródłem niejasności jest użycie terminu region, który nie jest tożsamy z segmentem, jednak różnica między tymi pojęciami nie została wyjaśniona. Brak jednoznacznych definicji powoduje trudność w poprawnym zinterpretowaniu omawianych parametrów oraz w śledzeniu logiki rozważań (np. " $R(s)$ to dystans pomiędzy dwoma regionami oddzielonymi od siebie segmentami.").
- i) Nie rozumiem jaki związek ma chromatyna z łańcuchem idealnym. Chromatyna nie jest homopolimerem.
- j) Uważam, że bardzo przydatne byłyby rysunki pokazujące (choćby schematycznie) jak wyglądają wszystkie modele opisane w Rozdziale 4.
- k) Warto odnotować, że autor pisze o konkursie CASP w czasie przeszłym, co zwraca uwagę, ponieważ do tej pory był to cykliczny i aktywny projekt. Piszę tę recenzję w momencie, gdy NIH ogłosił wstrzymanie finansowania CASP, więc konkurs faktycznie może przejść do historii. Tym samym użycie czasu przeszłego okazuje się nie tylko zaskakujące, ale też — być może — prorocze.
- l) Czy narzędzia CCD Caller i MetaTAD Caller zostały opisane w opublikowanym artykule? Opis ich działania w rozdziałach 6.1 i 6.2 ma raczej ogólny charakter, co byłoby zrozumiałe, gdyby szczegóły techniczne oraz zastosowanie zostały przedstawione w artykule naukowym.
- m) $r_{i,j}$ we wzorze na str. 89 jest opisany jako odległość między parą punktów, ale wydaje mi się, że powinien on oznaczać odległość między parą kul i , j lub - dokładniej - między punktami będącymi środkami kul i , j - zgadza się?
- n) W rozdziale "Dygresja geometryczna" Autor pracy odnosi się do tzw. liczby pocałunkowej, przywołując ją jako ograniczenie liczby kontaktów, jakie może tworzyć jedna sfera w przestrzeni trójwymiarowej. Wydaje się jednak, że zastosowanie tego pojęcia w omawianym kontekście nie jest w pełni uzasadnione. Liczba pocałunkowa określa maksymalną liczbę jednakowych, nieprzenikających się kul, które mogą stykać się z daną kulą o tym samym promieniu, co oznacza bezpośredni kontakt powierzchni. Tymczasem w opisywanym modelu kontakt między sferami niekoniecznie oznacza fizyczny styk — raczej chodzi o bliskość lub połączenie krawędzią w grafie (modelu), a nie rzeczywiste zetknięcie się brył w przestrzeni. Z tego powodu liczba pocałunkowa nie stanowi ścisłego górnego ograniczenia dla liczby kontaktów. Warto byłoby więc doprecyzować, czy i w jaki sposób autor oszacował maksymalną liczbę kontaktów, biorąc pod uwagę realistyczny zakres odległości między kulami reprezentującymi loci.
- o) Od 2019 roku obowiązującym i zalecanym formatem zapisu danych strukturalnych w bazie Protein Data Bank jest PDBx/mmCIF. Tymczasem Autor pracy zdecydował się stosować przestarzały

format PDB do zapisu informacji o strukturze chromatyny. Biorąc pod uwagę ograniczenia tego formatu — zwłaszcza w kontekście reprezentacji bardzo dużych i złożonych struktur, jakimi są modele chromatyny — taka decyzja wydaje się nietrafiona. Co więcej, przyjęcie formatu, który nie jest już wspierany jako standard, może w dłuższej perspektywie zmniejszyć użyteczność i popularność opracowanych przez Autora narzędzi, ograniczając ich integrację z nowoczesnym oprogramowaniem i obowiązującymi standardami bioinformatycznymi. Mam świadomość, że wielu użytkowników wciąż korzysta z formatu PDB ze względu na jego prostotę i długoletnią obecność w narzędziach bioinformatycznych. Dlatego też wszystkie współczesne, nowo powstające narzędzia obliczeniowe analizujące dane strukturalne oferują wsparcie zarówno dla formatu PDB, jak i PDBx/mmCIF. Dlaczego Autor nie zdecydował się na takie rozwiązanie?

p) Intriguje mnie nazwa klastra obliczeniowego. Dlaczego Eden?

Wnioski końcowe

Rozprawa doktorska mgr inż. Michała Kadlofa zawiera oryginalne rozwiązanie istotnego problemu naukowego i stanowi wartościowy wkład w rozwój komputerowych metod modelowania struktury chromatyny. Praca spełnia wszystkie wymagania stawiane rozprawom doktorskim w dziedzinie informatyki technicznej i telekomunikacji, łącząc oryginalność, interdyscyplinarność oraz wysoką użyteczność praktyczną uzyskanych wyników. Autor wykazał się samodzielnością badawczą, umiejętnością syntezy wiedzy z różnych dziedzin oraz wysokimi kompetencjami technicznymi.

Wyniki badań zostały opublikowane w dziewięciu współautorskich artykułach naukowych w renomowanych czasopismach z dziedziny, opublikowanych w latach 2016–2024.

Na podstawie powyższej oceny stwierdzam, że rozprawa mgr inż. Michała Kadlofa pt. „Komputerowe modelowanie trójwymiarowej struktury chromatyny” spełnia wymagania określone w ustawie o stopniach naukowych i tytule naukowym, obowiązującej obecnie w Polsce. Rekomenduję dopuszczenie rozprawy do dalszych etapów przewodu doktorskiego oraz przyznanie Autorowi stopnia doktora nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.

.....

prof. dr hab. inż. Marta Szachniuk