

Niniejsza praca stanowi wkład w dziedzinę modelowania komputerowego złożonych systemów biologicznych, jak również pełni rolę przystępnego wprowadzenia do symulacji dynamicznej struktury trójwymiarowej chromatyny w komórkach ludzkich. Zaczynając od podstawowych pojęć, czytelnik jest prowadzony przez kluczowe narzędzia matematyczne, na których opiera się modelowanie polimerów, w tym omówienie integratorów numerycznych, konstrukcji pól sił molekularnych oraz uzasadnienia stosowania symulacji stochastycznych w kontekście dynamiki struktury chromatynowej. Krytyczna dyskusja, podkreśla istniejące luki w modelowaniu chromatyny, zwracając uwagę na interdyscyplinarny charakter aktualnie wykorzystywanych modeli komputerowych oraz krytyczną rolę danych pochodzących z doświadczeń biologii molekularnej i genomiki w tworzeniu modeli trójwymiarowych oraz ich walidacji.

Doktorat rozpoczyna się od omówienia modelu LoopSage który jest komputerową implementacją biofizycznych obserwacji procesu ekstruzji pętli. W tej części pracy doktorskiej omówione zostają kluczowe biomolekuły, które biorą udział w formowaniu się pętli, takie jak kohezyna czy białko CTCF. Następnie wprowadzona zostaje metoda Monte Carlo oparta na biofizycznym Hamiltonianie, rygorystycznie zweryfikowana względem eksperymentalnych map interakcji pochodzących z doświadczeń typu 3C (HiC, HiChIP, ChIA-PET). Na tej podstawie opracowana została metoda RepliSage, która uwzględnia szereg procesów biologicznych zachodzących w trakcie podziału komórek ludzkich, tworząc bardziej kompleksowy i obliczeniowo wymagający model integrujący trzy kluczowe zjawiska biofizyczne: propagację stanu epigenetycznego prowadzącą do kompartmentalizacji, ekstruzję pętli oraz replikację DNA.

Istotnym wyzwaniem było opracowanie metody walidacji tak złożonego modelu komputerowego za pomocą obserwacji eksperymentalnych. Osiągnięto to przez porównanie przewidywań modelu komputerowego RepliSage z danymi eksperymentalnymi z doświadczeń HiChIP przeprowadzonych w laboratorium, w warunkach bez i ze stresem. Dodatkowo porównano wyniki symulacji z danymi Hi-C pojedynczych komórek. Porównanie symulacji biofizycznych oraz wyników doświadczeń pokazało jak integrować w rygorystycznym modelu komputerowym zarówno biologiczne spostrzeżenia dotyczące obserwowanego zjawiska, jak również biofizyczny opis oddziaływań molekularnych za pomocą funkcji Hamiltona reprezentującej sumę energii kinetycznej i potencjalnej układu chromatyny w jądrze komórkowym. Tylko realistyczne uwzględnienie zarówno biologii jak i fizyki układu umożliwiło opracowanie realistycznego modelu komputerowego dynamiki zmian konformacji trójwymiarowej chromatyny w trakcie cyklu komórkowego.

Na koniec w pracy przedstawiono nowatorski algorytm MultiMM, który jest wszechstronnym i wydajnym komputerowym modelem wielo-skalowym, który łączy organizację chromatyny od poziomu nukleosomów o rozmiarze 11 nm aż po całe chromosomy i pełną strukturę 3D jądra komórkowego zajmującą w przybliżeniu kulistą przestrzeń o średnicy od 5 do 10 μm . W tym podejściu nukleosomy reprezentowane są za pomocą sztywnych obiektów trójwymiarowych w ramach podejścia geometrycznego w celu zwiększenia szybkości obliczeń, zaś pętle chromatynowe symulowane są dzięki harmonicznym potencjałom wiązań. Kompartymenty natomiast modelowane są dzięki zastosowaniu symulacji kopolimerów blokowych oraz uwzględniając dodatkowe interakcje, np. między kompartmentem B a błoną jądra komórkowego. MultiMM dostarcza w ten sposób wszechstronne i przystępne narzędzie wielo-skalowego modelowania komputerowego dla biologów i biofizyków umożliwiające

śledzenie zachowania się architektury chromatyny we wszystkich istotnych skalach przestrzennych i czasowych.

Podsumowując, praca doktorska omawia wyzwania komputerowej reprezentacji genomu ludzkiego, wprowadza szereg innowacji do aktualnie wykorzystywanych modeli, stanowiąc istotny wkład w genomikę obliczeniową oraz wskazując na nowe możliwości, jakie otwierają się dla symulacji molekularnych złożonych systemów biologicznych, potencjalnie prowadząc do dalszych odkryć w biologii molekularnej. W szczególności zwrócono uwagę na ograniczenia obecnie stosowanych metod. Podkreślono, że podejścia hybrydowe i wielo-skalowe, łączące dynamikę molekularną z technikami uczenia maszynowego wykorzystywanymi w procesie optymalizacji hiperparametrów, mają potencjał, aby sprostać tym wyzwaniom i przyczynić się do dalszego rozwoju dziedziny modelowania komputerowego systemów biologicznych w skali całej komórki.